STATISTICAL ESTIMATION AND PREDICTION IN PROBABILISTIC MODELS, WITH APPLICATION TO STRUCTURAL RELIABILITY

by

DANIELE VENEZIANO

LAUREA IN ARCHITECTURE, University of Florence, Italy
(1970)

Submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology
September, 1974

Signature of Author . . . . . .
Department of Civil Engineering,
August 12, 1974

Certified by. . . . . .
Thesis Supervisor

Accepted by . . . . . .
Chairman, Departmental Committee on Graduate
Students of the Department of Civil Engineering

to my father

ABSTRACT

STATISTICAL ESTIMATION AND PREDICTION IN
PROBABILISTIC MODELS,
WITH APPLICATION TO STRUCTURAL RELIABILITY

by

DANIELE VENEZIANO

Submitted to the Department of Civil Engineering on
August 12, 1974 in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

When modeling engineering systems it is common practice to neglect statistical uncertainty on the model parameters. Such an uncertainty originates from one estimating the parameters of the model from limited data; as a result, given the estimates, the true parameters values remain unknown. Statistical uncertainty can be reduced (in the limit: removed) by collecting additional information.

The model itself may be probabilistic, in which case statistical uncertainty superimposes to the uncertainty of the model itself. The combination of statistical and model uncertainties is known in the statistics literature as a problem of statistical prediction. If instead the model is deterministic, all the uncertainty has statistical nature, leading to problems of statistical estimation.

Both problems of statistical estimation and prediction are studied; their relevance to the reliability of engineering systems is quantified and illustrated through examples. It is found that the uncertainty from poor statistical information may dominate the response of the model (say with regard to failure-no failure events). This is particularly true in the tail behavior of the model since the occurrence probability of "rare" events may be modified substantially. Problems of statistical prediction are studied for univariate and multivariate, memoriless and serially dependent probabilistic models, and from both the viewpoints of Classical and Bayesian statistics. Data are considered to be either in censored or in uncensored format. In discussing estimation, the important problem is addressed of designing sampling experiments in an optimal way. This problem arises when estimating unknown functions of continuous (possibly multidimensional) parameter from a finite number of (possibly

noisy) observations; a solution has potential applications to diverse fields like soils, hydrology, biological systems surveillance and quality control.

Thesis Supervisor: C. Allin Cornell

Title: Professor of Civil Engineering

Acknowledgements

The author takes pleasure in acknowledging Professor C.A. Cornell for providing stimulating and wise guidance during his three years of graduate work at M.I.T. This thesis is the outcome of a close cooperation with Professor Cornell, whose specific contributions to it are too many to mention explicitly here.

The author wishes also to thank Professor E.H. Vanmarcke who supervised his research work during the first year of graduate study at M.I.T., and Professor I. Rodriguez-Iturbe who gave him the opportunity to work on a Water Resource project.

The wish is expressed that this thesis is the start of an even more intense and fruitful cooperation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

$f(\cdot)$	- probability density function (PDF)
$F(\cdot)$	- probability cumulative distribution function (CDF)
$\Phi(\cdot)$	- standard normal CDF
$l(\cdot)$	- likelihood function
$\mu, \underline{\mu}$	- mean value, mean vector
$\sigma, \underline{\Sigma}$	- variance, covariance matrix
ρ	- correlation coefficient
ν	- degrees of freedom
P_f	- failure probability
b (subscript)	- prior (in Bayesian sense)
a (subscript)	- posterior (in Bayesian sense)
PL (subscript)	- Proof Loading
PI (subscript)	- Perfect Information
N (upperscript)	- Normal population
EX (upperscript)	- Exponential population
G (upperscript)	- Gamma population
El (upperscript)	- Extreme type I population
LN (upperscript)	- Lognormal population
P (upperscript)	- Poisson population

INTRODUCTION

Uncertainty in probabilistic modeling

As a necessary premise to a work dealing with statistical modeling, the uncertainties in mathematical analogs to natural mechanisms are reviewed and classified. We shall not differentiate probabilities on account of their meaning; i.e., depending on whether they are factual, logical or evaluative. Instead, our classification will concern the object of probability statements, and in particular how they relate to mathematical models of nature.

The classical question should be addressed first, whether nature is deterministic or stochastic. The prevalent belief of scientists is that nature is deterministic: to say it with Laplace, that all natural events follow "the great (deterministic) laws of nature" and that an intelligence provided with absolute knowledge could establish such laws, predict the future with certainty and reconstruct events in the past with the same degree of confidence.

But can one prove the existence of "the great laws of nature"? Can one test the hypothesis of a deterministic against a stochastic nature? The answer being no, the original problem belongs to the speculative domain of

metaphysics and has no "correct" general solution, in the sense of a demonstrable one. Indeed, having an answer to this problem is not important to what follows, except for its constituting a reference viewpoint, at all times interchangeable with the opposite one.

In this sense and with this arbitrariness, let us assume that natural phenomena are the output of a deterministic natural system whose future behavior will "resemble its past". This last assertion, called "the principle of uniformity of nature", is essential to natural sciences in that it provides the basis for scientific induction and prediction. The primary objective of natural sciences is to produce mathematical analogs to natural systems (system identification problem) and to describe their present state (state reconstruction or state estimation problem) at least to the extent which is relevant to the phenomena under study (definition of boundaries).

For the purpose of illustration consider a discrete sequence of events generated by a recursive "law of nature" of the type:

$$e_j = f(e_{j-1}, e_{j-2}, e_{j-3}) \ , \qquad (1)$$

where e_j denotes the magnitude of the j^{th} event and f is a deterministic function. Given $\{e_j\}$, but with f unknown, the mathematical modeling of the natural mechanism might start from assuming some boundaries (e.g., e_j does not depend directly on e_k, for $k < j - 8$) and then proceed hypothesizing a mathematical structure within these boundaries; for instance $e_j = f_1(e_k, \underline{\Theta}; j-8<k<j)$. f_1 is a given class of functions of $\{e_k\}$, indexed by the parameter $\underline{\Theta}$, and $\underline{\Theta}$ is estimated from the available data. Finally, in an experimental situation, predictions from the model are confronted with the outcome of nature. In absence of agreement one returns to a previous stage, redefining the boundaries, or the mathematical structure, or the parameters.

The foregoing example was kept simple on purpose. In most circumstances natural systems are quite complex, without well-defined boundaries, and little information is available on their mathematical structure. A more practical approach might then be to construct a <u>probabilistic</u> analog, in which the mathematical structure is defined in terms of only few parameters and reflects our level of understanding the natural system, while the randomness accounts for the lack of complete knowledge or for the unwillingness of modeling the system in greater detail. (This is only a gross classification of possible approaches; for instance it

does not differientiate between structured (parametric) and nonstructured (nonparametric) uncertainty, which distinction is considered unnecessary for the moment). Thus, in modeling the sequence introduced earlier one might restrict the search to a simple mathematical representation of the type:

$$e_j = \tilde{g}(e_{j-1}) \quad , \tag{2}$$

where $\tilde{g}$ is a stochastic relation to be determined. Because of the limitation in the arguments of $\tilde{g}$, the model (2) cannot describe exactly the actual process (1); hence its probabilistic character. The most one can do is to find the function $\tilde{g}$ which leads to the closest (in some probabilistic sense) representation of (1) through a stochastic first-order difference equation. As before one might start from hypothesizing a structured relation, say $\tilde{g}_1(e_{j-1}, \underline{\theta})$, and proceed sequentially towards the "best" form of $\tilde{g}$ and towards the "best" set of parameters $\underline{\theta}$. Let

$$e_j^* = \tilde{g}^*(e_{j-1}, \underline{\theta}^*) \tag{3}$$

denote the optimal solution, optimality being judged by an observer who knows the true law, equation (1). This final

condition ensures that (3) is a correct probabilistic model in the sense that long-run relative frequencies from (1) coincide with the probabilities in (3). Given e_{j-1}, e_j^* is a random variable; its uncertainty originates from the imperfect description of the constitutive law (1) by model (3); it might be called <u>probabilistic uncertainty</u>. A reduction in probabilistic uncertainty requires a modification of the model, say by using $e_j = \tilde{g}(e_{j-1}, e_{j-2})$ instead of $e_j = \tilde{g}(e_{j-1})$. If one chooses a third or higher order model: $e_j = \tilde{g}(e_{j-1}, e_{j-2}, e_{j-3}, \ldots)$ the optimal choice is $\tilde{g} = f$, with no probabilistic uncertainty.

Going back to a model of the type (2) with optimal choice (3), assume that $\tilde{g}^*$ is known. For instance $\tilde{g}^*$ might be a first-order autoregressive normal sequence, with correlation and variance as unknown parameters (vector $\underline{\theta}$). From a finite set of data the estimate $\hat{\underline{\theta}}$ of $\underline{\theta}^*$ is obtained. The stochastic model

$$e_j = \tilde{g}(e_{j-1}, \hat{\underline{\theta}}) \ , \qquad (4)$$

which is optimal under the condition $\hat{\underline{\theta}} = \underline{\theta}^*$, is no longer correct <u>for prediction purposes</u> due to the fact that $\underline{\theta}^* | \hat{\underline{\theta}}$ is a random vector in a fiducial or Bayesian sense. (Note that (4) may still be correct <u>for describing the past</u>.)

A correct predictive model would be:

$$e_j = \tilde{G}^*(e_{j-1}, \hat{\underline{\Theta}}) = \int_{\text{all } \underline{\Theta}^*} \tilde{g}^*(e_{j-1}, \underline{\Theta}^*) \, d\, F_{\underline{\Theta}^* | \hat{\underline{\Theta}}}(\underline{\Theta}^*), \qquad (5)$$

which accounts for the uncertainty in $\underline{\Theta}^*$ through a modification of the model ($\tilde{G}^* \neq \tilde{g}^*$). For instance, if the probabilistic model $\tilde{g}^*$ is a first-order autoregressive normal sequence with unknown variance, e_j has "Student's" t predictive distribution, for an appropriate choice of the variance estimator (see Chapter III).

The uncertainty of the model parameters, which originates from the finite amount of available data, might be called statistical uncertainty. In prediction, it combines with probabilistic uncertainty through a rule of the type (5). The qualitative result is an increased prediction uncertainty. Statistical uncertainty is reduced by processing more information; in fact this reduces the variance of $(\Theta_i^* - \hat{\Theta}_i)$ if $\hat{\Theta}_i$ is a consistent estimator. In other words, additional information reduces the (statistical) uncertainty on the probabilistic uncertainty which is given, say, by the "correct" but unknown stochastic relation (3).

Consider relaxing the assumption that $\tilde{g}^*$ is known, and replace the optimal model by some consistent estimator $\hat{\tilde{g}}$, e.g., within a class of first-order autoregressive models that contain the Gaussian model as a special case. Then $\hat{\underline{\theta}} \to \underline{\theta}^*$ and $\hat{\tilde{g}} \to \tilde{g}^*$ with probability 1 in the limit case of an infinite amount of available information. Nevertheless, for limited information the model obtained from equation (5) with $\hat{\tilde{g}}$ in place of $\tilde{g}^*$ is not correct for prediction, due to the neglected (say fiducial or Bayesian) uncertainty of $\tilde{g}^* | \hat{\tilde{g}}$. For instance one might accept on statistical ground the hypothesis that $\{e_j\}$ is a normal sequence, but this does not exclude that the true sequence has a different marginal distribution. If γ indexes the possible exact models, one should use the following as a prediction model:

$$e_j = \hat{\hat{G}}(e_{j-1}) = \int_{\text{all } \gamma} \tilde{G}^*_\gamma(e_{j-1}, \hat{\underline{\theta}}_\gamma) d\, F_\gamma(\gamma) \ , \qquad (6)$$

where $\tilde{G}^*_\gamma(\cdot)$ is given by equation (5) with $\tilde{g}^*$, $\underline{\theta}^*$ and $\hat{\underline{\theta}}$ indexed by γ.

The uncertainty on the model has the same <u>statistical</u> nature as the uncertainty on the parameters for a fixed model.

To summarize briefly: the assumption was made that natural systems are deterministic; however the use of descriptive probabilistic models may be necessitated by the complexity of a detailed deterministic analysis. Two types of uncertainties are always present when such models are inferred from a finite set of data: (i) a probabilistic uncertainty, which is inherent in the model itself, due to its being an approximation to the actual deterministic system; (ii) a statistical uncertainty, with origin in the finite amount of available data. Statistical uncertainty has itself two components, one at the level of the model, the other at the level of the parameters. Probabilistic and statistical uncertainties combine in a rather complex way to provide prediction models; see equations (5) and (6).

No essential modification is induced by assuming that natural systems are themselves random, except for the fact that deterministic models should be ruled out. Perfect modeling would produce mathematical analogs with probabilistic, but without statistical uncertainty. In models inferred from limited data statistical and probabilistic uncertainties would again combine, with the net result of increasing the prediction uncertainty.

Under the hypothesis of stochastic natural systems the notion of probabilistic uncertainty is related to what Good (1965) calls "physical" probability and Benjamin and Cornell (1970) call "fundamental" probability; these labels refer to an uncertainty which is intrinsic in the natural mechanism, and therefore irreducible in prediction. However, identity between the present definition of probabilistic uncertainty (which refers to the model) and "physical" or "fundamental" uncertainty (which refers to the natural system) holds conditionally on the model selected corresponding exactly to the stochastic law of nature. This was not the case, for instance, when a process like (1), with no physical or fundamental uncertainty, was approximated by a first-order autoregressive process of the type (2) which, according to the present definition, has nonzero probabilistic uncertainty.

Analyzing uncertainties on models as opposed to doing it directly on nature evades the question whether probability is a constitutive property of nature (this viewpoint is supported by the physical and fundamental probabilities of Good, Benjamin and Cornell), or it is due to our incomplete state of knowledge, as in Laplace's view.

Going one step further one might ask whether subjective probabilities can be viewed also as the superposition of "model" uncertainties and statistical uncertainty about

the correct "model". In this case "model" stands for physiological or psychological structure, being the physical or mental counterpart of a stochastic mathematical model (or class of models, like equation (2)). In this sense, "absolute certainty" (or "equiprobability of exclusive events", etc.) might be just one of such mental "models". Research in this area would be valuable to assess the relative merit (not the correctness, which cannot be established) of the model-uncertainty notion versus the idea of an irreducible uncertainty in nature.

Safety and statistical uncertainty

In the analysis of structural safety it is customary to assume some probabilistic model for the unknown quantities and to neglect statistical uncertainties both in the model and in the parameters, as second-order variations. The arbitrariness of this assumption discredits the claim that a probabilistic approach to safety is more rational than a nonprobabilistic approach. In the former one assumes a probabilistic model; in the latter one assumes a deterministic model; the difference might be called quantitative more than qualitative. The author addressed the problem of rescuing probabilistic methods from the accusation of

arbitrariness in a recent study (Veneziano (1974)). After criticizing the deductive character of the present theory of probabilistic safety, some lines were indicated along which a more general inductive theory might be developed. Although still within a strictly deductive logic, the inclusion of statistical uncertainty in probabilistic models is a step in this direction.

The relevance of statistical uncertainty in state estimation and prediction is studied quantitatively in the following three chapters, with particular emphasis on safety implications. A basic problem in structural safety is that of finding the predictive distribution of unknown quantities (such as resistances and loads) at a future time, when the system performance, say failure or non-failure, is of concern. To solve this problem, a mathematical model is assumed for the time evolution of these unknown quantities (state of the model). The model may be deterministic or stochastic; in any case it contains unknown parameters, which control the evolutive law, or the present state, or both. If no statistical information were available, the predictive distribution of the state (and so the safety of a system depending on it) would be undefined.

This prediction problem has important particular cases and extensions which depend mainly on the choice of the mathematical model.

A very special and important case is when the state evolution in time is assumed deterministic (in particular, state = const.), with unknown present state. Statistical information consists of partial, and/or noisy, and/or censored observations of the state. In "deterministic" models of this type any prediction problem reduces then to a problem of statistical estimation.

In Chapter I, classical estimation theory for noisy observations is reviewed and applied to several problems in the area of Civil Engineering. In the same chapter the theory is extended to censored data, such as those from acceptance tests and from proof loading of structural systems.

The optimal design of estimation experiments is discussed and examplified; in particular, a problem which is given consideration is that of designing optimal sampling networks for the estimation of spatially continuous, unknown state functions with possibly multidimensional parameter (for instance the compressibility of a soil as a function of three geographical coordinates).

Clearly, since the state is unknown but not random, all the estimation uncertainty is statistical in nature, and can be reduced (to zero, in the limit) by collecting additional information.

The simplest models with stochastic time dependence assume that the state evolves at discrete times like a stationary, independent sequence. The marginal distribution type of the sequence may be assumed known with unknown parameters (parametric or parameter-free prediction) or may be itself unknown (nonparametric or distribution-free prediction). In both cases statistical information consists of an independent sample from the same statistical population; again, the sample may be noisy or censored.

For these models the simplest prediction problem asks for the distribution of the state at a future time (say the distribution of the next value in the random sequence), but one can imagine situations in which the joint distribution of many future observations is needed; for instance, when computing the reliability of a system subjected to $m > 1$ independent loading events. Not only does statistical uncertainty cause the predictive distribution to be of different type than that assumed for the sequence, but it also introduces correlation among distinct future observations. Finding the joint predictive distribution of the state at different times is known as a problem of simultaneous (as opposed to "simple") statistical prediction.

Independent models and related prediction problems are studied in Chapter II. For each marginal distribution considered, a penalty measure for statistical uncertainty is defined and tabulated, allowing one to make reliability analyses and designs which account for statistical uncertainty. The penalty may be quite large both in terms of estimated failure probability, and of design safety factors for a given reliability. The penalty increases with the level of reliability and with decreasing statistical information.

An impartial position is held with respect to Classical and Bayesian approaches, although the Bayesian logic seems more natural for handling prediction problems.

A third class of models hypothesizes that the state evolves according to a stochastic process with memory. The same prediction problems which were mentioned for white sequences can be defined here, but the difficulty of finding exact analytical solutions for the predictive distribution increases considerably. In Chapter III attention is restricted to first-order autoregressive models. In this case, when the process is known and the unknown state is estimated from noisy observations one can use the well-known Kalman filter and Kalman prediction algorithms. The

extension of the algorithms to cover censored information is presented and applied to the prediction of resistance parameters in deteriorating systems. Some cases with unknown process parameters are also studied, and a penalty measure for statistical uncertainty is defined, as for the independent models in Chapter II.

REFERENCES

1. Benjamin, J.R. and Cornell, C.A. (1970): Probability, Statistics and Decision for Civil Engineers, Chapter 4, McGraw-Hill, New York.

2. Good, I.T. (1965): The Estimation of Probabilities, Research monograph No. 30, Massachusetts Institute of Technology Press, Cambridge, Massachusetts.

3. de Laplace, P.S.: A Philosophical Essay on Probabilities. Trans. by F.W. Truscott and F.L. Emory, from the 6th French Edition; Chapters 2 and 3, Dover Publications, New York, 1952.

4. Veneziano, D. (1974): "Note on an Inductive Approach to Safety". Submitted to: Proceedings of the American Society of Civil Engineers, J. Structural Division.

CHAPTER I

DETERMINISTIC MODELS

In this chapter we discuss the problem of estimating an unknown but non-random state of nature. The terminology "deterministic models" refers to the evolution of the state in time; the evolution being deterministic, estimation of the present leads automatically to prediction of the future. In structural reliability, data collection and statistical estimation can be used to reduce the uncertainty on natural properties (resistance, elastic moduli, etc.), on soil parameters of structural interest before construction (profile, compressibility, bearing capacity, etc.), on the overall characteristics of elements and subsystems after construction (member and joint stiffnesses, dead loads, geometrical variables, etc.).

"Statistical estimation" is used here in a non standard sense in that the object of estimation are not the parameters of a probabilistic model, but rather the present state of a physical system; the two problems are, however, mathematically identical.

A flexible and general formulation of the estimation problem is one which allows for information being available on the state of nature before experimenting. This leads naturally,

although not necessarily, to Bayesian theory, in which evidence from new data is combined with quantified prior beliefs through Bayes' theorem, yielding quantified posterior beliefs. Most of the results in this chapter are valid under Bayesian inference, but methods of non-Bayesian estimation theory are also considered.

Necessary for applying the techniques of statistical estimation theory is that the state of nature has at least partial observability. In the case of complete non-noisy observation, the state is known with certainty after the experiment; therefore the interest of estimation theory lies in the less trivial cases of noisy or partial non-noisy observations (Sections I.1, I.2, I.3), and of censored, noisy or non-noisy data (Sections I.4, I.5). These are the kinds of data which are available most often, and which are considered in this chapter. However, there is no logical necessity for experimental evidence not to conform to different formats.

The distinction between uncensored and censored data is mathematically relevant, and also corresponds to different goals of experimentation. Uncensored data are collected by devices which measure directly the quantities of interest, although possibly with disturbances. Examples are the measurement of mechanical properties of materials sampled from a structure after construction and the laboratory tests on soil specimens. In both cases experimental data are in

the form of direct point measurements of the state of nature, which is an unknown function of one or more natural variables. Local disturbances in the system caused by the sampling procedure are generally neglected. Even with non-noisy measurements (which is rarely a realistic hypothesis) the state of the system at nearby points remains uncertain. Making "best guesses" and quantifying the residual uncertainties are among the goals of estimation theory.

In other cases data are collected in a censored format. Typical is the proof loading of a structure under prefixed loading conditions. In general the system survives, which fact insures a minimum resistance level via truncation of the resistance distribution. The reasons for collecting information of this type are both to increase the reliability of those systems which survived the experiment, and to detect possible causes of future disastrous failures by observing an anomalous behavior. In some cases this is also a way of demonstrating the meeting of legal requirements for minimum resistance. An obvious economic advantage of proof loading over direct methods of measuring resistance is that the experiment is destructive selectively, with survival of those systems whose performance under experimentation was satisfactory.

The organization of the chapter is as follows. The first three sections deal with estimation from uncensored

data. After reviewing basic results of Bayesian and Classical estimation theory (Section I.1) some criteria of experiment evaluation are introduced in Section I.2. The evaluation criteria allow one to assign informativeness measures to a given experiment and to compare a priori different experiments. This leads naturally to problems of optimal experiment design which are discussed briefly in Section I.3. Examples are used both to illustrate possible applications of the quoted results to (structural) reliability problems, and to suggest developments in directions of engineering interest. In the latter sense they may appear to be incomplete, as indeed they are.

I.1 ELEMENTS OF BAYESIAN AND NON-BAYESIAN ESTIMATION THEORY; UNCENSORED DATA

Consider a system with an unknown finite dimensional state vector $\underline{X} = [X_1, X_2, \ldots, X_n]'$ (for instance, X_i may be the resistance of the ith structural member or the compressibility of the ith soil stratum). In many cases $\underline{X}$ is defined from a continuous and possibly vector-valued state function $X(\cdot)$ through appropriate discretization. Another meaning that $\underline{X}$ may have is as a collection of parameters in a polynomial approximation to $X(\cdot)$. However, the analysis which follows applies also (in the limit) to the full function $X(\cdot)$. For

this, $\underline{X}$ collects the values of $X(\cdot)$ at a finite number of points. Since the length of $\underline{X}$ and the points at which $X(\cdot)$ is sampled by $\underline{X}$ are arbitrary, this covers in the limit the case of continuous estimation. Since $\underline{X}$ is not random the uncertainty on its actual value is purely statistical (see Introduction). The purpose of making experiments is to reduce this uncertainty by providing additional information. In Section I.1 we give some state estimators and the uncertainty in the state after experimentation. This is done concisely, without proofs. Proofs can be found in classical texts on estimation theory, which the reader is referred to. For applications see Sections I.2 and I.3.

I.1.1 Bayesian and non-Bayesian estimators

The problem in this paragraph is to define common rules (estimators) for obtaining point estimates of $\underline{X}$ from a given amount of information, including in particular the observations of experiments.

The ingredients of a statistical estimation rule are:

(i) the prior uncertainty on the state vector $\underline{X}$;

(ii) the observation model;

(iii) the estimation criterion.

In Bayesian estimation theory the first two elements define the posterior probability distribution of $\underline{X}$. Let the prior probabilities be given in the form of the PDF $f(\underline{X})$,

and denote $l(\underline{Z}|\underline{X}) \propto f(\underline{Z}|\underline{X})$ the likelihood function corresponding to the measurement model ($\underline{Z}$ = vector of measurements). Then, from Bayes' theorem, the posterior PDF of the state is:

$$f(\underline{X}|\underline{Z}) \propto f(\underline{X}) \cdot l(\underline{Z}|\underline{X}) \qquad \text{(I.1)}$$

$f(\underline{X}|\underline{Z})$ is an exhaustive descriptor of the state of knowledge after experimentation and can be used for posterior reliability evaluations. If desired, posterior confidence regions (regions in state space which contain $\underline{X}$ with a given probability) can be found from it. Also, given the error loss function

$$L(\tilde{\underline{X}}) = L[\underline{X} - \hat{\underline{X}}(\underline{Z})];$$

$\underline{X}$ = true state,

$\hat{\underline{X}}(\underline{Z})$ = posterior estimate of $\underline{X}$,

$\tilde{\underline{X}} = \underline{X} - \hat{\underline{X}}(\underline{Z})$, error,

one can find an optimal estimator $\hat{\underline{X}}(\cdot)$ as any law which minimizes the Bayes risk (= the a priori expected loss):

$$B[\hat{\underline{X}}(\cdot)] = \underset{\underline{X},\underline{Z}}{E}[L(\tilde{\underline{X}})] = \int_{\text{all } \underline{X}} \int_{\text{all } \underline{Z}} L[\underline{X}-\hat{\underline{X}}(\underline{Z})] \cdot f(\underline{X},\underline{Z})\, d\underline{Z}\, d\underline{X}.$$

After writing $f(\underline{X},\underline{Z}) = f(\underline{X}|\underline{Z}) \cdot f(\underline{Z})$ it is easy to see that the same estimator minimizes the conditional Bayes risk (= the posterior expected loss):

$$B[\hat{\underline{X}}(\underline{Z})|\underline{Z}] = E_{\underline{X}}[L(\tilde{\underline{X}})|\underline{Z}] = \int_{\text{all } \underline{X}} L[\underline{X}-\hat{\underline{X}}(\underline{Z})]\, f(\underline{X}|\underline{Z})\, d\underline{X}.$$

There is no unique, best choice for the estimator of $\underline{X}$. Different choices of $L[\cdot]$ produce different optimal estimators. Commonly used Bayes loss functions and the associated optimal point estimators are:

(i) $L(\tilde{\underline{X}}) = ||\tilde{\underline{X}}||^2_{\underline{S}}$ (Squared error loss function),

where $||\tilde{\underline{X}}||^2_{\underline{S}}$ denotes the generalized squared norm of the estimation error with respect to the symmetric positive matrix $\underline{S}$:

$$||\tilde{\underline{X}}||^2_{\underline{S}} = [\underline{X}-\hat{\underline{X}}(\underline{Z})]'\underline{S}[\underline{X}-\hat{\underline{X}}(\underline{Z})].$$

The associated optimal estimator is called the minimum-variance-of-error estimator, denoted $\hat{\underline{X}}_{MV}(\cdot)$. It is easy to show (see for instance Sage and Melsa (1971), Par. 6.2) that $\hat{\underline{X}}_{MV}(\underline{Z})$ coincides

with the posterior mean estimator $\hat{\underline{X}}_{PM}(\cdot)$, and therefore is independent of $\underline{S}$.

(ii) $$L(\tilde{\underline{X}}) = \lim_{\varepsilon \to 0} \begin{cases} 0 & \text{for } ||\tilde{\underline{X}}|| = (\tilde{\underline{X}}'\tilde{\underline{X}})^{1/2} < \varepsilon/2 \\ 1/\varepsilon & \text{for } ||\tilde{\underline{X}}|| \geq \varepsilon/2 \end{cases}$$

(limit uniform cost function)

The conditional Bayes risk attains its minimum value when $\hat{\underline{X}}(\cdot)$ satisfies

$$\frac{\partial f(\underline{X}|\underline{Z})}{\partial \underline{X}} \Bigg|_{\underline{X} = \hat{\underline{X}}(\underline{Z})} = \underline{0},$$

i.e., when $\hat{\underline{X}}(\underline{Z})$ equals the <u>maximum posterior estimate</u>, $\hat{\underline{X}}_{MAP}(\underline{Z})$.

(iii) $L(\tilde{\underline{X}}) = \max_{\underline{S}} ||\tilde{\underline{X}}||$ (maximum-absolute-value-of-error loss function)

This criterion is applicable when $f(\underline{X}|\underline{Z}) \equiv 0$ outside a bounded region, in which case the optimal estimator is the <u>minimax estimator</u>, also called the <u>minimum-error estimator</u>.

(iv) $L(\tilde{\underline{X}}) = ||\tilde{\underline{X}}||_{\underline{S}}$ (absolute-value-of-error loss function)

When $\underline{X}$ is a scalar the associated optimum estimator $\hat{X}_{ABS}(\cdot)$ corresponds to a median value of the posterior density. Analogous results do not exist for the multidimensional case, since there is no standard definition of the median of multivariate distributions.

In general these Bayesian point estimators differ the one from the others. However, for symmetrical and unimodal posterior densities it is true that: $\hat{\underline{X}}_{MV}(\cdot) = \hat{\underline{X}}_{MAP}(\cdot) = \hat{\underline{X}}_{ABS}(\cdot)$.

All the loss functions considered above are "homogeneous" in the sense that they penalize the error vector $\tilde{\underline{X}}$ irrespective of the actual value of the state vector $\underline{X}$ (and of the estimate $\hat{\underline{X}}(\underline{Z}) = \underline{X} - \tilde{\underline{X}}$). Although this is common in estimation theory and leads to statistically meaningful "best" estimates of $\underline{X}$ (like the posterior mean, the maximum posterior estimate, and so on), it is not a necessity of Bayesian estimation. One can easily think of estimation problems in which $L(\cdot)$ is markedly nonhomogeneous, reflecting the unequal consequences of an estimation error for different true states of nature.

Some non-Bayesian point estimators

(v) Linear-minimum-variance estimator

Let $\underset{\sim}{S}_{\underline{X}} = E[(\underline{X} - \hat{\underline{X}}(\underline{Z}))(\underline{X} - \hat{\underline{X}}(\underline{Z}))']$ be the second moment matrix of the estimate error. An estimator which minimizes the trace of $\underset{\sim}{S}_{\underline{X}}$ is called a minimum-mean-square-error estimator. The Gauss-Markov theorem states that if $\underline{X}$ and $\underline{Z}$ are random vectors with second moment matrices:

$$E[\underline{X}\ \underline{X}'] = \underset{\underline{X}}{S} ,$$

$$E[\underline{Z}\ \underline{Z}'] = \underset{\underline{Z}}{S} ,$$

$$E[\underline{X}\ \underline{Z}'] = \underset{\underline{X}\ \underline{Z}}{S} ,$$

and if $\det \underset{\underline{Z}}{S} \neq 0$, the linear (in $\underline{Z}$) minimum-mean-square-error estimator $\hat{\underline{X}}(\cdot)$ is:

$$\hat{\underline{X}}(\underline{Z}) = \underset{\underline{X}\ \underline{Z}}{S}\ \underset{\underline{Z}}{S}^{-1}\ \underline{Z} , \qquad \text{(I.2)}$$

with associated second moment error matrix:

$$E[\tilde{\underline{X}}\ \tilde{\underline{X}}'] = E[(\underline{X}-\hat{\underline{X}})(\underline{X}-\hat{\underline{X}})'] = \underset{\underline{X}}{\underset{\sim}{S}} = \underset{\underline{X}}{S} - \underset{\underline{X}\underline{Z}}{S}\ \underset{\underline{Z}}{S}^{-1}\ \underset{\underline{X}\underline{Z}}{S}' . \qquad \text{(I.3)}$$

Also, if the condition for unbiasedness

$$E[\hat{\underline{X}}] = \underline{S}_{\underline{X}\,\underline{Z}}\,\underline{S}_{\underline{Z}}^{-1}\,E[\underline{Z}] \quad (= E[\underline{X}]) \tag{I.4}$$

is satisfied, then $\hat{\underline{X}}$ is the (unbiased) linear-minimum-variance estimator, denoted $\hat{\underline{X}}_{LMV}(\cdot)$. For a proof of the theorem, see Liebelt (1967), Par. 5.3.

(vi) Maximum likelihood estimator

When the conditional measurement density $f(\underline{Z}|\underline{X})$ is known, the maximum likelihood estimator $\hat{\underline{X}}_{ML}(\underline{Z})$ is defined to be the vector which maximizes $f(\underline{Z}|\underline{X})$. Under conditions of differentiability and unimodality a necessary and sufficient condition for $\hat{\underline{X}}_{ML}(\underline{Z})$ is:

$$\frac{\partial\, F(\underline{Z}|\underline{X})}{\partial\, \underline{X}}\,\Bigg|_{\underline{X} = \hat{\underline{X}}_{ML}(\underline{Z})} = \underline{0} \tag{I.5}$$

This estimator is particularly appealing when no a priori information on $\underline{X}$ is available, or when the posterior Bayesian density $f(\underline{X}|\underline{Z})$ is difficult to compute.

(vii) Least-squares estimators

Consider the measurement model

$$\underline{Z} = g(\underline{X}) + \underline{\varepsilon}$$

where $\underline{\varepsilon}$ is a zero-mean random noise vector. A least-squares estimator $\hat{\underline{X}}_{LS}(\cdot)$ can be defined by the quadratic condition:

$$J[\hat{\underline{X}}(\underline{Z})] = \{\underline{Z} - g[\hat{\underline{X}}(\underline{Z})]\}'\underline{R}\{\underline{Z} - g[\hat{\underline{X}}(\underline{Z})]\} = \min_{\hat{\underline{X}}(\underline{Z})} ,$$

where $\underline{R} = \underline{R}' > 0$.

A more general form of the optimality criterion for least-squares estimators will be considered at the end of the next paragraph.

I.1.2 Linear observation model; measurement error covariance known

The measurement vector $\underline{Z}$ being a stochastic linear function of $\underline{X}$ defines a particularly simple and important case. From what will be said in this paragraph the linear observation model is mathematically very convenient and should be used always when meaningful in the physical context. It is also the obvious model when $\underline{Z}$ measures $\underline{X}$ directly.

Let

$$\underline{Z} = \underline{Z}_o + \underline{H}\,\underline{X} + \underline{\varepsilon} \quad , \tag{I.6}$$

where $\underline{Z}_o$ is a given m-vector, $\underline{H}$ is a given (m×n) measurement matrix and $\underline{\varepsilon}$ is a zero-mean noise vector. The posterior Bayesian density is given by equation (I.1), where $(\underline{Z}|\underline{X})$ follows the same distribution type as $\underline{\varepsilon}$, differing only in the location parameter. The special case when both the prior density $f(\underline{X})$ and the noise density $f(\underline{\varepsilon})$ are normal is of considerable practical and theoretical interest. Let

$$\underline{X} \sim N_n(\underline{X}_b;\underline{\Sigma}_b) \quad , \quad \underline{\varepsilon} \sim N_m(\underline{0};\underline{\Theta}) \quad ,$$

$$\mathrm{Cov}[\underline{X},\underline{\varepsilon}] = \underline{0} \quad .$$

Then, from equation (I.6):

$$(\underline{Z}|\underline{X}) \sim N_m(\underline{Z}_o + \underline{H}\,\underline{X};\underline{\Theta}) \quad ,$$

$$\underline{Z} \sim N_m(\underline{Z}_o + \underline{H}\,\underline{X}_b;\ \underline{H}\,\underline{\Sigma}_b\,\underline{H}' + \underline{\Theta})$$

From direct substitution into equation (I.1) one finds (see for instance Bryson and Ho (1969), Par. 12.7) that the posterior distribution of $\underline{X}$ is also Gaussian with parameters:

$$E[\underline{X}|\underline{Z}] = \underline{X}_a = \underline{\Sigma}_a\{\underline{H}'\ \underline{\Theta}^{-1}(\underline{Z} - \underline{Z}_o) + \underline{\Sigma}_b^{-1}\underline{X}_b\} \qquad (a)$$

(I.7)

$$= \underline{X}_b + \underline{\Sigma}_a\ \underline{H}'\ \underline{\Theta}^{-1}(\underline{Z} - \underline{Z}_o - \underline{H}\ \underline{X}_b) \qquad (b)$$

$$= \underline{X}_b + \underline{\Sigma}_b\ \underline{H}'(\underline{H}\ \underline{\Sigma}_b\ \underline{H}' + \underline{\Theta})^{-1}(\underline{Z} - \underline{Z}_o - \underline{H}\ \underline{X}_b); \qquad (c)$$

and

$$\text{Var}\ [\underline{X}|\underline{Z}] = \underline{\Sigma}_a = (\underline{\Sigma}_b^{-1} + \underline{H}'\ \underline{\Theta}^{-1}\underline{H})^{-1} \qquad (a)$$

(I.8)

$$= \underline{\Sigma}_b - \underline{\Sigma}_b\ \underline{H}'(\underline{\Theta} + \underline{H}\ \underline{\Sigma}_b\ \underline{H}')^{-1}\underline{H}\ \underline{\Sigma}_b\ , \qquad (b)$$

where the indicated inverses are assumed to exist. The last identity in equations (I.8) comes directly from the matrix inversion lemma, see Householder (1953). For the case of present interest the lemma proves the equality:

$$(\underline{A} + \underline{B}'\ \underline{C}\ \underline{B})^{-1} = \underline{A}^{-1} - \underline{A}^{-1}\ \underline{B}'(\underline{B}\ \underline{A}^{-1}\ \underline{B}' + \underline{C}^{-1})^{-1}\ \underline{B}\ \underline{A}^{-1}\ .$$

Again, the indicated inverses are assumed to exist (for the use of generalized inverses in the foregoing results, see Deutsch (1965) Par. 7.2).

The relative advantage of using either of equations (I.8) comes from comparing the dimensions of the matrices under the inversion signs, which are the same as the

dimensions of $\underline{X}$ and $\underline{Z}$. This point is discussed more below.

For the special case of a scalar state and a scalar noisy measurement (n = m = 1); i.e., for

$$Z = Z_o + h\,X + \varepsilon;\; X \sim N(X_b;\, \sigma_b^2);\; \varepsilon \sim N(0;\, \sigma_\varepsilon^2);\; \mathrm{Cov}[X, \varepsilon] = 0,$$

equations (I.7b) and (I.8) become:

$$X_a = X_b + h\,\frac{\sigma_a^2}{\sigma_\varepsilon^2}\,(Z - Z_o - h\,X_b)\;; \qquad \text{(I.9)}$$

$$\sigma_a^2 = \left(\frac{1}{\sigma_b^2} + h^2\,\frac{1}{\sigma_\varepsilon^2}\right)^{-1} = \frac{\sigma_\varepsilon^2}{h^2 + \sigma_\varepsilon^2/\sigma_b^2} \qquad \text{(I.10)}$$

Equations (I.7) and (I.8) play an important role in Bayesian and Classical linear estimation theory, and will be used repeatedly in the sequel. It seems therefore fruitful to pause for some comments.

The posterior mean (I.7b) or (I.7c) is a linear function of the "innovation vector" $(\underline{Z} - \underline{Z}_o - \underline{H}\,\underline{X}_b)$, meaning that the difference $(\underline{X}_a - \underline{X}_b)$ is proportional to the deviation of the observations from their prior expected value $(\underline{Z}_o + \underline{H}\,\underline{X}_b)$; see equation (I.6). The revision of the state mean, $(\underline{X}_a - \underline{X}_b)$, also depends on the matrix product $\underline{\Sigma}_a\,\underline{H}'\,\underline{\Theta}^{-1}$ (on $h\,\frac{\sigma_a^2}{\sigma_\varepsilon^2}$ in the scalar case), being larger for a smaller

error covariance matrix $\underline{\Theta}^{(*)}$. This means that measurements which are statistically affected by small errors are given more credibility than measurements with highly dispersed errors. The difference $(\underline{X}_a - \underline{X}_b)$ is also "proportional" to the signal-to-noise strength which is measured by the observation or modulation matrix $\underline{H}$; see equation (I.6). Finally the presence of $\underline{\Sigma}_a$ in equation (I.7b) indicates that Bayes' rule weights more the observations when the prior uncertainty about the value of $\underline{X}$ is larger (when $\underline{\Sigma}_a$ is larger); a statement which makes good sense.

Passing to the posterior error covariance, equations (I.8), if $\underline{\Theta}$ is positive definite, $\underline{H}' \, \underline{\Theta}^{-1} \, \underline{H}$ is at least positive semidefinite, so that $\underline{\Sigma}_a^{-1} \geq \underline{\Sigma}_b^{-1}$ (meaning that $\underline{\Sigma}_a^{-1} - \underline{\Sigma}_b^{-1}$ is positive semidefinite). Assuming $\underline{\Sigma}_b > 0$, the same is true for $\underline{\Sigma}_b^{-1}$ and $\underline{\Sigma}_a \leq \underline{\Sigma}_b$ (meaning that $\underline{\Sigma}_b - \underline{\Sigma}_a$ is positive semidefinite). The foregoing inequalities state formally that the uncertainty does not increase by processing more information, again an intuitively obvious fact.

An important feature of equations (I.7) and I.8) is that while the computation of the posterior mean requires knowing the outcome $\underline{Z}$ of the experiment, the posterior error covariance matrix does <u>not</u> depend on the numerical value

---

(*) If $\underline{A}$ and $\underline{B}$ are square matrices of the same dimension we say that $\underline{A}$ is larger than $\underline{B}$, written $\underline{A} > \underline{B}$ if the difference $(\underline{A} - \underline{B})$ is positive definite.

of the observations, and therefore is precomputable (before the experiment).

If $\underline{\Theta} = \text{diag}\,[\sigma^2_{\varepsilon_i}]$ equation (I.8b) presents the practical advantage over equation (I.8a) that one can process recursively one scalar observation at a time, thus reducing the matrix inversion to the trivial operation of inverting scalars. Another advantage of using equation (I.8b) is that when $\underline{\Sigma}_b^{-1}$ is positive semidefinite (which corresponds to the case of partially "diffuse" prior) the matrix $(\underline{\Sigma}_b^{-1} + \underline{H}\ \underline{\Theta}^{-1}\underline{H})$ may be singular. In any case, numerical difficulties are expected in the inversion of $\underline{\Sigma}_a^{-1}$ if both the prior and the observations carry little information on some linear combination of the state variables.

Intuitively speaking, if the absolute entries of $\underline{H}$ are not much larger than 1 and the prior distribution has little dispersion as compared to the measurement noise covariance, $\underline{\Sigma}_b^{-1}$ is "large" with respect to $\underline{H}'\ \underline{\Theta}^{-1}\ \underline{H}$, and $\underline{\Sigma}_a \approx \underline{\Sigma}_b$. This is immediately clear in the scalar case. Also, since the product $\underline{\Sigma}_a\ \underline{H}'\ \underline{\Theta}^{-1}$ is "small", it is $\underline{X}_a \tilde{\simeq} \underline{X}_b$, meaning that little is learned from measurements with large noise perturbations.

At the other extremum, if the prior uncertainty is "large" when compared to the covariance of the measurement error (accurate experiments), it is $\underline{\Sigma}_a \approx (\underline{H}'\ \underline{\Theta}^{-1}\ \underline{H})^{-1}$ (assume existance of the inverse) and, from equation (I.7b):

$$\underline{X}_a \approx \underline{X}_b + (\underline{H}'\underline{\Theta}^{-1}\underline{H})^{-1}\ \underline{H}'\ \underline{\Theta}^{-1}(\underline{Z}-\underline{Z}_o)-(\underline{H}'\underline{\Theta}^{-1}\underline{H})^{-1}\ \underline{H}'\ \underline{\Theta}^{-1}\ \underline{H}\ \underline{X}_b$$

$$= (\underline{H}'\underline{\Theta}^{-1}\underline{H})^{-1}\ \underline{H}'\ \underline{\Theta}^{-1}(\underline{Z}-\underline{Z}_o)$$

$$= \underline{H}'(\underline{H}\ \underline{H}')^{-1}\ (\underline{Z}-\underline{Z}_o) \quad .$$

The last equality holds under the assumption: det $(\underline{H}\underline{H}')\neq 0$. The point of this derivation is to show that for accurate experiments prior beliefs receive little weight in the posterior probabilities. (In processing accurate measurements one should use equations (I.7c) and (I.8c), which do not require inverting $\underline{\Theta}$).

From equations (I.8a) and (I.10) it is seen that the critical comparison which decides on the effectiveness of an experiment is between $\underline{\Sigma}_b^{-1}$ (or $1/\sigma_b^2$) and $\underline{H}'\ \underline{\Theta}^{-1}\ \underline{H}$ (or h^2/σ_ε^2). The latter matrix product, denoted $\underline{M}$, is known as the Fisher information matrix of the experiment. It is easy to show that $\underline{M}$ is positive semidefinite. In fact, after assuming $\underline{\Theta} \geq 0$, one may write:

$$\underline{M} = \underline{H}'\ \underline{\Theta}^{-1}\ \underline{H} = (\underline{\Theta}^{-1/2}\ \underline{H})'\ (\underline{\Theta}^{-1/2}\ \underline{H}) \quad .$$

Since any matrix of the form $\underline{A}'\ \underline{A}$ is positive semidefinite, the property $\underline{M} \geq 0$ follows.

Two generalizations of equations (I.7) and (I.8) are indicated next; they concern the case of correlation between the state $\underline{X}$ and the measurement errors $\underline{\varepsilon}$, and the case of repeated observations from possibly different linear models. Finally, equations (I.7) and (I.8) are specialized for a noninformative ("diffuse") prior.

(i) Correlation between the state vector and the measurement errors

Uncorrelation between $\underline{X}$ and $\underline{\varepsilon}$ is a rather "robust" assumption, in the sense of being justified in most problems of state estimation. However, depending on the modalities of the experiment, there may be cases in which this is not true.

When the measurement error vector $\underline{\varepsilon}$ and the state vector $\underline{X}$ are correlated a priori, say $\mathrm{Cov}[\underline{X}, \underline{\varepsilon}] = \underline{\Gamma} \neq \underline{0}$, some easy algebra generalizes equations (I.7c) and (I.8b) into:

$$\underline{X}_a = \underline{X}_b + (\underline{\Sigma}_b \underline{H}' + \underline{\Gamma})(\underline{H} \underline{\Sigma}_b \underline{H}' + \underline{H} \underline{\Gamma} + \underline{\Gamma}' \underline{H}' + \underline{\Theta})^{-1} (\underline{Z} - \underline{Z}_o - \underline{H} \underline{X}_b); \quad \text{(I.11)}$$

$$\underline{\Sigma}_a = \underline{\Sigma}_b - (\underline{\Sigma}_b \underline{H}' + \underline{\Gamma})(\underline{H} \underline{\Sigma}_b \underline{H}' + \underline{H} \underline{\Gamma} + \underline{\Gamma}' \underline{H}' + \underline{\Theta})^{-1} (\underline{H} \underline{\Sigma}_b + \underline{\Gamma}'). \quad \text{(I.12)}$$

For a scalar state variable and $\mathrm{Cov}[X, \varepsilon] = \gamma$ these equations yield:

$$X_a = X_b + \frac{h + \gamma/\sigma_b^2}{h^2 + 2h\gamma/\sigma_b^2 + \sigma_\varepsilon^2/\sigma_b^2} (Z - Z_o - h X_b), \quad \text{(I.13)}$$

$$\sigma_a^2 = \sigma_b^2 \left[1 - \frac{(h + \gamma/\sigma_b^2)^2}{h^2 + 2h\gamma/\sigma_b^2 + \sigma_\varepsilon^2/\sigma_b^2} \right] . \quad \text{(I.14)}$$

(ii) Repeated linear measurements

A second easy generalization of equations (I.7) and (I.8) concerns the case of multiple observations according to the model:

$$\underline{Z}_i = \underline{Z}_{o_i} + \underline{H}_i \underline{X} + \underline{\varepsilon}_i \; ; \quad i = 1,2,\ldots,K \; , \quad \text{(I.15)}$$

where $\underline{Z}_i$ is an m_i-measurement vector, $\underline{H}_i$ is an $(m_i \times n)$ observation matrix, $\underline{Z}_{o_i}$ is a given vector and the $\underline{\varepsilon}_i$'s are noise vectors with first and second-moments:

$$E[\underline{\varepsilon}_i] = \underline{0} \; ,$$
$$Cov[\underline{\varepsilon}_i, \underline{\varepsilon}_j] = \underline{\Theta}_i \, \delta_{ij} \; ,$$
$$Cov[\underline{\varepsilon}_i, \underline{X}] = \underline{0} \; .$$

Define:

$$\underline{Z} = \begin{pmatrix} \underline{Z}_1 \\ \vdots \\ \underline{Z}_K \end{pmatrix} ; \; \underline{Z}_o = \begin{pmatrix} \underline{Z}_{o_1} \\ \vdots \\ \underline{Z}_{o_K} \end{pmatrix} ; \; \underline{H} = \begin{pmatrix} \underline{H}_1 \\ \vdots \\ \underline{H}_K \end{pmatrix} ; \; \underline{\varepsilon} = \begin{pmatrix} \underline{\varepsilon}_1 \\ \vdots \\ \underline{\varepsilon}_K \end{pmatrix} \; .$$

Then:

$$E[\underline{\varepsilon}] = \underline{0} \; ; \quad Cov[\underline{\varepsilon},\underline{\varepsilon}] = \underline{\Theta} = diag[\underline{\Theta}_i] \quad .$$

Entering equations (I.7a,b) and (I.8a) with the foregoing notations yields:

$$\underline{X}_a = \underline{\Sigma}_a \left(\sum_{i=1}^{K} \underline{H}_i' \, \underline{\Theta}_i^{-1} \, (\underline{Z}_i - \underline{Z}_{o_i} + \underline{\Sigma}_b^{-1} \, \underline{X}_b) \right) \qquad (a)$$

(I.16)

$$= \underline{X}_b + \underline{\Sigma}_a \left(\sum_{i=1}^{K} \underline{H}_i \, \underline{\Theta}_i^{-1} \, (\underline{Z}_i - \underline{Z}_{o_i} - \underline{H}_i \, \underline{X}_b) \right) ; (b)$$

$$\underline{\Sigma}_a = (\underline{\Sigma}_b^{-1} + \sum_{i=1}^{K} \underline{H}_i' \, \underline{\Theta}_i^{-1} \, \underline{H}_i)^{-1} \quad . \qquad (I.17)$$

Alternatively (and with operational advantages) the equivalents of equations (I.7c) and (I.8b) can be written in recursive form:

$$\underline{X}_{a_i} = \underline{X}_{b_i} + \underline{\Sigma}_{b_i} \, \underline{H}_i' (\underline{H}_i \, \underline{\Sigma}_{b_i} \, \underline{H}_i' + \underline{\Theta}_i)^{-1} (\underline{Z}_i - \underline{Z}_{o_i} - \underline{H}_i \, \underline{X}_{b_i}) \; ; \; i=1,\ldots,K,$$

$$\underline{X}_{b_i} = \underline{X}_{a_{i-1}} \; ; \; i=2,\ldots,K,$$

(I.18)

$$\underline{\Sigma}_{a_i} = \underline{\Sigma}_{b_i} - \underline{\Sigma}_{b_i} \, \underline{H}_i' (\underline{\Theta}_i + \underline{H}_i \, \underline{\Sigma}_{b_i} \, \underline{H}')^{-1} \underline{H}_i \, \underline{\Sigma}_{b_i} \; ; \; i=1,\ldots,K,$$

$$\underline{\Sigma}_{b_i} = \underline{\Sigma}_{a_{i-1}} \quad ; \quad i=2,\ldots,K, \tag{I.18}$$

with given initial conditions: $\underline{X}_{b_1} = \underline{X}_b$; $\underline{\Sigma}_{b_1} = \underline{\Sigma}_b$.

(iii) Diffuse prior state probabilities

The prior distribution of $\underline{X}$ is said to be diffuse if $f(\underline{X}) \propto$ constant. Then from equation (I.1):

$$f(\underline{X}|\underline{Z}) \propto l(\underline{Z}|\underline{X})$$

$$= (2\pi)^{-n/2} \, |\underline{\Theta}|^{-1/2} \exp\{-\tfrac{1}{2}(\underline{Z}-\underline{Z}_o-\underline{H}\,\underline{X})'\underline{\Theta}^{-1}(\underline{Z}-\underline{Z}_o-\underline{H}\,\underline{X})\}. \tag{I.19}$$

By direct substitution one can verify the following identy for the quadratic form in the exponent of equation (I.19):

$$(\underline{Z}-\underline{Z}_o-\underline{H}\,\underline{X})'\underline{\Theta}^{-1}(\underline{Z}-\underline{Z}_o-\underline{H}\,\underline{X}) = (\underline{Z}-\hat{\underline{Z}})'\underline{\Theta}^{-1}(\underline{Z}-\hat{\underline{Z}})+(\underline{X}-\hat{\underline{X}})'\underline{H}'\underline{\Theta}^{-1}\underline{H}(\underline{X}-\hat{\underline{X}}) \ ,$$

where

$$\hat{\underline{X}} = (\underline{H}'\underline{\Theta}^{-1}\underline{H})^{-1}\underline{H}'\underline{\Theta}^{-1}(\underline{Z}-\underline{Z}_o) \ ,$$

$$\hat{\underline{Z}} = \underline{H}\,\hat{\underline{X}} + \underline{Z}_o$$

and the Fisher information matrix $\underline{M} = \underline{H}'\underline{\Theta}^{-1}\underline{H}$ is assumed to be positive definite. Since $(\underline{Z}-\hat{\underline{Z}})'\underline{\Theta}^{-1}(\underline{Z}-\hat{\underline{Z}})$ is not a function

of $\underline{X}$ for given $\underline{Z}$, we conclude:

$$f(\underline{X}|\underline{Z}) \propto \exp\{-\frac{1}{2}(\underline{X}-\hat{\underline{X}})'\underline{H}'\underline{\Theta}^{-1}\underline{H}(\underline{X}-\hat{\underline{X}})\}, \qquad (I.20)$$

meaning that $(\underline{X}|\underline{Z})$ has normal distribution with parameters:

$$E[\underline{X}|\underline{Z}] = \underline{X}_a = \hat{\underline{X}}$$

$$= (\underline{H}'\underline{\Theta}^{-1}\underline{H})^{-1}\underline{H}\,\underline{\Theta}^{-1}(\underline{Z}-\underline{Z}_o) = \underline{H}'(\underline{H}\,\underline{H}')^{-1}(\underline{Z}-\underline{Z}_o) \qquad (I.21)$$

$$Var[\underline{X}|\underline{Z}] = \underline{\Sigma}_a = (\underline{H}'\underline{\Theta}^{-1}\underline{H})^{-1} . \qquad (I.22)$$

(The last equality in equation (I.21) requires the condition $\det(\underline{H}\,\underline{H}') \neq 0$). Note that analogous results were found above as approximations to the case of informative prior and accurate measurements. For $\underline{Z}_o = \underline{0}$ and for measurement errors having spherical normal distribution ($\underline{\Theta} = \sigma^2\underline{I}_m$) equation (I.20) reproduces a well-known result in linear regression theory (see Box and Tiao (1973) page 115).

$$f(\underline{X}|\underline{Z}) = (\sqrt{2\pi}\cdot\sigma)^{-n}|\underline{H}'\underline{H}|^{1/2}\exp\{-\frac{1}{2}(\underline{X}-\hat{\underline{X}})'\underline{H}'\underline{H}(\underline{X}-\hat{\underline{X}})\} ,$$

where now $\underline{X} = (\underline{H}'\underline{H})^{-1}\underline{H}'\underline{Z}$.

The posterior parameters, equations (I.21) and (I.22), could

also be derived by letting $\underline{\Sigma}_b^{-1} = \underline{0}$ in equations (I.7) and (I.8). The posterior covariance coincides with the inverse of the information matrix, and the posterior mean is independent of the measurement dispersion.

For the linear Gaussian model considered above $\underline{X}_a$ is the best Bayesian estimator under several criteria: posterior mean ($\hat{\underline{X}}_{PM} = \underline{X}_a$); maximum posterior estimate ($\hat{\underline{X}}_{MAP} = \underline{X}_a$); absolute-value-of-error ($\hat{\underline{X}}_{ABS} = \underline{X}_a$). For all these cases the Bayesian estimator error covariance is $\underline{\Sigma}_a$.

If a priori the first two moments of $\underline{X}$ are known, but not its full prior probability distribution, the estimators (I.7) and (I.11) no longer possess the properties that we found for the Gaussian model. However, the estimators (I.7) and (I.11) are attractive because they use only the first two moments of $\underline{X}$ and $\underline{\varepsilon}$, and because the resulting error covariance is independent of the measurement vector $\underline{Z}$, and therefore is precomputable. Later in this chapter we shall see that they correspond to the linear-minimum-mean-square-of-error estimator. For non-normal prior and/or non-normal measurement error some nonlinear estimators may exist with lower estimate error mean square. Under a second moment characterization of $\underline{X}$ and $\underline{\varepsilon}$ the estimators (I.7) and (I.11) are often called pseudo-Bayes estimators.

A brief discussion of non-Bayesian point estimators for the linear measurement model (I.6) follows.

Some non-Bayesian linear estimators

(i) Linear-minimum-variance estimator

The Gauss-Markov theorem (cfr. equations (I.2) and (I.3)) does not require any functional relationship between $\underline{Z}$ and $\underline{X}$. Here we specialize the claim of the theorem for the linear observation model (I.6), for which we define:

$$\underline{Z}^* = \underline{Z}-\underline{Z}_o-\underline{H}\ \underline{X}_b \ ; \quad \underline{X}^* = \underline{X} - \underline{X}_b$$

Then, in the notations which precede equation (I.2):

$$\underline{S}_{\underline{X}^*\underline{Z}^*} = \mathrm{Cov}[\underline{X},\underline{Z}] = \underline{\Sigma}_b\ \underline{H}'+\underline{\Gamma} \quad ,$$

$$\underline{S}_{\underline{Z}^*} = \mathrm{Var}[\underline{Z}] = \underline{H}\ \underline{\Sigma}_b\ \underline{H}'+\underline{\Gamma}'\ \underline{H}'+\underline{H}\ \underline{\Gamma}+\underline{\Theta} \quad ,$$

where $\underline{\Gamma} = E[\underline{X}^*\ \underline{\varepsilon}'] = \mathrm{Cov}[\underline{X},\underline{\varepsilon}]$.

From equations (I.2) and (I.3):

$$\hat{\underline{X}}^* = (\underline{\Sigma}_b\underline{H}'+\underline{\Gamma})(\underline{H}\ \underline{\Sigma}_b\underline{H}'+\underline{\Gamma}'\underline{H}'+\underline{H}\ \underline{\Gamma}+\underline{\Theta})^{-1}\underline{Z}^* \ ; \qquad \text{(a)} \qquad \text{(I.23)}$$

$$(\hat{\underline{X}} = \hat{\underline{X}}^*+\underline{X}_b) \quad ; \qquad \text{(b)}$$

$$\underline{S}_{\underset{\sim}{\underline{X}^*}} = \underline{S}_{\underset{\sim}{\underline{X}}} = \underline{\Sigma}_b-(\underline{\Sigma}_b\ \underline{H}'+\underline{\Gamma})(\underline{H}\ \underline{\Sigma}_b\underline{H}'+\underline{\Gamma}'\underline{H}'+\underline{H}\ \underline{\Gamma}+\underline{\Theta})^{-1}(\underline{H}\ \underline{\Sigma}_b+\underline{\Gamma}') \ . \quad \text{(I.24)}$$

The reason for introducing the new variables $\underline{X}^*$ and $\underline{Z}^*$ is to have $E[\underline{Z}^*] = E[\hat{\underline{X}}^*] = E[\underline{X}^*] = \underline{0}$, so that equation (I.4) holds and $\hat{\underline{X}}$ is the linear-minimum-variance estimator (and not only the linear minimum-mean-square-error estimator).

As already anticipated in discussing Bayesian estimation equations (I.23) and (I.24) are identical with equations (I.11) and (I.12). They hold under no distribution assumption, a property which makes $\hat{\underline{X}}_{LMV}$ a very convenient distribution-free estimator (see Anderson (1972) for some structural engineering applications of the scalar form (I.9) (I.10)).

Equations (I.23) and (I.24) reproduce equations (I.7) and (I.8) when $\underline{\Gamma} = \underline{0}$, and the Bayesian results for diffuse prior, equations (I.21) and (I.22), when, in addition, $\underline{\Sigma}_b^{-1} = \underline{0}$.

(ii) Maximum likelihood estimator

Equation (I.5) is the condition for maximum likelihood when $(\underline{Z}|\underline{X})$ has unimodal distribution. For the linear observation model (I.6) with Gaussian errors the conditional distribution of the observations is $N_m(\underline{Z}_o+\underline{H}\ \underline{X};\ \underline{\Theta})$. Therefore maximizing $f(\underline{Z}|\underline{X})$ is equivalent to minimizing the quadratic form:

$$||\underline{Z}-\underline{Z}_o-\underline{H}\ \underline{X}||^2_{\underline{\Theta}^{-1}} = (\underline{Z}-\underline{Z}_o-\underline{H}\ \underline{X})'\underline{\Theta}^{-1}(\underline{Z}-\underline{Z}_o-\underline{H}\ \underline{X}) \ .$$

This condition is satisfied if

$$\underline{X} = \hat{\underline{X}}_{ML}(\underline{Z}) = (\underline{H}'\underline{\Theta}^{-1}\underline{H})^{-1}\underline{H}'\underline{\Theta}^{-1}(\underline{Z}-\underline{Z}_o) = \underline{H}'(\underline{H}\ \underline{H}')^{-1}(\underline{Z}-\underline{Z}_o) \ .$$

We conclude that in this case the maximum likelihood estimator coincides with the Bayes posterior mean for diffuse prior, equation (I.21). The estimate error covariance is the same as in equation (I.22).

Note that these results are independent of the prior information on $\underline{X}$, so that from a Bayesian viewpoint the maximum likelihood estimator is efficient only if $\underline{\Sigma}_b^{-1} = \underline{0}$. (An estimator is said to be efficient if it minimizes the error mean square). Otherwise the Bayes posterior mean and the linear-minimum-variance estimators have smaller error covariance.

(iii) (Weighted) least-squares estimators

Given the linear observation model (I.6), let the optimality condition for the estimator have the weighted quadratic form:

$$J = \frac{1}{2}\left\{(\hat{\underline{X}}-\underline{X}_b)'\underline{\Sigma}_b^{-1}(\hat{\underline{X}}-\underline{X}_b)+(\underline{Z}-\underline{Z}_o-\underline{H}\ \hat{\underline{X}})'\underline{\Theta}^{-1}(\underline{Z}-\underline{Z}_o-\underline{H}\ \hat{\underline{X}})\right\} = \min_{\hat{\underline{X}}} \quad \text{(I.25)}$$

As before, $\underline{X}_b$ is the prior mean vector and $\underline{\Sigma}_b$ is the prior covariance matrix of $\underline{X}$, $\underline{\Theta}$ is the covariance of the zero-mean

measurement error vector $\underline{\varepsilon}$, and $\mathrm{Cov}[\underline{X},\underline{\varepsilon}] = \underline{0}$. In equation (I.25), J is the sum of a penalty term for the weighted distance of the observation from its expected posterior value. Of course, $\underline{\Sigma}_b$ and $\underline{\Theta}$ may be replaced by different matrices, but when they are chosen as indicated above, the weighted-least-squares estimator $\hat{\underline{X}}_{LS}(\underline{Z})$ coincides with the posterior Bayes mean, equation (I.7), and with the linear-minimum-variance estimator. In fact, in order that $d\,J/d\,\hat{\underline{X}} = \underline{0}$ it must be:

$$\hat{\underline{X}}(\underline{Z}) = \hat{\underline{X}}_{LS}(\underline{Z}) = \underline{X}_b + \underline{\Sigma}_a \underline{H}' \underline{\Theta}^{-1} (\underline{Z} - \underline{Z}_o - \underline{H}\,\underline{X}_b) \quad , \qquad \text{(I.26)}$$

with the error covariance $\underline{\Sigma}_a$ in equation (I.8).

If $\underline{\Sigma}_b$ and $\underline{\Theta}$ in equation (I.25) are replaced by different matrices the weighted-least-squares estimator is still given by equation (I.26) with the obvious substitutions, but now $\hat{\underline{X}}_{LS}$ is inferior to the estimators $\hat{\underline{X}}_{LMV}$ and $\underline{X}_a$.

In summary, for the linear observation model (I.6) the estimator (I.7) satisfies the linear-minimum-mean-square-of-error and the least-squares criteria, the latter in the form (I.25). If in addition the prior distribution of $\underline{X}$ and the distribution of the measurement error $\underline{\varepsilon}$ are multivariate normal, then (I.7) shares some desirable Bayesian properties, such as being the maximum posterior (density)

estimator, the absolute-value-of-error estimator, and the posterior mean. For a diffuse prior $\underline{X}_a$ in equation (I.21) has analogous properties being, in addition, the maximum likelihood estimator.

I.1.3 Linear observation model; measurement error covariance unknown

The estimation problem with the linear experiment model (I.6) and unknown measurement error covariance $\underline{\Theta}$ is significantly more involved than the same problem when $\underline{\Theta}$ is known. Nevertheless, simple Bayesian solutions can be found in particular situations.

One case which lends itself to an easy, although approximate, treatment is when the prior information on $\underline{\Theta}$ has the discrete form:

$$P(\underline{\Theta}) = \sum_{j=1}^{nj} P_j \, \delta \, (\underline{\Theta}-\underline{\Theta}_j) \qquad \text{(I.27)}$$

where $\sum_j P_j = 1$, $\delta(\underline{A}) = \begin{cases} 0 \text{ if } \underline{A} \neq \underline{0} \\ 1 \text{ if } \underline{A} = \underline{0} \end{cases}$, and $\{\underline{\Theta}_j\}$ is a set of known matrices.

Under the conditions $\underline{\varepsilon} \sim N_m(\underline{0}, \underline{\Theta})$; $\underline{X} \sim N_n(\underline{X}_b, \underline{\Sigma}_b)$; $\text{Cov}[\underline{X}, \underline{\varepsilon}] = \underline{0}$, and under the simplifying assumption that (I.27) is also the posterior distribution of $\underline{\Theta}$ (this assumption is legitimate for measurements carrying little

new information) the posterior mean of the state is given by

$$\underline{X}_a = \sum_{j=1}^{nj} P_j \underline{X}_{a_j} = \underline{X}_b + \sum_{j=1}^{nj} P_j \left(\underline{\Sigma}_{a_j} \underline{H}' \underline{\Theta}_j^{-1} (\underline{Z} - \underline{Z}_o - \underline{H} \underline{X}_b) \right) , \quad \text{(I.28)}$$

$$\text{where} \quad \underline{\Sigma}_{a_j} = (\underline{\Sigma}_b^{-1} + \underline{H}' \underline{\Theta}_j^{-1} \underline{H})^{-1} \quad \text{(I.29)}$$

The associated error covariance of the estimate is:

$$\underline{\Sigma}_a = \sum_{j=1}^{nj} \left(P_j \underline{\Sigma}_{a_j} + (\underline{X}_a - \underline{X}_{a_j})(\underline{X}_a - \underline{X}_{a_j})' \right) \quad \text{(I.30)}$$

For a diffuse prior on $\underline{X}$ equations (I.28), (I.29), (I.30) still hold with $\underline{\Sigma}_b^{-1} = \underline{0}$.

The posterior mean (I.28) is not any longer the maximum-posterior estimate, but it still retains the property of being the linear-minimum-variance estimator. The foregoing problem is solved exactly in Aoki (1967) Par. III.2.D.d for the scalar case with $n_j = 2$.

The difficulty of solving the general problem is illustrated by considering the maximum likelihood criterion. Using the general linear observation model (I.15) with $m_i = m$ and $\underline{\theta}_i = \underline{\theta}$ $(i = 1,\ldots,K)$ and following Sage and Melsa (1971) Par. 6.5, one finds:

$$\hat{\underline{X}}_{ML} = \left(\sum_{i=1}^{K} \underline{H}_i' \hat{\underline{\Theta}}_{ML}^{-1} \underline{H}_i \right)^{-1} \sum_{i=1}^{K} \underline{H}_i' \hat{\underline{\Theta}}_{ML}^{-1} (\underline{Z}_i - \underline{Z}_{o_i}) ; \quad \text{(a)}$$

(I.31)

$$\hat{\underline{\Theta}}_{ML} = \frac{1}{K} \sum_{i=1}^{K} (\underline{Z}_i - \underline{Z}_{o_i} - \underline{H}_i \hat{\underline{X}}_{ML})(\underline{Z}_i - \underline{Z}_{o_i} - \underline{H}_i \hat{\underline{X}}_{ML})' , \quad \text{(b)}$$

which is a system of coupled nonlinear equations in the coefficients of $\hat{\underline{\Theta}}_{ML}$ and of $\hat{\underline{X}}_{ML}$. The special case K=1 is degenerate since under the condition $\det(\underline{H}_1 \underline{H}_1') \neq 0$ equation (I.31b) yields: $\hat{\underline{\Theta}}_{ML} = \underline{0}$.

Nevertheless (and quite importantly) if one assumes that $\underline{\Theta}$ has the diagonal form

$$\underline{\Theta} = \sigma^2 \underline{I}_m , \quad (\sigma^2 \text{ unknown}) \qquad \text{(I.32)}$$

$\hat{\underline{X}}_{ML}$ does <u>not</u> depend on $\hat{\underline{\Theta}}_{ML}$. This special case is studied next in greater detail within the context of Bayesian estimation.

Consider first a <u>noninformative</u> prior density $f(\underline{X}, \sigma^2)$ of the type:

$$f(\underline{X}, \sigma^2) \propto 1/\sigma^2 . \qquad \text{(I.33)}$$

From equation (I.1) the joint posterior density is:

$$f(\underline{X}, \sigma^2|\underline{Z}) \propto f(\underline{Z}|\underline{X}, \sigma^2) \cdot f(\underline{X}, \sigma^2) \ , \qquad \text{(I.34)}$$

where

$$f(\underline{Z}|\underline{X}, \sigma^2) \propto \frac{1}{\sigma^m} \exp\left\{ - \frac{1}{2\sigma^2} (\underline{Z}-\underline{Z}_o-\underline{H}\ \underline{X})'(\underline{Z}-\underline{Z}_o-\underline{H}\ \underline{X}) \right\}$$

$$= \frac{1}{\sigma^m} \exp\left\{ - \frac{1}{2\sigma^2} \left(\nu S^2+(\underline{X}-\hat{\underline{X}})'\underline{H}'\underline{H}(\underline{X}-\hat{\underline{X}})\right) \right\} \qquad \text{(I.35)}$$

with the notations:

$$\hat{\underline{X}} = (\underline{H}'\underline{H})^{-1}\underline{H}'(\underline{Z}-Z_o) \quad \text{(if det } (\underline{H}'\underline{H}) \neq 0) \ ;$$

$$\hat{\underline{Z}} = \underline{H}\ \hat{\underline{X}}+Z_o \ ;$$

$$\nu = m-n \ ;$$

$$S^2 = \left((\underline{Z}-\hat{\underline{Z}})'(\underline{Z}-\hat{\underline{Z}})\right)/\nu \ .$$

Substituting (I.33) and (I.35) into (I.34) yields (Box and Tiao (1973) Par. 2.7.2):

$$f(\underline{X},\sigma^2|\underline{Z}) \propto \frac{1}{\sigma^{m+2}} \exp\left\{ - \frac{1}{2\sigma^2}\left(\nu S^2+(\underline{X}-\hat{\underline{X}})'\underline{H}'\underline{H}(\underline{X}-\hat{\underline{X}})\right) \right\} \qquad \text{(I.36)}$$

This joint density can be factored into a marginal posterior density for σ^2 and a conditional posterior density for $\underline{X}$, the latter being in multivariate normal form with mean $\hat{\underline{X}}$ and covariance matrix $\sigma^2(\underline{H}'\underline{H})^{-1}$. Integrating out σ^2 yields the marginal posterior distribution of $(\underline{X}-\hat{\underline{X}})$ as an n-variate

t-distribution with $\nu = m-n$ degrees of freedom and covariance matrix $\frac{\nu}{\nu-2} S^2 (\underline{H}'\underline{H})^{-1}$.

It is immediate to generalize this result for a conjugate prior distribution of $\underline{X}$ and σ^2. In fact, let

$$f(\underline{X},\sigma^2) \propto \frac{1}{\sigma^{m_1+2}} \exp\left\{ - \frac{1}{2\sigma^2}\left(\nu_1 S_1^{\,2} + (\underline{X}-\hat{\underline{X}}_1)'\underline{H}_1'\underline{H}_1(\underline{X}-\hat{\underline{X}}_1)\right)\right\} ,$$

where:

$$\begin{aligned} \nu_1 &= m_1 - n \quad ; \\ \hat{\underline{X}}_1 &= (\underline{H}_1'\underline{H}_1)^{-1}\underline{H}_1'(\underline{Z}_1 - \underline{Z}_{o_1}) \quad (\det(\underline{H}_1'\underline{H}_1)\neq 0) \quad ; \\ \nu_1 S_1^{\,2} &= (\underline{Z}_1 - \underline{Z}_{o_1} - \underline{H}_1\hat{\underline{X}}_1)'(\underline{Z}_1 - \underline{Z}_{o_1} - \underline{H}_1\hat{\underline{X}}_1) \quad . \end{aligned}$$

Then the posterior density is of the same type, namely:

$$f(\underline{X},\sigma^2 \mid \underline{Z}) \propto \frac{1}{\sigma^{m_1+m-2}} \cdot \exp\left\{ - \frac{1}{2\sigma^2}\left(\nu S^2 + (\underline{X}-\hat{\underline{X}})'\underline{M}(\underline{X}-\hat{\underline{X}})\right)\right\} , \qquad (I.37)$$

where:

$$\begin{aligned} \underline{M} &= \underline{H}_1'\underline{H}_1 + \underline{H}'\underline{H} \quad ; \\ \hat{\underline{X}} &= \underline{M}^{-1}(\underline{H}_1'\underline{Z}_1 + \underline{H}'\underline{Z}) \quad ; \\ \nu S^2 &= (\underline{Z}_1 - \underline{Z}_{o_1} - \underline{H}_1\hat{\underline{X}})'(\underline{Z}_1 - \underline{Z}_{o_1} - \underline{H}_1\hat{\underline{X}}) \\ &\qquad + (\underline{Z} - \underline{Z}_o - \underline{H}\,\hat{\underline{X}})'(\underline{Z} - \underline{Z}_o - \underline{H}\,\hat{\underline{X}}) \quad ; \\ \nu &= m_1 + m - n \quad . \end{aligned}$$

For a detailed derivation see Zellner (1971) Par. 3.2.3. The result can be extended by induction to any number of measurements $\underline{Z}_j$, $j=1,2,\ldots$. Since the posterior density (I.37) has the same form as the posterior density (I.36)

the considerations in the marginal state probability density for diffuse prior carry over to the present case; in particular $(\underline{X}-\hat{\underline{X}}) \sim t_n[S^2\underline{M}^{-1}; m_1+m-n]$.

I.1.4 Some additional considerations and extensions

(i) Error analysis

Equations (I.7) and (I.8) were derived under the hypothesis that the parameters $\underline{X}_b$ and $\underline{\Sigma}_b$ of the prior normal distribution and the measurement error covariance $\underline{\Theta}$ are the "correct" ones. For instance, $\underline{X}$ might be a realization from a normal process with parameters $\underline{X}_b$ and $\underline{\Sigma}_b$; in this context one may say that the values assumed a priori for the parameters are the "correct" ones or not. The sensitivity to this type of errors is studied by Sage and Melsa (1971) Par. 6.5, with reference to the maximum-likelihood and to the Bayes maximum-posterior-density estimators. Their main result for the Bayes estimator, equations (I.7), is as follows.

Let $\overline{\underline{X}}_b$, $\overline{\underline{\Sigma}}_b$, $\overline{\underline{\Theta}}$ be the assumed parameter values, and $\underline{X}_b$, $\underline{\Sigma}_b$, $\underline{\Theta}$ the true values. If one uses equations (I.7) with the assumed parameters the estimator is biased. In fact, denoting $\overline{\underline{X}}_a$ the resulting estimate it is:

$$E[\overline{\underline{X}}_a] = (\underline{H}'\overline{\underline{\Theta}}^{-1}\underline{H}+\overline{\underline{\Sigma}}_b^{-1})^{-1}(\underline{H}'\overline{\underline{\Theta}}\,\underline{H}\,\underline{X}_b+\overline{\underline{\Sigma}}_b^{-1}\overline{\underline{X}}_b)$$

$$= \underline{X}_b + (\underline{H}'\overline{\underline{\Theta}}^{-1}\underline{H} + \overline{\underline{\Sigma}}_b^{-1})^{-1}\overline{\underline{\Sigma}}_b^{-1}(\overline{\underline{X}}_b - \underline{X}_b) \quad , \tag{I.38}$$

which equals $\underline{X}_b$ only if $\overline{\underline{X}}_b = \underline{X}_b$ or if $\overline{\underline{\Sigma}}_b^{-1} = \underline{0}$. The mean square error of the computed estimate $\overline{\underline{X}}_a$ is:

$$E\left((\underline{X}-\overline{\underline{X}}_a)(\underline{X}-\overline{\underline{X}}_a)'\right) = (\underline{H}'\overline{\underline{\Theta}}^{-1}\underline{H} + \overline{\underline{\Sigma}}_b^{-1})^{-1}$$

$$\left\{\overline{\underline{\Sigma}}_b^{-1}\left(\underline{\Sigma}_b + (\overline{\underline{X}}_b - \underline{X}_b)(\overline{\underline{X}}_b - \underline{X}_b)'\right)\overline{\underline{\Sigma}}_b^{-1} + \underline{H}'\overline{\underline{\Theta}}^{-1}\underline{\Theta}^{-1}\overline{\underline{\Theta}}^{-1}\underline{H}\right\}$$

$$(\underline{H}'\overline{\underline{\Theta}}^{-1}\underline{H} + \overline{\underline{\Sigma}}_b^{-1})^{-1} \quad . \tag{I.39}$$

which is not smaller than the error variance (I.8). Equations (I.38) and (I.39) reduce to equations (I.7) and (I.8) if $\overline{\underline{X}}_b = \underline{X}_b$; $\overline{\underline{\Sigma}}_b = \underline{\Sigma}_b$ and $\overline{\underline{\Theta}} = \underline{\Theta}$.

These results are useful to judge the relevance of errors if the prior parameters are estimated from previous statistical data, without going through a full Bayesian analysis. Indeed, this is a recurrent situation in problems of engineering reliability; for instance, in estimating the quality of serially produced structural elements when limited data are available a priori from the statistical population (e.g., data are the resistances of previously tested elements). A similar problem arises in two-stage sampling (say of soil properties), in which one needs to estimate the posterior probabilities after a preliminary

sampling, in order to plan the second exploration in some optimal way.

Sage and Melsa also give an approximate expression for the increase in the estimate error covariance due to imperfect knowledge of $\underline{\Sigma}_b$ and of $\underline{\theta}$, valid for small errors.

(ii) H unknown

The estimation of the state of a physical system from noisy observations is mathematically analogous to finding the "best" parameters of a regression relation. Within the latter framework Fedorov (1972) Par. 1.6, studies the problem of (nonlinear) parameters estimation when the control matrix $\underline{H}$ (our measurement matrix) has random entries. In estimation, $\underline{H}$ being random means that the observed quantities and the state of the system are not related deterministically, even after removing the additive measurement noise term (see equation (I.6)). Working with the first few moments of $\underline{H}$ Fedorov finds the first two moments of the regression parameters by using a finite series expansion of $\underline{H}$ around its expected value.

(iii) Robustness

In estimation, Bayesian robustness is concerned with the sensitivity of the posterior distribution to variations in the prior distribution $f(\underline{X})$ and to variations in the

distribution of the measurement errors. Note the difference with error analysis, in which one studies the effect of deviations from fixed, "correct", parameter values. Much of the literature on estimation robustness in Classical and Bayesian inference refers to population parameters, but the same results can be applied to the estimation of a state of nature.

Since Bayesian estimation requires information in which the analyst has rarely complete confidence, the estimators robustness is a question of great practical interest. A detailed account of the available results is beyond the scope of this study; as guides to the literature in this area the reader is referred to the review chapter on robustness in Lindley (1971), to Chapter 3 in Box and Tiao (1973), and to the references therein. We mention briefly some relevant studies.

Box and Tiao (1973) varied the type of distribution of the measurement error, and studied the effects of that on the posterior density of the (simple) regression coefficients (the analogous of a two-dimensional state vector). $f(\varepsilon)$ was assumed within a family of symmetrical distributions, indexed by a continuous parameter $\beta\varepsilon[-1,1]$. β controls the Kurthosis coefficient, and is respectively -1, 0, and 1 for the uniform, the normal and the double exponential distribution. Processing a set of real data they found that

with respect to the normal case ($\beta = 0$) values of β smaller than -0.3 affected considerably the shape of the parameters posterior density as well as its location, while positive values up to $\beta = 0.9$ did not. However, the relaxation of the normality assumption on $\underline{\varepsilon}$ is not of critical importance when estimating a non-random state of nature from noisy measurements, since the actual distribution of $(\underline{Z}|\underline{X})$ is generally close to the normal. The assumption of normality may instead be unjustified when $\underline{\varepsilon}$ has not the meaning of a measurement error, as in the case of regression analysis.

With reference to the estimation of a state of nature robustness with respect to the prior density $f(\underline{X})$ is perhaps more important. On the effects of departures from a normal density, see Edwards, et al. (1963).

Finally, we mention another area of robustness analysis, namely the sensitivity of the estimator to a small presence of outliers in the data. In the framework of Bayesian analysis this problem was studied by Box and Tiao (1968).

(iv) Nonlinear observation model

Nonlinear estimation has been the object of much research during the last decade, particularly in the area of nonlinear optimal control, resulting in a variety of non-linear filtering algorithms. The "static" problem consists

in estimating a vector $\underline{X}$ from noisy nonlinear observations $\underline{Z} = g(\underline{X},\underline{\varepsilon})$. The practical need for nonlinear models is apparent; however, for our present goals it is considered sufficient to address the reader to existing literature. Also, since statistical estimation theory is relatively new to the reliability engineer, it seems appropriate to start from the relatively simple linear results. Their use in solving practical problems will suggest the appropriate extension into the nonlinear theory.

A general discussion of nonlinear estimation is in Balakrishnam (1964). Sage and Melsa (1971) devote a whole chapter to nonlinear Bayesian estimation. Although they refer to discrete and continuous dynamical systems, steps of the dynamic algorithms can be used for "static" estimation. Deutsch (1965) deals with "static" nonlinear estimation problems in a Bayesian framework, and indicates several approximate procedures. Other useful readings are Aoki (1971), and the pertinent references in Sage and Melsa (1971).

EXAMPLE 1. EFFECT OF NOISY MEASUREMENTS ON THE RELIABILITY OF SERIES, PARALLEL AND SERIES-PARALLEL SYSTEMS

Consider a construction plant which serially produces pre-cast reinforced concrete members. The resistance of the

i-th member is denoted R_i; the sequence R_i is assumed to be Gaussian and white. (The assumption of independence is in no sense essential, but it simplifies the computations without weakening the results.) Before combining the elements into a more complex structural configuration or system, we decide to measure nondestructively a secondary characteristic of each of the elements, such as their hardnesses or stiffnesses which are known to be correlated with their resistances. Then only those elements which have been tested are used in the system.

Denote R_b and σ_b^2 the mean and the variance of the production sequence $\{R_i\}$, which are assumed to be known. For the ith element in the system the measurement Z_i of the secondary characteristic and the resistance R_i are related through the linear stochastic equation

$$Z_i = Z_o + h\, R_i + \varepsilon_i \ , \ (i = 1,\ldots,n)$$

where Z_o and h are known constants and n is the number of elements in the system. The zero-mean error vector $\underline{\varepsilon}' = [\varepsilon_1,\ldots,\varepsilon_n]$ has multinormal distribution with covariance matrix:

$$\underline{\Theta} = C^2\, \sigma_b^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} ,$$

where C is a known constant. The equicorrelated structure of $\underline{\Theta}$ results from modeling the measurement error ε_i as the sum of an uncertain bias term ε_o common to all the observations, and an independent error component ε_i^*. In fact the covariance matrix $\underline{\Theta}$ corresponds to:

$$\varepsilon_i = \varepsilon_o + \varepsilon_i^*$$

with $E[\varepsilon_o] = E[\varepsilon_i^*] = 0$; $\sigma^2_{\varepsilon_o} = E[\varepsilon_o^2] = \rho C^2 \sigma_b^2$; $\sigma^2_{\varepsilon_i^*} = E[\varepsilon_i^{*2}] = (1-\rho) C^2 \sigma_b^2$; $E[\varepsilon_o \varepsilon_i^*] = 0$.

If $\sigma^2_{\varepsilon_o}$ and $\sigma^2_{\varepsilon_i^*}$ are given as part of the modeling, the necessary parameters $C^2\sigma_b^2$ and ρ can be determined. Because of the whiteness assumption on $\{R_i\}$ the prior covariance matrix of the resistances is: $\underline{\Sigma}_b = \sigma_b^2 \underline{I}_n$ and the posterior covariance matrix is, from equations (I.8):

$$\underline{\Sigma}_a = \sigma_b^2 \underline{I}_n - h^2 \sigma_b^4 (\underline{\Theta} + h^2 \sigma_b^2 \underline{I}_n)^{-1} \qquad \text{(I.40a)}$$

Due to the correlation in the measurements, $\underline{\Sigma}_a$ is not diagonal. Its entries are:

$$(\underline{\Sigma}_a)_{ij} = \begin{cases} \sigma_a^2 = \sigma_b^2 \left\{ 1 - \dfrac{h^2}{(h^2+C^2)} \dfrac{1 + (n-2)\,\rho^*}{(1-\rho^*)[1+(n-1)\rho^*]} \right\} & \text{for } i=j; \\ \\ \rho_a \sigma_a^2 = \sigma_b^2 \left\{ \dfrac{h^2}{(h^2+C^2)} \dfrac{\rho^*}{(1-\rho^*)[1+(n-1)\rho^*]} \right\} & \text{for } i \neq j \end{cases} \qquad \text{(I.40b)}$$

where

$$\rho^* = \rho \frac{C^2}{C^2 + h^2} \quad .$$

For making possible a comparison between the prior and the posterior reliability of systems which use these elements, one has to know the actual experimental outcomes, i.e., the numerical values of the measurements. In what follows we assume for simplicity and clarity that each outcome turns out to be simply its expected value:

$$Z_i = E[Z_i] = Z_o + h \cdot R_b \qquad (i = 1,\ldots,n)$$

(in the next section we shall comment further on this assumption); for this assumption the posterior mean resistances are the same as the prior mean resistances; i.e., $R_{a_i} = R_b$ from equations (I.7), simplifying the reliability comparisons to follow.

Series system. When the elements are connected in series as in fig. 1A and the system is acted on by a deterministic load S, failure may occur in any of the n modes corresponding to the conditions: $R_i < S$ $(i = 1,2,\ldots,n)$. Due to the fact that the posterior means are $R_{a_i} = R_b$ and to the equicorrelated structure of the posterior covariance matrix (I.40), the posterior reliability of the system depends only on the number of elements, n, on the normalized modal safety margin $\beta_a = (S-R_b)/\sigma_a$, and on the modal correlation coefficient ρ_a, which is the same for all pairs of modal resistances. From the assumption that $\{R_i\}$ is a priori an independent sequence, the prior modal correlation ρ_b is zero. The prior normalized modal safety margin is $\beta_b = (S-R_b)/\sigma_b$.

For a given number of bars, n, a normalized safety margin β, and a modal correlation coefficient ρ_M the probability of failure of the system is the probability that at least one bar's resistance is less than the applied load, or:

$$P_f = 1 - \Phi_n(\rho_M, \beta) \quad ,$$

where $\Phi_n(\rho_M, \beta)$ is the probability that n standard normal variates with common correlation coefficient ρ_M are simultaneously less than or equal to β. It can be shown (see

Gupta (1963)) that

$$\Phi_n(\rho_M,\beta) = \int_{-\infty}^{\infty} \Phi^n\left((\beta+\rho_M^{1/2}u)(1-\rho_M)^{-1/2}\right)\,\phi(u)\,du \qquad (I.41)$$

(with the obvious particular case: $\Phi_n(0,\beta) = \Phi^n(\beta)$), where $\Phi[\cdot]$ and $\phi(\cdot)$ denote the standard normal CDF and PDF, respectively. Gupta tabulated $\Phi_n(\rho_M,\beta)$ for n = 1(1)12, for β(his h) = -3.5(0.1)3.5, and for discrete values of ρ_M in the range [0.1, 0.9]. Plots of the reliability for selected values of β, ρ_M and for 1 to 12 elements, are shown in fig. 1. In accordance with a theorem by Slepian (1962) P_f is a non-increasing function of ρ_M. Noticeable features of these plots are (see fig. 1):

(i) Small modal correlation coefficients (say $\rho_M < 0.5$) are negligible in safely designed systems (fig. 1 D), while the opposite is true for unreliable systems (fig. 1A);

(ii) The sensitivity of P_f to the number of elements increases with the system reliability and decreases with the modal correlation. In particular, fig. 1D suggests that for reliable systems with slightly correlated modes the probability of failure decays with the number of elements in an approximately geometrical fashion: $P_f \approx 1 - \Phi^n(\beta)$.

Returning to the reliability effects of noisy measurements, for S deterministic the prior reliability corresponds to the straight lines in fig. 1 ($\rho_M = 0$). Given n, β_b and the measurement constants h, C and ρ, the posterior parameters are:

$$\beta_a = \beta_b \frac{\sigma_b}{\sigma_a} = \beta_b \left\{ 1 - \frac{h^2}{(h^2+C^2)} \frac{1 + (n-2)\ \rho^*}{(1-\rho^*)[1+(n-1)\rho^*]} \right\}^{-1/2} ;$$

$$\rho_a = \frac{h^2}{h^2+C^2} \frac{\rho^*}{(1-\rho^*)[1+(n-1)\rho^*]} \left\{ 1 - \frac{h^2}{(h^2+C^2)} \frac{1 + (n-2)\ \rho^*}{(1-\rho^*)[1+(n-1)\rho^*]} \right\}^{-1}$$

$$= \frac{h^2\rho^*}{C^2[1+(n-2)\rho^*] - (C^2+h^2)(n-1)\rho^{*2}} .$$

From these one can calculate the posterior reliability. For example, if n = 2, h = C = 1 and ρ = 0.5 one finds

$$\beta_a = 1.46\ \beta_b, \text{ and } \rho_a = 0.286.$$

For the same parameter values the prior and the posterior failure probabilities for selected values of β_b are given in the following table.

β_b	P_{f_b}	P_{f_a}
0	0.75	0.705
1.0	0.29	0.13
1.5	0.13	0.027
2.0	0.046	0.004

Table 1. Prior normalized safety margin β_b and prior and posterior (after the estimation experiment) failure probabilities P_{f_b} and P_{f_a} for the series system in fig. 1 under deterministic loading. Refer to the text for parameters values and for the definition of the estimation experiment.

Ductile parallel system (fig. 2A). System failure occurs if the applied load S excedes the sum of the individual element resistances, $R = \sum_{i=1}^{n} R_i$. A priori the distribution of R is $N(n\,R_b;\ \sigma^2_{R_b})$; a posteriori: $R \sim N(n\,R_b;\ \sigma^2_{R_a})$, where

$$\sigma^2_{R_b} = n\,\sigma^2_b \quad ;$$

$$\sigma^2_{R_a} = \sum_i \sum_j (\underline{\Sigma}_a)_{ij} = n\sigma^2_b \left\{ 1 - \frac{h^2}{(h^2+C^2)\,[1+(n-1)\,\rho^*]} \right\} .$$

For a given load S, let β_b' be the normalized prior safety margin; then the normalized posterior safety margin is:

$$\beta_a' = \beta_b' \, \sigma_{R_b}/\sigma_{R_a} = \beta_b' \left\{ 1 - \frac{h^2}{(h^2+C^2)\,[1+(n-1)\,\rho^*]} \right\}^{-1/2} .$$

If S is random, say $S \sim N(\bar{S}, \sigma_S^2)$, and independent of R, the prior and the posterior normalized safety margins become:

$$\beta_b = \beta_b' \frac{\sigma_{R_b}^2}{\sigma_{R_b}^2 + \sigma_S^2} \quad ; \quad \beta_a = \beta_b \frac{\sigma_{R_b}^2 + \sigma_S^2}{\sigma_{R_a}^2 + \sigma_S^2} ,$$

where β_b' has the meaning of narmalized prior safety margin for $S = E[S] = \bar{S}$.

The prior and the posterior failure probabilities are the values of the standard normal CDF at β_b' and at β_a' for given S, at β_b and at β_a for random S.

Series-parallel system. Consider the system in fig. 2 and let R_j^G denote the resistance of the jth parallel group of n/N elements. The (N×N) prior covariance matrix of the group resistances, $\underline{\Sigma}_b^G$, is diagonal with non-zero elements $\frac{n}{N}\,\sigma_b^2$. The posterior covariance matrix is not diagonal; in fact straightforward algebra yields

$$(\underline{\Sigma}_a^G)_{ij} = \begin{cases} \frac{n}{N}\sigma_b^2 \left\{1 - \frac{h^2}{(h^2+C^2)} \frac{1 + (n-1-n/N)\ \rho^*}{(1-\rho^*)[1+(n-1)\rho^*]}\right. & \text{for } i=j\ , \\ \left(\frac{n}{N}\right)^2 \sigma_b^2 \left\{ \frac{h^2}{(h^2+C^2)} \frac{\rho^*}{(1-\rho^*)[1+(n-1)\rho^*]}\right. & \text{for } i\neq j\ . \end{cases}$$

If S is deterministic and β_b is the prior safety margin for a single group (normalized with respect to $\left(\frac{n}{N}\right)^{1/2} \cdot \sigma_b$), then the situation is analogous to that for series systems, except that now:

$$\beta_a = \beta_b (1-q)^{-1/2}\ , \text{ and } \rho_a = \frac{n}{N} \frac{q}{1-q} \frac{\rho^*}{1+(n-1-n/N)\rho^*}\ ,$$

where

$$q = \frac{h^2}{(h^2+C^2)} \frac{1+(n-1-n/N)\rho^*}{(1-\rho^*)[1+(n-1)\rho^*]}\ .$$

The prior and the posterior failure probabilities can be obtained easily from these expressions and from the aforementioned tables in Gupta (1963).

Additional examples of applications of estimation theory to reliability problems are presented in the following two sections.

I.2 EXPERIMENT EVALUATION

The effectiveness of an experiment depends on the measurement model, on the quality of prior information and on the goals of experimentation. Two situations should be distinguished:

(a) The effectiveness of the experiment is evaluated a posteriori, in which case no decision problem arises;

(b) The effectiveness is evaluated a priori, in order to compare different alternatives and finally to choose the best course of action. The latter is a typical problem of decision making under uncertainty and will be considered in more depth in Section I.3.

Situation (a) is illustrated by the example in the last section: in that case, given the prior distribution, the difference between the prior and the posterior reliabilities

was taken as a measure of effectiveness. A preposterior application of the same criterion for decision purposes, i.e., for choosing a priori the best experiment, based on considerations about the posterior state of knowledge, is senseless because $E[P_{f_a}] = P_{f_b}$ (the expectation being over all possible experiment outcomes). For the purpose of a qualitative comparison, a simple-minded way of avoiding the difficulty was used in the last section: namely it was assumed that the measurements would correspond to their prior expected values. Under this condition the posterior mean is the same as the prior mean and in a Gaussian model the difference $\underline{\Sigma}_a - \underline{\Sigma}_b$ gives a complete description of the gain from experimentation.

In decision problems a quite general preposterior criterion is based on the notion of expected utility. For instance, the expected utility of an experiment ε might be (for simplicity assume additivity):

$$E[U(\varepsilon)] = E[U_\varepsilon] + \int_{\text{all } \underline{Z}} \int_{\text{all } \underline{X}} E[U_{\underline{X}}]\, f(\underline{X}|\underline{Z})\, f(\underline{Z})\, d\underline{X}\, d\underline{Z} \quad \text{(I.42)}$$

where U_ε is the expected utility of making the experiment with no account for its outcome (this term includes the actual cost of the experiment, the legal consequences of non-experimenting, etc.), $U_{\underline{X}}$ is the utility of the state of nature being $\underline{X}$, $f(\underline{X}|\underline{Z})$ is the PDF of the state of nature

having observed $\underline{Z}$, and $f(\underline{Z})$ is the PDF of the observation. A widely accepted optimality criterion for ε is that it maximizes $E[U(\varepsilon)]$.

A decision analysis based on the maximum expected utility criterion is practically feasible only in very special cases. Often one cannot express utilities in an analytical form, if one can quantify them at all. In addition, there are cases in which an "optimal" sampling policy must be decided for a whole class of different future situations (for example in structural codes writing), making the expected utility criterion even less practical. These difficulties motivate the search for simpler evaluation criteria.

In second-moment probability theory the dispersion of a random vector is measured by its covariance matrix $\underline{\Sigma}$: the "larger" $\underline{\Sigma}$, the "larger" the dispersion. If the utility of an experiment increases with "decreasing" posterior estimate error covariance, a class of second-moment criteria can be constructed by defining a preference order in $\underline{\Sigma}$ space. In this sense some commom definitions are given below, together with their geometrical interpretation.

Let E_1 and E_2 be two alternative experiments. We write $E_1 > E_2$ if E_1 is preferred to E_2 according to a given criterion. Also, let $\underline{\Sigma}_a$ denote the posterior covariance or

the estimation error mean square, whichever applies. Some criteria for comparing E_1 and E_2 on a second-moment basis are (see also Fedorov (1972) Par. 1.8):

$$\text{(i)} \quad E_1 > E_2 \text{ if } \underline{\Sigma}_a(E_1) < \underline{\Sigma}_a(E_2) \tag{I.43}$$

The associated geometric condition is that the (origin centered) dispersion ellipsoid of $E_1: \underline{X}' \underline{\Sigma}_a^{-1}(E_1)\underline{X} \leq C^2$ is inscribed within the ellipsoid of $E_2: \underline{X}'\underline{\Sigma}_a^{-1}(E_2)\underline{X} \leq C^2$ for all C. This is a very strict requirement, and the criterion is applicable only when the dispersion of <u>any</u> linear combination of the state under one experiment is less than the corresponding dispersion under the other experiment.

$$\text{(ii)} \quad E_1 > E_2 \text{ if } |\underline{\Sigma}_a(E_1)| < |\underline{\Sigma}_a(E_2)| \tag{I.44}$$

Since the eigenvalues of $\underline{\Sigma}_a(E)$ are the squares of the principal semiaxes of the ellipsiod $\underline{X}'\underline{\Sigma}_a^{-1}(E)\underline{X} \leq 1$, $|\underline{\Sigma}_a(E)|$ increases with the volume of the dispersion ellipsoid. An experiment which is optimal in the sense of (I.44) is called a <u>D-optimal experiment</u>.

$$\text{(iii)} \quad E_1 > E_2 \text{ if } \mathrm{Tr}\ \underline{\Sigma}_a(E_1) < \mathrm{Tr}\ \underline{\Sigma}_a(E_2) \tag{I.45}$$

In this case one penalizes an invariant sum of directional dispersions. In some instances the trace operator is analytically less tractable than the determinant operator, but it has the conceptual advantage of associating non-zero penalty to singular matrices. An experiment which minimizes the mean dispersion $\frac{1}{n}$ Tr $\underline{\Sigma}_a(E)$ is called an <u>A-optimal experiment</u>.

$$\text{(iv)} \quad E_1 > E_2 \text{ if } \max_i [\underline{\Sigma}_a(E_1)]_{ii} < \max_i [\underline{\Sigma}_a(E_2)]_{ii} \qquad \text{(I.46)}$$

Here only the maximum posterior variance is penalized.

(v) $E_1 > E_2$ if one of the conditions (i) to (iv) applies to

$$\underline{\Sigma}_{a*}(\cdot) = \underline{L}\ \underline{\Sigma}_a(\cdot)\ \underline{L}' \text{ ; i.e., to } \underline{X}^* = \underline{L}\ \underline{X} \ .$$

The last criterion is quite flexible; in particular it may be used to judge the effectiveness of experiments along specific directions in state space.

The foregoing criteria are appropriate for ranking experiments in preferential order. Sometimes the question arises whether or not to make an experiment when prior information is available. At the level of second moments the informativeness of prior knowledge can be quantified by

a scalar dispersion measure $L(\underline{\Sigma}_b)$ (such as $|\underline{\Sigma}_b|$, Tr $\underline{\Sigma}_b$, $\max_i [\underline{\Sigma}_b]_{ii}$, or other). If an experiment E is informative the posterior dispersion is smaller than the prior under any criterion $L(\cdot)$.

In the examples which follow the effectiveness of an experiment is measured by the ratio:

$$r(E) = \left\{ \frac{|\underline{\Sigma}_b|}{|\underline{\Sigma}_a(E)|} \right\}^{1/2n} \tag{I.47}$$

The geometrical meaning of (I.47) is best illustrated by the case in which $f(\underline{X})$ and $f(\underline{X}|\underline{Z})$ are spherical normal densities; then r is the ratio between the radii of any two n-dimensional mean-centered spheres with the same probability content. The more an experiment is informative, the larger the effectiveness ratio r. Also, if $r(E_1) > r(E_2)$, it is $E_1 > E_2$ according to the criterion (I.43). Occasionally different effectiveness measures will be introduced.

An explicit expression for r(E) can be obtained in terms of the prior state covariance and of the measurement parameters, if $\underline{\Sigma}_a(E)$ is computed through equations (I.8). For a scalar state:

$$r(E) = \sigma_b/\sigma_a(E) = \frac{\sigma_b}{\sigma_b(1+\sigma_b^2\underline{H}'\underline{\Theta}^{-1}\underline{H})^{-1/2}} = (1+\sigma_b^2 m)^{1/2} , \tag{I.48}$$

where $m = \underline{H}'\underline{\Theta}^{-1}\underline{H}$ is the Fisher information matrix, here simply a scalar. When $\sigma_b \to \infty$ and $\underline{H}'\underline{\Theta}^{-1}\underline{H} \neq 0$ (diffuse prior and informative experiment) equation (I.48) gives $r \to \infty$. At the other extremum, when $\sigma_b = 0$ or when the experiment is noninformative ($\underline{\Theta}^{-1} = \underline{0}$, or $\underline{H} = \underline{0}$), one finds $r = 1$, which indicates that the experiment is useless. The quantity $\sigma_b^2 m$ is the fraction by which $1/\sigma_b^2$ is increased by making the experiment.

If K measurements are made, possibly with different linear models, equation (I.48) generalizes to:

$$r_K(E) = (1+\sigma_b^2 m_K)^{1/2}, \text{ where } m_K = \sum_{i=1}^{K} \underline{H}_i \underline{\Theta}_i^{-1} \underline{H}_i \ . \qquad \text{(I.49)}$$

When the unknown state of nature is an n-vector and equations (I.8) hold, the effectiveness ratio (I.47) is:

$$r(E) = (|\underline{\Sigma}_b| \cdot |\underline{\Sigma}_a^{-1}|)^{1/2n} = |\underline{\Sigma}_b(\underline{\Sigma}_b^{-1}+\underline{H}'\underline{\Theta}\ \underline{H})|^{1/2n}$$

$$= |\underline{I}_n+\underline{\Sigma}_b\underline{H}'\underline{\Theta}^{-1}\underline{H}|^{1/2n}$$

which can be written, in analogy with equation (I.48):

$$r(E) = |\underline{I} + \underline{\Sigma}_b\ \underline{M}|^{1/2n}, \text{ where } \underline{M} = \underline{H}'\underline{\Theta}^{-1}\underline{H}. \qquad \text{(I.50)}$$

Finally, for a set of K linear measurements the same ratio is:

$$r_K(E) = |\underline{I} + \underline{\Sigma}_b \underline{M}_K|^{1/2n}, \text{ where } \underline{M}_K = \sum_{i=1}^{K} \underline{H}_i \underline{\Theta}_i^{-1} \underline{H}_i \ . \qquad (I.51)$$

One can also find the effectiveness ratio with respect to a set of transformed variables $\underline{X}^* = \underline{L}\ \underline{X}$ ($\underline{X}^*$ has dimension $n^* \leq n$). For the most general case considered here equation (I.51) becomes:

$$r_K^* = |\underline{I} + \underline{\Sigma}_b^* \underline{M}_K^*|^{1/2n^*}$$

where $\underline{\Sigma}_b^* = \underline{L}\ \underline{\Sigma}_b\ \underline{L}'$ and $\underline{M}_K^* = \underline{L}\ \underline{M}_K\ \underline{L}'$.

Two examples follow. The first one refers to a scalar state for which equations (I.48) and (I.49) are used; the second one to a three-dimensional vector $\underline{X}$. In particular, the first example shows the relative importance of "systematic" versus "random" measurement errors as a function of the number of repeated observations.

EXAMPLE 2. SYSTEMATIC AND RANDOM MEASUREMENT ERRORS IN SCALAR ESTIMATION

Let X denote an unknown scalar state of nature, whose physical meaning can be decided at will. For reducing

the uncertainty on X one can make one or more measurements of X with a given facility (machinery and personnel) in which case the measurement errors are correlated by the presence of uncertain but systematic bias. Alternatively one may use more than one facility (possibly also changing the type of experiment), in which case the errors when using different facilities are uncorrelated. In general the experiment program may involve K different facilities (or K different types of experiment) with n_i measurements from the ith facility. This corresponds to a measurement vector $\underline{Z}$ of the form:

$$\underline{Z} = \begin{pmatrix} \underline{Z}_1 \\ \underline{Z}_2 \\ \vdots \\ \underline{Z}_K \end{pmatrix}, \text{ where } \underline{Z}_i = \begin{pmatrix} \underline{Z}_{i_1} \\ \underline{Z}_{i_2} \\ \vdots \\ \underline{Z}_{i_{n_i}} \end{pmatrix}$$

The following model is considered to be appropriate:

- Prior information (at a second-moment level):

$$X \sim (X_b;\ \sigma_b^2) \tag{I.52a}$$

- Experiment model:

$$\underline{Z}_i = \underline{H}_i X + \underline{\varepsilon}_i = \begin{pmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{pmatrix} X + \begin{pmatrix} \varepsilon_{i_1} \\ \varepsilon_{i_2} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_{i_{n_i}} \end{pmatrix} \quad (i=1,\ldots,K) \qquad (I.52b)$$

- The measurement error vector $\underline{\varepsilon}' = [\underline{\varepsilon}_1', \underline{\varepsilon}_2', \ldots\ldots, \underline{\varepsilon}_K']$ has zero mean and block-diagonal covariance matrix $\underline{\Theta} = \text{diag}\ (\underline{\Theta}_i)$, where:

$$\underline{\Theta}_i = \sigma_i^2 \begin{array}{c} \begin{matrix} 1 & & & & n_i \end{matrix} \\ \begin{matrix} 1 \\ \\ \\ \\ n_i \end{matrix} \left[\begin{array}{ccccc} 1 & & & & \\ & \ddots & & \rho_i & \\ & & \ddots & & \\ & \rho_i & & \ddots & \\ & & & & 1 \end{array}\right] \end{array} , \quad \rho_i \geq 0 . \qquad (I.52c)$$

Note that since the n_i measurements with the ith facility have the same distribution, no generality was lost by choosing $\underline{H}_i = [1,\ldots,1]'$ in equation (I.52b).

Let $\underline{\Theta}^{(n)}$ denote the matrix (I.52c) with $n_i = n$,

$\sigma_i^2 = \sigma^2$ and $\rho_i = \rho$. Then:

(I.53)

$$\left[\underline{\Theta}^{(n)}\right]^{-1} = \frac{1}{\sigma^2(1-\rho)[1+(n-1)\rho]} \begin{pmatrix} [1+(n-2)\rho] & -\rho & \cdots & -\rho \\ -\rho & [1+(n-2)\rho] & \cdots & \vdots \\ \vdots & \ddots & \ddots & -\rho \\ -\rho & \cdots & -\rho & [1+(n-2)\rho] \end{pmatrix}$$

Since $\underline{H}$ is the unit column vector, when using the pseudo-Bayes estimator (I.7) the Fisher information matrix of the experiment (here a scalar) equals the sum of the entries of $\left[\underline{\Theta}^{(n)}\right]^{-1}$:

$$m = \underline{H}' \left[\underline{\Theta}^{(n)}\right]^{-1} \underline{H} = \frac{1}{\sigma^2} \frac{n}{1+(n-1)\rho} \quad (*) \qquad \text{(I.54)}$$

The informativeness ratio of the experiment, r, increases with m, according to equation (I.48). The quantity $\sigma^2 m$ is plotted in fig. 3 for varying n and ρ. With the exception of ρ being 0 or 1, $\sigma^2 m$ is a nonlinear function of n and the difference $[(\sigma^2 m)_n - (\sigma^2 m)_{n-1}]$ decreases with n. This means

---

(*) If one removes the restriction $\rho \geq 0$, equation (I.54) yields $m = \infty$ when $\rho = -1/(n-1)$ $(n \geq 2)$. In fact $-1/(n-1)$ is the minimum value attainable by ρ for a set of n equicorrelated variables. If $\rho = -1/(n-1)$, X is a deterministic function of $\underline{Z}$.

that there is little convenience in making many correlated observations with the same facility, the advantage of making one additional observation being a decreasing function of the number of past observations. If $\rho > 0$, m approaches a finite asymptote as $n \to \infty$:

$$m_\infty = \lim_{n\to\infty} m = \frac{1}{\rho\ \sigma^2} \qquad \text{(I.55)}$$

This fact is easily explained. The equicorrelated error covariance matrix $\underline{\theta}^{(n)}$ can be generated by the combination of n+1 independent error terms $e_0, e_1, \ldots, e_n$, as follows:

$$\varepsilon_j = e_0 + e_j \quad (j=1,\ldots,n)$$

where $e_0 \sim (0, \rho\sigma^2)$; $e_j \sim (0, (1-\rho)\sigma^2)$; $E[e_i\, e_j] = 0$ $(i \neq j;$ $i,j = 0,1,\ldots,n)$.

The error component e_0 is common to all the measurements and is responsible for the asymptotic $(n \to \infty)$ bias in the estimate of X. For this reason e_0 is called <u>systematic error</u>. The terms e_j, $j > 0$, represent the <u>random error</u> and are uncorrelated in different measurements. Infinitely many measurements reduce to zero the effect of the random component of the error in the posterior state covariance. In other words, an infinite sequence of equicorrelated measurements is equivalent to a single measurement with

error $\varepsilon_o = e_o$. In fact in this case $\underline{H} = 1$, $\underline{\theta}^{-1} = 1/\rho\sigma^2$, and $m = 1/\rho\sigma^2$, which corresponds to the asymptotic result (I.55).

Fig. 3 can be used to compare the informativeness of n correlated observations with small variance σ_1^2 with that of n uncorrelated (or less correlated) observations with a larger variance σ_2^2. Similarly one can decide which of the following alternatives is more rewarding: to make n_1 measurements with an accurate technique, or to make $n_2 > n_1$ measurements with a less accurate procedure.

Consider the first problem with the following data: $\sigma_2^2 = 2\sigma_1^2$, $\rho_1 = 0.2$, $\rho_2 = 0$. For each experiment the quantity $\sigma^2 \cdot m$ is a measure of informativeness (see equation (I.48)), and is plotted versus n in fig. 4A. It is found that n measurements of type 1 are more informative than n measurements of type 2 if $n < 6$. The reverse is true if $n > 6$, while the two experiments are equally informative if $n = 6$.

For the second problem suppose that $\sigma_2^2 = 2\sigma_1^2$ as before, and that $\rho_1 = \rho_2 = 0.2$. The quantity $\sigma^2 \cdot m$ is plotted versus n in fig. 4B, allowing to compare the effectiveness of the two experiments for given n_1 and n_2. Alternatively, an analytical solution can be found: in order that $r_2 \geq r_1$ it must be (from equations (I.48) and (I.53):

$$\frac{1}{\sigma_2^2}\,\frac{n_2}{1+(n_2-1)\rho_2} \geq \frac{1}{\sigma_1^2}\,\frac{n_1}{1+(n_1-1)\rho_1}\,, \qquad \text{(I.56a)}$$

or:
$$n_2 \geq \frac{\sigma_2^2(1-\rho_2)\ m_1}{1-\sigma_2^2\ \rho_2\ m_1} \qquad \text{(I.56b)}$$

where m_1 equals the right side of equation (I.56a). Equation (I.56B) holds under the condition: $\sigma_2^2\ \rho_2\ m_1 < 1$. If this inequality is not satisfied experiment 1 must be prefered. This follows from $m_{2_\infty} = 1/\sigma_2^2\ \rho_2$ being the maximum value of m which is attainable from experiment 2; see equation (I.55). When $\rho_2 = 1$ one should drop the equal sign from (I.56b).

For the most general experiment considered here the effectiveness ratio r_K is, from equation (I.49):

$$r_K = (1 + \sigma_b^2\ m)^{1/2} \qquad \text{(I.57)}$$

where $m = \sum_{i=1}^{K} m_i = \sum_{i=1}^{K} \frac{1}{\sigma_i^2}\,\frac{n_i}{1+(n_i+1)\rho_i}$

By using equation (I.57) one can study more complex problems than the ones considered so far, like finding the experiment with the optimal tradeoff between cost and effectiveness. The problem of optimal experiment design is addressed in

more general terms in Section I.3; Example 4 in the same section is a continuation of the present particular case.

EXAMPLE 3. STIFFNESS ESTIMATION OF A THREE BARS TRUSS

A three bars truss is shown in fig. 5. The problem is to reduce the prior uncertainty in the actual stiffness of the bars, denoted X_1, X_2, X_3. Let $\underline{X} = [X_1, X_2, X_3]'$ be the unknown state with prior distribution $\underline{X} \sim N(\underline{X}_b, \underline{\Sigma}_b)$.

The uncertainty on $\underline{X}$ has two sources: the variability of the material quality, which affects the stiffness of all the bars, and the manufacturing uncertainty, which is independent for different bars. Along each bar the material properties are constant. Under these assumptions the stiffness of the ith bar can be expressed as:

$$X_i = X_{b_i} + (\varepsilon_o + \varepsilon_i)/l_i \ , \quad (i=1,2,3) \qquad \text{(I.58)}$$

where l_i is the length of the ith bar, X_{b_i} is the best prior estimate of X_i, ε_o is the error associated with the material uncertainty, and ε_i is a manufacturing error. ε_o and ε_i have units of force. According to equation (I.58) the effect of the errors on the stiffness is proportional to the inverse bar length. All errors are assumed to be

independent and normally distributed:

$$\varepsilon_o \sim N(0;\sigma_o^2) \ ,$$

$$\varepsilon_i \sim N(0;\alpha^2\sigma_o^2) \ , \quad E[\varepsilon_o\varepsilon_i] = 0 \quad (i=1,2,3). \tag{I.59}$$

Since $l_3 = l$ and $l_1 = l_2 = \mu l$ ($\mu > 0.5$), the a prior state covariance matrix is:

$$\underline{\Sigma}_b = \frac{\sigma_0^2}{\mu^2 l^2} \begin{pmatrix} (1+\alpha^2) & 1 & \mu \\ 1 & (1+\alpha^2) & \mu \\ \mu & \mu & \mu^2(1+\alpha^2) \end{pmatrix} \tag{I.60}$$

The experiment may include one or more measurements of the following two types.

(i) Hardness measurements. If $\underline{Z}$ is the vector of hardness measurements, the following observation model is hypothesized:

$$\underline{Z} = \underline{Z}_o + \underline{H}\,\underline{X} + \underline{\varepsilon} \ , \tag{I.61}$$

where Z_o is a known vector and the matrix $\underline{H}$ has coefficients:

$$(\underline{H})_{ij} = \begin{cases} h\, l_j & \text{if } Z_i \text{ is a hardness measure of bar } j, \\ 0 & \text{otherwise.} \end{cases}$$

For instance the matrix

$$\underline{H} = h \begin{pmatrix} l_1 & 0 & 0 \\ 0 & 0 & l_3 \\ l_1 & 0 & 0 \end{pmatrix}$$

results when Z_1 and Z_3 measure the hardness of bar 1 and Z_2 measures the hardness of bar 3. The measurement errors have normal distribution:

$$\underline{\varepsilon} \sim N \left(\underline{0};\ \underline{\Theta} = \sigma_H^2 \begin{pmatrix} 1 & \rho_H \\ \rho_H & 1 \end{pmatrix} \right) \qquad \text{(I.62)}$$

The correlation coefficients ρ_H account for the uncertainty in the relationship between hardness and stiffness (a random bias term common to all the scalar observations in $\underline{Z}$ would introduce correlation), and for the correlation due to systematic measurement errors.

(ii) Stiffness measurements. Stiffness experiment may consist of imposing compatible nodal displacements δ_1, δ_2, δ_3 (see fig. 5) and of measuring the forces F_1, F_2, F_3 required for equilibrium. Then each measurement is of the type $(\underline{F}_i, \underline{\delta}_i)$, where

$$\underline{F}_i = \underline{K}\ \underline{\delta}_i + \underline{\varepsilon}_i \tag{I.63}$$

is the load vector corresponding to the displacement vector $\underline{\delta}_i$. The coefficients of the elastic stiffness matrix $\underline{K}$ are linear functions of $\underline{X}$; with reference to fig. 5:

$$\underline{K} = \begin{pmatrix} (X_1+X_2)\cos^2\beta & (X_1-X_2)\sin\beta\cos\beta & -X_2\cos^2\beta \\ (X_1-X_2)\sin\beta\cos\beta & (X_1+X_2)\sin^2\beta & X_2\sin\beta\cos\beta \\ -X_2\cos^2\beta & X_2\sin\beta\cos\beta & X_3+X_2\cos^2\beta \end{pmatrix} \tag{I.64}$$

where $\cos\beta = 1/2\mu$. Equation (I.63) can be rewritten with $\underline{F}_i$ as a function of $\underline{X}$:

$$\underline{F}_i = \underline{H}_i\ \underline{X} + \underline{\varepsilon}_i \tag{I.65}$$

where

$$\underline{H}_i = \begin{Vmatrix} (\delta_{1i}\cos^2\beta+\delta_{2i}\sin\beta\cos\beta) & [(\delta_{1i}-\delta_{3i})\cos^2\beta-\delta_{2i}\sin\beta\cos\beta] & 0 \\ (\delta_{1i}\sin\beta\cos\beta+\delta_{2i}\sin^2\beta) & [-(\delta_{1i}-\delta_{3i})\sin\beta\cos\beta+\delta_{2i}\sin^2\beta] & 0 \\ 0 & [-(\delta_{1i}-\delta_{3i})\cos^2\beta+\delta_{2i}\sin\beta\cos\beta] & (\delta_{3i}) \end{Vmatrix} \tag{I.66}$$

As to the experiment errors we assume:

$$\underline{\varepsilon}_i \sim N \left| \underline{0};\ \underline{\Theta} = \sigma_\varepsilon^2 \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix} \right. \tag{I.67}$$

Passing to a numerical application let $l = 1$, $\mu = \sqrt{2}/2$, so that from (I.60):

$$\underline{\Sigma}_b = 2\sigma_o^2 \begin{pmatrix} (1+\alpha^2) & 1 & \sqrt{2}/2 \\ 1 & (1+\alpha^2) & \sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 & (1+\alpha^2)/2 \end{pmatrix}, \tag{I.68}$$

and from (I.66):

$$\underline{H}_i = \begin{Vmatrix} h_{1i} & h_{2i} & 0 \\ h_{1i} & -h_{2i} & 0 \\ 0 & -h_{2i} & h_{3i} \end{Vmatrix}, \text{ where } \begin{cases} h_{1i} = (\delta_{1i}+\delta_{2i})/2\ , \\ h_{2i} = (\delta_{1i}-\delta_{2i}-\delta_{3i})/2, \\ h_{3i} = \delta_{3i}\ . \end{cases} \tag{I.69}$$

The three distinct non-zero entries of $\underline{H}_i$ are controllable experiment parameters.

We compare now the effectiveness of different experiments, as measured by the ratio r in equation (I.47).

(i) Hardness measurements. If one makes a single measurement from bar 1 the measurement matrix is:

$$\underline{H} = h[\sqrt{2}/2,\ 0,\ 0]$$

with information matrix and effectiveness ratio:

$$\underline{M} = \underline{H}'\underline{\Theta}^{-1}\underline{H} = \frac{h^2}{2\sigma_H^2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix};$$

$$r = |\underline{I}_3 + \underline{\Sigma}_b\ M|^{1/6} = [1 + (1+\alpha^2)\ \frac{\sigma_o^2}{\sigma_H^2}\ h^2]^{1/6} \qquad \text{(I.70)}$$

As one would expect, r increases with the prior state variance and with the signal-to-noise strength h, and decreases with the variance of the measurement error. The effectiveness ratio is the same for a single measurement from bars 2 or 3 (i.e., for $\underline{H} = h[0,\ \sqrt{2}/2,\ 0]$ or for $\underline{H} = h[0,\ 0,\ 1]$, respectively).

When n hardness measurements are made from the same bar equation (I.70) generalizes to:

$$r_n = \left[1 + (1+\alpha^2) \frac{\sigma_o^2}{\sigma_H^2} h^2 \frac{n}{1+(n-1)\rho_H} \right]^{1/6} \qquad (I.71)$$

with $\rho_H \geq -1/(n-1)$. When $n \to \infty$ equation (I.71) yields:

$$r_\infty = \left[1 + \frac{1+\alpha^2}{\rho_H} \frac{\sigma_o^2}{\sigma_H^2} h^2 \right]^{1/6} , \qquad (I.72)$$

which is finite for $\rho_H > 0$. Because of the similarity with the case studied in EXAMPLE 2 (compare equation (I.71) with equations (I.48) and (I.54)), the plots in fig. 3 can be used also here to compare the effectiveness of experiments obtained by varying ρ_H and n (replace ρ by ρ_H in fig. 3). Equations (I.71) and (I.72) are valid also when making n measurements from bar 2 or from bar 3.

A different situation is faced when <u>making measurements from different bars</u>. Suppose we measure once the hardness of bar 1, and once the hardness of bar 2. Then

$$\underline{H} = \frac{\sqrt{2}}{2} h \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

and $r = \left(1 + 2a(1+\alpha^2-\rho_H) + a^2\alpha^2(2+\alpha^2)(1-\rho_H^2)\right)^{1/6} \qquad (I.73)$

where $$a = \frac{h^2\sigma_o^2}{\sigma_H^2(1-\rho_H^2)} \quad .$$

It is interesting to compare the effectiveness of this experiment with the effectiveness of making two measurements from the same bar. Specializing equation (I.71) for n = 2 gives:

$$r_2 = \left(1 + 2a(1+\alpha^2-\rho_H-\rho_H\alpha^2)\right)^{1/6} \qquad (I.74)$$

For easy reference denote "experiment 1" measuring twice from the same bar and "experiment 2" measuring once from bar 1 and once from bar 2. Experiment 1 is more effective when $r_2/r > 1$, i.e., when

$$-1 \leq \rho_H \leq \frac{a(2+\alpha^2)}{a(2+\alpha^2)-2} \quad \text{and} \quad 0 \leq a(2+\alpha^2) \leq 1 \; ; \qquad (I.75)$$

otherwise experiment 2 should be preferred. In the plane XY where $X=a(2+\alpha^2)$ and $Y = \rho_H$ in $[-1,1]$, fig. 6 shows the regions where each experiment is more informative than the other, assuming $\alpha > 0$. When $\alpha = 0$ the stiffnesses are perfectly correlated and there is no difference between experimenting on different bars or on the same bar ($r = r_2$). As α increases, so does $a(2+\alpha^2)$; the correlation between

the bar stiffnesses decreases (see equation (I.60)) and the second experiment becomes more advantageous.

From this example one can intuitively extrapolate that when the measurement errors are positively correlated and the quantities to be measured are identically distributed the best policy is to allocate the measurements as evenly as possibly.

In EXAMPLE 2 the state of nature was a scalar. In that case any effectiveness measure based on the ratio $L(\sigma_b^2)/L(\sigma_a^2)$, where $L(\cdot)$ conforms one of the criteria (I.43) - (I.46), does not differ qualitatively from the ratio (I.47). (The determinant, the trace and the maximum diagonal entry of a scalar are all the same). This is no longer true when the state of nature is an unknown vector. For instance, the truss in fig. 5 might be designed to carry only vertical loads, so that its performance is best described by the stiffness coefficient $X_4 = F_2/\delta_2$ (while $F_1 = F_3 = 0$). The effectiveness of an experiment should then be judged from the reduction of uncertainty in X_4, as it is shown next in some detail. From linear elasticity and for $\mu = \sqrt{2}/2$ (see fig. 5) it is:

$$X_4 = 4 \frac{X_1 X_2 X_3}{X_1 X_2 + 2X_2 X_3 + 2X_3 X_1} \quad .$$

In order to use the results of linear estimation theory, one can truncate the series expression of X_4 aroung the prior mean state $\underline{X}_b$:

$$X_4 \approx X_4\Big|_{\underline{X}=\underline{X}_b} + \sum_{i=1}^{3} \frac{\partial X_4}{\partial X_i}\Big|_{\underline{X}=\underline{X}_b} (X_i - X_{i_b}) = X_4\Big|_{\underline{X}=\underline{X}_b} + \underline{d}'(\underline{X}-\underline{X}_b) \quad (I.76)$$

where $\underline{d}' = [d_1, d_2, d_3]$

and $$d_1 = \frac{\left(X_4\Big|_{\underline{X}=\underline{X}_b}\right)^2}{2X_{1b}^2} ; \; d_2 = \frac{\left(X_4\Big|_{\underline{X}=\underline{X}_b}\right)^2}{2X_{2b}^2} ; \; d_3 = \frac{\left(X_4\Big|_{\underline{X}=\underline{X}_b}\right)^2}{4X_{3b}^2} .$$

If the bars have the same prior mean stiffness per unit length, then

$$\underline{X}_b = X_{1_b} \begin{pmatrix} 1 \\ 1 \\ \sqrt{2}/2 \end{pmatrix} ; \quad X_4\Big|_{\underline{X}=\underline{X}_b} = \frac{2\sqrt{2}\, X_{1_b}}{1 + 2\sqrt{2}} ; \; d_1=d_2=d_3 = \frac{4}{9 + 4\sqrt{2}} .$$

Denote $\underline{M}$ the Fisher information matrix relative to $\underline{X}$; then the effectiveness ratio (I.48) for the linearized stiffness X_4 is (in this case we use a ranking criterion of the type (I.46)):

$$r = \sigma_{4_b}/\sigma_{4_a} = \sigma_{4_b}\left[\, \underline{d}'(\underline{\Sigma}_b^{-1}+\underline{M})^{-1}\underline{d}\right]^{-1/2} \qquad (I.77)$$

where $\underline{\Sigma}_b$ is the matrix defined in (I.68). Equation (I.77) has the inconvenience of requiring two matrix inversions. For a small number of measurements the posterior variance $\sigma^2_{4_a}$ can be found more easily through direct application of the Gauss-Markov theorem; namely of equation (I.3).

If n hardness measurements are made from bar 1 in the form $Z_i = Z_o + h\, l_1\, X_1 + \varepsilon_i$ (i=1,...,n) with $\mathrm{Cov}[X_1,\varepsilon_i]=0$, then one can follow the same procedure as for the general linear measurement model (see the steps which led to equation (I.24)), except that now

$$\underline{S}_{\underline{X}^*} = \sigma^2_{4_b} \quad ;$$

$$\underline{S}_{\underline{Z}^*} = \underline{H}\,\underline{\Sigma}_b\,\underline{H}' + \underline{\Theta} \quad ;$$

$$\underline{S}_{\underline{X}^*\underline{Z}^*} = \mathrm{Cov}[X_4,\underline{Z}] = d_1[(\underline{\Sigma}_b)_{11}+(\underline{\Sigma}_b)_{12}+(\underline{\Sigma}_b)_{13}][1,1,\ldots,1] \ ,$$

where, for $l = \sqrt{2}/2$, it is

$$\underline{H} = \frac{\sqrt{2}}{2}\, h \begin{array}{c} 1 \\ \\ \\ \\ \\ n \end{array} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & 0 & 0 \end{pmatrix} .$$

and $\underline{\Sigma}_b$ and $\underline{\theta}$ are the same as in equations (I.68) and (I.62).

We study first the case n = 1 for which

$$\underline{S}_{\underline{Z}^*} = \sigma_Z^2 = \sigma_H^2 + h^2 \sigma_o^2(1+\alpha^2) \ ,$$

$$\underline{S}_{\underline{X}^*,\underline{Z}^*} = \mathrm{Cov}[X_4,Z] = 2d_1 \sigma_o^2[(1+\alpha^2)+1+\sqrt{2}/2] \ ,$$

and from equation (I.3):

$$\sigma^2_{4a} = \sigma^2_{4b} - \frac{d_1^2 \sigma_o^4}{\sigma_H^2 + h^2 \sigma_o^2 (1+\alpha^2)} (4+\sqrt{2}+2\alpha^2)^2 \qquad \text{(I.78)}$$

Going through the same procedure, the posterior variance of X_4 after making one observation of the hardness of bar 3 is:

$$\sigma^2_{4a} = \sigma^2_{4b} - \frac{d_1^2 \sigma_o^4}{\sigma_H^2 + h^2 \sigma_o^2(1+\alpha^2)} (1+ 2\sqrt{2}+\alpha^2)^2 \qquad \text{(I.79)}$$

By comparison of equations (I.78) and (I.79) it is seen that for any given α the maximum variance reduction occurs when measuring the hardness of bar 1. The reason for that is the higher prior variance of X_1 as compared with the prior variance of X_3. Since $d_1 = d_3$, it is obvious that one should measure from the bar with maximum prior stiffness variance. This result contrasts with the previous finding

that measurements from any bar are equally effective when compared through the ratio (I.47).

If two measurements are made from bar 3, then

$$\underline{S}_{\underline{Z}^*} = \begin{pmatrix} \sigma_H^2 + h^2\,\sigma_o^2(1+\alpha^2) & \sigma_H^2\,\rho_H + h^2\,\sigma_o^2(1+\alpha^2) \\ \sigma_H^2\,\rho_H + h^2\,\sigma_o^2(1+\alpha^2) & \sigma_H^2 + h^2\,\sigma_o^2(1+\alpha^2) \end{pmatrix},$$

$$S_{\underline{X}^*,\underline{Z}^*} = d_1\,\sigma_o^2(2\sqrt{2} + 1 + \alpha^2)\ [1,1] \quad ,$$

and $$\sigma^2_{4a} = \sigma^2_{4b} - \frac{2\,d_1^2\,\sigma_o^4}{\sigma_H^2(1+\rho_H)+2h^2\,\sigma_o^2(1+\alpha^2)}\,(2\sqrt{2}+1+\alpha^2)^2 . \quad (I.80)$$

If $\rho_H = 1$ there is no advantage in making the second measurement, and equation (I.80) coincides with equation (I.79); if $\rho_H = -1$, the two measurements remove completely the statistical uncertainty on the hardness of bar 3. This last case gives the minimum value of σ^2_{4a} when using hardness measurements from bar 3 only.

(ii) Stiffness measurements. We compare the effectiveness of six different experiments, in which the displacements are the control variables (see fig. 5):

$$(1)\quad \begin{cases} \delta_1 = 1 \\ \delta_2 = 0 \\ \delta_3 = 0 \end{cases} \qquad (2)\quad \begin{cases} \delta_1 = 0 \\ \delta_2 = 1 \\ \delta_3 = 0 \end{cases} \qquad (3)\quad \begin{cases} \delta_1 = 0 \\ \delta_2 = 0 \\ \delta_3 = 1 \end{cases}$$

$$(4)\quad \begin{cases} \delta_1 = 1 \\ \delta_2 = 1 \\ \delta_3 = 0 \end{cases} \qquad (5)\quad \begin{cases} \delta_1 = 1 \\ \delta_2 = 0 \\ \delta_3 = 1 \end{cases} \qquad (6)\quad \begin{cases} \delta_1 = 0 \\ \delta_2 = 1 \\ \delta_3 = 1 \end{cases}$$

The deformed configurations corresponding to these displacements are shown in fig. 7. The associated measurement matrices $\underline{H}_{(1)},\ldots,\underline{H}_{(6)}$ are, from equation (I.69):

$$\underline{H}_{(1)} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & -1 & 0 \end{bmatrix} \quad ; \quad \underline{H}_{(2)} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad ; \quad \underline{H}_{(3)} = \frac{1}{2}\begin{bmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$$

$$\underline{H}_{(4)} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad ; \quad \underline{H}_{(5)} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad ; \quad \underline{H}_{(6)} = \frac{1}{2}\begin{bmatrix} 1 & -2 & 0 \\ 1 & 2 & 0 \\ 0 & 2 & 2 \end{bmatrix}$$

For simplicity we assume that the measurement error covariance matrix (I.67) is diagonal. With $\underline{\Sigma}_b$ given by equation (I.68) the effectiveness ratios

$r_{(i)} = |\underline{I}_3 + \underline{\Sigma}_b \underline{H}_i' \underline{\Theta}^{-1} \underline{H}_i|^{1/6}$ are:

$$r_{(1)} = r_{(2)} = \{1+5C(1+\alpha^2)+6C^2[-1+(1+\alpha^2)^2]\}^{1/6} ;$$

$$r_{(3)} = \{1+C[2\sqrt{2}+5(1+\alpha^2)]+4C^2[-1+(1+\alpha^2)^2]\}^{1/6} ;$$

$$r_{(4)} = \{1+8C(1+\alpha^2)\}^{1/6} ;$$

$$r_{(5)} = \{1+4C(1+\alpha^2)+4C^2[-1+(1+\alpha^2)^2]\}^{1/6} ;$$

$$r_{(6)} = \{1+C[3\sqrt{2}+16(1+\alpha^2)]+6C^2[-8-\sqrt{2}+\sqrt{2}(1+\alpha^2)+8(1+\alpha^2)^2]$$

$$+40C^3[2-3(1+\alpha^2)+(1+\alpha^2)^3]\}^{1/6} ;$$

where $C = \frac{1}{2}\sigma_o^2/\sigma_\varepsilon^2$.

As α^2 varies from 0 to ∞ the prior variances of the bar stiffnesses increase and their correlations decrease. Plots of $r_{(i)}^6$ (i=1,...,6) as functions of C for $\alpha = 0$ (perfect stiffnesses correlation) are shown in fig. 8. The same functions are plotted on semilogarithmic paper in fig. 9 for $\alpha^2 = 0.25$ and in fig. 10 for $\alpha^2 = 1$.

The effectiveness of the experiments increases with C (i.e., with the variance ratio $\sigma_o^2/\sigma_\varepsilon^2$) and with $(1+\alpha^2)$, which for fixed C is proportional to the prior stiffness

variances. When $\alpha^2 \neq 0$ the stiffnesses are not perfectly correlated (the correlation coefficient being $1/(1+\alpha^2)$), and this favors the experiments which induce deformations in more than one bar; so the experiment (4), which is the only one which deforms one bar gets the least reward from increasing α^2. At the other extreme (6) is by far the best experiment, since it deforms considerably and simultaneously all the bars; each of the experiments (1), (2), (3), (5) stresses only two bars (see fig. 7).

If $\alpha = 0$ the preference order of the experiments remains the same for all C; instead the convenience of performing one experiment versus another may change with C when $\alpha^2 > 0$. This is due to the higher powers of C in the expressions for $r_{(i)}$ having non-zero coefficients if $\alpha \neq 0$; nevertheless for all values of α^2 there is an initial range for C in which the high power terms are negligible and the preferential order is the same as for $\alpha = 0$. The physical explanation of this behavior is that as C increases the measurements become more accurate (smaller measurement error variance as compared to the prior variance) and more is gained in terms of the effectiveness ratio (I.50) from deforming more bars; at the same time it makes little difference which bars are deformed and to what extent. On the contrary, for small values of C the measurements are quite noisy and the effectiveness of the experiment is

controlled by the signal-to-noise strength, which increases with the deformation. From fig. 7, experiment (4) imposes a large deformation to bar 1, which fact makes (4) quite competitive in the low range of C. (Also note that experiment (4) can be performed on bar 1 before it is assembled into the system.)

In closing this paragraph on the evaluation of experiments we recall the major qualitative conclusions. Experiments can be judged according to different criteria; when using the ratio (I.47) as a measure of effectiveness it was found that:

(i) The presence of systematic errors prevents one from obtaining asymptotically deterministic estimates as the number of observations increases, see fig. 3.

(ii) For scalar states of nature the relative effectiveness of two experiments depends on their signal-to-noise ratios, on the number of measurements in each experiment, and on the percentages of systematic error in the total measurement error (fig. 4).

(iii) For vector-valued states the same considerations hold when measuring only one component of the state vector $\underline{X}$. When the measurements are functions of more than one state component general conclusions are difficult to derive. Nevertheless, if measurements are accurate it is convenient to have an observation vector which is

informative on the largest possible number of components of $\underline{X}$. Instead, if the measurements are unreliable one should try to increase the signal strength, while the measurements being noninformative on some state component becomes relatively unimportant (figs. 7 - 10).

(iv) For vector valued states the relative effectiveness of two experiments depends also on the ranking criterion; however, the qualitative considerations at point (iii) are believed to be valid under all sensible choices.

We considered here mainly experiment evaluation criteria which measure the reduction of uncertainty. Other criteria may be more appropriate, particularly for pre-posterior analyses in a decision process; see the comments which precede and follow equation (I.42).

I.3 OPTIMAL EXPERIMENT DESIGN

The first section of this chapter dealt with the analysis of estimation experiments. In Section I.2 some evaluation criteria were introduced and a few principles for improving the effectiveness of experiments were found by working out specific examples. The design of optimal experiments is considered briefly in the present section.

First, we give a more precise definition of "experiment", and we clarify in which sense the word "optimal" is used.

Given an n-dimensional state vector $\underline{X}$ as the object of estimation, and within the class of linear measurements we call "design of an experiment" (or simply an "experiment", denoted E) the collection of scalars

$$N,\ H_{ij} \quad , \quad i = 1,\ldots,N;\ j = 1,\ldots,n \qquad (I.81)$$

This set of variables defines the linear observation equation:

$$\underline{Z} = \underline{Z}_o + \underline{H}\,\underline{X} + \underline{\varepsilon}$$

where $\underline{Z} = [Z_1,\ldots,Z_N]'$ is the observation vector, $\underline{H} = [H_{ij}]$ is the observation matrix, $\underline{Z}_o$ is a known vector (possibly a function of E), and $\underline{\varepsilon}$ is a random vector of non-controllable errors. Two experiments E_1 and E_2 are said to be different, $E_1 \neq E_2$, if they differ by at least one variable in the set (I.81). This definition of experiment implicitly assumes that given (N, $\underline{H}$) both $\underline{Z}_o$ and the probability distribution of $\underline{\varepsilon}$ are defined uniquely. The design of an optimal experiment consists in finding the parameter values N* and $\underline{H}$* which extremize a function of E. There are several levels

at which this problem can be formulated, depending on the choice of the objective function. Two possible formulations are given below.

(i) Restricted formulation. Given a set S of feasible designs, an experiment E* is optimal if no experiment $E \in S$ exists such that $E > E^*$ according to a criterion of the type $L[\underline{\Sigma}_a(E)] = \min$, where $\underline{\Sigma}_a$ is the posterior state covariance. (See equations (I.43) - (I.46) for some examples). The criterion establishes a preference order for the elements (N, $\underline{H}$) of S, which makes the optimum problem well posed in a mathematical sense.

This formulation has the drawback of accounting in a simplistic way for the cost of experimentation, which defines the feasible set S. Optimization problems of this type are discussed at length by Fedorov (1972).

A dual (although not completely equivalent) restricted formulation exists, in which a maximum is imposed to the loss function $L[\underline{\Sigma}_a(E)]$, resulting in a constraint on E. The optimum experiment is the one which satisfies the constraint at minimum cost. While the former is a cost-constrained formulation, the latter might be called an informativeness-constrained formulation.

(ii) Maximum expected utility formulation. The maximization of the expected utility has become a classical objective in Bayesian statistical decision theory (see for

instance Raiffa and Schlaifer (1961); Raiffa (1968); DeGroot (1970) chapter 8). In this formulation the optimal experiment maximizes a function of the Bayes risk (see Paragraph I.1.1), of the experiment cost, as well as of more subjective attributes, called the expected utility. The loss function $L[\underline{\Sigma}_a(E)]$ in the restricted formulation becomes one of the arguments of the utility function.

The critical point of this formulation is the construction of the utility function. Much work has been done recently in this area (see, among others: Raiffa (1968); Fishburn (1968); Hampton, et al. (1973), but the problem has not only a technical aspect: it requires answering difficult questions such as what should be optimized and for whom the experiment should be optimal (Churchman (1961); Rosenblueth (1973).

Without going into the general theory of optimal estimation experiments (the interested reader will find a comprehensive review of methods and results in the aforementioned book by Fedorov) two examples are presented here. The first one completes Example 2 in Section I.2 by finding the optimal combination of three types of measurements which reaches a target level of informativeness at minimum cost (restricted approach). The second example concerns a problem of optimal estimation which is of interest in various fields of engineering; in this case the objective is to

maximize the expected utility.

EXAMPLE 4. A PROBLEM OF OPTIMAL EXPERIMENT DESIGN IN SCALAR ESTIMATION

For preliminaries and notations refer back to Example 2 in Section I.2. The problem of optimal experiment design, aimed at "best" estimating the scalar state X, is given the following restricted formulation. Each scalar measurement has a known cost associated with it, and an experiment is optimal if its effectiveness ratio (I.57) exceeds a prefixed target value at minimum cost. Let C_{i_1} (i = 1,...,K) denote the cost of the first measurement with facility i, and C_{i_2} denote the cost of any additional measurement with the same facility. The cost of the entire experiment, C, is a function of the number of measurements with each facility: n_i (i = 1,...,K). Explicitly:

$$C = \sum_{i=1}^{K} [C_{i_1} + (n_i - 1)\, C_{i_2} + (C_{i_2} - C_{i_1})\, \delta(n_i)] \; ; \qquad (I.82)$$

where

$$\delta(n) = \begin{cases} 0 \text{ if } n > 0 \; , \\ 1 \text{ if } n = 0 \; . \end{cases}$$

For the purpose of exemplification consider three facilities being available (K=3), the target: $r_3 \geq 5$ (or $\sigma_a^2 \leq \sigma_b^2/25$), and the following measurement data:

Facility, i	σ_i^2/σ_b^2	ρ_i	C_{i_1}	C_{i_2}
1	0.5	0	5	2
2	0.1	0.2	15	2
3	0.1	0.4	10	1

Table 2. Data for Example 4.

From equation (I.57) the condition $r_3 \geq 5$ is equivalent to

$$n_1 + \frac{25\ n_2}{4 + n_2} + \frac{25\ n_3}{3 + 2n_3} \geq 12 \qquad \text{(I.83)}$$

with n_1, n_2, n_3 non-negative integers. The objective function is:

$$25 + 2n_1 + 2n_2 + n_3 - 3\ \delta(n_1) - 13\ \delta(n_2) - 9\ \delta(n_3) = \min. \qquad \text{(I.84)}$$

The problem (I.83, I.84) has a single optimum at $n_1 = 0$, $n_2 = 4$, $n_3 = 0$, where $C = 21$ and $r_3 = \sqrt{26}$.

For each type of experiment Figure 11 displays the informativeness measure $\sigma_b^2 m_i$ (see equation (I.57)) versus the associated experiment cost C_i. Although these curves do not provide a direct way of finding the optimum strategy for a fixed minimum value of $\sigma_b^2 m = \sigma_b^2 (m_1 + m_2 + m_3)$, it is easy to see that experiments of type 1 predominate for small (less than 6) and for very large $\sigma_b^2 m$ values. Experiments of type 3 have best information-to-cost ratio for intermediate values, say for $6 < \sigma_b^2 m < 20$. Finally, when $\sigma_b^2 m$ exceeds 20 but is not very large, experiments of type 2 give the highest reward for fixed cost. The combined use of different types of experiments is optimal near the transition values, and certainly when $\sigma_b^2 m$ is very large.

EXAMPLE 5. OPTIMAL SAMPLING OF CONTINUOUS FUNCTIONS

Let $\underline{Y}$ be the generic point of an n-dimensional region $D \subset R^n$, and let $S(\underline{Y})$ be an unknown scalar function of $\underline{Y}$, defined in D.(*) From a Bayesian viewpoint $S(\underline{Y})$ is a random field with n-dimensional parameter. In second-moment theory the prior information on $S(\underline{Y})$ consists in the prior

---

(*) The case when $S(\cdot)$ is an unknown vector function can be treated similarly.

mean value function $E[S(\underline{Y})] = S_b(\underline{Y})$; $\underline{Y}$ D, and in the prior covariance function

$$Cov_b(\underline{Y}_1, \underline{Y}_2) = Cov[S(\underline{Y}_1), S(\underline{Y}_2)] \; ; \quad \underline{Y}_1, \underline{Y}_2 \in D.$$

(If $S(\cdot)$ is a vector function, Cov $(\cdot,\cdot)$ is a matrix function.) Let N be the number of points at which $S(\underline{Y})$ is measured in the experiment. These points are denoted $\underline{Y}_1^*, \underline{Y}_2^*, \ldots, \underline{Y}_N^*$ and are constrained to belong to a region $D^E \subset D$. For fixed N the set $\{\underline{Y}_i^*\}$ is called the "experiment trace". Assume that the purpose of sampling is to reduce the variance of the weighted spatial mean of S over a region $D^S \subset D$:

$$q = \int_{D^S} w(\underline{Y}) \; S(\underline{Y}) \; d\,\underline{Y}, \tag{I.85}$$

where $w(\cdot)$ is a deterministic weight function. The prior mean and variance of q are:

$$q_b = \int_{D^S} w(\underline{Y}) \; S_b(\underline{Y}) \; d\,\underline{Y} \; ; \tag{I.86a}$$

$$\sigma_b^2 = \int_{D^S} \int_{D^S} w(\underline{Y}_1) \; w(\underline{Y}_2) \; Cov_b(\underline{Y}_1, \underline{Y}_2) \; d\,\underline{Y}_1 \; d\,\underline{Y}_2 \, . \tag{I.86b}$$

Similarly, the posterior mean and variance are:

$$q_a = \int_{D^S} w(\underline{Y})\ S_a(\underline{Y})\ d\ \underline{Y}\ ; \qquad \text{(I.87a)}$$

$$\sigma_a^2 = \int_{D^S}\int_{D^S} w(\underline{Y}_1)\ w(\underline{Y}_2)\ Cov_a(\underline{Y}_1,\underline{Y}_2)\ d\ \underline{Y}_1\ d\ \underline{Y}_2\ , \qquad \text{(I.87b)}$$

where $S_a(\cdot)$ and $Cov_a(\cdot,\cdot)$ are the posterior Bayesian mean and the posterior error covariance functions of $S(\underline{Y})$. In accordance with equation (I.47) the effectiveness of the experiment is measured by the ratio

$$r = \sigma_b/\sigma_a\ . \qquad \text{(I.88)}$$

Under a cost-constraint (here associated with a constraint on N) the optimal experiment is defined by the trace $\{\underline{Y}_i^*\}$ (i=1,...,N) which maximizes r. Alternatively and more generally, N is let free and optimized with $\{\underline{Y}_i^*\}$.

Problems of this type arise when sampling material properties of spatially distributed structural systems (beams, plates, etc.), in soil sampling for settlement estimation (Diaz Padilla and Vanmarcke, 1973), in the design of raingage networks for best monitoring mean areal rainfalls

(Bras, et al., 1974), in forestry surveying (Matern, 1960), to name but a few fields of application.

For mathematical convenience, the problem is discretized by choosing a set of m representative points in D, which we denote $\underset{\sim}{Y}_1,\ldots,\underset{\sim}{Y}_m$. Also, for sake of notational simplicity, let $D^S = D^E = D$. To each point $\underset{\sim}{Y}_i$ we associate an influence coefficient a_i (i=1,...,m) such that:

$$\sum_i a_i = \int_D d\,\underline{Y} \ .$$

Then the prior and the posterior mean value and variance of q become, in approximation:

$$q_b \approx \underline{v}'\,\underline{S}_b \quad ; \quad \sigma_b^2 \approx \underline{v}'\,\underline{\Sigma}_b\,\underline{v} \quad ; \qquad \text{(I.89a)}$$

$$q_a \approx \underline{v}'\,\underline{S}_a \quad ; \quad \sigma_a^2 \approx \underline{v}'\,\underline{\Sigma}_a\,\underline{v} \quad ; \qquad \text{(I.89b)}$$

where

$$\underline{v} = \begin{Bmatrix} (i=1) \\ \vdots \\ w(\underset{\sim}{Y}_i)\,a_i \\ \vdots \\ (i=m) \end{Bmatrix} \quad ; \quad \underline{S}_b = \begin{Bmatrix} (i=1) \\ \vdots \\ S_b(\underset{\sim}{Y}_i) \\ \vdots \\ (i=m) \end{Bmatrix} \quad ; \quad \underline{\Sigma}_b = \begin{Bmatrix} \\ \mathrm{Cov}_b(\underset{\sim}{Y}_i, \underset{\sim}{Y}_j) \\ \\ i,j=1,\ldots,m \end{Bmatrix} \ .$$

Similarly for $\underline{S}_a$ and $\underline{\Sigma}_a$. The sample points $\underline{Y}_1^*,\ldots,\underline{Y}_N^*$ must coincide with points in the set $\{\underline{\tilde{Y}}_i\}$; but they need not be distinct (i.e., one can sample more than once at the same point). An example of discretization and of experiment spectrum on the plane is shown in Figure 12.

A linear measurement model is assumed in the form:

$$Z_j = Z_{o_j} + h \cdot S(\underline{Y}_j) + \varepsilon_j \; ; \quad \underline{Y}_j \in \{\underline{\tilde{Y}}_i\} \quad (j=1,\ldots,N), \qquad \text{(I.90a)}$$

or, in vector notation,

$$\underline{Z} = \underline{Z}_o + \underline{H}\,\underline{S} + \underline{\varepsilon} \quad , \qquad \text{(I.90b)}$$

where $\underline{Z}_o$ is a given N-vector, $\underline{H}$ is a given (N x m) matrix, and $\underline{\varepsilon}$ is a zero-mean error vector with diagonal covariance matrix $\underline{\Theta} = \text{diag}[\sigma^2_{\varepsilon_i}]$; $\underline{\varepsilon}$ is assumed uncorrelated with $\underline{S}$.

Within the class of linear estimators the error mean square is minimized by the estimator (I.23). Straightforward application of equations (I.23) and (I.26) with $\underline{\Gamma} = \underline{0}$ yields:

$$\underline{S}_a = \underline{S}_b + \underline{\Sigma}_b\,\underline{H}'(\underline{H}\,\underline{\Sigma}_b\,\underline{H}' + \underline{\Theta})^{-1}(\underline{Z} - \underline{Z}_o - \underline{H}\,\underline{S}_b) \quad , \qquad \text{(I.91a)}$$

$$\underline{\Sigma}_a = \underline{\Sigma}_b - \underline{\Sigma}_b \underline{H}' (\underline{H} \underline{\Sigma}_b \underline{H}' + \underline{\Theta})^{-1} \underline{H} \underline{\Sigma}_b \qquad \text{(I.91b)}$$

This completes the second-moment analysis of a given experiment. The optimal design problem requires finding the sample size N and the (N x m) matrix $\underline{H}$ which maximize a given utility function. The structure of $\underline{H}$ is constrained by the condition:

$$(\underline{H})_{ij} = h\ \delta(\underline{\tilde{Y}}_i - \underline{\tilde{Y}}_j) = \begin{cases} h & \text{if } Z_j \text{ measures } S(\underline{\tilde{Y}}_j)\ , \\ 0 & \text{otherwise}\ , \end{cases}$$

where h is a given constant. Under these constraints the measurement matrices $\underline{H}$ and the experiment traces $\{\underline{Y}^*_i\}$ are in one-to-one relationship, so that optimizing with respect to $\underline{H}$ is equivalent to optimizing with respect to $\{\underline{Y}^*_i\}$.

The utility function to be maximized over all the possible choices of $(N,\underline{H})$ is written parametrically in the form:

$$U[E] = C(N,\underline{H}) - C_\sigma \underline{V}' \underline{\Sigma}_a \underline{V}\ , \qquad \text{(I.92)}$$

where $C(\cdot)$ is the experiment cost and C_σ is a constant (but as yet undefined) parameter giving the monetary equivalent of a unitary change in the posterior variance of q. Although (I.92) is not by itself a very general expression for the utility function, solving the optimization problem parametrically with respect to C_σ gives a set of optimal solutions in the form of the minimum posterior variance per given experiment cost (the so-called transformation curve). This allows one to make easy sensitivity analyses of experiment cost versus variance reduction. Finally, depending on the value of the tradeoff coefficient C_σ the decision maker considers appropriate, the best course of action is chosen out of the set of potentially optimal solutions.

With this formulation, Bras, et al. (1974) have studied the optimal location of raingage stations over a given geographical region. The effectiveness of the design was measured by the posterior variance of the mean areal precipitation during a storm with specified general characteristics. The precipitation intensity was modeled as a two-dimensional random field with known mean and covariance function.

Some numerical results are presented in the aforementioned report, to which the reader is referred also for a detailed account of the optimization procedure. The maximization of (I.92) is carried out in two alternating

stages: first N is fixed and $\underline{H}$ is optimized under this constraint; then N is changed and $\underline{H}$ is optimized again, until the best combination $(N,\underline{H})$ is found. The most difficult and time consuming step is the optimization of $\underline{H}$ for given N. This is essentially a problem of integer programming. A theorem in Fedorov (1972) proves useful for defining a search strategy among all the possible choices of $\underline{H}$, which requires a relatively modest computation effort. The essential feature which makes the whole optimization algorithm operational is that it avoids using equation (I.91b) for updating the posterior error covariance matrix whenever $\underline{H}$ is modified to improve the design.

I.4. PROOF LOADING AND THE ANALYSIS OF CENSORED DATA

From the viewpoint of statistics, proof loading is a way of collecting censored data on an unknown state of nature. In the simplest situation a system with unknown "resistance" R is observed while responding to a control input. If the behavior is satisfactory it is $R \geq R_{PL}$, where R_{PL} is the minimum resistance for survival.

In the context of structural safety, censored data of this type are not collected to reduce the dispersion of the state estimate error (as was the case for uncensored data), but to assure a minimum range of safe performance

and to detect the possible presence of major construction defects.

The theory of proof loading for structural reliability has not gone further than the so-called fundamental case, which concerns single-mode-of-failure systems with failure condition $R - S < 0$ (R = modal resistance, S = modal load). In paragraph I.4.1 this case is given a more general formulation than it is commonly found in the literature. Some guidelines for the analysis of more complex proof loading problems are given in paragraph I.4.3. Paragraph I.4.2 introduces a general probabilistic model for construction defects in structural systems. This model is used in Example 6, which shows how the probability of different defects combinations and the reliability of a series system are modified by processing proof load data through Bayes' theorem.

I.4.1 The Fundamental Case

Classical formulation. The fundamental problem of reliability theory concerns a one-failure-mode system reaching its limit state when the random load S equals the random resistance R. As well known, the failure probability of such a system is, in a time-invariant situation and with obvious notations:

$$P_f = \int_0^\infty f_S(s)\ F_R(s)\ d\ s\ . \qquad \text{(I.93a)}$$

P_f is also the failure probability without proof loading, denoted P_{f_b}. Consider now the same problem with proof loading, and denote R_{PL} the minimum resistance required to survive the experiment. Survival induces a truncation in the probability density $f_R(\cdot)$ at $R = R_{PL}$. The reliability consequences of survival are well known for this simple case (Shinozuka, 1969; Shinozuka, Yang and Heer, 1969; Sexsmith, 1969), the posterior failure probability being (see Figure 13):

$$P_{f_a} = [1 - F_R(R_{PL})]^{-1} \int_{R_{PL}}^\infty f_S(s)[F_R(s) - F_R(R_{PL})]\ d\ s\ . \qquad \text{(I.93b)}$$

The probability of failure during proof loading is:

$$P_{f_{PL}} = F_R(R_{PL})\ . \qquad \text{(I.93c)}$$

It is also easy to prove the following inequalities:

$$P_{f_{PL}} + P_{f_a} \geq P_{f_b} \tag{I.94a}$$

$$P_{f_a} \geq P_{f_b} \tag{I.94b}$$

which become equalities if $P_{f_{PL}} = 0$. The larger R_{PL}, the larger $P_{f_{PL}}$. Therefore, high proof loads may be uneconomical due to the expected loss from proof loading failure; on the other hand, the larger R_{PL} the smaller P_{f_a}, which is desirable if failure under service conditions involves a major expected cost. The tradeoff between proof loading cost and increased posterior reliability, and the associated optimal choice of R_{PL} are studied in the aforementioned papers by Shinozuka and Sexsmith.

The fundamental problem in its classical formulation is not easily extended to proof loading in a multidimensional random space (several random load and/or resistance parameters). For this purpose the following generalization is useful.

A generalized formulation. Let $G_b(S)$ denote the probability of failure under a modal load S, before proof loading. Since failure may occur either for a too large positive load, or for a too small negative load (for instance a bar may fail in tension or compression, a beam may fail for excessive positive or negative moment, and so on), S is

not restricted in sign and $G_b(S)$ may look like the solid line in Figure 14. (Note that in general $G_b(S)$ is <u>not</u> a probability distribution.)

When negative values of S are neglected $G_b(\cdot)$ may coincide with the prior CDF $F_R(\cdot)$ of the classical formulation. This is true only if failure of a system with resistance R = r implies failure of all (otherwise identical) systems with resistance $0<R<r$, for any $r>0$. As a consequence $G_b(\cdot)$ is a nondecreasing function of S in the positive semiaxis.

In the general case, the proof loading experiment consists in applying a modal load S which attains all the values in the interval $[S_{PL}^m, S_{PL}^M]$, with no sign restriction on S_{PL}^m and S_{PL}^M. The analysis of this problem requires calculating the posterior probability of failure function $G_a(\cdot)$, and is considerably more complicated than the analysis of the classical proof loading model. Formally one can use Bayes' theorem:

$$G_a(s) = G_b(s) \frac{\Pr\{ L([S_{PL}^m, S_{PL}^M]) \mid C(s)\}}{\Pr\{ L([S_{PL}^m, S_{PL}^M])\}}, \quad \text{(I.95)}$$

where L([a,b]) denotes survival when the proof load ranges between a and b, and C(s) denotes failure under the load

$S = s$. The difficulty of updating the failure function G consists in finding explicit expressions for the fractional term in (I.95). A general treatment is quite complicated and perhaps unjustified for our purposes. The following developments rest on two assumptions which are met in most practical situations:

(i) (The safe interval is connected). If the system survives for $S = s_1$ and for $S = s_2 > s_1$, it also survives for all S in $[s_1, s_2]$; in particular:

$$L(S^m_{PL}) \cap L(S^M_{PL}) \text{ implies } L([S^m_{PL}, S^M_{PL}]). \qquad \text{(I.96a)}$$

Stated differently, this condition says that the (unknown) set of safe realizations of S is an interval on the S line.

(ii) (Scalar basic resistance parameter). There exists an (unknown) scalar resistance parameter $R^* \geq 0$ (for instance the resistance of the basic material) such that

(a) to each value of R^* corresponds a deterministic (possibly empty or infinite) safe interval $[S^m(R^*), S^M(R^*)]$ on the S line;

(b) $r^*_1 < r^*_2$ implies:

$$[S^m(r^*_1), S^M(r^*_1)] \subset [S^m(r^*_2), S^M(r^*_2)] \qquad \text{(I.96b)}$$

From conditions (a) and (b) (which correspond to the common notion that safety increases with the resistance parameters) one finds:

$$G_b[S^m(r^*)] = G_b[S^M(r^*)] = F_{R_b^*}(r^*) \ , \qquad \text{(I.96c)}$$

where $F_{R_b^*}(\cdot)$ denotes the prior CDF of R^*.

Under these assumptions the event $L([S_{PL}^m, S_{PL}^M])$ is equivalent to the event $L(S_{PL})$, where

$$S_{PL} = \begin{cases} S_{PL}^m \text{ if } G_b(S_{PL}^m) \geq G_b(S_{PL}^M) \ , \\ \\ S_{PL}^M \text{ otherwise} \ , \end{cases} \qquad \text{(I.97)}$$

and Bayes' equation (I.95) can be restated in the simpler form

$$G_a(s) = G_b(s) \frac{\Pr\{ L(S_{PL}) \mid C(s) \}}{\Pr\{ L(S_{PL}) \}} \ . \qquad \text{(I.98a)}$$

Since

$$\Pr \{ L(S_{PL}) \mid C(s) \} = \max \left\{ 0, \frac{G_b(s) - G_b(S_{PL})}{G_b(s)} \right\} \quad \text{(I.98b)}$$

and

$$\Pr \{ L(S_{PL}) \} = 1 - G_b(S_{PL}) \ , \quad \text{(I.98c)}$$

one can rewrite equation (I.98a) in the form:

$$G_a(s) = \max \left\{ 0, \frac{G_b(s) - G_b(S_{PL})}{1 - G_b(S_{PL})} \right\} . \quad \text{(I.99)}$$

The posterior probability of failure function $G_a(\cdot)$ is shown as a dashed line in the example of Figure 14. The prior, the posterior and the proof loading failure probabilities are given by the following expressions:

$$P_{f_b} = \int_{-\infty}^{\infty} f_s(s) \cdot G_b(s) \, d \, s \ ; \quad \text{(I.100a)}$$

$$P_{f_a} = \int_{-\infty}^{\infty} f_s(s) \cdot G_a(s)\, d\, s \; ; \qquad \text{(I.100b)}$$

$$P_{f_{PL}} = G_b(S_{PL}) \; , \qquad \text{(I.100c)}$$

S_{PL} being defined by equation (I.97).

The inequalities (I.94) and the remarks thereafter on the optimal choice of S_{PL} (or of $[S^m_{PL}, S^M_{PL}]$) still hold. When the function $G(\cdot)$ can be identified with the CDF of the system resistance equations (I.100) reproduce equations (I.93), while equation (I.99) becomes:

$$F_{R_a}(s) = \max \left\{ 0, \frac{F_{R_b}(s) - F_{R_b}(S_{PL})}{1 - F_{R_b}(S_{PL})} \right\},$$

a well known result in the classical approach.

One may also allow for errors in the estimate of the proof load intensity $\hat{S}_{PL}$ by writing $S_{PL} = \hat{S}_{PL} + \varepsilon$, where ε is a random estimation error term. Assuming that the PDF $f_\varepsilon(\cdot)$ is known and that it vanishes outside the interval [a,b], equations (I.98) become:

$$G_a(s) = G_b(s) \frac{Pr\{ L(\hat{S}_{PL} + \varepsilon) \mid C(s) \}}{Pr\{ L(\hat{S}_{PL} + \varepsilon) \}} \; , \qquad \text{(I.101)}$$

where

$$\Pr\{L(\hat{S}_{PL}+\varepsilon)\,|\,C(s)\} = \frac{1}{G_b(s)} \int_a^b \max\left\{0,[G_b(s)-G_b(\hat{S}_{PL}+e)]\right\} f_\varepsilon(e)\,d\,e \;;$$

$$\Pr\{L(\hat{S}_{PL}+\varepsilon)\} = \int_a^b [1-G_b(\hat{S}_{PL}+e)]\, f_\varepsilon(e)\,d\,e \;.$$

Equations (I.100a,b) still hold, while equation (I.100c) is replaced by:

$$P_{f_{PL}} = \int_a^b G_b(\hat{S}_{PL}+e)\, f_\varepsilon(e)\,d\,e \;.$$

The validity of inequalities (I.94) is retained.

Representation of the fundamental case in the space of the basic random variables

When the fundamental problem (1 random load, 1 unknown resistance) is considered in the generalized formulation, the results are more easily interpreted in the plane of the basic random variables R* and S; see Figure 15a.

In this plane the safe region D is defined:

$$D = \{(r^*,s)\,|\,s \in [S^m(r^*),S^M(r^*)]\} \;, \qquad \text{(I.102)}$$

with no bound in the direction of increasing resistance. The functions $S = S^m(R^*)$ and $S = S^M(R^*)$ correspond to limit state conditions; they depend on the particular system under consideration and on the definition of failure. Given the prior CDF, $F_{R_b^*}(\cdot)$, the function G_b defined by (I.96c) is the transform of $F_{R_b^*}$ under the laws $S = S^m(R^*)$ and $S = S^M(R^*)$; see Figure 15b. While this gives a geometrical interpretation to the previous analysis, when working in the SR* plane one can replace equation (I.100a) by

$$P_{f_b} = 1 - \int_D d\, F_{R_b^*}(r^*)\; d\, F_S(s) \; . \qquad \text{(I.103)}$$

In the two-dimensional formulation there is no need for introducing the function G, unless one wants to compute P_{f_b} in two stages:

(i) $G_b(s)$ = (failure probability given S=s) = $F_{R_b^*}[R^*(s)]$

(ii) equation (I.100a).

The effect of proof loading becomes clearer in the two-dimensional representation: if the system survives a proof loading range $[S^m_{PL}, S^M_{PL}]$, the minimum value that R* can have is:

$$R^*_{PL} = \max \{ R^*(S^m_{PL}), R^*(S^M_{PL}) \}$$

so that the shaded area in Figure 15a does not contribute to the posterior failure probability. The posterior CDF of R* is:

$$F_{R^*_a}(r^*) = \begin{cases} 0 & \text{for } r^* < R^*_{PL} \\ \dfrac{F_{R^*_b}(r^*) - F_{R^*_b}(R^*_{PL})}{1 - F_{R^*_b}(R^*_{PL})} & \text{otherwise} \end{cases}$$

from which the posterior failure probability can be computed:

$$P_{f_a} = 1 - \int_{D \cap (r^* \geq R^*_{PL})} d\, F_{R^*_a}(r^*)\, d\, F_S(s) \; . \qquad (I.104)$$

I.4.2. A Stochastic Model for Construction Defects

The reliability of a structural system is the result

of resistance and load uncertainties; "failure" results from unusual overloads, or from unusual construction defects, or from a combination of the two. The fact that failure is a rare event makes its probabilistic analysis difficult, and often questionable if the major causes of possible future misfunctionings are not enumerated and modeled properly. In particular it is important to realize that many types of construction defects may exist, ranging from a poor choice of the basic materials to gross manufacturing mistakes such as the use of an incorrect amount of reinforcement or the incorrect execution of a welded joint, and that these relatively rare, gross errors may be major causes of catastrophic failures.

So far little has been done to model resistance uncertainties other than those which arise from the usual variability of the basic materials. A more general stochastic model for construction errors is presented here. Defects are classified into two basic categories: the "standard" defects, which can be reproduced and studied in the laboratory under controlled conditions, and the "gross" defects which generate one or more subpopulations of "outliers" within the overall resistance population, and are caused primarily by human errors. Standard and gross defects do not necessarily differ in magnitude, but they originate from distinct mechanisms and belong to different

statistical populations.

As an example consider the process of concrete production. The basic materials shipped to the factory are controlled in such a way that the relevant properties, collected in the vector $\underline{X}^{(M)} = [X_1^{(M)},\ldots,X_{n_M}^{(M)}]'$, have the statistical distribution shown in Figure 16a for the case $n_M = 1$. The properties of the concrete produced by the factory, vector $\underline{X}^{(C)} = [X_1^{(C)},\ldots,X_{n_C}^{(C)}]'$, depend in part on $\underline{X}^{(M)}$, in part on the manufacturing procedure. If the manufacturing operations are executed correctly, the density of $(\underline{X}^{(C)} \mid \underline{X}^{(M)} = \underline{x}^{(M)})$ is shown hypothetically in Figure 16b, curve (a) (for $n_C = 1$); the dispersion is due exclusively to "standard" variations. If "gross" errors are made, the same conditional density is exemplified by curve (b).

Many random variables of interest to the safety of engineering systems (and not necessarily related to resistances) conform to composite models of this kind: for example, wind velocities in tornadoes do not follow the same probability law as wind velocities in convective storms; tsunami sea waves are not statistically identical with wind-induced waves. Similarly, the absence of reinforcing bars in a concrete beam is a quite distinct event from their misplacement; also, over- or under-production of a factory, the change of machinery or personnel, the shortage of raw

materials, may cause quality variations in the product, which are statistically different from those of normal production.

Returning to the concrete production problem, the unconditional (on the type of production errors) probability distribution of the generic concrete property $X_i^{(C)}$ is:

$$f(X_i^{(C)}|\underline{X}^{(M)}) = (1 - P_G) \cdot f(X_i^{(C)}|\underline{X}^{(M)}; \text{ only "minor" errors})$$

$$+ P_G \cdot f(X_i^{(C)}|\underline{X}^{(M)}; \text{ "minor" + "gross" errors}) ,$$

where P_G is the probability that gross errors were made in the manufacturing process. This unconditional density is shown in Figure 16c. If two or more types of gross errors need to be discriminated, a distinct conditional distribution like (b) in Figure 16b should be introduced for each type of error, and weighted by the associated error probability in the expression for the unconditional density.

The same model can be applied to the statistical variations of the properties of structural concrete elements, as well as to the resistance parameters of the entire system. For the generic element let $\underline{X}^{(E)} = [X_1^{(E)},\ldots,X_{n_E}^{(E)}]'$ summarize the relevant properties. Due to the possibility of gross manufacturing errors (poor concrete vibration, partial

absence of reinforcement, etc.) the density function of $\underline{X}^{(E)}$ for given concrete properties may look like the curve in Figure 16d. (Again, the shape may be modified to account for several different types of gross errors.) Finally, the system properties $\underline{X}^{(S)} = [X_1^{(S)}, \ldots X_{n_S}^{(S)}]'$ may be related to the elements properties as shown in Figure 16e.

The unconditional densities of the properties of concrete, of the structural elements, and of the system can be found through successive integration:

$$f(\underline{x}^{(C)}) = \int_{\text{all } \underline{x}^{(M)}} f(\underline{x}^{(C)} | \underline{X}^{(M)} = \underline{x}^{(M)}) \cdot f(\underline{x}^{(M)}) \cdot d\,\underline{x}^{(M)} \; ;$$

$$f(\underline{x}^{(E)}) = \int_{\text{all } \underline{x}^{(C)}} f(\underline{x}^{(E)} | \underline{X}^{(C)} = \underline{x}^{(C)}) \cdot f(\underline{x}^{(C)}) \cdot d\,\underline{x}^{(C)} \; ;$$

$$f(\underline{x}^{(S)}) = \int_{\text{all } \underline{x}^{(E)}} f(\underline{x}^{(S)} | \underline{X}^{(E)} = \underline{x}^{(E)}) \cdot f(\underline{x}^{(E)}) \cdot d\,\underline{x}^{(E)} \; .$$

Hypothetical configurations of these densities (assumed univariate) are sketched in Figure 17.

The failure probability may be contributed to substantially by gross errors; if this is the case, one purpose of experimenting is to detect the presence of gross errors

and possibly also to discriminate among different types of gross errors, if they exist. For instance, if the laboratory test of a concrete cylinder yields the axial resistance $X^{(C)} = \bar{x}^{(C)}$ shown in Figure 17b, the probability that the batch of concrete is grossly defective is quite small. Via Bayes' equation the posterior density of the concrete resistance might look like the curve in Figure 17b'. This would also have the effect of reshaping the densities of $\underline{X}^{(E)}$ and $\underline{X}^{(S)}$ in those parts which are contributed to most by gross errors in concrete manufacturing. However, for the purpose of discriminating gross errors at the structural level, material tests do not help much unless gross errors at early production stages occur at a much higher rate than gross errors during elements manufacturing and system assemblage. On the other hand, there are economical penalties for proof loading the structural elements or the entire system, so that the best data collection policy poses a problem of discrete optimal allocation of limited resources.

Processing of Experimental Data

In what follows we consider the classical one-sided formulation of the fundamental reliability problem (see Paragraph I.4.1), and we show how prior resistance distributions of the composite type suggested above are modified

through Bayesian processing of experimental data. (The parallel development for the generalized formulation only requires a more elaborate formalism.)

(a) Generalities

Let R denote any structural property of interest, and assume that the prior CDF of R can be written in the form:

$$F_b(R) = \sum_{i=1}^{n_e} P_{b_i} \int_{\underline{\Theta}\ \text{space}} F_b(R|\underline{\Theta})\ d\ F_b(\underline{\Theta}|e_i)\ , \qquad \text{(I.105)}$$

where $\underline{\Theta}$ is a vector of distribution parameters and $\{e_i\}$ is the set of possible construction errors affecting R. $e_1,\ldots,e_{n_e}$ are exhaustive and mutually exclusive events. According to equation (I.105) the distribution type of R is assumed known, while the parameters are random functions of the errors. The posterior CDF of R, given the information Z, is:

$$F_a(R) = \sum_{i=1}^{n_e} P_{a_i} \int_{\underline{\Theta}\ \text{space}} F_a(R|\underline{\Theta})\ d\ F_a(\underline{\Theta}|e_i)\ , \qquad \text{(I.106)}$$

where

$$d\,F_a(\underline{\Theta}|e_i) \propto d\,F_b(\underline{\Theta}|e_i) \cdot l(Z|\underline{\Theta}), \text{ for fixed } e_i; \qquad \text{(I.107a)}$$

$$P_{a_i} \propto P_{b_i} \int_{\underline{\Theta} \text{ space}} l(Z|\underline{\Theta})\, d\,F_a(\underline{\Theta}|e_i); \qquad \text{(I.107b)}$$

$$F_a(R|\underline{\Theta}) \propto F_b(R|\underline{\Theta})\; l(Z|R), \text{ for fixed } \underline{\Theta}; \qquad \text{(I.107c)}$$

$$l(Z|\underline{\Theta}) \propto \int_{\text{all } R} l(Z|R) \cdot d\,F_a(R|\underline{\Theta}). \qquad \text{(I.107d)}$$

and $l(\cdot)$ denotes likelihood function.

The integrals in equations (I.105) and (I.106) are the conditional prior and posterior distributions of R, $F_b(R|e_i)$ and $F_a(R|e_i)$. All the three posterior distributions:

P_{a_i} $(i = 1,\ldots,n_e)$, equation (I.107b);

$F_a(\underline{\Theta}|e_i)$ $(i = 1,\ldots,n_e)$, equation (I.107a);

$F_a(R)$, equation (I.106)

are of interest: the first one for discriminating among error types, the second one for making inferences about the distribution of the resistance population (resistance of

nontested structures under the present assumption that the same unknown types of errors have been made in manufacturing all the structures), the third one for making inferences on the resistance of the structure being tested directly. In fact, the posterior distribution of R for a system belonging to the same statistical population, but not being subjected to direct experimentation is:

$$F_a(R) = \sum_{i=1}^{n_e} P_{a_i} \int_{\underline{\Theta} \text{ space}} F_b(R|\underline{\Theta}) \, d \, F_a(\underline{\Theta}|e_i). \qquad \text{(I.108)}$$

The non-identity between the CDF (I.108) and the CDF (I.106) agrees with the intuitive notion that the experiment is more informative on the resistance of the system being tested directly, than it is on any other system from the same population.

The foregoing analysis simplifies if $F_b(\underline{\Theta}|e_i)$ is conjugate with respect to the likelihood function $l(Z|\underline{\Theta})$, and $F_b(R|\underline{\Theta})$ is conjugate with respect to $l(Z|R)$.

Another case which leads to considerable simplifications is when $F_b(\underline{\Theta}|e_i)$ is a step function defining the deterministic relationship: $\underline{\Theta}_i = \underline{\Theta}(e_i)$; then the integrals over $\underline{\Theta}$ are replaced by discrete summations and equations (I.105) and (I.106) become:

$$F_b(R) = \sum_{i=1}^{n_e} P_{b_i} \cdot F_b(R|\underline{\Theta}_i) \; ; \qquad \text{(I.109)}$$

$$F_a(R) = \sum_{i=1}^{n_e} P_{a_i} \cdot F_a(R|\underline{\Theta}_i) \; , \qquad \text{(I.110)}$$

where

$$P_{a_i} \propto P_{b_i} \cdot l(Z|\underline{\Theta}_i) \; ; \qquad \text{(I.111a)}$$

$$F_a(R|\underline{\Theta}_i) \propto F_b(R|\underline{\Theta}_i) \cdot l(Z|R) \; ; \qquad \text{(I.111b)}$$

$$l(Z|\underline{\Theta}_i) \propto \int_{\text{all } R} l(Z|R) \cdot d\, F_a(R|\underline{\Theta}_i) \; . \qquad \text{(I.111c)}$$

(b) Data from proof loading experiments

The outcome Z of a (non-noisy) proof loading experiment which measures R directly may have either of two forms:

$$Z: \begin{cases} R=\overline{R}, \text{ with } \overline{R}<R_{PL}, \text{ in case of failure at load } \overline{R}, \text{ or} & \text{(I.112a)} \\ R \geq R_{PL}, \text{ in case of survival,} & \text{(I.112b)} \end{cases}$$

where R_{PL} is the minimum resistance for survival of the test.

In the first case the likelihood functions of R and $\underline{\Theta}$ are:

$$l(Z|R) \propto \delta(R - \bar{R}) \tag{I.113a}$$

where $\delta(\cdot)$ is the Dirac delta function, and, if $F_a(R|\underline{\Theta})$ is differentiable at $\bar{R}$:

$$l(Z|\underline{\Theta}) \propto \left. \frac{d\, F_a(R|\underline{\Theta})}{d\, R} \right|_{R = \bar{R}} . \tag{I.113b}$$

In the second case:

$$l(Z|R) \propto \begin{cases} 0 & \text{if } R < R_{PL}, \\ 1 & \text{if } R \geq R_{PL}, \end{cases} \tag{I.114a}$$

and

$$l(Z|\underline{\Theta}) \propto 1 - F_a(R_{PL}|\underline{\Theta}) \tag{I.114b}$$

if $F_a(R|\underline{\theta})$ is continuous at R_{PL}.

These equations are used next in connection with the composite resistance model developed earlier for reinforced concrete structures. The proof loading of a bridge deck serves as a numerical example (Example 6).

(c) Application to proof loading of reinforced concrete elements connected in parallel or in series

Proof loading experiments can be made at different stages of construction in order to detect different types of gross errors and to improve the reliability of the survived items. So far we have considered one-mode-of-failure systems, proof loaded after completion of construction. However, the same analysis applies to experimental data from earlier construction stages, or from structural subsystems.

We examine now the reliability consequences of proof loading the elements of a series or of a parallel system. As before, four construction phases are distinguished: basic materials selection, concrete manufacturing, elements manufacturing, structure assemblage. When reference to a particular phase is needed, the following upperscripts are used: (M) for materials, (C) for concrete, (E) for elements, (S) for system. $R^{(S)}$ and $R_i^{(E)}$ denote the

resistance of the system and of the i^{th} element, respectively.

At each stage of construction "minor" errors are introduced, while "gross" errors may or may not be introduced. The following notation is used: G = presence of gross (and minor) errors; M = presence of minor errors only. When several such symbols appear in a sequence, they refer to sequential construction phases, starting from the selection of the basic materials; so $f(R_i^{(E)}|M,G,M)$ denotes the probability density function of the resistance of element i, given that minor variations are present in the material properties and in the construction of the elements while gross errors were introduced during the concrete production. When necessary, an upperscript letter will be used to identify the stage.

Prior distribution of the systems resistance. Prior probabilities are assigned to each possible combination of error types: (M), (G), (M,M), (G,M), (M,G), (G,G),...., (G,G,G,G). Conditional on the i^{th} error combination, e_i, the elements resistances are independent, with common prior probability distribution $F_b(R^{(E)}|e_i)$. The unconditional prior distribution of the resistance $R_j^{(E)}$ is therefore (assuming, as before, that the same but unknown error combination affects all the resistances):

$$F_b(R_j^{(E)}) = P_b(M,M,M) \cdot F_b(R^{(E)}|M,M,M)$$
$$+ P_b(M,M,G) \cdot F_b(R^{(E)}|M,M,G)$$
$$+ \ldots + P_b(G,G,G) \cdot F_b(R^{(E)}|G,G,G)$$

Assume for simplicity that $F_b(R^{(E)}|e_i) = F(R^{(E)}|e_i) = N(\mu_i,\sigma_i^2)$ and that the errors in assembling the system are independent of previous errors. For a parallel system the following model is used. If only minor assembling errors are made:

$$(R_p^{(S)}|M^{(S)}) = C_{p_M} \cdot \sum_{i=1}^{n} R_i^{(E)} + \varepsilon_{p_M} ; \qquad \text{(I.115a)}$$

where $\varepsilon_{p_M} \sim N(\mu_{p_M}, \sigma_{p_M}^2)$,

and the subscript p denotes "parallel"; if instead also gross errors are introduced in the final construction stage:

$$(R_p^{(S)}|G^{(S)}) = C_{p_G} \cdot \sum_{i=1}^{n} R_i^{(E)} + \varepsilon_{p_G} , \qquad \text{(I.115b)}$$

where $\varepsilon_{p_G} \sim N(\mu_{p_G}, \sigma_{p_G}^2)$.

Similarly, for a series system:

$$(R_S^{(S)} | M^{(S)}) = C_{S_M} \cdot \min_i \{ R_i^{(E)} \} + \varepsilon_{S_M} , \qquad \text{(I.116a)}$$

$$(R_S^{(S)} | G^{(S)}) = C_{S_G} \cdot \min_i \{ R_i^{(E)} \} + \varepsilon_{S_G} , \qquad \text{(I.116b)}$$

where

$$\varepsilon_{S_M} \sim N(\mu_{S_M}, \sigma^2_{S_M}) \; ; \; \varepsilon_{S_G} \sim N(\mu_{S_G}, \sigma^2_{S_G}) \; .$$

For a parallel system the unconditional prior distribution of $R_p^{(S)}$ is:

$$F_b(R_p^{(S)}) = P_b(M^{(S)}) \sum_{i=1}^{8} P_{b_i} \Phi(X_{i_M}) + P_b(G^{(S)}) \sum_{i=1}^{8} P_{b_i} \Phi(X_{i_G}) , \qquad \text{(I.117a)}$$

where $e_1 = (M,M,M)$, $e_2 = (M,M,G), \ldots, e_8 = (G,G,G)$; P_{b_i} is the prior probability of the error combination e_i; $P_b(A)$ is the prior probability of the event A;

$$X_{i_M} = \frac{R_p^{(S)} - n\,\mu_i - \mu_{p_M}}{(n \cdot C^2_{p_M} \cdot \sigma_i^2 + \sigma^2_{p_M})^{1/2}} \; ;$$

$$X_{i_G} = \frac{R_p^{(S)} - n\,\mu_i - \mu_{p_G}}{(n \cdot C_{p_G}^2 \cdot \sigma_i^2 + \sigma_{p_G}^2)^{1/2}} \quad ;$$

n is the number of elements connected in parallel, and $\Phi(\cdot)$ is the standard normal CDF.

For a series system the same probability distribution is:

$$F_b(R_S^{(S)}) = P_b(M^{(S)}) \int_0^{\infty} \left\{ \sum_{i=1}^{8} P_{b_i} [1-(1-\Phi(\frac{r-\mu_i}{\sigma_i}))^n] \right\}$$

$$\cdot \phi(\frac{R_S^{(S)} - C_{S_M} \cdot r - \mu_{S_M}}{\sigma_{S_M}}) d\,r$$

$$+ P_b(G^{(S)}) \int_0^{\infty} \left\{ \sum_{i=1}^{8} P_{b_i} [1-(1-\Phi(\frac{r-\mu_i}{\sigma_i}))^n] \right\}$$

$$\cdot \phi(\frac{R_S^{(S)} - C_{S_G} \cdot r - \mu_{S_G}}{\sigma_{S_G}}) d\,r \,, \qquad \text{(I.117b)}$$

where $\phi(\cdot)$ is the standard normal PDF. Equations (I.117) greatly simplify if no errors are introduced in the assemblage phase and $C_p = C_S = 1$, in which case:

$$F_b(R_p^{(S)}) = \sum_{i=1}^{8} P_{b_i} \Phi\left(\frac{R_p^{(S)} - n\,\mu_i}{n^{1/2} \cdot \sigma_i} \right) ; \qquad \text{(I.118a)}$$

$$F_b(R_S^{(S)}) = \sum_{i=1}^{8} P_{b_i} \left\{ 1-[1-\Phi\left(\frac{R_S^{(S)} - \mu_i}{\sigma_i} \right)]^n \right\}. \qquad \text{(I.118b)}$$

Posterior distribution of the systems resistance. We examine now how the last two equations are modified by proof loading one structural element. If failure occurs at a load intensity $\overline{R} < R_{PL}$, the element is replaced by a new one, and the experiment is not repeated; if the element survives, two strategies are considered: either the element is used, or it is replaced.

(a) The element fails at a load $\overline{R} < R_{PL}$

The probabilities P_{b_i} are updated as indicated by equation (I.107b):

$$P_{a_i} = \frac{P_{b_i} \cdot \phi[(\overline{R}-\mu_i)/\sigma_i]/\sigma_i}{\sum_{j=1}^{8} P_{b_j} \cdot \phi[(\overline{R}-\mu_j)/\sigma_j]/\sigma_j} . \qquad \text{(I.119)}$$

The posterior CDF of the system resistance is:

$$F_a(R_p^{(S)}) = \sum_{i=1}^{8} P_{a_i} \Phi\left(\frac{R_p^{(S)} - n\,\mu_i}{n^{1/2} \cdot \sigma_i} \right) \qquad \text{(I.120a)}$$

for a parallel configuration, and

$$F_a(R_S^{(S)}) = \sum_{i=1}^{8} P_{a_i} \left\{ 1-[1-\Phi\left(\frac{R_S^{(S)} - \mu_i}{n^{1/2} \cdot \sigma_i} \right)]^n \right\} \qquad \text{(I.120b)}$$

for a series configuration. (These posterior distributions are generalized in the obvious way for the prior CDF's (I.117).) Equations (I.120) differ from the prior expressions (I.118) only because of the revised error probabilities. Recall that all elements are assumed to suffer from the same (but unknown) set of error types. Therefore failure of one element carries information about the resistance of all the others.

(b) <u>The element survives</u>

The posterior error combination probabilities are:

$$p_{a_i} = \frac{p_{b_i} \{ 1 - \Phi[(R_{PL}-\mu_i)/\sigma_i] \}}{\sum_{i=1}^{8} p_{b_j} \{ 1 - \Phi[(R_{PL}-\mu_j)/\sigma_j] \}} . \qquad \text{(I.121)}$$

(b1) The element is replaced. Then the posterior CDF of the system resistance is formally identical with (I.120a) or (I.120b), with p_{a_i} given by equation (I.121).

(b2) The element is not replaced. The CDF of the resistance of the element being proof loaded, denoted $F_{a_{PL}}(\cdot)$, is:

$$F_{a_{PL}}(R^{(E)}) = \begin{cases} 0 & \text{if } R^{(E)} < R_{PL}, \\ \dfrac{F_a(R^{(E)}) - F_a(R_{PL})}{1 - F_a(R_{PL})} & \text{otherwise}, \end{cases}$$

where $F_a(\cdot)$ is the posterior CDF for the elements not being loaded:

$$F_a(r) = \sum_{i=1}^{8} p_{a_i} \cdot \Phi[(r-\mu_i)/\sigma_i] ,$$

and p_{a_i} is given by equation (I.121). Therefore the CDF of the system resistance results: for elements connected in parallel:

$$F_a(R_p^{(S)}) = \begin{cases} \sum_{i=1}^{8} p_{a_i} \int_0^{R_p^{(S)}-R_{PL}} \phi\left[\frac{r-(n-1)\mu_i}{(n-1)^{1/2}\sigma_i}\right] \cdot F_{a_{PL}}(R_p^{(S)}-r)\,dr\,, & \text{for } R_p^{(S)} \geq R_{PL}\,, \\ 0\,, & \text{otherwise;} \end{cases} \tag{I.122a}$$

and for a series configuration:

$$F_a(R_S^{(S)}) = \sum_{i=1}^{8} p_{a_i} \left\{ 1-\left[(1-\Phi\left[\frac{R_S^{(S)}-\mu_i}{\sigma_i}\right])^{n-1} \cdot (1-F_{a_{PL}}(R_S^{(S)}))\right] \right\}. \tag{I.122b}$$

Equations (I.119) through (I.122) can be easily generalized to the case when more than one element is proof loaded.

The example which follows is an application of the foregoing results.

EXAMPLE 6. PROOF LOADING OF A MULTI-SPAN BRIDGE DECK

Consider a simply-supported n-span bridge structure (series system with n elements). Before proof loading, the possible error combinations and the parameters of the conditional distribution of the "element" resistances (in terms

of maximum load carrying capacity) are shown in Table 3. The probability distribution of the generic element resistance $R^{(E)}$ conditional on a given error combination is normal; the resistances of different spans are independent. The design load, S_D, is fixed to be either 230 or 200 kips (second-moment reliability index $(\mu_1-S_D)/\sigma_2$=3.5 and, respectively, 5.0 for the error combination M,M,M). Any number $m \leq n$ of spans can be proof loaded; the proof load intensity is denoted S_{PL}. Given the distribution of the load, the posterior failure probability can be computed through straightforward generalization of equations (I.119) through (I.122). For n = 1,2,5,10,20; S_D = 230,200 kips; m = 0(1)n, and for several proof load values, the posterior failure probability corresponding to the design load S_D:

$$P_{f_a} = \Pr\left[\bigcup_{i=1}^{n} (R_i^{(E)} < S_D)\right]$$

is plotted in Figures 18 - 21. The values indicated in these figures are conditional on proof load survival and on a non-replacement policy. They were computed through equation (I.121) and a generalization of equation (I.122b) which allows for $m \neq 1$.

The failure probability after proof loading the whole bridge with $S_{PL} = S_D$ is zero, and the probability that

i	e_i	p_{b_i}	μ_i (kips)	σ_i (kips)
1	M,M,M	0.950800	300	20
2	M,M,G	0.019404	260	20
3	M,G,M	0.019404	250	30
4	G.M.M	0.009600	250	20
5	M,G,G	0.000396	230	30
6	G,M,G	0.000196	230	30
7	G,G,M	0.000196	220	30
8	G,G,G	0.000004	170	40

Table 3. Possible error combinations, $\{e_i\}$; prior errors probabilities $\{p_{b_i}\}$; and resistance parameters $\{\mu_i; \sigma_i\}$ for each set of error types.

at least one span fails during proof loading is $P_{f_{PL}} = P_{f_a} - P_{f_b}$, where P_{f_b} coincides with P_{f_a} for m = 0. The results show that for reliable systems (second-moment reliability index = 5.0) proof loading is not effective unless S_{PL} is close to S_D. It is also noted that for long bridges (Figure 21) survival of the first few spans reduces the posterior failure probability by decreasing the probability of gross errors, while survival of the last few spans reduces P_{f_a} by truncation of $f_a(R_S^{(S)})$ at S_{PL}.

For n = 10, S_D = 230 kips and S_{PL} = 180 or 220 kips the evolution of P_{f_a} as the number of proof loaded spans in ranges from 0 to 10 is shown in Tables 4 and 5. Also given are the posterior error probabilities $P_a(M,M,M)$; $P_a(M,M,G)$; $P_a(M,G,G)$ and $P_a(G,G,G)$, chosen as a representative subset of the eight possible error combinations. The evolution of $P_a(M,G,M)$ and of $P_a(G,M,M)$ was found to be similar to that of $P_a(M,M,G)$; so too was the evolution of $P_a(G,M,G)$ and of $P_a(G,G,M)$ similar to that of $P_a(M,G,G)$.

Although the posterior failure probability does not vary considerably with m when S_{PL} = 180 kips, the probability that gross errors are simultaneously present, $P_a(G,G,G)$, decreases significantly with m, meaning that the proof loading experiment was effective in discriminating between presence and absence of such errors. When S_{PL} = 220 kips this fact is even more accentuated.

S_{PL} = 180 kips; S_D = 230 kips; n = 10

m	p_{f_a}	p_a(M,M,M)	p_a(M,M,G)	p_a(M,G,G)	p_a(G,G,G)
0	0.3892 -01	0.9508	0.1940 -01	0.3960 -03	0.4000 -05
1	0.3869 -01	0.9510	0.1941 -01	0.3772 -03	0.1606 -05
2	0.3846 -01	0.9513	0.1941 -01	0.3593 -03	0.6445 -06
3	0.3824 -01	0.9515	0.1942 -01	0.3422 -03	0.2587 -06
4	0.3802 -01	0.9517	0.1942 -01	0.3259 -03	0.1038 -06
5	0.3781 -01	0.9519	0.1942 -01	0.3104 -03	0.4168 -07
6	0.3760 -01	0.9521	0.1943 -01	0.2957 -03	0.1673 -07
7	0.3739 -01	0.9523	0.1943 -01	0.2816 -03	0.6715 -08
8	0.3719 -01	0.9525	0.1943 -01	0.2682 -03	0.2695 -08
9	0.3699 -01	0.9527	0.1944 -01	0.2554 -03	0.1082 -08
10	0.3679 -01	0.9529	0.1944 -01	0.2433 -03	0.4342 -09

Table 4. Example 6: m = number of proof-loaded spans; p_{f_a} = posterior failure probability; p_a(e) = posterior probability of the errors combination e. S_{PL} = proof load intensity; S_D = design load; n = number of spans.

S_{PL} = 220 kips; S_D = 230 kips; n = 10

m	p_{f_a}	p_a(M,M,M)	p_a(M,M,G)	p_a(M,G,G)	p_a(G,G,G)
0	0.3892 -01	0.9508	0.1940 -01	0.3960 -03	0.4000 -05
1	0.3456 -01	0.9551	0.1905 -01	0.2508 -03	0.4243 -06
2	0.3083 -01	0.9587	0.1869 -01	0.1588 -03	0.4498 -07
3	0.2759 -01	0.9619	0.1832 -01	0.1004 -03	0.4766 -08
4	0.2476 -01	0.9647	0.1796 -01	0.6352 -04	0.5047 -09
5	0.2228 -01	0.9671	0.1759 -01	0.4016 -04	0.5344 -10
6	0.2009 -01	0.9692	0.1723 -01	0.2538 -04	0.5656 -11
7	0.1814 -01	0.9711	0.1687 -01	0.1603 -04	0.5984 -12
8	0.1640 -01	0.9728	0.1652 -01	0.1013 -04	0.6331 -13
9	0.1485 -01	0.9743	0.1617 -01	0.6396 -05	0.6696 -14
10	0.1344 -01	0.9757	0.1582 -01	0.4039 -05	0.7081 -15

Table 5. Example 6: m = number of proof-loaded spans; p_{f_a} = posterior failure probability; p_a(e) = posterior probability of the errors combination e. S_{PL} = proof load intensity; S_D = design load; n = number of spans.

Minor modifications would allow to replace the deterministic design load with a random service load.

I.4.3 Proof Loading in Multidimensional Random Spaces

The logical steps in the analysis of proof loading data are not affected by more than one load or resistance parameter being random or unknown. However, operational complications may become prohibitive.

Let $\underline{S} = [S_1,\ldots,S_{n_S}]'$ be the vector of basic random load parameters and $\underline{R} = [R_1,\ldots,R_{n_R}]'$ the vector of basic unknown resistances. By generalization of the function G(S) in Paragraph I.4.1, $G(\underline{S})$ denotes the probability of failure conditional on a given load vector; then the prior probability of failure is:

$$P_{f_b} = \int_{\underline{S}\ \text{space}} G_b(\underline{s})\ d\ F_{\underline{S}}(\underline{s}) \qquad \text{(I.123)}$$

where $F_{\underline{S}}(\cdot)$ is the CDF of the service load vector. If the proof loading experiment consists in having $\underline{S}$ attaining all the values inside the region $D_S \subset R^{n_S}$, the posterior probability of failure conditional on proof loading survival is:

$$P_{f_a} = \int_{\underline{s} \notin D_S} G_a(\underline{s})\, d\, F_{\underline{S}}(\underline{s}) \qquad \text{(I.124)}$$

where, from Bayes' theorem:

$$G_a(\underline{s}) = G_b(\underline{s}) \frac{Pr[L(D_S) \mid C(\underline{s})]}{Pr[L(D_S)]} \qquad \text{(I.125)}$$

In equation (I.125), $L(D_S)$ denotes survival when $\underline{S}$ attains all possible values in D_S, and $C(\underline{s})$ denotes collapse under the load $\underline{s}$.

The probability ratio in equation (I.125) can be computed conveniently after formulating the proof-loading problem in the space of the basic variables which, for simplicity of notations, we assume to coincide with $\underline{R}$ and $\underline{S}$. The following parallels the generalized analysis and the two-dimensional representation of the fundamental problem in Paragraph I.4.1.

Let $D_R(\underline{s})$ be the set of safe realizations of $\underline{R}$ if $\underline{S} = \underline{s}$. Then, conditional on survival under proof loading, $\underline{R}$ belongs with probability 1 to the region

$$D_R(D_S) = \{\underline{r} \mid \underline{r} \in D_R(\underline{s}) \text{ for all } \underline{s} \in D_S\}$$

and the last two probabilities in equation (I.125) can be written as:

$$\Pr[L(D_S)] = 1 - P_{f_{PL}} = \int_{D_R(D_S)} d\, F_b(\underline{r}) \; ;$$

$$\Pr[L(D_S) \,|\, C(\underline{s})] = \int_{D_R(D_S) \cap D_R^C(\underline{s})} d\, F_b(\underline{r}) ;$$

where the upperscript C denotes complementation.

One can also compute the posterior PDF of $\underline{R}$, being:

$$f_a(\underline{r}) = \begin{cases} 0 & , \text{ if } \underline{r} \notin D_R(D_S), \\ & \\ f_b(\underline{r}) \cdot \{\Pr[L(D_S)]\}^{-1} & , \text{ otherwise .} \end{cases} \qquad \text{(I.126)}$$

It is interesting to note that in general the posterior density function (I.126) cannot be written as the product of n_R marginal densities, even if a priori the components of $\underline{R}$ were independent. Stated differently, an informative experiment ($P_{f_{PL}} \neq 0$) introduces dependence between the resistance parameters. For example, consider

the case of two unknown resistance parameters R_1 and R_2, and of two random load parameters S_1 and S_2. During proof loading the loads are varied along the perimeter of (and possibly inside) the quadrilateral ABCD in Figure 22. The boundaries of the safe regions $D_R(A)$, $D_R(B)$, $D_R(C)$, $D_R(D)$ corresponding to the corner load points are shown in Figure 23. The intersection of these regions defines $D_R(D_S)$ if survival under $\underline{S} = \underline{s}_1$ and under $\underline{S} = \underline{s}_2$ implies survival under all convex combinations $\underline{S} = a\,\underline{s}_1 + (1-a)\underline{s}_2$; $0 \leq a \leq 1$. In the same figure the solid lines indicate the prior marginal densities of R_1 and R_2, and the dashed lines indicate qualitatively the corresponding posterior densities, conditional on survival. The posterior densities of $R_1|R_2$ and $R_2|R_1$ for a given prior joint density can also be found easily (Veneziano, 1972).

I.5. DESIGN OF PROOF LOADING EXPERIMENTS

Using the Bayesian analysis developed in Section I.4 it is straightforward to give a formal definition of the optimal proof loading experiment. As in any problem of decision making under risk two basic elements define the optimal experiment:

(i) The set of the feasible actions (here the set of feasible proof load experiments), A, with

generic action a_{PL}.

(ii) An expected utility function $U(a_{PL})$, which defines a preference order among the feasible actions. Typical arguments of $U(\cdot)$ are the prior, the posterior and the proof loading failure probabilities P_{f_b}, P_{f_a}, $P_{f_{PL}}$; the cost of the experiment $C(a_{PL})$; and the cost C_o of making no experiment. The expected utility function depends also on the level (e.g., materials, elements, subsystem, system) at which the experiment is made, and on the policy of replacing or not the parts of the system which survived the experiment.

Then the optimal experiment a^*_{PL} is defined to be the action in A which maximizes the expected utility.

If several experiments are made sequentially (for instance one on each element of a system) the proof load may vary from experiment to experiment, possibly accounting for previous results. In this case one has to optimize a <u>proof loading strategy</u>, say by defining the best rule of the type:

$$a^*_{PL_n} = \hat{a}_n(a^*_{PL_1}, a^*_{PL_2}, \ldots, a^*_{PL_{n-1}}, I_1, I_2, \ldots, I_{n-1}) \ ,$$

where I_j is an indicator of survival ($I_j = 0$) and failure ($I_j = 1$) for the j^{th} experiment. In this enlarged framework one might include the number of experiments in the set of design variables.

The problem of optimizing a single proof load experiment was discussed by Sexsmith (1969) and by Shinozuka (1969) with respect to the fundamental case in its classical formulation (see Paragraph I.4.1). The writer is not aware of studies of optimal proof loading experiments in a more general context, with the exception of Bouton, et al. (1967) and Campion, et al. (1972) who examine the problem at a rather qualitative level.

Conceptually, the optimization of a single proof loading experiment has the same degree of difficulty as finding the optimal design of a structural system for maximum expected utility. In the proof loading problem the definition of the optimal experiment a^*_{PL} requires identifying the optimal set $D_R(D_S)$ in which the posterior resistance density is non-zero (see Paragraph I.4.3); in the second problem the optimal design of a structural system consists in finding the optimal safe region D in $\underline{R}\ \underline{S}$ space (Gavarini and Veneziano, 1972). In both cases the objective is an implicit function of the design variables. Unfortunately these features make the problem practically unsolvable in its generality. Manageable formulations can be arrived at in

particular situations; a simple, yet nontrivial case is studied in Example 7.

EXAMPLE 7. OPTIMAL PROOF LOADING OF A SIMPLY-SUPPORTED, ONE-SPAN BRIDGE DECK

In some countries (e.g., in Italy), before opening a bridge to the traffic, the structure has to be inspected and proof loaded under load intensities which are close to the design values. In order to achieve high stresses in some parts of the deck, it is common practice to load the bridge with the maximum possible eccentricity. The "best" combination of proof load intensity and eccentricity is studied here with reference to a simply-supported, single-span bridge.

(a) Mechanics

The relevant structural elements of the deck are shown in Figure 24: the reinforced concrete slab is supported by four (statistically) identical beams, which in turn are connected by transverse beams to ensure a proper distribution of the loads. The response of the bridge to concentrated loads depends on the torsional rigidity of the deck, on the flexural rigidity of the beams and on the stiffness of the transversal connectors.

Simplified linear elastic models are often used in practice: at one extreme, the infinite-torsional-stiffness model (equiripartition of the load among the longitudinal beams); at the other extreme the independent-beams model in which each longitudinal beam carries completely and exclusively the loads applied in its "influence region". The influence regions cover the total surface of the deck without overlapping. Other more realistic models assume uneven load distributions among the beams. A simple and useful idealization views the transverse beams as infinitely rigid connections. Instead, the deck has finite torsional stiffness due to the finite bending stiffness of the longitudinal beams (the torsional stiffnesses of the beams and of the slab are neglected).

(b) Safe regions

Consider a concentrated load P applied at midspan, with eccentricity X from the deck centerline. If the load carrying capacities of the individual beams are the same, say $P_{M_1} = P_{M_2} = P_{M_3} = P_{M_4} = P_M$, the safe regions in the plane $|X|P$ corresponding to the three models introduced above are shown in Figure 25. For each model, survival and failure correspond respectively to points inside and outside its safe region.

For P being the resultant of a load at midspan which is uniformly distributed in the transversal interval

$[(2X - \ell/2), \ell/2]$ if $X \geq 0$, and in the transversal interval $[-\ell/2, (2X + \ell/2)]$ if $X < 0$, the safe regions are shown in Figure 26. (Only the safe region corresponding to the independent-beams model is modified with respect to Figure 25.)

Focusing now on the infinitely-rigid-transversal-beams model, the load carried by the most stressed beam is, for both concentrated and distributed loads:

$$S = P/4 + 0.3\ P\ \frac{|X|}{d}$$

where d is the distance between adjacent beams. Neglecting failure for negative moments the safe region is defined by the inequality:

$$P - \frac{P_M}{0.25 + 0.3\ \frac{|X|}{d}} \leq 0\ ,\ (P \geq 0)\ .$$

A sketch of the safe region in P, P_M, $|X|$ space is shown in Figure 27. (The region is truncated in the positive direction of the P_M axis for representation convenience.)

If the beam resistances are allowed to be different, the safe region D is properly defined in a space of higher dimension. An idea of the geometry of these regions can be given by drawing conditional safe regions (sections of D).

For instance, suppose that failure may be caused by excessive loading of the edge beams, but not of the interior beams. The maximum load carrying capacities of the edge beams are P_{M_1} and P_{M_4}. For any given value of P_{M_1} the safe region in P, P_{M_4}, X space is sketched in Figure 28. (Again, the safe region is unlimited in the positive P_{M_4} direction.)

Both regions in Figures 27 and 28 are concave.

(c) Probabilistic model for resistances and the effect of surviving a proof load experiment

The following analysis applies to failure being caused by overloading either one of the edge beams. A priori the maximum load carrying capacities P_{M_1} and P_{M_4} are assumed to follow a bivariate Extreme type III distribution (Gumbel, 1958, 1965), with CDF:

$$F_{b_{14}}(P_{M_1},P_{M_4})=\exp\{-[(-\log F_{b_1}(P_{M_1}))^m+(-\log F_{b_4}(P_{M_4}))^m]^{1/m}\};m\geq 1, \tag{I.127}$$

where the prior marginal CDF's $F_{b_1}(\cdot)$ and $F_{b_4}(\cdot)$ have the same Extreme type III form:

$$F_{b_1}(p) = F_{b_4}(p) = F_b(p) = 1 - e^{-(p/u)^K};\ P\geq 0,\ K>0. \tag{I.128}$$

The expected values and the variances of the marginal distributions are:

$$E[P_{M_1}] = E[P_{M_4}] = u\,\Gamma(1 + 1/K) \ ,$$

$$\sigma^2_{b_1} = \sigma^2_{b_4} = u^2[\Gamma(1 + 2/K) - \Gamma^2(1 + 1/K)] \ .$$

m, u and K are parameters of the joint distribution; m controls the correlation coefficient, which is zero for m = 1 (see equation I.127).

Survival to a proof load experiment with parameters $P = P_{PL}$ and $X = X_{PL}$ assures that the resistance point (P_{M_1}, P_{M_4}) belongs to the intersection of the safe region D in P, X, P_{M_1}, P_{M_4} space with the linear variety ($P = P_{PL}$; $X = X_{PL}$). Such intersection, denoted $D(P_{PL},X_{PL})$, is an open rectangle in the plane P_{M_1}, P_{M_4} (Figure 29). The resistances P_1^* and P_4^* in Figure 29 are defined:

$$P_1^* = \max \left\{ 0,\ P_{PL} \cdot \left(0.25 - 0.3\,\frac{X_{PL}}{d}\right) \right\} ;$$

$$P_4^* = \max \left\{ 0,\ P_{PL} \cdot \left(0.25 + 0.3\,\frac{X_{PL}}{d}\right) \right\} . \qquad \text{(I.129)}$$

Conditional on proof load survival the joint PDF of the resistances is

$$f_a(P_{M_1},P_{M_4}) = \begin{cases} 0 & \text{, if } (P_{M_1},P_{M_4}) \notin D(P_{PL},X_{PL}); \\ \dfrac{f_b(P_{M_1},P_{M_4})}{1-F_b(P_1^*)-F_b(P_4^*)+F_{b_{14}}(P_1^*,P_4^*)} & \text{, otherwise.} \end{cases} \qquad \text{(I.130)}$$

For later use we give also the prior and the posterior probabilities that the resistance of beam 1 exceeds P_1 and that simultaneously the resistance of beam 4 exceeds P_4:

$$\Pr\{(P_{M_1}>P_1) \cap (P_{M_4}>P_4)\} = 1 - F_b(P_1) - F_b(P_4) + F_{b_{14}}(P_1,P_4)$$

$$= e^{-(P_1/u)^K} + e^{-(P_4/u)^K} - 1$$

$$+ \exp\left(-\left\{[-\log(1-e^{-(P_1/u)^K})]^m+[-\log(1-e^{-(P_4/u)^K})]^m\right\}^{1/m}\right); \qquad \text{(I.130a)}$$

$$Pr_a\{(P_{M_1}>P_1) \cap (P_{M_4}>P_4)\} = 1-F_{a_1}(P_1)-F_{a_2}(P_4)+F_{a_{14}}(P_1,P_4)$$

$$= P_{f_b}\{(P_{M_1}>\bar{P}_1) \cap (P_{M_4}>\bar{P}_4)\} \Big/ \left\{ e^{-(P_1^*/u)^K}+e^{-(P_4^*/u)^K}-1 \right.$$

$$\left. +\exp\left(-\{[-\log(1-e^{-(P_1^*/u)^K})]^m+[-\log(1-e^{-(P_4^*/u)^K})]^m\}^{1/m}\right) \right\} \qquad \text{(I.130b)}$$

where $\bar{P}_i = \max\{P_i, P_i^*\}$, $i = 1, 4$; P_1^*, P_4^* are defined by equation (I.129).

(d) Probabilistic model for the maximum service loads

We assume that the maximum loads sustained by the edge beams during the design lifetime, P_1^M and P_4^M, have a joint Extreme type I distribution (of the maximum type). Following Gumbel (1965):

$$F(P_1^M,P_2^M) = \exp\{-[e^{-n\,\alpha_1(P_1^M-u_1)} + e^{-n\,\alpha_4(P_4^M-u_4)}]^{1/n}\},$$

$$n \geq 1; \quad \alpha_1,\alpha_4 > 0. \qquad \text{(I.131)}$$

In terms of the reduced variables:

$$r_1 = \alpha_1(P_1^M - u_1); \quad r_4 = \alpha_4(P_4^M - u_4) ,$$

the density function is:

$$f(r_1,r_4) = e^{-n(r_1+r_4)}(e^{-n\ r_1} + e^{-n\ r_4})^{-2+1/n}$$

$$\cdot\{n-1+(e^{-n\ r_1} + e^{-n\ r_4})^{1/n}\}$$

$$\cdot \exp\{-(e^{-n\ r_1} + e^{-n\ r_4})^{1/n}\} . \qquad \text{(I.132)}$$

n, α_1, α_4, u_1, u_4 are parameters of the distribution. The marginal mean values and standard deviations are (Gumbel, 1965):

$$E[P_i^M] \approx u_i + 0.577/\alpha_i , \quad i = 1, 4 ;$$

$$\sigma_{P_i^M} = \frac{\pi}{\sqrt{6}\ \alpha_i} , \ i = 1, 4 .$$

n controls the correlation coefficient, whose value is (Oliveira, 1961):

$$\rho = 1 - n^{-2} \quad .$$

(e) Failure probabilities

The following equations are self-explanatory.

Prior probability of failure:

$$P_{f_b} = \int_{-u_1\alpha_1}^{\infty} \int_{-u_4\alpha_4}^{\infty} f(r_1, r_4) \{1 - P_{r_b} [(P_{M_1} > \frac{r_1}{\alpha_1} + u_1) \cap (P_{M_4} > \frac{r_4}{\alpha_4} + u_4)]\} d\, r_1\, d\, r_4 \quad . \tag{I.133}$$

Posterior probability of failure:

$$P_{f_a} = \int_{-u_1\alpha_1}^{\infty} \int_{-u_4\alpha_4}^{\infty} f(r_1, r_4) \{1 - P_{r_a} [(P_{M_1} > \frac{r_1}{\alpha_1} + u_1) \cap (P_{M_4} > \frac{r_4}{\alpha_4} + u_4)]\} d\, r_1\, d\, r_4 \quad . \tag{I.134}$$

Proof loading probability of failure:

$$P_{f_{PL}} = 1 - Pr_b[(P_{M_1} > P_1^*) \cap (P_{M_4} > P_4^*)] \quad . \qquad (I.135)$$

For the probabilities $Pr_b[\cdot]$ and $Pr_a[\cdot]$, see equations (I.130).

(e) Numerical computations

For the problem under study the proof loading set D_S (see Paragraph I.4.3) coincides with the interval $[(X_{PL}, 0), (X_{PL}, P_{PL})]$ in X P space, so that X_{PL} and P_{PL} define completely the experiment. Instead of specifying a utility function and carrying out the optimization directly, one can solve the problem parametrically by calculating P_{f_b}, and finding P_{f_a} and $P_{f_{PL}}$ as functions of X_{PL} and P_{PL}. Since P_{f_a} depends implicitly on the design variables (equation I.134), a discrete parametrization was made with respect to X_{PL} and P_{PL}, using the following distribution parameters for the resistances and the loads:

In the joint CDF of the resistances, equation (I.127):

$$K = 10 \Rightarrow V_1 = V_4 \approx 0.12 \quad ,$$

$$u = 95 \Rightarrow E_b[P_{M_1}] = E_b[P_{M_4}] \approx 90.3 \text{ kips} \quad ,$$

$$m = 2 \quad ,$$

where V_j denotes the coefficient of variation of P_{M_j}.

In the joint CDF of the loads, equation (I.131):

$$\alpha_1 = \alpha_4 = 0.2419 \Rightarrow \sigma_{P_j^M} = 5.3 \text{ kips}$$

$$u_1 = u_4 = 50.61 \Rightarrow E[P_j^M] = 53 \text{ kips}$$

$$u = \sqrt{2} \Rightarrow \rho = 0.50$$

Corresponding to these parameter values, the prior failure probability is, from equation (I.133), $P_{f_b} = 0.0095$.

Since it is practically impossible to apply heavy concentrated loads, the intensity of the proof load resultant, P_{PL}, was constrained more severely for larger eccentricities (Figure 30a). Within the set of allowable experiments shown in this figure, six alternative proof tests with different eccentricities and load intensities were analyzed. The intensity and the position of the resultant are shown in Figure 30, b through g, for the six cases.

The cost of making the experiment increases with P_{PL}, but is rather insensitive to X_{PL} within the region of admissible designs. Also, the cost does not increase significantly with the number of proof loading experiments which make use of the same load P_{PL} with different eccentricities;

for this reason, two additional experiments were considered, in which the bridge was loaded twice with the maximum allowed eccentricities (Figure 30, h and i). The proof loading failure probabilities corresponding to the eight cases are collected in Table 6. From these results, one can deduce that:

(i) Small proof load intensities (say $P_{PL} < 100$ kips) have little effect on the posterior probability of failure (cases 1, 2, 3).

(ii) For larger proof load intensities (say $P_{PL} \approx 200$ kips) the posterior probability of failure is highly sensitive to the eccentricity (compare 4 with 5), decreasing with $|X_{PL}|$. Although not shown directly by the table, for large P_{PL} (say $P_{PL} > 300$ kips) the opposite dependence of P_{f_a} on $|X_{PL}|$ is expected.

(iii) Experiments with heavier loads (within the range considered here) and smaller eccentricities are more effective (compare 5 with 6).

(iv) Performing double tests with maximum eccentricities is highly rewarded (compare 3 with 7, and 5 with 8).

Experiment	Fig. 30	X_{PL}(ft)	p_{PL}(kips)	$p_{f_{PL}}$	p_{f_a}
1	b	0	100	0.0000019	0.0095
2	c	5	100	0.00057	0.0089
3	d	10	100	0.0220	0.0045
4	e	0	200	0.0031	0.0066
5	f	5	200	0.441	0.00074
6	g	0	300	0.146	0.00033
7	h	$\pm$10	100	0.040	0.0014
8	i	$\pm$ 5	200	0.568	0.000016

Table 6. Example 7: Proof loading failure probability $p_{f_{PL}}$ and posterior failure probability p_{f_a}. (Prior failure probability: $p_{f_b} = 0.0095$.)
Refer to Figure 30 for the definition of the experiments.

X_{PL} = proof load eccentricity.

Possibly using a larger set of experiments, and with a given optimality criterion, the decision maker can close the (nearly) optimal proof test parameters.

REFERENCES

1. Anderson, R.G. (1972): "Reduction of Uncertainty in Structural Design and Decision Making by First-Order Bayesian Information Analysis". M.S. Thesis, Dept. of Civil Engineering, M.I.T.

2. Aoki, M. (1967): Optimization of Stochastic Systems; Topics in Discrete-Time Systems. Academic Press.

3. Balakrishnam, A.V. (1964): "A General Theory of Non-linear Estimation Problems in Control Ssytems". J. Math. Anal. Appl., 8, No. 1, pp 4-30.

4. Bouton, I.; Fisk, M.;Trent, D.J.; Oder, H.S.; Hung, F.C. and Harris, L.A. (1967): "Quantitative Structural Design Criteria by Statistical Methods", Vol. II, Space Division, North American Rockwell Corporation.

5. Box, G.E.P. and Tiao, G.C. (1968): "A Bayesian Approach to Some Outlier Problems". Biometrika, 55, pp 119-129.

6. Box, G.E.P. and Tiao, G.C. (1973): Bayesian Inference in Statistical Analysis. Addison-Wesley.

7. Bras, R.L.; Veneziano, D.; Lenton, R.L.; Rodriguez-Iturbe, I., and Moore, S. (1974): "Notes on the Optimal Design of Sampling Networks, with an Application to Regional Rainfall Estimation". Res. Report, R.M. Parson Lab., M.I.T., September, 1974.

8. Bryson, A.E. and Ho, Y.C. (1969): Applied Optimal Control. Ginn and Company.

9. Campion, M.C.; Hanson, L.C.; Morcock, D.S., et al. (1972): "Implementation Studies for a Reliability-Based Static Strength Criteria System". Technical Report AFFDL-TR-71-178, Vol. I.

10. Churchman, C.W. (1961): Prediction and Optimal Decision - Philosophical Issues of a Science of Values. Prentice-Hall, Englewood Cliffs, N.J.

11. DeGroot, M.H. (1970): Optimal Statistical Decisions. McGraw-Hill, New York.

12. Deutsch, R. (1965): Estimation Theory. Prentice-Hall, New York.

13. Diaz Padilla, J. and Vanmarcke, E.H. (1973): "Settlement Prediction: a Probabilistic Approach". Res. Report R73-40, Dept. of Civil Engineering, M.I.T., August, 1973.

14. Edwards, W.; Lindman, H. and Savage, L.J. (1963): "Bayesian Statistical Inference for Psychological Research". Psychological Rev., 70, pp 193-242.

15. Fedorov, V.V. (1972): Theory of Optimal Experiments. Academic Press, New York.

16. Fishburn, P.C. (1968): "Utility Theory". Management Sci., 14, pp 335-378.

17. Gavarini, C. and Veneziano, D. (1972): "Minimum Weight Limit Design under Uncertainty". Meccanica, 7, No. 2 pp 98-104.

18. Gumbel, E.J. (1958): Statistics of Extremes. 2nd ed. New York: Columbia University Press.

19. Gumbel, E.J. (1965): "Two Systems of Bivariate Extremal Distributions". 35th Session of the International Statistical Institute, Beograd, No. 69.

20. Gupta, S.S. (1963): "Probability Integrals of Multivariate Normal and Multivariate t". Ann. Math. Stat., 34, pp 792-828.

21. Hampton, J.M.; Moore, P.G. and Thomas, H. (1973): "Subjective Probability and its Measurement". Journal Royal Statist. Soc. A, 136, Part 1.

22. Householder, A.S. (1953): Principles of Numerical Analysis. McGraw-Hill.

23. Liebelt, P.B. (1967): An Introduction to Optimal Estimation. Addison-Wesley.

24. Lindley, D.V. (1971): Bayesian Statistics, a Review. Regional Conference Series in Applied Mathematics. SIAM, Philadelphia, Pa.

25. Matern, B. (1960): Spatial Variation. Comm. Swed. Forestry Res. Inst. 49.

26. Raiffa, H. (1968): Decision Analysis - Introductory Lectures on Choices under Uncertainty. Addison-Wesley.

27. Raiffa, H. and Schlaifer, R. (1961): Applied Statistical Decision Theory. Division of Research, Graduate School of Business Administration, Harvard University, Cambridge, Mass.

28. Rosenblueth, E. (1973): "Ethical Optimization in Engineering", Proc. Am. Soc. of Civil Engineers, Vol. 99, No. PP2, pp 223-243.

29. Sage, A.P. and Melsa, J.L. (1971): Estimation Theory with Applications to Communication and Control. McGraw-Hill.

30. Sexsmith, R.G. (1969): "Structural Reliability Decisions Using Bayesian Techniques". Proceedings ASCE-EMD Specialty Conference at Purdue University on Probabilistic Concepts and Methods in Engineering.

31. Shinozuka, M. (1969): "Structural Safety and Optimum Proof Load". Technical Report No. 2, Dept. of Civil Engineering and Engineering Mechanics, Columbia University.

32. Shinozuka, M, Yang, J.N. and Heer, E. (1969): "Optimum Structural Design Based on Reliability Analysis". Proceedings of the Eighth International Symposium on Space Science and Technology, August 25-30, Tokyo, Japan.

33. Slepian, D. (1962): "The One-Sided Barrier Problem for Gaussian Noise". Bell System Tech. J., 41, pp 463-501.

34. Veneziano, D. (1972): "Calcolo a rottura probabilistico con più parametri casuali". Giornale del Genio Civile, No. 2, pp 129-149.

35. Zellner, A. (1971): An Introduction to Bayesian Inference in Econometrics. Wiley & Sons.

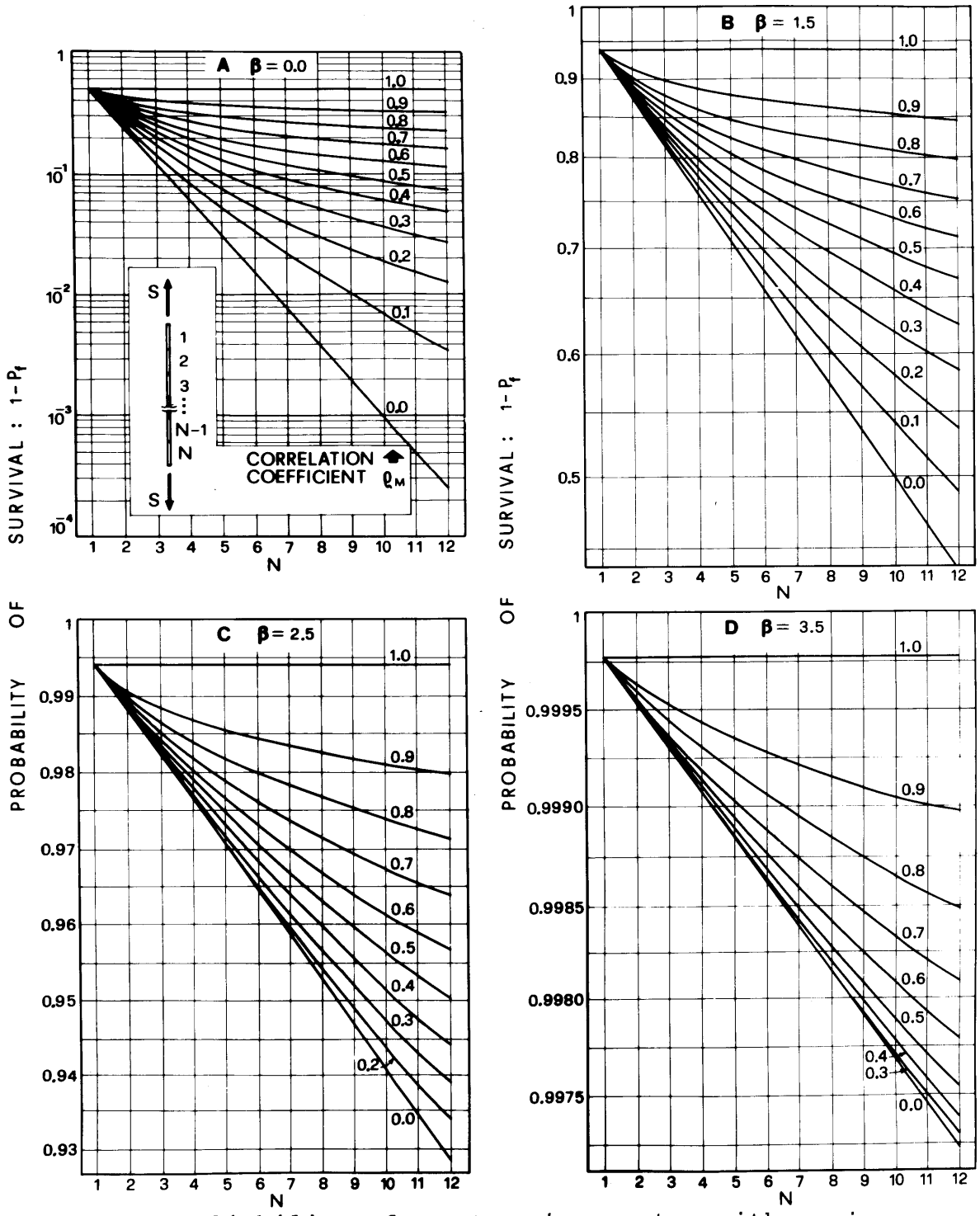
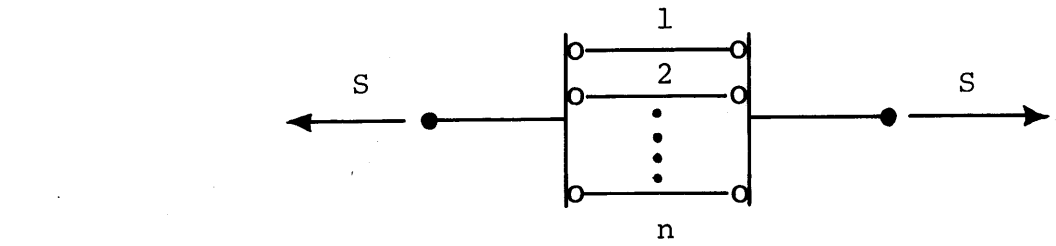
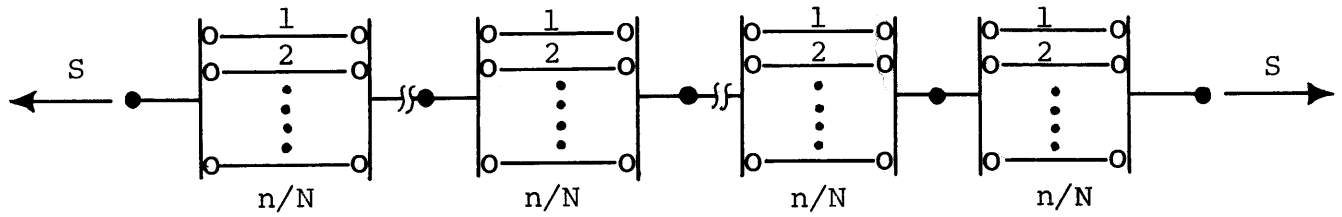


Figure 1. Reliability of an N-series system with equicorrelated normal element resistances under a deterministic load S. ρ_M = common modal (elements resistances) correlation coefficient; β = normalized modal safety margin = $(S-R_i)/\sigma_i$. Refer to Example 1 in the text.



(a)



(b)

GROUP → 1 ..... j ..... N-1 N

Figure 2. (a) n-parallel system

(b) n-series-parallel system with N parallel groupings of n/N bars

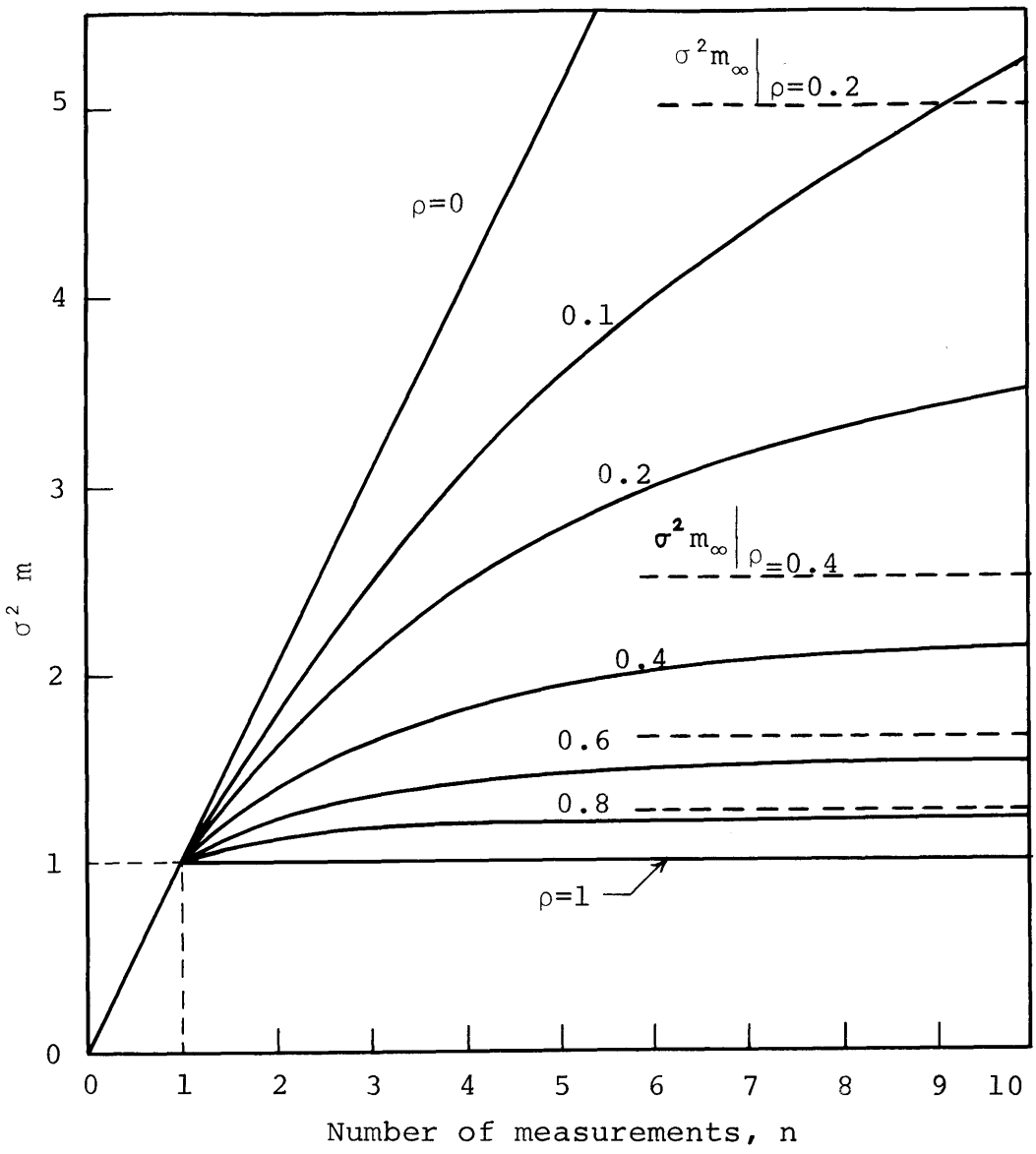


Figure 3. Informativeness of equicorrelated noisy measurements in scalar estimation. ρ = common correlation coefficient; σ^2 = prior estimation error variance; m = Fisher information "matrix", equation (I.54); $m_\infty = \lim_{n\to\infty} m$. Refer to Example 2 in the text.

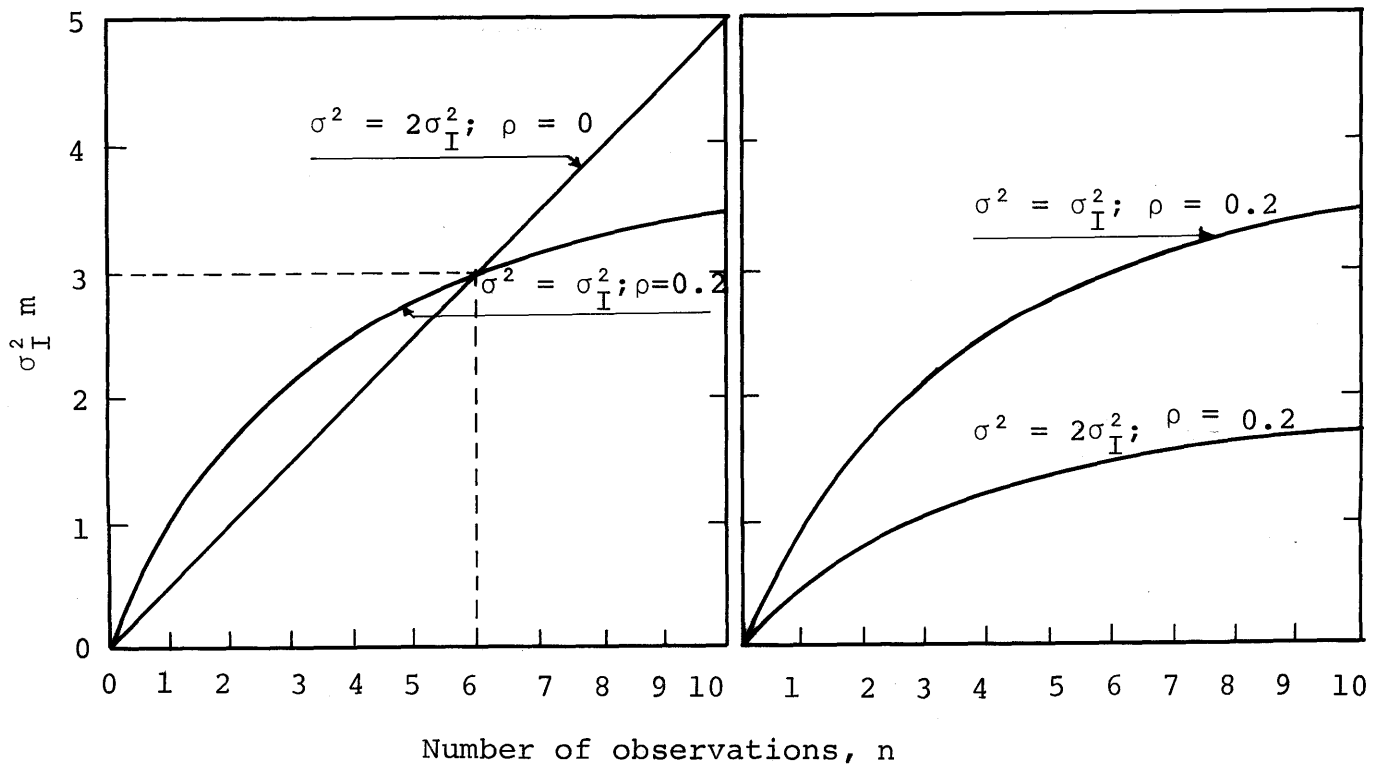


Figure 4. Comparison of the informativeness measure σ^2m of alternative experiments; for a description of the experiments see Example 2 in the text.

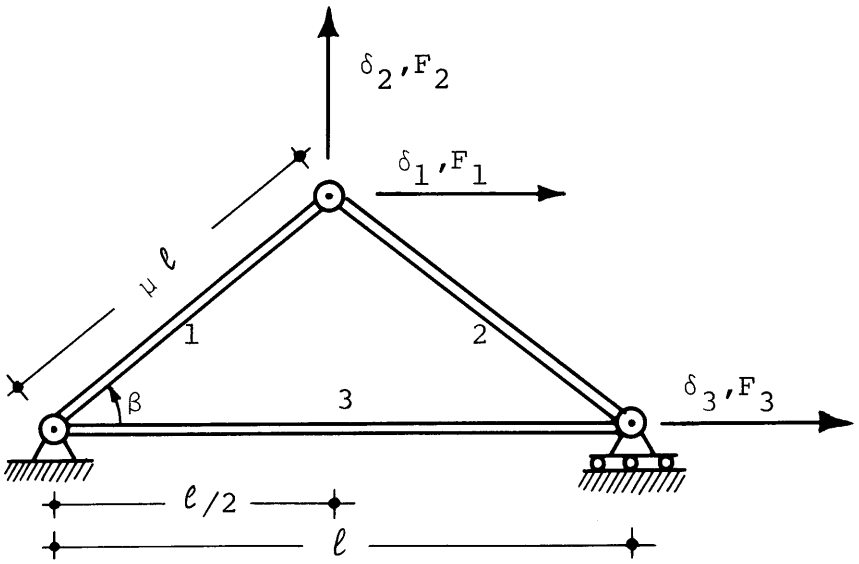


Figure 5. Simply supported 3-bars truss.
$\{\delta_i\}$ = set of compatible modal displacements;
$\{F_i\}$ = set of active external loads.
Example 3 in the text.

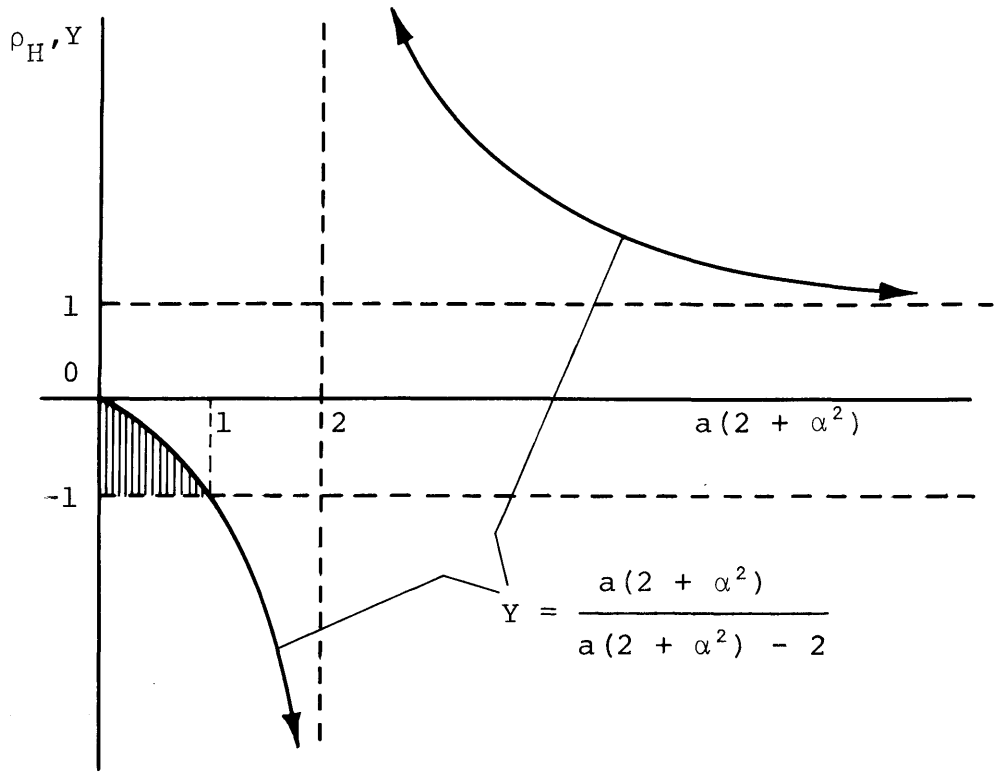


Figure 6. Example 3 (see also Figure 5). Relative effectiveness of measuring twice the hardness of the same bar versus measuring once the hardness of bar 1, and once the hardness of bar 2. ρ_H = correlation coefficient of measurement errors. For the definitions of α^2 and of a, see equations (I.59) and (I.73). The first experiment is more informative in the shaded region; the second experiment is more informative elsewhere.

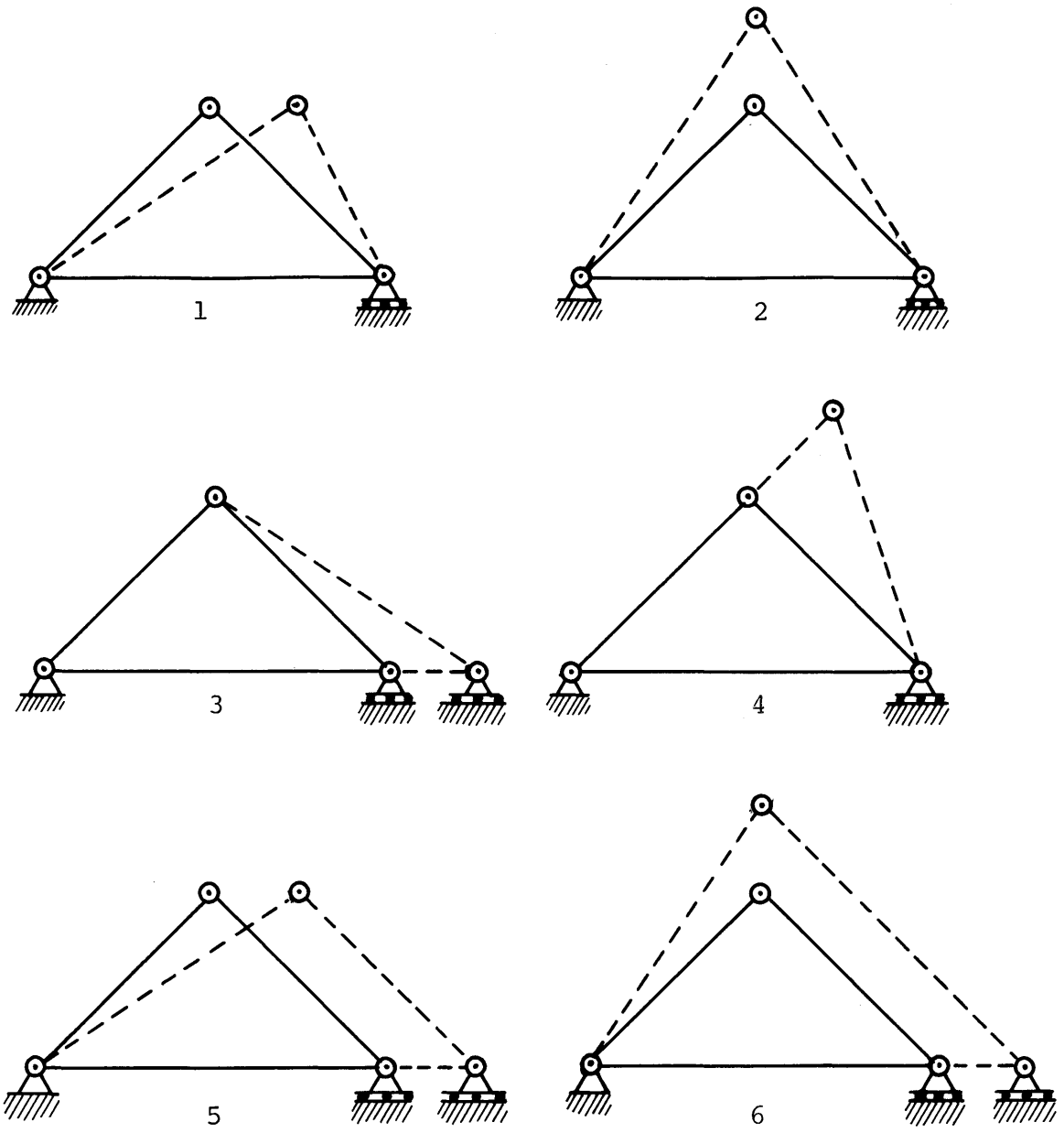


Figure 7. Six alternative load-deflection (stiffness) experiments considered in Example 3.

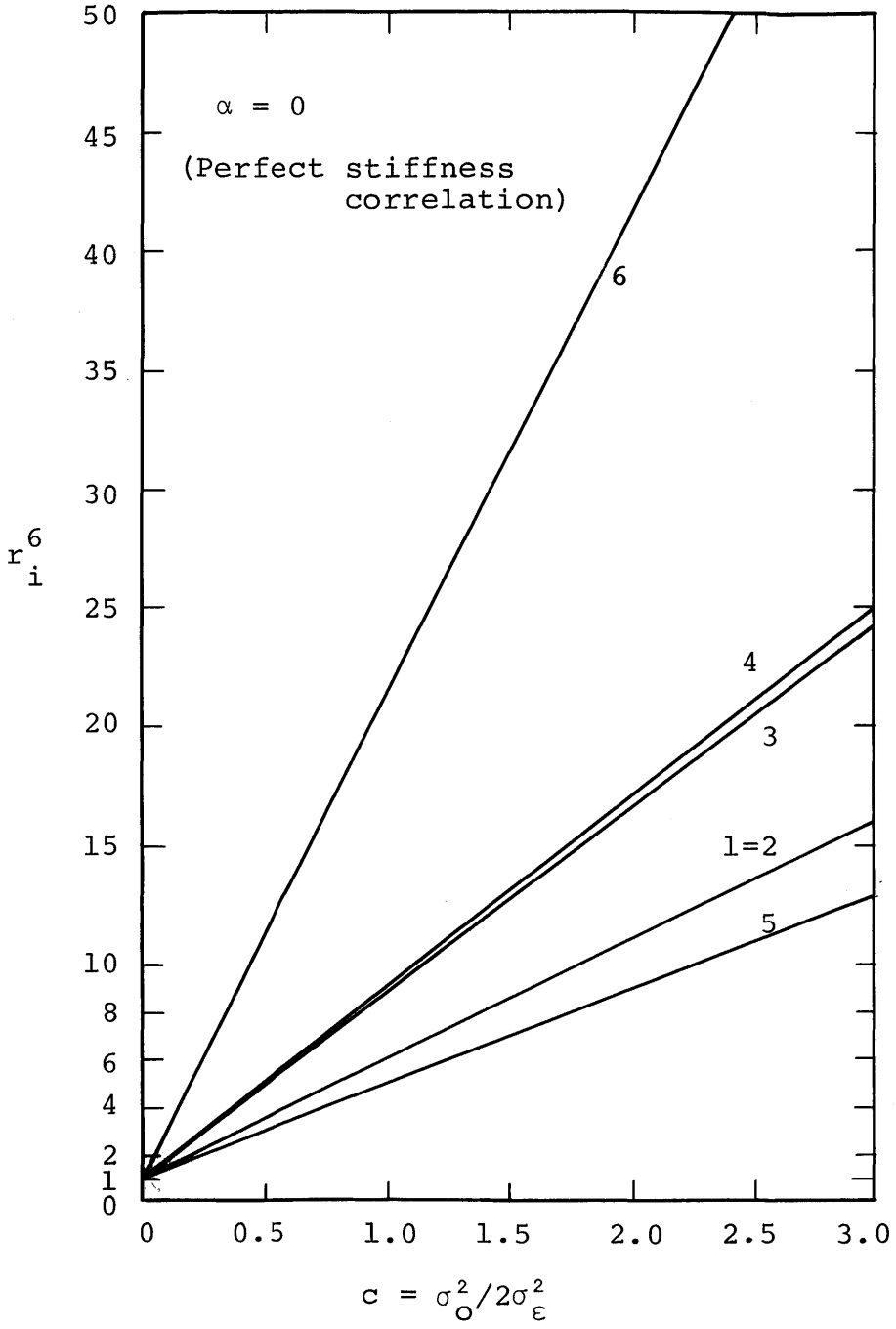


Figure 8. Informativeness of the experiments in Figure 7 as a function of the variance ratio c; no manufacturing error (perfect stiffness correlation).

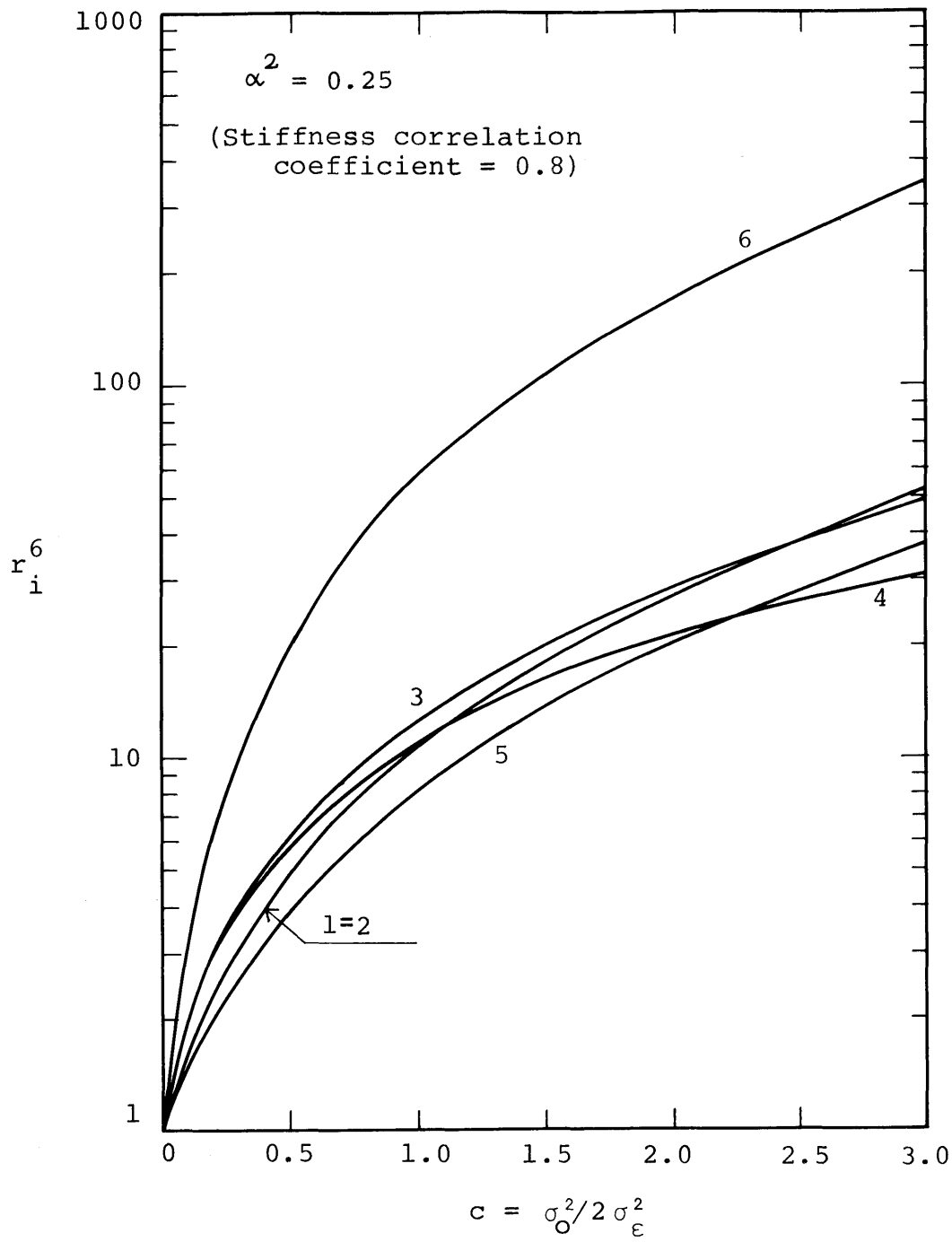


Figure 9. Informativeness of the experiments in Figure 7 as a function of the variance ratio c; manufacturing error variance = 0.25 X material stiffness variance (stiffness correlation coefficient = 0.8).

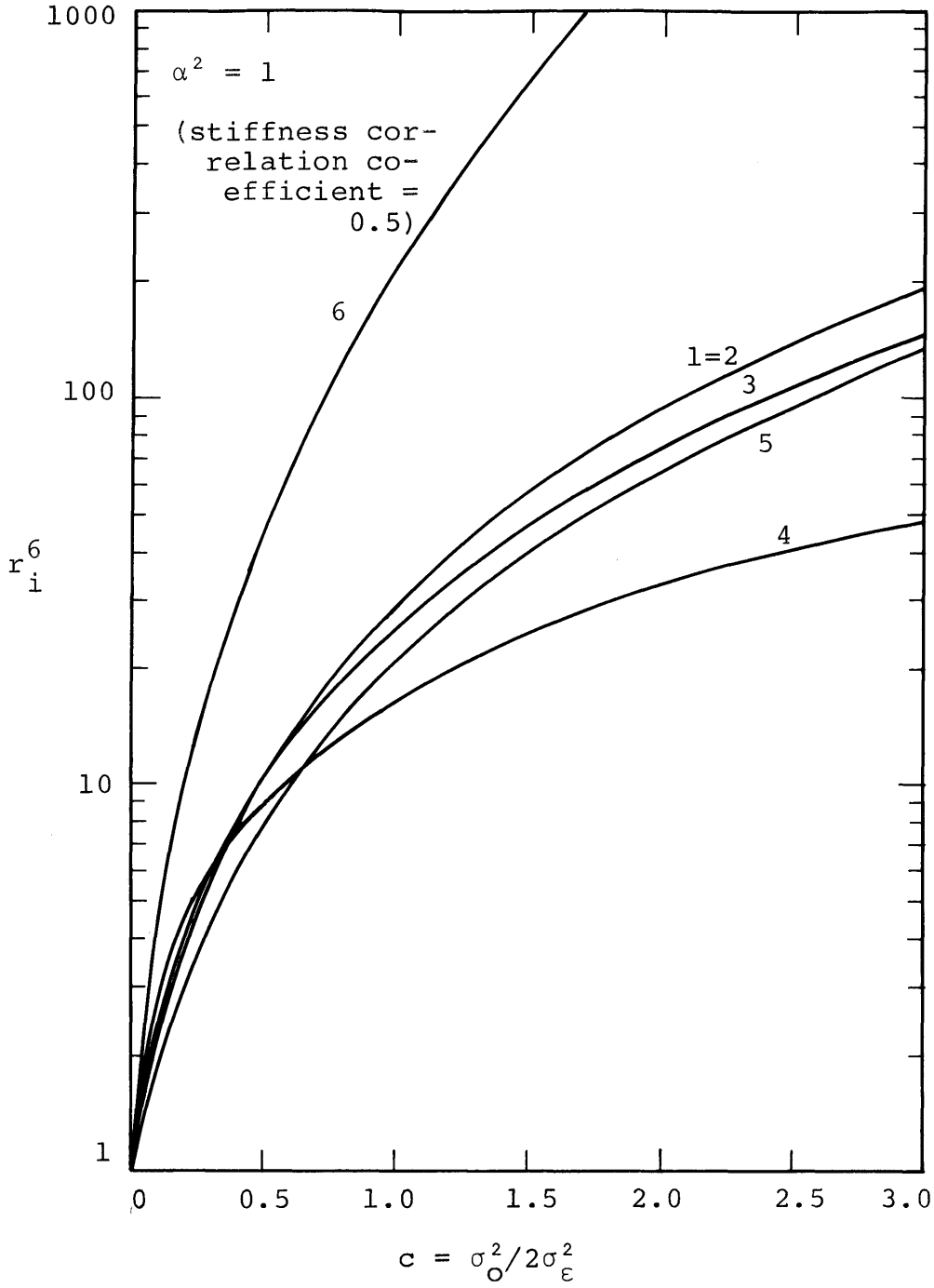


Figure 10. Informativeness of the experiments in Figure 7 as a function of the variance ratio c; manufacturing error variance = material stiffness variance (stiffness correlation coefficient = 0.5).

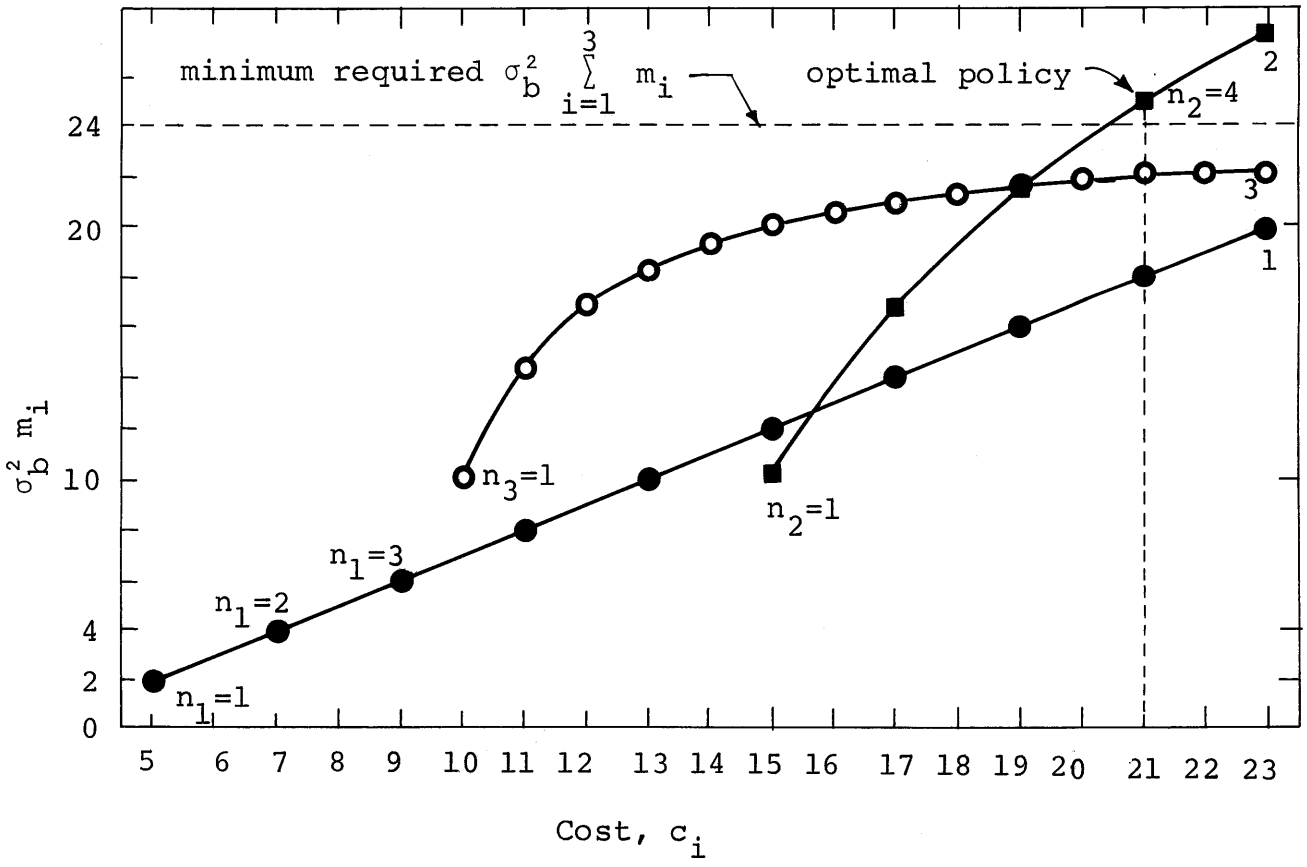


Figure 11. Informativeness versus cost of three alternative ezperiments. Each point refers to a specific facility and to a fixed number of measurements with that facility. Example 4 in the text.

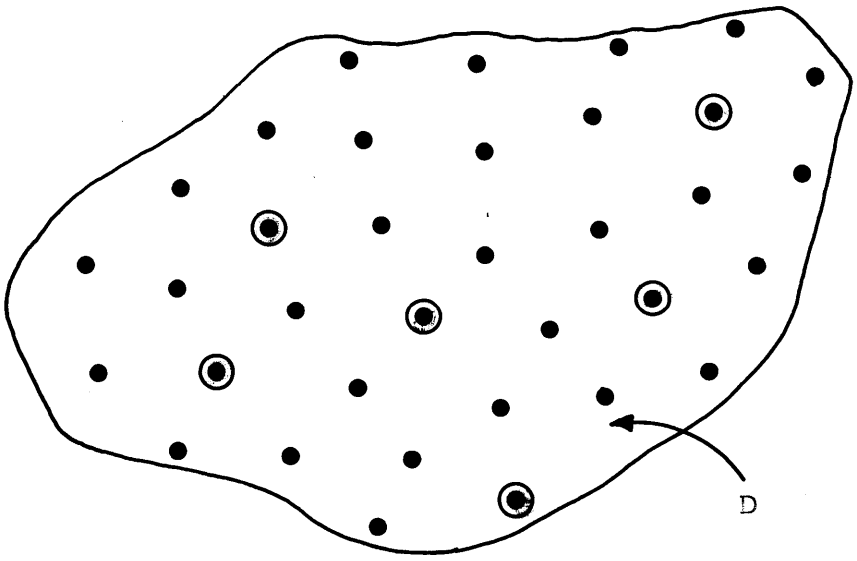


Figure 12. Discrete sampling of a two-dimensional continuous function; Example 5 in the text.

● - discretization points,

◉ - measurement points.

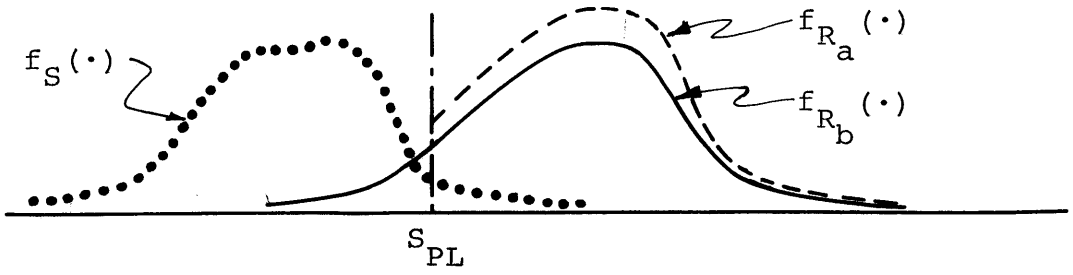


Figure 13. The fundamental case of proof loading in the classical formulation.

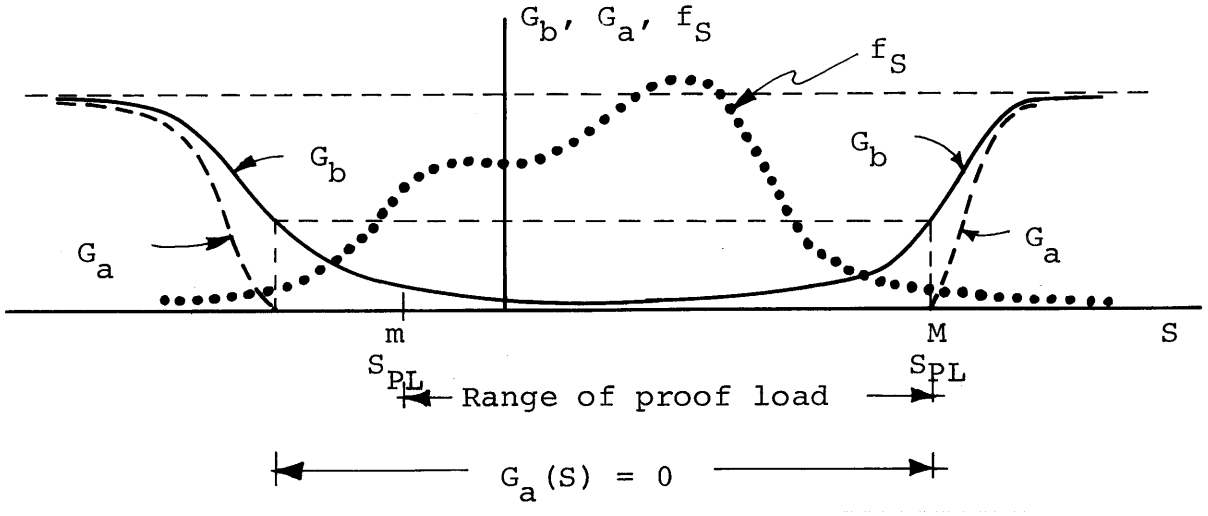


Figure 14. Generalized formulation of the fundamental problem. $G_b(S)$, $G_a(S)$ = prior and posterior probability of failure, given S.

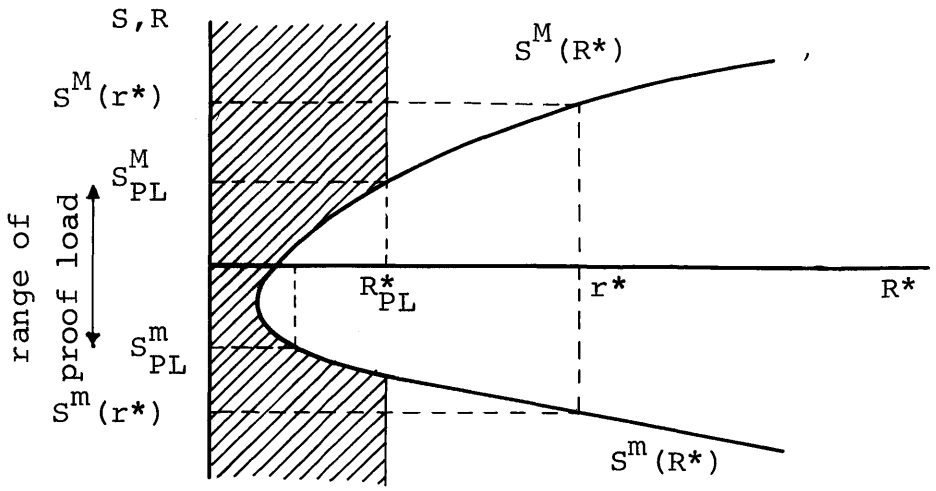


Figure 15a. Two-dimensional representation of the fundamental problem.

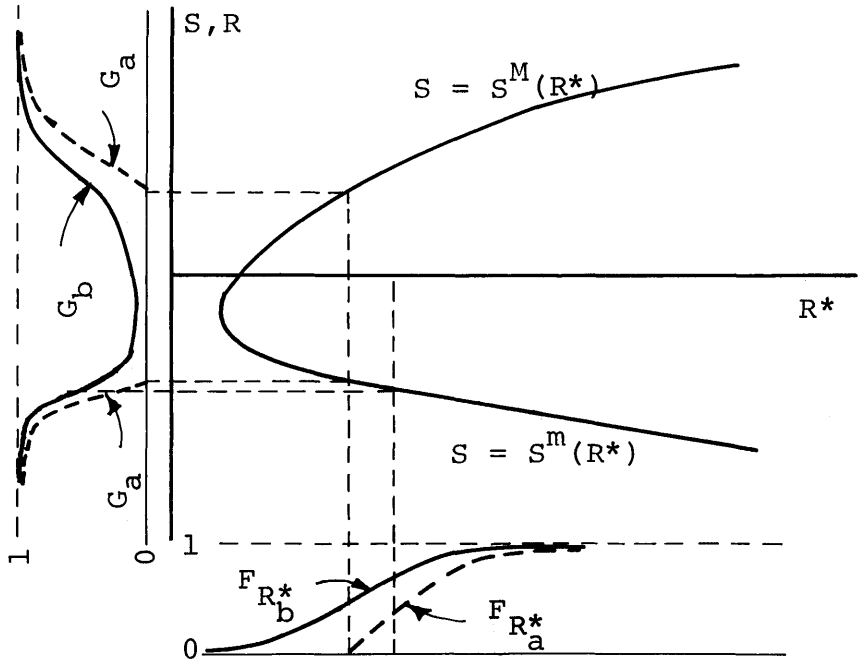


Figure 15b. Prior and posterior resistance CDF and prior and posterior failure functions.

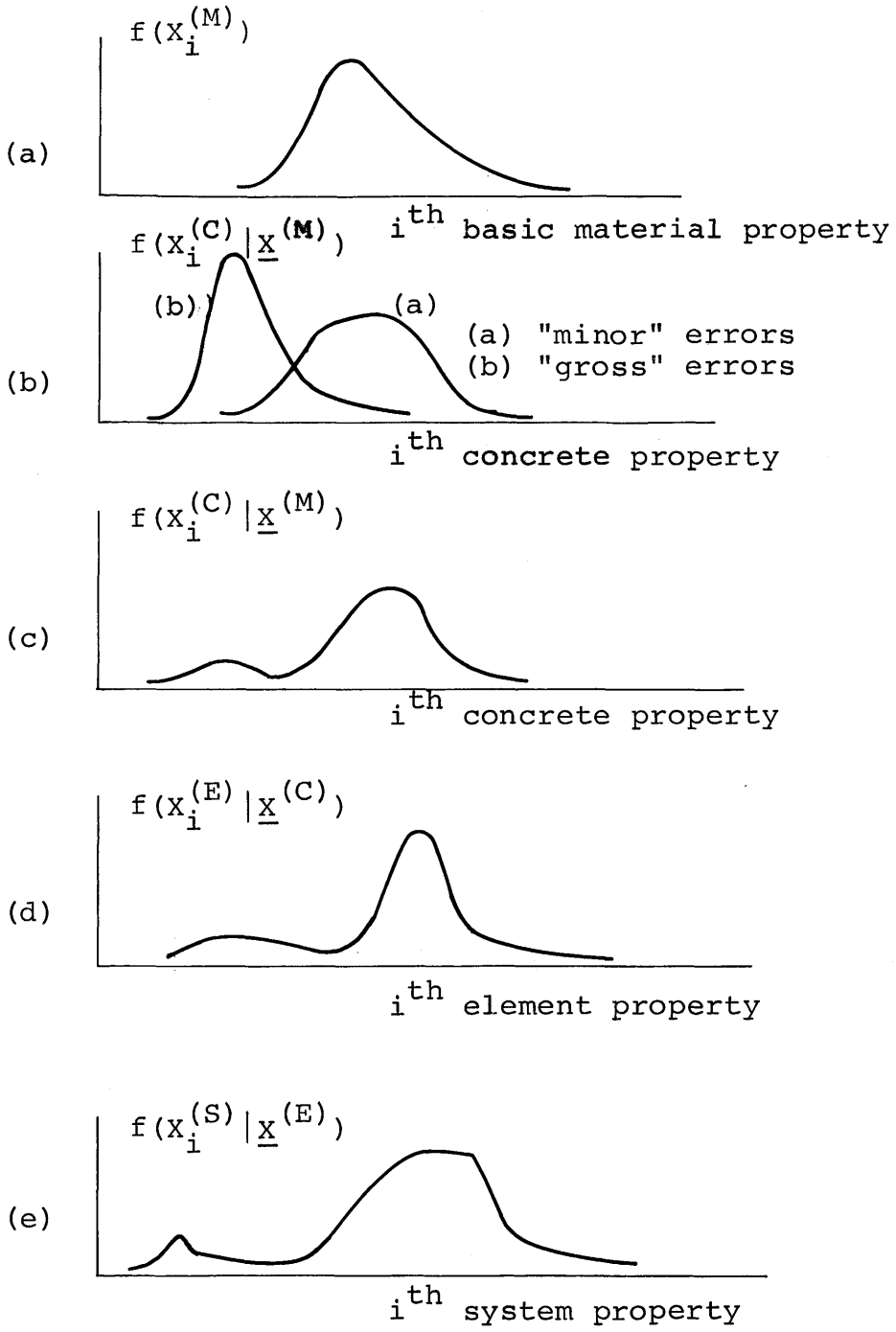


Figure 16. Conditional densities for the model in Paragraph I.4.2.

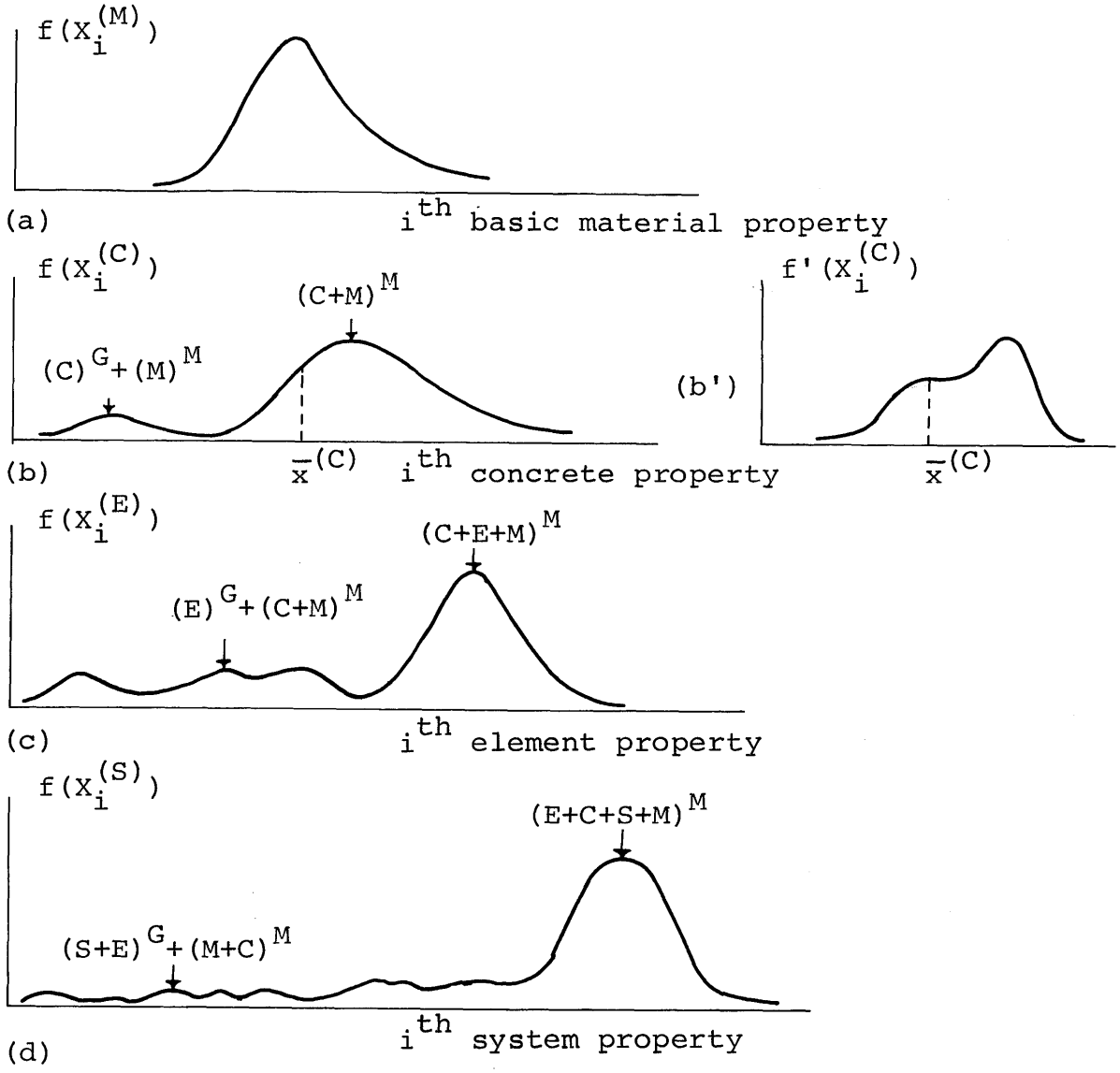


Figure 17. Unconditional densities for the model in Paragraph I.4.2.
Notation: $(\cdot)^{G}+(\cdot\cdot)^{M}$ = the density is mostly contributed by "gross" + "minor" errors in $(\cdot)$ and by only "minor" errors in $(\cdot\cdot)$.
(For simplicity only a few error combinations are indicated explicitly.)

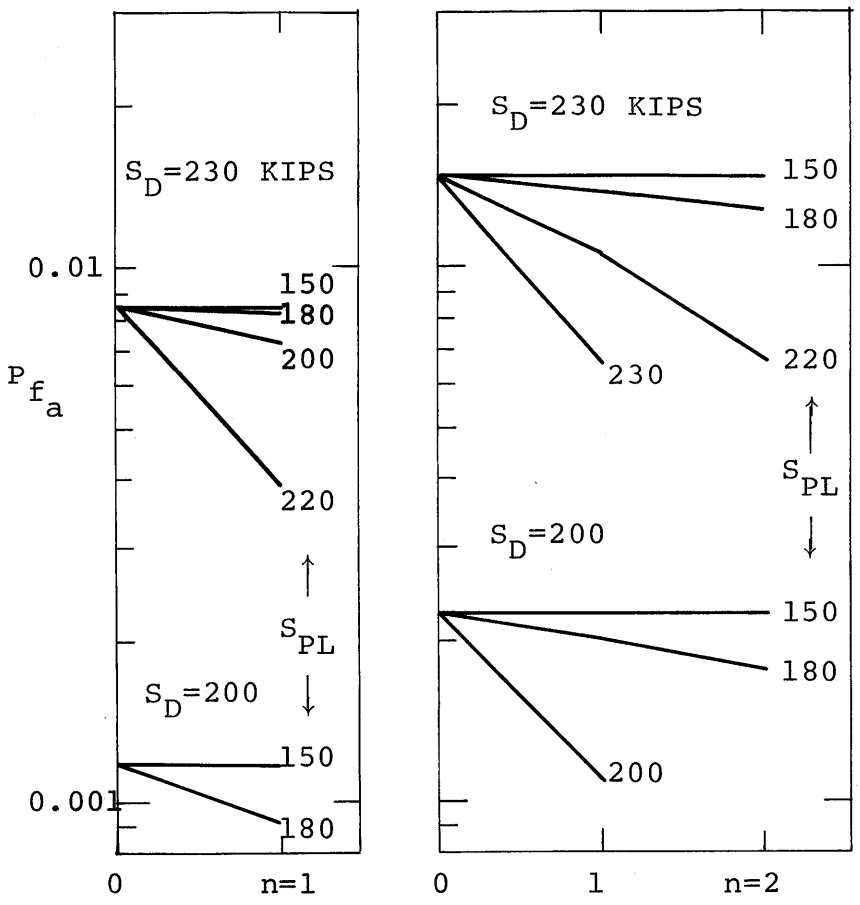


Figure 18. Posterior failure probability of an n-span bridge after survival of m spans under proof loading. S_D = design load; S_{PL} = proof load intensity.

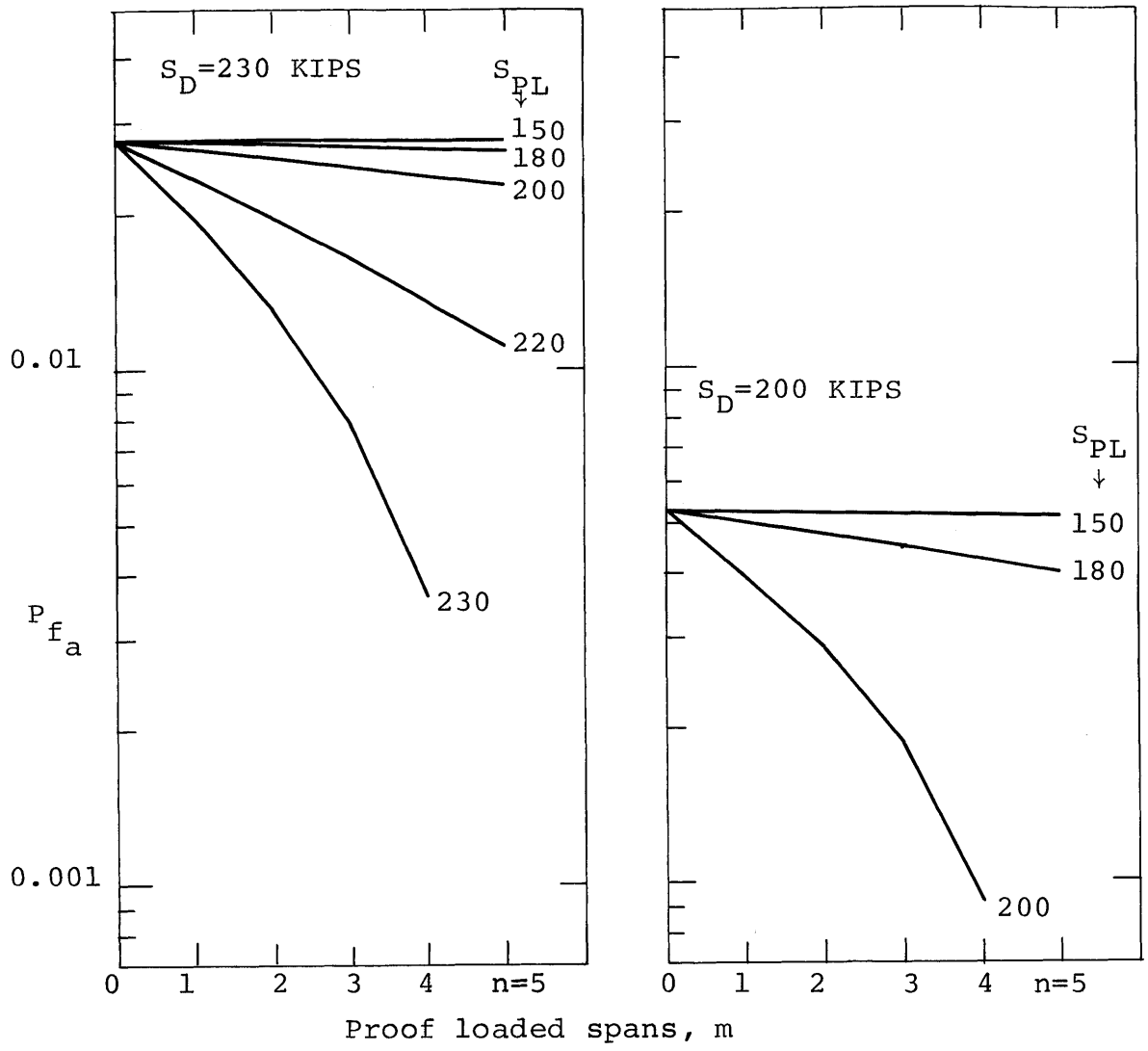


Figure 19. Posterior failure probability of a 5-span bridge after survival of m spans under proof loading.

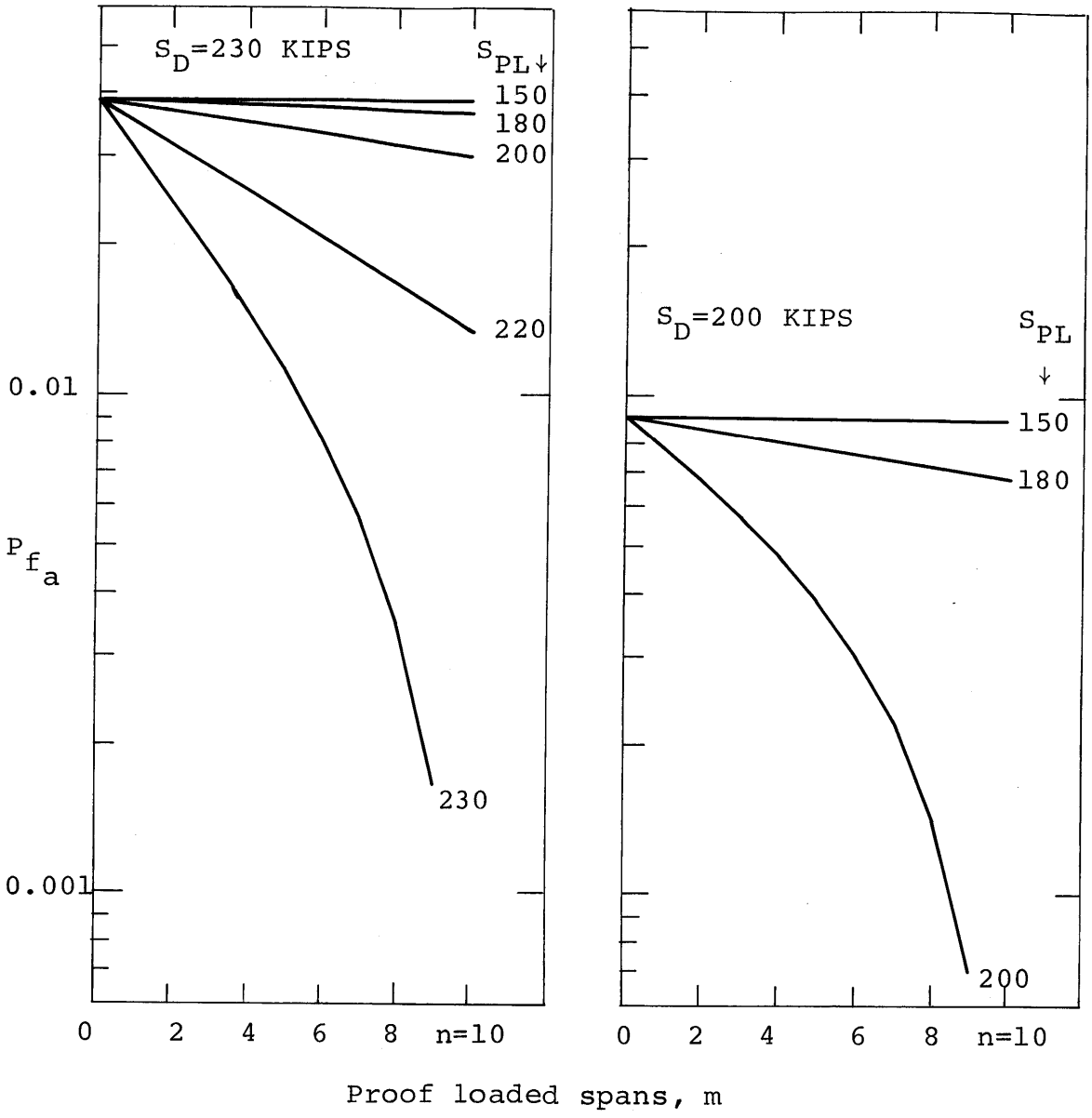


Figure 20. Posterior failure probability of a 10-span bridge after survival of m spans under proof loading.

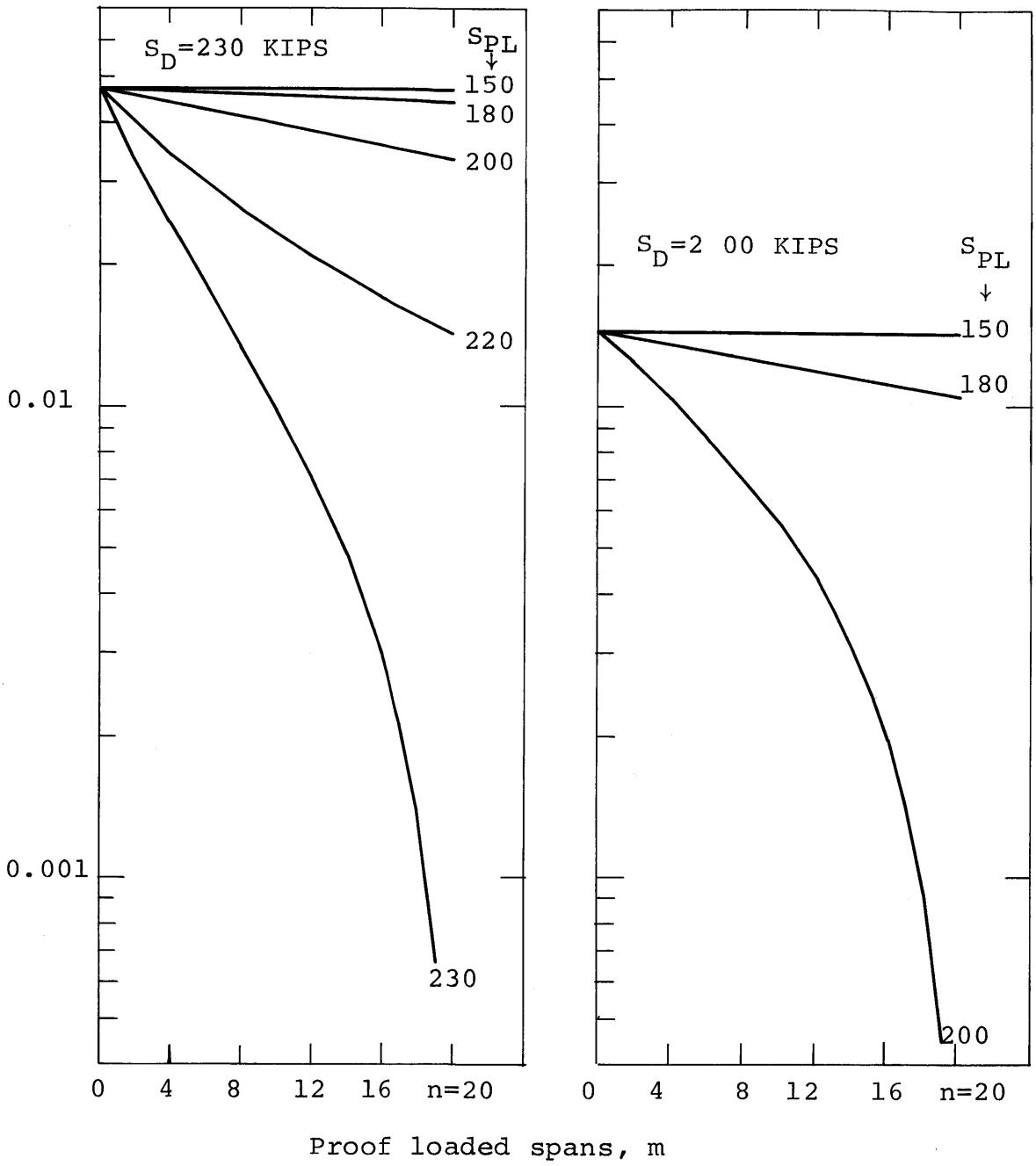


Figure 21. Posterior failure probability of a 20-span bridge after survival of m spans under proof loading.

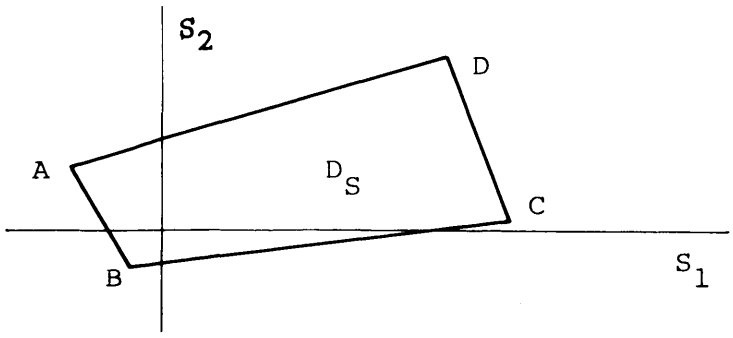


Figure 22. During proof loading $\underline{S} = [S_1, S_2]$ is given values inside the quadrilateral ABCD, including the corner load points.

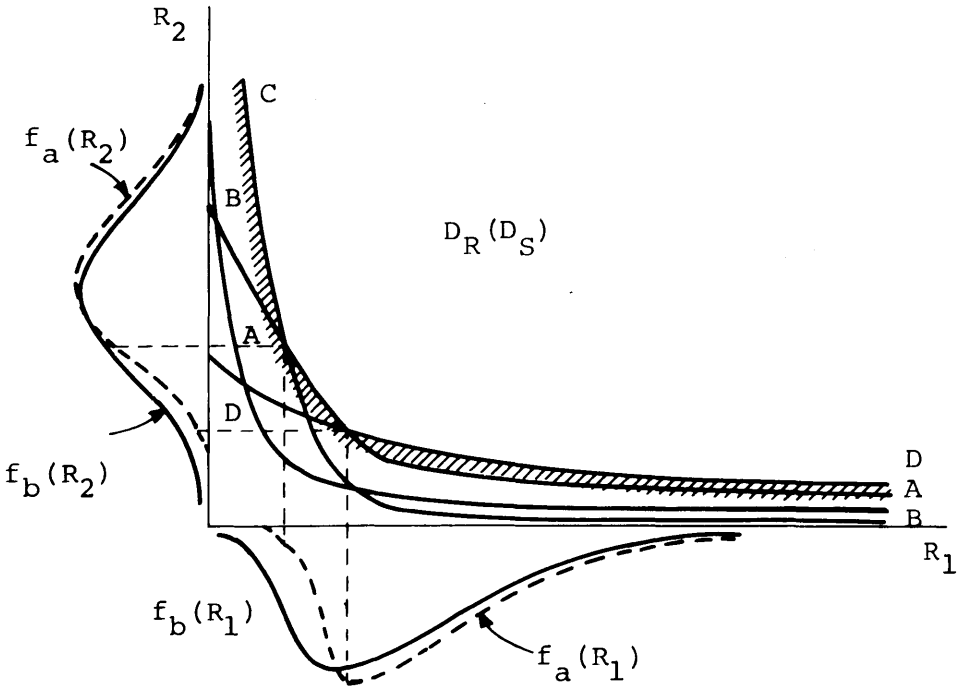


Figure 23. A posteriori and conditional on survival, $\underline{R}$ belongs to $D_R(D_S)$, intersection of the regions $D_R(A)$, $D_R(B)$, $D_R(C)$, $D_R(D)$. Also indicated are the prior and the posterior marginal distributions of the resistances R_1 and R_2.

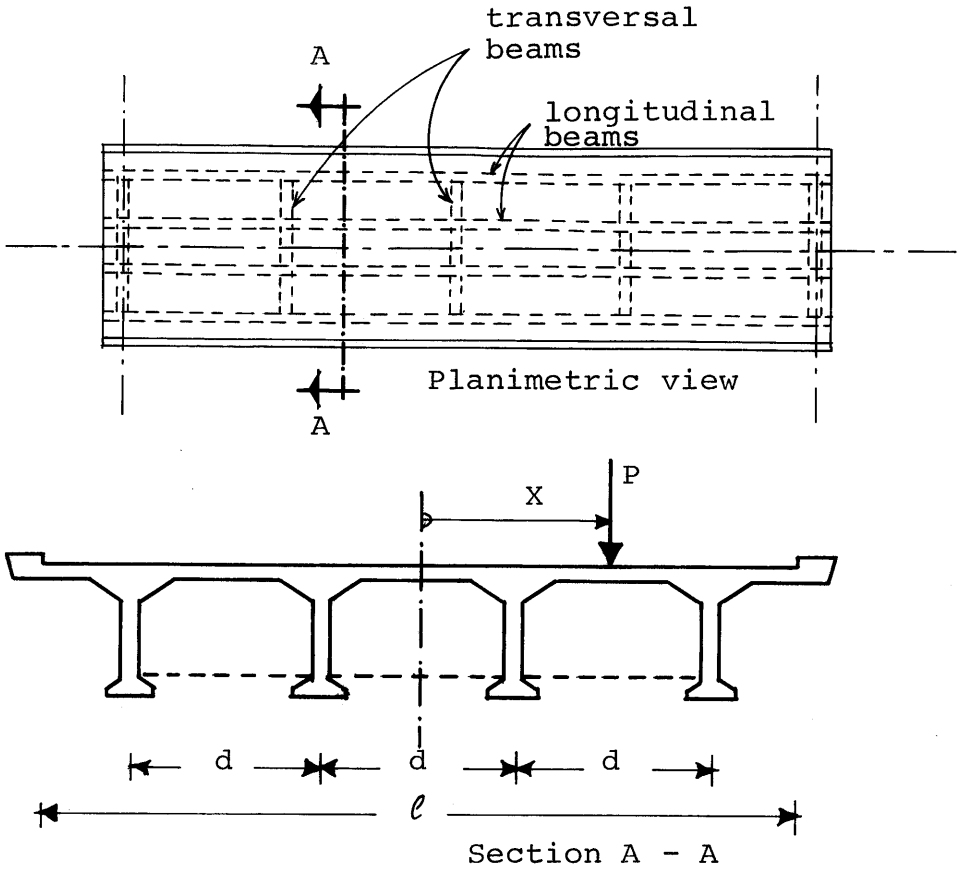
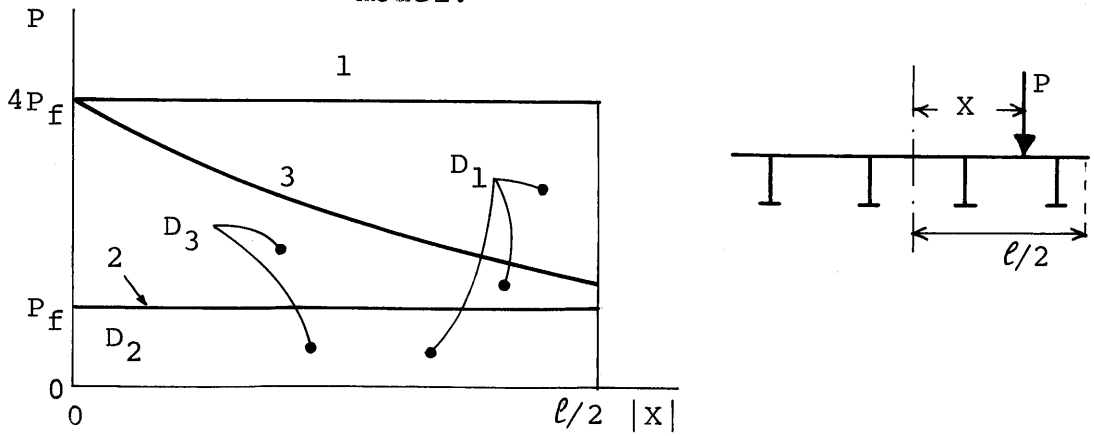


Figure 24. Structural members of the 1-span bridge deck in Example 7.

Figure 25. Example 7. Safe regions for concentrated loads;
1 - infinite-torsional stiffness-model;
2 - independent-beams-model;
3 - infinitely-rigid-transversal beams-model.



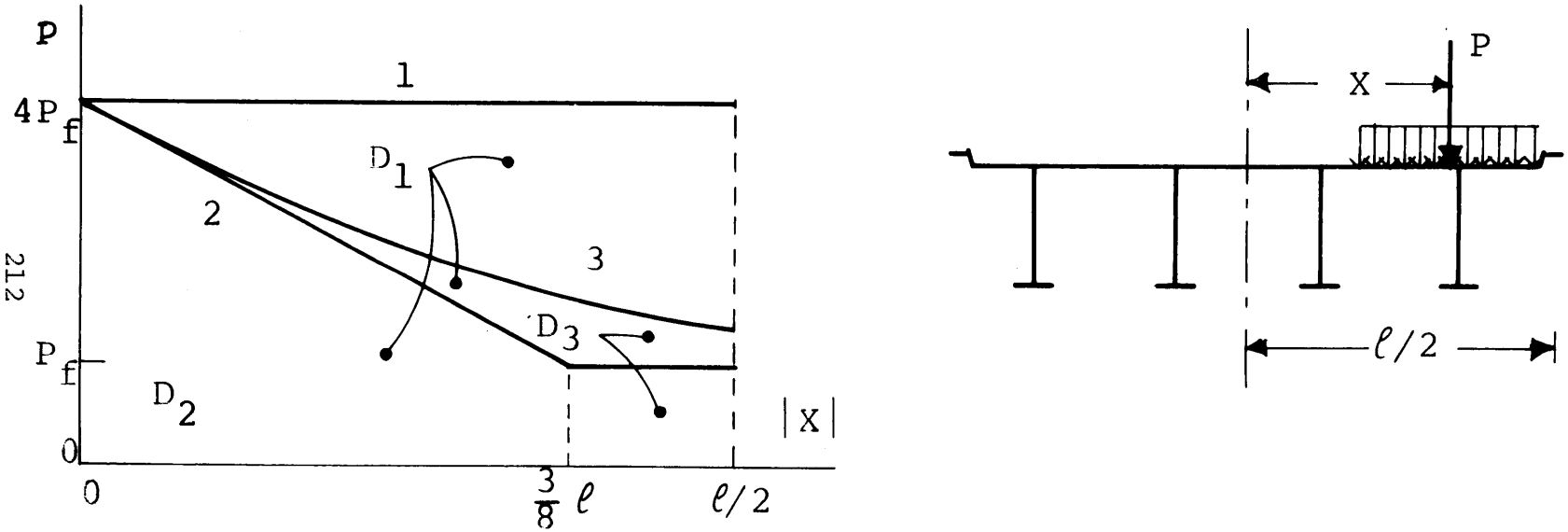


Figure 26. Example 7. Safe regions for distributed loads.
1 - infinite-torsional stiffness-model;
2 - independent-beams-model;
3 - infinitely-rigid-transversal beams-model.

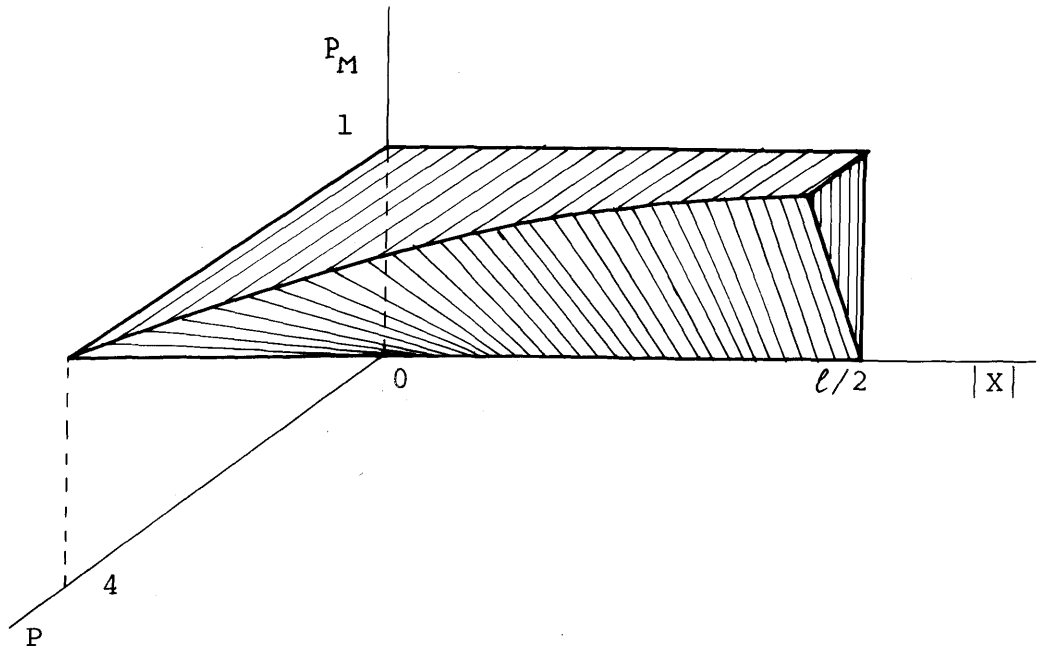


Figure 27. Safe region in |X| P P_M space; P_M = common beams resistance; infinitely-rigid-transversal beams-model.

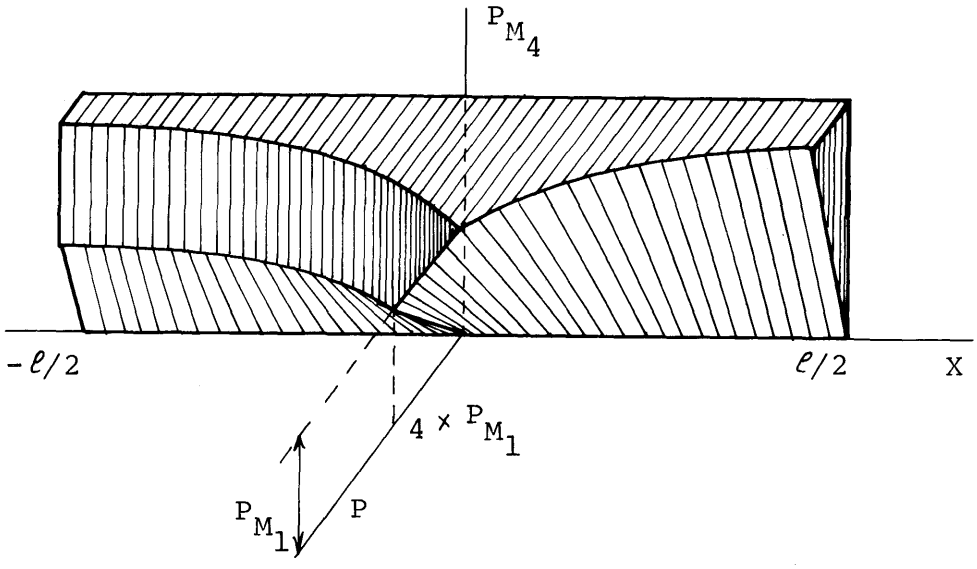


Figure 28. Unequal beams resistances. Intersection of the safe region at a fixed P_{M_1} value.

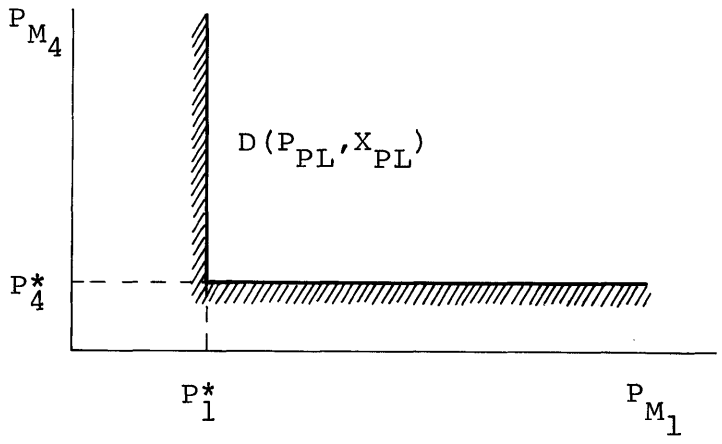


Figure 29. Intersection of D with the linear variety: $\begin{cases} P = P_{PL}\,, \\ X = X_{PL}\,. \end{cases}$

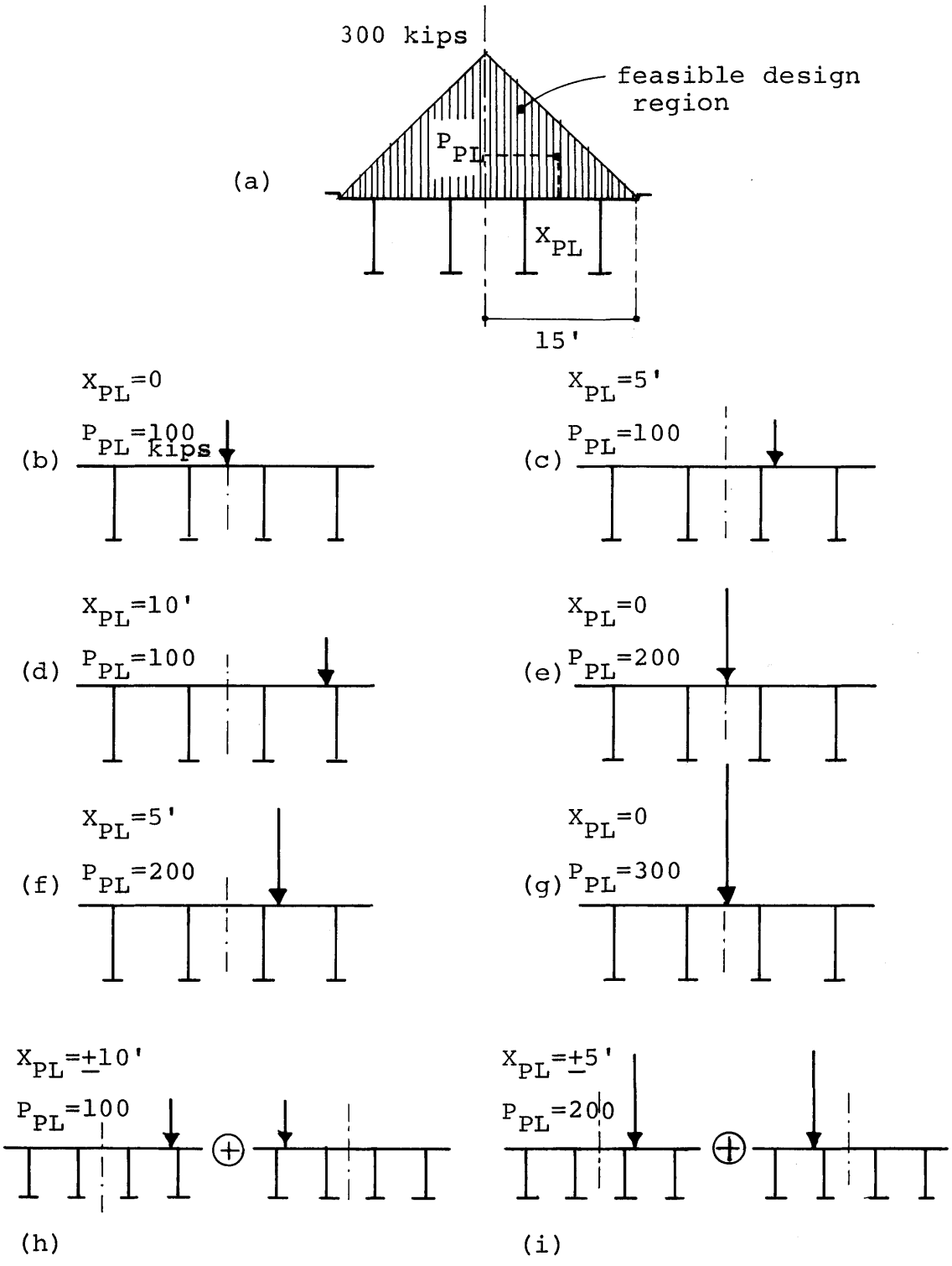


Figure 30. Feasible proof load design region, (a); and eight possible experiments, (b) through (i).

CHAPTER II

INDEPENDENT MODELS

The estimation problems studied in Chapter I are characterized by a wait-and-see attitude: first the state of nature specifies itself, then it is measured experimentally. The present chapter focuses on a different class of problems in which statistical evidence is used. Common to these problems is that the experiment does not measure directly the state of nature, either because reference is to the state of nature at a future time, or because direct measurements of it would have undesired consequences. Examples of the former case are the various meteorological loadings on a structure. As an example of the latter situation, one would not make destructive tests on structural elements which are to be used later for construction; however, one can collect useful information by testing destructively other elements from the same population.

The availability of indirect experimental information of this type is characteristic of decision problems in which the "future" has to be predicted from the "past". In mathematical terms the "past" is contained in the outcome of an experiment E, and the "future" is the outcome of an hypothetical future experiment F. Any structural design requires

prediction both of the properties of the system not yet in existence, and of the environment in which it will be operating.

Probabilistic models. Consider first the case in which E and F consist of, respectively, past and future realizations of the same stationary, independent (white) sequence $\{Y_i\}$. Indeed, because of the whiteness assumption, there is no need for keeping the notion of time: the "past" is simply a set of independent random variables with the same probability distribution as any "future" realization (consider the foregoing example of making destructive tests on structural members). Conditional on the population distribution F(Y) the "past" and the "future" are independent, and in fact if $F(\cdot)$ is known there is no reason for collecting statistical data. On the contrary, if $F(\cdot)$ is unknown, observing the past is informative on the distribution of the sequence, and so also on its future realizations. The case of E and F being regression experiments with independent residuals fits the same scheme since the notion of time is again irrelevant. This type of prediction problem is studied in this chapter.

A different situation is faced when the observed past and the predicted future belong to the same realization of a process with memory. In this case observation of the

past is informative on the future in two ways:

(i) If the process is ergodic or if data are available from different realizations, one can estimate the probabilistic structure of the process (mean value, autocovariance function, and so on), just as discussed above for white sequences.

(ii) The available data are used to estimate the past and the present and, through the memory properties of the process, also to reduce the prediction error. If the probabilistic structure of the process is given a priori, past measurements are used for prediction only in the second modality; this brings one back to an estimation problem of the type discussed in Chapter I.

Problems of prediction from first-order autoregressive processes are studied in Chapter III.

It is perhaps worth saying that estimation and prediction are not mutually exclusive notions. In the example of testing structural members, destructive tests on members from the same population and nondestructive tests on the members used for construction can be combined to reduce most effectively the uncertainty of the resistance of the latter group. A numerical example will be made in Paragraph II.2.2.

Example 6 in Chapter I dealt with a similar situation.

Prediction in restricted and extended forms. The word "prediction" is used in the statistical literature with different meanings. As remarked by Aitchison and Sculthorpe (1965), most of the time "prediction" is given the meaning of "prediction inference", the purpose being to make statements about the outcome of F. Denote this outcome $\underline{Y}$. Based on a past experiment E, prediction in this sense may consist in finding either the best point predictor $\hat{\underline{Y}}(E)$, or a prediction region $D_p(E)$ in $\underline{Y}$-space containing the outcome of F with given probability P, or the full distribution of $\underline{Y}$ conditional on the outcome of E (Bayesian approach). We call this the restricted form of prediction. The restricted form is appropriate for structural reliability prediction.

The second way in which one can make statistical predictions is to associate a utility measure $U(D,\underline{Y})$ with each future outcome $\underline{Y}$ and with each prediction region D in $\underline{Y}$-space. Then the problem is to define D as a function of the outcome of E such that the expected utility is maximized. Clearly the optimal decision D(E) depends on the utility function. For instance, in a structural design problem U(D,Y) might be defined as follows:

$$U(D,Y) = \begin{cases} \bar{U}(D) + U_S(D) & \text{if } \underline{Y} \in D \ ; \\ \\ \bar{U}(D) + U_F(D) & \text{if } \underline{Y} \notin D, \end{cases}$$

where D is the safe domain of the system in the space of the basic random vector $\underline{Y}$ (note: D is a design "variable"); $-\bar{U}(D)$ is the cost of a system with safe region D; $U_S(D)$ is the utility of such a system not experiencing failure, and $U_F(D)$ is the (negative) penalty for failure. We call this the extended form of prediction (Aitchison and Sculthorpe use the terminology: prediction problems of decision type). The extended form is appropriate for structural design and code-writing purposes. The available results are few (see Aitchison and Sculthorpe, 1965; Aitchison, 1966; Guenther and Terragno, 1964; Dunsmore, 1966, 1968, 1969). Our main concern will be with the restricted formulation.

"Prior" knowledge. One additional preliminary remark will be useful. Prediction problems with independent sequences differ according to the degree of knowledge of $F(\underline{Y})$ before data are collected. A common assumption in classical prediction theory is that the distribution type is known, but not the parameters (parametric, or parameter-free prediction). Results in the form of prediction regions

are available for particular types of population distributions (Proshan, 1953; Fraser and Guttman, 1956; Guttman, 1959; Goodman and Madansky, 1962; Thatcher, 1964; Chew, 1966; Aitchison and Sculthorpe, 1965; among others). Another case which is considered by classical statistics is that of complete ignorance of the distribution $F(\underline{Y})$ (i.e., not even the shape of the distribution is known). This assumption leads to the theory of nonparametric (or distribution-free) statistical prediction. Important contributions to it were made by Wilks(1941), Wald (1943), Tukey (1967), Fraser (1951,1953), Danziger and Davis (1964), Kemperman (1956); see also the review in Guttman (1970), Chapter I.2.

From a Bayesian viewpoint the aforementioned states of knowledge are only extreme cases out of a continuous spectrum which ranges from complete ignorance to perfect knowledge of $F(\cdot)$. Intermediate states of knowledge are described through a "prior" probability distribution of the parameters (which case parallels classical parametric inference; see Aitchison (1964,1966); Aitchison and Sculthorpe (1965); Guttman (1970)), or through a prior distribution over all possible $F(\cdot)$ functions (which case parallels classical nonparametric inference; see Ferguson (1973)).

In this chapter an impartial viewpoint is taken, and results from both Classical and Bayesian inference are

given. However, for conceptual as well as for practical reasons the author's preference stays with the Bayesian approach; some motivations for this preference are given in Section II.2.

Simple and simultaneous prediction. When F consists of a single realization of the process at a specified "future time", we say that prediction is "simple".

Statistical prediction for structural reliability when the random quantities (say of the load-type) change with time requires a more elaborate analysis. For a discrete-time stochastic model the problem consists of predicting the maximum of m future realizations from an available sample. m may be fixed, or may be a random function of the "life time" of the system. For a given m the problem is usually called one of "simultaneous prediction". We shall deal with simple prediction in the first part of the chapter (Sections II.1 and II.2), and with simultaneous prediction in the second part (Section II.3).

As already mentioned, the sample may consist of censored observations; a simple problem of this type was studied in Chapter I, Example 6. More will be said in this chapter on a general Bayesian approach to prediction from censored data (Paragraph II.2.5).

The organization of the chapter is as follows. After an exposition of the basic elements of frequentist and Bayesian prediction (Section II.1), Section II.2 deals with simple-prediction problems (one future realization) for univariate (Paragraphs II.2.1 and II.2.2) and for multivariate random sequences (Paragraphs II.2.3 and II.2.4). In both cases the parametric and nonparametric frequentist approaches are reviewed first and parallel Bayesian results follow. Paragraph II.2.5 deals briefly with prediction from censored data. The last section (Section II.3) treats problems of simultaneous prediction for univariate sequences.

The chapter has the format of a review of distribution theory; as a result, the reader who is interested in the application-side of statistical prediction will find this chapter less well "motivated" than Chapter I. On the other hand such a reader will perhaps find it easier to apply the results in this chapter, as opposed to the results in the last chapter. The extensive tables in Appendices A, B and C are intended to ease such applications. Examples are used to clarify results more than to further develop the theory; an exception in this sense is the application of simultaneous prediction to damage accumulation and to fatigue failure, a problem which is studied in Paragraph II.3.4.

II.1 PRINCIPLES OF FREQUENTIST AND BAYESIAN PREDICTION

Consider the following problem of decision under risk. A structural system must be designed to survive a future vector-valued load $\underline{Y}$ with given probability P. The properties of the system can be controlled deterministically by the designer; instead, $\underline{Y}$ is modeled as a random realization point of a stationary independent sequence $\{\underline{Y}_i\}$, at a given future time. The marginal distribution of the sequence is not known exactly. As an additional piece of information, a sample of size n is available from the same sequence. For instance one might have observed $\underline{Y}_1$, $\underline{Y}_2$, ..., $\underline{Y}_n$, and the problem might be to design a system which will survive the next load event $\underline{Y} = \underline{Y}_{n+1}$ with probability P.

For any given system the set of points in $\underline{Y}$-space corresponding to survival is called the "safe region" (with respect to $\underline{Y}$). Let D denote this region. The problem of design under reliability constraint consists in finding a system whose safe region D contains $\underline{Y}$ with probability P. A region with this property is said to have "P-content", or "P-coverage"; in turn P is called the (probability) content or coverage of D. Because $\underline{Y}$ is a "prediction variable" whose value gets specified at a later time, D is often referred to as a "prediction region".

In problems of reliability analysis D is given and the objective is to find the probability content of D.

A similar formulation for structural reliability analysis and design holds when survival depends on a set of future load realizations, say $\underline{Y}_{n+1}, \underline{Y}_{n+2}, \ldots, \underline{Y}_{n+m}$. If survival requires that all these load vectors belong to a region D in $\underline{Y}$-space, then D is called a simultaneous prediction region (of P-content, if survival occurs with probability P).

The usefulness of prediction regions for probabilistic structural analysis and design is therefore apparent. Their construction follows different lines, depending on whether one moves within the frequentist or the Bayesian framework. Both these approaches to prediction are elegantly reviewed by Aitchison and Sculthorpe (1965) under the assumption that the generating sequence is white with known distribution type but unknown parameters (parametric prediction), or that both the distribution type and the parameters are unknown (nonparametric prediction). A concise exposition of the frequentist and Bayesian approaches follows.

Frequentist approach

Under this approach, for each distribution parameter there are only two possible states of knowledge before

sampling: either the parameter's value is known, or we are totally ignorant of it. Let $\underline{\Theta}$ denote the vector of unknown parameters. The experiment E consists of observing $\underline{Y}_1, \underline{Y}_2, \ldots, \underline{Y}_n$, while the next realization $\underline{Y} = \underline{Y}_{n+1}$ is the object of prediction.

Collect $\underline{Y}_1$ through $\underline{Y}_n$ into the observation vector $\underline{Z}$. With respect to the next observation the probability content of a region D in $\underline{Y}$-space is, for any given $\underline{\Theta}$:

$$P(D|\underline{\Theta}) = \int_D d\, F(\underline{Y}|\underline{\Theta}) \qquad \text{(II.1)}$$

where $F(\cdot|\underline{\Theta})$ is the conditional CDF of the sequence $\{\underline{Y}_i\}$. Since here $\underline{\Theta}$ is not known, equation (II.1) cannot be used. At most $\underline{\Theta}$ can be estimated from $\underline{Z}$. With this type of statistical uncertainty the problem of constructing a prediction region is that of finding a rule $D = D(\underline{Z})$ such that <u>for any given $\underline{\Theta}$</u> D has probability content P:

$$E_{\underline{Z}}[P(D(\underline{Z})|\underline{\Theta})] = \int_{\text{all } \underline{Z}} f(\underline{Z}|\underline{\Theta}) \int_{D(\underline{Z})} d\, F(\underline{Y}|\underline{\Theta})\, d\, \underline{Z} = P. \qquad \text{(II.2)}$$

$D(\underline{Z})$ is called a <u>parameter-free (or parametric) prediction region of P-content</u>. For the parallel formulation when

prediction problems are in extended form see Aitchison and Sculthorpe (1965).

From equation (II.2) it is apparent that $D = D(\underline{Z})$ is a decision rule which produces prediction regions having expected content P, the expectation being taken over all the possible observations $\underline{Z}$. Logically this is not the proper attitude if decisions are made after observing $\underline{Z}$. On the other hand classical statistics is incapable of accounting for parameters uncertainty in a different form. The impass can be avoided by revising the frequentist interpretation of probability and by using the so-called "fiducial argument". By this argument, if before making any observation the quantity $(\underline{Z}|\underline{\Theta})$ is uncertain with distribution $F(\underline{Z}|\underline{\Theta})$, after observing $\underline{Z}$, say $\underline{Z} = \underline{Z}^*$, the parameter vector $\underline{\Theta}$ is random with distribution

$$d\,F(\underline{\Theta}|\underline{Z}^*) \propto d\,F(\underline{Z}^*|\underline{\Theta}) \quad \text{(now a function of } \underline{\Theta}).$$

Then a prediction region $D(\underline{Z}^*)$ is of P-content if

$$P[D(\underline{Z}^*)] = \int_{\text{all } \underline{\Theta}} \int_{D(\underline{Z}^*)} d\,F(\underline{Y}|\underline{\Theta})\; d\,F(\underline{\Theta}|\underline{Z}^*) = P.$$

(The fiducial approach will be given no further consideration.)

We turn now to the case when nothing is known about the marginal distribution of the sequence, $F(\underline{Y})$. Following the same reasoning as in the case of parametric prediction, the probability content of a prediction region D for given $F(\cdot)$ is:

$$P[D|F(\cdot)] = \int_D d\, F(\underline{Y}). \qquad \text{(II.3)}$$

In the frequentist approach the problem is to define a decision rule $D = D(\underline{Z})$ such that the probability content (expected value with respect to $\underline{Z}$) is P <u>for any given CDF $F(Y)$</u>:

$$E_{\underline{Z}}\left(P[D(\underline{Z})|F(\cdot)]\right) = \int_{\text{all } \underline{Z}} f[\underline{Z}|F(\cdot)] \int_{D(\underline{Z})} d\, F(\underline{Y})\, d\, \underline{Z} = P. \qquad \text{(II.4)}$$

When defined through equation (II.4), $D(\underline{Z})$ will be called a <u>distribution-free (or nonparametric) prediction region of P content</u>.

The same difficulties of parametric frequentist prediction are present also in nonparametric frequentist

prediction. Conceptually, the fiducial argument might be used also when $F(\underline{Y})$ is unknown, although additional complications are created by working in function space.

Bayesian approach

We consider only the case with known distribution type and unknown parameters. Indeed, little work has been done for the situation in which also the distribution type is unknown, even within a given distribution family. For methodological suggestions and some results see Smallwood (1968), Wood, et al. (1974) and Box and Tiao (1973), Chapter 3. The very similar problem of model selection is considered by Fedorov (1972) in the context of experiment design.

For a known distribution type, say $F_1(\underline{Y}|\underline{\Theta})$, let $f(\underline{\Theta})$ be the prior density function of the unknown parameters. Then the prediction distribution of $\underline{Y}$ is:

$$F(\underline{Y}|\underline{Z}) = \int_{\text{all } \underline{\Theta}} F_1(\underline{Y}|\underline{\Theta})\, f(\underline{\Theta}|\underline{Z})\, d\,\underline{\Theta}\,, \qquad \text{(II.5)}$$

where the conditional density $f(\underline{\Theta}|\underline{Z})$ is given by Bayes' theorem:

$$f(\underline{\theta}|\underline{z}) = f(\underline{\theta}) \frac{f(\underline{z}|\underline{\theta})}{f(\underline{z})} . \qquad \text{(II.6)}$$

The posterior predictive distribution of $\underline{Y}$, equation (II.5), contains all the necessary information for reliability analysis and design. In particular, prediction regions of given content are easily found.

For problems of Bayesian prediction in extended form, see Aitchison and Sculthorpe (1965) and Dunsmore (1966, 1968, 1969).

The general considerations above hold also for problems of simultaneous prediction. Two formulations can be given:

(a) Define $\underline{Y}$ to be the collection of the next $m > 1$ realizations, say:

$$\underline{Y} = [Y'_{n+1}, \ldots, Y'_{n+m}]' .$$

The associated prediction regions of P-content, $D(\underline{Z})$, are defined in $\underline{Y}$-space.

(b) For application to reliability problems it is more convenient to have $D(\underline{Z})$ defined in $\underline{Y}_i$-space; then P is the probability that m future realization points in $\underline{Y}_i$-space will all

fall inside $D(\underline{Z})$. Equations (II.1) to (II.4) are easily generalized to cover this definition of prediction region.

In Section II.3 consideration will be given to both definitions of $D(\underline{Z})$.

II.2 SIMPLE PREDICTION

In this section the "future" experiment consists of a single observation from a univariate or multivariate population. Prediction is made in terms of regions which contain the outcome of the experiment with a given probability. For the Bayesian approach, the predictive probability distribution is also derived.

II.2.1. Frequentist Prediction Intervals for Univariate Sequences

Consider a univariate white sequence $\{Y_i\}$ for which a sample of size n is available. The observation vector $\underline{Z}$ is defined:

$$\underline{Z} = \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{pmatrix} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}$$

and the prediction variable is $Y = Y_{n+1}$.

Since at least one distribution parameter is unknown, no prediction interval can be constructed before sampling. Parametric prediction intervals of P-expectation, see equation (II.2), are explicitly given below for several

independent probabilistic models: Normal, Exponential, Extreme type I, Poisson, Gamma, Lognormal. For the binomial case see Thatcher (1964). At the end of the paragraph the construction of nonparametric prediction intervals is also reviewed. In each case a measure of penalty for statistical uncertainty, r, is calculated. r is defined as the ratio between the length of the prediction interval under limited information and the length of the same interval when the distribution parameters are known (or the sample size becomes infinite).

(a) NORMAL SEQUENCES - $N(\mu; \sigma^2)$

μ and σ unknown. The problem of constructing a prediction interval of expected P content in the sense of equation (II.2) was solved by Wilks (1941). Let

$$(\hat{Z}, S^2) = \left(\frac{1}{n} \sum_{i=1}^{n} Z_i ; \frac{1}{n-1} \sum_{i=1}^{n} (Z_i - \hat{Z})^2 \right)$$

be a set of sufficient statistics for μ and σ^2. Wilks proved that the interval:

$$(\hat{Z} - \beta^N_{\mu,\sigma} \cdot S, \hat{Z} + \beta^N_{\mu,\sigma} \cdot S] \qquad \text{(II.7)}$$

is of P expectation in the sense of equation (II.2) if and only if

$$\beta^{N}_{\mu,\sigma} = (1 + 1/n)^{1/2}\, t_{n-1}[(1+P)/2] \qquad \text{(II.8)}$$

where $t_{\nu}(\gamma)$ is the γ-fractile of the "Student's" t-distribution with ν degrees of freedom. In other words, the variable $R = (Y-\hat{Z})/S(1+1/n)^{1/2}$ has t-distribution with (n-1) degrees of freedom. A number of similar results are derived in this chapter for normal populations with only μ or only σ unknown, as well as for different distribution types. In each case the upperscript on β refers to the population type (here, for instance, N stands for "normal") and the subscripts indicate the parameters assumed unknown (here both the location and the scale parameters).

Under perfect information (PI) on the population parameters μ and σ, the central prediction interval of P-content is:

$$(\mu - \Phi[(1+P)/2]\ \sigma,\ \mu + \Phi[(1+P)/2]\ \sigma], \qquad \text{(II.9)}$$

where $\Phi(\cdot)$ is the cumulative standard normal distribution. A measure of penalty for imperfect information is therefore

given by the ratio between $\beta^N_{\mu,\sigma}$ and $\beta^N_{PI} = [(1+P)/2]$, or:

$$r^N_{\mu,\sigma}(P,n) = \frac{\beta^N_{\mu,\sigma}(P,n)}{\beta^N_{PI}(P)} = (1 + 1/n)^{1/2} \frac{t_{n-1}[(1+P)/2]}{\Phi[(1+P)/2]} \qquad \text{(II.10)}$$

If $\hat{Z} = E[\hat{Z}] = \mu$, and $S^2 = E[S^2] = \sigma^2$, the quantity (II.10) has also the geometric meaning of length ratio between the central prediction intervals for limited and for perfect statistical information. For any fixed value of P, $r^N_{\mu,\sigma}(P,n)$ is a decreasing function of n, approaching 1 as $n \to \infty$.

For <u>one-sided prediction intervals</u> equations (II.7) are replaced by:

$$(\hat{Z} - \overline{\beta}^N_{\mu,\sigma} \cdot S, \infty), \text{ or by } (-\infty, \hat{Z} + \overline{\beta}^N_{\mu,\sigma} \cdot S] \qquad \text{(II.11a)}$$

and

$$\overline{\beta}^N_{\mu,\sigma} = (1 + 1/n)^{1/2} \cdot t_{n-1}(P). \qquad \text{(II.11b)}$$

Since for perfect information one-sided prediction intervals

have the form:

$$(\mu - \Phi(P)\ \sigma,\ \infty),\ \text{and}\ (-\infty,\ \mu + \Phi(P)\ \sigma],$$

the penalty ratio for finite sample sizes is, in accordance with equation (II.10):

$$\bar{r}^{N}_{\mu,\sigma}(P,n) = \frac{\bar{\beta}^{N}_{\mu,\sigma}(P,n)}{\Phi(P)} = (1 + 1/n)^{1/2}\ \frac{t_{n-1}(P)}{\Phi(P)}\ . \qquad \text{(II.12)}$$

μ unknown, σ known. This case (as well as the following one) was studied by Proshan (1953) and by Guttman (1955). When σ is known, sufficient statistics are n and the sample mean $\hat{Z} = \sum_{i=1}^{n} Z_i/n$, which define the central prediction interval of P-expectation:

$$[\hat{Z} - \beta^{N}_{\mu} \cdot \sigma,\ \hat{Z} + \beta^{N}_{\mu} \cdot \sigma], \qquad \text{(II.13a)}$$

where

$$\beta^{N}_{\mu} = (1 + 1/n)^{1/2} \cdot \Phi[(1+P)/2]. \qquad \text{(II.13b)}$$

The associated penalty for lack of information depends only on the sample size n, being

$$r_{\mu}^{N} = \frac{\beta_{\mu}^{N}(P,n)}{\Phi[(1+P)/2]} = (1 + 1/n)^{1/2}. \qquad \text{(II.14)}$$

By comparison with the ratio $r_{\mu,\sigma}^{N}$ in equation (II.10) one can judge the importance of the population variance being also unknown when μ is unknown. As expected, for all $P > 0$ and for all finite n it is:

$$r_{\mu,\sigma}^{N}(P,n) > r_{\mu}^{N}(n).$$

If one-sided prediction intervals are desired, equations (II.13) are replaced by

$$[\hat{z} - \bar{\beta}_{\mu}^{N} \cdot \sigma, \infty), \text{ or by } (-\infty, \hat{z} + \bar{\beta}_{\mu}^{N} \cdot \sigma]; \qquad \text{(II.15a)}$$

$$\bar{\beta}_{\mu}^{N} = (1 + 1/n)^{1/2} \Phi(P). \qquad \text{(II.15b)}$$

The penalty ratio is still given by equation (II.14).

μ known, σ unknown. In terms of the sufficient statistics n and $S^2 = \frac{1}{n} \sum_{i=1}^{n} (Z_i - \mu)^2$,

$$(\mu - \beta_\sigma^N \cdot S, \ \mu + \beta_\sigma^N \cdot S] \qquad \text{(II.16a)}$$

is a central prediction interval of P-expectation if

$$\beta_\sigma^N = t_n[(1+P)/2]. \qquad \text{(II.16b)}$$

This corresponds to the penalty ratio

$$r_\sigma^N(P,n) = \frac{\beta_\sigma^N(P,n)}{\Phi[(1+P)/2]} = \frac{t_n[(1+P)/2]}{\Phi[(1+P)/2]} . \qquad \text{(II.17)}$$

The one-sided prediction intervals of P-expectation are:

$$[\mu - \bar{\beta}_\sigma^N \cdot S, \ \infty) \text{ and } (-\infty, \ \mu + \bar{\beta}_\sigma^N \cdot S]; \qquad \text{(II.18a)}$$

$$\bar{\beta}_\sigma^N = t_n(P) \ , \qquad \text{(II.18b)}$$

with associated penalty ratio:

$$\bar{r}^{N}_{\sigma}(P,n) = t_{n}(P)/\Phi(P). \qquad \text{(II.19)}$$

The ratios $r^{N}_{\mu,\sigma}(P,n)$, $r^{N}_{\mu}(n)$ and $r^{N}_{\sigma}(P,n)$ in equations (II.10), (II.14) and (II.16) are related as:

$$r^{N}_{\mu,\sigma}(P,n) = r^{N}_{\mu}(n) \cdot r^{N}_{\sigma}(P,n-1). \qquad \text{(II.20)}$$

(the same relationship holds for one-sided predictions).

Tabulated values of the penalty ratios and of the coefficients $\beta^{N}_{(\cdot)}$ are collected in Appendix A (Tables A1 through A11). For selected values of P and for n ranging from 1 to 18 the penalty ratios are displayed in Figures 1, 2 and 3 (solid lines). The plots refer to central predictions. From the figures and from the more extensive tables in Appendix A it is seen that when σ is unknown and n is fixed the penalty increases with P. This would not be so if the β values of the prediction intervals were always related to normal CDF's (as opposed to cumulative t-distributions). The importance of the distribution type can

be measured by comparing the exact penalty ratios with the corresponding "first-order"(*) penalty ratios, which are defined here as those resulting from replacing the t - distribution by the normal distribution with the same first two moments, $N(0, \nu/(\nu-2))$, in equations (II.10), (II.12), (II.17, (II.19). Using the prime sign to denote "first order":

$$r^{N'}_{(\cdot)}(n) = \begin{cases} 1 & \text{for } \mu \text{ and } \sigma \text{ known,} \quad (a) \\ (1 + 1/n)^{1/2} & \text{for } \mu \text{ unknown,} \quad (b) \\ (1 + \frac{2}{n-2})^{1/2}, \ (n>2), & \text{for } \sigma \text{ unknown,} \quad (c) \\ [1 + \frac{3n-1}{n(n-3)}]^{1/2}, \ (n>3), & \text{for } \mu \text{ and } \sigma \text{ unknown.} \quad (d) \end{cases} \qquad (II.21)$$

For the case of only μ unknown, it is: $r^{N}_{\mu}(n) = r^{N'}_{\mu}(n)$. The expressions (II.21c) and (II.21d) are tabulated in Appendix A (Table A12) and plotted in Figures 1, 2, 3 (dashed lines). The "first-order" penalty ratios are sometimes larger, sometimes smaller than the exact ratios, but they are increasingly

---

(*) The present definition of "first-order" is not standard in the statistical literature.

unconservative as P gets larger (i.e., in the range of interest to safety and reliability theory).

This discloses a problem of "statistical invariance" in the sense of Ditlevsen (1973) and Veneziano (1974a), since in the presence of statistical uncertainty it is important whether one makes the assumption of the distribution type before or after sampling.

Other "first-order" solutions to this prediction problem are found in the literature. Recently Anderson (1972), working with a Bayesian formulation, concluded that the uncertainty in the variance has prediction effects which are negligible when compared to those from an uncertain mean. His results for the two extreme cases (1) of μ being known, and (2) of "diffuse" prior information on the mean yield the following penalty factors:

$$r(n) = \begin{cases} 1 & \text{for } \mu \text{ known,} \\ (1 + 1/n)^{1/2} & \text{for } \mu \text{ unknown,} \end{cases} \qquad \text{(II.22)}$$

irrespective of the state of knowledge about σ. In the frequentist approach the same approximation would lead to identical r values. The plots in Figures 1, 2 and 3 show that this approximation is quite unsatisfactory; in fact,

as we already pointed out, the variance being unknown is generally a much more critical condition for prediction than the mean value being unknown. Anderson's assumption has been used by Ang (1973) to derive probabilistic design formulas for simple systems in presence of statistical uncertainty (for additional considerations, see Veneziano (1974b)).

The effect of σ being unknown may also be considered from a different viewpoint. Suppose that the target of sampling is to reach a penalty ratio $r \leq 1.2$. For central prediction intervals the required minimum sample sizes are:

UNKNOWN PARAMETER(S)	Expected Content, P			
	0.75	0.90	0.99	0.999
μ and σ	7	9	15	21
σ	4	6	12	18
μ	3	3	3	3

TABLE 1. Minimum sample size for a penalty ratio $r \leq 1.2$ (equations (II.10), (II.14), (II.17)).

which values show again the importance of accounting correctly for the state of knowledge about σ.

(b) EXPONENTIAL SEQUENCE

Fraser and Guttman (1956) related the problem of constructing parametric prediction regions to a problem of hypothesis testing. The idea was applied later by Guttman (1959) to sampling from the exponential population with density function

$$f(Y) = \frac{1}{\sigma} e^{-Y/\sigma} \qquad \text{(II.23)}$$

and σ unknown. In terms of the statistic $\hat{Z} = \frac{1}{n} \sum_{i=1}^{n} Z_i$, Guttman found that the right-hand prediction interval of P-expectation for the next observation is:

$$[\hat{Z} \cdot [1 - \beta_{\sigma}^{EX}(P,n)], \infty) \ , \qquad \text{(II.24a)}$$

where $\beta_{\sigma}^{EX}(P,n)$ satisfies the equation

$$\int_{1-\beta_{\sigma}^{EX}(P,n)}^{\infty} [n/(n+w)]^{n+1} \, d\, w = P.$$

After integrating and solving explicitly for β_σ^{EX}:

$$\beta_\sigma^{EX}(P,n) = (n+1) - n\, P^{-1/n} \qquad \text{(II.24b)}$$

(see Table A13, part 1, in Appendix A).

In accordance with previous definitions for normal sequences, the penalty for imperfect information can be measured by the ratio

$$r_\sigma^{EX}(P,n) = \frac{\beta_\sigma^{EX}(P,n)}{1 + \log(P)} = \frac{(n+1) - n\, P^{-1/n}}{1 + \log(P)} \,. \qquad \text{(II.25a)}$$

However, since $\beta_\sigma^{EX}(P,n)$ is very close to 1 for the (large) values of P of interest to us, the inverse ratio of the complements to 1 is a more indicative measure of penalty; in this case:

$$r_\sigma^{EX}(P,n) = \frac{-\log P}{1 - \beta_\sigma^{EX}(P,n)} = \frac{\log (P)}{n(1-P^{-1/n})} \,. \qquad \text{(II.25b)}$$

The ratio (II.25b) is tabulated in Appendix A for right-hand prediction intervals (Table A14, part 1). Interestingly, for fixed P, $r_\sigma^{EX}(P,n)$ in equation (II.25b) is an increasing function of n, approaching 1 as $n \to \infty$. (The opposite trend was found for Normal sequences.) The minimum values for P = 0.90 and for P = 0.99 are $r_\sigma^{EX}(0.90,1)=0.949$ and $r_\sigma^{EX}(0.99,1)=1.000$, showing that little would be gained by collecting further information (according to the present definition of penalty).

Quite different is the case of left-hand prediction intervals of P-expectation:

$$[0,\ \hat{z}[1 + \beta_\sigma^{EX}(P,n)]] \ , \qquad \text{(II.26a)}$$

for which one finds:

$$\beta_\sigma^{EX}(P,n) = n(1-P)^{-1/n} - (n+1) \ . \qquad \text{(II.26b)}$$

The penalty ratio (II.25a) becomes:

$$r_\sigma^{EX}(P,n) = \frac{(n+1) - n(1-P)^{-1/n}}{1 + \log(1-P)} \ . \qquad \text{(II.27)}$$

Values of β_{σ}^{EX} and of r_{σ}^{EX} for left-hand predictions are given in Tables A13 and A14 (parts 2). In contrast with right-hand predictions the penalties are now quite large, often exceeding those for normal populations.

Guttman (1959, 1970) considered the analogous problem when sampling from the shifted exponential population

$$f(Y) = \frac{1}{\sigma} \exp \{-(Y-\mu)/\sigma\}$$

when μ or μ and σ are unknown.

The exponential distribution is widely used for modeling the time-to-failure of systems and subsystems (and more generally the interarrival time of "rare" events). Practical reasons impose limitations on the duration of a laboratory life-test, so that it is typical that one observes only the lower $r (r \leq n)$ times to failure (the lower r order statistics) from a sample of size n. Goodman and Madansky (1962) proved that for this situation the interval

$$[C(P,r) \cdot \hat{\lambda}^{-1}(n,r), \infty)$$

is of P-expectation if the statistic $\hat{\lambda}^{-1}(n,r)$ is defined as follows in terms of the order statistics $Z_{(1)}<Z_{(2)}<\cdots<Z_{(r)}$:

$$\lambda^{-1}(n,r) = r^{-1}\left[\sum_{i=1}^{r} Z_{(i)} + (n-r)\, Z_{(r)}\right],$$

and $$C(P,r) = r(P^{-1/r} - 1) .$$

By setting $\beta_{\sigma}^{EX}(P,n,r) = 1 - C(P,r)$ one verifies that this is a generalization of equations (II.26), which are reproduced in the special case: $r = n$. More will be said on prediction from censored data in Paragraph II.2.5.

(c) EXTREME TYPE I SEQUENCE

The extreme type I distribution is of frequent use in structural reliability since many potentially critical events are extremes of random processes within its domain of attraction. (Benjamin and Cornell (1971) pp 271-279; Gumbel (1958) pp236-254). The form of the distribution is:

$$F(Y) = \exp\{-\exp[-(Y-a)/b]\} , \; b > 0.$$

The parameters a and b can be estimated graphically or analytically by the maximum likelihood criterion. The latter approach, yielding the estimates $\hat{a}$ and $\hat{b}$, was pursued by Harter and Moore (1967, 1968). Following Antle and Rademaker (1972), the statistic

$$E = (Y_{n+1} - \hat{a})/\hat{b}$$

has a distribution which depends only on the number of past observations, n. Write the left-hand prediction interval of P-expectation in the form

$$(-\infty, \hat{a} + \beta^{El}_{a,b}(P,n) \cdot \hat{b}] \ . \qquad \text{(II.28)}$$

Representative values of $\beta^{El}_{a,b}(P,n)$ were computed by Antle and Rademaker through simulation. They tabulated $\beta^{El}_{a,b}$ for P = 0.90, 0.95, 0.975, 0.98, 0.99; n = 10(10)70, 100, ∞; and for P = 0.995 and n = 20(10)70, 100, ∞ (their Table 1). Some of these values are reproduced in Appendix A (Table A15). The maximum likelihood estimates of a and b are biased. After defining

$$E[(\hat{a}-a)/b] = g(n) \ ,$$

$$E[\hat{b}/b] = 1/h(n) \ ,$$

Antle and Rademaker found the values of $h(n)$ and $g(n)h(n)$ for $n = 10(10)100$ (their Table 2) via simulation. From these data one can calculate the penalty ratio:

$$r^{El}_{a,b}(P,n) = \frac{E[\hat{a}] + \beta^{El}_{a,b}(P,n) \cdot E[\hat{b}]}{a + \beta^{El}_{a,b}(P,\infty) \cdot b} = \frac{g(n) + \frac{\beta^{El}_{a,b}(P,n)}{h(n)} + \frac{a}{b}}{\frac{a}{b} + \beta^{El}_{a,b}(P,\infty)} \ . \qquad \text{(II.30)}$$

The ratio a/b depends only on the coefficient of variation V_Y, being (see for example Benjamin and Cornell (1971), p 274):

$$\frac{a}{b} \approx \frac{1.282}{V_Y} - 0.577.$$

Table A16 gives the values of $r^{El}_{a,b}(P,n)$ for $a/b = 1(1)10$, $n = 10, 20, 30$ and for selected probability contents, P. For each value of P and n the penalty ratio increases with the dispersion of the population distribution.

Unfortunately, Antle and Rademaker did not tabulate $\beta^{El}_{a,b}$ for sample sizes smaller than 10, which are frequent if the extreme sequence collects rare events.

(d) POISSON SEQUENCE

Let λ be the parameter of the Poisson process. The observation consists of a single Poisson variable Z with mean value $n\lambda$; the variable to be predicted, Y, has Poisson distribution with mean $m\lambda$. If Z and Y count the number of arrivals in a Poisson process, n and m are proportional to the duration of the past and of the future observation periods, t_p and t_f. The left-hand prediction interval of P-expectation is (see Aitchison and Sculthorpe (1965)):

$$[0, \beta^{P}_{\lambda}(P,n,m) \cdot S] \ , \qquad \text{(II.31)}$$

where $S^2 = Z\ m/n$, (II.32a)

$$\beta^{P}_{\lambda}(P,n,m) = [\min \{\beta : I_{n/(n+m)}\ (Z+1,\ \beta) \geq P\} - 1]/S \ , \qquad \text{(II.32b)}$$

and I is the incomplete beta function tabulated by Pearson (1934). Since $\beta^{P}_{\lambda}(\cdot)S$ is an integer, it was found more

convenient to tabulate the upper limit of the prediction interval, instead of $\beta_\lambda^P(\cdot)$. The parameters Z, $m/n = t_f/t_p$, and P in equations (II.31) and (II.32) were assumed as the independent variables and were given a set of representative values. The corresponding values of $\beta_\lambda^P \cdot S$ are collected in Table A17. For instance, if the ratio t_f/t_p is 0.5 and the prediction interval must have 0.99 probability content, it is $\beta_\lambda^P \cdot S = 4$ if $Z = 0$; $\beta_\lambda^P \cdot S = 5$ if $Z = 1$; $\beta_\lambda^P \cdot S = 6$ if $Z = 2$; $\beta_\lambda^P \cdot S = 9$ if $Z = 5$, and so on. It is interesting to note that even when $Z = 0$ the prediction interval may contain nonzero integers.

The penalty ratio is also best referred to the upper limit of the prediction interval; for $Z = E[Z] = n\lambda$:

$$r_\lambda^P(P,n\lambda,m\lambda) = \frac{\beta_\lambda^P(P,n,m) \cdot S}{\min\left\{Y : \sum_{i=0}^{Y} \frac{(m\lambda)^i e^{-m\lambda}}{i!} \geq P\right\}} ; \quad n\lambda > 0. \tag{II.33}$$

The denominator in equation (II.33) can be calculated from tables of the cumulative Poisson distribution or, indirectly, from inverse tables of the χ^2 distribution (Harter (1919)). It gives the upper limit of the prediction interval for perfect knowledge of λ. Table A18 collects the values of

the denominator in equation (II.33) for the same values of m λ and P considered in Table A17. The associated penalty ratios are given in Table A19. Since $r_{\lambda}^{P}(\cdot)$ is the ratio of two integers, only qualitative conclusions can be drawn from this table: for instance, the penalty tends to increase with the ratio t_f/t_p (which, in a sense, weights the uncertain future, t_f, versus the known past, t_p) and with P, although a number of "local" exceptions can be found in the table.

The right-hand interval of P-expectation was also found by Aitchison and Sculthorpe. Their result is:

$$[K,\infty), \quad \text{where } K = \max \{\gamma : I_{n/(n+m)}(Z,\gamma) \leq 1 - P\} . \tag{II.34}$$

(e) GAMMA SEQUENCE

The main reference is again the fundamental paper by Aitchison and Sculthorpe (1965). Let λ be the unknown parameter of the underlying exponential process (which process needs not have any physical meaning). The experiment consists of observing a variable Z with Gamma distribution $G(m, \lambda)$, where m is known. The problem is to find a prediction interval of P-expectation for a future observation

Y distributed like $G(K, \lambda)$, with K known. When the "past" experiment consists of observing n independent variables from the statistical population of Y, it is $m = n\,K$. In the sequel we assume that this is the case, so that m is a multiple of K (this implies no loss of generality for the results).

The mean and the standard deviation of Y are K/λ and $K^{1/2}/\lambda$. They are estimated by Z/n and by $Z/(n\,K^{1/2})$ respectively. Given Z, the left-hand prediction interval of P-expectation can be written:

$$[0, \frac{Z}{n} + \beta^{G}_{\lambda}(n,K,P)\ \frac{Z}{n\ K^{1/2}}]. \qquad \text{(II.35a)}$$

Using Aitchison and Sculthorpe's results, one finds the following expression for β^{G}_{λ}:

$$\beta^{G}_{\lambda}(n,K,P) = \left[\frac{n}{B(m,K;P)} - (n+1) \right] \cdot K^{1/2}\ , \qquad \text{(II.35b)}$$

where $B(m,K;P)$ is the P-fractile of the incomplete beta distribution. It satisfies $I_{B(m,K;P)}(m,K) = P$, and this last function is tabulated by Pearson (1934). Under perfect information the same interval would be:

$$\left[0,\ \frac{K}{\lambda} + \beta_{PI}^{G} \cdot \frac{K^{1/2}}{\lambda}\right] , \qquad \text{(II.36)}$$

where β_{PI}^{G} is given by:

$$\beta_{PI}^{G} = K^{-1/2}\left[\frac{1}{2}\,\beta_{\chi^2}(K,P) - K\right] \qquad \text{(II.37)}$$

and $\beta_{\chi^2}(\cdot)$ satisfies:

$$F_{\chi^2(2K)}\ [\beta_{\chi^2}(K,P)] = P \qquad \text{(II.38)}$$

Therefore $\beta_{\chi^2}(K,P)$ is obtained from the inverse cumulative χ^2 tables in Harter (1919).

The penalty ratio for imperfect information is

$$r_{\lambda}^{G}(n,K,P) = \frac{\beta_{\lambda}^{G}(n,K,P)}{\beta_{PI}^{G}(K,P)} = \frac{\dfrac{n}{B(m,K;P)} - (n+1)}{\dfrac{\beta_{\chi^2}(K,P)}{2K} - 1} . \qquad \text{(II.39)}$$

All the foregoing results hold also for the right-hand prediction interval of P-expectation:

$$\left[\frac{Z}{n} - \beta_{\lambda}^{G}(n,K,P) \frac{Z}{n K^{1/2}} , \infty \right) \qquad \text{(II.40)}$$

after replacing P by (1-P).

For several values of K, n and P equations (II.35b), (II.37) and (II.39) are tabulated in Appendix A (Tables A20, A21, A22 for left-hand predition intervals; Tables A23, A24, A25 for right-hand intervals). As K increases, the shape of the distribution passes gradually from an exponential (K = 1) to a normal ($K \to \infty$); at the same time the penalty ratios decrease for left-hand predictions and increase for right-hand predictions.

(f) LOGNORMAL SEQUENCE

Let Y belong to a lognormal population with mean μ and variance σ^2. By definition $Z = \ln Y$ has normal distribution with mean μ_Z and variance σ^2_Z. The relationships between the parameters of the two distributions are as follows:

$$\mu = \exp\left(\frac{1}{2}\sigma_Z^2 + \mu_Z\right) , \qquad \text{(II.41a)}$$

$$\sigma^2 = \exp(\sigma_Z^2 + 2\mu_Z)[\exp(\sigma_Z^2) - 1] , \qquad \text{(II.41b)}$$

and inversely:

$$\mu_Z = \ln(\mu^2) - \frac{1}{2}\ln(\sigma^2 + \mu^2) , \qquad \text{(II.42a)}$$

$$\sigma_Z^2 = -\ln(\mu^2) + \ln(\sigma^2 + \mu^2) . \qquad \text{(II.42b)}$$

Let $Z_1 = \ln Y_1$, $Z_2 = \ln Y_2$, ..., $Z_n = \ln Y_n$ be the logarithms of n independent observations from the lognormal sequence. When both μ and σ^2 are unknown μ_Z and σ_Z^2 are also unknown (see equations (II.42)). The latter quantities are estimated by $\hat{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i$ and by $S_Z^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Z_i - \hat{Z})^2$. One-sided prediction intervals for $Z = \ln Y$ are then (see results for the normal process):

$$(-\infty, \hat{Z} + \bar{\beta}^N_{\mu,\sigma}(P,n) \cdot S_Z] ,$$

and

$$[\hat{Z} - \bar{\beta}^N_{\mu,\sigma}(P,n) \cdot S_Z, \infty) .$$

Since the logarithmic transformation is continuous and one-to-one, one-sided prediction intervals of P-expectation for Y are found by transforming exponentially the limits of the corresponding intervals for Z. This yields:

$$[0,\ \exp\{\hat{Z} + \bar{\beta}^{N}_{\mu,\sigma}(P,n) \cdot S_Z\}] \ ,$$

and

$$[\exp\{\hat{Z} - \bar{\beta}^{N}_{\mu,\sigma}(P,n) \cdot S_Z\ ,\ \infty). \qquad \text{(II.43)}$$

These prediction intervals can be written in a more convenient form in terms of the following estimates of the mean and of the variance of Y:

$$\hat{Y} = \exp\left(\frac{1}{2} S_Z^2 + \hat{Z}\right) \ ,$$

$$S_Y^2 = \exp(S_Z^2 + 2\hat{Z})[\exp(S_Z^2) - 1] \ .$$

Then the intervals (II.43) are the same as

$$[0,\ \hat{Y} + \beta^{LN}_{\mu,\sigma}(P,n) \cdot S_Y] \quad \text{and} \quad [\hat{Y} - \beta^{LN}_{\mu,\sigma}(P,n) \cdot S_Y,\ \infty) \qquad \text{(II.44)}$$

if $\beta^{LN}_{\mu,\sigma}$ is chosen as indicated below. For left-hand predictions:

$$\beta^{LN}_{\mu,\sigma}(P,n) = \frac{\exp\{\hat{z} + \bar{\beta}^{N}_{\mu,\sigma}(P,n) \cdot S_z\} - \exp\{\frac{1}{2} S_z^2 + \hat{z}\}}{\exp\{\hat{z}\}[\exp\{S_z^2\}(\exp\{S_z^2\} - 1)]^{1/2}}$$

$$= \frac{\exp\{\bar{\beta}^{N}_{\mu,\sigma}(P,n) \cdot S_z\} - \exp\{S_z^2/2\}}{\exp\{S_z\}(\exp\{S_z^2\} - 1)^{1/2}} \qquad \text{(II.45)}$$

where $\bar{\beta}^{N}_{\mu,\sigma}(P,n) = (1 + 1/n)^{1/2}\, t_{n-1}(P)$ is tabulated in Appendix A (Table All).

For right-hand predictions:

$$\beta^{LN}_{\mu,\sigma}(P,n) = \frac{\exp\{S_z^2/2\} - \exp\{-\bar{\beta}^{N}_{\mu,\sigma}(P,n) \cdot S_z\}}{\exp\{S_z\}(\exp\{S_z^2\} - 1)^{1/2}} \cdot \qquad \text{(II.46)}$$

In both cases the penalty ratio $r^{LN}_{\mu,\sigma}(P,n)$ is computed from the following expression:

$$r^{LN}_{\mu,\sigma}(P,n) = \beta^{LN}_{\mu,\sigma}(P,n)/\beta^{LN}_{PI}(P) \ , \qquad \text{(II.47)}$$

where $$\beta_{PI}^{LN}(P) = \lim_{n\to\infty} \beta_{\mu,\sigma}^{LN}(P,n).$$

Both $\beta_{\mu,\sigma}^{LN}$ and $r_{\mu,\sigma}^{LN}$ depend on S_Z (besides varying with n and with P). The variance of the normal sequence $\{\ln Y_i\}$ is related directly to the coefficient of variation of the lognormal population V_Y being, as well known:

$$\sigma_{\ln Y} = [\ln(V_Y^2 + 1)]^{1/2}$$

After relating in a similar way the estimate $\hat{V}_Y$ of V_Y to the estimate S_Z of $\sigma_{\ln Y}$:

$$S_Z = [\ln(\hat{V}_Y^2 + 1)]^{1/2},$$

$\beta_{\mu,\sigma}^{LN}(\cdot)$ and $r_{\mu,\sigma}^{LN}(\cdot)$ depend on $\hat{V}_Y$, n and P. Tables A26 and A27 give values of $\beta_{\mu,\sigma}^{LN}(\hat{V}_Y,n,P)$ and of $r_{\mu,\sigma}^{LN}(\hat{V}_Y,n,P)$ for left-hand prediction intervals; Tables A28 and A29 collect the same quantities for right-hand prediction intervals. It is interesting to note that for large coefficients of variation (say $\hat{V}_Y > 0.3$) and for left-hand predictions the penalty ratios are substantially higher than those for normal populations with μ and σ unknown, due to the

logarithmic transformation.

If $\mu_Z = \mu_{\ln Y}$ or $\sigma_Z = \sigma_{\ln Y}$ are known, equations (II.45), (II.46) and (II.47) still hold for $\beta^{LN}_{\sigma_Z}(P,n)$ or for $\beta^{LN}_{\mu_Z}(P,n)$, and for the associated penalty ratios after replacing $\overline{\beta}^{N}_{\mu,\sigma}(P,n)$ by $\overline{\beta}^{N}_{\sigma}(P,n)$ or by $\overline{\beta}^{N}_{\mu}(P,n)$. Tabulated values for these cases are also given in Appendix A (Tables A30 through A37).

SOME APPLICATIONS TO RELIABILITY ANALYSIS AND DECISION

Some applications of the preceeding results (including Tables A1 - A37 in Appendix A) are suggested. For convenience of exposition and for comparing the present results with the "first-order" approximations (II.21) and with Anderson's approximation, we shall refer to normal populations (Tables A1 - A13). However, the same use can be made of the results for different populations.

The ratios in Tables C1 - C5 are penalty measures with respect to perfect information, decreasing with the sample size for fixed probability content P (for the values of P considered there). Let U_n be the expected utility of making n observations. U_n may be considered to be a function of n through the penalty ratio $r^{N}_{(\cdot)}(P,n)$, with a typical trend as shown in Figure 4. On the other

hand the cost of sampling, C_n, may evolve according to the straight solid line in the same figure. The optimal sample size is the (integer) value n* of n for which $(U_n - C_n)$ is maximum. If sample values are historical data the function C_n may look quite different, associating small costs with values of n less or equal to the present availability of data (n_p in Figure 4), and substantially higher costs with sample sizes larger than n_p (collection of additional data).

Another use of the penalty ratios is as multiplicative factors for design values which are valid under perfect parameters knowledge. For instance, if a design value for given distribution parameters is the upper 99% fractile of the normal distribution (i.e., $\mu + 2.326\ \sigma$), the penalty factors and the design values when one or both the location and the scale parameters are estimated from a sample of finite size n are shown in Table 2 for selected values of n.

The "first-order" solution does not provide finite design values when σ is unknown and n = 2, but for larger sample sizes is preferable to Anderson's approximation. When only μ is unknown the exact and the approximate design values coincide.

One can also use the tables of $r^N_{(\cdot)}(P,n)$ and $\beta^N_{(\cdot)}(P,n)$ for reliability analysis. Consider a structure whose state of survival depends only on the normal variable

UNKNOWN PARAMETER(S)	SAMPLE SIZE n	PENALTY RATIO	NEW DESIGN VALUE	"FIRST-ORDER" APPROXIMATION EQ(II.21)	"FIRST-ORDER"; VARIABILITY OF S NEGLECTED (Anderson(1972))
μ and σ	2	16.759	$\hat{Z} + 38.98\ S$	———	$\hat{Z} + 2.849\ S$
	5	1.766	$\hat{Z} + 4.107\ S$	$\hat{Z} + 3.600\ S$	$\hat{Z} + 2.549\ S$
	10	1.272	$\hat{Z} + 2.959\ S$	$\hat{Z} + 2.762\ S$	$\hat{Z} + 2.440\ S$
	20	1.119	$\hat{Z} + 2.602\ S$	$\hat{Z} + 2.515\ S$	$\hat{Z} + 2.384\ S$
σ	2	2.994	$\mu + 6.965\ S$	∞	$\mu + 2.326\ S$
	5	1.447	$\mu + 3.365\ S$	$\mu + 3.000\ S$	$\mu + 2.326\ S$
	10	1.188	$\mu + 2.764\ S$	$\mu + 2.640\ S$	$\mu + 2.326\ S$
	20	1.087	$\mu + 2.528\ S$	$\mu + 2.451\ S$	$\mu + 2.326\ S$
μ	2	1.225	$\hat{Z} + 2.849\ \sigma$	$\hat{Z} + 2.849\ \sigma$	$\hat{Z} + 2.849\ \sigma$
	5	1.096	$\hat{Z} + 2.549\ \sigma$	$\hat{Z} + 2.549\ \sigma$	$\hat{Z} + 2.549\ \sigma$
	10	1.049	$\hat{Z} + 2.440\ \sigma$	$\hat{Z} + 2.440\ \sigma$	$\hat{Z} + 2.440\ \sigma$
	20	1.025	$\hat{Z} + 2.384\ \sigma$	$\hat{Z} + 2.384\ \sigma$	$\hat{Z} + 2.384\ \sigma$

TABLE 2. Design values which include statistical uncertainty (see text).

Y. Let β_1 and β_2 denote the distances of the extremes of the system safe interval on the Y axis from the estimated or known mean of Y, in units of estimated or known standard deviations of Y. The probability of failure is:

$$P_f = 2 - Pr(\beta_1, n) - Pr(\beta_2, n) ,$$

where $Pr(\beta, n)$ is the probability content of the 1-tailed region corresponding to $\beta^N_{(\cdot)} = \beta$ for a sample of size n and for the appropriate combination of known - unknown parameters. Alternatively P_f can be calculated from tables of the standard normal CDF, being:

$$P_f = \Phi[-\beta_1/r^N_{(\cdot)}(P_1, n)] + \Phi[-\beta_2/r^N_{(\cdot)}(P_2, n)] ,$$

where P_1 and P_2 are defined implicitly:

$$P_i = \Phi[\beta_i/r^N_{(\cdot)}(P_i, n)] \; ; \; i = 1,2.$$

Suppose for example that the bending resistance of a beam population has normal distribution. The applied moment M is known, so that the safe interval on the resistance

axis is $[M,\infty)$. From the population, five beams are chosen at random and their resistances are measured. With the notations introduced earlier β_1 is the estimated value of $(\mu-M)/\sigma$ (μ and σ^2 are the mean and the variance of the resistance), and $\beta_2 = \infty$. The failure probability of a beam from the same population depends on β_1 and on the state of knowledge about the distribution parameters. A few values of P_f are collected in Table 3.

KNOWLEDGE OF PARAMETERS	β_1			
	1.00	2.00	3.00	4.00
μ,σ known	0.159	0.023	0.00135	0.000032
μ unknown	0.181	0.034	0.0031	0.00013
σ unknown	0.182	0.051	0.015	0.0052
μ,σ unknown	0.207	0.063	0.026	0.011

TABLE 3. Probability of failure for selected values of β_1, $\beta_2 = \infty$ and a sample of size n = 5 (see also text).

The values of P_f in Tables 4 and 5 correspond to the same values of β_1 and to the same combinations of known - unknown parameters. They result respectively from using the "first-order" approximation (II.21) and from neglecting

the uncertainty of S in the same approximation (i.e., when using the approximation proposed by Anderson (1972)). The "first-order" approximations (II.21) (Table 4) are conservative for small β_1 values and become unconservative as β_1 increases. If, in addition, one neglects the statistical variability of S, the estimated reliability increases, this making the approximations in Table 5 always unconservative if σ is unknown.

KNOWLEDGE OF PARAMETERS	β_1			
	1.00	2.00	3.00	4.00
μ unknown	0.181	0.034	0.0031	0.00013
σ unknown	0.219	0.061	0.0101	0.0010
μ,σ unknown	0.260	0.099	0.026	0.0051

TABLE 4. Probabilities of failure from the "first-order" approximation (II.21).

KNOWLEDGE OF PARAMETERS	β_1			
	1.00	2.00	3.00	4.00
μ unknown	0.181	0.034	0.0031	0.00013
σ unknown	0.159	0.023	0.00135	0.000032
μ,σ unknown	0.181	0.034	0.0031	0.00013

TABLE 5. Probabilities of failure; "first-order" approximation and variability of S neglected (Anderson (1972)).

As anticipated before, the coefficients β and the penalty ratios for non-normal populations can be used in the same way.

It is interesting to compare the penalty ratios $r = \beta/\beta_{PI}$ when a sample of the same size n is available from different populations. In Figure 5 the penalty ratios for several distribution types and for various combinations of known - unknown parameters are plotted as functions of n. The curves refer to a left-hand prediction interval of 0.999-expectation and indicate clearly that the penalty ratio is highly sensitive to the population type and to the state of knowledge about the parameters. Curves (3), (6), (7) and (8) refer to cases with both location and scale

parameters unknown; only the location parameter is unknown for curves (1), (4) and (5); curve (3) is for a case with unknown variance. For all the remaining curves the location and the scale parameters are both unknown, but they are related deterministically (for instance through a known coefficient of variation).

Although some overlapping exists and the results are rather dispersed, the general tendency is that the penalty increases from the location parameter being unknown to the scale parameter being also unknown but with a deterministic functional relationship between the two, to the case with only the scale parameter unknown, and finally to both parameters being unknown and independent.

Unfortunately, it does not seem that a distribution-free approximation for r as a function of n can be found easily without restricting substantially the distribution type. For left prediction intervals and for exponentially decaying densities like, say, $\exp\{Y^{\alpha}\}$ with α given, the penalty ratio increases with decreasing α (for example it is larger for the exponential distribution than it is for the normal). Of course, comparisons are meaningful only if the state of knowledge about the parameters is similar.

(g) DISTRIBUTION-FREE PREDICTION INTERVALS

So far the distribution type and possibly some distribution parameters were assumed known (=> parametric prediction). Under some conditions on the sample size, a prediction interval for the next observation can be constructed also if no information is available on the population distribution; i.e., neither on the parameters, nor on the shape. In this type of statistical (nonparametric) prediction, inferences rely exclusively on the information from the sample. The only distributional property which is usually required (and to which the following results are conditional) is the continuity of the population CDF.

Given an ordered set of n independent observations from the same unknown population: $Z_{(1)} < Z_{(2)} < \ldots < Z_{(n)}$, consider the (n+1) intervals: $(-\infty, Z_{(1)}]$, $(Z_{(1)}, Z_{(2)}]$,..., $(Z_{(n)}, \infty)$. After setting $Z_{(o)} = -\infty$ and $Z_{(n+1)} = \infty$, the difference $[F(Z_{(j)}) - F(Z_{(j-1)})]$, with $F(\cdot)$ being the unknown population CDF, is called the coverage of the j^{th} interval. Tukey (1947) proved that the sum of any $r < n+1$ coverages has Beta distribution $I(r, n-r+1)$, where

$$I_t(r, n-r+1) = \int_0^t \frac{n!}{(r-1)!(n-r)!} \alpha^{r-1} (1-\alpha)^{n-r} d\alpha, \quad 0 \leq t \leq 1 \qquad \text{(II.48)}$$

is tabulated by Pearson (1934). In particular the expected sum of any r coverages is

$$P(r,n) = r/(n+1) \qquad \text{(II.49)}$$

(Less general statements were proved earlier by Wilks (1941) and by Wald (1943).) Equation (II.49) enables one to construct distribution-free prediction intervals of P-expectation <u>if</u> enough data are available. For instance, $(Z_{(1)}, Z_{(n)}]$ is a central nonparametric prediction interval of expectation $(n-1)/(n+1) = 1 - \frac{2}{n+1}$, and $(Z_{(1)}, \infty)$ is a right-hand prediction interval of expectation $(1 - \frac{1}{n+1})$.

Suppose that sampling is from a <u>Normal population</u>, but that the experimenter has a poor knowledge of the distribution type; thus he decides to use a distribution-free predictor. Keeping n, the sample size, to a minimum, a few prediction intervals of P-expectation are considered in Table 6 for different values of P. The entries of the table are the values of n which make the nonparametric intervals in the left column of P-expectation.

PREDICTION INTERVAL	P					
	0.75	0.90	0.95	0.99	0.995	0.999
$(Z_{(1)}, Z_{(n)}]$	7	19	39	199	399	1999
$(-\infty, Z_{(n)}]$ or $(Z_{(1)}, \infty)$	3	9	19	99	199	999

TABLE 6. Minimum sample sizes for the construction of nonparametric prediction intervals.

The quantity $[Z_{(n)}-Z_{(1)}]/\sigma$, where σ^2 is the population variance, is called the "normalized range" of the sample. For normal populations the expected normalized range (ENR) was computed and tabulated by Tippett (1925) for n = 2(1)1000. For instance it is: ENR(39) = 4.30117 and ENR(199) = 5.48876. The ratio

$$r_F^N(P,n) = \frac{ENR(n)}{2\Phi[(1+P)/2]} \qquad \text{(II.50a)}$$

for central predictions and the ratio

$$r_F^N(P,n) = \frac{ENR(n)}{2\Phi(P)} \qquad \text{(II.50b)}$$

for 1-sided predictions are measures of the expected penalty for not knowing the distribution shape when in fact the distribution is normal. For the prediction intervals and for the sample sizes in Table 6, r_F^N has the following values.

PREDICTION INTERVAL	P					
	0.75	0.90	0.95	0.99	0.995	0.999
CENTRAL	1.176	1.121	1.097	1.065	1.057	-
1-SIDED	1.256	1.158	1.121	1.077	1.065	1.049

TABLE 7. Penalty factors for nonparametric prediction intervals when sampling is from normal populations.

One might plot the points $[n, r_F^N(P,n)]$ in Figures 1, 2, 3 to compare with the analogous penalties when the distribution type is known but some or all the parameters are not; however the major penalty in nonparametric prediction is the very large required sample size (for central prediction intervals: $n \geq (1+P)/(1-P)$; for 1-sided predictions: $n \geq P/(1-P)$).

Suppose now that the population has <u>exponential distribution</u> and that the purpose of sampling is to construct

a 1-sided nonparametric prediction interval of P-expectation. If the distribution type is unknown to the experimenter, the minimum required sample sizes are those given in Table 6 for 1-sided predictions.

The mean of the smallest value among n independent observations with exponential density: $f(Z) = \lambda \exp(-\lambda Z)$ is (see Gumbel (1958), p 117):

$$\overline{Z}_{(1)} = 1/\lambda\, n \qquad \text{(II.51a)}$$

and the mean of the largest value is, again from Gumbel, p 116:

$$\overline{Z}_{(n)} = \sum_{i=1}^{n} 1/\lambda \cdot i \qquad \text{(II.51b)}$$

$$\approx \ln n + 0.57722 \text{ (large n)}$$

Since the P-fractile of the exponential distribution is $-\lambda^{-1} \ln(1-P)$, the penalty for not knowing the distribution type can be defined:

(II.52a)

$$r_F^{EX}(P,n) = -n \ln(P)$$

for prediction intervals of the type $(Z_{(1)},\infty)$; and

$$r_F^{EX}(P,n) = \frac{1 - \sum_{i=1}^{n} \frac{1}{i}}{1 + \ln(1-P)}$$

for prediction intervals of the type $[0,Z_{(n)}]$

(II.52b)

For n having the values in Table 6 for 1-sided predictions, the penalty ratios are given in Table 8.

PREDICTION INTERVAL	P					
	0.75	0.90	0.95	0.99	0.995	0.999
$(Z_{(1)},\infty)$	0.863	0.948	0.975	0.995	0.9975	0.9995
$[0,Z_{(n)}]$	2.157	1.404	1.304	1.157	1.133	1.098
SAMPLE SIZE	3	9	19	99	199	999

TABLE 8. Penalty factors for distribution-free prediction intervals when sampling is from exponential distributions.

The penalty ratios for right-hand predictions are always smaller than the corresponding ratios when the distribution type is known, and for the same sample size. This is due to the particular shape of the left "tail" of the exponential density. Conversely, the penalty ratios for left-predictions always exceed the corresponding parametric values. For the latter case (as for normal populations) the lack of information on the distribution type is penalized mostly by the large minimum sample size which is required to construct prediction intervals of high expected content ($n \geq P/(1-P)$).

The "paradox" of having penalty ratios smaller than 1 and of them decreasing when assuming no knowledge of the distribution type motivates some skepticism on the validity of the ratio r as an index of penalty for the "pathological" case of the exponential distribution (see also comments on this point in Goodman and Madansky (1962)).

Along the same lines one can define penalty ratios for nonparametric prediction when the actual distribution is Gamma, Extreme type I, lognormal or other, if the expected range is known. Useful references are Gupta (1960), David (1954), Cox (1954), Singh (1967, 1972).

II.2.2 Bayesian Prediction for Univariate Sequences

Not until recently has statistical prediction profited from Bayes' theorem. Important contributions are authored by Aitchison (1964), Thatcher (1964), Lindley (1965), Aitchison and Sculthorpe (1965), Guttman (1970).

The Bayesian approach to prediction was outlined briefly in Section II.1. The quantity to be predicted, Y, is considered to be random, with uncertainty contributed both by the probabilistic model and by the lack of statistical information. If the population distribution type is known but not the vector of parameters $\underline{\Theta}$, a prior distribution for $\underline{\Theta}$, $f(\underline{\Theta})$, is established which quantifies the status of knowledge before sampling. This prior knowledge is then combined with the information from the sample data, $\underline{Z}$, yielding a posterior distribution for $\underline{\Theta}$, $f(\underline{\Theta}|\underline{Z})$. This is accomplished formally through equation (II.6), which is rewritten here in terms of the likelihood function $l(\underline{Z}|\underline{\Theta})$ (for any given observation vector, $l(\underline{Z}|\underline{\Theta})$ is proportional to $f(\underline{Z}|\underline{\Theta})$):

$$f(\underline{\Theta}|\underline{Z}) \propto f(\underline{\Theta}) \cdot l(\underline{Z}|\underline{\Theta}) \qquad \text{(II.53)}$$

The posterior predictive density of Y is then calculated through equation (II.5). The same equation gives also the prior predictive density f(Y) after replacing $f(\underline{\theta}|\underline{Z})$ by $f(\underline{\theta})$. If desired, point and interval predictions can be found a posteriori after specifying a prediction criterion or a utility function (see Chapter I, Section 2).

The power of Bayesian analysis resides mainly in its capability to combine information of different types from different sources. However, closed-form results for the posterior predictive distribution are known only for particular likelihood functions (probabilistic models) and prior distributions. In this sense, the case of vague initial information ("diffuse" or "noninformative" prior) and the case when the posterior distribution is of the same type as the prior distribution (conjugate prior) are of special interest.

The problem of selecting "correctly" the prior distribution is not free of difficulties, and is still the object of active research and of much debate. Savage (1962), Jeffreys (1961) and Novick (1962) among others made early and fundamental contributions in this area; conjugate prior distributions were introduced by Raiffa and Schlaifer (1961) and discussed by Lindley (1965) and by Draper and Guttman (1968a,b). Recent reviews with vast bibliographies are in Lindley (1971) and in Hampton, et al. (1973). Also

informative is Chapter 6 in DeGroot (1970).

The results presented in this chapter are mostly for conjugate prior distributions. Particular relevance is given to the fact that the cases with conjugate priors can be reduced to problems with diffuse priors by artificially increasing the sample size and by modifying the sufficient sample statistics. The characterization of the prior information through an equivalent sample may have useful practical implications, as suggested by the experiments conducted by Winkler (1967) and others. A concise account of Bayesian prediction theory for Normal, Exponential, Poisson and Gamma populations follows.

(a) NORMAL SEQUENCES

μ unknown, σ known. From Raiffa and Schlaifer (1961), p 55, the likelihood function in terms of the sufficient statistics $\hat{z} = \frac{1}{n} \sum_{i=1}^{n} z_i$ and n is:

$$l(\mu|\hat{z},n) \propto \exp\{-\frac{n}{2\sigma^2}(\mu-\hat{z})^2\} . \qquad \text{(II.54)}$$

For a prior in the form of the normal density

$$f(\mu) = \frac{1}{\sqrt{2\pi}\,\sigma_o} \exp\{-\frac{1}{2\sigma_o^2}(\mu-\mu_o)^2\} \qquad \text{(II.55)}$$

the posterior density of μ is also normal. Let $\sigma_o^2 = \sigma^2/n'$; then

$$f(\mu|\hat{Z},n) = \frac{1}{\sqrt{2\pi}\,\sigma/(n+n')^{1/2}} \exp\{-\frac{n+n'}{2\sigma^2}[\mu - \frac{n'\mu_o + n\hat{Z}}{n+n'}]^2\} .$$

(II.56)

From equation (II.5), after "integrating out" μ:

$$f(Y|\underline{Z}) = \frac{1}{\sqrt{2\pi}\,\sigma\,(\frac{n+n'+1}{n+n'})^{1/2}} \cdot \exp\left\{-\frac{n+n'}{2(n+n'+1)\sigma^2}\left[Y-\frac{n'\mu_o+n\hat{Z}}{n+n'}\right]^2\right\}.$$

(II.57)

This is again in the form of a normal density with parameters:

$$\mu_Y = \frac{n'\mu_o + n\hat{Z}}{n+n'} \; ; \; \sigma_Y^2 = \sigma^2 \frac{n+n'+1}{n+n'} . \qquad \text{(II.58)}$$

The prior density (II.55) becomes more and more "diffuse" as σ_o increases or, equivalently, as n' decreases. The noninformative situation is approached asymptotically as $n' \to 0$. Although (II.55) degenerates into an improper density when $n' \to 0$, the predictive density (II.57) does not, becoming simply: $N(\hat{Z};\ \sigma^2(n+1)/n)$.

It is interesting to observe that for integer n'

$$N\left(\frac{n'\mu_o + n\hat{Z}}{n + n'} ;\quad \sigma^2 \frac{n + n' + 1}{n + n'} \right)$$

would be the posterior predictive distribution if, starting with a diffuse prior, a sample of size (n+n') were available, with sample mean $(n'\mu_o + n\hat{Z})/(n + n')$. This shows, as anticipated earlier, that the effect of assuming an informative conjugate prior with integer n' is to "increase the sample size" and to modify the sufficient statistics. When n' is not an integer the prior information "is worth more" than [n'] observations and less than [n']+1 observations with sample mean μ_o. ($[\cdot]$ = max integer function.)

If one is interested in prediction intervals of probability content P, the following applies.

Central prediction interval:

$$\left[\frac{n'\mu_o + n\hat{Z}}{n+n'} - \Phi[(1+P)/2] \cdot \sigma\left(\frac{n+n'+1}{n+n'} \right)^{1/2} , \right.$$

$$\left. \frac{n'\mu_o + n\hat{Z}}{n+n'} + \Phi[(1+P)/2] \cdot \sigma\left(\frac{n+n'+1}{n+n'} \right)^{1/2} \right]. \qquad \text{(II.59)}$$

One-sided prediction intervals:

$$\left[-\infty, \frac{n'\mu_o + n\hat{Z}}{n+n'} + \Phi(P)\ \sigma\left(\frac{n+n'+1}{n+n'} \right)^{1/2} \right] ; \text{ or} \qquad \text{(II.60a)}$$

$$\left[\frac{n'\mu_o + n\hat{Z}}{n+n'} - \Phi(P)\ \sigma\left(\frac{n+n'+1}{n+n'} \right)^{1/2} , \infty \right] \qquad \text{(II.60b)}$$

These intervals coincide with the frequentist prediction intervals of P-expectation when the prior is diffuse

($n' = 0$), and coincide with the frequentist results for a sample size $(n + n')$ and a sample mean $\frac{n'\mu_o + n\hat{Z}}{n + n'}$ when the prior is informative of the conjugate type and n' is integer.

This means that one can use the tables in Appendix A also for Bayesian prediction, by entering them with a sample size $(n + n')$. Interpolation should be used for non-integer n'.

<u>μ known, σ unknown</u>. Again from Raiffa and Schlaifer (1961) p 54, the likelihood function of $h = 1/\sigma^2$ is:

$$l(h|S,n) \propto h^{n/2} \exp\{-\tfrac{1}{2} n h S^2\}$$

where n is the sample size and $S^2 = \frac{1}{n} \sum_{i=1}^{n} (Z_i-\mu)^2$ is the sample variance. The family of conjugate distributions is Gamma-2:

$$f(h) \propto h^{(\frac{1}{2} n'-1)} \exp\{-\tfrac{1}{2} h n' S'^2\}; \quad h > 0; \quad n', S'^2 > 0,$$

for which the posterior density of h results:

$$f(h|S,n) \propto h^{(\frac{1}{2} n''-1)} \exp\{-\tfrac{1}{2} h n'' S''^2\} , \qquad \text{(II.61)}$$

where $n'' = n + n'$,

$$S''^2 = (n\,S^2 + n'\,S'^2)/(n + n').$$

Since $f(Y|h) \propto h^{1/2} \exp\{-\frac{1}{2} h(Y-\mu)^2\}$, the predictive posterior density is:

$$f(Y|S,n) = \int_0^\infty f(Y|h)\; f(h|S,n)\; d\,h$$

$$\propto \int_0^\infty h^{1/2} \exp\{-\tfrac{1}{2} h(Y-\mu)^2\} \cdot h^{(\frac{1}{2} n''-1)}$$

$$\cdot \exp\{-\tfrac{1}{2} h\, n''\, S''^2\}\; d\,h.$$

The integrand is proportional to a joint Normal-Gamma density for Y and h. After integrating with respect to h one finds that $(Y - \mu)/S''$ has a "Student's" t-distribution with n" degrees of freedom; i.e.:

$$f(Y|S,n) \propto [n'' + (\tfrac{Y-\mu}{S''})^2]^{-(n''+1)/2} \qquad \text{(II.62)}$$

The considerations following equation (II.58) hold also in the present case. In fact, as the prior density becomes less informative ($n' \to 0$, S' finite), a posteriori $(Y-\mu)/S$ becomes t-distributed with n degrees of freedom, and the Bayesian prediction intervals of P-content coincide with the frequentist prediction intervals of P-expectation. The frequentist results can be used also when the conjugate prior is informative, in which case one should consider an "equivalent" sample size $n'' = n + n'$ and an "equivalent" sample variance $S''^2 = (n S^2 + n' S'^2)/(n + n')$. In other words, the prior information "is worth" n' additional sample data with sample variance S'^2.

The notion of an informative conjugate prior being equivalent to a diffuse prior plus a set of sample statistics (here n', μ_o) when only μ is unknown, (n', S'^2) when only σ is unknown) is recurrent in Bayesian prediction. It may also prove relevant to ease the application of Bayesian ideas to engineering decisions; after all, it is not unusual that prior knowledge derives from previous sampling. In a sense, μ_o and S'^2 represent <u>what</u> one believes prior to sampling, while n' measures one's <u>degree of belief</u> in those values (how many data points one would be willing to exchange for his prior information). In addition, this interpretation of the prior knowledge (when applicable)

makes clear the consequences of misestimating the prior distribution within the conjugate family. This point is elaborated further in what follows.

Consider first the case when the population mean is known but the variance is not, and let n'_o and S'^2_o be the "correct" prior parameters. If one choses the correct value for n', but the incorrect value $S'^2 = S'^2_o + (\Delta S')^2$ for the equivalent prior sample variance, the posterior variance of Y is, from equation (II.62):

$$\sigma_Y^2 = S''^2 \frac{n + n'_o}{n + n'_o - 2} = \sigma_{Y_o}^2 + \frac{n'_o}{n + n'_o - 2} (\Delta S')^2 , \tag{II.63}$$

where $\sigma_{Y_o}^2$ is the value of σ_Y^2 for $S'^2 = S'^2_o$. Therefore an error in the equivalent prior sample variance is transferred into the posterior predictive variance with a weighting factor $n'_o/(n + n'_o - 2)$. The quantity

$$\gamma = \left[1 + \frac{n'_o}{n + n'_o + 2} \left(\frac{\Delta S'}{\sigma_{Y_o}} \right)^2 \right]^{1/2} \tag{II.64}$$

is the factor by which the extremes of any prediction interval are multiplied because of an error $(\Delta S')^2$ in the estimation of S'^2. For instance, for $\Delta S'/\sigma_{Y_o} = 0.8$ and $n = n'_o = 5$ (implying $S' = 1.8\ S'^2_o$ if $S'_o = S$) one finds $\gamma = 1.183$.

In the same way, when assuming the "correct" value for S'^2, but a "wrong" value for n', say $n' = n'_o + \Delta n'$, one finds:

$$\sigma_Y^2 = \frac{nS^2 + (n'_o + \Delta n')\ S'^2_o}{n + n'_o + \Delta n' - 2} \quad ,$$

and

$$\gamma = \left[\frac{nS^2 + (n'_o + \Delta n')\ S'^2_o}{nS^2 + n'_o\ S'^2_o} \frac{n + n'_o - 2}{n + n'_o + \Delta n' - 2} \right]^{1/2} . \quad \text{(II.65)}$$

Similar results are easily derived also for the case: μ unknown, σ known.

μ and σ unknown. This case was studied by Aitchison and Sculthorpe (1965) and is reported in Guttman (1970). With respect to the likelihood function:

$$l(\mu,\sigma^2|\underline{Z}) = (2\pi)^{-n/2}\ \sigma^{-n}\ \exp\{-[(n-1)S^2 + n(\hat{Z}-\mu)^2]/2\sigma^2\}$$

where $\hat{Z} = n^{-1} \sum_{i=1}^{n} Z_i$; $S^2 = (n-1)^{-1} \sum_{i=1}^{n} (Z_i-\hat{Z})^2$,

the conjugate family is of the Normal-Gamma type (see Raiffa and Schlaifer (1961) p 55). If the prior parameters are $[\mu',n',(n'-1)S'^2]$, the joint prior density of μ and σ^2 is:

$$f(\mu,\sigma^2) \propto \sigma^{-(n'+2)}\ \exp\{-[(n'-1)S'^2 + n'(\mu-\mu')^2]/2\sigma^2\}. \quad \text{(II.66)}$$

As n' and S'^2 tend to zero, this density becomes more and more "flat", approaching the noninformative form (in the sense of Jeffreys (1961)):

$$f(\mu,\sigma^2) \propto 1/\sigma^2\ . \quad \text{(II.67)}$$

The posterior distribution of (μ,σ^2) is also of the Normal-Gamma type, with parameters:

$$n'' = n + n';$$

$$\mu'' = n''^{-1}(n'\mu' + n\hat{Z});$$

$$q = S''^2(n''-1) = (n'-1)S'^2 + (n-1)S^2 + n'n(n'')^{-1}(\hat{Z}-\mu')^2$$

$$= [(n'-1)S'^2 + n'\mu'^2] + [(n-1)S^2 + n\hat{Z}^2]$$

$$- (n+n')\mu''^2 .$$

From equation (II.5) the posterior density of the next observation from the same normal population is

$$f(Y|\underline{Z}) = \left(\frac{n''}{n''+1} \right)^{1/2} \frac{\Gamma\frac{n''}{2}}{\Gamma\left(\frac{n''-1}{2} \right)(\pi q)^{1/2}} \left[1 + \frac{n''(Y-\mu'')^2}{(n''+1)\ q} \right]^{-n''/2} \tag{II.69}$$

which implies that $\left(\frac{n"}{n"+1} \right)^{1/2} \frac{Y-\mu"}{[q/(n"-1)]^{1/2}}$ has t-distribution with (n"-1) degrees of freedom.

Prediction intervals are readily obtained:

Central prediction interval of P-content:

$$(\mu" - K, \mu" + K], \tag{II.70a}$$

where $$K = \left(\frac{q(n"+1)}{n"(n"-1)} \right)^{1/2} \cdot t_{n"-1}[(1+P)/2]. \tag{II.70b}$$

One-sided prediction intervals of P-content:

$$(-\infty, \mu" + K']; \text{ or } (\mu" - K', \infty), \tag{II.71a}$$

where $$K' = \left(\frac{q(n"+1)}{n"(n"-1)} \right)^{1/2} \cdot t_{n"-1}(P), \tag{II.71b}$$

and $t_\nu(P)$ denotes the P-fractile of the t-distribution with ν degrees of freedom.

In the case of noninformative prior, equation (II.67), these intervals reduce to the corresponding frequentist intervals, equations (II.7) and (II.11). More generally the prediction intervals for informative conjugate priors correspond to the frequentist intervals when using an equivalent sample size n", a sample mean μ" and a sample variance S" = q/(n"-1). This parallels analogous findings for the cases when μ or σ are known. Again, one can use the tables in Appendix A.

(b) EXPONENTIAL SEQUENCE

When sampling is from the exponential distribution (II.23) the likelihood function is

$$l(\sigma|\underline{z}) = \sigma^{-n} \exp\{-n\,\hat{z}/\sigma\} ,$$

where $\hat{z} = n^{-1} \sum_{i=1}^{n} z_i$. The associated conjugate prior family is (see Guttman (1970), p 130):

$$f(\sigma) \propto \sigma^{-(n'+1)} \exp\{-n'\mu'/\sigma\}$$

which approaches the diffuse density $f(\sigma) \propto 1/\sigma$ as n' and

μ' tend to zero. From equation (II.53) the posterior density has the form:

$$f(\sigma|\underline{Z}) = \frac{(n\hat{Z} + n'\mu')^{n+n'}}{\Gamma(n + n')} \sigma^{-(n+n'+1)} \exp\{-(n\hat{Z} + n'\mu')/\sigma\} ,$$

meaning that $2(n\hat{Z} + n'\mu')/\sigma$ is distributed like Chi-square with $2(n + n')$ degrees of freedom. From equation (II.5) the predictive density is:

$$f(Y|\underline{Z}) = (n + n')[1 + Y/(n\hat{Z} + n'\mu')]^{-(n+n'+1)} \qquad \text{(II.72)}$$

<u>One-sided prediction intervals of P-content</u> are:

for right predictions:

$$(K, \infty), \text{ where } K = (n\hat{Z} + n'\mu')(\underline{P}^{-1/(n+n')} - 1); \qquad \text{(II.73)}$$

for left predictions:

$$[0, K], \text{ where } K = (n\hat{Z} + n'\mu')[(1-P)^{-1/(n+n')} - 1]. \qquad \text{(II.74)}$$

Like in sampling from normal populations, Bayesian prediction intervals of given probability content for a diffuse prior coincide with the corresponding frequentist intervals, equations (II.24) and (II.26). For an informative conjugate prior the frequentist results (and the tables in Appendix A) hold after replacing n and $\hat{Z}$ by the "equivalent" sufficient statistics (n + n') and $(n\hat{Z} + n'\mu')/(n+n')$.

The following results for Poisson and for Gamma sequences are due to Aitchison and Sculthorpe (1965).

(c) POISSON SEQUENCE

A Poisson sequence generates independent, identically distributed exponential variates Z_i with density function $\lambda e^{-\lambda Z}$; $Z \geq 0$, $\lambda > 0$. The likelihood function for λ in terms of the statistics n and $\hat{Z} = \frac{1}{n} \sum_{i=1}^{n} Z_i$ is:

$$l(\lambda|\underline{Z}) = e^{-\lambda n\hat{Z}} \cdot \lambda^n \qquad \text{(II.75a)}$$

with Gamma conjugate family (see Raiffa and Schlaifer (1961), p 53). For the prior density

$$f(\lambda) \propto e^{-\lambda n'\mu'} \lambda^{n'-1} ; \quad \lambda \geq 0; \quad n',\mu' > 0. \qquad \text{(II.75b)}$$

the associated posterior results in the form:

$$f(\lambda|\underline{Z}) = \frac{n'\mu' + n\hat{Z}}{\Gamma(n+n')} \exp\{-\lambda(n'\mu' + n\hat{Z})\}[(n'\mu' + n\hat{Z})\lambda]^{(n+n'-1)},$$

$$\lambda \geq 0 . \qquad \text{(II.75c)}$$

From equation (II.5) the density of the prediction variable Y having Poisson distribution with mean m λ is:

$$f(Y|\underline{Z}) = \frac{\Gamma(n+n'+Y)}{Y!\Gamma(n+n')} \left(\frac{m}{n'\mu'+n\hat{Z}+m}\right)^{Y} \left(\frac{n'\mu'+n\hat{Z}}{n'\mu'+n\hat{Z}+m}\right)^{n+n'} . \qquad \text{(II.76)}$$

Therefore the <u>left-hand prediction interval</u> of P-probability content is

$$[0, K], \text{ where } K = \min\{\gamma: I_{b/(b+m)}(n+n', \gamma) \geq P\} - 1. \qquad \text{(II.77)}$$

In equation (II.77), $b = n'\mu' + n\hat{Z}$, and I is the incomplete beta function (tables in Pearson (1934)).

The corresponding <u>right-hand prediction interval</u> is:

$$[K, \infty), \text{ where } K = \max\{\gamma : I_{b/(b+m)}\ (n+n', \gamma) \leq 1-P\}. \qquad \text{(II.78)}$$

When the prior is diffuse ($n' \to 0$, μ' finite) the prediction interval (II.78) reproduces the frequentist result, equation (II.34). (Replace $(n'\mu'+n\hat{Z})$ by $(n'+n)$ for comparing the results.) On the contrary the interval (II.77) does <u>not</u> reproduce equations (II.31) and (II.32) unless for left-hand predictions one assumes $n'\mu' = 1$ as a noninformative condition and $n'\mu' > 1$ as an informative one. This makes good sense; in fact for $n\hat{Z} = 0$ and for a diffuse prior the prediction interval [0, K] in equation (II.77) should have a non-zero upper limit for high probability contents (see also earlier comments on the frequentist results).

Under this condition on left-hand predictions the tables in Appendix A can be used also for Bayesian predictions.

(d) GAMMA SEQUENCE

The results for Gamma sequences are closely related to those for the Poisson process. If λ is the parameter of the underlying exponential process, the "past" experiment E consists in observing a variable Z with distribution $G(m, \lambda)$ and m known. λ is a Bayesian random variable. The

"future" experiment F concerns the realization of a variable with distribution $G(K,\lambda)$ and known K. Viewing E as a set $\underline{Z}$ of n replicates of F it is: $m = n\,K$. The likelihood function has the form of equation (II.75a) with n K in place of n. Consequently the conjugate family is Gamma as in equation (II.75b) where, for consistency of notations, one must replace n' by n'K'. With the same substitutions the posterior density of λ is given by equation (II.75c). Finally, the posterior predictive distribution has the form:

$$f(Y|\underline{Z}) = \frac{b^a}{B(a,K)} \frac{Y^{K-1}}{(b+Y)^{a-1}} \qquad \text{(II.79)}$$

where $b = nK\hat{Z} + n'K'\mu'$; $a = nK + n'K'$; and $B(p,q)$ is the complete beta function:

$$B(p,q) = \Gamma(p)\Gamma(q)/\Gamma(p+q).$$

From equation (II.79) the following one-sided prediction intervals of P-content result:

$$[0,\ K],\ \text{where } K = b\,\frac{1 - B(a,K;1-P)}{B(a,K;1-P)} \qquad \text{(II.80)}$$

(B(a,K;P) is the P-fractile of the incomplete beta distribution); and

$$[K,\infty), \text{ where } K = b\,\frac{1 - B(a,K;P)}{B(a,K;P)} \quad . \tag{II.81}$$

When the prior distribution becomes diffuse ($n' \to 0$; K',μ' finite) b approaches $n\hat{KZ}$ and a approaches nK, which coincide with what was denoted Z and, respectively, K in the frequentist approach (equations (II.35a), (II.35b), (II.40)).

Also for Gamma sequences the Bayesian intervals coincide with the frequentist intervals after replacing the sample statistics, here $(nK, n\hat{KZ})$, by "equivalent" sample statistics, here $(nK+n'K', n\hat{KZ}+n'K'\mu')$. As a consequence, with these substitutions one can use the tables for Gamma populations in Appendix A.

SOME REMARKS AND EXEMPLIFIED APPLICATIONS

As we saw for a few population distributions, when the conjugate prior degenerates to a noninformative density the Bayesian prediction intervals of P-content coincide with the corresponding frequentist intervals. We also noted

that in the (conjugate) informative case one can still use the frequentist results, referred to an "equivalent" set of sufficient statistics. More exactly we found that the information contained in the conjugate prior is the same as that contained in an "equivalent" prior sample, which Bayes' theorem "pools together" with the actual sample. From a practical viewpoint, since the pooled statistics are sufficient, tables for frequentist prediction and for frequentist penalty ratios can be used also for Bayesian analysis. Two examples follow.

EXAMPLE 1. FREQUENTIST VERSUS BAYESIAN PREDICTION

Consider sampling from the normal population $N(\mu,\sigma^2)$ with both parameters unknown. Five samples are chosen at random ($n = 5$), with statistics $\hat{Z} = 10$ and $S^2 = 9$. The prior information is judged to be worth four additional samples ($n' = 4$), with statistics $\mu' = 14$ and $S'^2 = 16$. The pooled statistics are, from equations (II.68):

$$n'' = n + n' = 9;$$

$$\mu'' = (n\hat{Z} + n'\mu')/(n + n') = 11.778;$$

$$S''^2 = q/(n'' - 1) = 14.9.$$

If a central prediction interval is desired, with content $P = 0.99$, the frequentist answer is (use Table A8 for $P = 0.99$ and $n = 5$):

$$(\hat{Z} - 5.044\ S,\ \hat{Z} + 5.044\ S];\quad \text{or}\quad (-5.13,\ 25.13].$$

The Bayesian answer is (enter the same table with $n = 9$):

$$(\mu'' - 3.537\ S'',\ \mu'' + 3.537\ S'];\ \text{or}\quad (-1.86,\ 25.42].$$

From the Bayesian viewpoint one can also construct a prediction interval of 0.99-content based exclusively on prior knowledge. In this case, using Table A8 with $n = 4$ one finds:

$$(\mu' - 6.53\ S',\ \mu' + 6.53\ S'];\ \text{or}\quad (-11.12,\ 39.12].$$

Note that, although the frequentist prediction interval usually contains the Bayesian prediction interval when the same sample is used for both, this is not a general rule since Bayesian predictions depend also on the prior statistics. It should also be said that if the sample

statistics are "very different" from the equivalent prior statistics one might want to check the experimental data and/or to revise his prior beliefs. (This problem is discussed by Savage (1972); see also Veneziano (1974c).)

EXAMPLE 2. COMBINED USE OF BAYESIAN ESTIMATION AND PREDICTION

Reinforced concrete structures (parallel examples can be made for timber and steel structures) often use precast beams. The properties of the basic materials, the geometry of the beams and the position of the reinforcement are controlled during production. Nevertheless, even an accurate control cannot remove all the statistical variability, so that the load-carrying capacity of each individual beam is a random variable.

In order to reduce the uncertainty of the product, a "here-and-now" predictive strategy is actuated at the plant level, where the scope of sampling and of making destructive tests is to evaluate the distribution parameters of the resistance population with "minimum" uncertainty. Later on, a "wait-and-see" (estimative) strategy is followed by the consumer who tests nondestructively the beams delivered to him (for instance by measuring the beams' stiffnesses and by relating them to their ultimate resistances.

The consumer's purpose is to reduce the uncertainty on the resistance of the beams he is going to use.

Therefore, the posterior predictive distribution of the producer becomes prior information for the consumer. At the final stage of construction all the tests (both those performed by the producer and those performed by the consumer) contribute to reduce the uncertainty on the resistance of the beams.

Two problems arise: (i) how to combine the information from different experiments; and (ii) what is the "best" overall testing policy. We restrict these questions to the case of the producer and the consumer controlling only the number of experiments, which are performed with fixed modalities.

Let us focus on a single beam with unknown resistance, R. The initial distribution of R (before any testing) is assumed to be $N(\mu;\sigma^2)$ with σ^2 given and μ unknown. After the producer has measured the resistances $R_i (i=1,\ldots,n_1)$ of n_1 beams chosen at random from the statistical population, the (predictive) distribution of R is:

$$(R|R_1,\ldots,R_{n_1}) \sim N(\hat{R};(1+1/n_1)\sigma^2) , \qquad (II.82)$$

where $\hat{R} = \frac{1}{n_1} \sum_{i=1}^{n_1} R_i$. Correspondingly the right-hand prediction interval of P-content is:

$$(\hat{R} - \beta_{\mu}^{N}(P,n_1) \cdot \sigma, \infty) \ , \qquad \text{(II.83a)}$$

where $$\beta_{\mu}^{N}(P,n_1) = (1 + 1/n_1)^{1/2} \ \Phi(P) \qquad \text{(II.83b)}$$

Suppose now that the consumer's measurements have the form:

$$\underline{Z} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} R + \underline{\varepsilon}, \qquad \text{(II.84)}$$

where $\underline{Z}$ and $\underline{\varepsilon}$ are n_2-vectors, and

$$\underline{\varepsilon} \sim N_{n_2}(\underline{0};\underline{\Theta}), \text{ where } \underline{\Theta} = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{pmatrix} \qquad \text{(II.85)}$$

After testing, the right-hand interval of P-content becomes:

$$(\hat{R}^* - \beta_\mu^{N^*}(n_1,n_2,P) \cdot \sigma, \infty) \ , \qquad \text{(II.86a)}$$

where, from Chapter I, equations (I.8a):

$$\beta_\mu^{N^*}(n_1,n_2,P) = \beta_\mu^{N}(n_1,P)\left[1 + \frac{\sigma^2(1+1/n_1)}{\sigma_\varepsilon^2} \ \frac{n_2}{1 + (n_2-1)\rho}\right]^{-1/2}$$

$$= \Phi(P)\left[\frac{n_1}{n_1 + 1} + \frac{\sigma^2}{\sigma_\varepsilon^2} \ \frac{n_2}{1 + (n_2-1)\rho}\right]^{-1/2} . \qquad \text{(II.86b)}$$

Since a priori it is $E[\hat{R}] = E[\hat{R}^*] = \mu$ = prior mean resistance, the key quantity for deciding upon n_1 and n_2 is:

$$h = \frac{1}{\Phi(P)} \beta_\mu^{N^*}(n_1,n_2,P) = \left[\frac{n_1}{n_1 + 1} + \frac{\sigma^2}{\sigma_\varepsilon^2} \ \frac{n_2}{1 + (n_2-1)\rho}\right]^{-1/2} . \qquad \text{(II.87)}$$

The smaller h, the smaller the posterior variance. h is a decreasing function of n_1 and n_2, with limiting value

$$\lim_{n_1, n_2 \to \infty} h = h_\infty = \left(1 + \frac{\sigma^2}{\sigma_\varepsilon^2} \frac{1}{\rho}\right)^{-1/2}. \qquad \text{(II.88)}$$

Different testing strategies (n_1, n_2) can be compared through the ratio

$$\frac{h}{h_\infty} = \left(\frac{1 + \frac{1}{\rho} \frac{\sigma^2}{\sigma_\varepsilon^2}}{\frac{n_1}{n_1 + 1} + \frac{\sigma^2}{\sigma_\varepsilon^2} \frac{n_2}{1 + (n_2 - 1)\rho}} \right)^{1/2} \qquad \text{(II.89)}$$

Assume for instance that $\sigma^2/\sigma_\varepsilon^2 = 4$ and $\rho = 0.4$; also assume a target ratio $h/h_\infty \leq 1.1$. The solutions which are candidates for optimality are collected in the following table.

n_1	n_2
0	15
1	10
2	9
3	8
6	7

Any other combination is either unfeasible (for instance (n_1,n_2) with $n_2 < 7$), or suboptimal (for instance (n_1,n_2) with $n_2 > 15$). If testing costs are given together with the number of beams (the consumer's tests have to be repeated on each beam, while the producer's tests do not), one can chose the best strategy by direct comparison of the costs.

(e) DISTRIBUTION-FREE PREDICTION INTERVALS

The difficulty of applying Bayesian ideas to non-parametric prediction is to describe appropriately the prior knowledge over a family of distribution functions. The family may comprise all the possible distributions over a given sample space, y, or a subset of them (like the subset of the continuous distributions). In part the difficulty can be overcome by considering a partition $(B_1,\ldots,B_K)$ of y

such that $B_i \cap B_j = \phi$ for $i \neq j$ and $\bigcup_{i=1}^{K} B_i = y$. Then all the distributions over y can be generated by specifying the probability contents $P(B_1),\ldots,P(B_K)$ <u>for all possible partitions</u>, under the restriction $P(B_i) \geq 0$; $\sum_{i=1}^{K} P(B_i) = 1$. If one fixes the number of subsets K and the partition $(B_1,\ldots,B_K)$ of y, the joint distribution of the $P(B_j)$'s is not sufficient information to characterize the probability distribution of P(A), where A is any subset of y.

Let us consider this restricted and simpler case first. In the univariate case y coincides with the real line, or with a subset of it. The conjugate prior for the discrete model in which only the K exhaustive and mutually exclusive events: $Y \in B_1,\ldots,Y \in B_K$, are considered, has Dirichlet form with parameters, say, $\alpha_1,\ldots,\alpha_K$. (This means that $P(B_1),\ldots P(B_K)$ have the joint Dirichlet distribution $D(\alpha_1,\ldots,\alpha_K)$.) Let α_i (i=1,...,K) be the prior parameters and consider an observation $Z \in y$ which belongs to the i^{th} set B_i. The posterior distribution of $P(B_1),\ldots,P(B_K)$ is again jointly Dirichlet (and marginally Beta) with parameters $\alpha_1^{(i)},\ldots,\alpha_K^{(i)}$, where:

$$\alpha_j^{(i)} = \alpha_j \qquad \text{for } i \neq j,$$

$$\alpha_j^{(i)} = \alpha_j + 1 \text{ for } i = j.$$

This Bayesian property is easily generalized to n observations and holds for any finite dimension of y. The sum $\alpha = \sum_{i=1}^{K} \alpha_i$ is a measure of the confidence one has in the prior best guess of the distribution. In fact, when the parameters α_i are integer-valued, α lends itself to the interesting interpretation of an equivalent prior sample size, with α_1 observations located (anywhere) in B_1, α_2 observations anywhere in B_2, and so on. The first two moments of $P(B_i)$ are (see for instance Johnson and Kotz (1972), Chapter 40 for a review of the properties of the Dirichlet distribution):

$$E[P(B_i)] = \alpha_i/\alpha \quad ,$$

$$E[P(B_i)^2] = \frac{\alpha_i(\alpha_i + 1)}{\alpha(\alpha + 1)} \quad . \qquad (i = 1,\ldots,K)$$

This discrete approach has been extended by Ferguson (1973) to cover continuous distributions over y. Ferguson starts with the following definition of the Dirichlet process:

"Let α be a non-null finite measure (non-negative and finitely additive) on $(y, \mathcal{A})$[(*)]. We say P is a Dirichlet process on $(y, \mathcal{A})$ with parameter α if for every $K = 1,2,\ldots$ and measurable partition $(B_1,\ldots,B_K)$ of y the distribution of $(P(B_1),\ldots,P(B_K))$ is Dirichlet, $D(\alpha(B_1),\ldots,\alpha(B_K))$."

Therefore, the prior information (which for the Dirichlet process turns out to be in conjugate form) may be defined through the joint distribution of $P(B_1),\ldots,P(B_K)$ for all K and for all measureable partitions $(B_1,\ldots,B_K)$ of y. In the Dirichlet process this is done by defining the parameter α as a set-to-scalar function. The expected probability content of any set $B \subset y$ is

$$E[P(B)] = \frac{\alpha(B)}{\alpha(y)} \quad \frac{\alpha(B)}{\sum_{i=1}^{K} \alpha(B_i)} .$$

(*) $\mathcal{A}$ is a σ-field of subsets of y.

As for the discrete Dirichlet distribution, $\alpha(y)$ measures the degree of confidence in the prior best estimate of the probability distribution (in units of number of observations), which is:

$$F(Y) \propto \alpha((-\infty, Y]).$$

A number of properties of the Direchlet process are proved by Ferguson. In particular his Theorem 1 states that the Bayesian property of the Direchlet distribution carries over to the Dirichlet process as follows. Let $P(B)$ denote the probability content of the generic set $B \subset y$ for the true distribution. If the prior Dirichlet process on $(y, \mathscr{A})$ has parameter α_o, and $Z_1, Z_2, \ldots, Z_n$ is a population sample, then the posterior distribution of $P(\cdot)$ is a Dirichlet process with parameter $\alpha_o + \sum_{i=1}^{n} \delta_{Z_i}$, where δ_Z denotes a measure on $(y, \mathscr{A})$ giving mass 1 to the point Z; i.e.:

$$\delta_Z(A) = 1 \quad \text{if } Z \in A,$$

$$= 0 \quad \text{if } Z \notin A.$$

Given the posterior parameter $\alpha_o + \sum_{i=1}^{n} \delta_{Z_i}$, the posterior best estimate of the distribution function can be obtained for any given optimality criterion. For instance, when the process is defined over the real axis R, Ferguson shows that the best estimate of the distribution function under a quadratic loss criterion is the expected distribution function:

$$F(Y) = \frac{\alpha_o(R)}{\alpha_o(R) + n} F_o(Y) + \left[1 - \frac{\alpha_o(R)}{\alpha_o(R) + n}\right] F_n(Y|Z_1,\ldots,Z_n), \tag{II.90}$$

where

$$F_o(Y) = \frac{\alpha_o((-\infty,Y])}{\alpha_o(R)}$$

is the best prior estimate, and

$$F_n(Y|Z_1,\ldots,Z_n) = \frac{1}{n} \sum_{i=1}^{n} \delta_{Z_i}((-\infty,Y])$$

is the "empirical CDF" of the sample. The posterior degree of belief in the best estimate is measured by $\alpha_o(R) + n$, showing again that $\alpha_o(R)$ can be interpreted as an equivalent

prior sample size.

Analogous results hold when y is a set in a higher dimensional space.

Intervals (or regions in the multivariate analog) of expected cover P are easily constructed if the parameter α of the Dirichlet process is known. In fact, for every interval $A \subset y$ it is: $E[P(A)] = \alpha(A)/\alpha(y)$, where $\alpha(y)$ is a normalization constant. For $y = R$ prediction intervals of P-content are:

Central intervals:

$$[d_1, d_2] ,$$

where $$d_1 = \min\{d:F(d) \geq (1-P)/2\}$$

$$d_2 = \min\{d:F(d) \geq (1+P)/2\}$$

One-sided intervals:

$$(-\infty, d_2] ,$$

where $d_2 = \min\{d:F(d) \geq P\}$; and

$$[d_1, \infty) ,$$

where $d_1 = \min\{d:F(d) \geq 1-P\}$.

In all cases it is:

$$F(d) = \frac{1}{\alpha_o(R)+n} \left(\alpha_o((-\infty,d]) + \sum_{i=1}^{n} Z_i((-\infty,d])\right) .$$

As $\alpha_o(R) \to 0$ the prior information becomes more and more "diffuse", and the Bayesian prediction intervals approach the corresponding frequentist intervals (see Paragraph II.2.1(g)). When discussing nonparametric frequentist prediction intervals we observed that a large penalty had to be paid in terms of the mimimum required sample size. If the Bayesian prior distribution is discrete with a probability mass $1/\alpha_o(R)$ concentrated at each of $\alpha_o(R)$ points ($\alpha_o(R)$ integer), the required minimum sample size for the construction of prediction regions of P-expectation is reduced by $\alpha_o(R)$ with respect to the frequentist requirements (see

Table 6).

EXAMPLE: CONSTRUCTION OF A NONPARAMETRIC BAYESIAN CENTRAL PREDICTION INTERVAL OF 0.90 - EXPECTATION

Two types of prior distributions are considered and specified through different prior parameters α_o of a Dirichlet process: a concentrated-probability-mass distribution as in Figure 6(a), and a distributed-probability distribution as in Figure 6(b). The prior (b) can be viewed as a smoothed version of the prior (a). In both cases we assume $\alpha_o(R) = 14$.

In order to construct a central interval of exactly 0.90 - expectation it must be $\alpha_o(R) + n \geq 20$, so that the sample size n must be at least 6. An hypothetical measure α_{SAMPLE} from a sample of size 6 is shown in Figure 6(c). The parameter $\alpha_o + \alpha_{SAMPLE}$ of the posterior Dirichlet process is sketched in Figure 6(d) for the prior (a) and in Figure 6(e) for the prior (b). In both cases the prior information is "weighted" in the posterior information by the factor $\alpha_o(R)/[\alpha_o(R)+n] = 14/20 = 0.7$, and the empirical (sample) information by the factor $n/[\alpha_o(R)+n] = 0.3$. The central prediction intervals of 0.90 - expectation are indicated on the same figures.

II.2.3. Frequentist Prediction Intervals for Multivariate Sequences

Parametric results for multivariate prediction are generally restricted to multinormal populations. Nevertheless, under the condition of independence, rectangular prediction regions for non-normal vectors are easily obtained from the univariate results. For instance, suppose that the prediction vector $\underline{Y}$ has two independent components Y_1 and Y_2. Y_1 has exponential distribution $EX(\lambda)$, with λ unknown, and Y_2 has lognormal distribution $LN(\mu,\sigma^2)$ with both parameters unknown. Then the region

$$D = \{(Y_1,Y_2) \mid Y_1 \in I^{EX}(n,P_1);\ Y_2 \in I^{LN}(n,P_2)\}$$

($I^{EX}(n,P_1)$ is a prediction interval of expectation P_1 for Y_1, and $I^{LN}(n,P_2)$ is a prediction interval of expectation P_2 for Y_2, given n independent observations from the two populations) is of P-expectation if and only if $P_1P_2 = P$. In fact the events $E_1: y_1 \in I^{EX}(n,P_1)$ and $E_2: Y_2 \in I^{LN}(n,P_2)$ are independent and

$$P = Pr(E_1 \cap E_2) = Pr(E_1) \cdot Pr(E_2) = P_1P_2 \quad .$$

The idea is generalized in the obvious way to $\underline{Y}$ having any finite number of components.

For the bivariate case one can also construct rectangular probability charts. One such chart is shown in Figure 7 (lower half) for Y_1 being $EX(\lambda)$ with λ unknown and Y_2 being $N(\mu;\sigma^2)$ with only μ unknown. The sample size is $n = 5$. Each non-dotted rectangle has expected probability content 0.001 and each dotted rectangle has expected content 0.003, so that any selection of m undotted and $\overline{m}$ dotted rectangles has the property of $0.001(m + 3\overline{m})$-expectation. Rectangular probability charts of this type depend on the sample size (which needs not be the same for Y_1 and Y_2), on the distribution types, and on the parameters being known or unknown.

The chart in the upper half of Figure 7 corresponds to perfect knowledge of the distribution parameters ($n \to 0$). Since little penalty is associated with the normal population when only the location parameter is unknown, the rectangles of the lower part of Figure 7 are elongated mainly in the direction of the exponential variate, particularly for large values of Y_1.

In each quadrant the unbounded regions have an expected probability content 0.006 (the equivalent of six undotted rectangles). Considering also the specular images

of the external contours of the upper and lower charts, the ratio between the areas with expected content (1-2x0.006) = 0.988 for the cases n = 5 and n → ∞ is 1.93, which is a two-dimensional penalty measure for limited statistical information at the probability level 0.988.

More complete results are given now for multinormal populations.

(a) MULTINORMAL SEQUENCES

Let $\underline{Y} \sim N_K(\underline{\mu};\underline{\Sigma})$ be a K-dimensional normal prediction vector with mean $\underline{\mu}$ and covariance matrix $\underline{\Sigma}$. If $\underline{\mu}$ and $\underline{\Sigma}$ are known, the maximum-probability-density (and therefore minimum-volume) central prediction region of P-content is

$$D^N(P) = \{\underline{Y} \mid (\underline{Y}-\underline{\mu})' \ \underline{\Sigma}^{-1} \ (\underline{Y}-\underline{\mu}) \leq \chi^2_K(P)\}, \qquad (II.91)$$

where $\chi^2_K(P)$ denotes the P-fractile of the Chi-square distribution with K degrees of freedom. $D^N(P)$ as defined by equation (II.91) is a K-dimensional ellipsoid, centered in $\underline{\mu}$. The points which satisfy the condition (II.91) as an equality belong to a contour surface of the normal density of $\underline{Y}$.

The construction of similar central prediction

regions of P-expectation when either $\underline{\mu}$, or $\underline{\Sigma}$, or both $\underline{\mu}$ and $\underline{\Sigma}$ are unknown was first studied by Fraser and Guttman (1956) (see also the review paper by Chew (1966)). For the derivation of the results collected here the reader is referred to the original paper.

$\underline{\mu}$ and $\underline{\Sigma}$ unknown. Let $\underline{Z}_1 = \underline{Y}_1, \ldots, \underline{Z}_n = \underline{Y}_n$ be a sample of size n from the population of $\underline{Y}$, with sufficient statistics:

$$n;\quad \hat{\underline{Z}} = n^{-1} \sum_{i=1}^{n} \underline{Z}_i;\quad \underline{S} = (n-1)^{-1} \sum_{i=1}^{n} (\underline{Z}_i - \hat{\underline{Z}})(\underline{Z}_i - \hat{\underline{Z}})' \ .$$

A prediction region of P-expectation is:

$$D^N_{\underline{\mu},\underline{\Sigma}}(P) = \{\underline{Y} \mid (\underline{Y}-\hat{\underline{Z}})' \underline{S}^{-1} (\underline{Y}-\hat{\underline{Z}}) \leq (\beta^N_{\underline{\mu},\underline{\Sigma}})^2\} \quad , \qquad \text{(II.92)}$$

where

$$(\beta^N_{\underline{\mu},\underline{\Sigma}})^2 = (n-1) \frac{K}{n-K} (1 + 1/n) F_{K,n-K}(P) \ ,$$

$K \geq 1$ = dimension of $\underline{Y}$,

$n > K$ = sample size, and

$F_{K,n-K}(P)$ = P-fractile of the F-distribution with K and (n-K) degrees of freedom.

When K = 1, equation (II.92) coincides with the prediction interval (II.7) since

$$\beta^N_{\mu,\sigma}\cdot(1 + 1/n)^{-1/2} = t_{n-1}[(1+P)/2] = [F_{1,n-1}(P)]^{1/2} .$$

The coefficient $\beta^N_{\underline{\mu},\underline{\Sigma}}(n,K,P)$ controls the size of the predictive region (the volume of $D^N_{\underline{\mu},\underline{\Sigma}}(P)$ is proportional to $(\beta^N_{\underline{\mu},\underline{\Sigma}})^K$). Under perfect information its value is, from equation (II.91):

$$\beta^N_{PI}(K,P) = \chi_K(P) .$$

Paralleling the univariate case, a penalty ratio for imperfect information can be defined as follows:

$$r^N_{\underline{\mu},\underline{\Sigma}}(n,K,P) = \frac{\beta^N_{\underline{\mu},\underline{\Sigma}}(n,K,P)}{\beta^N_{PI}(K,P)} = \frac{\left(\frac{n^2-1}{n}\ \frac{K}{n-K}\ F_{K,n-K}(P)\right)^{1/2}}{\chi_K(P)} . \quad \text{(II.93)}$$

$\underline{\mu}$ known, $\underline{\Sigma}$ unknown. The central prediction ellipsoid of P-expectation results:

$$D_{\underline{\Sigma}}^{N}(P) = \{\underline{Y} | (\underline{Y}-\underline{\mu})' \underline{S}^{-1} (\underline{Y}-\underline{\mu}) \leq (\beta_{\underline{\Sigma}}^{N})^2\} , \tag{II.94}$$

where $$(\beta_{\underline{\Sigma}}^{N})^2 = \frac{n\ K}{n - K + 1} F_{K,n-K+1}(P) ; (n > K-1)$$

$$\underline{S} = n^{-1} \sum_{i=1}^{n} (\underline{Z}_i - \underline{\mu})(\underline{Z}_i - \underline{\mu})' .$$

The associated penalty ratio is:

$$r_{\underline{\Sigma}}^{N}(n,K,P) = \frac{\beta_{\underline{\Sigma}}^{N}(n,K,P)}{\beta_{PI}^{N}(K,P)} = \frac{\left(\frac{n\ K}{n-K+1} F_{K,n-K+1}(p)\right)^{1/2}}{\chi_K(P)} \tag{II.95}$$

$\underline{\mu}$ unknown, $\underline{\Sigma}$ known. The central prediction region of expected P-content is now:

$$D_{\underline{\mu}}^{N}(P) = \{\underline{Y} | (\underline{Y}-\hat{\underline{Z}})' \underline{\Sigma}^{-1} (\underline{Y}-\hat{\underline{Z}}) \leq (\beta_{\underline{\mu}}^{N})^2\} , \tag{II.96}$$

where $$(\beta^N_{\underline{\mu}})^2 = (1 + 1/n)\ \chi^2_K(P) \ .$$

Interestingly enough, the penalty ratio when only $\underline{\mu}$ is unknown is independent of K and P, and therefore coincides with the univariate value:

$$r^N_{\underline{\mu}}(n) = (1 + 1/n)^{1/2} \ . \qquad \text{(II.97)}$$

The penalty ratios (II.93), (II.95), (II.97) are related as

$$r^N_{\underline{\mu},\underline{\Sigma}}(n,K,P) = r^N_{\underline{\mu}}(n) \cdot r^N_{\underline{\Sigma}}(n-1,K,P) \ ,$$

which equation generalizes the univariate result, equation (II.20).

Tables of $\beta^N_{(\cdot)}$ and $r^N_{(\cdot)}$ for selected values of n,K = 2(1)5,7,10,15,20,40 and for P = 0.99, 0.995, 0.999 are collected in Appendix B. The use of the tables does not differ from the univariate case. It is interesting to observe how the penalty ratios depend on the dimension of $\underline{Y}$. Plots of $r^N_{\underline{\mu},\underline{\Sigma}}$ and of $r^N_{\underline{\Sigma}}$ for P = 0.99 and

and $n \leq 20$ are shown in Figures 8 and 9 (note difference in scales). Both penalties increase considerably with K for any fixed n; for instance, given a sample of size n = 10, the first few penalty ratios for $\underline{\mu}$ and $\underline{\Sigma}$ unknown are:

K	$r^{N}_{\underline{\mu},\underline{\Sigma}}(10,K,0.99)$
1	1.323
2	1.525
3	1.778
4	2.133
5	2.683

Rectangular prediction regions

When K > 1 no obvious equivalents to one- or two-sided prediction intervals exist. Definitions of one-sided rectangular regions can be given, such as

$$D(P) = \{\underline{Y} | Y_i \leq \overline{Y}_i(n,K,P);\ i=1,\ldots,K\}$$

(for each i the sign $\leq$ may be replaced by $\geq$); also, two-sided rectangular prediction regions can be defined as:

$$D(P) = \{\underline{Y} | Y_i^m(n,K,P) < Y_i \leq Y_i^M(n,K,P);\ i=1,\ldots,K\} .$$

However, in no case do results exist with the generality of those for ellipsoidal regions. The appeal of rectangular regions is that they uncouple the prediction intervals of the components of $\underline{Y}$.

When the covariance matrix $\underline{\Sigma}$ is diagonal, one can use the one-dimensional results, as described earlier.

If $\underline{\Sigma}$ is given and has equicorrelated structure:

$$(\underline{\Sigma})_{ij} = \begin{cases} \sigma_i^2 \quad , \text{ for } i = j \ , \\ \\ \rho\sigma_i\sigma_j \ , \text{ for } i \neq j \ , \end{cases} \tag{II.98}$$

prediction regions of the type

$$D(P) = \{\underline{Y} | Y_i \leq \hat{Z}_i + \beta_{\underline{\mu}}^N \sigma_i\} \ , \text{ or}$$

$$D(P) = \{\underline{Y} | Y_i \geq \hat{Z}_i - \beta_{\underline{\mu}}^N \sigma_i\} \tag{II.99a}$$

can be used, where:

$$\hat{Z}_i = n^{-1} \sum_{j=1}^{n} Z_{ij}$$

(Z_{ij} = i^{th} component of the j^{th} observation),

$$\beta^N_{\underline{\mu}}(n,K,P) = (1 + 1/n)\ \beta^N_{PI}(K,P),$$

and $\beta^N_{PI}(K,P)$ satisfies:

$$\int_{-\infty}^{\beta^N_{PI}(\cdot)} \cdots \int_{-\infty}^{\beta^N_{PI}(\cdot)} f(\underline{y}^*)\ d\ \underline{y}^*$$

$$= \int_{-\infty}^{\infty} \Phi^K[(\beta^N_{PI}(\cdot) + \rho^{1/2} \cdot u)/(1-\rho)^{1/2}]\,\phi(u)\,du = P$$

($\Phi[\cdot]$ and $\phi(\cdot)$ are the standard normal CDF and PDF, and $Y^*_i = (Y_i-\mu_i)/\sigma_i$.)

Gupta (1963) tabulated P as a function of K (his n), ρ, and $\beta^N_{PI}(\cdot)$ (his h) for K = 1(1)12, ρ ranging discretely from 0.1 to 0.9, and for $\beta^N_{PI}(\cdot)$ = -3.5(0.1)3.5. The result

(II.99) is important in connection with some known inequalities of the normal integral over multidimensional rectangles (Slepian (1962); Sidak (1968); Tong (1970)) which allow one to obtain bounds on the probability content of rectangular regions in K-space for correlation structures other than (II.98).

For any correlation structure and for all combinations of $\underline{\mu}$ and $\underline{\Sigma}$ known or unknown, conservative two-sided rectangular regions can be constructed in several ways (see Chew (1968) for a different use of the same results).

1. Circumscribed rectangles. The (hyper)rectangle which circumscribes the ellipsoid $\{\underline{Y} | (\underline{Y}-\underline{\mu})' \underline{\Sigma}^{-1} (\underline{Y}-\underline{\mu}) \leq q^2\}$ and has sides parallel to the coordinate axes is:

$$D^* = \{\underline{Y} | (\mu_i - q\sigma_i) \leq Y_i \leq (\mu_i + q\sigma_i);\ i=1,\ldots,K\}.$$

If one replaces the regions (II.92), (II.94) and (II.96) by the two-sided rectangles:

$$D^{N*}_{\underline{\mu},\underline{\Sigma}}(P) = \{\underline{Y} | Y_i - \hat{Z}_i | \leq \beta^N_{\underline{\mu},\underline{\Sigma}}(n,K,P) \cdot S_i;\ (i=1,\ldots,K)\}$$

(where $\hat{Z}_i$ is the i^{th} component of $\underline{\hat{Z}}$ and S_i^2 is the i^{th} diagonal element of $\underline{S}$);

$$D_{\underline{\Sigma}}^{N*}(P) = \{\underline{Y} \Big| |Y_i - \mu_i| \leq \beta_{\underline{\Sigma}}^{N}(n,K,P) \cdot S_i;\ (i=1,\ldots,K)\}\ ;$$

and

$$D_{\underline{\mu}}^{N*}(P) = \{\underline{Y} \Big| |Y_i - \hat{Z}_i| \leq \beta_{\underline{\mu}}^{N}(n,K,P) \cdot \sigma_i;\ (i=1,\ldots,K)\}\ ,$$

the new regions have expected content greater than P (the values of $\beta_{(\cdot)}^{N}$ in the preceeding expressions are those tabulated in Appendix B). Nevertheless, for practical applications to problems of structural safety (for instance $\underline{Y}$ might be a load vector) this underestimation of the expected content may not be a critical factor, particularly if the components of $\underline{Y}$ are not highly correlated. When Y_i and Y_j are uncorrelated for $i \neq j$, the expected content of the circumscribed rectangle for a given expected content P of the inscribed ellipsoid is, for a few selected values of K and P, and for $n \to \infty$:

K	P = 0.75	P = 0.90	P = 0.99	r = 0.999
2	0.8174	0.9372	0.9952	0.999598
3	0.8773	0.9632	0.99773	0.999834
5	0.95075	0.9882	0.999485	0.9999703
10	0.99605	0.999364	0.9999855	$0.9^{6}470$
20	0.9999788	$0.9^{5}804$	$0.9^{7}823$	$0.9^{9}678$

TABLE 9. Expected content of circumscribed rectangles; P = expected content of the inscribed K-dimensional ellipsoid.

Sometimes the penalty associated with this conservatism is balanced by the simplicity of the rectangular geometry which, as observed earlier, uncouples the prediction intervals for each component of $\underline{Y}$. (Of course, one-sided rectangular regions which are tangent to the prediction ellipsoids are even more on the conservative side.)

2. Use of a Bonferroni's inequality. (see Lieberman (1961)). When $\underline{\mu}$ and $\underline{\Sigma}$ are both unknown, the event

$$E_i : |Y_i - \hat{Z}_i| \leq \left(\frac{n-1}{n}\right)^{1/2} S_i \, t_{n-K}\left(\frac{2K-1+P}{2K}\right) \qquad \text{(II.100)}$$

occurs with expected probability (K-1+P)/K. When using a Bo ferroni's inequality (see Feller (1967), p 110) one can prove that

$$P \leq \Pr\{\bigcap_{i=1}^{K} E_i\} \leq P + \Delta P$$

where

$$\Delta P = \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} \Pr\{\bar{E}_i \cap \bar{E}_j\}$$

(the bar denotes complementation) and $\Pr\{\bar{E}_i \cap \bar{E}_j\}$ can be found from tables of the bivariate t-integral (see references in Chew (1968) p 327; see also Hahn and Hendrickson (1971), and references therein).

In analogous way, prediction intervals for $\underline{\mu}$ known and $\underline{\Sigma}$ unknown are:

$$|Y_i - \mu_i| \leq t_{n-K+1}\left(\frac{2K-1+P}{2K}\right) S_i \,, \qquad \text{(II.101)}$$

and for $\underline{\mu}$ unknown and $\underline{\Sigma}$ known:

$$|Y_i-\mu_i| \leq (1 + 1/n)^{1/2} \Phi\left(\frac{2K-1+P}{2K} \right)\sigma_i \ . \qquad \text{(II.102)}$$

3. Other results. When Σ is unknown and S has equicorrelated structure with correlation coefficient ρ, one can use known percentage points of the largest absolute value of K equicorrelated "Student's" t-variates (Hahn and Hendrickson (1971)) to construct exact two-sided rectangular prediction regions. Similarly one can use tables of the largest signed value of K equicorrelated t-variates (Krishnaiah and Armitage (1966)) to construct one-sided rectangular prediction regions.

Denote $u(K,\nu,\rho;P)$ the P-fractile of these maximum distributions (ν = number of degrees of freedom of the t-distribution). Two-sided prediction intervals for each component of Y are:

$$|Y_i-\hat{Z}_i| \leq S_i \cdot u(K,n-K,\rho;P) \cdot (1 + 1/n)^{1/2}$$

when μ is unknown, and

$$|Y_i-\mu_i| \leq S_i \cdot u(K,n-K+1,\rho;P)$$

when $\underline{\mu}$ is known. Similar results hold for one-sided prediction intervals.

EXAMPLE

Eight independent samples are available from a bivariate normal population, with both $\underline{\mu}$ and $\underline{\Sigma}$ unknown. The sample statistics are:

$$n = 8; \qquad \hat{\underline{Z}} = \begin{pmatrix} 5 \\ \\ 6 \end{pmatrix} ; \qquad \underline{S} = \begin{pmatrix} 4 & 2 \\ \\ 2 & 4 \end{pmatrix} .$$

From equation (II.92) and from Table B1 in Appendix B, the elliptical central prediction region of 0.99 expected content is:

$$D^{N}_{\underline{\mu},\underline{\Sigma}}(0.99) = \left\{\underline{Y} \middle| [(Y_1-5),(Y_2-6)] \begin{pmatrix} 1/3 & -1/6 \\ \\ -1/6 & 1/3 \end{pmatrix} \begin{pmatrix} (Y_1-5) \\ \\ (Y_2-6) \end{pmatrix} \leq 28.68 \right\}. \qquad \text{(II.103)}$$

The squares of the principal semiaxes of the prediction ellipse are $\{S_{11} + S_{22} \pm [(S_{11}-S_{22})^2 + 4\, S_{12}]^{1/2}\} \times 28.68/2$, or $(12.461)^2$ and $(8.612)^2$. The area of the ellipse is 337.14. After assuming $\underline{\Sigma} = \underline{S}$, the penalty ratio (II.93) is, from Table B 1 : $r^N_{\underline{\mu},\underline{\Sigma}}(8,2,0.99) = 1.765$. This means that the area of the prediction ellipse under perfect information and with $\underline{\Sigma} = \underline{S}$ is $337.14/(r^N_{\underline{\mu},\underline{\Sigma}})^2 = 108.22$.

The circumscribed rectangle is defined by the inequalities:

$$5 - 2\sqrt{28.68} \leq Y_1 \leq 5 + 2\sqrt{28.68}$$

$$6 - 2\sqrt{28.68} \leq Y_2 \leq 6 + 2\sqrt{28.58}$$

with an area of 458.88; i.e., 1.361 times the area of the prediction ellipse.

Using Bonferroni's inequality the prediction rectangle is defined by:

$$5 - 2\sqrt{9/8}\; t_6\left(\frac{3.99}{4}\right) \leq Y_1 \leq 5 + 2\sqrt{9/8}\; t_6\left(\frac{3.99}{4}\right) ;$$

$$6 - 2\sqrt{9/8}\; t_6\left(\frac{3.99}{4}\right) \leq Y_2 \leq 6 + 2\sqrt{9/8}\; t_6\left(\frac{3.99}{4}\right) ,$$

or:

$$-4.158 \leq Y_1 \leq 14.158 ,$$

$$-3.158 \leq Y_2 \leq 15.158 ,$$

with area 335.46, being 0.995 times the area of the ellipse (II.103).

Finally, if one uses the tables in Hahn and Hendrickson for $\rho = 0.5$, $K = 2$, $\nu = 6$ and $P = 0.99$ one finds $u(2,6,0.5;0.99) = 4.21$, so that a prediction rectangle of 0.99 expectation is

$$5 - 8.42\sqrt{9/8} \leq Y_1 \leq 5 + 8.42\sqrt{9/8} \text{ , or } -3.931 \leq Y_1 \leq 13.931$$

$$6 - 8.42\sqrt{9/8} \leq Y_2 \leq 6 + 8.42\sqrt{9/8} \text{ , or } -2.931 \leq Y_2 \leq 14.931$$

The area of this rectangle is 319.03, i.e., 0.946 times the area of the ellipse (II.103).

These four prediction regions are shown in Figure 10.

(b) **DISTRIBUTION-FREE PREDICTION REGIONS**

Multivariate nonparametric prediction generalizes logically the univariate theory. However, it presents additional difficulties, the most part of geometrical nature. Particularly worth of mention in this area are early studies by Wald (1943) and by Tukey (1947).

Wald proposed a method, later called that of "successive elimination", for the definition of prediction regions of given expectation, provided that enough data are available and that the population has continuous (but otherwise unknown) joint CDF.

Wald's method is introduced through a simple bi-dimensional application (see also Guttman (1970) p 8).

Let $(X_1,Y_1),\ldots,(X_n,Y_n)$ be a set of n observation pairs, with the order statistics $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$ with respect to X. Choose two positive integers r_1 and s_1 such that

$$r_1 \geq 1, \quad s_1 \leq n, \quad \text{and } r_1 \leq s_1 - 3 .$$

Eliminate the observations (X_i,Y_i) with $X_i \leq X_{(r_1)}$ or with $X_i \geq X_{(s_1)}$. Order the remaining observations with respect

to Y: $Y_{(1)} < Y_{(2)} < \cdots < Y_{(s_1-r_1-1)}$. Choose two more positive integers r_2 and s_2 satisfying:

$$r_2 \geq 1, \quad s_2 \leq s_1-r_1-1, \quad \text{and } r_2 \leq s_2-1.$$

Then Wald proves that the probability content of the region

$$D = \{(X,Y) \mid X_{(r_1)} < X \leq X_{(s_1)};\ Y_{(r_2)} < Y \leq Y_{(s_2)}\}$$

has Beta distribution with parameters $(p,q) = (s_2-r_2;\ n-s_2+r_2+1)$. In particular the expected probability content of D is

$$E[P(D)] = \frac{s_2 - r_2}{n + 1} \quad .$$

EXAMPLE:

Construct a two-dimensional region of expected content E[P] = 2/3 using the 29 data points in Figure 11.

Since we want that $s_2 - r_2 = 20$ in order that $E[P(D)] = 20/30 = 2/3$, a possible choice is:

$$r_1 = 2,\ s_1 = 29;$$

$$r_2 = 3,\ s_2 = 23.$$

These values satisfy the inequalities given earlier. The associated rectangular prediction region is shown in Figure 11. The construction is generalized easily to the case of regions with more than two dimensions.

Tukey (1947) extended Wald's method to the construction of regions of any prefixed shape, also including unbounded regions (such as one-sided rectangles). Tukey's starting point was the definition of (n+1) "statistically equivalent blocks" as sample statistics; then he proved that any collection of r such blocks has Beta-distributed probability content with parameters (p,q) = (r,n+1) (and therefore the expected content is r/(n+1)). For the construction of the "blocks" refer to the original paper by Tukey and to the later developments in Fraser (1953) and in Kemperman (1956).

Sample size requirements and penalty factors. The minimum sample size for the construction of nonparametric prediction regions of given content P depends on the shape of the region. For instance, when using Wald's method as in the example above, it must be:

$$\frac{n-3}{n+1} \geq P, \quad \text{or } n \geq \frac{P+3}{1-P} \qquad (K = 2)$$

which for P = 0.90 gives $n \geq 39$ and for P = 0.99 gives $n \geq 399$. The sample size requirements for the construction of similar regions in a K-dimensional space is: $n \geq (P + 2K - 1)/(1 - P)$, increasing linearly with K.

Computing expected penalty ratios when sampling from a certain distribution but using a nonparametric prediction region is more complicated than for the univariate case. When the components of the K-dimensional prediction vector $\underline{Y}' = [Y_1,\ldots,Y_K]$ are normal and independent, two-sided prediction regions of the type discussed earlier, with $n = (P + 2K - 1)/(1 - P)$ have expected volume:

$$E[\text{Vol}] = E[(Y_{1(n)} - Y_{1(1)})(Y_{2(n-2)} - Y_{2(1)})\ldots(Y_{K(n-2K+2)} - Y_{K(1)})]$$

$$= \prod_{j=1}^{K} E[(Y_{j(n-2j+2)} - Y_{j(1)})] = \prod_{j=1}^{K} \sigma_j \cdot \text{ENR}(n-2j+2) \ ,$$

where σ_j is the r.m.s. of Y_j and ENR(n) is the expected normalized range for a normal sample of size n, as tabulated by Tippett (1925).

For instance, for rectangular prediction regions of 0.90 expectation with $n = (P + 2K - 1)/(1 - P)$, one finds the following expected volume or length for K = 1, or area for K = 2) as K ranges from 1 to 4 (in units of $\prod_{j=1}^{K} \sigma_j$).

K	n	Expected normalized volume, E[Vol]	Volume of prediction ellipsoid-perfect information, V_{PI}	Penalty ratio r(K)
1	19	3.69	3.29	1.12
2	39	18.32	14.47	1.13
3	59	97.27	65.47	1.14
4	79	537.67	298.70	1.16

TABLE 10. K-variate nonparametric prediction when sampling from multinormal populations.

Table 10 gives also the normalized volumes of central ellipsoid regions of 0.99-content under perfect information.

These volumes are:

$V_{PI} = 2\Phi(0.95)$, for K =1; $V_{PI} = \pi\ \chi_2^2(0.90)$, for K=2;

$V_{PI} = \frac{4}{3}\ \pi[\chi_3^2(0.90)]^{3/2}$, for K=3; $V_{PI} = \frac{\pi^2}{2}\ [\chi_4^2(0.90)]^2$, for K=4.

The penalty ratios in the last column are defined:

$$r(K) = (E[Vol]/V_{PI})^{1/K} .$$

In comparing the penalty ratios one should note that they refer to different sample sizes and that in fact the main penalty for not knowing the distribution type is the high minimum value of n.

II.2.4. Bayesian Prediction for Multivariate Sequences

The Bayesian literature on multivariate parametric prediction (like the frequentist literature on the same subject) is confined to multinormal populations (see Guttman (1970); Zellner (1971), p 72; see also Geisser (1965) for the Bayesian estimation of multinormal parameters).

On the other hand nonparametric Bayesian prediction regions can be constructed through a simple generalization of the univariate theory in Paragraph II.2.2(e).

(a) MULTINORMAL SEQUENCE

Prediction for the normal population $N_K(\underline{\mu},\underline{\Sigma})$ when both $\underline{\mu}$ and $\underline{\Sigma}$ are unknown is studied by Guttman (1970), p 135. No reference was found for the somewhat simpler cases when only the location or the scale parameter is unknown; these cases are worked out below.

$\underline{\mu}$ unknown, $\underline{\Sigma}$ known. This is the simplest case. The likelihood function in terms of the sufficient sample statistics:

$$n;\quad \hat{\underline{Z}} = n^{-1} \sum_{i=1}^{n} \underline{Z}_i$$

is:

$$l(\underline{\mu}|n,\hat{\underline{Z}}) \propto \exp\{-\frac{n}{2}(\underline{\mu}-\hat{\underline{Z}})'\underline{\Sigma}^{-1}(\underline{\mu}-\hat{\underline{Z}})\} \quad .$$

For a prior in the form of a normal density:

$$f(\underline{\mu}|n_o,\underline{\mu}_o) = n_o^{1/2}(2\pi)^{-K/2}|\underline{\Sigma}^{-1}|^{1/2}\exp\{-\frac{n_o}{2}(\underline{\mu}-\underline{\mu}_o)'\underline{\Sigma}^{-1}(\underline{\mu}-\underline{\mu}_o)\},$$

(II.104)

the posterior density is also normal, of the type $f(\underline{\mu}|n_1,\underline{\mu}_1)$ with parameters:

$$n_1 = n + n_o \quad ;$$

$$\underline{\mu}_1 = (n + n_o)^{-1}(n_o\underline{\mu}_o + n\,\hat{\underline{Z}}). \qquad \text{(II.105)}$$

From equation (II.5), after integrating with respect to $\underline{\mu}$, one finds that $\underline{Y}$ has multinormal distribution:

$$\underline{Y} \sim N_K(\underline{\mu}_1; \frac{n_1 + 1}{n_1}\underline{\Sigma}) \quad , \text{ or:}$$

$$\sim N_K\left(\frac{n_o\underline{\mu}_o + n\hat{\underline{Z}}}{n_o + n}; \frac{n + n_o + 1}{n + n_o}\underline{\Sigma}\right) . \qquad \text{(II.106)}$$

$\underline{\mu}$ known, $\underline{\Sigma}$ unknown. The likelihood function of $\underline{\Sigma}^{-1}$, given a sample $\underline{Z}_1,\ldots,\underline{Z}_n$, is:

$$l(\underline{\Sigma}^{-1}|n,\underline{S}) \propto |\underline{\Sigma}^{-1}|^{-n/2} \exp\{-\frac{n}{2} tr[\underline{\Sigma}^{-1}\underline{S}]\} ,$$

where

$$\underline{S} = n^{-1} \sum_{i=1}^{n} (\underline{Z}_i - \underline{\mu})(\underline{Z}_i - \underline{\mu})' .$$

The conjugate prior family, indexed by n_o and $\underline{S}_o$, is:

$$f(\underline{\Sigma}^{-1}|n_o,\underline{S}_o) \propto |\underline{\Sigma}^{-1}|^{-\frac{1}{2}n_o+1} \exp\{-\frac{n_o}{2} tr[\underline{\Sigma}^{-1}\underline{S}_o]\}$$

for which the posterior density results:

$$f(\underline{\Sigma}^{-1}|n,\underline{S},n_o,\underline{S}_o) \propto |\underline{\Sigma}^{-1}|^{-\frac{1}{2}n_1+1} \exp\{-\frac{n_1}{2} tr[\underline{\Sigma}^{-1}\underline{S}_1]\} ,$$

where $n_1 = n_o + n;$

$$\underline{S}_1 = (n\underline{S} + n_o\underline{S}_o)/(n + n_o). \qquad \text{(II.107)}$$

After the necessary integration, the predictive posterior density is:

$$f(\underline{Y}|n_1,\underline{S}_1) \propto |\underline{S}_1^{-1}|^{1/2} [1 + \frac{1}{n_1}(\underline{Y}-\underline{\mu})'\underline{S}_1^{-1}(\underline{Y}-\underline{\mu})]^{-\frac{1}{2}(n_1-1)} , \tag{II.108}$$

which means that the quantity $[(n+n_o-K+1)/(n+n_o)]^{1/2}(\underline{Y}-\underline{\mu})$ has K-variate t-distribution with $\nu = (n+n_o-K+1)$ degrees of freedom amd covariance matrix $\frac{\nu}{\nu-2}\underline{S}_1$.

$\underline{\mu}$ and $\underline{\Sigma}$ unknown. (Guttman (1970).) Given a set of independent observations $\underline{Z}_1,\ldots,\underline{Z}_n$ from the normal population $N_K(\underline{\mu},\underline{\Sigma})$, the likelihood function of $\underline{\mu}$ and $\underline{\Sigma}^{-1}$ is:

$$l(\underline{\mu},\underline{\Sigma}^{-1}|n,\hat{\underline{Z}},\underline{S}) \propto |\underline{\Sigma}^{-1}|^{n/2} \exp\{-\frac{1}{2}tr\underline{\Sigma}^{-1}[(n-1)\underline{S}+n(\hat{\underline{Z}}-\underline{\mu})(\hat{\underline{Z}}-\underline{\mu})']\} ,$$

where

$$\hat{\underline{Z}} = n^{-1}\sum_{i=1}^{n}\underline{Z}_i ;$$

$$\underline{S} = (n-1)^{-1}\sum_{i=1}^{n}(\underline{Z}_i-\hat{\underline{Z}})(\underline{Z}_i-\hat{\underline{Z}})' .$$

The conjugate prior, also in terms of $\underline{\mu}$ and $\underline{\Sigma}^{-1}$ is:

$$f(\underline{\mu},\underline{\Sigma}^{-1}|n_o,\underline{\mu}_o,\underline{S}_o) \propto |\underline{\Sigma}^{-1}|^{(n_o-K-1)/2}$$

$$\cdot\exp\{-\frac{1}{2}\,\mathrm{tr}\,\underline{\Sigma}^{-1}[(n_o-1)\underline{S}_o + n_o(\underline{\mu}-\underline{\mu}_o)(\underline{\mu}-\underline{\mu}_o)']\} \quad ,$$

where n_o = equivalent prior sample size;

$\underline{\mu}_o$ = prior estimate of $\underline{\mu}$;

$\underline{S}_o$ = prior estimate of $\underline{\Sigma}$.

The posterior density has the same form: $f(\underline{\mu},\underline{\Sigma}^{-1}|n_1,\underline{\mu}_1,\underline{S}_1)$, with parameters:

$$n_1 = n_o + n \;;$$

$$\underline{\mu}_1 = \frac{n_o\underline{\mu}_o + n\,\hat{\underline{Z}}}{n_o + n} \;; \qquad \text{(II.109)}$$

$$\underline{S}_1 = \frac{1}{n_o+n-1}[(n_o-1)\underline{S}_o+(n-1)\underline{S} + \frac{n_o\,n}{n_o+n}(\hat{\underline{Z}}-\underline{\mu}_o)(\hat{\underline{Z}}-\underline{\mu}_o)'].$$

Guttman (1970) gives also the posterior predictive density for the next observation $\underline{Y}$. He finds that

$$\left[\frac{n_1(n_1-K)}{n_1+1} \right]^{1/2} (\underline{Y}-\underline{\mu}_1)$$

has multivariate t-distribution with (n_1-K) degrees of freedom and covariance matrix $\frac{n_1-K}{n_1-K-2} \underline{S}_1$.

In all the three cases analyzed before, an initial knowledge in the conjugate form is equivalent to a set of sufficient sample statistics:

$(n_o, \underline{\mu}_o)$ for $\underline{\mu}$ unknown and $\underline{\Sigma}$ known, (II.110a)

$(n_o, \underline{S}_o)$ for $\underline{\mu}$ known and $\underline{\Sigma}$ unknown, (II.110b)

$(n_o, \underline{\mu}_o, \underline{S}_o)$ for $\underline{\mu}$ and $\underline{\Sigma}$ unknown. (II.110c)

This agrees with earlier findings in univariate Bayesian prediction, Paragraph II.2.2. When the prior becomes diffuse, i.e., when $n_o \to 0$ and $\underline{S}_o \to \underline{0}$, Bayesian and frequentist prediction regions coincide. In any case, if the prior knowledge is expressed in the form of an equivalent prior sample,

equations (II.110), the frequentist results and the tables in Appendix B can be used also for Bayesian prediction, by entering them with the following sets of equivalent posterior statistics:

$$[n_o+n;(n_o\underline{\mu}_o+n\hat{\underline{Z}})/(n_o+n)], \quad \text{when } \underline{\mu} \text{ is unknown and } \underline{\Sigma} \text{ is known;} \tag{II.111a}$$

$$[n_o+n;(n_o\underline{S}_o+n\underline{S})/(n_o+n)], \quad \text{when } \underline{\mu} \text{ is known and } \underline{\Sigma} \text{ is unknown;} \tag{II.111b}$$

$$\left\{n_o+n;\frac{n_o\underline{\mu}_o+n\hat{\underline{Z}}}{n_o+n};\frac{1}{n_o+n-1}\left[(n_o-1)\underline{S}_o+(n-1)\underline{S}\right.\right.$$

$$\left.\left.+\frac{n\,n_o}{n_o+n}(\hat{\underline{Z}}-\underline{\mu}_o)(\hat{\underline{Z}}-\underline{\mu}_o)'\right]\right\},$$

when $\underline{\mu}$ and $\underline{\Sigma}$ are unknown. (II.111c)

(If n_o is fractional, interpolate between the closest integral approximations.)

EXAMPLE

A central prediction region of 0.999-expectation is desired for a normal vector $\underline{Y} = [Y_1,Y_2]'$. Both $\underline{\mu}$ and $\underline{\Sigma}$ are

unknown. The prior information is judged to be worth six sample points (n_o=6), with mean value:

$$\underline{\mu}_o = \begin{pmatrix} 10 \\ 15 \end{pmatrix}, \quad \text{and covariance matrix:} \quad \underline{S}_o = \begin{pmatrix} 4 & 4 \\ 4 & 9 \end{pmatrix}.$$

A sample of size n = 5 is observed, with sufficient statistics:

$$\hat{\underline{Z}} = \begin{pmatrix} 8 \\ 14 \end{pmatrix} \quad \text{and} \quad \underline{S} = \begin{pmatrix} 4 & 3 \\ 3 & 7 \end{pmatrix}.$$

The posterior mean and the posterior covariance matrix are found from equations (II.109):

$$\underline{\mu}_1 = \begin{pmatrix} 9.091 \\ 14.545 \end{pmatrix}, \quad \text{and} \quad \underline{S}_1 = \begin{pmatrix} 4.691 & 3.745 \\ 3.745 & 7.573 \end{pmatrix}.$$

From the tables in Appendix B one finds:

$$\beta^{N}_{\underline{\mu},\underline{\Sigma}}(5,2,0.999) = 21.799;$$

$$\beta^{N}_{\underline{\mu},\underline{\Sigma}}(6,2,0.999) = 13.366;$$

$$\beta^{N}_{\underline{\mu},\underline{\Sigma}}(11,2,0.999) = 6.303.$$

Therefore the frequentist, the prior Bayesian, and the posterior Bayesian prediction regions are:

Frequentist: $\{\underline{Y} | (\underline{Y}-\hat{\underline{Z}})' \underline{S}^{-1} (\underline{Y}-\hat{\underline{Z}}) \leq 21.799^2\}$;

Prior Bayesian: $\{\underline{Y} | (\underline{Y}-\underline{\mu}_o)' \underline{S}_o^{-1} (\underline{Y}-\underline{\mu}_o) \leq 13.366^2\}$;

Posterior Bayesian: $\{\underline{Y} | (\underline{Y}-\underline{\mu}_1)' \underline{S}_1^{-1} (\underline{Y}-\underline{\mu}_1) \leq 6.303^2\}$.

These three regions are shown in Figure 12. The penalty ratios associated with them are, from the tables in Appendix B:

Frequentist: $r^{N}_{\underline{\mu},\underline{\Sigma}}(5,2,0.999) = 5.864$;

Prior Bayesian: $r^{N}_{\underline{\mu},\underline{\Sigma}}(6,2,0.999) = 3.595$;

Posterior Bayesian: $r^{N}_{\underline{\mu},\underline{\Sigma}}(11,2,0.999) = 1.696$.

(b) DISTRIBUTION-FREE PREDICTION REGIONS

Along the lines of Ferguson (1973) it is not difficult to generalize the theory of univariate nonparametric Bayesian prediction to the multivariate case. In fact the Dirichlet process defined by Ferguson applies equally well when sampling is from a K-variate population with continuous density function. Formally one needs to replace y in Paragraph II.2.2 by R^K, the K-dimensional Euclidean space. Then, if α denotes the parameter of the Dirichlet process, and $(B_1,\ldots,B_m)$ is a measurable partition of R^K, the joint distribution of the probability coverages $(P(B_1),\ldots,P(B_m))$ is Dirichlet, with parameters $\alpha(B_1),\ldots,\alpha(B_m)$.

Since the Dirichlet process indexed by α is the conjugate family for nonparametric Bayesian inference, the prior knowledge can be usefully assumed in the form of a Dirichlet process with parameter α_o. Then, given a set of n independent observations $\underline{Z}_1,\ldots,\underline{Z}_n$, the posterior parameter is, from Ferguson's theorem 3:

$$\alpha_1 = \alpha_o + \sum_{i=1}^{n} \delta_{\underline{Z}_i} ,$$

where $\delta_{\underline{Z}_i}$ is defined as in Paragraph II.2.2(e). The measures $\alpha_o(R^K)$ and $\alpha_1(R^K) = \alpha_o(R^K)+n$ quantify (in units of observations) the prior and the posterior degrees of belief in the best estimates of the probability distribution over R^K.

A result similar to (II.90) holds for the best estimate of the posterior distribution under a quadratic loss criterion, giving this function as a weighted sum of the prior best estimate of the distribution and of the empirical distribution function of the sample.

Prediction regions of P-expectation can be constructed easily, since for any measurable set $A \subset R^K$ it is:

$$E[P(A)] = \alpha(A)/\alpha(R^K)$$

if α is the parameter of the Dirichlet process.

II.2.5. Prediction from Censored Data

In Chapter I, Section 4, we analyzed censored data in the context of estimation theory, with particular emphasis on the applications to proof loading of structural systems. Expressions were found for the prior, for the posterior and for the "proof loading" failure probabilities.

Here we present the general formulation and two examples of processing censored data for statistical (Bayesian) prediction. An additional case, concerning censored sampling from exponential populations, was covered in Paragraph II.2.1(b).

Consider having n censored observations from a univariate statistical population with CDF $F_Y(\cdot)$. The censored information on the generic realization Y_i is:

$$Y_i = Z_i \qquad \text{if } Y_i \leq Z_{max} \,, \tag{II.112a}$$

$$Y_i > Z_{max} \qquad \text{otherwise} \,. \qquad (i=1,\ldots,n) \tag{II.112b}$$

In a parametric Bayesian approach ($F_Y(\cdot|\underline{\Theta})$ known, but not the vector of parameters $\underline{\Theta}$) all the observations of the

type (II.112a) can be processed as indicated in Paragraph II.2.2. The resulting posterior distribution of the unknown parameters $\underline{\Theta}$ is denoted $F_{a'}(\underline{\Theta})$. The posterior distribution of $\underline{\Theta}$ which accounts also for one observation of the type (II.112b) is, from Bayes' theorem:

$$d\,F_{a''}(\underline{\Theta}) = d\,F_{a'}(\underline{\Theta})\,\frac{Pr[Y>Z_{max}|\underline{\Theta}]}{Pr_{a'}[Y>Z_{max}]}$$

$$= d\,F_{a'}(\underline{\Theta})\,\frac{1 - F_Y(Z_{max}|\underline{\Theta})}{1 - F_{Y_{a'}}(Z_{max})}\quad , \qquad \text{(II.113)}$$

where $$F_{Y_{a'}}(\cdot) = \int_{\text{all } \underline{\Theta}} F_Y(\cdot|\underline{\Theta})\,d\,F_{a'}(\underline{\Theta})\ .$$

The posterior predictive CDF, $F_{Y_{a''}}(\cdot)$, is therefore:

$$F_{Y_{a''}}(\cdot) = \int_{\text{all } \underline{\Theta}} F_Y(\cdot|\underline{\Theta})\,d\,F_{a''}(\underline{\Theta})$$

$$= \frac{\int_{\text{all } \underline{\Theta}} F_Y(\cdot|\underline{\Theta})[1 - F_Y(Z_{max}|\underline{\Theta})]\,d\,F_{a'}(\underline{\Theta})}{\int_{\text{all } \underline{\Theta}} [1 - F_Y(Z_{max}|\underline{\Theta})]\,d\,F_{a'}(\underline{\Theta})} \qquad \text{(II.114)}$$

If $\bar{n} \leq n$ is the number of observations of the type (II.112b), the posterior distribution of $\underline{\Theta}$ is found by applying inductively equation (II.113):

$$d\,F_a(\underline{\Theta}) = d\,F_{a'}(\underline{\Theta}) \left(\frac{1 - F_Y(Z_{max}|\underline{\Theta})}{1 - F_{Y_{a'}}(Z_{max})} \right)^{\bar{n}}, \qquad \text{(II.115)}$$

from which the posterior predictive CDF, $F_{Y_a}(\cdot)$, is:

$$F_{Y_a}(\cdot) = \int_{\text{all } \underline{\Theta}} F_Y(\cdot|\underline{\Theta})\, d\,F_a(\underline{\Theta}) \; . \qquad \text{(II.116)}$$

If Y denotes a "resistance" variable, and the distribution of the associated "load" S, $F_S(\cdot)$, is known, the probability of failure in a time-invariant situation is:

$$P_f = \int_0^\infty F_{Y_a}(Z)\, d\,F_S(Z) \; .$$

A completely analogous formalism applies to multivariate situations in which, however, one may have also partially censored sample data.

EXAMPLE 1. NORMAL SEQUENCE

Suppose sampling is from a normal white sequence with unknown mean and variance ($\underline{\Theta}' = [\mu,\sigma^2]$) and that the prior $F(\underline{\Theta})$ is in the form of a conjugate Normal-Gamma (NG) distribution with parameters $[\mu',n',(n-1)S'^2]$; see equation (II.66). After processing the $(n-\bar{n})$ observations with values $Y_i = Z_i \leq Z_{max}$ $(i=1,\ldots,n-\bar{n})$, the posterior distribution $F_{a'}(\underline{\Theta})$ is again NG with parameters $[\mu'',n'',(n''-1)S''^2]$, where:

$$n'' = n' + n - \bar{n} \quad ;$$

$$\mu'' = [n'\mu' + (n-\bar{n})\,\hat{Z}]/n'' \quad ;$$

$$(n''-1)S''^2 = (n'-1)\,S'^2 + (n-\bar{n}-1)\,S^2 + \frac{n'(n-\bar{n})}{n'+n-\bar{n}}\,(\hat{Z}-\mu')^2 \quad ,$$

and $$\hat{Z} = (n-\bar{n})^{-1} \sum_{i=1}^{n-\bar{n}} Z_i \quad ;$$

$$S^2 = (n-\bar{n}-1)^{-1} \sum_{i=1}^{n-\bar{n}} (Z_i-\hat{Z})^2 \quad .$$

Denoting $NG[\mu,\sigma^2|\mu_1,n_1,q]$ the Normal-Gamma density with parameters μ_1,n_1,q, evaluated at (μ,σ^2); and denoting $\Phi(\cdot)$ the standard normal CDF, the posterior density of $\underline{\Theta}$ is, from equation (II.115):

$$f_a(\mu,\sigma^2) \propto NG[\mu,\sigma^2|\mu",n",(n"-1)S"^2]\{1 - \Phi[(Z_{max}-\mu)/\sigma]\}^{\overline{n}}$$

$$\propto \sigma^{-(n"+\frac{1}{2})} \exp\{-[(n"-1)S"^2 + n"(\mu-\mu")^2]/2\sigma^2\}$$

$$\cdot \int_{\frac{(Z_{max}-\mu)}{\sigma}}^{\infty} \exp(-t^2/2)\, d\, t$$

and the predictive posterior CDF results, from equation (II.116):

$$F_{Y_a}(y) = \int_{\underline{\Theta} \text{ space}} \Phi[(y-\mu)/\sigma]\, f_a(\mu,\sigma^2)\, d\mu\, d\sigma^2 \quad .$$

EXAMPLE 2. EXPONENTIAL SEQUENCE

In a recent paper Pierce (1973) computed the posterior density function of the predictive CDF, $F_a(Y)$, when a censored sample is known from the two-parameter negative exponential population:

$$f(Y|\mu,\sigma) = \sigma^{-1} \exp[-(Y-\mu)/\sigma] \; ; \quad \mu,\sigma \geq 0; \; Y \geq \mu,$$

with both μ and σ unknown. For the case with $\mu = 0$ and σ unknown, see Paragraph II.2.1(b).

Prediction for the exponential distribution is of special interest in reliability theory because the time-to-failure of stochastic systems is often associated with Poisson events. Also, since these events are rare, data are typically available in censored form (i.e., censored by the present time).

Pierce made a Bayesian analysis for the case when the prior is noninformative: $f(\mu,\sigma) \propto \sigma^{-1}$, and at least two "failures" (two uncensored data) are available. (If not, the posterior density is also improper.) Denote $Z_{(1)} < Z_{(2)} < \ldots < Z_{(n-\bar{n})}$ the first $(n-\bar{n}) \geq 2$ order statistics. In terms of the sufficient statistics:

$$Z_{(1)} \ ;$$

$$S = \sum_{i=1}^{n-\bar{n}} (Z_{(i)} - Z_{(1)}) + \bar{n}(Z_{(n-\bar{n})} - Z_{(1)})$$

Pierce found:

$$f(P_S | n, \bar{n}, Z_{(1)}, S) = K(Z_{(1)}, S)(n-\bar{n}-2)! \ T^{-n+\bar{n}-1} \ P_S^{n-1}$$

$$\cdot \{P_S^{T/Y} \sum_{i=0}^{n-\bar{n}-2} \frac{1}{i!} [-T \log P_S/Y]^i$$

$$- P_S^{T/(Y-Z_{(1)})} \sum_{i=0}^{n-\bar{n}-2} \frac{1}{i!} [-T \log P_S/(Y-Z_{(1)})]^i\} \ , \qquad \text{(II.117)}$$

where $P_S = 1 - F_a(Y)$;

$$T = S - n(Y-Z_{(1)}) \ ;$$

$$K(Z_{(1)}, S) = \{(n-\bar{n}-2)! n^{-1} [S^{-n+\bar{n}+1} - (S+n \ Z_{(1)})^{-n+\bar{n}+1}]\}^{-1} \ .$$

This result can be generalized to those informative cases in which the prior knowledge can be expressed through a set of equivalent prior sample statistics (see Paragraph II.2.2(b)): n^o, $\bar{n}^o$, $Z^o_{(i)}$ $(i=1,\ldots,n^o-\bar{n}^o)$, S^o. Then equation (II.117) holds after the following substitutions:

$$n \to n' = n + n^o \ ;$$

$$\bar{n} \to \bar{n}' = \bar{n} + \bar{n}^o \ ; \quad (n'-\bar{n}' \geq 2)$$

$$Z_{(1)} \to Z'_{(1)} = \min \{Z_{(1)}, Z^o_{(1)}\} \ ;$$

$$S \to S' = \sum_{i=1}^{n-\bar{n}} (Z_{(i)}-Z'_{(1)}) + \sum_{i=1}^{n^o-\bar{n}^o} (Z^o_{(i)}-Z'_{(1)})$$

$$+ \bar{n}(Z_{(n-\bar{n})}-Z'_{(1)}) + \bar{n}^o(Z_{(n^o-\bar{n}^o)}-Z'_{(1)}).$$

II.3. SIMULTANEOUS PREDICTION

Calculating the reliability of a structural system with respect to future events is a problem of prediction.

Suppose that nature is modeled (for the aspect of interest) as a sequence of independent random events with known distribution. Then finding $P_f(T) = \Pr\{\text{system failure in } [0,T]\}$ as a function of T might be called a problem of probabilistic prediction. If instead uncertainties of a statistical nature exist on the parameters of the distribution (and possibly also on the distribution type), the same problem becomes one of statistical prediction.

Conditional on a given probabilistic model being a correct representation of "nature" the reliability of a system depends on the statistical uncertainty of the model parameters. The purpose of the present section is to study the effect that statistical uncertainty has on (structural) reliability when failure occurs either as a no-memory first-crossing event (e.g., when the load process overcomes first the system resistance), or as the conclusion of a damage accumulation process (fatigue). In both cases calculating the system reliability requires predicting several future observations from the same process; in the jargon of statistics theory this is a problem of "simultaneous prediction".

In the special case when the reliability function $R(T) = 1 - r_f(T)$ is constant with time, the problem becomes one of simple prediction (see Section II.2).

A general approach to time-dependent reliability analysis, including statistical uncertainty.

Depending on the structural system (e.g., brittle or ductile), on the failure criterion (e.g., first-crossing, fatigue, low-cycle damage accumulation, maximum economic loss), and on the disturbances (earthquakes, high wind pressures, foundations settlements, live loads, etc.) a variety of stochastic models can be formulated for reliability analysis, with different levels of sophistication and complexity.

Also when the model parameters, vector $\underline{\Theta}$, are given, it is seldom possible to derive the probability of failure function $P_f(T|\underline{\Theta})$ in closed form. When the statistical uncertainty on $\underline{\Theta}$ is added, even fewer analytical solutions are known. However, the general methodology is straightforward, as is shown next by following the Bayesian approach.

Let A be the present state of knowledge, and denote $F_{\underline{\Theta}}(\underline{\theta}|A)$ the probability distribution of the model parameters conditional on it. $F_{\underline{\Theta}}(\cdot|A)$ can be derived from Bayesian statistics methods, or else may be given directly. The (statistical) reliability function: $R(T|A) = \Pr\{\text{survival in } [0,T], \text{ given } A\}$ is obtained through convolution:

$$R(T|A) = 1 - P_f(T|A) = 1 - \int_{\underline{\Theta} \text{ space}} P_f(T|\underline{\theta})\, dF_{\underline{\Theta}}(\underline{\theta}|A). \quad \text{(II.118)}$$

If A = perfect information, supporting the value $\underline{\Theta} = \underline{\Theta}^*$, then $f_{\underline{\Theta}}(\underline{\Theta}|A) = \delta(\underline{\Theta}-\underline{\Theta}^*)$ (Dirac-delta function), and equation (II.118) yields the (probabilistic) reliability function:

$$R(T|A) = 1 - P_f(T|A) = 1 - P_f(T|\underline{\Theta}^*). \qquad \text{(II.119)}$$

In the following paragraph independent reliability models are considered for the case of failure occurring as a first-crossing event with no damage accumulation, and an alternative procedure for statistical reliability analysis is presented. A more detailed (frequentist and Bayesian) analysis of the same models is the object of Paragraphs II.3.2 and II.3.3.

Correlated sequences are studied in Chapter III.

II.3.1. Independent Models for Time-Dependent Reliability

Consider a structural system with deterministic "resistance" R and 0-1 damage function (i.e., either the system performs satisfactorily, or else it fails, without passing through intermediate damage conditions). The environment is modeled as a stationary white sequence $\{Y_i\}$

of either sustained "loads" (model S in Figure 13), or of instantaneous "loads" (model P in the same figure). The exact times at which the load intensity changes in model S, or at which the point events occur in model P, may be deterministic or random. If $\{Y_i\}$ is independent of the interarrival times, one can replace [0,T] in the definition of failure by the number N of events during the same time interval. Again, N may be deterministic or stochastic. In both cases the first problem which needs to be solved is to find the probability of failure $P_f(N)$ as a function of the number of events. In fact if the time t_i at which the i^{th} event occurs is known it is:

$$P_{f_S}(T) = P_f(N_{T,S}); \text{ and } P_{f_P}(T) = P_f(N_{T,P}), \qquad \text{(II.120)}$$

where the subscripts S and P refer to the model (see Figure 13), and $N_{T,S}$ and $N_{T,P}$ are the values of N such that:

$$t_N \leq T < t_{N+1} \qquad \text{(II.121)}$$

(in the models S and P, respectively).

If instead the arrival times are random, the relation (II.121) defines the probability distributions of $N_{T,S}$ and $N_{T,P}$ given T, say:

$$P_{T,S}(N) = \Pr\{\text{exactly N events occur in model S during } [0,T]\} ;$$

$$P_{T,P}(N) = \Pr\{\text{exactly N events occur in model P during } [0,T]\} .$$

In this case the probability of failure as a function of time is:

$$P_{f_S}(T) = \sum_{N=1}^{\infty} P_{f_S}(N) \cdot P_{T,S}(N) ; \qquad \text{(II.122a)}$$

$$P_{f_P}(T) = \sum_{N=1}^{\infty} P_{f_P}(N) \cdot P_{T,P}(N) . \qquad \text{(II.122b)}$$

Note that in model S at least one occurrence must be considered for any T since $t_1 = 0$, while this is not true for model P. As a consequence in model S there is a nonzero probability of failure at time T = 0.

As shown above, the evolution in time of P_{f_S} and P_{f_P} depends on the variability of the occurrence times. For

deterministic arrivals, for highly regular (but random) arrival times, for highly regular (but random) event clusters, and for Poisson-type occurrences the qualitative evolution of P_{f_S} is sketched in Figure 14, cases a,b,c,d respectively.

All these cases are generated through equations (II.122) by the same function $P_f(N)$ and by different discrete distributions $P_{T,S}(N)$. From here arises the importance of studying the functions $P_{f_S}(N)$ and $P_{f_P}(N)$.

After replacing [0,T] by the number of events N in [0,T], computing the structural reliability R(N|A) as a function of N does not necessarily require using equation (II.118). In a Bayesian approach (a parallel analysis could be done for the frequentist viewpoint) an alternative and equally general procedure consists in finding the unconditional (with respect to $\underline{\Theta}$) predictive distribution of the next N realizations $[Y_1,\ldots,Y_n]' = \underline{Y}$ from the sequence $\{Y_i\}$:

$$F_{\underline{Y}}(\underline{y}|A) = \int_{\text{all } \underline{\Theta}} F_{\underline{Y}}(\underline{y}|\underline{\Theta}) \, d\, F_{\underline{\Theta}}(\underline{\Theta}|A) \ , \qquad \text{(II.123)}$$

Then the reliability is computed as:

$$R(N|A) = \int_{\substack{\text{Safe region in} \\ \underline{Y} \text{ space}}} d\, F_{\underline{Y}}(\underline{y}|A). \qquad \text{(II.124)}$$

In both procedures (equation (II.118) with N in place of T, and equations (II.123) and (II.124)) one has to integrate over $\underline{\theta}$; however equations (II.123) and (II.124) may have advantages over equation (II.118) if the integral (II.123) has closed-form solution. This happens for particular states of knowledge A, and for particular probabilistic models; one such case is when A is the collection of n independent realizations from the same normal sequence which generates the future observations $\underline{Y}$ (Paragraph II.3.2).

If $\{Y_i\}$ is, as assumed here, an independent stationary sequence, and the marginal distribution of Y_i is known, the probabilistic reliability R(N) after N "Bernoulli trials" is:

$$R(N) = 1 - P_f(N) = [1 - P_f(1)]^N , \quad \text{(II.125)}$$

where $P_f(1)$ is the probability that $Y_i > R$ for any i. On semilog paper the reliability R(N) plots linearly against N.

Paragraphs II.3.2 and II.3.3 deal (implicitly or explicitly) with the modifications that statistical uncertainty induces on equation (II.125). The concluding paragraph considers the effects of statistical uncertainty on reliability models with fatigue-type failure modes.

II.3.2. Simultaneous Frequentist Prediction for Univariate Sequences

The independent models in Paragraph II.3.1 are studied here under the assumption that the distribution parameters (possibly also the distribution shape) are estimated from limited data with the methods of classical statistics.

For univariate sequences the problem of simultaneous prediction consists in finding an open or closed interval on the real line (or more generally a region in R^N) in which all N future realizations (or the N-dimensional future realization) of the process will fall with given probability. This is in complete analogy with what was done in Section II.2 for simple prediction. When the simultaneous prediction interval (region) coincides with the safe interval (region) of the system, the same probability may be interpreted as a reliability with respect to the next N random events.

We first discuss the construction of simultaneous prediction intervals and regions for normal sequences and then we present some results for nonnormal populations.

(a) NORMAL SEQUENCES (see Lieberman (1961); Chew (1968).)

Let a sample of size n be available from the independent normal sequence $\{Y_i\}$, $Y_i \sim N(\mu,\sigma^2)$, and denote

$Z_1,\ldots,Z_N$ the observed values.

μ and σ unknown. When both the location and the scale parameters are unknown, the sufficient sample statistics are

$$n;\ \hat{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i \quad ; \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n} (Z_i-\hat{Z})^2.$$

Denote $\underline{Y}' = [Y_1,\ldots,Y_N]$ the vector of future observations and let $\hat{\underline{Z}}' = \hat{Z}[1,\ldots,1]$, with N components. Lieberman (1961) found that the statistic $\underline{d} = \underline{Y}-\hat{\underline{Z}}$ has multivariate t-distribution with (n-1) degrees of freedom and covariance matrix $\frac{n^2-1}{n(n-3)} S^2 \underline{R}$ $(n > 3)$, $\underline{R}$ being the correlation matrix:

$$\underline{R} = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & & & \\ \vdots & & & \vdots \\ \rho & & & 1 \end{pmatrix}, \quad \rho = 1/(n+1)$$

The structure of $\underline{R}$ shows that in presence of statistical uncertainty on μ and σ the components of $\underline{d}$ are not independent, as instead they would be if μ and σ were known (in fact $\underline{R} \to \underline{I}$ as $n \to \infty$).

A direct consequence of the result above is that the ellipsoid in $\underline{Y}$-space:

$$(\underline{Y}-\hat{\underline{Z}})' \underline{R}^{-1} (\underline{Y}-\hat{\underline{Z}}) \leq \frac{n+1}{n} N S^2 F_{N,n-1}(P) \qquad \text{(II.126)}$$

has expected content P.

For the purpose of reliability analysis and design, having a prediction region in R^N like the ellipsoid (II.126) is not of much help. If failure occurs when the first event of the type $Y_i > R$ (or $|Y_i| > R$) occurs, it would be more useful to dispose of a one-sided (or of a two-sided) interval on the real line, containing all the next N realizations $Y_1,\ldots,Y_N$ with a given probability P. This corresponds to constructing cuboidal prediction regions of P-expectation in R^N, and to this end the following results are useful.

Gupta and Sobel (1957) showed that the integral

$$P(h) = \int_{-\infty}^{h} \cdots \int_{-\infty}^{h} f(t_1,\ldots,t_N)\, d\, t_1 \cdots d\, t_N \ ,$$

where $t_1,\ldots,t_N$ have jointly N-variate t-distribution with (n-1) degrees of freedom and common correlation coefficient $\rho_{ij} = \rho$, $i \neq j$, can be written as

$$P(h) = \int_0^\infty \left\{ \int_{-\infty}^{\infty} \left[\Phi\left(\frac{h\,S + \rho^{1/2}\,y}{(1-\rho)^{1/2}} \right) \right]^N \phi(y)\,d\,y \right\} g_{n-1}(S)\,d\,S. \quad \text{(II.127)}$$

In equation (II.127) $\phi(\cdot)$ and $\Phi(\cdot)$ are the PDF and the CDF of the standard normal variate. Also, $g_m(S)$ is the probability density function of S, where $m\,S^2$ is a Chi-square variate with m degrees of freedom:

$$g_m(S) = \frac{m^{m/2}}{\Gamma\left(\frac{m}{2}\right) 2^{(m/2)-1}}\; S^{m-1}\; e^{-m\,S^2/2}\,.$$

Extending previous tables and using equation (II.127), Hahn (1970) computed the one-sided 100 P percent prediction limit for P = 0.90, 0.95 and 0.99; for n = 4(1)12, 15(5)30, 40, 60, ∞ and for a future sample size N = 1(1)6(2)12, 15, 20.

Figure 15 shows the factors $\beta^N_{\mu,\sigma}(P,n,N)$ such that

$$\Pr\left\{ \bigcap_{i=1}^{N} Y_i \leq \hat{\mu} + \beta^N_{\mu,\sigma}(P,n,N)\,S \right\} = P$$

for P = 0.99 and for several values of n and N (solid lines). These factors are taken from the tables computed by Hahn.

Under perfect information ($n \to \infty$) it would be

$$\beta^N_{\mu,\sigma}(P,\infty,N) = \Phi(P^{1/N}) \tag{II.128}$$

where $\Phi(\gamma)$ is the γ-fractile of the standard normal distribution. Therefore the penalty ratio for statistical uncertainty, defined in accordance with Section II.2, is:

$$r^N_{\mu,\sigma}(P,n,N) = \frac{\beta^N_{\mu,\sigma}(P,n,N)}{\Phi(P^{1/N})} \quad . \tag{II.128}$$

The computation of $\beta^N_{\mu,\sigma}(\cdot)$ is tedious, especially for large values of P. For practical applications it would be advisable to have some simple approximations. One way of obtaining approximate values for $\beta^N_{\mu,\sigma}$ is suggested by the well-known fact (see, e.g., Tong (1970)) that the probability

$$\Pr\{\bigcap_{i=1}^{N} (Y_i \leq R)\} \tag{II.129}$$

is a decreasing function of the correlation coefficients ρ_{ij} between Y_i and Y_j. Since these coefficients are all equal and positive, replacement of the true correlation matrix $\underline{R}$ by the identity matrix (uncorrelated variables) gives a conservative estimate of the reliability. The approximate value for $\beta^N_{\mu,\sigma}(\cdot)$ is:

$$\overline{\beta}^N_{\mu,\sigma}(P,n,N) = t_{n-1}(P^{1/N})(1 + 1/n)^{1/2} , \qquad \text{(II.130)}$$

where $t_\nu(\gamma)$ is the γ-fractile of the "Student's" t-distribution with ν degrees of freedom.

For all values of P, n, and N it is $\overline{\beta}^N_{\mu,\sigma}(\cdot) \geq \beta^N_{\mu,\sigma}(\cdot)$. The approximation (II.130) is believed to be quite good for large values of P (i.e., in the range of interest to structural reliability) and for moderate to large sample sizes, n. In fact the correlation coefficient is $(n+1)^{-1}$ (which is quite small for, say, $n > 10$), and it is known that the probability (II.129) when $\underline{Y}$ has equicorrelated normal distribution is insensitive to the presence of small correlation coefficients when P is large (see Chapter I, Example 1). The same is believed to be true when $\underline{Y}$ has N-variate t_{n-1}-distribution.

In Tables C1 - C4 (Appendix C) the values of $\overline{\beta}^N_{\mu,\sigma}(P,n,N)$ are computed for P = 0.90, 0.99, 0.999, 0.9999;

n = 2(1)12, 15(5)30, 40, 60, ∞ and for a "future" sample size N = 1(1)5, 10, 20. For P = 0.90 and P = 0.99 (Tables C1 and C2) the approximate values (in parenthesis) are compared with the exact values from Hahn. For P = 0.99 some approximate values are also plotted in Figure 15 (dashed lines).

These comparisons show that replacing $\underline{R}$ by $\underline{I}$ yields quite good (and always conservative for reliability purposes) approximations to $\beta^N_{\mu,\sigma}(\cdot)$. The only cases in which the discrepancy between $\beta^N_{\mu,\sigma}(\cdot)$ and $\overline{\beta}^N_{\mu,\sigma}(\cdot)$ is considerable is when n and P are both "small", a joint condition which rarely occurs in structural reliability.

Tables C5 - C7 collect the values of the penalty ratio $r^N_{\mu,\sigma}$ for P = 0.99, 0.999, 0.9999 and for the same combinations of n and N. For P = 0.99 these ratios use the exact $\beta^N_{\mu,\sigma}$ values from Hahn's tables (with the exception of n = 2, 3), while for P = 0.999 and P = 0.9999 and for n = 2, 3 the ratios are calculated from the approximation (II.130). The same ratios for P = 0.999, 0.9999 are plotted in Figures 16 and 17 against the available sample size n for a future number of observations N = 1, 3, 10, 20. The curves for N = 1 correspond to simple prediction intervals.

For the case of two-sided prediction intervals centered on the sample mean, Chew (1968) proposed two approximations to $\beta^{N}_{\mu,\sigma}(P,n,N)$. Both of them are conservative, suggesting to replace the exact value with an upper-bound. The stricter of the two approximations is based on a Bonferroni's inequality (see Chapter II, Paragraph II.2.3(a)) which for one-sided prediction intervals yields:

$$\beta^{N}_{\mu,\sigma}(P,n,N) \leq (1 + 1/n)^{1/2} t_{n-1}\left(\frac{N - 1 + P}{N}\right) .$$

Hahn compared this approximation with the exact values for P = 0.90, 0.95, 0.99 and for several choices of n and N. He found that the approximation is in general accurate, if one excludes the combination: N large - n small. He also found that the same approximation improves with increasing P, in analogy with the presently proposed upper bound, equation (II.130).

<u>μ known, σ unknown</u>. Define $\underline{d} = \underline{Y}-\underline{\mu}$ with generic component $d_i = Y_i-\mu$. From Chew (1968), $\underline{d}$ has N-variate t-distribution with n degrees of freedom, mean zero, and covariance matrix $\frac{n}{n-2} S^2 \underline{I}_N$, where

$$S^2 = \frac{1}{n} \sum_{i=1}^{n} (Z_i - \mu)^2$$

is the sample variance. Therefore the prediction interval

$$(-\infty, \ \mu + \beta_\sigma^N(P,n,N) \cdot S] \qquad \text{(II.131a)}$$

contains the future N realizations with probability P if $\beta_\sigma^N(\cdot)$ satisfies:

$$\int_{-\infty}^{\beta_\sigma^N(\cdot)} \cdots \int_{-\infty}^{\beta_\sigma^N(\cdot)} \frac{\Gamma\left(\frac{N+n}{2}\right)}{(n\pi)^{N/2}\Gamma(n/2)} \left(1 + \frac{1}{n} \sum_{i=1}^{N} t_i^2\right)^{-(N+n)/2}$$

$$\cdot \, d\, t_1 \cdots d\, t_N = P \ . \qquad \text{(II.131b)}$$

Unfortunately the multivariate t-density in the integrand cannot be factored into N marginal densities, although t_i and t_j are uncorrelated for $i \neq j$.

For the uncorrelated case, inverse tables of $\beta_\sigma^N(P,n,N)$ for P = 0.95, 0.99, n = 5(1)35, and N = 1(1)10 have been published by Krishnaiah and Armitage (1966); the same tables include the values of $\beta_\sigma^N(\cdot)$ for equicorrelated

t-distribution with $\rho_{ij} = \rho = 0.0(0.1)0.9$ $(i \neq j)$. More extensive tables, including the values $P = 0.90$ and $P = 0.975$ have been collected by the same authors in an unpublished report.

As for the cases considered previously, simple approximations can be devised, which use more accessible and extensive tables. Three of them are given and compared below.

(i) Assume independence of the future observations; then:

$$\beta_{\sigma}^{N}(P,n,N) \approx \beta_{\sigma_i}^{N}(P,n,N) = t_n(P^{1/N}) \quad . \qquad \text{(II.132a)}$$

The subscript i stands for "independence".

(ii) (Chew (1968)). From the cuboid which circumscribes the mean-centered ellipsoid with exact content P:

$$\beta_{\sigma}^{N}(P,n,N) \leq \beta_{\sigma_1}^{N}(P,n,N) = [N\, F_{N,n}(P)]^{1/2} \quad . \qquad \text{(II.132b)}$$

This value for $\beta_{\sigma}^{N}(\cdot)$ was proposed by Chew for two-sided prediction intervals; more conservatism is therefore expected when using the same approximation

in one-sided prediction, as is done here.

(iii) (Chew (1968)). Using a Bonferroni's inequality yields, for one-sided prediction intervals:

$$\beta_{\sigma}^{N}(P,n,N) \leq \beta_{\sigma_2}^{N}(P,n,N) = t_n\left(\frac{N - 1 + P}{N}\right). \qquad \text{(II.132c)}$$

The subscripts 1 and 2 in (II.132b, c) identify the approximations.

In Table 11 these three approximations are compared with the exact values for P = 0.99; n = 1, 2, 5, 10 and N = 1, 2, 5, 10, 20. Approximations (a) and (c) are quite accurate and practically identical. On the contrary approximation (b) is usually very conservative, particularly for large values of N and n.

Tables C8 - C10 in Appendix C collect the values of $\beta_{\sigma}^{N}(P,n,N)$ for P = 0.99, 0.999, 0.9999; n = 1(1)11, 14(5)29, 39, 59, ∞ and for N = 1(1)5, 10, 20. Tables C9 and C10 were obtained from equation (II.132a), and therefore their values are approximate. Instead, Table C8 collects the exact values computed by Krishnaiah and Armitage (values in parenthesis) and compares them with the approximation (II.132a) . The

APPROXIMATION (see equations (II.132))

N / n	(a), (c)	(b)	(a), (c)	(b)	(a), (c)	(b)	(a), (c)	(b)	(a), (c)	(b)
	1		2		5		10		20	
1	31.821	63.655	63.657	100.000	—	169.765	318.309	246.09	636.6	352.39
2	6.965	9.925	9.925	14.071	—	22.282	22.327	31.528	31.60	44.587
5	3.365	4.037	4.032	5.158	5.030	7.416	5.893	10.050	6.869	13.820
10	2.764	3.162	3.169	3.888	3.720	5.310	4.144	6.964	4.587	9.391

(a)

N / n	1	2	5	10	20
1	—	—	—	—	—
2	—	—	—	—	—
5	3.36	4.00	4.90	5.61	—
10	2.76	3.16	3.69	4.09	—

(b)

TABLE 11. (a) Approximate values of $\beta_{\sigma}^{N}(0.99,N,n)$ from equations (II.132)
(b) Exact values from Krishnaiah and Armitage (1966).

agreement is quite satisfactory and supports the use of equation (II.132a), particularly for large values of P (for high "reliabilities"). The associated (approximate) penalty factors for statistical uncertainty:

$$r_{\sigma}^{N}(P,n,N) = t_n(P^{1/N})/\Phi(P^{1/N})$$

are displayed in Tables C11 - C13 and plotted versus n in Figures 18, 19 and 20 for selected values of N. The penalty increases with N, although it "stabilizes" as the number of future events becomes large. Of course, for obtaining $\beta_{\sigma}^{N}(\cdot)$, the penalty factors must be multiplied by the value of $\beta_{\sigma}^{N}(\cdot)$ under perfect information: $\beta_{PI}^{N}(P,N) = \Phi(P^{1/N})$, which for given P is also an increasing function of N (see last row in Tables C8 - C10).

<u>μ unknown, σ known</u>. Again from Chew (1968) the distribution of $\underline{d} = \underline{Y} - \hat{\underline{Z}}$ ($d_i = Y_i - \frac{1}{n}\sum_{i=1}^{n} Z_i$) is multivariate normal with zero mean, variances $\sigma_{d_i}^2 = (1 + 1/n)\sigma^2$ and common correlation coefficient

$$\frac{\mathrm{Cov}[d_i,d_j]}{\sigma^2_{d_i}} = \rho_{ij} = \frac{1}{n+1} \qquad (i \neq j) \ .$$

The value of $\beta_{\mu}^{N}(\cdot)$ for which the content of the simultaneous prediction interval

$$(-\infty,\ \hat{z} + \beta_{\mu}^{N}(P,n,N) \cdot \sigma] \qquad \text{(II.133a)}$$

is P can be found as:

$$\beta_{\mu}^{N}(P,n,N) = (1 + 1/n)^{1/2}\ \beta(P,N,\ \frac{1}{n+1}\ ), \qquad \text{(II.133b)}$$

where $\beta(P,N,\rho)$ is the value of β in equation (I.41) of Chapter I, such that $\Phi_N(\rho,\beta) = P$. Of course, equation (II.133b) can be used also to find P for given $\beta_{\mu}^{N}(\cdot)$, in which case the integral (I.41) in Chapter I should be evaluated for

$$\beta = (1 + 1/n)^{-1/2}\ \beta_{\mu}^{N}(\cdot) = (1 + 1/n)^{-1/2}\ \frac{P - \hat{z}}{\sigma}$$

and for $\rho = 1/(n + 1)$.

P is given in Tables C14 and C15 for n = 1(1)4, 7, 9; β_{μ}^{N} = 2.50, 3.50, and for N = 1(1)12. The values under perfect information ($n \to \infty$) are:

$$P_{PI} = \Phi^N(\beta_\mu^N)$$

and are given in the last column of the same tables.

For values of n larger than 3 or 4 (correlation coefficient between d_i and d_j less than 0.25), an accurate and conservative approximation to r_μ^N is obtained by assuming $\rho_{ij} = 0$, in which case:

$$r_\mu^N(P,n,N) = r_\mu^N(n) = (1 + 1/n)^{1/2}.$$

The values of P which correspond to the independence assumption,

$$P_i = \Phi^N[\beta_\mu^N \cdot (1 + 1/n)^{-1/2}] \qquad \text{(II.134)}$$

are smaller than the exact values. P_i is given in parenthesis in Tables C14 and C15. The agreement of P_i with the correct values is in general good and increases with n (with decreasing correlation, as expected) and with β_μ^N (with increasing content P). For values of P very close to 1, say for $P \geq 1 - 10^{-4}$, the independence approximation (II.134) is quite satisfactory for applications to structural reliability.

For given n, N and P, and for two-sided prediction intervals Chew (1968) proposed two approximations to β_{μ}^{N}.

The first approximation, written here for one-sided predictions, is:

$$\beta_{\mu_1}^{N}(P,n,N) = [(1 + 1/n)\ \chi_N^2(P)]^{1/2}$$

where $\chi_N^2(P)$ is the P-fractile of the Chi-square distribution with N degrees of freedom. This expression for β_{μ}^{N} corresponds to a prediction cuboid in the outcome space R^N of $\underline{Y}$, circumscribing the mean-centered ellipsoid of P-content. $\beta_{\mu_1}^{N}$ is therefore a conservative approximation (i.e., larger than the exact value).

The second approximation uses a Bonferroni's inequality, yielding, for one-sided predictions:

$$\beta_{\mu_2}^{N}(P,n,N) = (1 + 1/n)^{1/2}\ \Phi\left(\frac{N - 1 + P}{N}\right) .$$

A third approximation can be generated by assuming independence between different components of $\underline{d}$. In this case:

$$\beta_{\mu_i}^{N}(P,n,N) = (1 + 1/n)^{1/2}\ \Phi(P^{1/N}) . \qquad \text{(II.135)}$$

For P = 0.995, n = 4, N = 1(1)12, Table 12 compares the exact value of β_{μ}^{N} (which can be obtained from the tables in Gupta, 1963) with the approximations $\beta_{\mu_1}^{N}$, $\beta_{\mu_2}^{N}$ and $\beta_{\mu_i}^{N}$, and with the value of β^{N} under perfect statistical information ($n \to \infty$):

$$\beta_{PI}^{N} = \Phi(P^{1/N}) \ .$$

The approximations $\beta_{\mu_2}^{N}$ and $\beta_{\mu_i}^{N}$ are both very accurate (the accuracy increases with P). Instead, the approximation β_{μ_1} is much less satisfactory.

EXAMPLE

Consider a system with deterministic resistance R, which is designed to resist a sequence of ten normal independent load events with a β value of 4.265. In other words, the resistance is chosen so that $R = \mu + 4.265\ \sigma$, where μ and σ^2 are the (estimated or known) mean and variance of the load sequence. Under perfect statistical information, μ and σ^2 are known and the reliability of the system evolves with N as indicated by the curve "n = ∞" in Figure 21.

Col.→	1	2	3	4	5
N	β_{μ}^{N}	$\beta_{\mu_i}^{N}$	$\beta_{\mu_1}^{N}$	$\beta_{\mu_2}^{N}$	β_{PI}^{N}
1	2.88	2.88	3.97	2.88	2.58
2	3.13	3.14	3.99	3.14	2.81
3	3.28	3.28	4.13	3.28	2.94
4	3.38	3.38	4.32	3.38	3.02
5	3.45	3.46	4.48	3.46	3.09
6	3.51	3.51	4.65	3.51	3.14
7	3.56	3.56	4.82	3.56	3.19
8	3.61	3.61	4.98	3.61	3.23
9	3.64	3.65	5.12	3.65	3.26
10	3.67	3.68	5.27	3.68	3.29
11	3.71	3.71	5.41	3.71	3.32
12	3.73	3.73	5.54	3.73	3.34

TABLE 12. Values of $\beta_{\mu}^{N}(0.995,4,N)$ is equation (II.133a). Exact value: col. 1; approximations: cols. 2, 3, 4; perfect statistical information: col. 5.

Suppose now that μ and σ are unknown and that R is to be chosen as

$$R = \hat{Z} + 4.265\ S\ ,$$

where $\hat{Z}$ and S^2 are the usual unbiased estimates from a given sample of size n. Using the approximation (II.130) the expected reliability of the system is:

$$R(N) = \{t_{n-1}[4.265(1 + 1/n)^{-1/2}]\}^N\ ,$$

or: for n = 5: $R(N) = 0.991154^N$,

for n = 10: $R(N) = 0.998579^N$,

for n = 20: $R(N) = 0.999734^N$.

These three reliability functions are also plotted in Figure 21a. The high sensitivity of R(N) to the available statistical information is apparent.

For the case with μ known and σ unknown the resistance is

$$R = \mu + 4.265\ S$$

where $S^2 = \frac{1}{n} \sum_{i=1}^{n} (Z_i - \mu)^2$. The evolution of the reliability with N for the sample sizes n = 5, 10, 20, ∞ is shown in Figure 21b. Again, the curve "n = ∞" corresponds to perfect knowledge of the load process parameters. R(N) was calculated from the approximate expression (corresponding to the independence assumption leading to equation (II.132c)):

$$R(N) = [t_n(4.265)]^N .$$

Finally, if μ is unknown and σ is known it is: $R = \hat{Z} + 4.265\ \sigma$. The reliability evolves with N as shown in Figure 21c for the available sample sizes: n = 1, 2, 4, 9, ∞.

Note that Figures 21a and 21b are quite similar. The results in Figure 21c, however, are much less dependent on n. This says again that it is the uncertainty in σ and not that in μ that really matters.

(b) EXTREME TYPE I SEQUENCE

The problem of finding the one-sided prediction interval of P-content for the maximum $Y_{(N)}$ of N future observations from a white extreme type I sequence with distribution

$$F_Y(y) = \exp\{-\exp[-(y-a)/b]\} \; ; \quad b > 0, \; -\infty < a < \infty. \quad \text{(II.136)}$$

was investigated by Antle and Rademaker (1972). In their study a and b were assumed unknown and had to be estimated from an available sample of size n.

The problem is of interest in reliability theory if the potentially critical events are generated by some limiting process within the domain of attraction of the distribution (II.136); for example by an exponential, normal or gamma process.

Using the maximum likelihood estimators $\hat{a}$ and $\hat{b}$ proposed by Harter and Moore (1967, 1968), the statistic $E = [(Y_{(N)}-\hat{a})/\hat{b}-\log N]$ has a distribution which depends only on n. Antle and Rademaker calculated through simulation the P-fractile $e_{P,n}$ of E for n = 20(10)70, 100, ∞; and for P = 0.90, 0.95, 0.975, 0.98, 0.99; and also for n = 20(10)70, 100, ∞; and P = 0.995. The upper P-confidence limit for

$Y_{(N)}$ was given by them in the form:

$$Y_{(N)}(P) = \hat{a} + e_{P,n}\,\hat{b} + \hat{b} \log N.$$

For perfect information ($n \to \infty$) it is: $\hat{a} \to a$; $\hat{b} \to b$, and

$$Y_{(N)}(P) = a + b[El(P) + \log N],$$

where El(P) is the P-fractile of the extreme type 1 distribution.

The penalty for imperfect information (n finite) is defined, in accordance with equation (II.30) (but not in accordance with the definition for normal sequences):

$$r^{El}_{a,b}(P,n,N) = \frac{E[\hat{a}] + E[\hat{b}](e_{P,n} + \log N)}{a + b(e_{P,\infty} + \log N)}$$

$$= \frac{g(n) + \frac{1}{h(n)}\,(e_{P,n} + \log N) + \frac{a}{b}}{\frac{a}{b} + (e_{P,n} + \log N)} \quad ,$$

where the functions g(n) and h(n), defined by equation (II.29), are tabulated by Antle and Rademaker for n = 10(10)100.

Table C16 in Appendix C gives the values of $r^{El}_{a,b}(P,n,N)$ for a/b = 1(1)10; for n = 10(10)40, 60; for P = 0.99, 0.995 (n = 10 is not considered for P = 0.995) and for N = 1(1)10(5)30, 40, 60.

Prediction intervals for values of P greater than 0.995 are not calculated by Antle and Rademaker. For P > 0.995 one might, in approximation, replace the prediction interval of P-content for N future realizations with the prediction interval of $P^{N'/N}$-content for N' > N realizations. N' can be chosen so that $P^{N'/N}$ equals 0.99 or 0.995. The approximation is believed to be on the unconservative side.

EXAMPLE

Find the 0.999 upper confidence limit for the next six realizations of an extreme type I sequence of which forty past sample values are available.

By choosing N' = 10N = 60, it is $P^{N'/N} \approx 0.99$ for which the upper confidence limit is:

$$\hat{a} + \hat{b}(e_{0.99,40} + \log 60) = \hat{a} + \hat{b}(4.79+4.08) = \hat{a} + 8.87\,\hat{b}.$$

By choosing $N' = 5N = 30$, it is $P^{N'/N} \simeq 0.995$ and the upper prediction limit becomes:

$$\hat{a} + \hat{b}(e_{0.995,40} + \log 30) = \hat{a} + \hat{b}(5.68+3.40) = \hat{a} + 9.08\,\hat{b}.$$

The simulation approach used by Antle and Rademaker is quite appealing when exact analytical prediction methods are unfeasible. If independent sequences with given marginal distribution can be generated easily (see for instance Fishman (1973) Chapter 8) and parameter estimators are available, prediction confidence levels for the maximum of the next N realizations based on n past observations can be found by simulation. The simulation approach is even more attractive when the sequence of past and future realizations has a (possibly complex) probabilistic correlation structure, for which case an exact analytical solution is very rarely workable (see Chapter III).

(c) OTHER INDEPENDENT MODELS

In white normal sequences neglecting the correlation between distinct future events caused by limited statistical information leads to slightly conservative but otherwise accurate results. This fact has important implications. If the same is true also for nongaussian sequences, then approximate but accurate simultaneous prediction intervals for the next N realizations can be constructed using the simple-prediction results in Section II.2. With respect to perfect statistical information the penalty ratio for a prediction region with content P, a past sample size n, and a future sample size N would be simply:

$$r(P,n,N) = r(P^{1/N},n,1), \qquad \text{(II.137)}$$

and the function r(P,n,1) is tabulated in Appendix A for a wide variety of distributions. In agreement with the normal case the approximation (II.137) should improve as P and/or n increase; in any case it should give slightly conservative penalties.

The use of the penalty factor $r(P^{1/N},n,1)$ does not differ from that of r(P,n,1); see Section II.2.

Using the approximation (II.137) Table C17 compares the penalty factors for white sequences with different marginal distributions, for P = 0.99, 0.999; for N = 2, 10; and for an available sample size n = 1(1)10(5)30(10)60.

For the lognormal sequence $\hat{V}_Y$ is the sample coefficient of variation. The true value is related to the variance of the normal sequence $\{\ln Y_i\}$ as:

$$\sigma_{\ln Y} = [\ln(V_Y^2 + 1)]^{1/2} .$$

For usual values of $\hat{V}_Y$ the penalty factor for lognormal sequences is considerably larger than the corresponding penalty for normal sequences. A comparison with the exponential and Gamma processes is not easy, due to the different parameters being assumed known or unknown.

(d) DISTRIBUTION-FREE PREDICTION INTERVALS

Nonparametric simultaneous prediction is considered only for completeness, in that the problem can be found treated in the literature by several authors; see for instance Danziger and Davis (1964). Their results are quoted by Chew (1968) and are reported here.

The only distributional assumption is that the population has continuous but otherwise unknown density function. With $Z_{(1)} < Z_{(2)} < \ldots < Z_{(n)}$ denoting the available order statistics from a sample of size n, Danziger and Davis show that the probability that at least K out of the next N observations will fall between $Z_{(r_1)}$ and $Z_{(r_2)}$ is:

$$P(N,r_1,r_2,K,N) = \frac{\sum_{i=K}^{N} \left(\begin{matrix} N-i+r_1+n-r_2 \\ N-i \end{matrix} \right) \left(\begin{matrix} i+r_2-r_1-1 \\ i \end{matrix} \right)}{\left(\begin{matrix} n+N \\ N \end{matrix} \right)} \qquad \text{(II.138)}$$

if $1 \leq r_1 < r_2 \leq n$. This general result can be specialized to find the probability that all the next N events will fall between $Z_{(r_1)}$ and $Z_{(r_2)}$:

$$P(n,r_1,r_2,N,N) = \left(\begin{matrix} N+r_2-r_1-1 \\ N \end{matrix} \right) \Big/ \left(\begin{matrix} n+N \\ N \end{matrix} \right) . \qquad \text{(II.139)}$$

Of course, for N = 1 one finds that the probability that the next observation will fall between $Z_{(r_1)}$ and $Z_{(r_2)}$ is:

$$P(n,r_1,r_2,1,1) = (r_2 - r_1)/(n + 1) ,$$

a familiar result from simple prediction (see Paragraph II.2.1(g)).

For instance, let n = 100, r_1 = 5, r_2 = 95. The probability that all N future observations will fall inside $(Z_{(5)}, Z_{(95)}]$ is given in the following table for N = 1(1)10.

N	P(100,5,95,N,N)	N	P(100,5,95,N,N)
1	0.891	6	0.509
2	0.795	7	0.457
3	0.710	8	0.411
4	0.635	9	0.369
5	0.568	10	0.332

TABLE 13. Simultaneous prediction probabilities for the nonparametric interval $(Z_{(5)}, Z_{(95)}]$ and n = 100.

II.3.3. Simultaneous Bayesian Prediction for Univariate Sequences

In simultaneous Bayesian prediction with a conjugate prior one can take full advantage of the initial information being equivalent to a prior sample. The size n' of the sample equals the number of observations that one would exchange for his prior information; n', together with the remaining sufficient sample statistics, characterizes fully the initial state of knowledge.

The immediate consequence of this observation is that when the prior is of the conjugate type the frequentist results apply after pooling together the actual sample statistics with the equivalent prior sample statistics. The combination rules for this "pooling operation" are given in Paragraph II.2.2 for various independent sequences. For instance equations (II.58), (II.61) and (II.68) for normal sequences, equations (II.73) and (II.74) for exponential sequences, and so on.

As for simple prediction, Bayesian simultaneous prediction intervals for noninformative priors of particular types given in Section II.2 coincide with the corresponding frequentist intervals.

EXAMPLE

Consider sampling from the normal population $N(\mu,\sigma^2)$ with both μ and σ^2 unknown. The prior density $f(\mu,\sigma^2)$ is assumed in conjugate form (see equation (II.66)) with parameters:

$$n' = 9; \quad \mu' = 80; \quad S'^2 = 16.$$

A sample of size 6 is available, with sufficient statistics:

$$n' = 6; \quad \hat{Z} = 75; \quad S^2 = 10.$$

Using equation (II.68) the prior parameters and the sufficient sample statistics are combined to give the following sufficient posterior statistics:

$$n'' = 15; \quad \mu'' = 78; \quad S''^2 = 4.375^2.$$

The left-sided prediction interval of 0.9999 content for the next ten observations is, using the approximation

(II.130) with n" = 15 in place of n (see also Table C4 in Appendix C):

$$(-\infty,\ 78 + 6.507 \cdot 4.375],\ \text{or}\ (-\infty,\ 106.468].$$

II.3.4. Applications of Simultaneous Prediction to Probabilistic Fatigue Failure

Up to this point the critical event was considered to be of the first-crossing type: given N future realizations of an independent "load" process: $Y_1, Y_2, \ldots, Y_N$, failure occurs if $\max \{Y_i\}$ or $\max_i \{|Y_i|\}$ exceeds a given "resistance" threshold. For this type of failure events an important finding was that in most situations the assumption of independence of the future realizations (regarded here as Bayesian variables) was adequate and slightly conservative.

In general, both these properties are lost in fatigue-type failure modes, as indicated by the following example. This is explained by the fact that sums of random variables have distributions which are highly affected by correlation. Another result is that the uncertainty in the mean of the process, which was often negligible for first-

crossing failure, becomes a determinant factor against fatigue-failure events.

Consider a stationary white sequence $\{Y_i\}$ of "load" variables. The damage caused by each load accumulates in the system, until failure occurs when the total accumulated damage exceeds a fatigue-resistance limit.

In a linear, memoryless (in the sense specified next) model, the damage state variable d experiences jumps at the load events occurrence times, and each jump depends exclusively on the intensity of the load which is causing it. In more sophisticated and perhaps realistic models the system performance (constitutive laws, damage accumulation rate, threshold of failure, etc.) may depend in a more complex fashion on the past history and on the present level of accumulated damage.

Consider the very simple case of the damage state variable being a linear combination of the past load intensities:

$$d(t) = \sum_{i=1}^{N(t)} \alpha_i \, Y_i . \qquad \text{(II.140)}$$

Here N(t) is the number of load "cycles" to which the system has been exposed up to time t, and $\alpha_1,\ldots,\alpha_{N(t)}$ are given constants. For fixed N(t), d(t) is a random variable which depends linearly on the intensity of the events in [0,t]. If d_{MAX} denotes the level of damage at which failure occurs, the reliability at time t is

$$R[N(t)] = \Pr\{d(t) < d_{MAX}\} = \Pr\{\sum_{i=1}^{N(t)} \alpha_i Y_i < d_{MAX}\} .$$

To illustrate the effect of parametric statistical uncertainty of the load sequence on R(t), suppose that $\{Y_i\}$ is a white stationary normal process, with $Y_i \sim N(\mu,\sigma^2)$.

If <u>μ and σ are known</u> and $\alpha_i = \alpha$ for all i, it is:

$$R_{PI}(N) = \Phi[(\frac{d_{MAX}}{\alpha} - N \cdot \mu)/\sigma\sqrt{N})] . \qquad \text{(II.141a)}$$

If <u>μ is unknown, σ is known</u>, and a sample of size n is available from the load population with sample mean $\hat{Z}$:

$$R_{\mu}(N) = \Phi[(\frac{d_{MAX}}{\alpha} - N\,\hat{Z})/\sigma\sqrt{\frac{N(N+n)}{n}}] \qquad \text{(II.141b)}$$

If μ is known and σ is unknown one finds (using well-known properties of the multivariate t-distribution; see for instance Box and Tiao (1973), p 118):

$$R_{\sigma}(N) = t_n[(\frac{d_{MAX}}{\alpha} - N\,\mu)/S\sqrt{N}] \ , \qquad \text{(II.141c)}$$

where S^2 is the usual unbiased variance estimator.

Finally, if μ and σ are both unknown and $\hat{Z}$ and S^2 are unbiased estimates of μ and σ^2:

$$R_{\mu,\sigma}(N) = t_{n-1}[(\frac{d_{MAX}}{\alpha} - N\,\hat{Z})/S\sqrt{\frac{N(N+n)}{n}}] \ . \qquad \text{(II.141d)}$$

In equations (II.141c,d) $t_{\nu}(\cdot)$ is the CDF of the t-distribution with ν degrees of freedom.

All the expressions above are obtained readily from the simultaneous prediction results in Paragraph II.3.2. Using these expressions for design, one can choose the

resistance level d_{MAX} which corresponds to a prefixed probability of failure P within N load events. For a design equation

$$\frac{d_{MAX}}{\alpha} = N \cdot \mu + \beta \sigma\sqrt{N} \qquad \text{(II.142)}$$

(or analogously, with μ replaced by $\hat{Z}$ and for σ replaced by S if these parameters are unknown) the penalty factors on β with respect to its value β_{PI} under perfect knowledge of μ and σ are:

for <u>μ unknown and σ known</u>:

$$r_{\mu}(N,n) = \sqrt{\frac{N+n}{n}} \quad \text{(independent of P);} \qquad \text{(II.143a)}$$

for <u>μ known and σ unknown</u>:

$$r_{\sigma}(P,n) = \frac{t_n(P)}{\Phi(P)} \quad \text{(independent of N);} \qquad \text{(II.143b)}$$

for <u>μ and σ unknown</u>:

$$r_{\mu,\sigma}(P,N,n) = \sqrt{\frac{N+n}{n}} \quad \frac{t_{n-1}(P)}{\Phi(P)} \; . \qquad \text{(II.143c)}$$

When N = 1 these penalty factors coincide with those found in Paragraph II.2.1 for simple prediction.

Typically, in "high cycle" fatigue problems both N and n are large. This means that $r_\sigma(P,n)$ is very close to 1. On the contrary the penalty for unknown location parameter, $r_\mu(N,n)$, may be much larger than 1, depending on the ratio N/n. This is, in a sense, a situation opposite to that of simple prediction (N = 1) and of simultaneous prediction with a first-crossing failure mode. In the last two cases σ being unknown generally has a much larger penalty than μ being unknown.

Numerical results. Consider the resistance design equation (the same as in equation (II.142), with N_d in place of N):

$$\frac{d_{MAX}}{\alpha} = N_d\,\mu + \beta\,\sigma\sqrt{N_d} \quad , \tag{II.144}$$

where N_d is a given design number of load repetitions and β is a safety coefficient for the case of perfect knowledge of the load process parameters. For selected values of N_d, μ (or $\hat{Z}$), σ (or S) and β, the arguments of the cumulative normal and t-distribution in equations (II.141) are plotted

versus N in Figures 22a,b,c. $V = \sigma/\mu$ is the coefficient of variation of the loads. In each figure the steepest line is the argument of Φ and of t_n in equations (II.141a,c), while the remaining lines give the argument of Φ and of t_{n-1} in equations (II.141b,d), for different values of n. In each case the failure probability can be read on the appropriate lateral scale (for very large n, the scales of t_n and t_{n-1} almost coincide with the normal scale, and are omitted).

For instance one finds (see Figure 22a) that corresponding to a reliability $R_{PI}(100) = 0.999$ for perfect knowledge it is:

n	$R_\mu(100)$	$R_\sigma(100)$	$R_{\mu,\sigma}(100)$
10	0.817	0.9933	0.805
30	0.925	0.9973	0.920
70	0.973	0.9981	0.971

The (common) value of N for which all the arguments plotted in Figures 22 are 0 and the reliability is 0.5 is $N(0.5) = N_d + \beta \cdot V \cdot \sqrt{N_d}$, increasing linearly with β and, for

fixed β, with the coefficient of variation. The fact that for $N = N_d$ the argument values (β if μ is known, $\beta(1+N/n)^{-1/2}$ if μ is unknown) do not depend on V and are linear in β implies that in the vicinity of N_d the sensitivity of $R_{(\cdot)}(N)$ to N decreases with V. In turn this says that if N is random its uncertainty has more effect on the system reliability for smaller V.

A quadratic model of damage accumulation. The linearity of the damage increment in the load intensity postulated by the foregoing model may be rather unrealistic. Often the damage increment caused by one load cycle increases more than linearly with the intensity of the load.

Consider the quadratic damage accumulation law:

$$d(N) = \sum_{i=1}^{N} Y_i^2 .$$

As before, $\{Y_i\}$ is a sequence of normal variates from an independent process. The resistance d_{MAX} is known.

In the case of perfect knowledge of the process parameters μ and σ^2 the distribution of $d(N)/\sigma^2$ is noncentral Chi-square with N degrees of freedom and noncentrality parameter $\lambda = N\,\mu^2/\sigma^2 = N/V^2$, denoted $\chi'^2_N(\lambda)$. The properties of this distribution are reviewed in Johnson and Kotz (1970)

Chapter 28. Therefore, for perfect knowledge of the process parameters, the reliability is

$$R_{PI}(N) = \chi'^{2}_{N}(N/V^2, d_{MAX}/\sigma^2), \qquad \text{(II.145)}$$

where $\chi'^{2}_{N}(\lambda,\cdot)$ is the CDF of the $\chi'^{2}_{N}(\lambda)$ distribution.

For convenience of language, the cases with unknown parameters are treated using the Bayesian terminology. In all cases the prior distribution is assumed noninformative (see Paragraph II.2.2), so that the same conclusions would be reached by following the frequentist approach.

μ unknown, σ known. If a sample of size n is available, the next N realizations have multinormal predictive distribution with mean

$$\underline{\mu} = \hat{Z} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad ; \quad \hat{Z} = \text{sample mean},$$

and covariance matrix

$$\underline{\Sigma} = \begin{bmatrix} (1+1/n)\sigma^2 & \sigma^2/n & \cdots & \sigma^2/n \\ \sigma^2/n & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \sigma^2/n \\ \sigma^2/n & \cdots & \sigma^2/n & (1+1/n)\sigma^2 \end{bmatrix}$$

Equivalently, the prediction vector $\underline{Y} = [Y_1,\ldots,Y_N]'$ can be written:

$$Y_i = Y_i^* + Y_o \qquad i = 1,\ldots,N$$

where $Y_i^* \sim N(\hat{Z};\sigma^2)$; $Y_o = (\mu-\hat{Z}) \sim N(0;\sigma^2/n)$,

$$\mathrm{Cov}[Y_i^*,Y_j^*] = \mathrm{Cov}[Y_i^*,Y_o] = 0 \quad \text{for } i,j = 1,\ldots,N.$$

Then one can write:

$$d(N) = \sum_{i=1}^{N} (Y_i^* + Y_o)^2 = \sum_{i=1}^{N} Y_i^{*2} + N\, Y_o^{\,2} + 2Y_o \sum_{i=1}^{N} Y_i^* . \quad \text{(II.146)}$$

The three terms in the right hand side are analyzed separately and their contributions to the uncertainty of d(N) are compared.

When divided by σ^2, the <u>first term</u> in the right hand side of equation (II.146), $\sum_{i=1}^{N} Y_i^{*2}$, is distributed like $\chi_N'^2(N\,\hat{Z}^2/\sigma^2) = \chi_N'^2(N/\hat{V}^2)$. Since the mean and the variance of $\chi_\nu'^2(\lambda)$ are:

$$E[\chi_\nu'^2(\lambda)] = \lambda + \nu; \quad Var[\chi_\nu'^2(\lambda)] = 2(\nu + 2\lambda) ,$$

it is :

$$E\left[\sum_{i=1}^{N} Y_i^{*2}\right] = \sigma^2(N + N/\hat{V}^2) = N(\sigma^2 + \hat{Z}^2) \quad ;$$

$$Var\left[\sum_{i=1}^{N} Y_i^{*2}\right] = 2\sigma^4(N + 2N/\hat{V}^2) = 2N\sigma^2(\sigma^2 + 2\hat{Z}^2). \quad \text{(II.147a)}$$

The distribution of $\frac{n}{\sigma^2} Y_o^2$ is χ_1^2; therefore the second term in the right hand side of equation (II.146), $N\ Y_o^2$, has mean and variance:

$$E[N\ Y_o^2] = \frac{N}{n}\sigma^2; \quad Var[N\ Y_o^2] = 2\ \frac{N}{n}\ \sigma^2. \tag{II.147b}$$

Consider now the third term, $2\ Y_o \sum_{i=1}^{N} Y_i^*$. Since N is in general quite large (meaning a small coefficient of variation for $\sum_{i=1}^{N} Y^*_i$) and Y_o is a zero-mean variable, the distribution of this term is dominated by the (normal) distribution of Y_o. This is even more so if Y_i^* has itself small sample coefficient of variation: $\hat{V} = \sigma/\hat{Z}$.

The distribution of $\sum_{i=1}^{N} Y_i^*$ is $N(N\ \hat{Z};\ N\ \sigma^2)$; the distribution of $2\ Y_o$ is $N(0;\ 4\sigma^2/n)$. Since Y_o and Y_j^* are independent for j = 1,...,N, it is:

$$E[2\ Y_o \sum_{i=1}^{N} Y_i^*] = 0$$

$$Var[2\ Y_o \sum_{i=1}^{N} Y_i^*] = N^2\ \hat{Z}^2(4\sigma^2/n) + N\ \sigma^2(4\sigma^2/n)$$

$$= \frac{4}{n} N \sigma^2 (N \hat{Z}^2 + \sigma^2)$$

$$\approx \frac{4}{n} N^2 \sigma^2 \hat{Z}^2 \qquad \text{(large N)}. \qquad \text{(II.147c)}$$

As observed earlier, the shape of the distribution is strongly affected by the shape of the distribution of Y_o (normal). Therefore, in approximation:

$$2 Y_o \sum_{i=1}^{N} Y_i^* \sim N(0, \frac{4}{n} N^2 \sigma^2 \hat{Z}^2).$$

A comparison of the variances, equations (II.147a, b,c), shows that the variance of the third term dominates for moderate to large N values. Instead, the mean value of d(N) in equation (II.146) is mostly contributed by the first term. Therefore for large N we find again, as in the linear damage accumulation model, that the mean of d(N) is not significantly affected by the uncertainty of the parameters, while the variance of d(N) depends strongly on the uncertainty of the mean (here on the posterior variance

$\mathrm{Var}[\mu|\hat{Z}] = \mathrm{Var}[Y_o] = \sigma^2/n)$. This corresponds to linearizing the quadratic model as follows:

$$d(N) = E[d(N)|\mu=\hat{Z}] + E\left[\left.\frac{\partial\{d(N)|\mu\}}{\partial\mu}\right|_{\mu=\hat{Z}}\right]\cdot(\mu-\hat{Z}). \qquad \text{(II.148)}$$

According to equation (II.148) all the uncertainty of $d(N)$ comes from the term $Y_o = (\mu-\hat{Z}) \sim N(0,\sigma^2/n)$. The approximation (II.148) improves with N and is believed to be applicable also to damage accumulation laws other than the quadratic one (and perhaps also to nonnormal sequences). However, for damage laws like

$$d(N) = \sum_{i=1}^{N} Y_i^{\ell} \quad , \quad \ell \text{ given},$$

the accuracy of the linearized model (II.148) decreases with ℓ. In particular, for the combination: large-ℓ, moderate-N one should account also for the uncertainty in Y_i^* and for higher order terms in Y_o.

μ known, σ unknown. Given an independent sample of size n, the posterior distribution of $\underline{Y}$ is multivariate-t with n degrees of freedom, mean

$$\underline{\mu} = \mu \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

and covariance matrix

$$\underline{\Sigma} = \frac{n}{n-2} S^2 \underline{I}_N \quad ; \quad S^2 = n^{-1} \sum_{i=1}^{n} (Z_i - \mu)^2 .$$

In analogy with equation (II.146) one can express the damage after N load cycles as:

$$d(N) = \sum_{i=1}^{N} Y_i^2 = \sum_{i=1}^{N} (Y_i - \mu + \mu)^2$$

$$= \sum_{i=1}^{N} (Y_i - \mu)^2 + N \mu^2 + 2\mu \sum_{i=1}^{N} (Y_i - \mu) .$$

(II.149)

The first term in the right hand side of equation (II.149) is related to an F-variate; in fact

$$\frac{\sum_{i=1}^{N} (Y_i - \mu)^2}{N S^2} \sim F_{N,n}$$

Therefore:

$$E\left[\sum_{i=1}^{N} (Y_i - \mu)^2\right] = N S^2 \frac{n}{n-2} \quad ;$$

$$\mathrm{Var}\left[\sum_{i=1}^{N} (Y_i - \mu)^2\right] = N^2 S^4 \frac{2n^2(N + n - 2)}{N(n-2)^2(n-4)} \quad . \qquad \text{(II.150a)}$$

The second term in (II.149) is a known constant, and the third term is related to a t-variate, being:

$$\frac{\sum_{i=1}^{N} (Y_i - \mu)}{S\sqrt{N}} \sim t_n \quad .$$

Therefore:

$$E[2\mu \sum_{i=1}^{N} (Y_i - \mu)] = 0 \quad ;$$

$$\text{Var}[2\mu \sum_{i=1}^{N} (Y_i - \mu)] = 4\mu^2 \, N \, S^2 \, \frac{n}{n-2} = 4N \, \frac{S^4}{\hat{V}^2} \, \frac{n}{n-2}, \quad \text{(II.150b)}$$

where $\hat{V} = S/\mu$.

From these results and from equation (II.149) the expected damage is

$$E[d(N)] = N(\mu^2 + S^2 \, \frac{n}{n-2}) \; .$$

For large N (e.g., N = 10,000), relatively small n (say n < 100) and large $\hat{V}$ (say $\hat{V} > 0.5$), the variance of the first term in equation (II.149) dominates, increasing with the square of N. For small N (e.g., N = 100), relatively large n (say n > 50) and small coefficients of variation (say $\hat{V} < 0.2$), the variance of the third term dominates, with linear dependence on N.

For the first combination of parameters the variance of d(N) becomes comparable (in order of magnitude, although remaining generally smaller) with the variance of the damage when μ is unknown and σ is known. For the second combination the variance of d(N) when σ is unknown is substantially smaller than when μ is unknown. These results are in qualitative agreement with earlier conclusions for the linear damage accumulation model. The case with both location and scale parameters unknown is expected to behave roughly like the case with only μ unknown.

As a final consideration, one might choose other types of independently additive damage functions, such as

$$d(N) = \sum_{i=1}^{N} g(Y_i) \ , \qquad \text{(II.151)}$$

where $g(\cdot)$ has, say, one of the forms sketched in Figure 23.

With reference to Figure 23, whenever $Y_i \leq Y_E$ the damage level does not increase, generating an "endurance limit" at Y_E; instead, a discrete jump occurs if $Y_i > Y_E$. In Figure 23(a) the jump is linear in $(Y_i - Y_E)$, in 23(b) it is constant (d(N) counts the number of exceedances of the threshold value Y_E in N trials), and in Figure 23(c) the damage contributed by the i^{th} cycle is constant beyond the

transition interval $[Y_E, Y_M]$. In case (b), if the "crossing events": $Y_i > Y_E$ are rare one can use the approximation of independent trials which proved itself accurate for first-crossing failure events (=>predictive Bernoulli process, see Paragraph II.3.2).

For the damage accumulation law (II.151) and the functions $g(\cdot)$ in Figure 23 the importance of μ being unknown should be somewhere between the cases of fatigue-failure and of first-crossing failure, depending on the endurance limit Y_E.

REFERENCES

1. Aitchison, J. (1964): "Bayesian Tolerance Regions". J. Royal Statist. Soc. B, 26, pp 161-210.

2. Aitchison, J. (1966): "Expected-Cover and Linear-Utility Tolerance Intervals". J. Royal Statist. Soc. B, 28, pp 57-62.

3. Aitchison, J. and Sculthorpe, D. (1965): "Some Problems of Statistical Prediction". Biometrika, 52, 3 and 4, 469.

4. Anderson, R.G. (1972): "Reduction of Uncertainty in Structural Design and Decision Making by First-Order Bayesian Information Analysis". M.S. Thesis, Dept. of Civil Engineering, M.I.T.

5. Ang, A.H.S. (1973): "Structural Risk Analysis and Reliability-based Design". Journal of the Structural Division, ASCE, Vol 99, No ST9, pp 1891-1909.

6. Antle, C.E. and Rademaker, F. (1972): "An Upper Confidence Limit on the Maximum of m Future Observations from a Type I Extreme Distribution". Biometrika, 59, pp 475-477.

7. Benjamin, J.A. and Cornell, C.A. (1971): Probability, Statistics and Decision Theory for Civil Engineers, McGraw-Hill, New York.

8. Box, G.E.P. and Tiao, G.C. (1973): Bayesian Inference in Statistical Analysis. Addison-Wesley.

9. Chew, V. (1966): "Confidence, Prediction and Tolerance Regions for the Multivariate Normal Distribution". Ann. Statist. Ass., 61, pp 605-617.

10. Chew, V. (1968): "Simultaneous Prediction Intervals". Technometrics, 10, pp 323-330.

11. Cox, D.R. (1954): "The Mean and Coefficient of Variation of Range in Small Samples from Nonnormal Populations". Biometrika, 41, pp 469-481.

12. Danziger, L. and Davis, S.A. (1964): "Tables of Distribution-Free Tolerance Limits". Ann. Math. Statist., 35, pp 1361-1365.

13. David, H.A. (1954): "The Distribution of Range in Certain Nonnormal Populations". Biometrika, 41, pp 463-468.

14. DeGroot, M.H. (1970): Optimal Statistical Decisions. McGraw-Hill, New York.

15. Ditlevsen, O. (1973): "Structural Reliability and the Invariance Problem". Report NO 22, Solid Mechanics Division, University of Waterloo, Ontario, Canada.

16. Draper, N.R. and Guttman, I. (1968a): "Some Bayesian Stratified Two-phase Sampling Results". Biometrika, 55, pp 131-140.

17. Draper, N.R. and Guttman, I. (1968b): "Bayesian Stratified Two-phase Sampling Results: K Characteristics". Biometrika, 55, pp 587-588.

18. Dunsmore, I.R. (1966): "A Bayesian Approach to Classification". J. Royal Statist. Soc. B, 28, pp 568-577.

19. Dunsmore, I.R. (1968): "A Bayesian Approach to Calibration". J. Royal Statist. Soc. B, 30, pp 396-405.

20. Dunsmore, I.R. (1969): "Regulation and Optimization". J. Royal Statist. Soc. B, 31, pp 160-170.

21. Fedorov, V.V. (1972): Theory of Optimal Experiments. Academic Press, New York.

22. Feller, W. (1967): An Introduction to Probability Theory and its Applications: Vol I. (third edition), Wiley & Sons, New York.

23. Ferguson, T.S. (1973): "A Bayesian Analysis of Some Nonparametric Problems". The Annals of Statistics, 1, No 2, pp 309-330.

24. Fraser, D.A.S. (1951): "Sequentially Determined Statistically Equivalent Blocks". Ann. Math. Statist., 22, pp 372-381.

25. Fraser, D.A.S. (1953): "Nonparametric Tolerance Regions". Ann. Math. Statist., 24, pp 44-55.

26. Fraser, D.A.S. and Guttman, I. (1956): "Tolerance Regions". Ann. Math. Statist., 27, pp 162-179.

27. Geisser, S. (1965): "Bayesian Estimation in Multivariate Analysis". Ann. Math. Statist., 36, pp 150-159.

28. Goodman, L.A. and Madansky, A. (1962): "Parameter-free and Nonparametric Tolerance Limits: the Exponential Case". Technometrics, 4, pp 75-96.

29. Guenther, W.C. and Terragno, P.J. (1964): "A Review of the Literature on a Class of Coverage Problems". Ann. Math. Statist., 35, pp 232-260.

30. Gumbel, E.J. (1958): Statistics of Extremes, Columbia University Press, New York.

31. Gupta, S.S. (1960): "Order Statistics from Gamma Distribution". Technometrics, 2, pp 243-262.

32. Gupta, S.S. (1963): "Probability Integrals of Multivariate Normal and Multivariate t". Ann. Math. Statist., 34, pp 792-828.

33. Gupta, S.S. and Sobel, M. (1957): "On a Statistic which Arises in Selection and Ranking Problems". Ann. Math. Statist., 28, pp 957-967.

34. Guttman, I. (1959): "Optimum Tolerance Regions and Power when Sampling from Some Non-normal Universes". Ann. Math. Statist., 30, pp 926-938.

35. Guttman, I. (1970): Statistical Tolerance Regions: Classical and Bayesian. Griffin's Statistical Monographs & Courses, No 26, Stuart Ed., Hafner Publ. Co., London.

36. Hahn, G.J. (1970): "Additional Factors for Calculating Prediction Intervals for Samples from a Normal Distribution". J. Am. Statist. Ass., 65, pp 1668-1676.

37. Hahn, G.J. and Hendrickson, R.W. (1971): "A Table of Percentage Points of the Distribution of the Largest Absolute Value of K Student t Variates and its Applications". Biometrika, 58, pp 323-332.

38. Hampton, J.M., Moore, P.G. and Thomas, H. (1973): "Subjective Probability and its Measurement". J. Royal Statist. Soc. A, 136, Part 1, pp 21-42.

39. Harter, H.L. (1919): New Tables of the Incomplete Gamma - Function Ratio and of Percentage Points of the Chi-square and Beta Distributions. U.S. Aerospace Research Labs, Wright-Patterson Air Force Base, Ohio.

40. Harter, H.L. and Moore, A.H. (1967): "A Note on Estimation from a Type 1 Extreme Value Distribution". Technometrics, 9, pp 325-331.

41. Harter, H.L. and Moore, A.H. (1968): "Maximum Likelihood Estimation from Doubly Censored Samples of the Parameters of the First Asymptotic Distribution of Extreme Values". J. Am. Statist. Ass., 63, pp 889-901.

42. Jeffreys, H. (1961): Theory of Probability. Oxford University Press, third edition.

43. Johnson, L.J. and Kotz, S. (1970): Distributions in Statistics: Continuous Univariate Distributions - 2. Houghton Mifflin, New York.

44. Johnson, L.J. and Kotz, S. (1972): Distributions in Statistics: Continuous Multivariate Distributions. Wiley & Sons, New York.

45. Kemperman, J.H.B. (1956): "Generalized Tolerance Limits". Ann. Math. Statist., 27, pp 180-186.

46. Krishnaiah, P.R. and Armitage, J.V. (1966): "Tables for Multivariate t-Distributions". Sankhyā B, 28, pp 31-56.

47. Lieberman, G.J. (1961): "Prediction Regions for Several Predictions from a Single Regression Line". Technometrics, 3, pp 21-27.

48. Lindley, D.V. (1965): Introduction to Probability and Statistics. Part 2: Inference. Cambridge University Press.

49. Lindley, D.V. (1971): Bayesian Statistics, A Review. SIAM, Philadelphia, Pa.

50. Novick, M.R. (1962): "A Bayesian Indifference Postulate". Institute of Statistics Mimeographic Series, No 319, University of North Carolina.

51. Pearson, K. (1934): Tables of the Incomplete Beta Function. Cambridge University Press.

52. Pierce, D.A. (1973): "Fiducial, Frequency and Bayesian Inference on Reliability for the Two Parameter Negative Exponential Distribution". Technometrics, 15, pp 249-253.

53. Proshan, F. (1953): "Confidence and Tolerance Interval for the Normal Distribution". J. Am. Statist. Ass., 48, pp 550-564.

54. Raiffa, H. and Schlaifer, R. (1961): Applied Statistical Decision Theory. Harvard University Press.

55. Savage, L.J. (1962): The Foundations of Statistical Inference: a Discussion. Methuen, London.

56. Savage, L.J. (1972): English summary of: DeFinetti, B. and Savage, L.J., "Sul modo di scegliere le probabilità iniziali", Biblioteca del Metron, Series C, vol 1, 1962, pp 81-147, Rome. The summary is in DeFinetti, B.: Probability, Induction and Statistics, Chapter 8, Wiley, New York.

57. Sidak, Z. (1968): "On Multivariate Normal Probabilities of Rectangles: their Dependence on Correlation". Ann. Math. Statist., 39, pp 1425-1434.

58. Singh, C. (1967): "On the Extreme Value and Range of Samples from Nonnormal Populations". Biometrika, 54, pp 541-550.

59. Singh, C. (1972): "Order Statistics from Nonnormal Populations". Biometrika, 59, pp 229-233.

60. Slepian, D. (1962): "The One-sided Barrier Problem for Gaussian Noise". Bell System Tech. J., 41, pp 463-501.

61. Smallwood, R.D. (1968): "A Decision Analysis of Model Selection". IEEE Trans. on Systems Science and Cybernetics, pp 333-342.

62. Thatcher, A.R. (1964): "Relationships between Bayesian and Confidence Limits for Prediction". J. Royal Statist. Soc. B, 26, pp 176-210.

63. Tippett, L.H.C. (1925): "On the Extreme Individuals and the Range of Samples taken from a Normal Population". Biometrika, 17, p 364.

64. Tong, Y.L. (1970): "Some Probability Inequalities of Multivariate Normal and Multivariate t". J. Am. Statist. Ass., 65, pp 1243-1247.

65. Tukey, J.W. (1947): "Nonparametric Estimation II: Statistically Equivalent Blocks and Tolerance Regions - the Continuous Case". Ann. Math. Statist., 18, pp 529-539.

66. Veneziano, D. (1974a): "Invariance in Second Moment Reliability". In Res. Report R74-33. Dept. of Civil Engineering, Mass. Institute of Technology.

67. Veneziano, D. (1974b): "Statistical Inference in Second Moment Reliability". In Res. Report R74-33, Dept. of Civil Engineering, M.I.T.

68. Veneziano, D. (1974c): "Note on an Inductive Approach to Safety". Submitted to: Proceedings of the American Society of Civil Engineers, J. Structural Division.

69. Wald, A. (1943): "An Extension of Wilk's Method for Setting Tolerance Limits". Ann. Math. Statist., 14, pp 45-55.

70. Wilks, S.S. (1941): "Determination of Sample Sizes for Setting Tolerance Limits". Ann. Math. Statist., 12, pp 91-96.

71. Wood, E.F., Rodriguez-Iturbe, I., Schaake, J.C. (1974): "The Methodology of Bayesian Inference and Decision Making Applied to Extreme Hydrologic Events". Report No 178, Dept. of Civil Engineering, Mass. Institute of Technology, January 1974.

72. Zellner, A. (1971): An Introduction to Bayesian Inference in Econometrics. Wiley & Sons, New York.

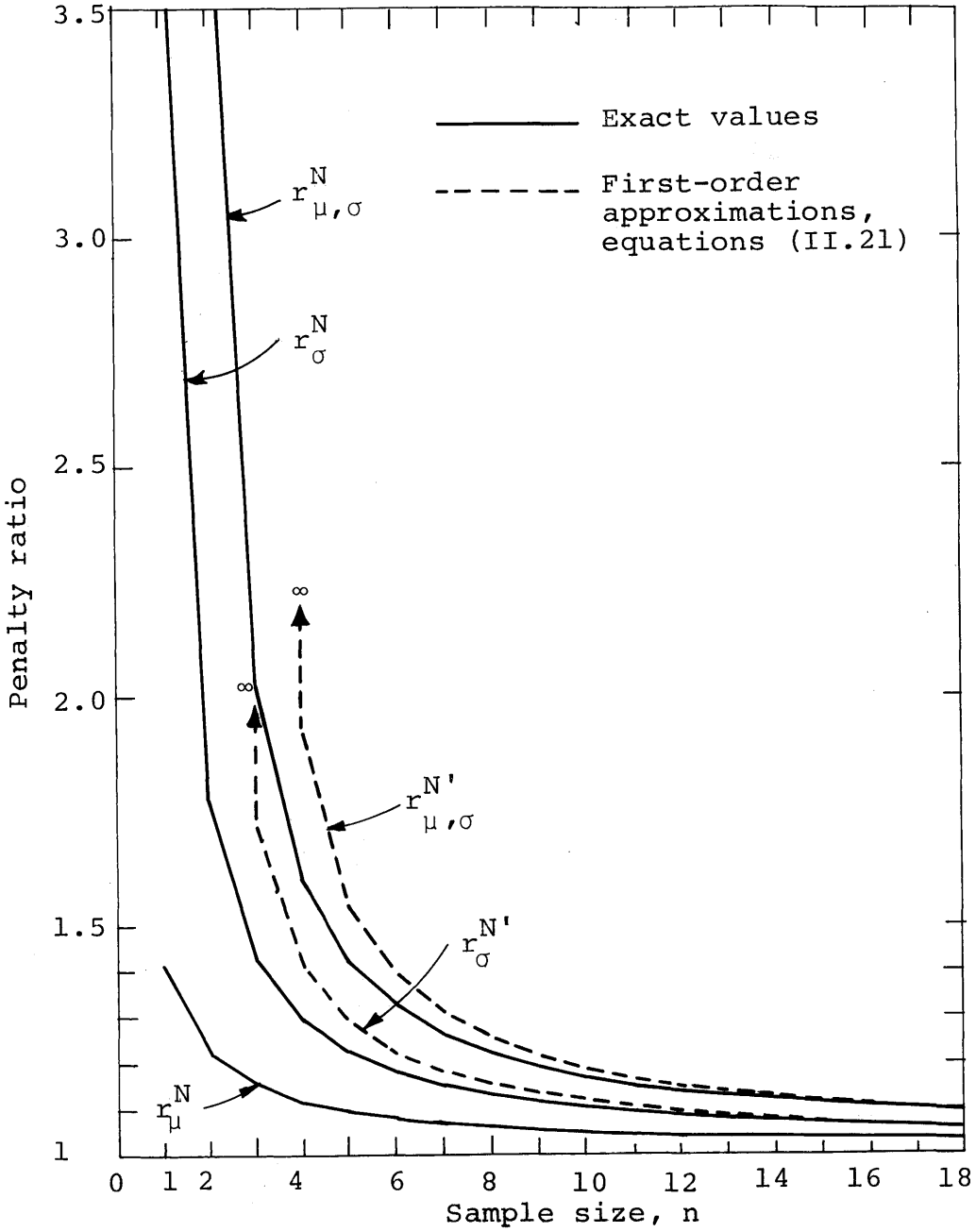


Figure 1. Penalty ratios for univariate normal sequences; central prediction intervals of 0.90 content.

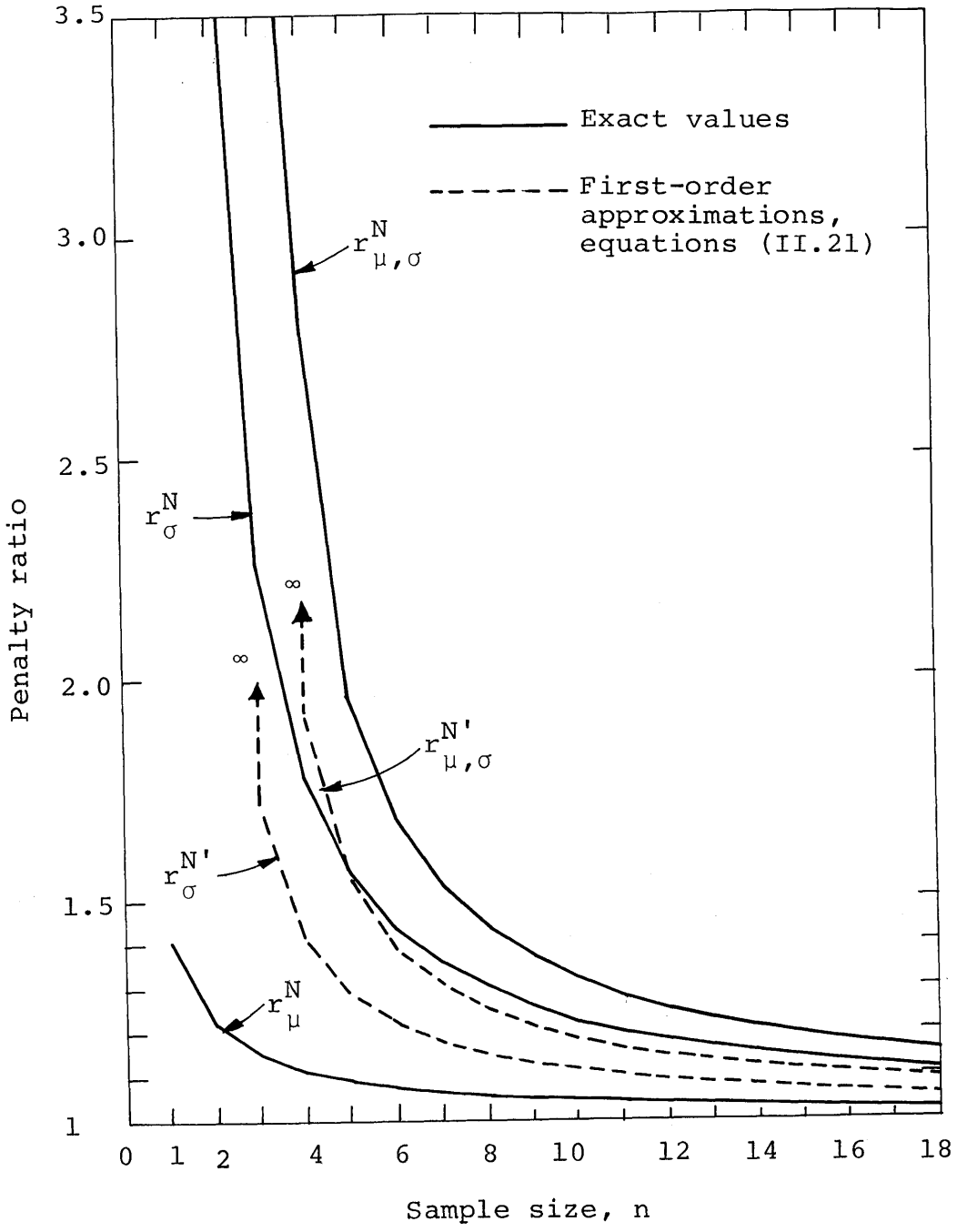


Figure 2. Penalty ratios for univariate normal sequences; central prediction intervals of 0.99 content.

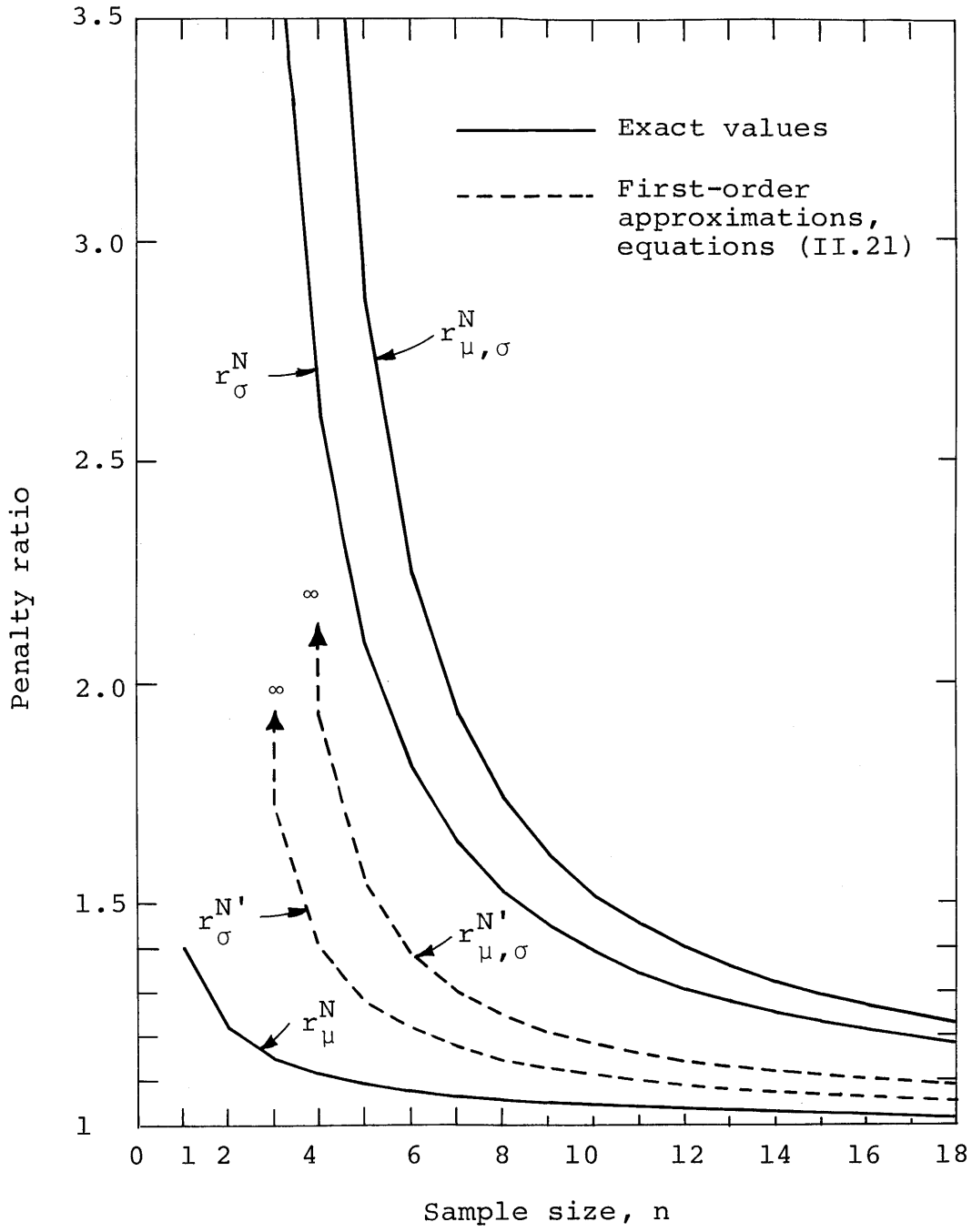


Figure 3. Penalty ratios for univariate normal sequences; central prediction intervals of 0.999 content.

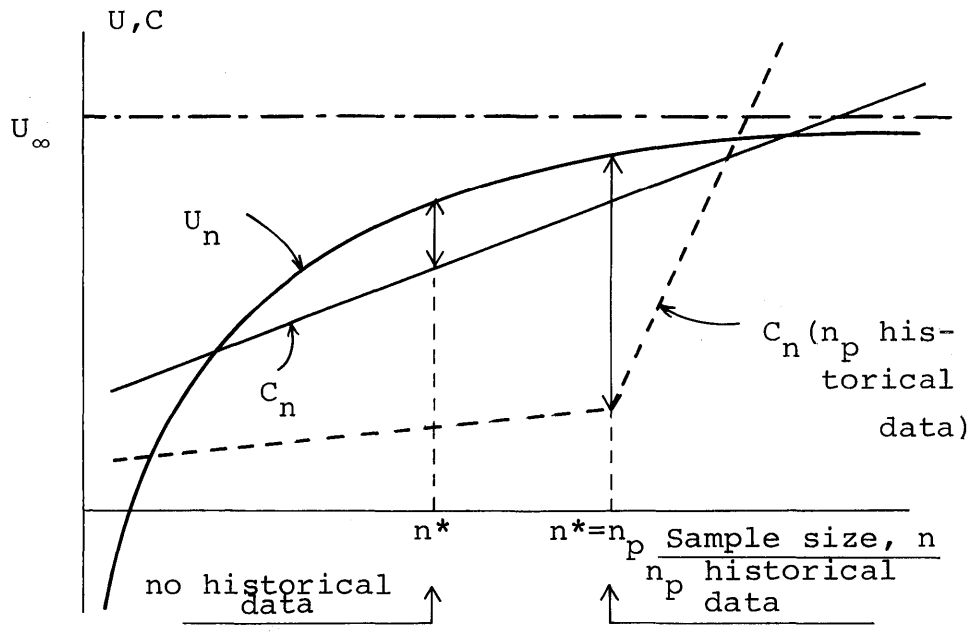


Figure 4. Choice of the optimal sample size. U_n = utility of a sample of size n; C_n = cost of a sample of size n; n* = optimal sample size.

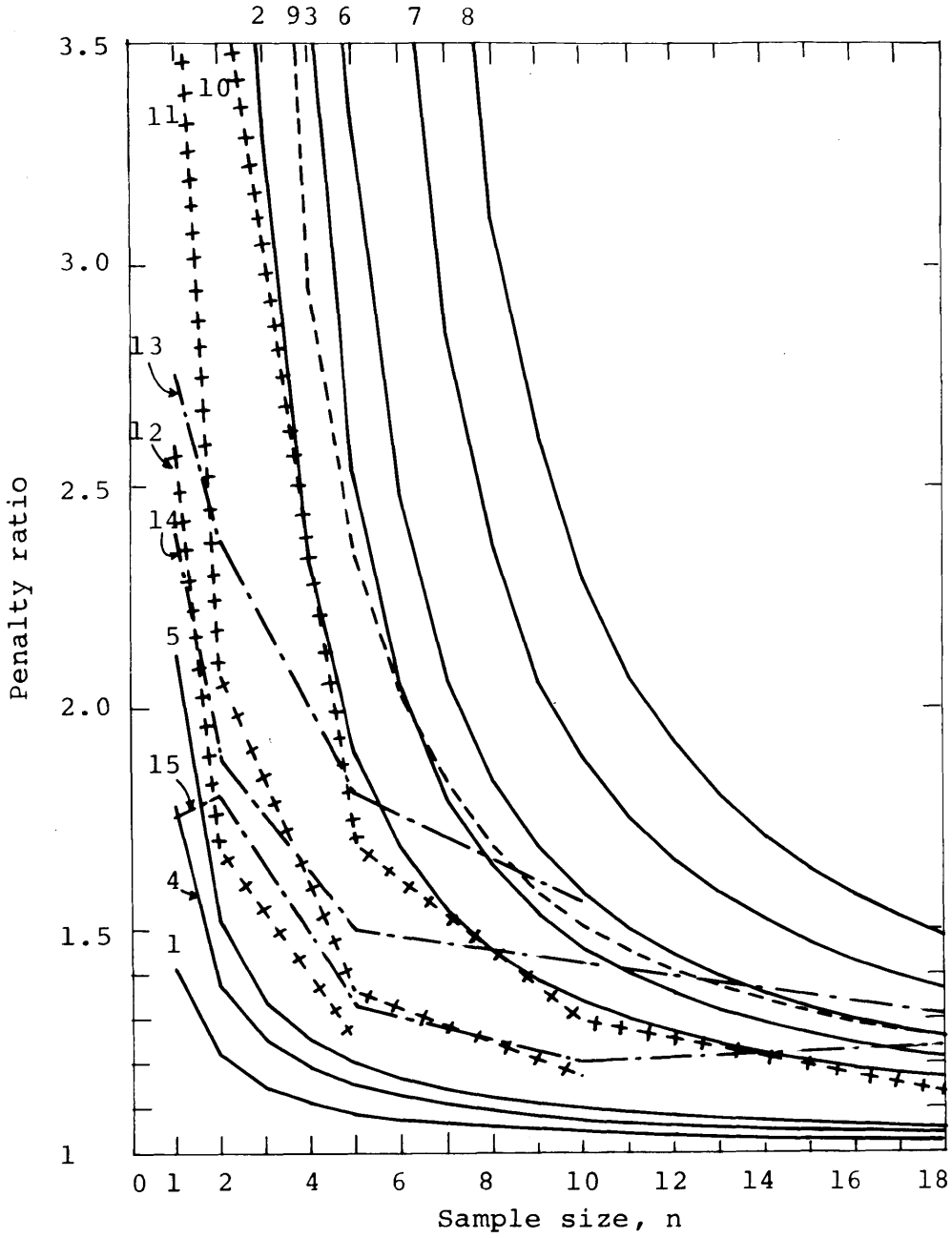


Figure 5. Penalty ratios for prediction intervals of 0.999 content. Sampling is from independent Normal, Log-normal, Exponential, Gamma and Poisson sequences. The following table associates each curve with a population distribution and with a set of unknown parameters.

CURVE (Fig. 5)	SAMPLING DISTRIBUTION AND UNKNOWN PARAMETERS (in parenthesis) (*)
1	$N(\mu)=LN(\mu_Z,\hat{V}\to 0)$
2	$N(\sigma)=LN(\sigma_Z,\hat{V}\to 0)$
3	$N(\mu,\sigma)=LN(\mu_Z,\sigma_Z,\hat{V}\to 0)$
4	$LN(\mu_Z,\hat{V}=0.3)$
5	$LN(\mu_Z,\hat{V}=0.5)$
6	$LN(\mu_Z,\sigma_Z,\hat{V}=0.1)$
7	$LN(\mu_Z,\sigma_Z,\hat{V}=0.3)$
8	$LN(\mu_Z,\sigma_Z,\hat{V}=0.5)$
9	$EX(\sigma)=G(\lambda,k=1)$
10	$G(\lambda,k=2)$
11	$G(\lambda,k=5)$
12	$G(\lambda,k=10)$
13	$P(\lambda,t_f/t_p=2)$
14	$P(\lambda,t_f/t_p=1)$
15	$P(\lambda,t_f/t_p=0.5)$

Association of the plots in Figure 5 with the population distribution. Unknown parameters and parameters with given values are indicated in parenthesis.

(*) N = Normal, LN = Lognormal, EX = Exponential
G = Gamma, P = Poisson.

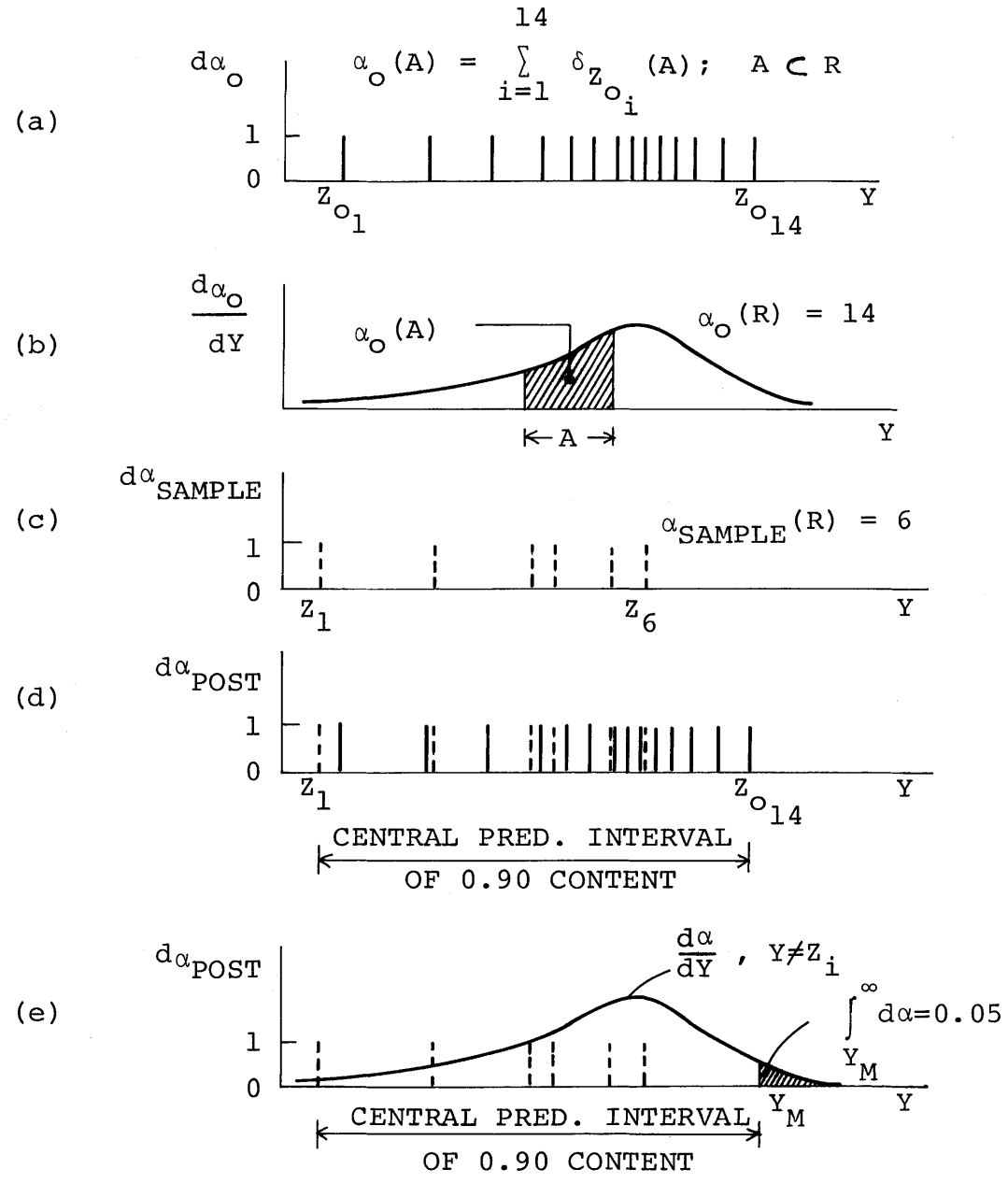


Figure 6. Construction of nonparametric Bayesian prediction intervals of 0.90 content. (a)-Prior; (b)-Prior; (c)-Sample; (d)-Posterior (a+c); (e)-Posterior (b+c).

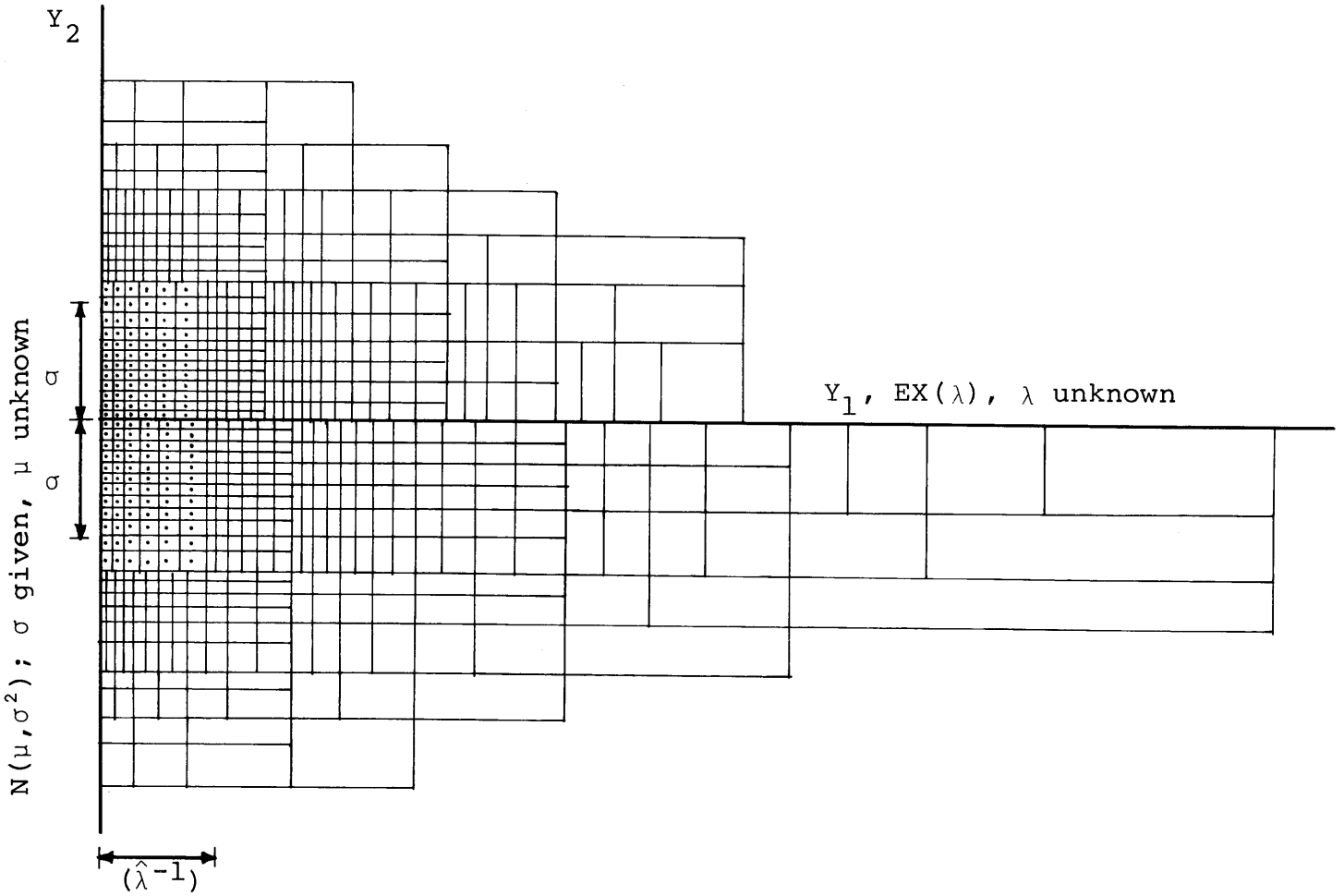


Figure 7. Lower half: rectangular probability chart; $Y_1 \sim EX(\lambda)$ with λ unknown; $Y_2 \sim N(\mu,\sigma^2)$ with μ unknown. Sample size = 5. Each non-dotted rectangle has probability content 0.001; each dotted rectangle has 0.003 content.
Upper half: the same chart for perfect knowledge of the distributions parameters.

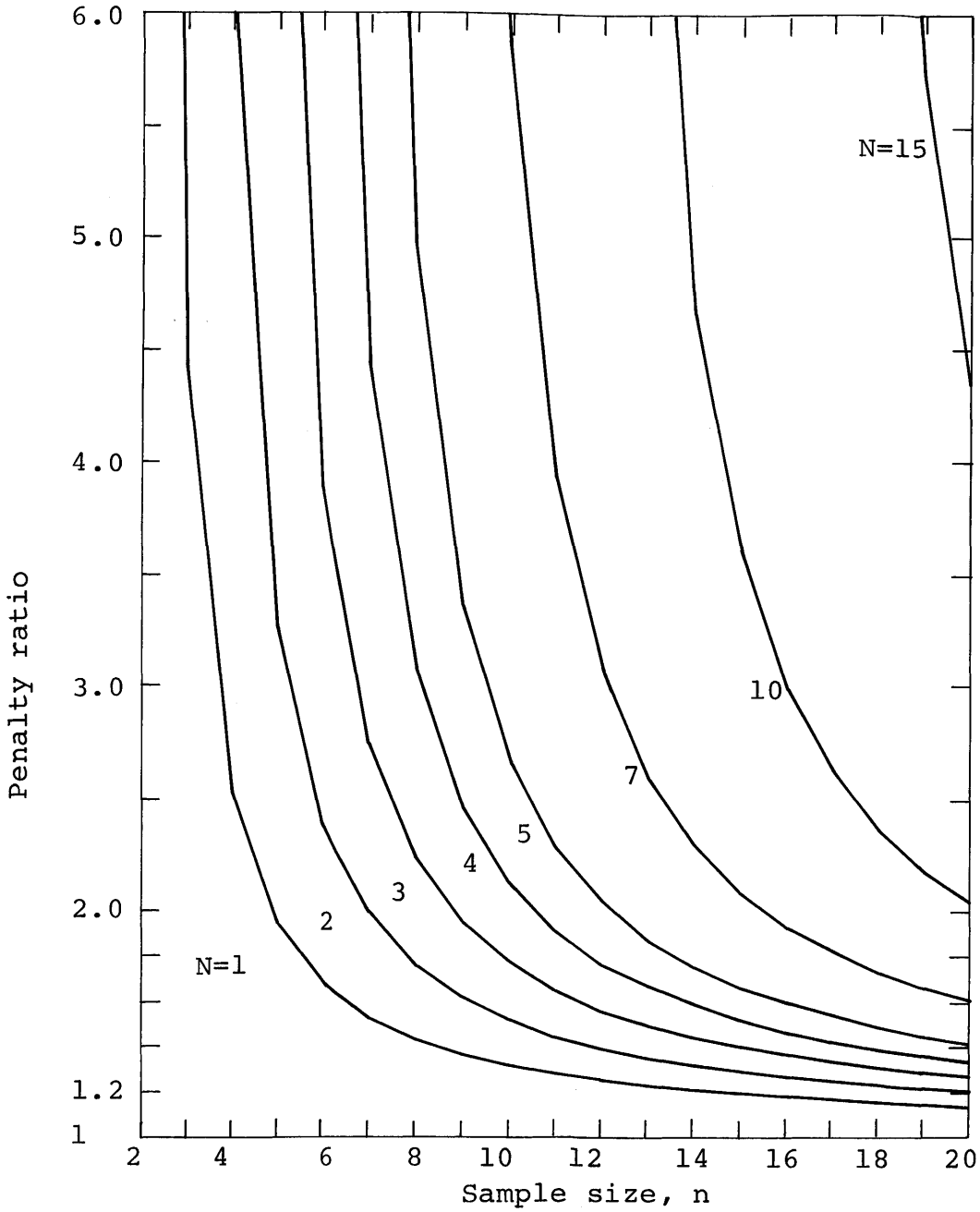


Figure 8. Penalty ratios for central prediction regions of 0.99 content. Sampling is from an N-variate normal population with $\underline{\mu}$ and $\underline{\Sigma}$ unknown.

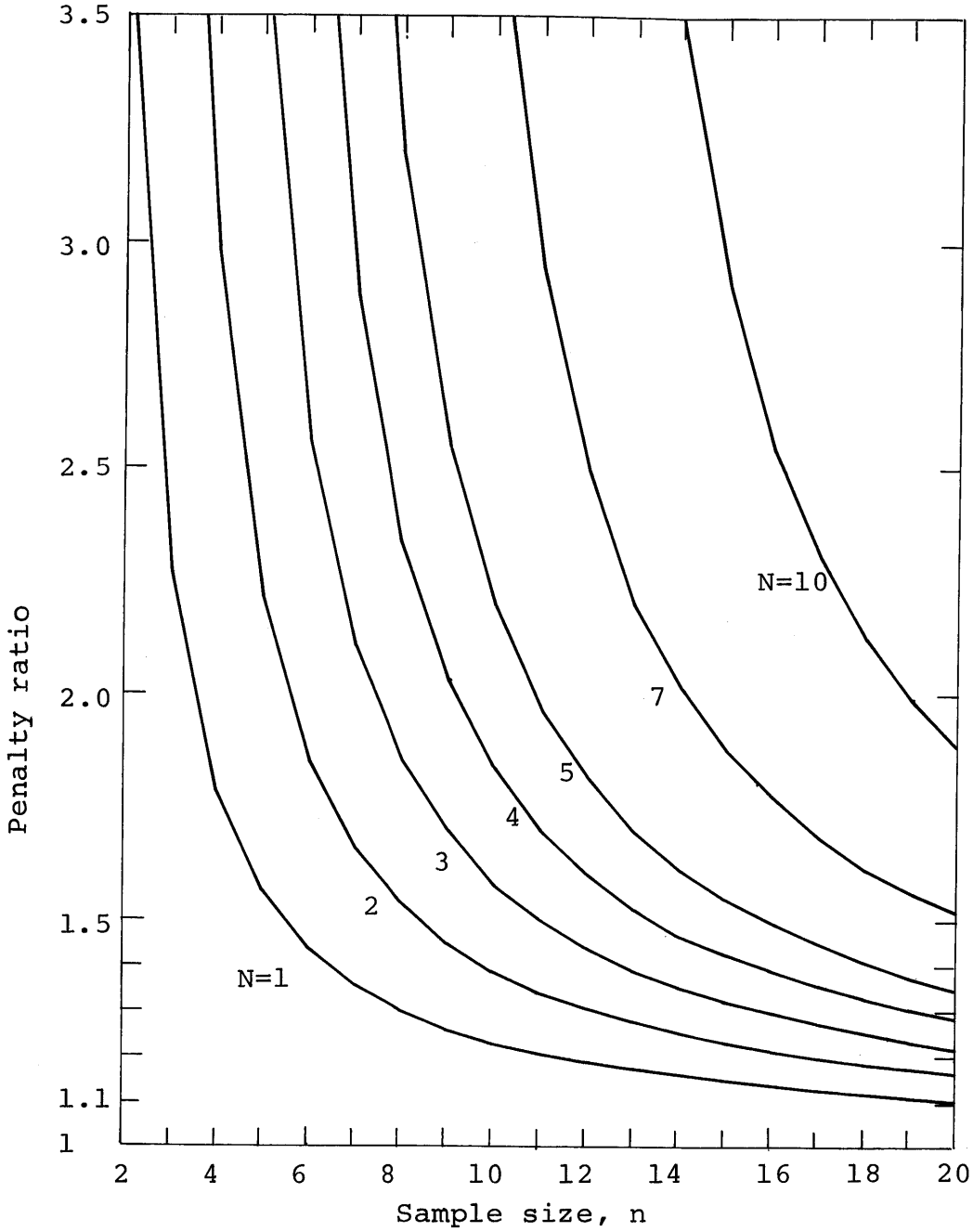


Figure 9. Penalty ratios for central prediction regions of 0.99 content. Sampling is from an N-variate normal population with $\underline{\mu}$ known and $\underline{\Sigma}$ unknown.

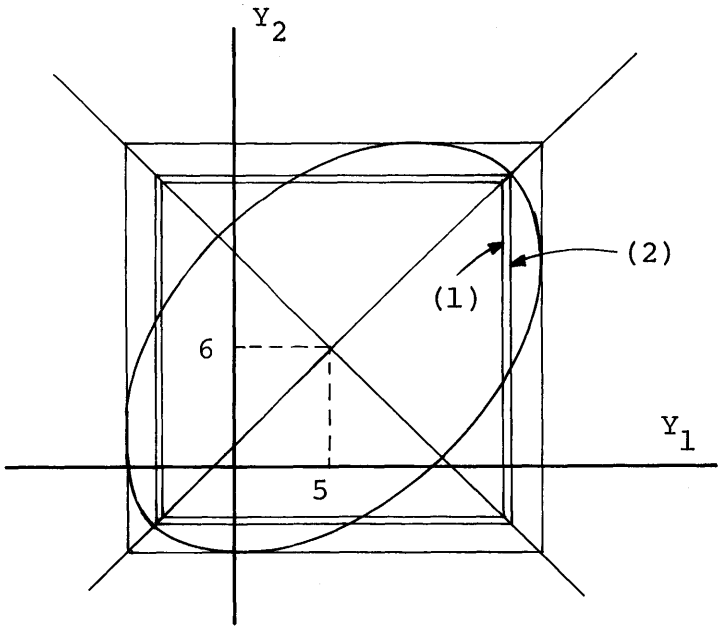


Figure 10. Two-dimensional prediction regions. The ellipse and the inner square are of 0.99 content; the other two squares are conservative approximations. Refer to Paragraph II.2.3 in the text.
(1) from tables in Hahn and Hendrickson (1971);
(2) from a Bonferroni's inequality.

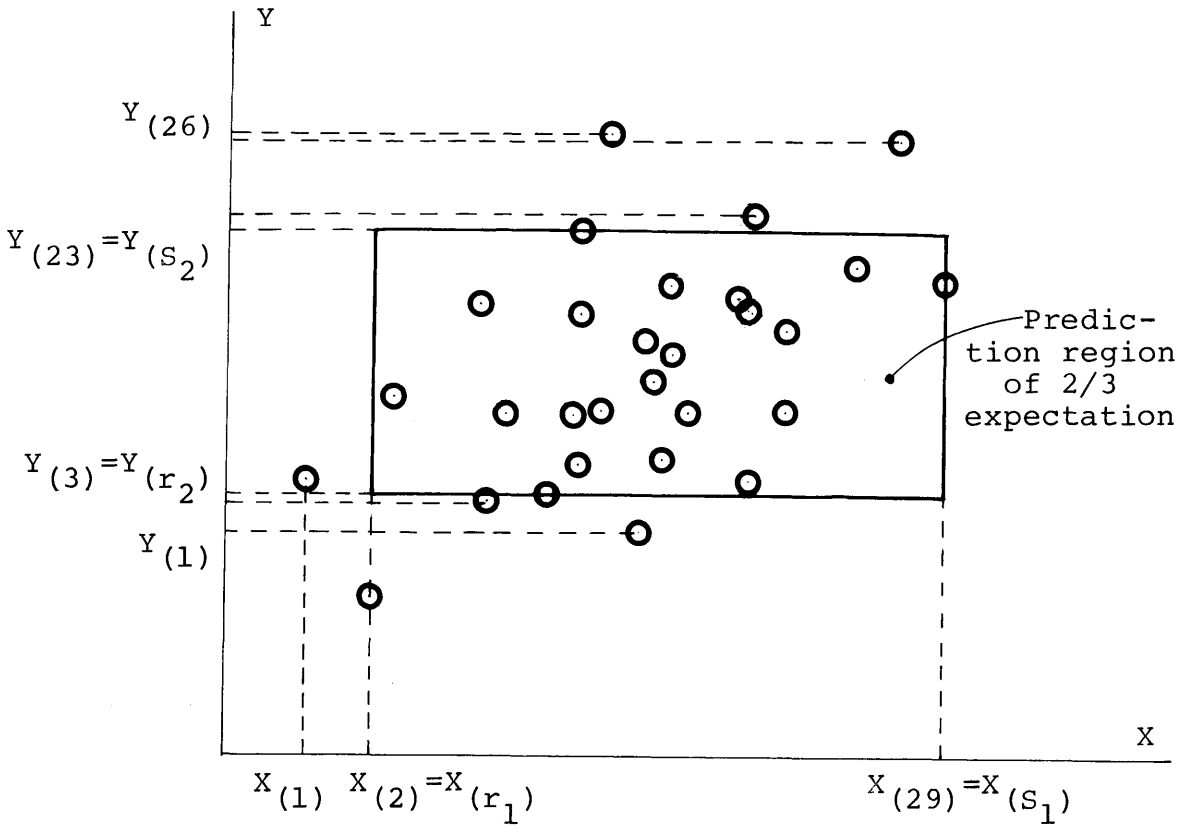


Figure 11. Example of Wald's "Successive Elimination Method" for the construction of nonparametric prediction rectangles of given expected content.

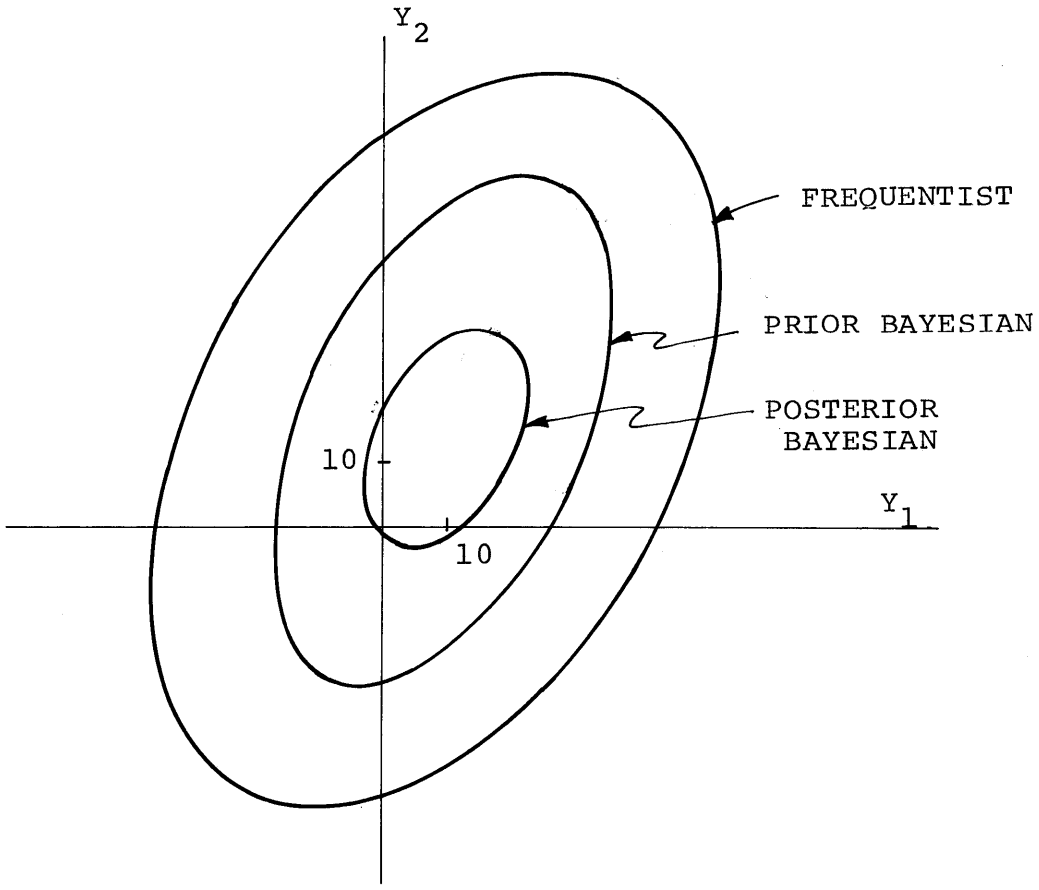


Figure 12. Frequentist and Bayesian prediction ellipses of 0.99 expectation. Example in Paragraph II.2.4.

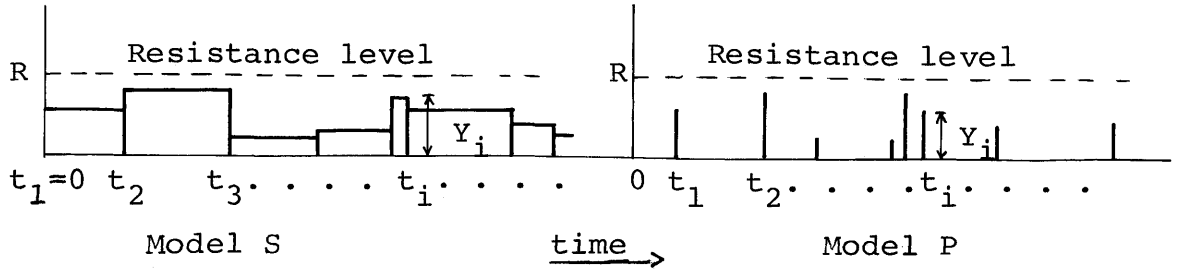


Figure 13. Realizations of simple independent "load" processes $\{Y_i\}$.

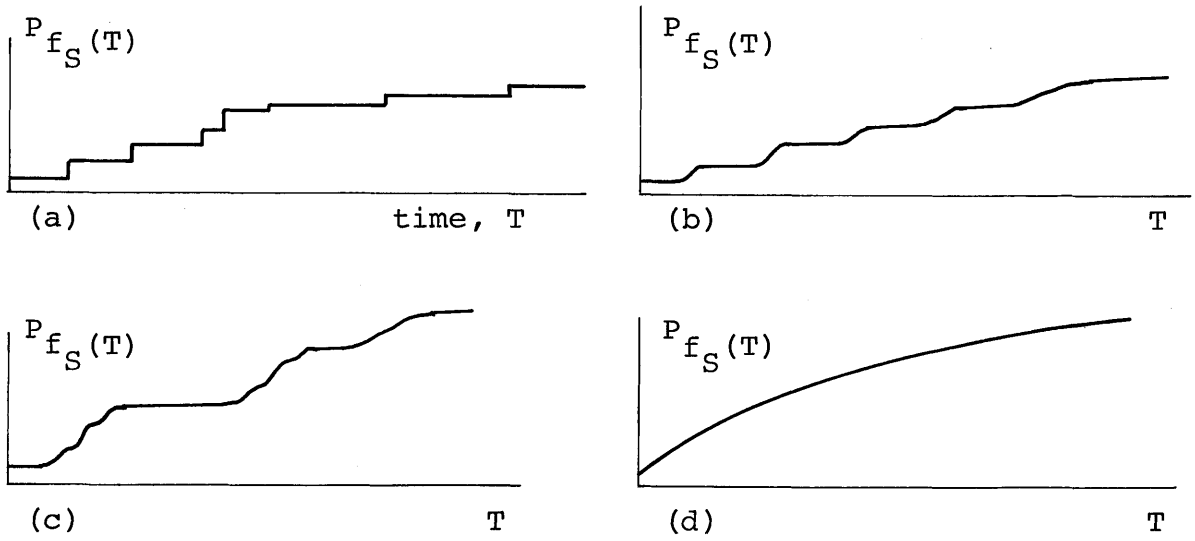


Figure 14. Qualitative evolution of the failure probability for different correlation structures of the interarrival times; model S in Figure 13. (a) deterministic arrival times; (b) highly regular (but random) interarrival times; (c) highly regular (but random) events clusters; (d) Poisson arrivals.

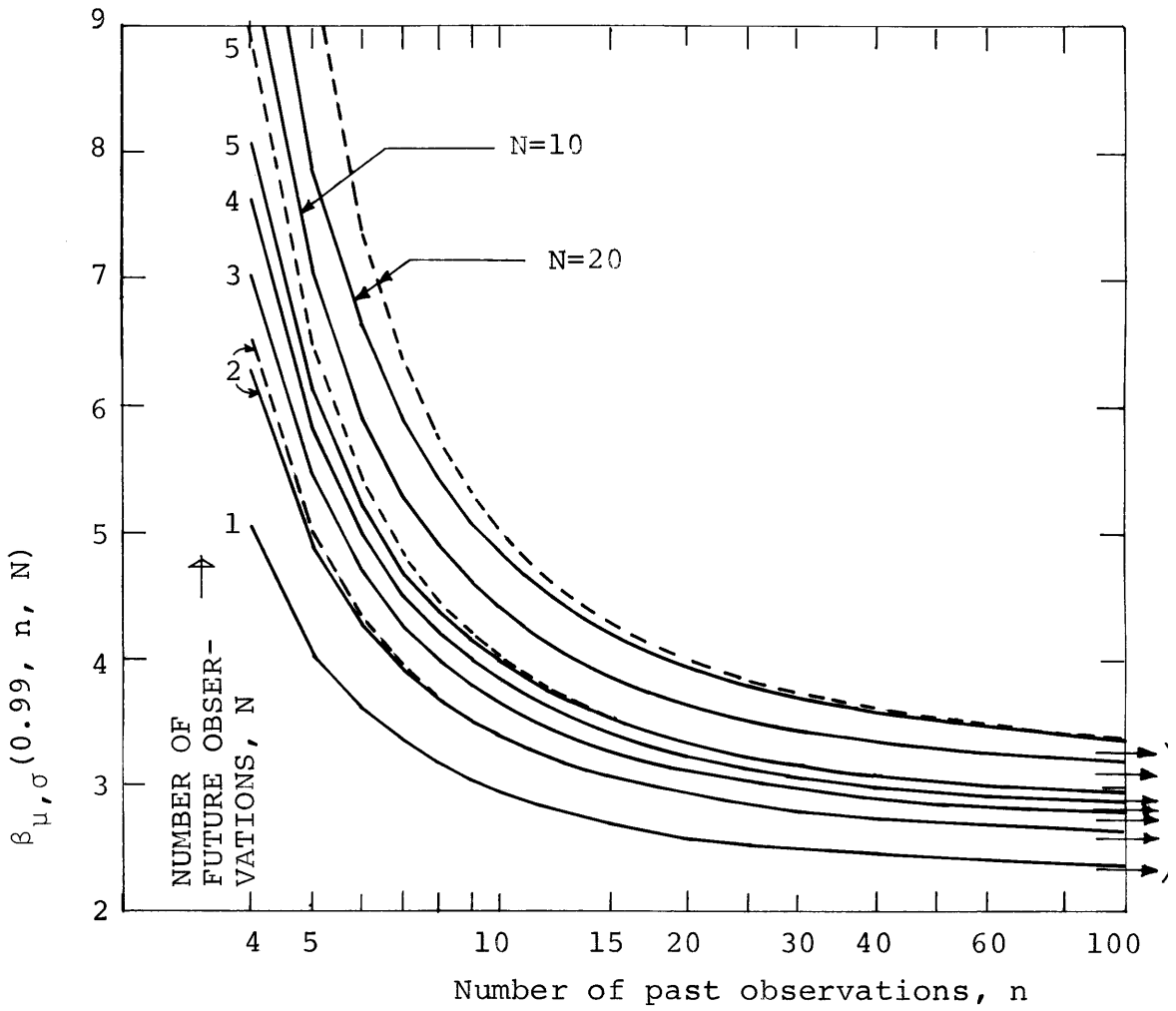


Figure 15. Simultaneous prediction from independent normal sequences with μ a σ unknown. One sided prediction intervals of 0.99 content.
Solid lines: exact results from Hahn (1970)
Dashed lines: independence approximation, equation (II.130)

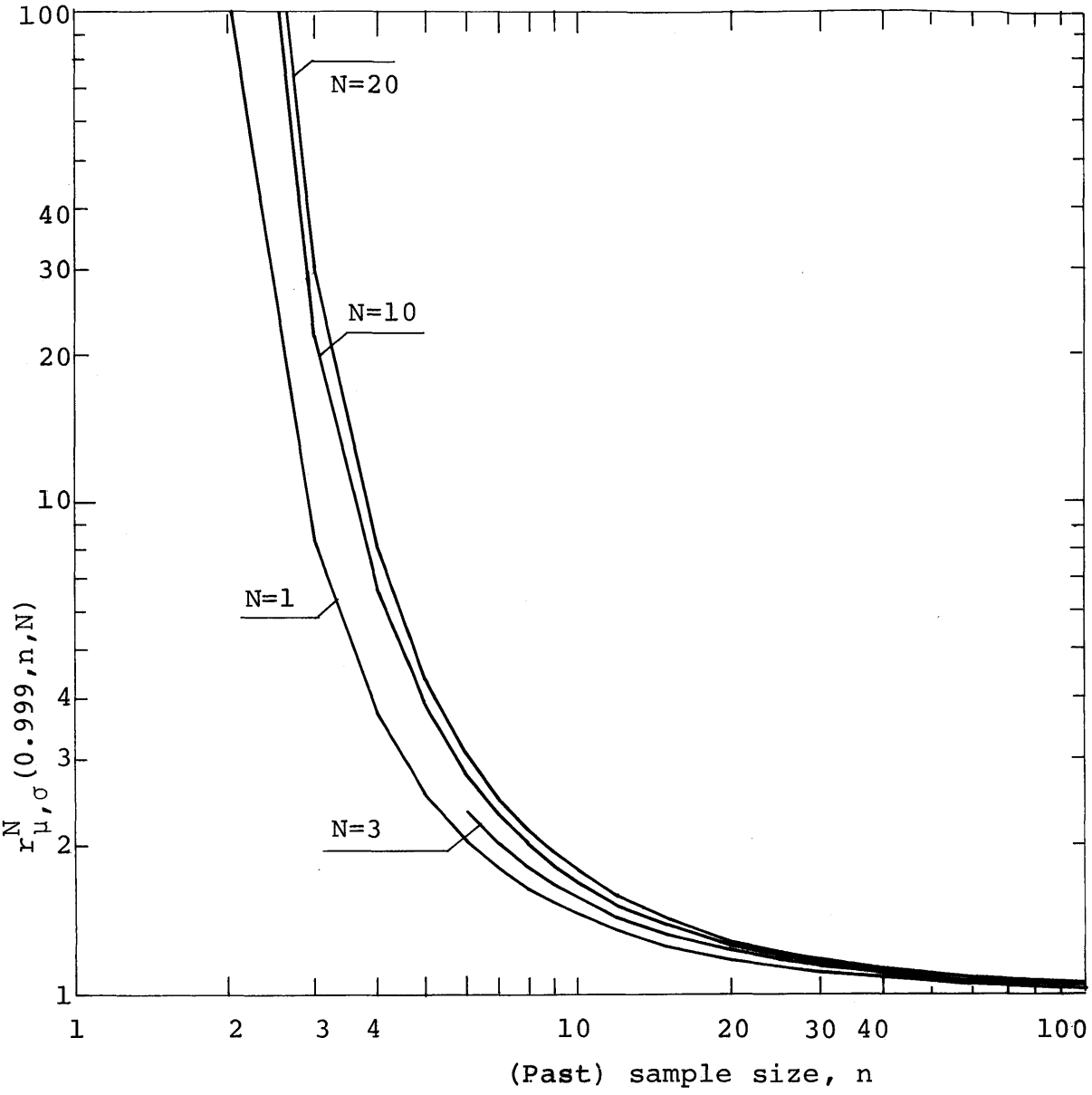


Figure 16. Approximate penalty ratios when constructing one-sided simultaneous prediction intervals of 0.999 content (N = number of future observations). Sampling is from normal populations with unknown mean and variance. See comments following equation (II.130).

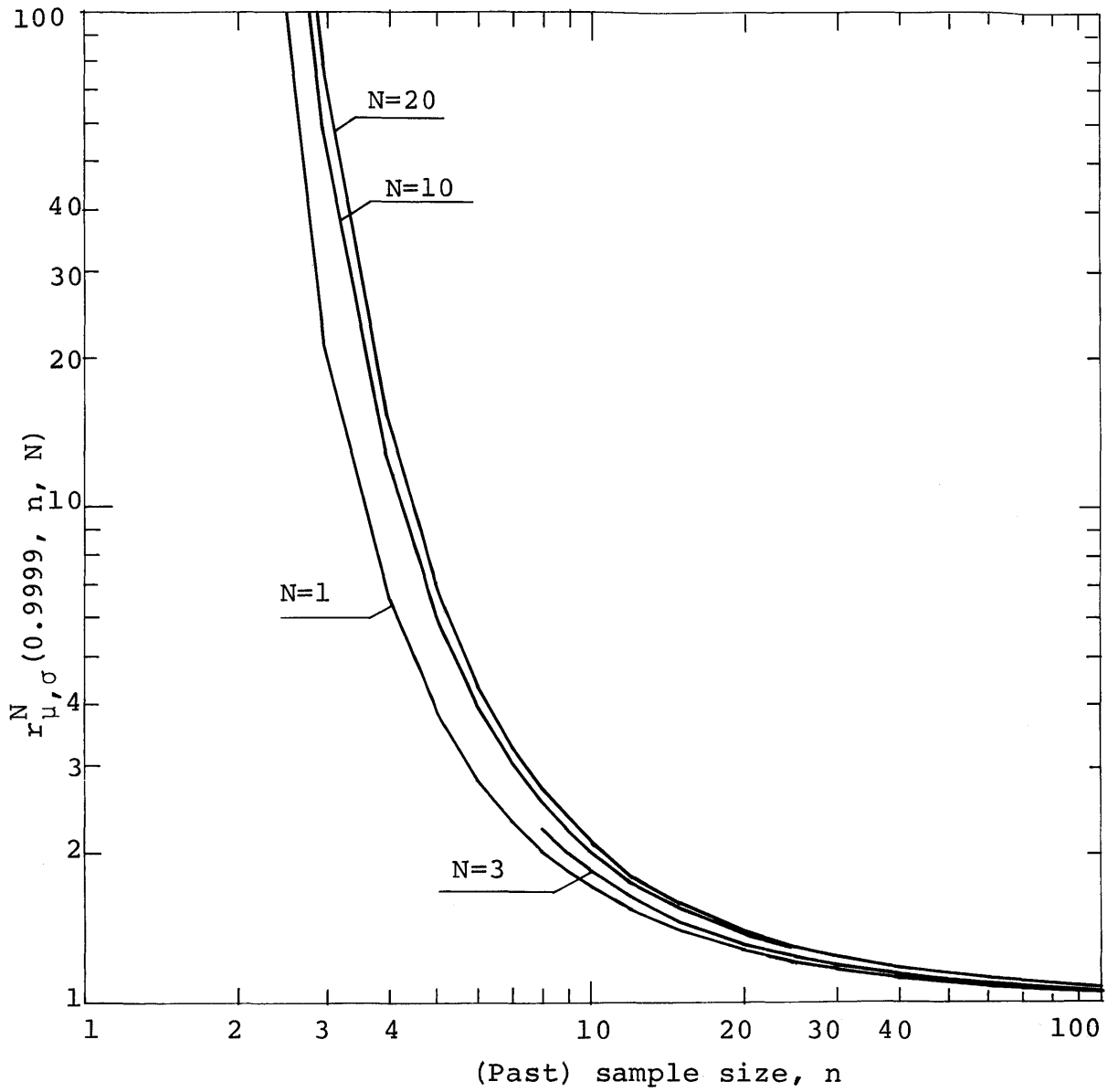


Figure 17. Approximate penalty ratios when constructing one-sided simultaneous prediction intervals of 0.9999 content (N = number of future observations). Sampling is from normal populations with unknown mean and variance. See comments following equation (II.130).

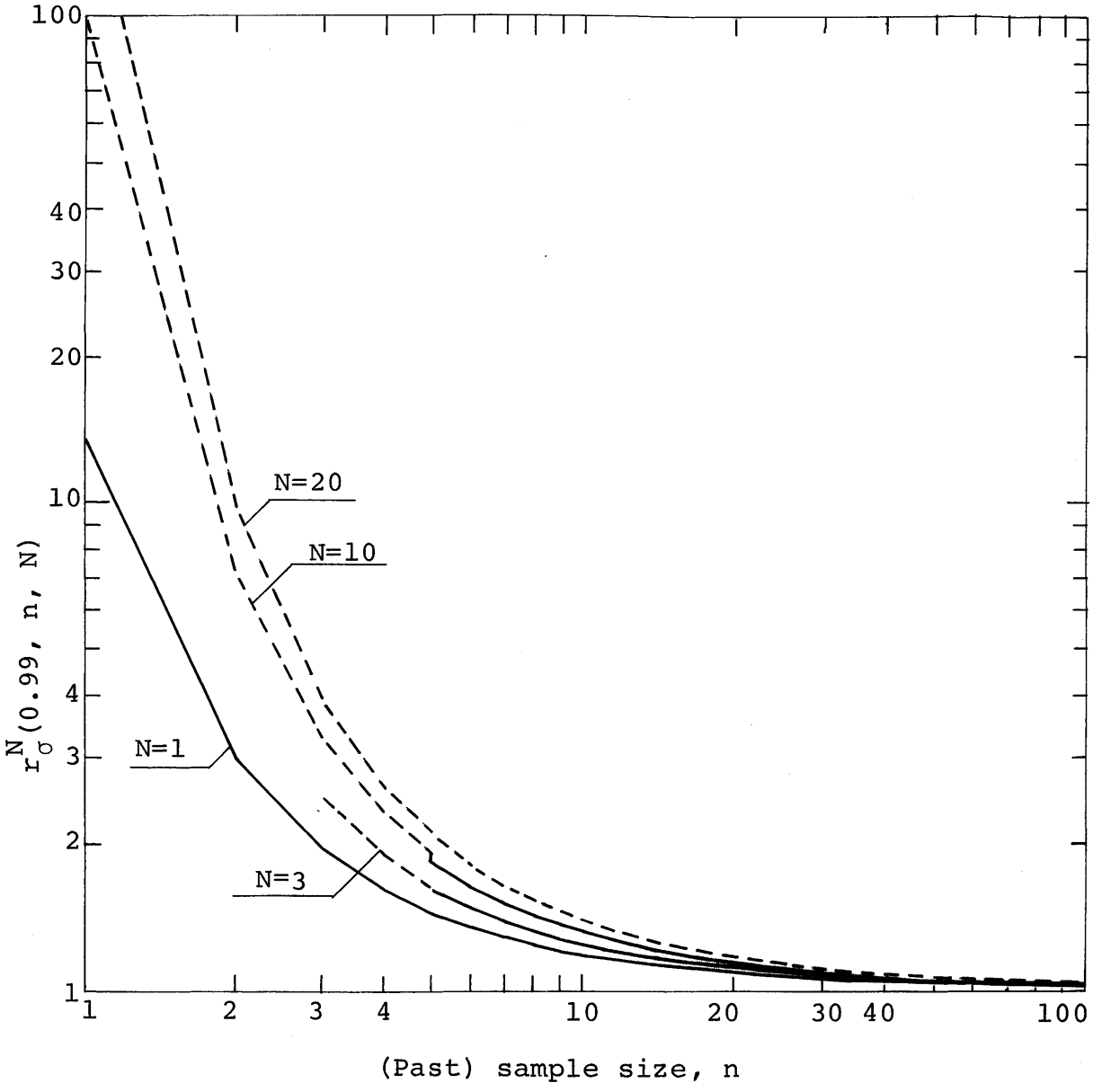


Figure 18. Penalty ratios when constructing one-sided simultaneous prediction intervals of 0.99 content (N = number of future observations). Sampling is from normal populations with known mean and unknown variance. Solid curves: exact values from Krishnaiah and Armitage (1966); dashed curves: approximate values.

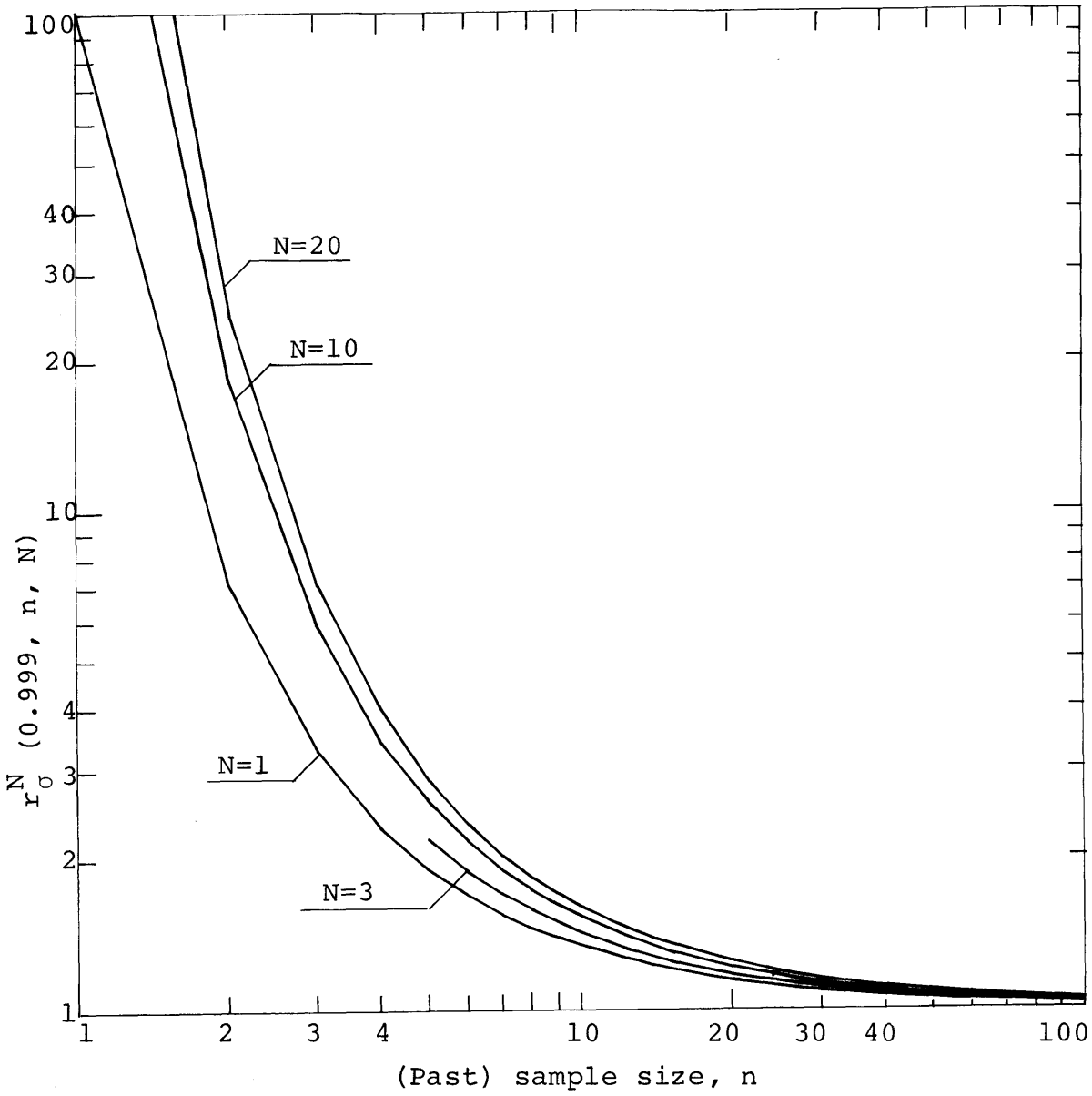


Figure 19. Approximate penalty ratios when constructing one-sided simultaneous prediction intervals of 0.999 content (N = number of future observations). Sampling is from normal populations with known mean and unknown variance.

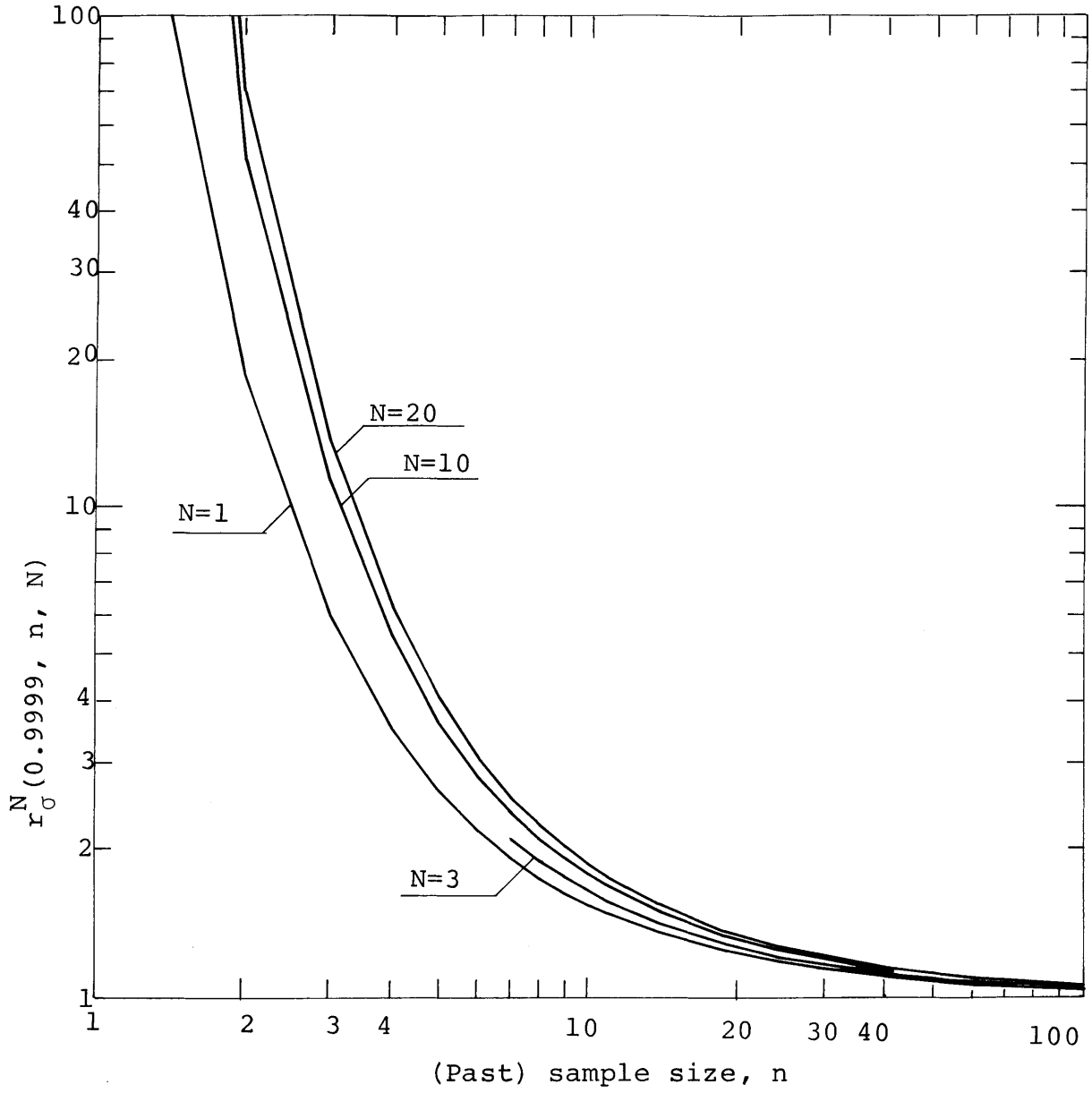


Figure 20. Approximate penalty ratios when constructing one-sided simultaneous prediction intervals of 0.9999 content (N = number of future observations). Sampling is from normal populations with known mean and unknown variance.

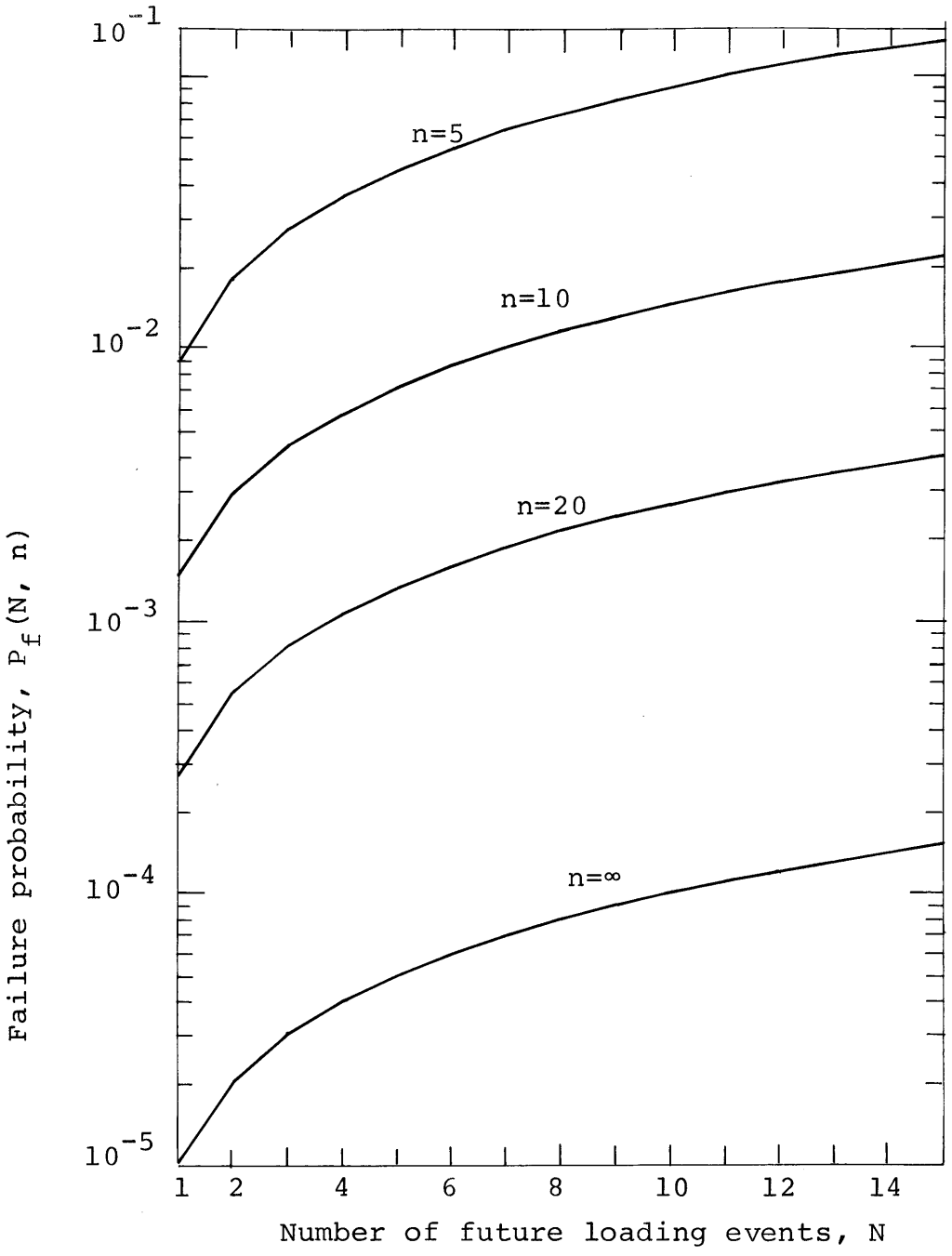


Figure 21(a). Probability of failure as a function of the number of future loading events, N. The load sequence is normal and independent, with unknown mean and variance. From the sequence, a sample of size n is available, yielding the unbiased parameters estimates $\hat{Z}$ and S^2. The resistance is $R = \hat{Z} + 4.265\ S$.

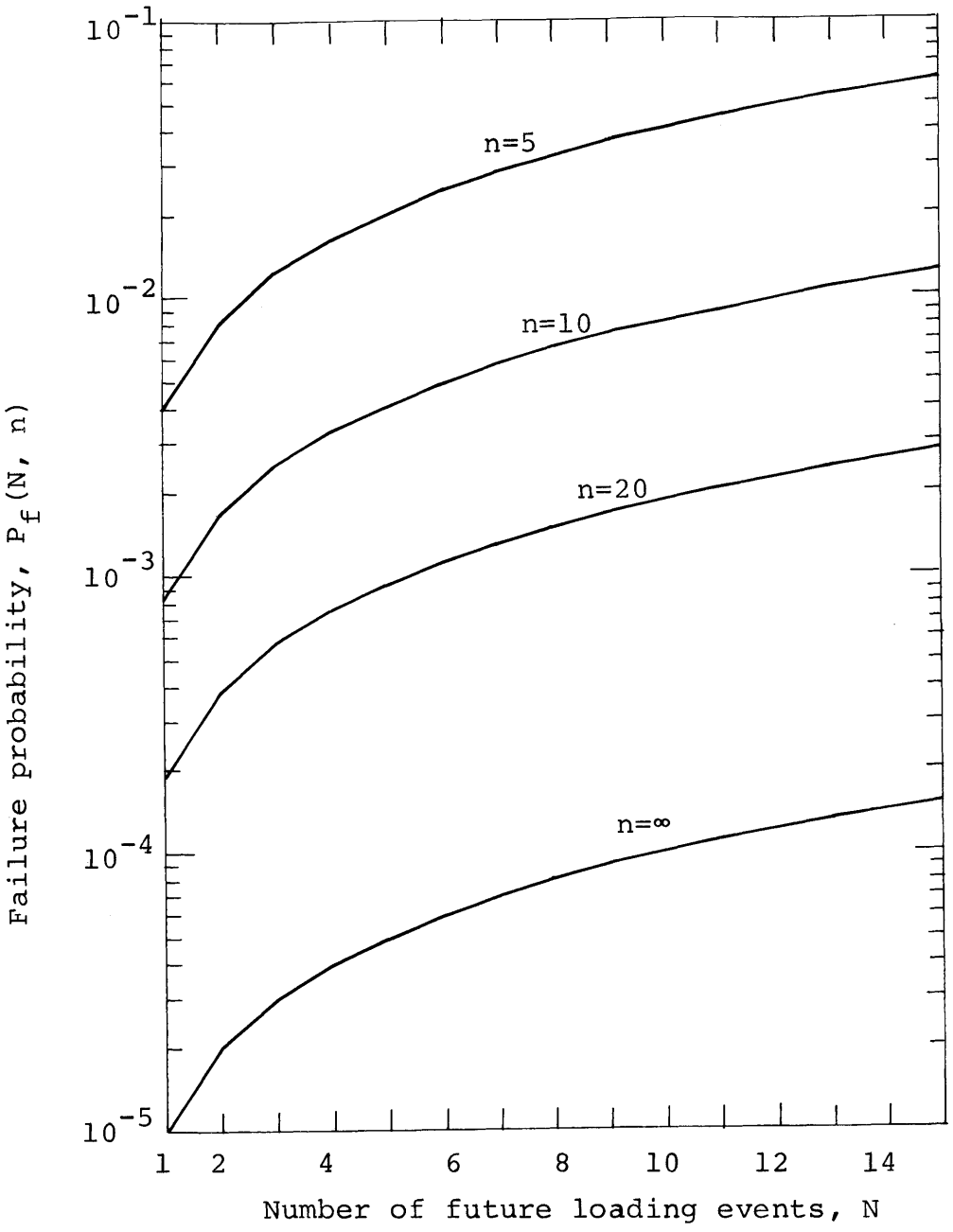


Figure 21(b). Probability of failure as a function of the number of future loading events, N. The load sequence is normal and independent, with known mean and unknown variance. From the sequence, a sample of size n is available, yielding the unbiased estimate $\hat{Z}$. The resistance is $R = \hat{Z} + 4.265\ \sigma$.

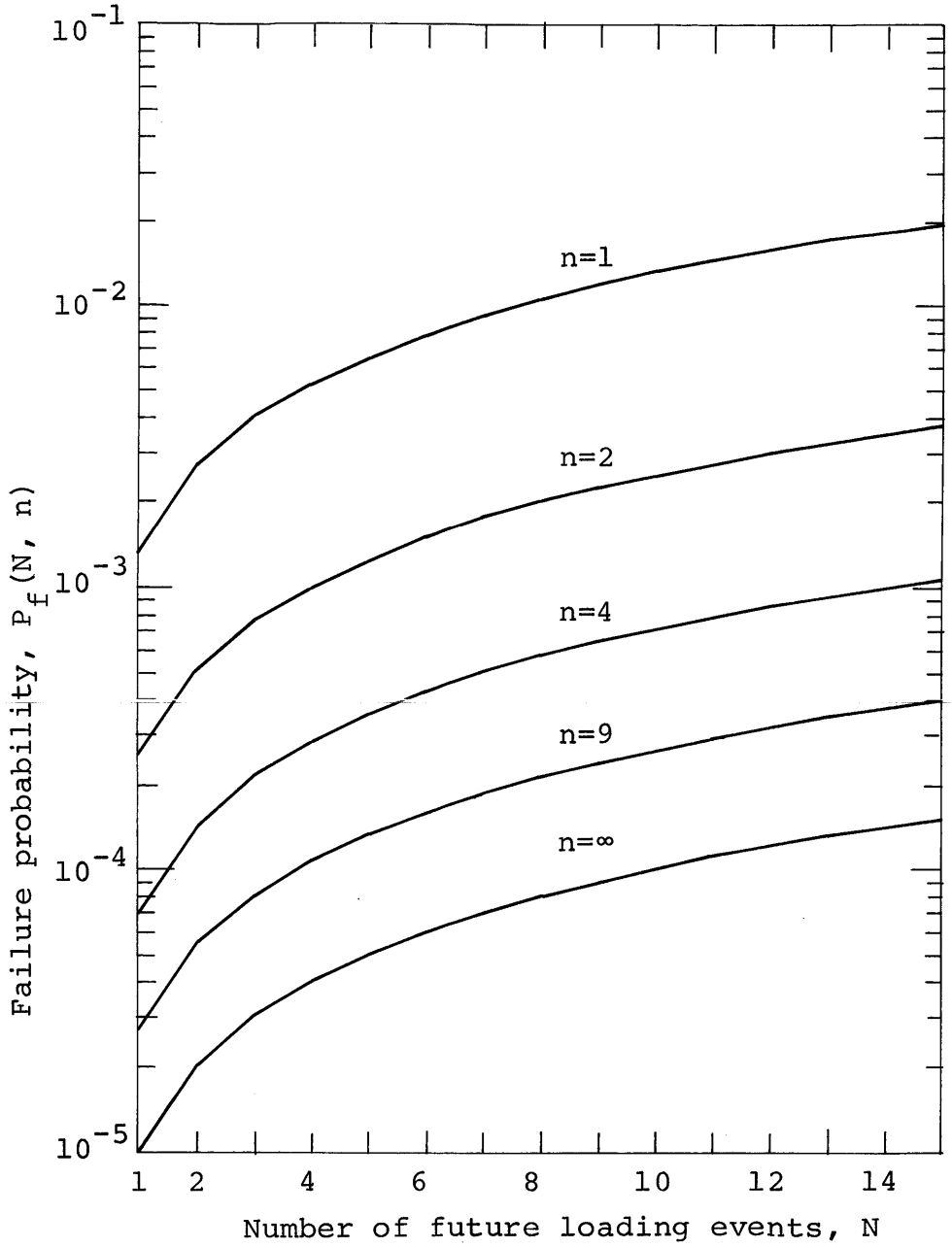


Figure 21(c). Probability of failure as a function of the number of future loading events, N. The load sequence is normal and independent, with unknown mean and known variance. From the sequence, a sample of size n is available, yielding the unbiased estimate s^2. The resistance is $R = \mu + 4.265\ S$.

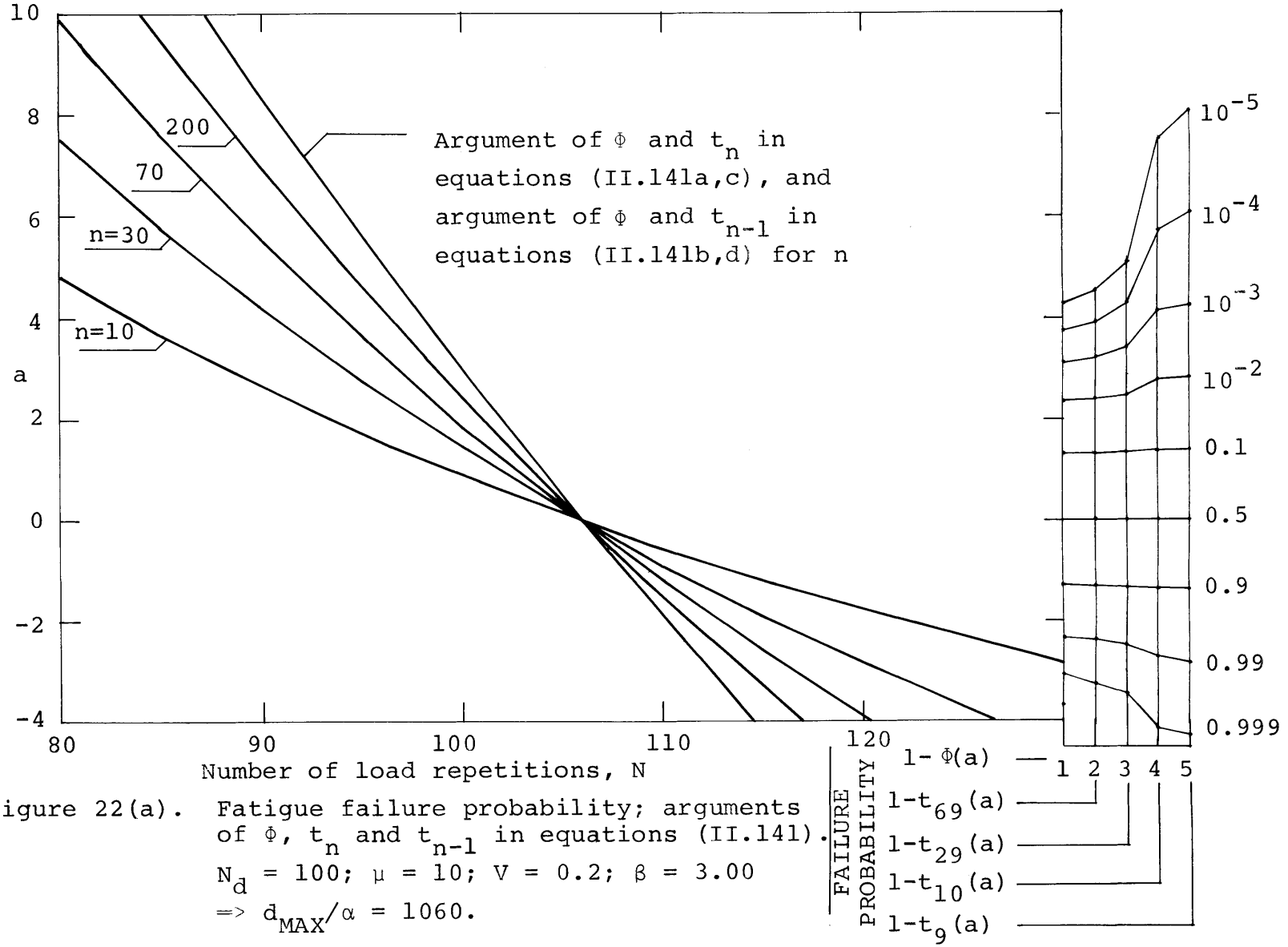


Figure 22(a). Fatigue failure probability; arguments of Φ, t_n and t_{n-1} in equations (II.141).
N_d = 100; μ = 10; V = 0.2; β = 3.00
=> d_{MAX}/α = 1060.

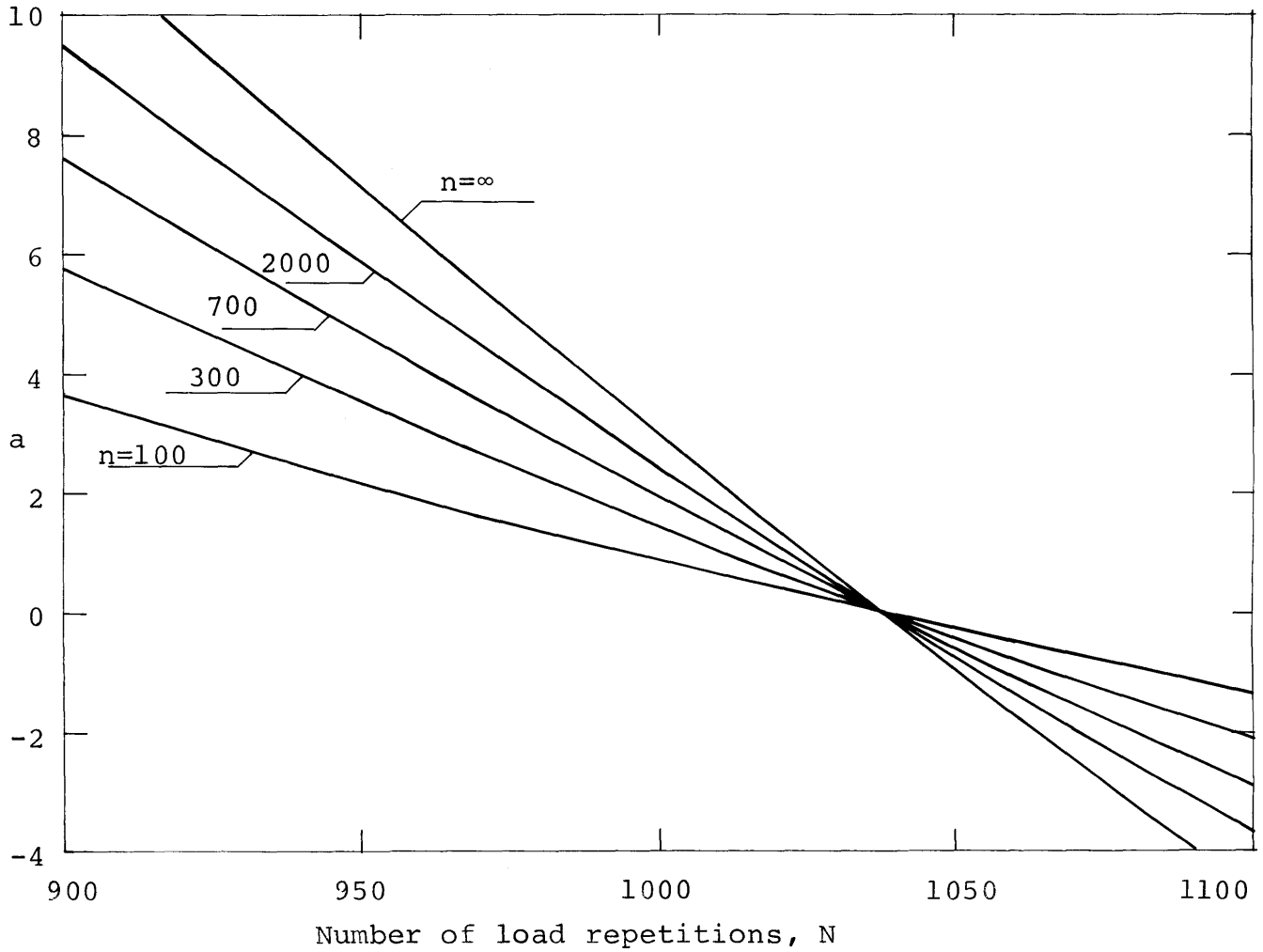


Figure 22(b). Fatigue failure. Arguments of Φ, t_n and t_{n-1} in equations (II.141).
$N_d = 1000$; $\mu = 10$; $V = 0.4$; $\beta = 3.00 \Rightarrow d_{MAX}/\alpha = 10{,}379$.

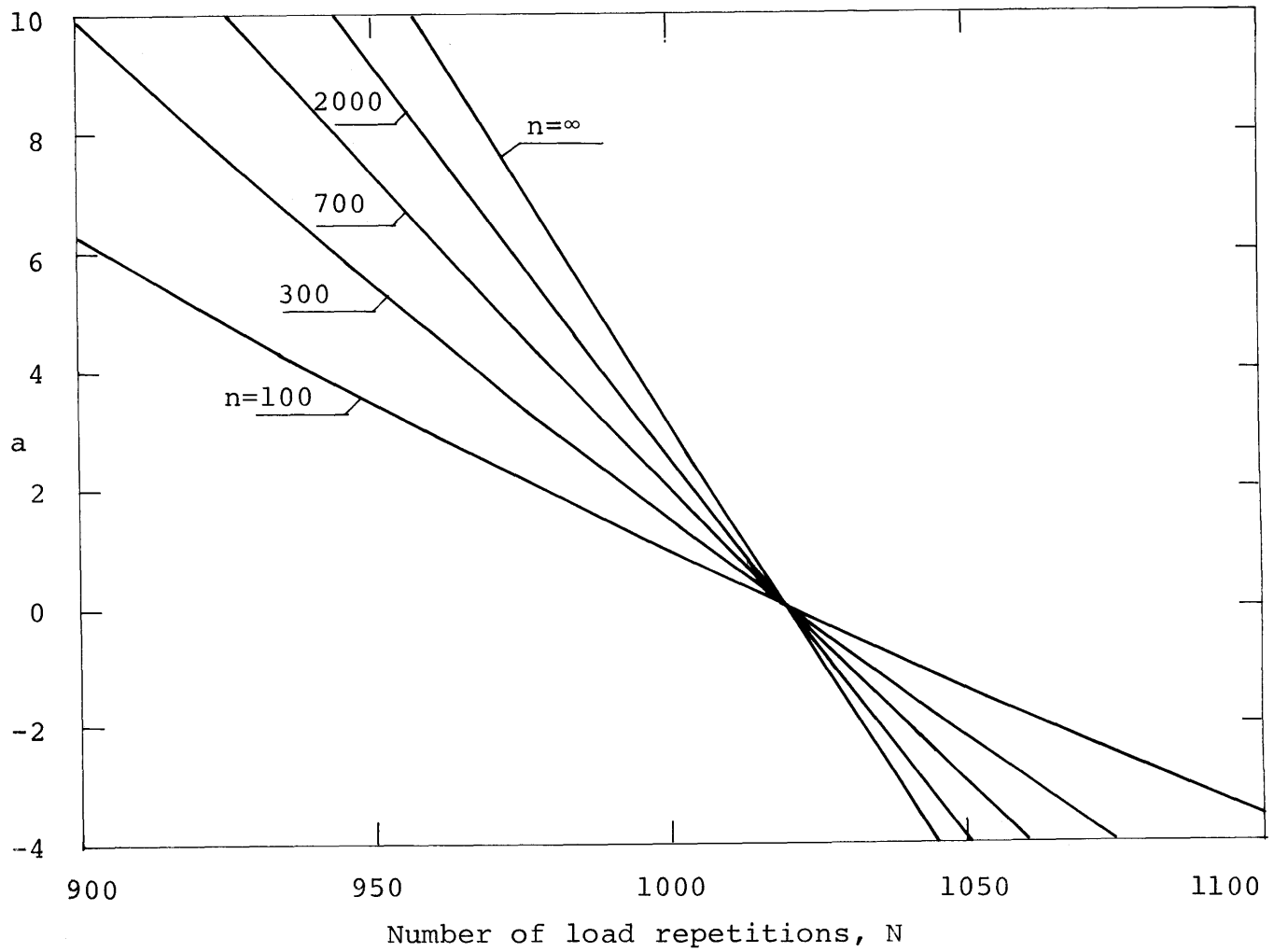


Figure 22(c). Fatigue failure. Arguments of Φ, t_n and t_{n-1} in equations (II.141). $N_d = 1000$; $\mu = 10$; $V = 0.2$; $\beta = 3.00 \Rightarrow d_{MAX}/\alpha = 10{,}190$.

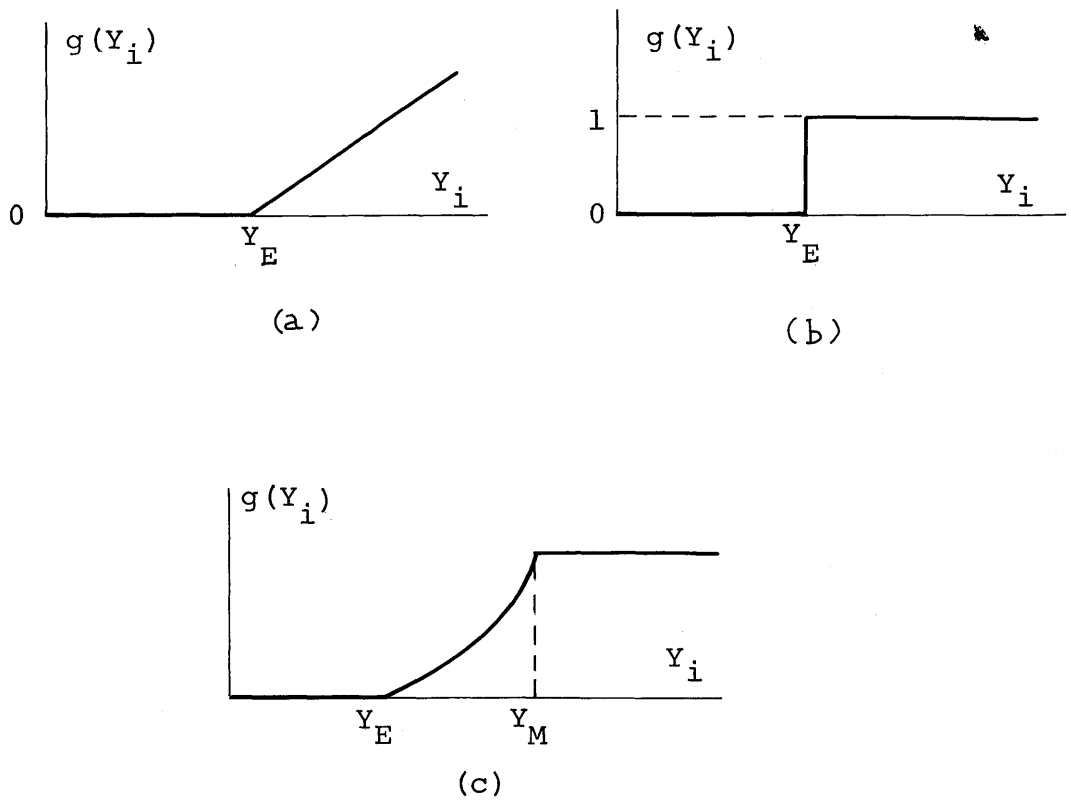


Figure 23. Damage accumulation functions. Refer to equation (II.151) in the text.

CHAPTER III

FIRST-ORDER AUTOREGRESSIVE MODELS

Knowledge of natural phenomena is seldom so accurate to justify a deterministic model in which, given the present, the future is known with certainty. On the other hand, independent models of the type studied in Chapter II (independence of the future from the present and from the past) may not be supported by solid statistical evidence. For example, in the area of structural safety, processes like the interarrival times and the intensities of meteorological loads on buildings and the time-dependent resistance of mechanical systems generally exhibit serial correlation and are best described by probabilistic models with memory. The degree of serial correlation may be assumed known, or may be estimated from statistical data.

One of the simplest and most commonly used probabilistic models with memory hypothesizes a Markovian dependence between the states of the system at different times. This chapter studies the relevance of statistical uncertainty in first-order autoregressive processes (a simple subclass of Markov processes), when some of the parameters are unknown. The statistical information is contained in the (known) realization of the process from an initial time,

t = 0, to the present. As for the case of independent sequences, simple and, to a lesser extent, simultaneous prediction are studied.

A distinctive characteristic of processes with memory is that the predictive distribution depends on the prediction lead. In fact, past information is used for prediction in two distinct ways. First, it allows one to estimate the unknown parameters of the process, and in this sense there is little difference with the case of independent sequences. Second, and peculiar to processes with memory, one can use knowledge of the present to estimate more accurately events in the near future, due to serial dependence. This fact motivates the distinction between prediction at lead 1 (Section III.1) and prediction at lead $\ell > 1$. The latter problem is studied in Section III.2.

A Bayesian approach is followed, using both "diffuse" and conjugate prior distributions of the unknown parameters.

III.1 SIMPLE PREDICTION OF THE NEXT OBSERVATION

Consider the Gauss Markov sequence $\{Y_i\}$ generated by the first-order linear difference equation:

$$Y_i = \beta_o + \beta_1 Y_{i-1} + \varepsilon_i; \quad i = 1,2,\ldots; \; Y_o \text{ given}, \quad \text{(III.1)}$$

where β_o and β_1 are constant parameters and $\{\varepsilon_i\}$ is an independent sequence of $N(0,\sigma^2)$ variates. Y_o and $\underline{Y} = [Y_1,\ldots,Y_n]'$ are known. Y_n denotes the present state, and the problem consists of finding the posterior distribution of Y_{n+1}, given Y_o, $\underline{Y}$ and the prior distribution of the unknown parameters. For both the cases of diffuse and conjugate prior the following conditions are studied:

(a) β_o, β_1, σ unknown;

(b) β_o, β_1 unknown, σ known.

III.1.1 "Diffuse" prior; β_o, β_1 and σ unknown

The noninformative prior distribution of the unknown parameters is assumed (with Zellner (1971), p 186) in the "diffuse" form:

$$f(\beta_o,\beta_1,\sigma) \propto 1/\sigma \ , \quad \sigma \geq 0. \tag{III.2}$$

(Note that considering the density (III.2) "noninformative" is not standard in the statistical literature.) Given the initial state Y_o and the sample $\underline{Y}$, the likelihood function for β_o, β_1 and σ is:

$$\ell(\beta_o,\beta_1,\sigma|Y_o,\underline{Y}) \propto \frac{1}{\sigma^n} \exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (Y_i-\beta_o-\beta_1 Y_{i-1})^2\} \quad \text{(III.3)}$$

which, when combined with the prior density (III.2), yields the posterior density:

$$f(\beta_o,\beta_1,\sigma|Y_o,\underline{Y}) \propto \frac{1}{\sigma^{n+1}} \exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (Y_i-\beta_o-\beta_1 Y_{i-1})^2\}, \quad \text{(III.4)}$$

meaning that the posterior distribution of β_o, β_1 and $h = \sigma^{-2}$ is of normal-gamma type (see Raiffa and Shlaifer (1961), p. 57).

Let

$$\underline{\beta} = \begin{pmatrix} \beta_o \\ \beta_1 \end{pmatrix} ; \quad \underline{H} = \begin{pmatrix} 1 & Y_o \\ 1 & Y_1 \\ \vdots & \vdots \\ 1 & Y_{n-1} \end{pmatrix} ; \quad \tilde{\underline{H}} = [1 \quad Y_n];$$

$$\underline{\varepsilon} = [\varepsilon_1,\ldots,\varepsilon_n]' .$$

Then:

$$\underline{Y} = \underline{H}\,\underline{\beta} + \underline{\varepsilon} \quad ,$$

$$Y_{n+1} = \tilde{\underline{H}}\,\underline{\beta} + \varepsilon_{n+1} \quad ,$$

and the posterior distribution of $\underline{\beta}$ (after integrating out σ from the joint density (III.4)) is bivariate-t with mean $\hat{\underline{\beta}}$ and covariance matrix $\frac{\nu}{\nu-2} S^2 (\underline{H}'\underline{H})^{-1}$, where $\hat{\beta}$, ν and S^2 are sample statistics, defined as follows:

$$\hat{\underline{\beta}} = (\underline{H}'\underline{H})^{-1}\underline{H}'\underline{Y} \qquad \text{(assume det } \underline{H}'\underline{H} \neq 0)$$

$$\nu = n-2$$

$$\nu\, S^2 = \underline{Y}'[\underline{I}-\underline{H}(\underline{H}'\underline{H})^{-1}\underline{H}']\underline{Y} = (\underline{Y}-\underline{H}\,\hat{\underline{\beta}})'(\underline{Y}-\underline{H}\,\hat{\underline{\beta}}).$$

With the posterior density function (III.4) for the unknown parameters, the posterior predictive distribution of Y_{n+1} (see Zellner (1971), p. 72) is t with $\nu = n-2$ degrees of freedom, mean value

$$Y_{n+1} = \tilde{\underline{H}}\,\hat{\underline{\beta}} \qquad \text{(III.5)}$$

and variance (for $n > 4$):

$$\sigma^2_{n+1} = \frac{\nu}{\nu-2} S^2[1+\tilde{\underline{H}}(\underline{H}'\underline{H})^{-1}\tilde{\underline{H}}'] = \frac{n-2}{n-4} S^2\left(1+ \frac{a-2bY_n+nY_n^2}{na - b^2}\right) \text{(III.6)}$$

where $a = \sum_{i=0}^{n-1} Y_i^2$; $b = \sum_{i=0}^{n-1} Y_i$.

From equation (III.6) it is seen that the prediction variance is the sum of two terms:

(1) $\frac{n-2}{n-4} S^2$, being the posterior predictive variance of ε_{n+1};

(2) $\frac{n-2}{n-4} S^2\tilde{\underline{H}}(\underline{H}'\underline{H})^{-1}\tilde{\underline{H}}'$, being the posterior variance of $\hat{\underline{H}}\,\underline{\beta}$.

From the posterior distribution of Y_{n+1} a left-hand prediction interval of P-content for Y_{n+1} is:

$$(-\infty, (\hat{\beta}_o + \hat{\beta}_1 Y_n) + \beta_{\beta_o,\beta_1,\sigma}(P,n,a,b,Y_n)\ S] \ , \qquad \text{(III.7)}$$

where

$$\beta_{\beta_o,\beta_1,\sigma}(P,n,a,b,Y_n) = t_{n-2}(P)\left[1 + \frac{a-2Y_n b+nY_n^2}{n\,a - b^2}\right]^{1/2}$$

For β_o, β_1 and σ known, the same interval would be:

$(-\infty, (\beta_o+\beta_1 Y_n) + \beta_{PI}(P)\ \sigma]$, where $\beta_{PI}(P) = \Phi(P)$.

Therefore one can define the penalty ratio for imperfect information as follows:

$$r_{\beta_o,\beta_1,\sigma}(P,n,a,b,Y_n) = \frac{t_{n-2}(P)}{\Phi(P)}\left[1 + \frac{a-2Y_n b+nY_n^2}{n\,a - b^2}\right]^{1/2} \qquad \text{(III.8)}$$

For example, if a = n and b = 0, it is:

$$r_{\beta_o,\beta_1,\sigma}(P,n,n,0,Y_n) = \frac{t_{n-2}(P)}{\Phi(P)}\left(1 + \frac{1+Y_n^2}{n}\right)^{1/2} \qquad \text{(III.9)}$$

The prediction penalty is minimum when $Y_n = 0$. For the case when $Y_n = 1$, n = 3(1)12,16(5)31,41,61, and P = 0.99, 0.999, 0.9999 the penalty factors (III.9) are collected in Table 1. For any given value of Y_n the penalty (III.9) exceeds the corresponding penalty for a white normal sequence with μ and σ^2 unknown (see Chapter II, Paragraph II.2.2(a) and Appendix A), due to the fact that β_1 is assumed unknown. However, this does not imply that the variances of Y_{n+1} in the two models are similarly ordered.

Parallel expressions and similar conclusions can be deduced for central and right-hand prediction intervals.

III.1.2 "Diffuse" prior; β_o, β_1 unknown, σ known

The condition "σ known" makes the prediction problem considerably simpler. Consistently with equation (III.2) we assume, as noninformative prior, the improper distribution:

$$d\, F(\beta_o,\beta_1) \propto d\, \beta_o\, d\, \beta_1 \ .$$

As in Paragraph III.1.1, define the statistic $\hat{\underline{\beta}} = (\underline{H}'\underline{H})^{-1}\underline{H}'\underline{Y}$ (assume det $(\underline{H}'\underline{H}) \neq 0$). Then the posterior distribution of $\underline{\beta} = [\beta_o,\beta_1]'$ is bivariate normal with mean $\hat{\underline{\beta}}$ and covariance matrix $\sigma^2(\underline{H}'\underline{H})^{-1}$, implying that the posterior predictive distribution of Y_{n+1} is also normal, with mean

$$\hat{Y}_{n+1} = \hat{\underline{H}}\ \hat{\underline{\beta}} = \hat{\beta}_o + \hat{\beta}_1\ Y_n$$

and variance

$$\sigma^2_{n+1} = \sigma^2[1 + \tilde{\underline{H}}(\underline{H}'\underline{H})^{-1}\hat{\underline{H}}'] = \sigma^2\left[\ 1 + \frac{a-2bY_n+nY_n^{\ 2}}{n\ a\ -\ b^2}\ \right]$$

(for the definition of a and b, see equation (III.6)).

If large realizations of Y_{n+1} are of concern, the left-hand prediction interval of P-content is:

$$(-\infty, (\hat{\beta}_o+\hat{\beta}_1\ Y_n) + \beta_{\beta_o,\beta_1}(P,n,a,b,Y_n)\cdot\sigma] \ ,$$

where

$$\beta_{\beta_o,\beta_1}(P,n,a,b,Y_n) = \Phi(P)\left(1 + \frac{a-2bY_n+nY_n^2}{n\,a - b^2}\right)^{1/2} .$$

The penalty ratio for imperfect information does not depend on the content P, being:

$$r_{\beta_o,\beta_1}(n,a,b,Y_n) = \left(1 + \frac{a-2bY_n+nY_n^2}{n\,a - b^2}\right)^{1/2} .$$

For example, when a = n and b = 0, this penalty becomes (compare with equation (III.9)):

$$r_{\beta_o,\beta_1}(n,n,0,Y_n) = \left(1 + \frac{1 + Y_n^2}{n}\right)^{1/2} ,$$

and is minimum for $Y_n = 0$.

III.1.3 Conjugate prior; β_o, β_1 and σ unknown

We follow again Zellner (1971), pages 70 and 186, and Raiffa and Shlaifer (1961), page 57.

For the likelihood function (III.3), the conjugate prior density is:

$$f(\beta_o,\beta_1,\sigma) \propto \frac{1}{\sigma^{n_1+1}} \exp\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n_1} (Y_{1_i}-\beta_o-\beta_1 Y_{1_{i-1}})^2\}$$

$$\propto \frac{1}{\sigma^{n_1+1}} \exp\{-\frac{1}{2\sigma^2} [\nu_1 S_1^2+(\underline{\beta}-\hat{\underline{\beta}}_1)'\underline{H}_1'\underline{H}_1(\underline{\beta}-\underline{\beta}_1)]\} \qquad \text{(III.10)}$$

where

$$\nu_1 = n_1 - 2 \quad , \quad \hat{\underline{\beta}}_1 = (\underline{H}_1'\underline{H}_1)^{-1}\underline{H}_1'\underline{Y}_1 \ ,$$

$$\nu_1 S_1^2 = (\underline{Y}_1-\underline{H}_1\hat{\underline{\beta}}_1)'(\underline{Y}_1-\underline{H}_1\hat{\underline{\beta}}_1), \text{ with } \underline{H}_1 = \begin{pmatrix} 1 & Y_{1_o} \\ 1 & Y_{1_1} \\ \vdots & \vdots \\ 1 & Y_{1_{n_1-1}} \end{pmatrix}$$

are "equivalent prior statistics".

The density (III.10) is formally similar to the posterior density (III.4) for diffuse prior.

The posterior density obtained by combining the prior (III.10) with the likelihood (III.3) is also of the same type. In fact, let $\underline{Y}_2 = [Y_1,\ldots,Y_{n_2}]'$ be the vector of observed sample values, and define

$$\underline{H}_2 = \begin{pmatrix} 1 & Y_{n_1} \\ 1 & Y_1 \\ \vdots & \vdots \\ 1 & Y_{n_2-1} \end{pmatrix} .$$

Then, in terms of

$$\underline{M} = \underline{H}_1'\underline{H}_1 + \underline{H}_2'\underline{H}_2 \ ;$$

$$\hat{\underline{\beta}} = \underline{M}^{-1}(\underline{H}_1\underline{Y}_1+\underline{H}_2\underline{Y}_2) = \underline{M}^{-1}(\underline{H}_1'\underline{H}_1\hat{\beta}_1+\underline{H}_2'\underline{H}_2\hat{\underline{\beta}}_2) \ ;$$

$$\hat{\underline{\beta}}_2 = (\underline{H}_2'\underline{H}_2)^{-1}\ \underline{H}_2'\underline{Y}_2 \ ;$$

$$\nu\ S^2 = (\underline{Y}_1-\underline{H}_1\hat{\underline{\beta}})(\underline{Y}_1-\underline{H}_1\hat{\underline{\beta}})+(\underline{Y}_2-\underline{H}_2\hat{\underline{\beta}})'(\underline{Y}_2-\underline{H}_2\hat{\underline{\beta}}) \ ;$$

$$\nu = n_1 + n_2 - 2 \ ;$$

The posterior density of $(\beta_o, \beta_1, \sigma)$ is:

$$f(\beta_o, \beta_1, \sigma | Y_o, \underline{Y}_1, \underline{Y}_2) \propto \frac{1}{\sigma^{n_1+n_2-1}} \exp\{-\frac{1}{2\sigma^2}[\nu S^2 + (\underline{\beta}-\hat{\underline{\beta}})' \underline{M} (\underline{\beta}-\hat{\underline{\beta}})]\} .$$

As a consequence, Y_{n_2+1} is distributed like t with $\nu = n_1 + n_2 - 2$ degrees of freedom, mean value

$$\hat{Y}_{n_2+1} = \tilde{\underline{H}}\, \hat{\underline{\beta}} = \hat{\beta}_o + \hat{\beta}_1\, Y_{n_2}$$

and variance (for $\nu > 2$):

$$\sigma^2_{n_2+1} = \frac{\nu}{\nu-2} S^2 (1 + \tilde{\underline{H}}\, \underline{M}^{-1} \tilde{\underline{H}}') = \frac{n_1+n_2-2}{n_1+n_2-4} S^2 \left[1 + \frac{a + 2bY_{n_2} + (n_1+n_2) Y_{n_2}^2}{(n_1+n_2)a - b^2} \right]$$

where

$$a = \sum_{i=0}^{n_1} Y_{1_i}^2 + \sum_{i=1}^{n_2-1} Y_i^2$$

and

$$b = \sum_{i=0}^{n_1} Y_{1_i} + \sum_{i=0}^{n_2-1} Y_i \quad .$$

After making the appropriate substitutions, prediction intervals of given content and penalty ratios are found in the same way as for a diffuse prior distribution (see Paragraph III.1.1). The practical effect of the conjugate prior is to enlarge the data set, the effective number of observations passing from n_2 to $n_1 + n_2$.

III.1.4 Conjugate prior; β_o and β_1 unknown, σ known

In this case the conjugate prior distribution is normal, with density:

$$f(\underline{\beta}) \propto \exp\{-\frac{1}{2\sigma^2} (\underline{\beta}-\hat{\underline{\beta}}_1)' \underline{H}_1' \underline{H}_1 (\underline{\beta}-\hat{\underline{\beta}}_1)\}$$

where $\underline{\beta} = [\beta_o, \beta_1]'$,

$$\underline{H}_1 = \begin{pmatrix} 1 & Y_{1_o} \\ \vdots & \vdots \\ 1 & Y_{1_{n_1-1}} \end{pmatrix},$$

$(\underline{H}_1'\underline{H}_1)^{-1}$ is a positive-definite symmetric matrix, and $\hat{\underline{\beta}}_1$ is the prior mean of $\underline{\beta}$: $\hat{\underline{\beta}}_1 = (\underline{H}_1'\underline{H}_1)^{-1}\underline{H}_1'[Y_{1_1}, \ldots, Y_{1_{n_1}}]'$.

Although one might give directly the prior mean vector $\hat{\underline{\beta}}_1$ and the prior covariance matrix $\underline{\Sigma}_1 = \sigma^2(\underline{H}_1'\underline{H}_1)^{-1}$, one would find some difficulties in interpreting the physical meaning of these prior parameters. Instead it seems more natural to specify the "best estimates" of the values $Y_{1_o}, Y_{1_1}, \ldots, Y_{1_{n_1}}$. (Besides, in any case one must give $Y_{1_{n_1}}$, the last realization before additional observations are made.)

Denote $\underline{Y} = [Y_1, \ldots, Y_{n_2}]'$ the actual sample and let:

$$\underline{H}_2 = \begin{pmatrix} 1 & Y_{1_{n_1}} \\ 1 & Y_1 \\ \vdots & \vdots \\ 1 & Y_{n_2} \end{pmatrix} .$$

Given $\underline{Y}$, the posterior distribution of $\underline{\beta}$ is also normal, with density

$$f(\underline{\beta}|\underline{Y}) \propto \exp\{-\frac{1}{2\sigma^2}(\underline{\beta}-\hat{\underline{\beta}})' \underline{M}(\underline{\beta}-\hat{\underline{\beta}})\} ,$$

where

$$\underline{M} = \underline{H}_1' \underline{H}_1 + \underline{H}_2' \underline{H}_2$$

$$\hat{\underline{\beta}} = \underline{M}^{-1}(\underline{H}_1' \underline{H}_1 \hat{\underline{\beta}}_1 + \underline{H}_2' \underline{H}_2 \hat{\underline{\beta}}_2)$$

As a consequence, the predictive distribution of Y_{n_2+1} is normal, with mean value $\hat{\beta}_o + \hat{\beta}_1 Y_{n_2}$ and variance

$$\sigma^2(1+\underset{\sim}{H}\ \underline{M}^{-1}\underset{\sim}{H}') = \sigma^2\left\{1 + \frac{a+2bY_n+(n_1+n_2)Y_n^2}{(n_1+n_2)a - b^2}\right\} \quad .$$

For the definition of a and b see the case with σ unknown in Paragraph III.1.3. Obvious replacements provide prediction intervals and penalty factors from the results for diffuse prior. Again, the prior being in conjugate form has the practical effect of increasing the number of observations, from n_2 to $n_1 + n_2$.

The foregoing results hold when the initial state Y_o and the following sequence of n realizations: $Y_1,\ldots,Y_n$ are known. No restriction is imposed on β_1 (as it is clear from the type of "diffuse" prior used); in particular $|\beta_1|$ could be larger than 1, corresponding to an unstable process (see, e.g., Box and Jenkins (1970), p. 346).

Sometimes it is more appropriate for physical reasons to assume that sampling is from a stationary and stable process, in which case the initial conditions at time t = 0 - so far given by Y_o and therefore deterministic - are expressed in probabilistic terms, through the (unknown) marginal normal distribution of the process (see Zellner

(1971), p. 188 and references therein). As might be expected, Zellner finds that in this case the prediction effects of modifying the initial conditions and of varying the prior distribution of the unknown parameters decrease as the sample size increases.

If the prior is assumed of the following truncated type, which is appropriate for a stable process:

$$f(\beta_o,\beta_1,\sigma) \propto \begin{cases} 1/\sigma & \text{for } |\beta_1| < 1 \ , \\ 0 & \text{for } |\beta_1| \geq 1 \ , \end{cases}$$

the posterior density of (β_o,β_1,σ) is of the same (but truncated) form as was found in Paragraph III.1.1 (see equation (III.3), with non-zero values only in the interval $|\beta_1| < 1$.

III.2 SIMPLE PREDICTION AT LEAD ℓ

If distinct realizations of the sequence, say Y_i and Y_j with $i \neq j$, were independent, the predictive distribution of $Y_{n+\ell}$, given Y_o, $Y_1,\ldots,Y_n$ would be the same for all $\ell \geq 1$. This is not the case with autoregressive sequences.

When fitting a first-order autoregressive model to statistical data the parameter of interest is generally β_1 (see equation (III.1)).

For the purpose of exemplifying prediction at a time lag ℓ from the last observation (prediction of $Y_{n+\ell}$), we consider the case when β_o is known (in particular assume $\beta_o = 0$), while both σ and β_1 are unknown. The following are unbiased estimators of β_1 and σ^2:

$$\hat{\beta}_1 = \frac{\underline{H}'\underline{Y}}{\underline{H}'\underline{H}} \quad , \text{ where } \quad \underline{H} = \begin{pmatrix} Y_o \\ Y_1 \\ \vdots \\ Y_{n-1} \end{pmatrix} \quad ; \quad \underline{Y} = \begin{pmatrix} Y_1 \\ \cdot \\ \cdot \\ \cdot \\ Y_n \end{pmatrix} \quad ,$$

and

$$S^2 = \frac{1}{\nu} (\underline{Y}-\underline{H}\ \hat{\underline{\beta}}_1)'(\underline{Y}-\underline{H}\ \hat{\underline{\beta}}_1) \quad ; \quad \text{where } \nu = n - 1.$$

Given Y_o, β_1 and σ, the joint distribution of $Y_1,\ldots,Y_n$ is:

$$f(\underline{Y}|Y_o,\beta_1,\sigma) \propto \frac{1}{\sigma^n} \exp\{-\frac{1}{2\sigma^2}(\underline{Y}-\underline{H}\,\beta_1)'(\underline{Y}-\underline{H}\,\beta_1)\}$$

$$\propto \frac{1}{\sigma^n} \exp\{-\frac{1}{2\sigma^2}[\nu S^2+(\beta_1-\hat{\beta}_1)^2\,\underline{H}'\underline{H}]\} \quad .$$

When seen as a function of β_1 and σ for given Y_o and $\underline{Y}$, this function is proportional to the parameters likelihood. For the "noninformative" prior density:

$$f(\beta_1,\sigma) \propto 1/\sigma$$

the joint posterior density of β_1 and σ is then:

$$f(\beta_1,\sigma|Y_o,\underline{Y}) \propto \frac{1}{\sigma^{n+1}} \exp\{-\frac{1}{2\sigma^2}[\nu S^2 + (\beta_1-\hat{\beta}_1)^2\,\underline{H}'\underline{H}\} \quad \text{(III.11)}$$

We introduce now the conditional independence approximation:

$$f(\beta_1,\sigma|Y_o,\underline{Y}) \approx f(\beta_1|Y_o,\underline{Y}) \cdot f(\sigma|Y_o,\underline{Y}) \quad .$$

Integrating the density (III.11) with respect to σ one finds that $(\beta_1 | Y_o, \underline{Y})$ has a marginal t-distribution with $\nu = n - 1$ degrees of freedom, mean value $\hat{\beta}_1$ and variance (for $n > 3$):

$$\frac{\nu}{\nu-2} s^2 (\underline{H}'\underline{H})^{-1} \quad ;$$

i.e., that

$$f(\beta_1 | Y_o, \underline{Y}) \propto [\nu s^2 + \frac{(\beta_1 - \hat{\beta}_1)^2}{\underline{H}'\underline{H}}]^{-n/2} \quad .$$

Integrating the density (III.11) with respect to β_1, $(\sigma | Y_o, \underline{Y})$ is found to have inverted-gamma density:

$$f(\sigma | Y_o, \underline{Y}) \propto \frac{1}{\sigma^n} \exp\left[- \frac{(n-1)s^2}{2\sigma^2} \right] \quad .$$

For given $\hat{\beta}_1$, $Y_{n+\ell}$ can be estimated by

$$\hat{Y}_{n+\ell} | \hat{\beta}_1 = \hat{\beta}_1^{\ell} \cdot Y_n \quad .$$

The mean value of $Y_{n+\ell}$, given the autoregressive parameter β_1, is

$$\hat{Y}_{n+\ell|\beta_1} = \beta_1^{\ell} Y_n \quad ,$$

and the prediction error when using the estimate $\hat{Y}_{n+\ell|\hat{\beta}_1}$ can be written

$$e_{\ell|\hat{\beta}_1} = Y_{n+\ell} - \hat{Y}_{n+\ell|\hat{\beta}_1} = e_{\ell|\beta_1} + (\beta_1^{\ell} - \hat{\beta}_1^{\ell})\, Y_n \quad ,$$

where $$e_{\ell|\beta_1} = Y_{n+\ell} - \beta_1^{\ell} Y_n \quad .$$

Since for given σ and β_1 the second moment of the prediction error is

$$E[e^2_{\ell|\hat{\beta}_1}|\sigma,\beta_1] = \sigma^2 \frac{1-\beta_1^{2\ell}}{1-\beta_1^2} + (\beta_1^{\ell} - \hat{\beta}_1^{\ell})^2 Y_n^2 \quad ,$$

the unconditional second moment of the prediction error can be written (assuming approximate independence of σ and β_1):

$$E[e_\ell^2] = \int_0^\infty \int_{-\infty}^\infty \{\sigma^2 \frac{1-\beta_1^{2\ell}}{1-\beta_1^2} + (\beta_1^\ell - \hat{\beta}_1^\ell)^2 Y_n^2\}$$

$$\cdot f(\beta_1|Y_o,\underline{Y}) f(\sigma|Y_o,\underline{Y}) d\beta_1 d\sigma$$

$$= \int_{-\infty}^\infty \{S^2 \frac{n-1}{n-3} \frac{1-\beta_1^{2\ell}}{1-\beta_1^2} + (\beta_1^\ell - \hat{\beta}_1^\ell)^2 Y_n^2\}$$

$$\cdot f(\beta_1|Y_o,\underline{Y}) d\beta_1 \quad .$$

The t-density of $(\beta_1|Y_o,\underline{Y})$ is nonzero in the range $(-\infty,\infty)$; therefore as ℓ increases the integral diverges (the second term in the integrand approaches infinity for $|\beta_1| > 1$). However, in many cases it is known that the process is stable, i.e., that $|\beta_1| < 1$. For a prior distribution of (β_1,σ) with constant density $1/\sigma$ over the interval $\beta_1 \varepsilon (-1,1)$ and zero density elsewhere the posterior density of β_1 is truncated t, being zero for $|\beta_1| \geq 1$. In this case the second moment of the prediction error is finite:

$$E[e_\ell^2] = S^2 \cdot \gamma_1(\ell,\hat{\beta}_1,n) + Y_n^2 \cdot \gamma_2(\ell,\hat{\beta}_1,n), \qquad \text{(III.12)}$$

where

$$\gamma_1(\cdot) = \frac{n-1}{n-3}\ \frac{1}{c}\int_{-1}^{1} \frac{1-\beta_1^{2\ell}}{1-\beta_1^2}\left[1 + \frac{(\beta_1-\hat{\beta}_1)^2}{(n-1)S_{\beta_1}^2}\right]^{-n/2} d\,\beta_1 \quad ,$$

$$\gamma_2(\cdot) = \frac{1}{c}\int_{-1}^{1} (\beta_1^{\ell}-\hat{\beta}_1^{\ell})^2\left[1 + \frac{(\beta_1-\hat{\beta}_1)^2}{(n-1)S_{\beta_1}^2}\right]^{-n/2} d\,\beta_1 \quad ,$$

$$c = \int_{-1}^{1}\left[1 + \frac{(\beta_1-\hat{\beta}_1)^2}{(n-1)S_{\beta_1}^2}\right]^{-n/2} d\,\beta_1 \quad .$$

S_{β_1} in equation (III.12) is the scale parameter of the posterior density of β_1, which satisfies:

$$(n-1)S_{\beta_1}^2 = \frac{(\underline{Y}'\underline{Y})(\underline{H}'\underline{H})-(\underline{H}'\underline{Y})^2}{(\underline{H}'\underline{H})^2} = \frac{\underline{Y}'\underline{Y}}{\underline{H}'\underline{H}} - \hat{\beta}_1^2 \approx 1 - \beta_1^2 \quad .$$

This final approximation is accurate for moderate to large values of n. When $\ell = 1$ one finds:

$$\gamma_1(\cdot) = \frac{n-1}{n-3} \quad ; \quad \gamma_2(\cdot) \approx \mathrm{Var}(\beta_1 | \hat{\beta}_1) \approx \frac{1}{n-3}(1 - \hat{\beta}_1^2) \quad .$$

Table 2 collects values of $\gamma_1(\cdot)$ and $\gamma_2(\cdot)$ for $\ell = 1(1)10$, $\hat{\beta}_1 = 0(0.3)0.9$, and for $n = 4(1)10,15,20,30,60$.

The second term in equation (III.12) accounts explicitly for the "initial conditions", Y_n, and is zero for $Y_n = E[Y_n] = 0$. In any case it is: $\gamma_2 \to 0$ as $\ell \to \infty$, due to the condition $|\beta_1| < 1$. On the contrary, when $\hat{\beta}_1$ and n are fixed, γ_1 increases with ℓ, approaching a finite asymptote when $\ell \to \infty$. One may therefore conclude that the main contribution to the prediction error variance comes from the first term in equation (III.12), unless Y_n is significantly different from zero and at the same time ℓ and n are very small.

When $n \to \infty$ it is: $\hat{\beta}_1 \to \beta_1$ and $S^2 \to \sigma^2$ with probability 1, and the second moment of the prediction error approaches the prediction variance under perfect information:

$$\lim_{n\to\infty} E[e_\ell^2] = \sigma^2_{PI}(\ell) = \sigma^2 \frac{1-\beta_1^{2\ell}}{1-\beta_1^2} ,$$

irrespective of the "initial conditions", Y_n.

For finite n, the penalty ratio (for imperfect statistical information) on the prediction error mean square is:

$$r^2_{\beta_1,\sigma}(\ell,\hat{\beta}_1,n,S^2,Y_n)$$

$$= \frac{1-\hat{\beta}_1^2}{1-\hat{\beta}_1^{2\ell}} \left[\gamma_1(\ell,\hat{\beta}_1,n) + \frac{Y_n^2}{S^2} \gamma_2(\ell,\hat{\beta}_1,n) \right] \quad .$$

Denoting $\overline{Y}_n = Y_n(1-\hat{\beta}_1^2)^{1/2}/S$ the deviation of Y_n from $E[Y_n] = 0$ in units of standard deviations (and in terms of the statistics S^2 and $\hat{\beta}_1$), the r.m.s. penalty ratio becomes:

$$r_{\beta_1,\sigma}(\ell,\hat{\beta}_1,n,\overline{Y}_n) = [r_1^2(\ell,\hat{\beta}_1,n) + \overline{Y}_n^2 \, r_2^2(\ell,\hat{\beta}_1,n)]^{1/2} \qquad \text{(III.13)}$$

where

$$r_1^2(\ell,\hat{\beta}_1,n) = \frac{1-\hat{\beta}_1^2}{1-\hat{\beta}_1^{2\ell}} \gamma_1(\ell,\hat{\beta}_1,n) \ ,$$

$$r_2^2(\ell,\hat{\beta}_1,n) = \frac{1}{1-\hat{\beta}_1^{2\ell}} \gamma_2(\ell,\hat{\beta}_1,n) \quad .$$

Values of the function $r_{\beta_1,\sigma}(\ell,\hat{\beta}_1,n,\overline{Y}_n)$ for selected values of the arguments are collected in Table 3.

It should be emphasized that the penalty ratio (III.13) refers to the r.m.s. of the prediction error. The construction of prediction regions of given content requires knowledge of the full distribution of $Y_{n+\ell}$. In approximation one might assume that the prediction error $e_\ell = Y_{n+\ell} - \hat{\beta}_1^{\ell} Y_n$ has t-distribution with n-1 degrees of freedom. This approximation is motivated by the fact that e_ℓ has indeed t-distribution for $\ell = 1$, and for any ℓ when β_1 is known to be zero (=>independent normal sequence). The distribution of e_ℓ is no longer of t-type when $\ell \neq 1$ and β_1 is unknown; however the approximation should be useful for practical purposes, at least when the sample size n is moderate to large, so that β_1 has small posterior variance.

Under this distribution assumption the penalty ratio for one-sided prediction intervals of P-content at lead ℓ is:

$$r(P,\ell,\hat{\beta}_1,n,\overline{Y}_n) = \left(\frac{n-3}{n-1} \right)^{1/2} \frac{t_{n-1}(P)}{\Phi(P)} \, r_{\beta_1,\sigma}(\ell,\hat{\beta}_1,n,\overline{Y}_n) \quad \text{(III.14)}$$

(for two-sided central predictions, replace P by P/2).

Table 4 collects the factors $[(n-3)/(n-1)]^{1/2} t_{n-1}(P)/\Phi(P)$ for n = 4(1)10,15,20,30,60 and for P = 0.99, 0.999, 0.9999.

An application follows, including an approximate extension to simultaneous prediction.

EXAMPLE. AN APPLICATION TO TIME-DEPENDENT STRUCTURAL RELIABILITY ANALYSIS

Consider a structural loading process, modeled as a stable normal Markov chain with known initial state Y_o, $\beta_o = 0$, unknown "noise" variance σ^2 and unknown regression coefficient β_1. The realization of the process from Y_1 to Y_n is available, yielding the unbiased parameters estimates S^2 and $\hat{\beta}_1$. Given the ultimate resistance of the system, Y_{MAX}, we want the probability of failure before or at time $n+\ell$, where $\ell = 1,2,\ldots$.

Before addressing this problem, let us first consider the prediction distribution at lead ℓ and the resistance level which corresponds to a given reliability P at time $n+\ell$. This is the P-fractile of the prediction distribution at lead ℓ; note that if $\ell > 1$, no consideration is given yet to the behavior (i.e., failure or not) of the system in the interval $[n+1, n+\ell-1]$.

Under perfect knowledge of σ^2 and β_1 the mean forecast at lead ℓ is: $\hat{Y}_{n+\ell} = \beta_1^{\ell} Y_n$, and the left-hand prediction interval of P-content is:

$$(-\infty,\ \beta_1^{\ell} Y_n + \sigma\ \Phi(P)\{(1 - \beta_1^{2\ell})/(1 - \beta_1^2)\}^{1/2}] \qquad \text{(III.15a)}$$

For the case of limited information with unknown σ^2 and β_1 the same interval is approximately:

$$(-\infty,\ \hat{\beta}_1^{\ell} Y_n + S\left(\frac{n-3}{n-1}\right)^{1/2} t_{n-1}(P) \cdot r_{\beta_1,\sigma}(\ell,\hat{\beta}_1,n,\bar{Y}_n)$$

$$\cdot\ (1 - \hat{\beta}_1^{2\ell})/(1 - \hat{\beta}_1^2)^{1/2}] \ . \qquad \text{(III.15b)}$$

For instance, if $P = 0.999$, $n = 10$, $S^2 = 1$, $\hat{\beta}_1 = 0.3$, $Y_n = 1.048$ ($\Rightarrow \bar{Y}_n = 1$), the interval (III.15b) becomes a function of ℓ only:

$$(-\infty,\ 1.048 \times 0.3^{\ell} + 3.973 \cdot r_{\beta_1,\sigma}(\ell,0.3,10,1)\{1-0.3^{2\ell}\}^{1/2}] \ .$$

The upper limits of these prediction intervals and the corresponding penalty ratios, equation (III.14), are collected

in Table 5 for $\ell = 1(1)10$. The values for perfect information are obtained by replacing $n = 10$ by $n = \infty$. As one might expect, the penalty increases with ℓ, approaching an asymptotic value as $\ell \to \infty$.

We pass now to the problem of simultaneous prediction which arises in the reliability analysis of the system during the time interval $(n, n+\ell]$. In Chapter II, in which white sequences were considered, it was shown that for simultaneous prediction intervals with large probability content (= reliability) P and with small correlation of the prediction variables, the hypothesis of independent future realizations produced accurate approximations. In that case correlation was due exclusively to the parameter uncertainty. If we make the same assumption of independence in this chapter, the predictive joint distribution of Y_{n+1}, $\ldots, Y_{n+\ell}$ is simply the product of the marginal predictive distributions which, due to the nonstationarity of the Markov model, depend on ℓ. For perfect knowledge of the parameters β_1 and σ, such as would result from a sample of size $n \to \infty$, and assuming approximate independence, the probability of failure before the $(n+\ell+1)^{th}$ loading event is:

$$P_{f_{n=\infty}}(\ell) = 1 - \prod_{i=1}^{\ell} \Phi(d_i) \quad ,$$

where $d_i = (Y_{max} - \beta_1^i \cdot Y_n)[(1 - \beta_1^2)/(1 - \beta_1^{2i})]^{1/2}/\sigma$.

The same probability when σ and β_1 are estimated from a sample of size n is, in approximation:

$$P_{f_n}(\ell) = 1 - \prod_{i=1}^{\ell} t_{n-1}(d_i) ,$$

where

$$d_i = (Y_{max} - \hat{\beta}_1^i Y_n) \frac{[\frac{n-1}{n-3}(1-\hat{\beta}_1^2)/(1-\hat{\beta}_1^{2i})]^{1/2}}{S \cdot r_{\beta_1,\sigma}(i,\hat{\beta}_1,n,\overline{Y}_n)} .$$

In the present case it is: $S = 1$, $\hat{\beta}_1 = 0.3$, $Y_n = 1.048$, $\overline{Y}_n = 1$. For a resistance $Y_{max} = 3.50$, 4.50, 5.50 the evolution of the failure probability with ℓ is shown in Table 6 for the cases $n = \infty$ (perfect information) and $n = 10$.

The ratio $P_{f_{n=10}}(\ell)/P_{f_{n=\infty}}(\ell)$ between the probabilities of failure in the two hypotheses of imperfect and perfect information is nonnegligible, and increases dramatically with Y_{max} (with the system reliability). The ratio is approximately 13 for $Y_{max} = 3.50$, 150 for $Y_{max} = 4.50$, and 5000 for $Y_{max} = 5.50$. These results agree qualitatively with earlier

findings for independent sequences.

The last development of the present example shows an approximate extension of the results for simple prediction to solve simultaneous prediction problems. The extension rests on the assumption of independence between Y_{n+i} and Y_{n+j} when $i \neq j$. This assumption should not be critical when $|\beta_1| << 1$ as in the example above, and when constructing simultaneous prediction intervals of large content P, as it is often the case in structural reliability analysis and design. For example, in the case studied numerically, the assumption of independence should have little (and conservative) effects on the analysis of the system with resistance $Y_{max} = 5.50$.

An easy, although tedious way of checking the error introduced by the assumption of independence might be to simulate numerically sequences of future realizations: $Y_{n+1}, Y_{n+2}, \ldots, Y_{n+\ell}$ for given σ^2, β_1 and Y_n, and to compare the resulting reliability for different system resistance levels with the approximate analytical results, as given by Table 6.

III.3 NOISY UNCENSORED AND CENSORED OBSERVATIONS IN MARKOV PROCESSES; AN APPLICATION TO DETERIORATING SYSTEMS

In this section we consider the estimation and prediction of the resistance R(t) of a deteriorating structural system. R(t) is a stochastic process, which is observed at discrete times through noisy sampling (generating uncensored noisy observations) or through proof loading tests (generating censored data). More precisely, proof loading experiments generate observations which are censored at an unknown level; in fact they are in the form

$R(t_i) > S_i$ if the system survives the application of a load S_i at time t_i; or

$R(t_i) = S_i$ if the system fails at time t_i under a load intensity S_i.

S_i may be deterministic or random, depending on the modalities of the experiment. The resistance R(t) is modeled as a Markov process, which is described next in greater detail.

III.3.1 The resistance model

A stochastic function $X(t)$ which is uniquely defined by its random initial value $X_o = X(t_o)$ is called a "purely deterministic random process". The estimation (and prediction) of this degenerate class of random processes was treated in Chapter I. A purely deterministic random model for the resistance of deteriorating systems was proposed and studied by Turkstra (1970). However, in presence of damage accumulation and wear-out, and particularly for systems operating under random environmental conditions, knowledge of $R(t_o)$ is not sufficient to define deterministically the process $R(t)$ at any later time. This fact leads to modeling $R(t)$ as a (nondeterministic) random process.

Two phenomena are particularly important in the deteriorating process: (i) the weakening of the structure due to damage accumulation; this phenomenon is caused by the repeated action of external loads; (ii) the natural wear-out (or strengthening) of the materials in absence of damage from external loads; the latter phenomenon is due to aging, and is controlled by environmental conditions like temperature, humidity, concentration of corrosive chemicals, etc. (see Gertsbakh and Kordonskiy (1969)). Depending on the circumstances one of these two phenomena may prevail on the other. In combination, they typically generate resistance

curves of the type shown in Figure 1(a) and (b).

In case (a) the structure weakens considerably during the initial life period (0 - 1), after which adaptation takes place (1 - 2). A third phase is characterized by rapid wear-out due to aging, and terminates with failure (point 3). In case (b) the same features are distinctive of the last two phases, but the system experiences an initial strengthening. These curves represent possible mean trends of the resistance process; uncertainty is present, however, particularly in the last phase of rapid wear-out.

A mathematical model for this behavior might be a nonstationary Markov process with an absorbing barrier at $R = R_{failure}$. If the process is observed only at discrete times, say at $t = 0, 1, 2, \ldots$ (for instance: 0 = after construction; 1 = after 1 year from construction; 2 = after 2 years, etc.), one might use the following continuous state, discrete time nonstationary model:

$$R_t = \bar{R}_t + a_t(R_{t-1} - \bar{R}_{t-1}) + \varepsilon_t \quad , \qquad \text{(III.16)}$$

where $\{\bar{R}_t\}$ is the sequence of a priori mean values at times $t = 0, 1, 2, \ldots$; $\{a_t\}$ is a sequence of known coefficients and $\{\varepsilon_t\} \sim \{(0, \sigma_{\varepsilon_t}^2)\}$ is a sequence of zero-mean second-moment variables, independent among themselves and independent of

the resistance process.

If $\sigma^2_{\varepsilon_t} = 0$ for all t, the model degenerates into a purely deterministic random process, in that a specific and unique history $\{R_t\}$ corresponds to each given initial value R_o. Depending on the coefficients $\{a_t\}$ the random process may be of explosive type ($a_t > 1$ for all t), may propagate the initial "error" $R_o - \bar{R}_o$ without modification ($a_t = 1$), or may damp out the initial "error" in a geometric fashion ($0 < a_t < 1$); see Figure 2. Cases with $a_t < 0$ are also shown in Figure 2.

When instead $\{\sigma^2_{\varepsilon_t}\} \neq 0$, the sequence $\{\varepsilon_t\}$ produces random perturbations around the trends indicated in Figure 2.

Another particular case generated by equation (III.16) is when $\{a_t\} = \{0\}$. The model becomes: $R_t = \bar{R}_t + \varepsilon_t$, corresponding to independent random deviations from the mean trend.

The foregoing discussion shows the wide spectrum of models generated by equation (III.16). Appropriate choices of the coefficients $\{a_t\}$ and the introduction of nonstationarities in the disturbances may induce qualitatively and quantitatively different characteristics within separate time intervals, if so desired.

A similar model for the random resistance of deteriorating structures has been proposed recently by Kameda and Koike (1974).

III.3.2 Propagation of the mean and of the variance

(a) No observation

Let equation (III.16) define a second-moment Markov process, with initial mean $\hat{R}_o$ and initial variance σ_o^2. If no observation is made in the interval (0,t], the mean and the variance at lead t are found through the recursive (predictive) relations:

$$\begin{cases} \hat{R}_j = \bar{R}_j + a_j(\hat{R}_{j-1} - \bar{R}_{j-1}) \quad , & \text{(III.17a)} \\ \\ \sigma_j^2 = a_j^2 \, \sigma_{j-1}^2 + \sigma_{\varepsilon j}^2 \quad ; \quad j = 1,2,\ldots,t, & \text{(III.17b)} \end{cases}$$

which, as it is easy to verify, yield:

$$\begin{cases} \hat{R}_t = \bar{R}_t + (\hat{R}_o - \bar{R}_o) \prod_{i=1}^{t} a_i \quad , & \text{(III.18a)} \\ \\ \sigma_t^2 = \sigma_o^2 \left(\prod_{i=1}^{t} a_i^2 \right) + \sum_{j=1}^{t-1} \sigma_{\varepsilon_j}^2 \left(\prod_{i=j+1}^{t} a_i^2 \right) + \sigma_{\varepsilon_t}^2 \quad . & \text{(III.18b)} \end{cases}$$

The covariance between R_i and R_j, with $i < j$, is:

$$\mathrm{Cov}[R_i, R_j] = \sigma_i^2 \prod_{k=i+1}^{j} a_k \quad . \tag{III.18c}$$

Note that the sign of the correlation depends on the number of negative coefficients a_k (with k ranging from i+1 to j).

For reasons of safety it may be advisable to reduce the variance σ_t^2 in equation (III.18b). This can be done in various ways: the two procedures considered herein are: (i) direct or indirect noisy measurements (the case of exact measurements being a trivial particular case); and (ii) proof loading experiments, yielding censored information. The processing of noisy measurements is studied hereafter, while the use of censored data is postponed until Paragraph III.3.3.

(b) Direct or indirect noisy measurements

The sampling procedure is assumed to be of the following type. At each discrete time t = 1,2,... two courses of action are possible:

(i) To make one or more measurements of the resistance R_t and of related quantities, resulting in a measurement vector $\underline{Z}_t$. In this case it is assumed that $\underline{Z}_t$ is related to R_t as:

$$\underline{Z}_t = \underline{Z}_{o_t} + \underline{h}_t R_t + \underline{\theta}_t \ , \qquad \text{(III.19)}$$

where $\underline{Z}_{o_t}$ and $\underline{h}_t$ are given vectors, and $\underline{\theta}_t \sim (\underline{0}, \underline{\Theta}_t)$ is a second-moment vector of possibly correlated, zero-mean error terms, with covariance matrix $\underline{\Theta}_t$. $\underline{\theta}_i$ and $\underline{\theta}_j$ are assumed independent when $i \neq j$, and independent of the resistance R_t.

(ii) To make no measurement, which corresponds to assuming $\underline{h}_t = \underline{0}$ in equation (III.19).

The scheme above leads to the so-called Kalman filter and Kalman prediction algorithms (for the original derivation see Kalman (1960) and Kalman and Bucy (1961)). Before quoting the explicit expressions for the mean and the variance of the resistance at time t, given the observation vectors $\{\underline{Z}_i, i=1, \ldots, t\}$, it is convenient to generalize equations (III.16), (III.17), (III.18) and (III.19) to the case of the system resistance being described by a vector sequence $\{\underline{R}_t\}$, instead of a scalar sequence as considered so far. Then the Kalman equations will be given for the vector case.

There are many instances in which the resistance is best described by a vector, the obvious one being when the

system is composed of n separate elements, the i^{th} element having resistance R_{i_t} at time t.

Another case in which the vector representation is useful is when at each time t = 0, 1, ... failure may occur in any of several modes, each associated with a specific loading condition. In this case the sequence $\{R_{i_t}\}$ monitors (at discrete times) the resistance in the i^{th} mode.

If equation (III.16) is rewritten:

$$\underline{R}_t = \overline{\underline{R}}_t + \underline{A}_t(\underline{R}_{t-1} - \overline{\underline{R}}_{t-1}) + \underline{B}_t \underline{\varepsilon}_t \qquad \text{(III.20)}$$

where $\{\overline{\underline{R}}_t\}$ is a given n-dimensional sequence; $\{\underline{A}_t\}$ is a given (n X n) matrix sequence, $\{\underline{B}_t\}$ is a known (n X m) matrix sequence, and $\{\underline{\varepsilon}_t\}$ is an independent sequence of $(\underline{0}, \underline{\Theta}_t)$ random m-vectors, equations (III.17) become:

$$\begin{cases} \hat{\underline{R}}_j = \overline{\underline{R}}_j + \underline{A}_j(\hat{\underline{R}}_{j-1} - \overline{\underline{R}}_{j-1}) \quad , & \text{(III.21a)} \\ \underline{\Sigma}_j = \underline{A}_j \underline{\Sigma}_{j-1} \underline{A}_j' + \underline{B}_j \underline{\Theta}_j \underline{B}_j' \quad . & \text{(III.21b)} \end{cases}$$

The stable or unstable, exponential or oscillatory character of the process shown graphically in Figure 2 for

the scalar case depends now on the eigenvalues of the matrices $\underline{A}_t$ (on the poles of the system); for instance, if $\underline{A}_t = \underline{A}$ for all t, the condition for asymptotic stability is that all the eigenvalues of $\underline{A}$ lie inside the unitary disk in the complex plane. A behavior of the type shown in Figure 2 could be associated separately with the component of $\underline{R}_t$ along each eigendirection in n-space.

The stability of the process is linked to the physical significance of the noise terms $\underline{\varepsilon}_t$. If they refer to the uncertainty in the time evolution of the material properties the matrix $\underline{A}$ might realistically produce an unstable process in which the perturbances increase with time. If instead the noise terms refer to the effects of "secondary loads", the matrix $\underline{A}$ might propagate the accumulated damage without magnification.

The generalization of equations (III.18) reads:

$$\hat{\underline{R}}_t = \bar{\underline{R}}_t + \left(\prod_{i=1}^{t} \underline{A}_i\right)\left(\hat{\underline{R}}_o - \bar{\underline{R}}_o\right) \qquad \text{(III.22a)}$$

$$\underline{\Sigma}_t = \left(\prod_{i=1}^{t} \underline{A}_i\right)\underline{\Sigma}_o\left(\prod_{i=1}^{t} \underline{A}_i\right)' + \sum_{j=1}^{t-1} \underline{B}_j\left(\prod_{i=j+1}^{t} \underline{A}_i\right)\underline{\Theta}_j\left(\prod_{i=j+1}^{t} \underline{A}_i\right)' \underline{B}_j' + \underline{B}_t\,\underline{\Theta}_t\,\underline{B}_t' \qquad \text{(III.22b)}$$

$$\left| \quad \mathrm{Cov}[\underline{R}_i, \underline{R}_j] = \left(\prod_{k=i+1}^{j} \underline{A}_k \right) \underline{\Sigma}_i \ , \quad i < j \right. \qquad \text{(III.22c)}$$

Equations (III.22) hold for the case of no observations.

When observations are made, the measurement vector at time t is written, generalizing equation (III.19):

$$\underline{Z}_t = \underline{Z}_{o_t} + \underline{H}_t \, \underline{R}_t + \underline{\theta}_t \ , \qquad \text{(III.23)}$$

where $\underline{Z}_{o_t}$ and $\underline{\theta}_t$ are as before, and $\underline{H}_t$ is a given (m X n) measurement matrix. If no measurement is made at time t, let $\underline{H}_t = \underline{0}$ in equation (III.23) (noninformative measurement).

Kalman filter for discrete-time processes

The Kalman filter for discrete-time processes is a recursive algorithm, each step of which consists of two phases: (i) a prediction phase through which, given the first two moments (or the distribution) of $\underline{R}_{j-1}$ conditional on the observations up to time j-1, one finds the predictive moments (or the predictive distribution) of $\underline{R}_j$, conditional on the same observations; (ii) an estimation (or correction) phase in which one uses the observations (if any) at time j to revise the outcome of the prediction phase. Analytically, for the Markov process (III.20) observed through (III.23)

these two phases at the generic step j are:

(i) Prediction phase

$$\begin{cases} \hat{\underline{R}}_{j|j-1} = \bar{\underline{R}}_j + \underline{A}_j (\hat{\underline{R}}_{j-1|j-1} - \bar{\underline{R}}_{j-1}) \quad , & \text{(III.24a)} \\ \\ \underline{\Sigma}_{j|j-1} = \underline{A}_j \, \underline{\Sigma}_{j-1|j-1} \, \underline{A}_j' + \underline{B}_j \, \underline{\Theta}_j \, \underline{B}_j' \quad , & \text{(III.24b)} \end{cases}$$

where the conditional sign $\cdot|\ell$ stands for:

$$\cdot|\ell \equiv \cdot|\{\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_\ell\} \ .$$

(ii) Estimation (correction) phase:

$$\begin{cases} \hat{\underline{R}}_{j|j} = \hat{\underline{R}}_{j|j-1} + \underline{\Sigma}_{j|j} \, \underline{H}_j' \, \underline{\Theta}_j^{-1} \, (\underline{z}_j - \underline{H}_j \, \hat{\underline{R}}_{j|j-1}), & \text{(III.25a)} \\ \\ \underline{\Sigma}_{j|j} = \underline{\Sigma}_{j|j-1} - \underline{\Sigma}_{j|j-1} \, \underline{H}_j' (\underline{H}_j \, \underline{\Sigma}_{j|j-1} \underline{H}_j' + \underline{\Theta}_j)^{-1} \underline{H}_j \, \underline{\Sigma}_{j|j-1} \cdot & \text{(III.25b)} \end{cases}$$

These last equations are structurally identical with those already encountered in estimation of time-independent models; see Chapter I, equations (I.7) and (I.8). Equations (III.25) above are written in a form which separates the correction (second term) from the forecast (first term). In the case of no measurement ($\underline{H}_j = \underline{0}$) these equations reduce to: $\hat{\underline{R}}_{j|j} = \hat{\underline{R}}_{j|j-1}$, and $\underline{\Sigma}_{j|j} = \underline{\Sigma}_{j|j-1}$. For comments on equations (III.25), see the remarks following equations (I.7) and (I.8) in Chapter I. We recall explicitly the important fact that equation (III.25b) does not contain the information set $\{\underline{Z}_1,\ldots,\underline{Z}_j\}$, so that $\underline{\Sigma}_{j|j}$ is computable before knowing the outcome of the experiment. This fact allows one to compare a priori different sampling strategies and to optimize the sampling experiment if the objective is to minimize, in some sense, the posterior variance.

If the initial moments are in the form $\hat{\underline{R}}_{o|o}$ and $\underline{\Sigma}_{o|o}$ the algorithm starts with a prediction step; if instead possible measurements at time t = 0 are not included in the initial moments, the algorithm starts with an estimation step at t = 0. When the initial conditions are given in the form of a multinormal distribution, and $\{\underline{\varepsilon}_t\}$ and $\{\underline{\theta}_t\}$ are normal sequences, then $\{\underline{R}_{j|j-1}\}$ and $\{\underline{R}_{j|j}\}$ are also normal sequences, so that equations (III.24) and (III.25) characterize fully the resistance process at time j = 1,2,... .

It should also be said that the case of $\underline{R}$ being a continuous Markov process in time can be analyzed in a similar way: if measurements are made continuously in time, the mean of $\underline{R}(t)$ evolves according to a linear vector differential equation of the first order, while the covariance matrix satisfies a forward quadratic differential equation of the Riccati type (see Kalman (1960) and Kalman and Bucy (1961)). Again, while $\underline{\Sigma}(t)$ is precomputable, $\hat{\underline{R}}(t)$ is not. An extension of this classical formulation of the Kalman filter to process censored data is proposed in the next paragraph.

III.3.3 Propagation of the resistance distribution when observations are censored

Consider for simplicity the case with scalar resistance. The logic of the Kalman algorithm is retained when observations are censored; in fact also in this case two computational phases can be distinguished at each discrete observation time: a prediction phase, and an estimation or correction phase. Nevertheless, some basic differences have to be noted.

Since observations have truncating effects on the resistance distribution, the informativeness of an observation like $R_t > S_t$ (or: $R_t > S_t + \theta_t$, with θ_t a noise term) depends strongly on the shape of the distribution of R_t

conditional on the previous observations, and cannot be assessed from the first two moments only. Contrary to the case with uncensored data, also when measurement errors, initial conditions and propagation errors have normal distribution, $\{R_{t|t}\}$ is not, in general, a normal sequence. The conclusion is that when data are in censored form one must work with full probability distributions, which fact makes the estimation and prediction analyses much more elaborate.

Physically one can view the sequence of proof loadings as the creation - at discrete times - of an artificial absorbing barrier for the failure - no failure process (see Figure 3).

The prediction phase is the same whether data are noisy or not. It consists of finding the distribution of $R_{j|j-1}$, given the distribution of $R_{j-1|j-1}$ and the process equation (III.15). In general this requires convolution:

$$F_{R_{j|j-1}}(r) = \int_{-\infty}^{\infty} f_{\varepsilon_j}(\varepsilon) \cdot F_{R_{j-1|j-1}}(\overline{R}_{j-1} + \frac{r-\overline{R}_j-\varepsilon}{a_j}) d\varepsilon. \quad \text{(III.26)}$$

The estimation phase depends on the observation model. If the load intensity S_j is known, the CDF of $R_{j|j}$ is simply:

$$F_{R_j|j}(r) = \begin{cases} F_{R_j|j-1}(r)[1 - F_{R_j|j-1}(S_j)]^{-1} & \text{for } r > S, \\ 0 & \text{otherwise.} \end{cases} \tag{III.27a}$$

If instead the proof load intensity is uncertain, say $S_j = \bar{S}_j + \theta_j$, where S_j and $\bar{S}_j$ are the true and the observed intensity values, respectively, and θ_j denotes the estimation error, a convolution operation is again required:

$$F_{R_j|j}(r) = F_{R_j|j-1}(r) \cdot \int_{-\infty}^{r} [1 - F_{R_j|j-1}(s)]^{-1} f_{S_j}(s) d\, s. \tag{III.27b}$$

In this second case the "truncation effect" decreases with increasing dispersion of θ_j.

The resistance sequence satisfies equations (III.26) and (III.27) if the structure is subjected to no other loads than proof loads. If other environmental actions are present, this would correspond to neglecting them whenever the system survives, and to reconstructing the system if it happens to fail, with the same resistance distribution it had at the moment of collapse. This is, of course, a quite artificious

assumption. Moreover, one would like to profit from the information about the performance of the system during the periods between proof loading experiments. E.g., if the system is known to have survived a violent windstorm, one might want to use this piece of information in judging a posteriori the system resistance. (For an application of this concept to offshore platform structures, see Marshall and Bea (1974).) In general, events of this type produce censored information like $R(t) > S(t)$, and thus can be treated in the same way as was indicated earlier for proof loading. A quantitative difference between "natural" and "artificial" proof loading is that for the former events the uncertainty on the load intensity is larger (e.g., the maximum wind pressure during a windstorm, the peak ground acceleration during an earthquake, and so on), besides the fact that modeling natural events as single load applications may be an oversimplification.

If a natural event of large magnitude has occurred close to (but before) a planned proof loading experiment, one may decide to cancel the experiment or to modify it in order to make it more informative when combined with the data already available. If instead the natural event was of small magnitude, its effect on the subsequent reliability may be neglected, and the artificial proof loading may be performed as on schedule.

The study of first-order autoregressive models does not exhaust the problem of quantifying statistical uncertainty in random processes with memory. This class of Markov models was chosen because analytically tractable and because recurrent in stochastic modeling of engineering systems. The vast literature on Kalman filtering and Kalman prediction (see for example the Special Issue on Linear-Quadratic Gaussian Problem, IEEE Transactions on Automatic Control, Vol. AC-16, No. 6, Dec. 1971) addresses to the same type of models.

REFERENCES

1. Box, G.E.P. and Jenkins, G.M. (1970): Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco.

2. Gertsbakh, I.B. and Kordonskiy, Kh.B. (1969): Models of Failure, Springer-Verlag, New York.

3. Kalman, R.E. (1960): "A New Approach to Linear Filtering and Prediction Problems", ASME J. Basic Eng., 82, 34-45.

4. Kalman, R.E. and Bucy, R.S. (1961): "New Results in Linear Filtering and Prediction Theory", ASME J. Basic Eng., 83, 95-107.

5. Kameda, H. and Koike, T. (1974): "Reliability Analysis of Deteriorating Structures". The Japan - U.S. Joint Seminar on Reliability Approach in Structural Engineering; 13-17 May, 1974, Tokyo.

6. Marshall, P.W. and Bea, R.G. (1974): "Failure Modes of Offshore Platforms". ASCE-EMD Specialty Conference, Probabilistic Methods in Engineering, Stanford Univ., June 24-26, 1974.

7. Raiffa, H. and Schlaifer, R. (1961): Applied Statistical Decision Theory. Harvard University Press.

8. Turkstra, C.J. (1970): Theory of Structural Design Decisions. Study No. 2, Solid Mechanics Division, Univ. of Waterloo, Waterloo, Ontario, Canada, page 87.

9. Zellner, A. (1971): An Introduction to Bayesian Inference in Econometrics. Wiley and Sons, New York.

10. Special Issue on Linear-Quadratic-Gaussian Problem, IEEE Trans. on Automatic Control, Vol. AC-16, No. 6, December, 1971.

n	P=0.99	P=0.999	P=0.9999
3	17.662	132.99	1106.94
4	3.667	8.850	23.282
5	2.310	3.912	7.063
6	1.860	2.680	4.046
7	1.641	2.162	2.952
8	1.510	1.884	2.410
9	1.425	1.712	2.098
10	1.364	1.596	1.897
11	1.319	1.512	1.757
12	1.283	1.448	1.653
16	1.196	1.300	1.426
21	1.143	1.212	1.295
26	1.111	1.164	1.228
31	1.092	1.134	1.179
41	1.068	1.098	1.132
61	1.045	1.063	1.084
∞	1	1	1

TABLE 1. Penalty for one-sided prediction intervals of content P at lead 1. First-order autoregressive process with β_o, β_1 and σ unknown. n = available sample size.

$\widehat{BETA1}$ = 0.0

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	3.000	2.000	1.667	1.500	1.400	1.333	1.286	1.167	1.118	1.074	1.035
2	3.623	2.377	1.952	1.733	1.598	1.506	1.438	1.260	1.177	1.111	1.052
3	3.911	2.543	2.072	1.827	1.674	1.568	1.490	1.283	1.186	1.115	1.053
4	4.082	2.640	2.139	1.878	1.713	1.599	1.515	1.293	1.188	1.115	1.053
5	4.197	2.703	2.183	1.910	1.738	1.619	1.531	1.298	1.189	1.115	1.053
6	4.279	2.749	2.214	1.933	1.756	1.632	1.541	1.301	1.189	1.115	1.053
7	4.341	2.782	2.237	1.949	1.768	1.642	1.548	1.303	1.189	1.115	1.053
8	4.389	2.808	2.254	1.962	1.777	1.649	1.554	1.305	1.189	1.115	1.053
9	4.426	2.828	2.268	1.972	1.785	1.654	1.558	1.306	1.189	1.115	1.053
10	4.456	2.844	2.279	1.979	1.790	1.659	1.561	1.307	1.189	1.115	1.053

(a) $\gamma_1(\ell, 0, n)$

$\widehat{BETA1}$ = 0.0

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	1.000	0.500	0.333	0.250	0.200	0.167	0.143	0.083	0.059	0.037	0.018
2	0.096	0.083	0.072	0.062	0.054	0.046	0.040	0.020	0.008	0.003	0.001
3	0.057	0.048	0.041	0.034	0.029	0.024	0.020	0.008	0.002	0.000	0.000
4	0.038	0.032	0.026	0.022	0.018	0.015	0.012	0.004	0.000	0.000	0.000
5	0.027	0.023	0.018	0.015	0.012	0.010	0.008	0.003	0.000	0.000	0.000
6	0.021	0.017	0.014	0.011	0.009	0.007	0.006	0.002	0.000	0.000	0.000
7	0.016	0.013	0.010	0.008	0.007	0.005	0.004	0.001	0.000	0.000	0.000
8	0.013	0.010	0.008	0.007	0.005	0.004	0.003	0.001	0.000	0.000	0.000
9	0.010	0.008	0.007	0.005	0.004	0.003	0.003	0.001	0.000	0.000	0.000
10	0.008	0.007	0.005	0.004	0.003	0.003	0.002	0.001	0.000	0.000	0.000

(b) $\gamma_2(\ell, 0, n)$

Table 2. First-order autoregressive sequence. Coefficients $\gamma_1(\ell, \hat{\beta}_1, n)$ and $\gamma_2(\ell, \hat{\beta}_1, n)$ in the expression for the second moment of the prediction error at lead ℓ, equation (III.12).

$\widehat{BETA1}$ = 0.30

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	3.000	2.000	1.667	1.500	1.400	1.333	1.286	1.167	1.118	1.074	1.035
2	3.684	2.432	2.009	1.795	1.664	1.576	1.512	1.346	1.273	1.207	1.145
3	4.010	2.631	2.163	1.923	1.776	1.675	1.602	1.408	1.319	1.237	1.163
4	4.207	2.749	2.252	1.996	1.838	1.730	1.650	1.438	1.338	1.247	1.166
5	4.339	2.828	2.310	2.043	1.878	1.764	1.680	1.456	1.347	1.251	1.167
6	4.434	2.884	2.351	2.076	1.905	1.788	1.700	1.467	1.353	1.252	1.167
7	4.506	2.925	2.382	2.101	1.926	1.805	1.715	1.475	1.356	1.253	1.168
8	4.561	2.957	2.405	2.119	1.941	1.818	1.726	1.480	1.359	1.253	1.168
9	4.605	2.983	2.424	2.134	1.953	1.828	1.734	1.485	1.360	1.254	1.168
10	4.640	3.003	2.439	2.145	1.962	1.835	1.741	1.488	1.361	1.254	1.168

(a) $\gamma_1(\ell, 0.3, n)$

$\widehat{BETA1}$ = 0.30

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	0.910	0.455	0.303	0.227	0.182	0.152	0.130	0.076	0.054	0.034	0.016
2	0.076	0.069	0.063	0.058	0.054	0.050	0.046	0.034	0.024	0.014	0.006
3	0.062	0.055	0.049	0.044	0.040	0.036	0.033	0.022	0.014	0.007	0.002
4	0.042	0.038	0.034	0.030	0.027	0.025	0.022	0.014	0.008	0.003	0.001
5	0.032	0.028	0.025	0.022	0.019	0.017	0.016	0.009	0.005	0.001	0.000
6	0.024	0.021	0.018	0.016	0.014	0.013	0.011	0.007	0.003	0.001	0.000
7	0.018	0.016	0.014	0.012	0.011	0.010	0.009	0.005	0.002	0.000	0.000
8	0.015	0.013	0.011	0.010	0.009	0.007	0.007	0.004	0.001	0.000	0.000
9	0.012	0.010	0.009	0.008	0.007	0.006	0.005	0.003	0.001	0.000	0.000
10	0.010	0.008	0.007	0.006	0.005	0.005	0.004	0.002	0.001	0.000	0.000

(b) $\gamma_2(\ell, 0.3, n)$

Table 2 (continued). First-order autoregressive sequence. Coefficients $\gamma_1(\ell, \hat{\beta}_1, n)$ and $\gamma_2(\ell, \hat{\beta}_1, n)$ in the expression for the second moment of the prediction error at lead ℓ, equation (III.12).

$\widehat{BETA1}$ = 0.60

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	3.000	2.000	1.667	1.500	1.400	1.333	1.286	1.167	1.118	1.074	1.035
2	3.906	2.622	2.206	1.996	1.870	1.796	1.739	1.599	1.541	1.482	1.419
3	4.378	2.946	2.489	2.255	2.113	2.036	1.975	1.816	1.749	1.670	1.577
4	4.674	3.148	2.666	2.416	2.263	2.185	2.119	1.946	1.868	1.770	1.648
5	4.878	3.286	2.787	2.526	2.364	2.286	2.217	2.031	1.944	1.829	1.683
6	5.026	3.385	2.876	2.607	2.437	2.359	2.287	2.091	1.996	1.868	1.701
7	5.138	3.460	2.943	2.667	2.492	2.415	2.341	2.136	2.035	1.894	1.712
8	5.224	3.519	2.996	2.714	2.535	2.458	2.383	2.170	2.064	1.913	1.717
9	5.293	3.565	3.038	2.752	2.568	2.493	2.416	2.197	2.088	1.928	1.721
10	5.349	3.602	3.072	2.783	2.596	2.521	2.443	2.220	2.106	1.939	1.723

(a) $\gamma_1(\ell, 0.6, n)$

$\widehat{BETA1}$ = 0.60

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	0.640	0.320	0.213	0.160	0.128	0.107	0.091	0.053	0.038	0.024	0.011
2	0.070	0.068	0.066	0.064	0.062	0.060	0.058	0.050	0.042	0.032	0.016
3	0.068	0.062	0.058	0.055	0.053	0.053	0.052	0.047	0.041	0.031	0.014
4	0.044	0.044	0.046	0.045	0.044	0.046	0.045	0.041	0.037	0.027	0.011
5	0.039	0.038	0.039	0.039	0.037	0.039	0.039	0.036	0.032	0.022	0.008
6	0.030	0.030	0.033	0.033	0.031	0.033	0.033	0.030	0.027	0.018	0.005
7	0.026	0.026	0.028	0.027	0.026	0.028	0.028	0.025	0.022	0.015	0.004
8	0.021	0.021	0.023	0.023	0.022	0.024	0.024	0.021	0.019	0.012	0.003
9	0.018	0.017	0.019	0.019	0.018	0.020	0.020	0.018	0.016	0.010	0.002
10	0.015	0.014	0.016	0.016	0.015	0.017	0.017	0.015	0.013	0.008	0.001

(b) $\gamma_2(\ell, 0.6, n)$

Table 2 (continued). First-order autoregressive sequence. Coefficients $\gamma_1(\ell, \hat{\beta}_1, n)$ and $\gamma_2(\ell, \hat{\beta}_1, n)$ in the expression for the second moment of the prediction error at lead ℓ, equation (III.12).

$\widehat{BETA1}$ = 0.90

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	3.000	2.000	1.667	1.500	1.400	1.333	1.286	1.167	1.118	1.074	1.035
2	4.576	3.166	2.706	2.470	2.332	2.240	2.176	2.020	1.964	1.911	1.865
3	5.595	3.957	3.432	3.161	3.003	2.902	2.832	2.671	2.624	2.577	2.539
4	6.311	4.535	3.974	3.680	3.513	3.408	3.337	3.184	3.153	3.117	3.092
5	6.841	4.976	4.395	4.087	3.916	3.810	3.739	3.597	3.585	3.563	3.552
6	7.248	5.324	4.733	4.416	4.242	4.136	4.067	3.939	3.945	3.937	3.938
7	7.568	5.606	5.011	4.687	4.512	4.407	4.341	4.225	4.249	4.254	4.266
8	7.824	5.839	5.244	4.915	4.739	4.636	4.572	4.469	4.510	4.526	4.546
9	8.033	6.033	5.441	5.109	4.933	4.832	4.770	4.680	4.737	4.762	4.788
10	8.205	6.198	5.611	5.276	5.101	5.001	4.942	4.863	4.935	4.969	4.998

(a) $\gamma_1(\ell, 0.9, n)$

$\widehat{BETA1}$ = 0.90

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	0.190	0.095	0.063	0.048	0.038	0.032	0.027	0.016	0.011	0.007	0.003
2	0.144	0.107	0.082	0.068	0.058	0.050	0.045	0.028	0.020	0.014	0.008
3	0.165	0.128	0.104	0.089	0.079	0.070	0.063	0.043	0.032	0.024	0.015
4	0.162	0.132	0.111	0.098	0.088	0.080	0.073	0.053	0.042	0.032	0.021
5	0.151	0.127	0.110	0.099	0.090	0.083	0.077	0.059	0.048	0.038	0.026
6	0.135	0.117	0.104	0.095	0.088	0.082	0.077	0.061	0.052	0.043	0.030
7	0.119	0.105	0.095	0.088	0.083	0.078	0.074	0.061	0.054	0.046	0.034
8	0.103	0.093	0.086	0.081	0.077	0.073	0.070	0.060	0.055	0.048	0.036
9	0.088	0.081	0.077	0.073	0.070	0.068	0.066	0.058	0.055	0.049	0.038
10	0.075	0.071	0.068	0.066	0.064	0.062	0.061	0.056	0.054	0.049	0.039

(b) $\gamma_2(\ell, 0.9, n)$

Table 2 (continued). First-order autoregressive sequence. Coefficients $\gamma_1(\ell, \hat{\beta}_1, n)$ and $\gamma_2(\ell, \hat{\beta}_1, n)$ in the expression for the second moment of the prediction error at lead ℓ, equation (III.12).

$\widehat{BETA1} = 0.0 \qquad \overline{YN} = 0.0$

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	1.732	1.414	1.291	1.225	1.183	1.155	1.134	1.080	1.057	1.036	1.017
2	1.903	1.542	1.397	1.317	1.264	1.227	1.199	1.122	1.085	1.054	1.026
3	1.978	1.595	1.439	1.352	1.294	1.252	1.220	1.133	1.089	1.056	1.026
4	2.020	1.625	1.463	1.370	1.309	1.265	1.231	1.137	1.090	1.056	1.026
5	2.049	1.644	1.478	1.382	1.318	1.272	1.237	1.139	1.090	1.056	1.026
6	2.069	1.658	1.488	1.390	1.325	1.278	1.241	1.141	1.091	1.056	1.026
7	2.084	1.668	1.496	1.396	1.330	1.281	1.244	1.142	1.091	1.056	1.026
8	2.095	1.676	1.501	1.401	1.333	1.284	1.246	1.142	1.091	1.056	1.026
9	2.104	1.682	1.506	1.404	1.336	1.286	1.248	1.143	1.091	1.056	1.026
10	2.111	1.687	1.510	1.407	1.338	1.288	1.249	1.143	1.091	1.056	1.026

$\widehat{BETA1} = 0.0 \qquad \overline{YN} = 1.00$

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	2.000	1.581	1.414	1.323	1.265	1.225	1.195	1.118	1.085	1.054	1.026
2	1.928	1.568	1.423	1.340	1.285	1.246	1.216	1.131	1.089	1.056	1.026
3	1.992	1.610	1.453	1.364	1.305	1.262	1.229	1.136	1.090	1.056	1.026
4	2.030	1.635	1.472	1.378	1.316	1.270	1.236	1.139	1.090	1.056	1.026
5	2.055	1.651	1.484	1.388	1.323	1.276	1.240	1.140	1.090	1.056	1.026
6	2.074	1.663	1.493	1.394	1.328	1.280	1.244	1.141	1.091	1.056	1.026
7	2.087	1.672	1.499	1.399	1.332	1.283	1.246	1.142	1.091	1.056	1.026
8	2.098	1.679	1.504	1.403	1.335	1.286	1.248	1.143	1.091	1.056	1.026
9	2.106	1.684	1.508	1.406	1.337	1.287	1.249	1.143	1.091	1.056	1.026
10	2.113	1.688	1.511	1.408	1.339	1.289	1.250	1.143	1.091	1.056	1.026

Table 3. First-order autoregressive sequence. Penalty ratios $r_{\beta_1,\sigma}(\ell,\hat{\beta}_1,n,\overline{Y}_n)$ in equation (III.13).

$\widehat{BETA1} = 0.0$ $\overline{YN} = 2.00$

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	2.646	2.000	1.732	1.581	1.483	1.414	1.363	1.225	1.163	1.106	1.051
2	2.002	1.646	1.497	1.408	1.346	1.300	1.264	1.158	1.100	1.060	1.027
3	2.035	1.654	1.495	1.401	1.337	1.290	1.253	1.147	1.093	1.057	1.026
4	2.058	1.663	1.498	1.402	1.336	1.288	1.250	1.145	1.091	1.056	1.026
5	2.075	1.671	1.502	1.404	1.337	1.288	1.250	1.144	1.091	1.056	1.026
6	2.088	1.678	1.506	1.406	1.338	1.289	1.250	1.144	1.091	1.056	1.026
7	2.099	1.683	1.509	1.408	1.340	1.290	1.251	1.144	1.091	1.056	1.026
8	2.107	1.688	1.512	1.410	1.341	1.290	1.252	1.144	1.091	1.056	1.026
9	2.113	1.691	1.515	1.412	1.342	1.291	1.252	1.144	1.091	1.056	1.026
10	2.119	1.694	1.516	1.413	1.343	1.292	1.253	1.144	1.091	1.056	1.026

$\widehat{BETA1} = 0.0$ $\overline{YN} = 3.00$

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	3.464	2.550	2.160	1.936	1.789	1.683	1.604	1.384	1.283	1.186	1.092
2	2.118	1.768	1.612	1.514	1.443	1.387	1.341	1.200	1.118	1.068	1.029
3	2.104	1.726	1.561	1.461	1.389	1.335	1.292	1.165	1.097	1.058	1.026
4	2.104	1.711	1.542	1.440	1.369	1.316	1.274	1.154	1.092	1.056	1.026
5	2.108	1.705	1.533	1.430	1.360	1.307	1.266	1.150	1.091	1.056	1.026
6	2.113	1.703	1.529	1.426	1.355	1.302	1.262	1.148	1.091	1.056	1.026
7	2.118	1.702	1.527	1.423	1.352	1.300	1.259	1.147	1.091	1.056	1.026
8	2.122	1.703	1.526	1.421	1.351	1.298	1.258	1.146	1.091	1.056	1.026
9	2.125	1.703	1.525	1.421	1.350	1.297	1.257	1.146	1.091	1.056	1.026
10	2.128	1.704	1.525	1.420	1.349	1.297	1.257	1.145	1.091	1.056	1.026

Table 3 (continued).

$\widehat{EETA1} = 0.30$ $\quad \overline{YN} = 0.0$

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	1.732	1.414	1.291	1.225	1.183	1.155	1.134	1.080	1.057	1.036	1.017
2	1.838	1.494	1.358	1.283	1.236	1.202	1.178	1.111	1.081	1.052	1.025
3	1.911	1.548	1.403	1.323	1.272	1.235	1.208	1.133	1.096	1.062	1.029
4	1.957	1.582	1.431	1.348	1.293	1.255	1.225	1.144	1.103	1.065	1.030
5	1.987	1.604	1.450	1.364	1.307	1.267	1.236	1.151	1.107	1.067	1.031
6	2.009	1.620	1.463	1.375	1.317	1.275	1.244	1.155	1.110	1.068	1.031
7	2.025	1.632	1.472	1.383	1.324	1.281	1.249	1.158	1.111	1.068	1.031
8	2.037	1.640	1.480	1.389	1.329	1.286	1.253	1.161	1.112	1.068	1.031
9	2.047	1.647	1.485	1.393	1.333	1.290	1.256	1.162	1.113	1.068	1.031
10	2.055	1.653	1.490	1.397	1.336	1.292	1.259	1.164	1.113	1.068	1.031

$\widehat{BETA1} = 0.30$ $\quad \overline{YN} = 1.00$

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	2.000	1.581	1.414	1.323	1.265	1.225	1.195	1.118	1.085	1.054	1.026
2	1.859	1.517	1.381	1.306	1.257	1.223	1.197	1.127	1.092	1.059	1.028
3	1.927	1.566	1.421	1.340	1.287	1.250	1.221	1.142	1.102	1.065	1.030
4	1.967	1.594	1.443	1.359	1.304	1.264	1.234	1.150	1.107	1.067	1.031
5	1.995	1.613	1.458	1.372	1.315	1.274	1.243	1.155	1.109	1.068	1.031
6	2.015	1.626	1.469	1.380	1.322	1.280	1.248	1.158	1.111	1.068	1.031
7	2.029	1.636	1.477	1.387	1.328	1.285	1.253	1.161	1.112	1.068	1.031
8	2.041	1.644	1.483	1.392	1.332	1.289	1.256	1.162	1.113	1.068	1.031
9	2.050	1.651	1.488	1.396	1.336	1.292	1.258	1.164	1.113	1.068	1.031
10	2.057	1.656	1.492	1.399	1.338	1.294	1.260	1.165	1.113	1.068	1.031

Table 3 (continued).

$\widehat{BETA1} = 0.30$ $\quad \overline{YN} = 2.00$

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	2.646	2.000	1.732	1.581	1.483	1.414	1.363	1.225	1.163	1.106	1.051
2	1.920	1.584	1.449	1.372	1.321	1.283	1.255	1.171	1.125	1.080	1.037
3	1.975	1.618	1.472	1.389	1.333	1.293	1.262	1.171	1.121	1.074	1.033
4	2.000	1.629	1.478	1.392	1.335	1.293	1.261	1.168	1.118	1.071	1.032
5	2.019	1.638	1.483	1.395	1.337	1.294	1.261	1.167	1.116	1.069	1.031
6	2.032	1.645	1.488	1.398	1.338	1.295	1.262	1.167	1.115	1.069	1.031
7	2.043	1.651	1.491	1.400	1.340	1.296	1.263	1.167	1.115	1.068	1.031
8	2.051	1.656	1.494	1.403	1.342	1.298	1.264	1.167	1.115	1.068	1.031
9	2.058	1.660	1.497	1.404	1.343	1.299	1.264	1.167	1.114	1.068	1.031
10	2.064	1.663	1.499	1.406	1.344	1.300	1.265	1.168	1.114	1.068	1.031

$\widehat{BETA1} = 0.30$ $\quad \overline{YN} = 3.00$

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	3.464	2.550	2.160	1.936	1.789	1.683	1.604	1.384	1.283	1.186	1.092
2	2.017	1.690	1.555	1.475	1.420	1.378	1.345	1.241	1.178	1.113	1.052
3	2.052	1.701	1.553	1.467	1.407	1.362	1.326	1.216	1.152	1.089	1.038
4	2.052	1.686	1.534	1.445	1.385	1.340	1.304	1.198	1.136	1.078	1.033
5	2.057	1.680	1.524	1.434	1.373	1.327	1.292	1.187	1.127	1.073	1.032
6	2.061	1.677	1.518	1.427	1.365	1.320	1.284	1.181	1.122	1.070	1.031
7	2.065	1.675	1.515	1.422	1.360	1.315	1.279	1.177	1.119	1.069	1.031
8	2.069	1.675	1.513	1.420	1.357	1.312	1.276	1.175	1.118	1.069	1.031
9	2.073	1.675	1.512	1.418	1.356	1.310	1.275	1.174	1.117	1.069	1.031
10	2.076	1.675	1.511	1.417	1.354	1.309	1.273	1.173	1.116	1.068	1.031

Table 3 (continued).

$\widehat{BETA1} = 0.60$ $\quad \overline{YN} = 0.0$

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	1.732	1.414	1.291	1.225	1.183	1.155	1.134	1.080	1.057	1.036	1.017
2	1.695	1.389	1.274	1.212	1.173	1.149	1.131	1.084	1.065	1.044	1.021
3	1.714	1.406	1.293	1.230	1.191	1.169	1.151	1.104	1.083	1.059	1.029
4	1.744	1.431	1.317	1.254	1.214	1.193	1.174	1.125	1.103	1.073	1.036
5	1.772	1.454	1.340	1.275	1.234	1.213	1.195	1.143	1.119	1.085	1.041
6	1.795	1.474	1.358	1.293	1.250	1.230	1.211	1.158	1.132	1.094	1.045
7	1.814	1.489	1.373	1.307	1.263	1.244	1.224	1.170	1.142	1.101	1.047
8	1.829	1.501	1.385	1.318	1.274	1.254	1.235	1.179	1.150	1.107	1.049
9	1.841	1.510	1.394	1.327	1.282	1.263	1.244	1.186	1.156	1.111	1.050
10	1.850	1.518	1.402	1.335	1.289	1.270	1.250	1.192	1.161	1.114	1.050

$\widehat{BETA1} = 0.60$ $\quad \overline{YN} = 1.00$

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	2.000	1.581	1.414	1.323	1.265	1.225	1.195	1.118	1.085	1.054	1.026
2	1.718	1.416	1.303	1.242	1.202	1.179	1.160	1.110	1.087	1.061	1.030
3	1.735	1.429	1.316	1.254	1.214	1.193	1.175	1.126	1.103	1.074	1.036
4	1.757	1.447	1.335	1.272	1.232	1.212	1.194	1.144	1.119	1.086	1.041
5	1.783	1.467	1.354	1.291	1.249	1.229	1.211	1.159	1.133	1.096	1.045
6	1.804	1.484	1.370	1.306	1.263	1.244	1.225	1.171	1.143	1.103	1.047
7	1.821	1.497	1.383	1.317	1.274	1.255	1.236	1.180	1.151	1.108	1.049
8	1.835	1.508	1.393	1.327	1.282	1.264	1.245	1.188	1.158	1.112	1.050
9	1.845	1.516	1.401	1.334	1.289	1.271	1.252	1.194	1.163	1.115	1.050
10	1.854	1.523	1.408	1.341	1.295	1.277	1.257	1.198	1.167	1.118	1.051

Table 3 (continued).

$\widehat{BETA1} = 0.60$ $\quad \overline{YN} = 2.00$

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	2.646	2.000	1.732	1.581	1.483	1.414	1.363	1.225	1.163	1.106	1.051
2	1.787	1.496	1.388	1.328	1.288	1.264	1.244	1.185	1.152	1.111	1.056
3	1.796	1.495	1.383	1.321	1.281	1.261	1.243	1.190	1.161	1.118	1.057
4	1.795	1.492	1.386	1.326	1.286	1.268	1.251	1.198	1.169	1.123	1.057
5	1.816	1.506	1.397	1.335	1.293	1.277	1.259	1.204	1.174	1.126	1.056
6	1.829	1.514	1.406	1.343	1.300	1.283	1.265	1.209	1.178	1.127	1.055
7	1.842	1.523	1.413	1.348	1.304	1.288	1.270	1.212	1.180	1.128	1.054
8	1.852	1.529	1.418	1.353	1.308	1.292	1.273	1.214	1.182	1.128	1.053
9	1.860	1.533	1.422	1.356	1.310	1.294	1.275	1.216	1.183	1.128	1.053
10	1.866	1.537	1.425	1.359	1.312	1.296	1.277	1.217	1.184	1.129	1.052

$\widehat{BETA1} = 0.60$ $\quad \overline{YN} = 3.00$

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	3.464	2.550	2.160	1.936	1.789	1.683	1.604	1.384	1.283	1.186	1.092
2	1.896	1.621	1.519	1.460	1.418	1.395	1.372	1.299	1.253	1.190	1.098
3	1.892	1.600	1.489	1.427	1.385	1.368	1.349	1.289	1.251	1.188	1.092
4	1.856	1.565	1.468	1.410	1.370	1.357	1.340	1.283	1.246	1.182	1.083
5	1.869	1.568	1.466	1.406	1.364	1.352	1.335	1.277	1.240	1.174	1.075
6	1.870	1.564	1.463	1.402	1.359	1.347	1.329	1.270	1.234	1.167	1.068
7	1.877	1.564	1.461	1.398	1.354	1.342	1.324	1.264	1.227	1.160	1.063
8	1.880	1.562	1.458	1.394	1.349	1.337	1.318	1.258	1.221	1.154	1.059
9	1.883	1.562	1.456	1.391	1.345	1.332	1.314	1.252	1.216	1.150	1.057
10	1.885	1.561	1.454	1.388	1.341	1.329	1.310	1.248	1.212	1.146	1.055

Table 3 (continued).

$\widehat{BETA1} = 0.90 \qquad \overline{YN} = 0.0$

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	1.732	1.414	1.291	1.225	1.183	1.155	1.134	1.080	1.057	1.036	1.017
2	1.590	1.323	1.223	1.168	1.135	1.113	1.096	1.057	1.042	1.027	1.015
3	1.506	1.267	1.180	1.132	1.104	1.085	1.072	1.041	1.032	1.022	1.015
4	1.451	1.230	1.151	1.108	1.083	1.066	1.055	1.031	1.026	1.020	1.016
5	1.413	1.205	1.132	1.092	1.069	1.054	1.044	1.024	1.023	1.020	1.018
6	1.385	1.187	1.120	1.081	1.060	1.046	1.038	1.021	1.022	1.021	1.021
7	1.365	1.175	1.111	1.075	1.054	1.042	1.034	1.020	1.023	1.024	1.025
8	1.351	1.167	1.106	1.071	1.051	1.040	1.033	1.021	1.026	1.027	1.030
9	1.340	1.161	1.103	1.069	1.050	1.039	1.033	1.023	1.029	1.032	1.035
10	1.332	1.158	1.102	1.068	1.050	1.040	1.034	1.026	1.033	1.037	1.040

$\widehat{BETA1} = 0.90 \qquad \overline{YN} = 1.00$

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	2.000	1.581	1.414	1.323	1.265	1.225	1.195	1.118	1.085	1.054	1.026
2	1.717	1.436	1.316	1.250	1.207	1.177	1.154	1.095	1.069	1.047	1.027
3	1.619	1.370	1.270	1.213	1.177	1.152	1.133	1.084	1.065	1.047	1.030
4	1.546	1.321	1.233	1.183	1.152	1.130	1.115	1.075	1.061	1.047	1.033
5	1.492	1.283	1.205	1.159	1.132	1.113	1.100	1.068	1.058	1.048	1.037
6	1.452	1.254	1.182	1.141	1.116	1.100	1.088	1.062	1.057	1.050	1.042
7	1.421	1.232	1.165	1.127	1.104	1.090	1.080	1.059	1.057	1.052	1.046
8	1.397	1.215	1.153	1.116	1.095	1.082	1.073	1.057	1.058	1.056	1.051
9	1.378	1.202	1.143	1.108	1.089	1.077	1.069	1.056	1.060	1.059	1.056
10	1.364	1.192	1.136	1.103	1.084	1.074	1.067	1.056	1.063	1.064	1.061

Table 3 (continued).

$\widehat{BETA1}$ = 0.90 $\overline{YN}$ = 2.00

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	2.646	2.000	1.732	1.581	1.483	1.414	1.363	1.225	1.163	1.106	1.051
2	2.051	1.731	1.564	1.468	1.401	1.351	1.312	1.203	1.149	1.104	1.061
3	1.917	1.642	1.510	1.430	1.374	1.332	1.299	1.205	1.158	1.117	1.074
4	1.801	1.562	1.452	1.385	1.338	1.304	1.276	1.198	1.160	1.124	1.085
5	1.710	1.493	1.399	1.342	1.303	1.274	1.251	1.188	1.159	1.129	1.093
6	1.635	1.435	1.354	1.303	1.270	1.246	1.228	1.177	1.156	1.132	1.101
7	1.575	1.387	1.315	1.270	1.241	1.221	1.206	1.166	1.153	1.134	1.107
8	1.526	1.348	1.283	1.242	1.217	1.200	1.188	1.157	1.150	1.136	1.113
9	1.487	1.315	1.256	1.219	1.197	1.183	1.172	1.149	1.148	1.138	1.118
10	1.454	1.289	1.235	1.200	1.181	1.168	1.160	1.142	1.146	1.140	1.123

$\widehat{BETA1}$ = 0.90 $\overline{YN}$ = 3.00

ℓ \ n	4	5	6	7	8	9	10	15	20	30	60
1	3.464	2.550	2.160	1.936	1.789	1.683	1.604	1.384	1.283	1.186	1.092
2	2.511	2.134	1.907	1.773	1.675	1.600	1.540	1.364	1.270	1.193	1.116
3	2.332	2.016	1.841	1.732	1.651	1.588	1.537	1.383	1.299	1.224	1.144
4	2.159	1.897	1.757	1.668	1.602	1.550	1.508	1.379	1.309	1.242	1.165
5	2.020	1.789	1.674	1.600	1.546	1.504	1.470	1.365	1.309	1.252	1.181
6	1.901	1.695	1.599	1.536	1.492	1.457	1.430	1.346	1.304	1.257	1.193
7	1.803	1.613	1.532	1.479	1.442	1.414	1.392	1.326	1.297	1.259	1.202
8	1.720	1.544	1.474	1.428	1.397	1.375	1.357	1.307	1.289	1.259	1.209
9	1.651	1.486	1.425	1.385	1.359	1.340	1.327	1.289	1.281	1.258	1.214
10	1.594	1.436	1.383	1.348	1.326	1.311	1.300	1.274	1.273	1.257	1.219

Table 3 (continued).

n \ P	0.99	0.999	0.9999
4	1.127	1.909	3.446
5	1.139	1.641	2.478
6	1.121	1.477	2.016
7	1.103	1.376	1.760
8	1.089	1.309	1.604
9	1.078	1.262	1.500
10	1.070	1.227	1.425
15	1.044	1.135	1.244
20	1.033	1.095	1.170
30	1.021	1.060	1.103
60	1.010	1.028	1.049

TABLE 4. Values of $\left(\frac{n-3}{n-1} \right)^{1/2} t_{n-1}(P)/\Phi(P)$ in equation (III.1).

ℓ	PERFECT INFORMATION $n = \infty$	$n = 10$	PENALTY RATIO eq. (III.14)
1	3.404	4.845	1.466
2	3.320	4.833	1.469
3	3.268	4.881	1.498
4	3.247	4.912	1.514
5	3.241	4.943	1.525
6	3.240	4.960	1.531
7	3.240	4.980	1.537
8	3.239	4.992	1.541
9	3.239	5.001	1.544
10	3.239	5.008	1.546

TABLE 5. 0.999 fractiles of the predictive distributions at lead ℓ for a sample size $n = \infty$ and for $n = 10$. Example following equations (III.15) in the text.

ℓ	$Y_{MAX} = 3.50$		$Y_{MAX} = 4.50$		$Y_{MAX} = 5.50$	
	$P_{f_{n=\infty}}(\ell)$	$P_{f_{n=10}}(\ell)$	$P_{f_{n=\infty}}(\ell)$	$P_{f_{n=10}}(\ell)$	$P_{f_{n=\infty}}(\ell)$	$P_{f_{n=10}}(\ell)$
1	7.30 - 4	7.23 - 3	1.42 - 5	1.64 - 3	1.08 - 7	4.20 - 4
2	1.29 - 3	1.37 - 2	2.64 "	3.20 "	2.21 "	8.45 "
3	1.75 "	2.02 "	3.64 "	4.83 "	3.10 "	1.30 - 3
4	2.19 "	2.69 "	4.55 "	6.55 "	3.92 "	1.77 "
5	2.62 "	3.37 "	5.44 "	8.34 "	4.70 "	2.28 "
6	3.04 "	4.05 "	6.32 "	1.01 - 2	5.46 "	2.80 "
7	3.45 "	4.74 "	7.20 "	1.20 "	6.22 "	3.32 "
8	3.87 "	5.44 "	8.08 "	1.38 "	6.98 "	3.86 "
9	4.29 "	6.14 "	8.96 "	1.57 "	7.74 "	4.41 "
10	4.71 "	6.83 "	9.84 "	1.76 "	8.50 "	4.96 "

TABLE 6. Probability of system failure within the next ℓ loading events. Example in Section III.2.

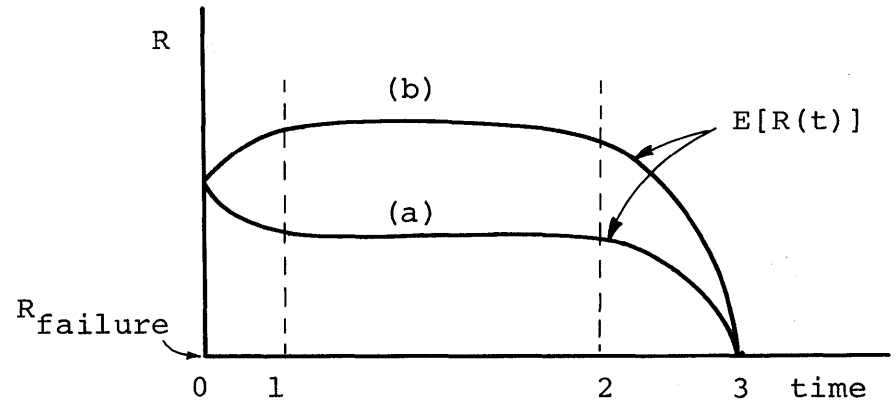


Figure 1. Mean resistance curves in deteriorating systems.

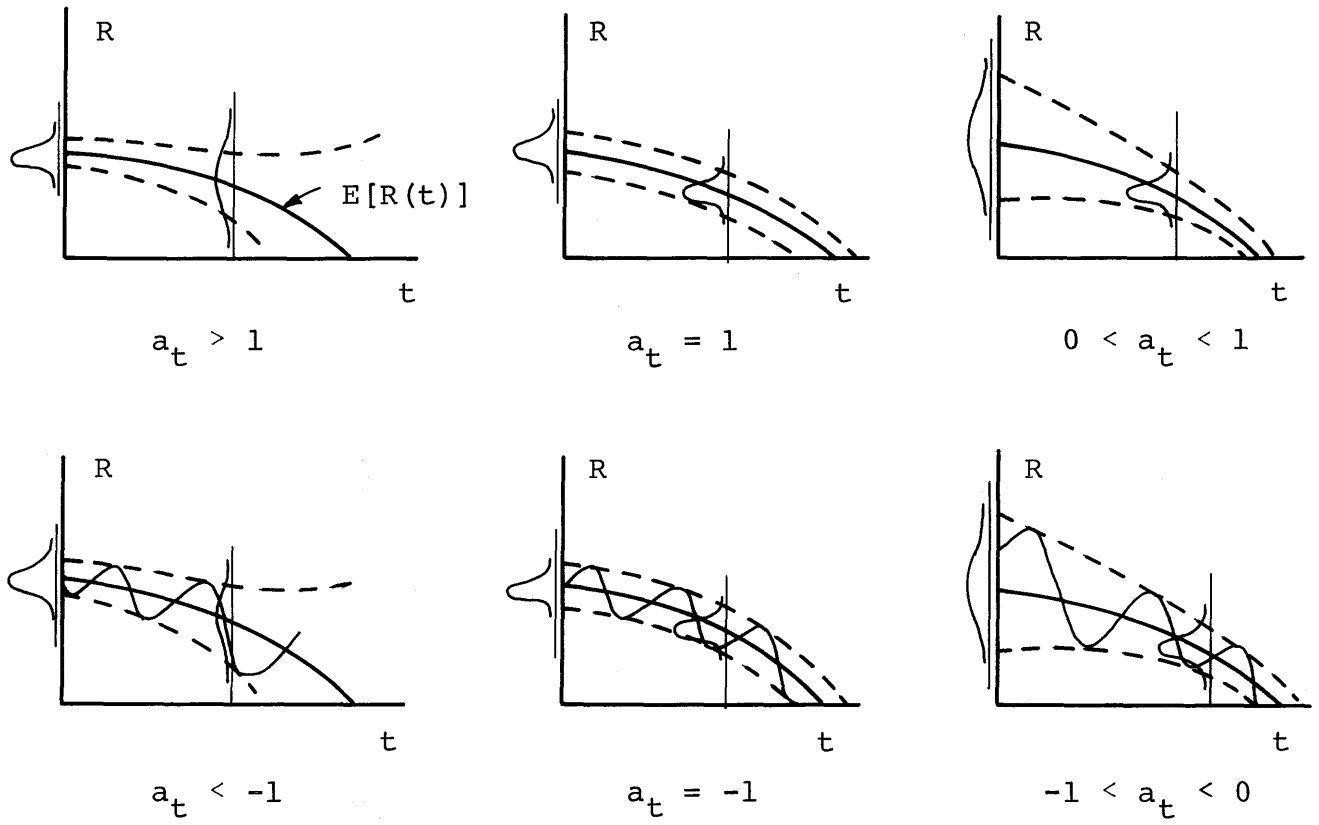


Figure 2. Evolution of the resistance with time in deteriorating systems; dependence on the autoregressive coefficients $\{a_t\}$. See equation (III.16) and following comments.

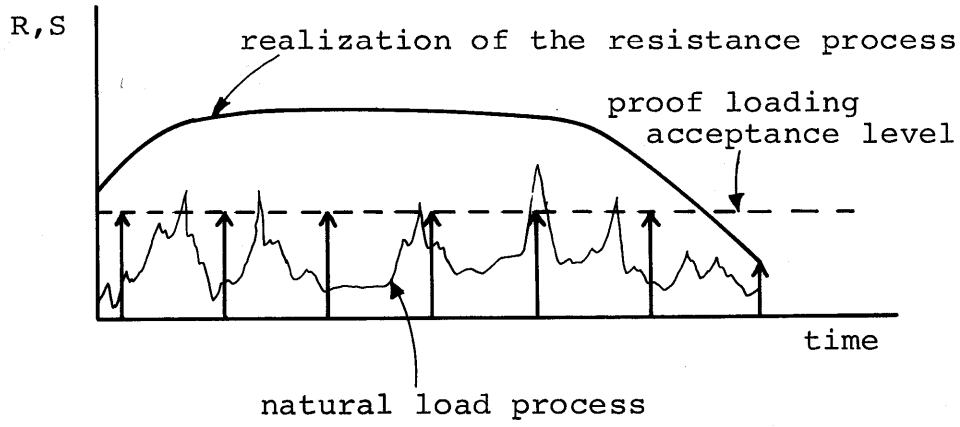


Figure 3. Proof loading of systems with time-dependent resistance.

Biography

The author was born in Florence, Italy, in 1946. From 1964 to 1969 he studied Architecture at the University of Florence where, after defending a thesis on the Control of the Progressive Replanning of Urban Districts, he was conferred the Laurea in Architecture (February, 1970).

In March, 1970 he was appointed Assistant (assistente incaricato) at the Institute of Materials and Structures in the Faculty of Engineering of L'Aquila, Italy. While holding that position, from March, 1970 to September, 1971, he worked on a research project on the reliability of inelastic structures with Prof. Carlo Gavarini.

In September, 1971 he came to M.I.T. as a graduate student in the Division of Structures (Department of Civil Engineering). During the three years of his study at M.I.T. he was granted a research assistantship (with the exception of the Fall term, 1972, during which he was a teaching assistant), allowing him to work on several research projects in the areas of random vibration, structural reliability, seismic risk analysis and optimal sampling, under the guidance of Professors E. H. Vanmarcke, C. A. Cornell and I. Rodriguez-Iturbe.

After winning a national competition held in July, 1972, in February, 1973 he was appointed "assistente

ordinario" at the Institute of Materials and Structures of L'Aquila, mentioned above. He resigned from that office in July, 1974, in the prospective of an offered faculty position in the Division of Constructed Facilities at M.I.T.

The following is the list of publications and Research Reports he has authored or co-authored.

LIST OF PUBLICATIONS

(a) Technical journals

Gavarini, C. and Veneziano, D.: "Calcolo a rottura e programmazione stocastica. Problemi con una variabile casuale", Giornale del Genio Civile, April, 1970.

--: "Sulla teoria probabilistica degli stati limite delle strutture", Giornale del Genio Civile, Nov.-Dec., 1970.

--: "Minimum Weight Limit Design under Uncertainty", Meccanica, Vol. 7, No. 2, 1972.

Veneziano, D.: "Calcolo a rottura probabilistico con più parametri casuali", Giornale del Genio Civile, February, 1971.

(b) Congresses and Conferences

Gavarini, C. and Veneziano, D.: "On the Safety Domain of Structures", Proceedings, 1st National Congress of Theoretical and Applied Mechanics, AIMETA, Udine, June, 1971.

--: "Probabilistic Structural Analysis and Synthesis", 15th Polish Solid Mechanics Conference, Zakopane, September, 1973.

Vanmarcke, E.H. and Veneziano, D.: "Probabilistic Seismic Response of Simple Inelastic Systems", Proceedings, Fifth World Conference on Earthquake Engineering, Rome, 1973.

Veneziano, D.: "Analysis and Design of Hysteretic Structures for Probabilistic Seismic Resistance", Proceedings, Fifth World Conference on Earthquake Engineering, Rome, 1973.

--: "Second Moment Reliability", ASCE, EM Division, Specialty Conference on Probabilistic Methods in Engineering, Stanford Univ., Stanford, Cal., June, 1974.

--: "Safe Regions and Second Moment Reliability Vectors", ASCE, EM Division, Specialty Conference on Probabilistic Methods in Engineering, Stanford Univ., Stanford, Cal., June, 1974.

(c) Research Reports

Gavarini, C. and Veneziano, D.: "Ricerca di un metodo numerico per lo studio del comportamento delle strutture snelle nel campo degli spostamenti finiti:, Publication No. 8, Istituto di Costruzioni, Facoltá di Ingegneria, L'Aquila, Italy, December, 1971.

Veneziano, D. and Vanmarcke, E.H.: "Seismic Damage of Inelastic Systems: a Random Vibration Approach", Res. Report R73-5, Dept. of Civil Engineering, M.I.T., May, 1973.

Veneziano, D.: "Analysis and Design of Simple Inelastic Structures in a Probabilistic Seismic Environment", Res. Report R70-40, Dept. of Civil Engineering, M.I.T., June, 1973.

--: "Contribution to Second Moment Reliability Theory", Res. Report R74-33, Dept. of Civil Engineering, M.I.T., April, 1974.

Veneziano, D. and Cornell, C.A.: "Earthquake Models with Spatial and Temporal Memory for Engineering Seismic Risk Analysis", Res. Report R74-18, Dept. of Civil Engineering, M.I.T., May, 1974.

APPENDIX A

This appendix contains tables of simple-prediction limits for univariate stationary independent sequences with given marginal distribution but one or more unknown parameters. P denotes the desired probability content of the prediction interval and n is the available sample size. In the frequentist approach the tables can be used directly; see Paragraph II.2.1. For using the same tables in Bayesian prediction with "diffuse" or conjugate prior, refer to Paragraph II.2.2.

Values of the penalty ratios corresponding to the same prediction intervals are given also. For their definition and use the reader is referred to Section II.2 in the text.

n	$r_{\mu}^{N}(n)$
1	1.414
2	1.225
3	1.155
4	1.118
5	1.095
6	1.080
7	1.069
8	1.061
9	1.054
10	1.049
11	1.044
12	1.041
13	1.038
14	1.035
15	1.033
16	1.031
17	1.029
18	1.027
19	1.026
20	1.025
30	1.017
40	1.012
60	1.008
120	1.004
∞	1.000

Table A1. Penalty ratios in simple prediction when a sample of size n is available from a normal population with unknown mean and known variance.

n	Probability content, P					
	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
1	2.099	3.839	6.482	24.713	45.358	193.452
2	1.394	1.775	2.195	3.853	5.019	9.602
3	1.237	1.431	1.624	2.268	2.655	3.927
4	1.169	1.296	1.417	1.787	1.994	2.616
5	1.131	1.225	1.311	1.565	1.701	2.087
6	1.107	1.181	1.248	1.439	1.538	1.811
7	1.090	1.152	1.206	1.359	1.436	1.643
8	1.078	1.131	1.177	1.303	1.365	1.532
9	1.069	1.115	1.154	1.262	1.314	1.453
10	1.062	1.102	1.137	1.230	1.276	1.394
11	1.056	1.092	1.123	1.206	1.246	1.348
12	1.051	1.084	1.112	1.186	1.221	1.312
13	1.047	1.077	1.102	1.169	1.201	1.282
14	1.043	1.071	1.094	1.156	1.185	1.258
15	1.040	1.066	1.087	1.144	1.171	1.238
16	1.038	1.062	1.082	1.134	1.159	1.220
17	1.036	1.058	1.076	1.125	1.148	1.205
18	1.033	1.054	1.072	1.117	1.139	1.192
19	1.032	1.051	1.068	1.111	1.131	1.180
20	1.030	1.049	1.064	1.105	1.123	1.170
30	1.020	1.032	1.042	1.068	1.079	1.108
40	1.015	1.024	1.031	1.050	1.059	1.079
60	1.010	1.016	1.020	1.033	1.038	1.051
120	1.005	1.008	1.010	1.016	1.019	1.025
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table A2. Penalty ratios $r_{\sigma}^{N}(P,n)$ for central prediction intervals of P-content. Sampling is from a normal population with known mean and unknown variance; n is the sample size. See equation (II.17) in the text.

n	Probability content, P					
	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
2	2.571	4.702	7.939	30.267	55.552	236.930
3	1.610	2.050	2.535	4.449	5.796	11.087
4	1.383	1.600	1.815	2.535	2.969	4.390
5	1.280	1.420	1.552	1.958	2.185	2.866
6	1.222	1.323	1.417	1.691	1.837	2.254
7	1.183	1.263	1.335	1.539	1.644	1.936
8	1.157	1.222	1.280	1.441	1.523	1.743
9	1.137	1.192	1.240	1.373	1.439	1.615
10	1.121	1.169	1.210	1.323	1.379	1.524
11	1.109	1.151	1.187	1.285	1.333	1.456
12	1.099	1.136	1.169	1.255	1.297	1.403
13	1.091	1.125	1.154	1.231	1.268	1.361
14	1.084	1.115	1.141	1.210	1.244	1.327
15	1.077	1.106	1.130	1.194	1.224	1.299
16	1.072	1.099	1.121	1.179	1.207	1.276
17	1.068	1.092	1.113	1.167	1.192	1.255
18	1.064	1.087	1.106	1.156	1.179	1.238
19	1.060	1.082	1.100	1.146	1.168	1.223
20	1.057	1.077	1.094	1.138	1.159	1.209
21	1.054	1.073	1.089	1.131	1.150	1.197
31	1.036	1.049	1.059	1.085	1.097	1.125
41	1.027	1.036	1.044	1.063	1.071	1.092
61	1.018	1.024	1.029	1.041	1.047	1.060
121	1.009	1.012	1.014	1.020	1.023	1.029
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table A3. Penalty ratios $r^{N}_{\mu,\sigma}(P,n)$ for central prediction intervals of P-content. Sampling is from a normal population with unknown mean and variance; n is the sample size. See equation (II.10).

n	Probability content, P					
	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
1	1.484	2.401	3.838	13.681	24.712	103.013
2	1.211	1.471	1.775	2.994	3.853	7.225
3	1.135	1.278	1.430	1.952	2.267	3.305
4	1.099	1.196	1.296	1.611	1.787	2.321
5	1.079	1.151	1.225	1.447	1.565	1.907
6	1.065	1.123	1.181	1.351	1.439	1.685
7	1.055	1.104	1.152	1.289	1.358	1.549
8	1.047	1.090	1.131	1.245	1.302	1.457
9	1.043	1.079	1.114	1.213	1.262	1.391
10	1.039	1.070	1.102	1.188	1.230	1.341
11	1.034	1.063	1.092	1.169	1.206	1.303
12	1.031	1.058	1.083	1.153	1.186	1.272
13	1.030	1.053	1.077	1.139	1.169	1.247
14	1.027	1.049	1.071	1.128	1.156	1.226
15	1.025	1.046	1.066	1.119	1.144	1.208
16	1.024	1.043	1.061	1.110	1.134	1.193
17	1.022	1.040	1.058	1.104	1.125	1.180
18	1.021	1.037	1.054	1.097	1.117	1.168
19	1.021	1.036	1.051	1.092	1.111	1.158
20	1.019	1.034	1.049	1.087	1.104	1.150
30	1.013	1.022	1.032	1.056	1.068	1.095
40	1.010	1.016	1.024	1.042	1.050	1.070
60	1.007	1.011	1.016	1.028	1.033	1.046
120	1.004	1.005	1.008	1.014	1.016	1.023
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table A4. Penalty ratios $\bar{r}^{N}_{\sigma}(P,n)$ for one-sided prediction intervals of P-content. Sampling is from a normal population with known mean and unknown variance; n is the sample size. See equation (II.19)

n	Probability content, P					
	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
2	1.817	2.941	4.701	16.755	30.265	126.165
3	1.398	1.699	2.050	3.458	4.449	8.343
4	1.269	1.429	1.599	2.183	2.535	3.695
5	1.204	1.310	1.420	1.765	1.958	2.543
6	1.165	1.244	1.323	1.563	1.691	2.060
7	1.139	1.201	1.263	1.445	1.538	1.802
8	1.119	1.171	1.222	1.367	1.441	1.642
9	1.104	1.149	1.192	1.312	1.373	1.535
10	1.094	1.131	1.169	1.272	1.323	1.458
11	1.085	1.118	1.150	1.241	1.285	1.401
12	1.076	1.107	1.136	1.216	1.255	1.356
13	1.070	1.098	1.124	1.196	1.231	1.320
14	1.066	1.090	1.114	1.179	1.210	1.290
15	1.060	1.084	1.106	1.165	1.194	1.266
16	1.057	1.078	1.098	1.153	1.179	1.245
17	1.053	1.073	1.092	1.143	1.167	1.227
18	1.050	1.068	1.087	1.134	1.156	1.212
19	1.047	1.064	1.081	1.126	1.146	1.199
20	1.046	1.061	1.077	1.119	1.138	1.187
21	1.043	1.058	1.073	1.112	1.130	1.177
31	1.030	1.038	1.048	1.073	1.085	1.113
41	1.023	1.029	1.036	1.054	1.062	1.083
61	1.016	1.019	1.024	1.036	1.041	1.054
121	1.009	1.010	1.012	1.018	1.020	1.027
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table A5. Penalty ratios $\bar{r}^{N}_{\mu,\sigma}(P,n)$ for one-sided prediction intervals of P-content. Sampling is from a normal population with unknown mean and variance; n is the sample size. See equation (II.12).

n	Probability content, P					
	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
1	1.627	2.326	2.772	3.643	3.970	4.654
2	1.409	2.014	2.401	3.155	3.438	4.031
3	1.328	1.899	2.263	2.974	3.241	3.800
4	1.286	1.839	2.191	2.880	3.138	3.679
5	1.260	1.802	2.147	2.822	3.075	3.605
6	1.242	1.776	2.117	2.782	3.032	3.555
7	1.230	1.758	2.095	2.754	3.001	3.518
8	1.220	1.744	2.079	2.732	2.977	3.491
9	1.212	1.734	2.066	2.715	2.959	3.469
10	1.206	1.725	2.056	2.702	2.944	3.452
11	1.201	1.718	2.047	2.690	2.932	3.437
12	1.197	1.712	2.040	2.681	2.922	3.425
13	1.194	1.707	2.034	2.673	2.913	3.415
14	1.191	1.702	2.029	2.666	2.905	3.406
15	1.188	1.699	2.024	2.660	2.899	3.399
16	1.186	1.695	2.020	2.655	2.893	3.392
17	1.184	1.692	2.017	2.651	2.888	3.386
18	1.182	1.690	2.014	2.646	2.884	3.381
19	1.180	1.687	2.011	2.643	2.880	3.376
20	1.179	1.685	2.009	2.639	2.876	3.372
30	1.169	1.672	1.993	2.618	2.853	3.345
40	1.165	1.665	1.984	2.608	2.842	3.332
60	1.160	1.658	1.976	2.597	2.830	3.318
120	1.155	1.652	1.968	2.587	2.819	3.305
∞	1.150	1.645	1.960	2.576	2.807	3.291

Table A6. Coefficients $\beta_{\mu}^{N}(P,n)$ in equations (II.13); to be used for constructing central prediction intervals of P-content when a sample of size n is available from a normal population with unknown mean and known variance.

n	Probability content, P					
	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
1	2.414	6.314	12.706	63.657	127.319	636.632
2	1.603	2.920	4.303	9.925	14.089	31.599
3	1.423	2.353	3.182	5.841	7.453	12.923
4	1.344	2.132	2.777	4.604	5.598	8.610
5	1.301	2.015	2.571	4.032	4.773	6.869
6	1.273	1.943	2.447	3.707	4.317	5.959
7	1.254	1.894	2.365	3.499	4.029	5.408
8	1.240	1.860	2.306	3.355	3.832	5.042
9	1.230	1.833	2.262	3.250	3.690	4.781
10	1.221	1.812	2.228	3.169	3.581	4.587
11	1.214	1.796	2.201	3.106	3.497	4.437
12	1.209	1.782	2.179	3.055	3.428	4.317
13	1.204	1.771	2.160	3.012	3.373	4.220
14	1.200	1.761	2.145	2.977	3.326	4.140
15	1.197	1.753	2.131	2.947	3.286	4.073
16	1.194	1.746	2.120	2.921	3.252	4.015
17	1.191	1.740	2.110	2.898	3.222	3.965
18	1.189	1.734	2.101	2.878	3.197	3.922
19	1.187	1.729	2.093	2.861	3.174	3.883
20	1.185	1.725	2.086	2.845	3.153	3.850
30	1.173	1.697	2.042	2.750	3.030	3.646
40	1.167	1.684	2.021	2.704	2.971	3.551
60	1.161	1.671	2.000	2.660	2.915	3.460
120	1.156	1.658	1.980	2.617	2.860	3.373
∞	1.150	1.645	1.960	2.576	2.807	3.291

Table A7. Coefficients $\beta_g^N(P,n)$ in equations (II.16); to be used for constructing central prediction intervals of P-content when a sample of size n is available from a normal population with unknown mean and known variance.

n	Probability content, P					
	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
2	2.957	7.733	15.562	77.963	155.933	779.712
3	1.851	3.372	4.968	11.460	16.269	36.487
4	1.591	2.631	3.558	6.530	8.333	14.448
5	1.473	2.335	3.042	5.044	6.132	9.432
6	1.405	2.176	2.777	4.355	5.156	7.419
7	1.361	2.077	2.616	3.963	4.615	6.370
8	1.330	2.009	2.508	3.712	4.274	5.736
9	1.307	1.960	2.431	3.537	4.040	5.315
10	1.290	1.922	2.372	3.408	3.870	5.015
11	1.276	1.893	2.327	3.310	3.741	4.791
12	1.264	1.869	2.291	3.233	3.639	4.619
13	1.254	1.850	2.261	3.170	3.558	4.480
14	1.246	1.833	2.236	3.118	3.491	4.368
15	1.239	1.819	2.215	3.075	3.435	4.276
16	1.233	1.807	2.197	3.037	3.387	4.198
17	1.228	1.796	2.181	3.005	3.346	4.131
18	1.224	1.787	2.168	2.978	3.311	4.073
19	1.220	1.779	2.156	2.953	3.280	4.024
20	1.216	1.772	2.145	2.932	3.252	3.979
21	1.213	1.765	2.135	2.912	3.228	3.940
31	1.192	1.725	2.075	2.794	3.078	3.704
41	1.182	1.704	2.046	2.737	3.007	3.594
61	1.171	1.684	2.017	2.682	2.938	3.488
121	1.161	1.665	1.988	2.628	2.872	3.387
∞	1.150	1.645	1.960	2.576	2.807	3.291

Table A8. Coefficients $\beta^{N}_{\mu,\sigma}(P,n)$ in equations (II.7); to be used for constructing central prediction intervals of P-content when a sample of size n is available from a normal population with unknown mean and variance.

n	Probability content, P					
	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
1	0.953	1.813	2.326	3.289	3.643	4.370
2	0.825	1.570	2.015	2.849	3.155	3.784
3	0.778	1.480	1.899	2.686	2.975	3.568
4	0.754	1.433	1.839	2.601	2.880	3.455
5	0.738	1.404	1.802	2.548	2.822	3.385
6	0.728	1.385	1.777	2.512	2.782	3.338
7	0.721	1.371	1.759	2.487	2.754	3.303
8	0.715	1.360	1.745	2.467	2.732	3.277
9	0.710	1.351	1.734	2.452	2.715	3.257
10	0.707	1.345	1.725	2.440	2.702	3.241
11	0.704	1.339	1.718	2.429	2.691	3.227
12	0.702	1.334	1.712	2.421	2.681	3.216
13	0.699	1.330	1.707	2.414	2.673	3.207
14	0.698	1.327	1.703	2.408	2.666	3.198
15	0.696	1.324	1.699	2.402	2.660	3.191
16	0.695	1.321	1.696	2.398	2.655	3.185
17	0.694	1.319	1.693	2.393	2.651	3.180
18	0.692	1.317	1.690	2.390	2.647	3.175
19	0.692	1.315	1.688	2.386	2.643	3.170
20	0.691	1.314	1.686	2.383	2.640	3.166
30	0.685	1.303	1.672	2.364	2.619	3.141
40	0.682	1.298	1.665	2.355	2.608	3.128
60	0.680	1.293	1.659	2.345	2.597	3.116
120	0.677	1.287	1.652	2.336	2.587	3.103
∞	0.674	1.282	1.645	2.326	2.576	3.090

Table A9. Coefficients $\bar{\beta}_{\mu}^{N}(P,n)$ in equations (II.15); to be used for constructing one-sided prediction intervals of P-content when a sample of size n is available from a normal population with unknown mean and known variance.

n	Probability content, P					
	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
1	1.000	3.078	6.314	31.821	63.657	318.310
2	0.816	1.886	2.920	6.965	9.925	22.326
3	0.765	1.638	2.353	4.541	5.841	10.213
4	0.741	1.533	2.132	3.747	4.604	7.173
5	0.727	1.476	2.015	3.365	4.032	5.893
6	0.718	1.440	1.943	3.143	3.707	5.208
7	0.711	1.415	1.895	2.998	3.499	4.785
8	0.706	1.397	1.860	2.896	3.355	4.501
9	0.703	1.383	1.833	2.821	3.250	4.297
10	0.700	1.372	1.812	2.764	3.169	4.144
11	0.697	1.363	1.796	2.718	3.106	4.025
12	0.695	1.356	1.782	2.681	3.055	3.930
13	0.694	1.350	1.771	2.650	3.012	3.852
14	0.692	1.345	1.761	2.624	2.977	3.787
15	0.691	1.341	1.753	2.602	2.947	3.733
16	0.690	1.337	1.746	2.583	2.921	3.686
17	0.689	1.333	1.740	2.567	2.898	3.646
18	0.688	1.330	1.734	2.552	2.878	3.610
19	0.688	1.328	1.729	2.539	2.861	3.579
20	0.687	1.325	1.725	2.528	2.845	3.552
30	0.683	1.310	1.697	2.457	2.750	3.385
40	0.681	1.303	1.684	2.423	2.704	3.307
60	0.679	1.296	1.671	2.390	2.660	3.232
120	0.677	1.289	1.658	2.358	2.617	3.160
∞	0.674	1.282	1.645	2.326	2.576	3.090

Table A10. Coefficients $\overline{\beta}^{N}_{\sigma}(P,n)$ in equations (II.18); to be used for constructing one-sided prediction intervals of P-content when a sample of size n is available from a normal population with known mean and unknown variance.

n	Probability content, P					
	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
2	1.225	3.770	7.733	38.973	77.964	389.848
3	0.942	2.178	3.372	8.042	11.460	25.780
4	0.855	1.831	2.631	5.077	6.530	11.418
5	0.812	1.679	2.335	4.105	5.043	7.858
6	0.785	1.594	2.176	3.635	4.355	6.365
7	0.768	1.539	2.077	3.360	3.963	5.568
8	0.754	1.501	2.010	3.180	3.711	5.075
9	0.744	1.473	1.961	3.053	3.536	4.744
10	0.737	1.451	1.922	2.959	3.409	4.507
11	0.731	1.433	1.893	2.887	3.310	4.328
12	0.725	1.419	1.869	2.829	3.233	4.189
13	0.721	1.407	1.849	2.782	3.170	4.078
14	0.718	1.397	1.833	2.743	3.118	3.987
15	0.715	1.389	1.819	2.710	3.075	3.911
16	0.712	1.382	1.807	2.682	3.038	3.848
17	0.710	1.376	1.797	2.658	3.006	3.793
18	0.708	1.370	1.788	2.637	2.977	3.746
19	0.706	1.365	1.779	2.618	2.953	3.704
20	0.705	1.361	1.772	2.602	2.932	3.667
21	0.703	1.356	1.766	2.587	2.912	3.636
31	0.694	1.331	1.724	2.496	2.794	3.439
41	0.689	1.319	1.704	2.452	2.737	3.347
61	0.685	1.307	1.685	2.410	2.682	3.258
121	0.680	1.294	1.665	2.368	2.628	3.173
∞	0.674	1.282	1.645	2.326	2.576	3.090

Table A11. Coefficients $\overline{\beta}^{N}_{\mu,\sigma}(P,n)$ in equations (II.11); to be used for constructing one-sided prediction intervals of P-content when a sample of size n is available from a normal population with unknown mean and variance.

n	$r_{\sigma}^{N'}$	$r_{\mu,\sigma}^{N'}$
3	1.7321	******
4	1.4142	1.9365
5	1.2910	1.5492
6	1.2247	1.3944
7	1.1832	1.3093
8	1.1547	1.2550
9	1.1339	1.2172
10	1.1180	1.1892
11	1.1055	1.1677
12	1.0954	1.1507
13	1.0871	1.1368
14	1.0801	1.1253
15	1.0742	1.1155
16	1.0690	1.1072
17	1.0646	1.1000
18	1.0607	1.0938
19	1.0572	1.0882
20	1.0541	1.0833
30	1.0351	1.0535
40	1.0260	1.0394
60	1.0171	1.0258
120	1.0084	1.0127
∞	1.0000	1.0000

Table A12. "First-order" penalty ratios when predicting from independent normal sequences with unknown variance and known or unknown mean; see equations (II.21c,d).

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.6667	0.8889	0.9474	0.9899	0.9950	0.9990	0.9995	0.9999
2	0.6906	0.8918	0.9480	0.9899	0.9950	0.9990	0.9995	0.9999
3	0.6981	0.8928	0.9483	0.9899	0.9950	0.9990	0.9995	0.9999
4	0.7017	0.8932	0.9484	0.9899	0.9950	0.9990	0.9995	0.9999
5	0.7039	0.8935	0.9484	0.9899	0.9950	0.9990	0.9995	0.9999
6	0.7053	0.8937	0.9485	0.9899	0.9950	0.9990	0.9995	0.9999
7	0.7063	0.8938	0.9485	0.9899	0.9950	0.9990	0.9995	0.9999
8	0.7071	0.8939	0.9485	0.9899	0.9950	0.9990	0.9995	0.9999
9	0.7077	0.8940	0.9486	0.9899	0.9950	0.9990	0.9995	0.9999
10	0.7081	0.8941	0.9486	0.9899	0.9950	0.9990	0.9995	0.9999
∞	0.7123	0.8946	0.9487	0.9899	0.9950	0.9990	0.9995	0.9999

Table A13 (Part 1). Coefficients $\beta_{\sigma}^{EX}(P,n)$ in equations (II.24); to be used for right-hand prediction when a sample of size n is available from an exponential population with unknown parameter.

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	2.000	8.000	18.000	98.000	198.00	998.01	1997.9	9996.3
2	1.000	3.325	5.944	17.000	25.284	60.246	86.441	196.98
3	0.762	2.463	4.143	9.925	13.544	26.000	33.797	60.629
4	0.657	2.113	3.459	7.649	10.042	17.494	21.749	34.998
5	0.598	1.924	3.103	6.559	8.427	13.905	16.865	25.547
6	0.560	1.807	2.885	5.927	7.510	11.974	14.297	20.849
7	0.533	1.726	2.739	5.515	6.922	10.779	12.733	18.093
8	0.514	1.668	2.634	5.226	6.514	9.971	11.688	16.298
9	0.499	1.624	2.555	5.013	6.215	9.390	10.942	15.043
10	0.487	1.589	2.493	4.849	5.986	8.953	10.385	14.118
15	0.452	1.489	2.316	4.390	5.355	7.773	8.898	11.717
20	0.435	1.440	2.232	4.178	5.066	7.251	8.247	10.698
25	0.425	1.412	2.183	4.057	4.902	6.956	7.883	10.136
30	0.419	1.393	2.150	3.977	4.795	6.768	7.650	9.780
40	0.411	1.370	2.111	3.881	4.665	6.540	7.371	9.357
50	0.406	1.356	2.087	3.824	4.589	6.408	7.209	9.113
∞	0.386	1.303	1.996	3.605	4.298	5.908	6.601	8.210

Table A13 (Part 2). Coefficients $\beta_{\sigma}^{EX}(P,n)$ in equations (II.26); to be used for left-hand prediction when a sample of size n is available from an exponential population with unknown parameter.

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.863	0.948	0.975	0.995	0.998	1.000	1.000	1.000
2	0.930	0.974	0.987	0.997	0.999	1.000	1.000	1.000
3	0.953	0.983	0.991	0.998	0.999	1.000	1.000	1.000
4	0.964	0.987	0.994	0.999	0.999	1.000	1.000	1.000
5	0.972	0.990	0.995	0.999	1.000	1.000	1.000	1.000
6	0.976	0.991	0.996	0.999	1.000	1.000	1.000	1.000
7	0.980	0.993	0.996	0.999	1.000	1.000	1.000	1.000
8	0.982	0.993	0.997	0.999	1.000	1.000	1.000	1.000
9	0.984	0.994	0.997	0.999	1.000	1.000	1.000	1.000
10	0.986	0.995	0.997	1.000	1.000	1.000	1.000	1.000
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A14 (Part 1). Penalty ratios $r_{\sigma}^{EX}(P,n)$ (see equation (II.25b)) for right-hand prediction intervals of content P. Sampling is from an exponential population with unknown parameter.

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	5.177	6.142	9.019	27.183	46.065	168.93	302.67	1217.6
2	2.589	2.552	2.978	4.715	5.882	10.198	13.095	23.993
3	1.973	1.891	2.076	2.753	3.151	4.401	5.120	7.385
4	1.700	1.622	1.733	2.122	2.336	2.961	3.295	4.263
5	1.547	1.477	1.555	1.819	1.961	2.354	2.555	3.112
6	1.448	1.387	1.446	1.644	1.747	2.027	2.166	2.539
7	1.380	1.325	1.372	1.530	1.610	1.825	1.929	2.204
8	1.330	1.281	1.320	1.450	1.515	1.688	1.771	1.985
9	1.291	1.247	1.280	1.390	1.446	1.589	1.658	1.832
10	1.261	1.220	1.249	1.345	1.393	1.515	1.573	1.720
15	1.171	1.143	1.160	1.218	1.246	1.316	1.348	1.427
20	1.127	1.106	1.118	1.159	1.179	1.227	1.249	1.303
25	1.101	1.084	1.094	1.125	1.140	1.178	1.194	1.235
30	1.084	1.070	1.078	1.103	1.116	1.146	1.159	1.191
40	1.063	1.052	1.058	1.076	1.085	1.107	1.117	1.140
50	1.050	1.041	1.046	1.061	1.068	1.085	1.092	1.110
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A14 (Part 2). Penalty ratios $r_{\sigma}^{EX}(P,n)$ (see equation (II.27)) for left-hand prediction intervals of content P. Sampling is from an exponential population with unknown parameter.

	Probability content, P				
n	0.900	0.950	0.975	0.990	0.995
10	2.64	3.59	4.51	6.00	
20	2.41	3.24	4.04	5.12	6.18
30	2.33	3.06	3.89	4.90	5.86
	2.25	2.97	3.68	4.60	5.30

Table A15. Coefficients $\beta^{El}_{a,b}(P,n)$ in equation (II.28). (From Antle and Rademaker, 1972).

	a/b	V_y \ P	0.900	0.950	0.975	0.990	0.995
n=10	1	0.813	1.270	1.319	1.349	1.438	
	2	0.497	1.207	1.255	1.287	1.372	
	3	0.358	1.167	1.212	1.244	1.323	
	4	0.280	1.141	1.182	1.213	1.285	
	5	0.230	1.121	1.159	1.188	1.256	
	6	0.195	1.106	1.141	1.169	1.231	
	7	0.169	1.095	1.127	1.153	1.212	
	8	0.149	1.086	1.116	1.140	1.195	
	9	0.134	1.078	1.106	1.129	1.180	
	10	0.121	1.072	1.098	1.119	1.168	
n=20	1	0.813	1.114	1.137	1.149	1.168	1.219
	2	0.497	1.087	1.109	1.123	1.142	1.189
	3	0.358	1.070	1.091	1.104	1.124	1.167
	4	0.280	1.059	1.078	1.091	1.109	1.149
	5	0.230	1.051	1.068	1.080	1.098	1.134
	6	0.195	1.045	1.061	1.072	1.089	1.122
	7	0.169	1.040	1.055	1.065	1.081	1.112
	8	0.149	1.036	1.050	1.060	1.075	1.104
	9	0.134	1.033	1.045	1.055	1.069	1.097
	10	0.121	1.030	1.042	1.051	1.064	1.090
n=30	1	0.813	1.065	1.064	1.089	1.099	1.137
	2	0.497	1.049	1.052	1.073	1.084	1.118
	3	0.358	1.040	1.043	1.062	1.073	1.104
	4	0.280	1.034	1.037	1.054	1.065	1.093
	5	0.230	1.029	1.032	1.048	1.058	1.084
	6	0.195	1.025	1.029	1.043	1.053	1.076
	7	0.169	1.023	1.026	1.039	1.048	1.070
	8	0.149	1.020	1.023	1.036	1.044	1.065
	9	0.134	1.019	1.021	1.033	1.041	1.060
	10	0.121	1.017	1.020	1.030	1.038	1.056

Table A16. Penalty ratios $r^{El}_{a,b}(P,n)$ from equation (II.30). n is the size of the sample from an Extreme type I population with unknown parameters a and b.

Z=0	Probability content, P							
t_f/t_p	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
10.00	14	24	31	48	**	**	**	**
5.00	7	12	16	25	29	38	41	**
2.00	3	5	7	11	13	17	18	22
1.00	1	3	4	6	7	9	10	13
0.50	1	2	2	4	4	6	6	8
0.20	**	1	1	2	2	3	4	5
0.10	**	**	1	1	2	2	3	3
0.05	**	**	**	1	1	2	2	3
0.02	**	**	**	1	1	1	1	2
0.01	**	**	**	**	1	1	1	1

Table A17. Upper limits of left prediction intervals when sampling from a Poisson Population with unknown parameter.
Z is the number of events observed during the "past" time interval t_p; t_f is the "future" time interval during which the number of occurrences, Y, must be predicted. For other values of Z see continuation of the table.

Z=1		Probability content, P							
Z t_f/t_p	t_f/t_p	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
10.00	10.00	27	40	49	**	**	**	**	**
5.00	5.00	14	20	25	35	40	**	**	**
2.00	2.00	6	9	11	16	17	22	24	28
1.00	1.00	3	5	6	9	10	12	13	16
0.50	0.50	2	3	3	5	6	7	8	10
0.20	0.20	1	1	2	3	3	4	5	6
0.10	0.10	**	1	1	2	2	3	3	4
0.05	0.05	**	**	1	1	2	2	2	3
0.02	0.02	**	**	**	1	1	2	2	2
0.01	0.01	**	**	**	1	1	1	1	2

Z=2		Probability content, P							
Z t_f/t_p	t_f/t_p	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
20.00	10.00	40	**	**	**	**	**	**	**
10.00	5.00	20	28	33	45	49	**	**	**
4.00	2.00	8	12	14	19	21	26	28	33
2.00	1.00	4	6	8	11	12	15	16	19
1.00	0.50	2	4	4	6	7	9	10	11
0.40	0.20	1	2	2	3	4	5	5	6
0.20	0.10	**	1	1	2	3	3	4	5
0.10	0.05	**	1	1	2	2	3	3	3
0.04	0.02	**	**	1	1	1	2	2	2
0.02	0.01	**	**	**	1	1	1	2	2

Table A17 (cont.). Upper limits of left prediction intervals for Poisson sequences with unknown parameter; Z=1 and Z=2.

Z=5		Probability content, P							
Zt_f/t_p	t_f/t_p	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
50.00	10.00	**	**	**	**	**	**	**	**
25.00	5.00	38	48	**	**	**	**	**	**
10.00	2.00	15	20	23	29	32	38	40	45
5.00	1.00	8	11	12	16	18	21	22	25
2.50	0.50	4	6	7	9	10	12	13	15
1.00	0.20	2	3	3	5	5	7	7	8
0.50	0.10	1	2	2	3	4	5	5	6
0.25	0.05	1	1	1	2	3	3	4	4
0.10	0.02	**	1	1	1	2	2	2	3
0.05	0.01	**	**	1	1	1	2	2	2

Z=10		Probability content, P							
Zt_f/t_p	t_f/t_p	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
20.00	2.00	27	33	37	44	48	**	**	**
10.00	1.00	14	17	19	24	26	30	31	35
5.00	0.50	7	9	11	14	15	17	18	20
2.00	0.20	3	4	5	7	8	9	10	11
1.00	0.10	2	3	3	4	5	6	6	7
0.50	0.05	1	2	2	3	3	4	5	5
0.20	0.02	**	1	1	2	2	3	3	4
0.10	0.01	**	1	1	1	2	2	2	3

Table A17 (cont.). Upper limits of left prediction intervals for Poisson sequences with unknown parameter; Z=5 and Z=10.

Z=20		Probability content, P							
Zt_f/t_p	t_f/t_p	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
20.00	1.00	25	30	32	38	40	45	47	**
10.00	0.50	13	16	18	21	22	26	27	30
4.00	0.20	6	7	8	10	11	13	14	15
2.00	0.10	3	4	5	6	7	8	9	10
1.00	0.05	2	2	3	4	5	6	6	7
0.40	0.02	1	1	2	3	3	4	4	5
0.20	0.01	**	1	1	2	2	3	3	3

Z=49		Probability content, P							
Zt_f/t_p	t_f/t_p	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
24.50	0.50	29	33	36	41	43	48	49	**
9.80	0.20	12	15	16	19	20	23	24	26
4.90	0.10	6	8	9	11	12	14	15	16
2.45	0.05	3	5	5	7	8	9	10	11
0.98	0.02	2	2	3	4	4	5	6	7
0.49	0.01	1	1	2	3	3	4	4	5

Table A17 (cont.). Upper limits of left prediction intervals for Poisson sequences with unknown parameter; Z=20 and Z=49.

	Probability content, P							
m λ	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
0.01	0	0	0	0	1	1	1	1
0.02	0	0	0	1	1	1	1	2
0.04	0	0	0	1	1	1	2	2
0.05	0	0	0	1	1	2	2	2
0.10	0	0	1	1	1	2	2	3
0.20	0	1	1	2	2	3	3	3
0.25	0	1	1	2	2	3	3	4
0.40	1	1	2	2	3	3	4	4
0.49	1	1	2	3	3	4	4	5
0.50	1	1	2	3	3	4	4	5
0.98	2	2	3	4	4	5	6	6
1.00	2	2	3	4	4	5	6	6
2.00	3	4	5	6	6	8	8	9
2.45	3	5	5	7	7	8	9	10
2.50	3	5	5	7	7	9	9	10
4.00	5	7	8	9	10	11	12	13
4.90	6	8	9	11	11	14	15	15
5.00	6	8	9	11	12	14	15	15
9.80	12	14	15	18	19	21	22	23
10.00	12	14	15	18	19	21	22	24
20.00	23	26	28	31	32	35	36	39
24.50	28	31	33	37	38	41	42	46
25.00	29	32	33	37	39	42	43	47

Table A18. Upper limits of left prediction intervals for Poisson sequences with known parameter λ. (denominator in equation (II.33).) m λ is the expected value of the Poisson distribution.

Z=1		Probability content, P							
Zt_f/t_p	t_f/t_p	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
10.00	10.00	2.25	2.86	3.27	****	****	****	****	****
5.00	5.00	2.33	2.50	2.78	3.18	3.33	****	****	****
2.00	2.00	2.00	2.25	2.20	2.67	2.83	2.75	3.00	3.11
1.00	1.00	1.50	2.50	2.00	2.25	2.50	2.40	2.17	2.67
0.50	0.50	2.00	3.00	1.50	1.67	2.00	1.75	2.00	2.00
0.20	0.20	****	1.00	2.00	1.50	1.50	1.33	1.67	2.00
0.10	0.10	****	****	1.00	2.00	2.00	1.50	1.50	1.33
0.05	0.05	****	****	****	1.00	2.00	1.00	1.00	1.50
0.02	0.02	****	****	****	1.00	1.00	2.00	2.00	1.00
0.01	0.01	****	****	****	****	1.00	1.00	1.00	2.00

Z=2		Probability content, P							
Zt_f/t_p	t_f/t_p	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
20.00	10.00	1.74	****	****	****	****	****	****	****
10.00	5.00	1.67	2.00	2.20	2.50	2.58	****	****	****
4.00	2.00	1.60	1.71	1.75	2.11	2.10	2.36	2.33	2.54
2.00	1.00	1.33	1.50	1.60	1.83	2.00	1.88	2.00	2.11
1.00	0.50	1.00	2.00	1.33	1.50	1.75	1.80	1.67	1.83
0.40	0.20	1.00	2.00	1.00	1.50	1.33	1.67	1.25	1.50
0.20	0.10	****	1.00	1.00	1.00	1.50	1.00	1.33	1.67
0.10	0.05	****	****	1.00	2.00	2.00	1.50	1.50	1.00
0.04	0.02	****	****	****	1.00	1.00	2.00	1.00	1.00
0.02	0.01	****	****	****	1.00	1.00	1.00	2.00	1.00

Table A19. Values of the penalty ratio (II.33) for Poisson sequences with unknown parameter. A is the number of events observed during the "past" time interval t_p; prediction is for the number of events during the "future" time interval t_f. For other values of Z see continuation of the table.

Z=5		Probability content, P							
Zt_f/t_p	t_f/t_p	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
50.00	10.00	****	****	****	****	****	****	****	****
25.00	5.00	1.31	1.50	****	****	****	****	****	****
10.00	2.00	1.25	1.43	1.53	1.61	1.68	1.81	1.82	1.88
5.00	1.00	1.33	1.38	1.33	1.45	1.50	1.50	1.47	1.67
2.50	0.50	1.33	1.20	1.40	1.29	1.43	1.33	1.44	1.50
1.00	0.20	1.00	1.50	1.00	1.25	1.25	1.40	1.17	1.33
0.50	0.10	1.00	2.00	1.00	1.00	1.33	1.25	1.25	1.20
0.25	0.05	****	1.00	1.00	1.00	1.50	1.00	1.33	1.00
0.10	0.02	****	****	1.00	1.00	2.00	1.00	1.00	1.00
0.05	0.01	****	****	****	1.00	1.00	1.00	1.00	1.00

Z=10		Probability content, P							
Zt_f/t_p	t_f/t_p	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
20.00	2.00	1.17	1.27	1.32	1.42	1.50	****	****	****
10.00	1.00	1.17	1.21	1.27	1.33	1.37	1.43	1.41	1.46
5.00	0.50	1.17	1.13	1.22	1.27	1.25	1.21	1.20	1.33
2.00	0.20	1.00	1.00	1.00	1.17	1.33	1.13	1.25	1.22
1.00	0.10	1.00	1.50	1.00	1.00	1.25	1.20	1.00	1.17
0.50	0.05	1.00	2.00	1.00	1.00	1.00	1.00	1.25	1.00
0.20	0.02	****	1.00	1.00	1.00	1.00	1.00	1.00	1.33
0.10	0.01	****	****	1.00	1.00	2.00	1.00	1.00	1.00

Table A19 (cont.). Values of the penalty ratio (II.33) for Poisson sequences with unknown parameter; Z=5 and Z=10.

Z=20		Probability content, P							
Zt_f/t_p	t_f/t_p	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
20.00	1.00	1.09	1.15	1.14	1.23	1.25	1.29	1.31	****
10.00	0.50	1.08	1.14	1.20	1.17	1.16	1.24	1.23	1.25
4.00	0.20	1.20	1.00	1.00	1.11	1.10	1.18	1.17	1.15
2.00	0.10	1.00	1.00	1.00	1.00	1.17	1.00	1.13	1.11
1.00	0.05	1.00	1.00	1.00	1.00	1.25	1.20	1.00	1.17
0.40	0.02	1.00	1.00	1.00	1.50	1.00	1.33	1.00	1.25
0.20	0.01	****	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Z=49		Probability content, P							
Zt_f/t_p	t_f/t_p	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
24.50	0.50	1.04	1.06	1.09	1.11	1.13	1.17	1.17	****
9.80	0.20	1.00	1.07	1.07	1.06	1.05	1.10	1.09	1.13
4.90	0.10	1.00	1.00	1.00	1.00	1.09	1.00	1.00	1.07
2.45	0.05	1.00	1.00	1.00	1.00	1.14	1.13	1.11	1.10
0.98	0.02	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.17
0.49	0.01	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table A19 (cont.). Values of the penalty ratio (II.33) for Poisson sequences with unknown parameter; Z=20 and Z=49.

n = 1

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	2.00	8.00	18.00	98.00	198.00	998.00	1998.00	9998.00
2	1.51	4.39	7.57	21.14	31.66	73.62	105.96	243.12
3	1.35	3.58	5.65	12.88	17.40	32.62	42.12	71.84
4	1.28	3.19	4.85	9.99	13.09	23.03	25.85	42.51
5	1.23	2.96	4.40	8.53	10.74	17.34	20.79	31.02
6	1.19	2.80	4.14	7.73	9.46	14.70	17.37	24.97
7	1.18	2.71	3.96	7.13	8.71	13.08	15.22	21.17
8	1.16	2.64	3.77	6.67	8.14	11.91	13.72	18.52
9	1.15	2.57	3.66	6.35	7.67	11.05	12.63	16.99
10	1.14	2.51	3.56	6.13	7.31	10.41	11.85	15.64
15	1.10	2.34	3.26	5.38	6.29	8.60	9.62	12.22
20	1.08	2.26	3.10	4.94	5.77	7.74	8.59	10.67
25	1.06	2.20	2.99	4.68	5.50	7.21	7.99	9.76
30	1.05	2.18	2.91	4.56	5.25	6.89	7.49	9.18
40	1.04	2.13	2.82	4.35	5.02	6.40	6.99	8.36
50	1.03	2.09	2.75	4.22	4.77	6.06	6.66	7.95

Table A20. Values of the parameter $\beta_{\lambda}^{G}(n,k,P)$ in equation (II.35) for the construction of left-hand prediction intervals when sampling is from a Gamma sequence. The shape parameter of the Gamma distribution, k, is known while the parameter λ of the underlying exponential distribution is unknown. n is the available sample size; here n=1. For other values of n, see continuation of the table.

n = 2

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	1.00	3.32	5.94	17.00	25.28	60.26	86.44	197.00
2	0.94	2.56	4.02	8.47	10.96	18.94	23.49	37.97
3	0.91	2.30	3.46	6.60	8.23	12.80	15.11	21.66
4	0.90	2.16	3.17	5.76	7.03	10.39	12.02	16.47
5	0.89	2.09	3.01	5.30	6.35	9.09	10.37	13.81
6	0.89	2.03	2.90	4.96	5.93	8.30	9.37	12.25
7	0.88	1.99	2.82	4.75	5.62	7.72	8.73	11.17
8	0.88	1.96	2.74	4.57	5.40	7.36	8.25	10.47
9	0.88	1.94	2.69	4.42	5.20	7.04	7.85	9.87
10	0.87	1.91	2.65	4.30	5.04	6.76	7.54	9.39
15	0.86	1.84	2.53	3.98	4.72	6.03	6.65	8.10
20	0.86	1.81	2.44	3.78	4.38	5.61	6.17	7.41
25	0.85	1.78	2.38	3.68	4.19	5.37	5.86	6.98

Table A20 (cont.). Values of the parameter $\beta_{\hat{\lambda}}^{G}(n,k,P)$ in equation (II.35) for left-hand prediction from Gammâ sequences. Sample size: n=2.

n = 5

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.60	1.92	3.10	6.56	8.43	13.91	16.87	25.55
2	0.65	1.76	2.63	4.86	5.89	8.64	9.88	13.27
3	0.68	1.69	2.45	4.29	5.10	7.14	8.07	10.40
4	0.69	1.66	2.36	3.97	4.69	6.41	7.18	9.05
5	0.69	1.64	2.28	3.80	4.43	5.97	6.64	8.23
7	0.70	1.60	2.20	3.56	4.14	5.45	6.02	7.35
10	0.71	1.56	2.12	3.37	3.88	5.04	5.52	6.67

Table A20 (cont.). Values of the parameter $\beta_{\lambda}^{G}(n,k,P)$ in equation (II.35) for left-hand prediction from Gamma sequences. Sample size: n=5.

n = 10

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.49	1.59	2.49	4.85	5.99	8.95	10.38	14.12
2	0.57	1.54	2.27	4.00	4.76	6.64	7.50	9.56
3	0.60	1.51	2.17	3.68	4.32	5.83	6.51	8.12
4	0.62	1.49	2.11	3.48	4.08	5.40	5.99	7.32
5	0.63	1.48	2.07	3.35	3.89	5.13	5.66	6.90

n = 20

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.44	1.44	2.23	4.18	5.07	7.25	8.25	10.70
2	0.51	1.44	2.10	3.65	4.25	5.81	6.50	8.09

n = 50

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.41	1.36	2.09	3.82	4.59	6.41	7.21	9.11

Table A20 (cont.). Values of the parameter $\beta_{\hat{\lambda}}^{G}(n,k,P)$ in equation (II.35) for left-hand prediction from Gamma sequences. Sample sizes: n=10, n=20, n=50.

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.386	1.303	1.996	3.605	4.299	5.908	6.601	8.210
2	0.490	1.336	1.940	3.280	3.840	5.115	5.656	6.899
3	0.531	1.341	1.903	3.121	3.622	4.751	5.226	6.309
4	0.555	1.340	1.877	3.023	3.489	4.531	4.967	5.957
5	0.570	1.339	1.858	2.954	3.396	4.380	4.790	5.717
6	0.581	1.337	1.842	2.902	3.327	4.268	4.658	5.539
7	0.589	1.335	1.830	2.861	3.273	4.181	4.556	5.401
8	0.594	1.333	1.820	2.828	3.229	4.111	4.474	5.290
9	0.600	1.332	1.811	2.801	3.193	4.052	4.406	5.198
10	0.605	1.330	1.804	2.777	3.162	4.003	4.348	5.121
15	0.620	1.324	1.778	2.697	3.056	3.835	4.152	4.858
20	0.628	1.320	1.762	2.649	2.993	3.735	4.036	4.703
25	0.633	1.317	1.750	2.615	2.949	3.667	3.956	4.597
30	0.638	1.314	1.742	2.591	2.917	3.616	3.898	4.519
40	0.648	1.311	1.730	2.556	2.871	3.545	3.815	4.410
50	0.656	1.308	1.721	2.532	2.840	3.497	3.759	4.336
∞	0.674	1.282	1.645	2.326	2.576	3.090	3.291	3.719

Table A21. Values of the parameter $\beta^{G}_{PI}(k,P)$ in equations (II.36) and (II.37) for the construction of left-hand prediction intervals when sampling is from a Gamma sequence with known parameters.

n = 1

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	5.18	6.14	9.02	27.18	46.06	168.92	302.68	1217.71
2	3.08	3.29	3.90	6.45	8.25	14.39	18.73	35.24
3	2.54	2.67	2.97	4.13	4.80	6.87	8.06	11.39
4	2.31	2.38	2.58	3.31	3.75	5.08	5.20	7.14
5	2.16	2.21	2.37	2.89	3.16	3.96	4.34	5.43
6	2.05	2.09	2.25	2.66	2.84	3.44	3.73	4.51
7	2.00	2.03	2.16	2.49	2.66	3.13	3.34	3.92
8	1.95	1.98	2.07	2.36	2.52	2.90	3.07	3.50
9	1.92	1.93	2.02	2.27	2.40	2.73	2.87	3.27
10	1.88	1.89	1.97	2.21	2.31	2.60	2.73	3.05
15	1.78	1.77	1.83	1.99	2.06	2.24	2.32	2.52
20	1.72	1.71	1.76	1.86	1.93	2.07	2.13	2.27
25	1.67	1.67	1.71	1.79	1.87	1.97	2.02	2.12
30	1.65	1.66	1.67	1.76	1.80	1.91	1.92	2.03
40	1.60	1.63	1.63	1.70	1.75	1.81	1.83	1.90
50	1.57	1.60	1.60	1.67	1.68	1.73	1.77	1.83

Table A22. Values of the penalty ratio $r_{\lambda}^{G}(n,k,P)$ in equation (II.39) for left-hand prediction from Gamma sequences. The shape parameter k is known, while the parameter λ of the underlying exponential distribution is not. n is the available sample size; here n = 1. For other values of n, see continuation of the table.

n = 2

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	2.59	2.55	2.98	4.72	5.88	10.20	13.09	23.99
2	1.92	1.92	2.07	2.58	2.85	3.70	4.15	5.50
3	1.71	1.72	1.82	2.11	2.27	2.69	2.89	3.43
4	1.62	1.61	1.69	1.91	2.02	2.29	2.42	2.76
5	1.56	1.56	1.62	1.79	1.87	2.08	2.17	2.42
6	1.53	1.52	1.57	1.71	1.78	1.94	2.01	2.21
7	1.49	1.49	1.54	1.66	1.72	1.85	1.92	2.07
8	1.48	1.47	1.51	1.62	1.67	1.79	1.84	1.98
9	1.47	1.46	1.48	1.58	1.63	1.74	1.78	1.90
10	1.44	1.44	1.47	1.55	1.59	1.69	1.73	1.83
15	1.39	1.39	1.42	1.48	1.54	1.57	1.60	1.67
20	1.37	1.37	1.38	1.43	1.46	1.50	1.53	1.58
25	1.34	1.35	1.36	1.41	1.42	1.46	1.48	1.52

Table A22 (cont.). Values of the penalty ratio $r_{\lambda}^{G}(n,k,P)$ in equation (II.39) for left-prediction from Gamma sequences. Sample size: n = 2.

n = 5

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	1.55	1.47	1.55	1.82	1.96	2.35	2.56	3.11
2	1.33	1.32	1.36	1.48	1.53	1.69	1.75	1.92
3	1.28	1.26	1.29	1.37	1.41	1.50	1.54	1.65
4	1.24	1.24	1.26	1.31	1.34	1.41	1.45	1.52
5	1.21	1.22	1.23	1.29	1.30	1.36	1.39	1.44
7	1.19	1.20	1.20	1.24	1.26	1.30	1.32	1.36
10	1.17	1.17	1.18	1.21	1.23	1.26	1.27	1.30

Table A22 (cont.). Values of the penalty ratio $r_{\lambda}^{G}(n,k,P)$ in equation (II.39) for left-prediction from Gamma sequences. Sample size: n = 5.

n = 10

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	1.27	1.22	1.25	1.35	1.39	1.51	1.57	1.72
2	1.16	1.15	1.17	1.22	1.24	1.30	1.33	1.39
3	1.13	1.13	1.14	1.18	1.19	1.23	1.25	1.29
4	1.12	1.11	1.12	1.15	1.17	1.19	1.21	1.23
5	1.11	1.11	1.11	1.13	1.15	1.17	1.18	1.21

n = 20

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	1.14	1.11	1.12	1.16	1.18	1.23	1.25	1.30
2	1.04	1.08	1.08	1.11	1.11	1.14	1.15	1.17

n = 50

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	1.06	1.04	1.05	1.06	1.07	1.08	1.09	1.11

Table A22 (cont.). Values of the penalty ratio $r_{\lambda}^{G}(n,k,P)$ in equation (II.39) for left-prediction from Gamma sequences. Sample sizes: n = 10, n = 20, n = 50.

n = 1

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.67	0.89	0.95	0.99	0.995	0.999	0.9995	0.9999
2	0.73	1.07	1.19	1.33	1.35	1.39	1.40	1.41
3	0.76	1.17	1.33	1.53	1.58	1.64	1.66	1.69
4	0.78	1.23	1.42	1.67	1.74	1.84	1.86	1.91
5	0.79	1.27	1.48	1.77	1.85	1.98	2.02	2.09
6	0.80	1.31	1.54	1.86	1.95	2.10	2.15	2.23
7	0.81	1.34	1.59	1.93	2.03	2.20	2.25	2.35
8	0.82	1.36	1.62	1.99	2.10	2.29	2.34	2.45
9	0.83	1.38	1.65	2.04	2.16	2.36	2.42	2.57
10	0.83	1.40	1.67	2.09	2.21	2.43	2.50	2.63
15	0.85	1.46	1.77	2.25	2.40	2.67	2.76	2.94
20	0.87	1.49	1.82	2.35	2.52	2.83	2.94	3.15
25	0.88	1.53	1.87	2.42	2.62	2.95	3.07	3.31
30	0.89	1.56	1.91	2.49	2.69	3.05	3.17	3.43
40	0.89	1.58	1.95	2.58	2.79	3.18	3.32	3.60
50	0.90	1.59	1.98	2.64	2.86	3.26	3.42	3.74

Table A23. Values of the parameter $\beta_{\lambda}^{G}(n,k,P)$ in equation (II.40) for the construction of right-hand prediction intervals when sampling is from a Gamma sequence. The shape parameter, k, is known, while the parameter λ of the underlying exponential distribution is not. n is the available sample size, here n = 1. For other values of n, see continuation of the table.

n = 2

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.69	0.89	0.95	0.99	0.995	0.999	0.9995	0.9999
2	0.73	1.06	1.18	1.32	1.35	1.39	1.39	1.41
3	0.75	1.14	1.30	1.51	1.56	1.64	1.66	1.69
4	0.76	1.19	1.38	1.64	1.71	1.82	1.85	1.89
5	0.77	1.22	1.43	1.73	1.82	1.94	1.99	2.07
7	0.78	1.26	1.50	1.86	1.97	2.15	2.21	2.32
10	0.79	1.32	1.58	1.99	2.12	2.35	2.43	2.57
15	0.80	1.36	1.65	2.10	2.27	2.55	2.65	2.83
20	0.80	1.39	1.70	2.19	2.37	2.69	2.79	3.02
25	0.80	1.41	1.73	2.27	2.45	2.78	2.91	3.15

Table A23 (cont.). Values of the parameter $\beta_{\lambda}^{G}(n,k,P)$ in equation (II.40) for right-prediction from Gamma sequences. Sample size: n = 2.

n = 5

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.70	0.89	0.95	0.99	0.995	0.999	0.9995	0.9999
2	0.73	1.05	1.17	1.32	1.34	1.38	1.39	1.41
3	0.74	1.12	1.28	1.49	1.55	1.63	1.65	1.69
4	0.75	1.15	1.34	1.62	1.69	1.80	1.84	1.90
5	0.75	1.18	1.39	1.70	1.79	1.93	1.98	2.05
7	0.75	1.22	1.44	1.80	1.91	2.11	2.18	2.31
10	0.75	1.25	1.50	1.93	2.06	2.30	2.38	2.51

Table A23 (cont.). Values of the parameter $\beta_{\lambda}^{G}(n,k,P)$ in equation (II.40) for right-prediction from Gamma sequences. Sample size: n = 5.

n = 10

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.71	0.89	0.95	0.99	0.995	0.999	0.9995	0.9999
2	0.74	1.04	1.17	1.31	1.34	1.38	1.39	1.40
3	0.74	1.10	1.27	1.48	1.54	1.63	1.65	1.68
4	0.75	1.15	1.33	1.60	1.67	1.79	1.82	1.88
5	0.75	1.18	1.38	1.68	1.77	1.90	1.95	2.04

n = 20

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.71	0.90	0.95	0.99	0.995	0.999	0.9995	0.9999
2	0.74	1.04	1.16	1.31	1.34	1.38	1.39	1.40

n = 50

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.71	0.90	0.95	0.99	0.995	0.999	0.9995	0.9999

Table A23 (cont.). Values of the parameter $\beta_{\lambda}^{G}(n,k,P)$ in equation (II.40) for right-prediction from Gamma sequences. Sample sizes: n = 10, n = 20, n = 50.

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.712	0.895	0.949	0.990	0.995	0.9990	0.9995	0.9999
2	0.734	1.038	1.163	1.309	1.341	1.382	1.392	1.404
3	0.735	1.096	1.260	1.480	1.537	1.622	1.646	1.682
4	0.732	1.128	1.317	1.588	1.664	1.786	1.822	1.884
5	0.730	1.148	1.355	1.664	1.754	1.905	1.953	2.037
6	0.727	1.163	1.383	1.721	1.822	1.998	2.055	2.158
7	0.725	1.174	1.404	1.765	1.876	2.071	2.136	2.256
8	0.723	1.182	1.421	1.801	1.920	2.132	2.203	2.338
9	0.721	1.189	1.435	1.831	1.956	2.182	2.260	2.408
10	0.719	1.195	1.447	1.856	1.987	2.226	2.309	2.468
15	0.713	1.214	1.486	1.942	2.093	2.377	2.478	2.678
20	0.709	1.224	1.508	1.994	2.157	2.469	2.582	2.808
25	0.706	1.231	1.524	2.029	2.201	2.533	2.654	2.899
30	0.704	1.236	1.535	2.055	2.233	2.580	2.708	2.967
40	0.700	1.243	1.550	2.092	2.279	2.647	2.783	3.064
50	0.697	1.248	1.561	2.117	2.310	2.693	2.836	3.131
∞	0.674	1.282	1.645	2.326	2.576	3.090	3.291	3.719

Table A24. Values of the parameter $\beta^{G}_{PI}(k,P)$ for the construction of right-hand prediction intervals when sampling is from a Gamma sequence with known parameters.

n = 1

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.94	0.99	1.00	1.00	1.00	1.00	1.00	1.00
2	0.99	1.03	1.02	1.02	1.01	1.01	1.01	1.00
3	1.03	1.07	1.06	1.03	1.03	1.01	1.01	1.00
4	1.06	1.09	1.08	1.05	1.05	1.03	1.02	1.01
5	1.08	1.11	1.09	1.06	1.05	1.04	1.03	1.03
6	1.10	1.13	1.11	1.08	1.07	1.05	1.05	1.03
7	1.12	1.14	1.13	1.09	1.08	1.06	1.05	1.04
8	1.13	1.15	1.14	1.10	1.09	1.07	1.06	1.05
9	1.15	1.16	1.15	1.11	1.10	1.08	1.07	1.07
10	1.15	1.17	1.15	1.13	1.11	1.09	1.08	1.07
15	1.19	1.20	1.19	1.16	1.15	1.12	1.11	1.10
20	1.23	1.22	1.21	1.18	1.17	1.15	1.14	1.12
25	1.25	1.24	1.23	1.19	1.19	1.16	1.16	1.14
30	1.26	1.26	1.24	1.21	1.20	1.18	1.17	1.16
40	1.27	1.27	1.26	1.23	1.22	1.20	1.19	1.18
50	1.29	1.27	1.27	1.25	1.24	1.21	1.21	1.19

Table A25. Values of the penalty ratio $r_{\lambda}^{G}(n,k,P)$ for <u>right-hand</u> prediction from Gamma sequences. The shape parameter k is known, while the parameter λ of the underlying exponential distribution is not. n is the available sample size, here n = 1. For other values of n, see continuation of the table.

n = 2

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.97	0.99	1.00	1.00	1.00	1.00	1.00	1.00
2	0.99	1.02	1.01	1.01	1.01	1.01	1.00	1.00
3	1.02	1.04	1.03	1.02	1.01	1.01	1.01	1.00
4	1.04	1.06	1.05	1.03	1.03	1.02	1.02	1.00
5	1.06	1.06	1.06	1.04	1.04	1.02	1.02	1.02
7	1.08	1.07	1.07	1.05	1.05	1.04	1.03	1.03
10	1.10	1.10	1.09	1.07	1.07	1.06	1.05	1.04
15	1.12	1.12	1.11	1.08	1.08	1.07	1.07	1.06
20	1.13	1.14	1.13	1.10	1.10	1.09	1.08	1.08
25	1.13	1.15	1.14	1.12	1.11	1.10	1.10	1.09

Table A25 (cont.). Values of the penalty ratio $r_{\lambda}^{G}(n,k,P)$ for right-prediction from Gamma sequences. Sample size: n = 2.

n = 5

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	0.98	0.99	1.00	1.00	1.00	1.00	1.00	1.00
2	0.99	1.01	1.01	1.01	1.00	1.00	1.00	1.00
3	1.01	1.02	1.02	1.01	1.01	1.00	1.00	1.00
4	1.02	1.02	1.02	1.02	1.02	1.01	1.01	1.01
5	1.03	1.03	1.03	1.02	1.02	1.01	1.01	1.01
7	1.03	1.04	1.03	1.02	1.02	1.02	1.02	1.02
10	1.04	1.05	1.04	1.04	1.04	1.03	1.03	1.02

Table A25 (cont.). Values of the penalty ratio $r_{\lambda}^{G}(n,k,P)$ for right-prediction from Gamma sequences. Sample size: n = 5.

n = 10

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2	1.00	1.00	1.01	1.00	1.00	1.00	1.00	1.00
3	1.01	1.01	1.01	1.00	1.00	1.00	1.00	1.00
4	1.02	1.02	1.01	1.01	1.00	1.00	1.00	1.00
5	1.03	1.03	1.02	1.01	1.01	1.00	1.00	1.00

n = 20

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

n = 50

k \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995	0.9999
1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table A25 (cont.). Values of the penalty ratio $r_{\lambda}^{G}(n,k,P)$ for right-prediction from Gamma sequences. Sample sizes: n = 10, n = 20, n = 50.

LEFT PREDICTION INTERVAL V_Y= 0.05

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	1.177	3.921	8.950	114.365	916.461	*******	*******
3	0.893	2.164	3.468	9.390	14.681	49.972	98.758
4	0.807	1.799	2.649	5.470	7.317	14.614	20.117
5	0.764	1.642	2.331	4.310	5.430	9.131	11.438
6	0.737	1.554	2.162	3.764	4.601	7.099	8.514
7	0.721	1.499	2.057	3.458	4.143	6.085	7.114
8	0.707	1.458	1.986	3.254	3.854	5.470	6.294
9	0.698	1.430	1.934	3.114	3.654	5.108	5.767
10	0.690	1.407	1.895	3.008	3.508	4.783	5.394
11	0.684	1.387	1.864	2.928	3.398	4.562	5.120
12	0.679	1.374	1.839	2.865	3.312	4.407	4.915
13	0.674	1.363	1.819	2.815	3.242	4.277	4.752
14	0.671	1.352	1.801	2.771	3.184	4.170	4.618
15	0.668	1.344	1.787	2.736	3.135	4.084	4.509
16	0.665	1.336	1.774	2.706	3.094	4.011	4.418
17	0.663	1.330	1.764	2.679	3.059	3.946	4.338
18	0.661	1.323	1.753	2.655	3.029	3.890	4.269
19	0.659	1.319	1.745	2.635	3.001	3.844	4.213
20	0.658	1.315	1.738	2.618	2.978	3.803	4.162
21	0.656	1.309	1.730	2.601	2.956	3.764	4.115
31	0.649	1.286	1.688	2.507	2.827	3.548	3.849
41	0.643	1.272	1.667	2.456	2.765	3.440	3.719
61	0.639	1.261	1.646	2.413	2.705	3.343	3.602
121	0.634	1.247	1.627	2.365	2.646	3.245	3.484
∞	0.628	1.235	1.606	2.321	2.590	3.153	3.377

Table A26. Values of the coefficient $\beta^{LN}_{\mu,\sigma}(V_Y,n,P)$ in equation (II.45) for left-hand prediction in lognormal sequences with unknown mean and variance. V_Y is the estimated coefficient of variation of the lognormal population and n is the sample size. For other values of V_Y, see continuation of the table.

LEFT PREDICTION INTERVAL V_Y = 0.10

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	1.131	4.088	10.478	432.784	*******	*******	*******
3	0.847	2.151	3.574	11.097	19.292	109.458	335.602
4	0.761	1.769	2.671	5.923	8.265	19.179	29.157
5	0.718	1.606	2.329	4.537	5.873	10.732	14.105
6	0.692	1.515	2.149	3.909	4.879	7.980	9.875
7	0.675	1.459	2.038	3.562	4.343	6.684	8.000
8	0.662	1.417	1.964	3.335	4.011	5.923	6.946
9	0.653	1.388	1.909	3.179	3.784	5.483	6.288
10	0.645	1.364	1.869	3.062	3.619	5.094	5.830
11	0.639	1.345	1.836	2.974	3.496	4.833	5.497
12	0.635	1.331	1.810	2.906	3.399	4.651	5.252
13	0.630	1.320	1.789	2.851	3.321	4.499	5.057
14	0.627	1.308	1.771	2.803	3.257	4.375	4.898
15	0.624	1.300	1.756	2.765	3.203	4.275	4.771
16	0.621	1.293	1.743	2.732	3.157	4.191	4.663
17	0.619	1.286	1.732	2.703	3.118	4.117	4.570
18	0.617	1.279	1.721	2.677	3.085	4.053	4.490
19	0.615	1.275	1.712	2.656	3.055	4.000	4.425
20	0.614	1.271	1.705	2.637	3.030	3.953	4.366
21	0.612	1.265	1.697	2.618	3.005	3.909	4.311
31	0.605	1.241	1.653	2.517	2.864	3.664	4.006
41	0.599	1.228	1.632	2.463	2.796	3.542	3.857
61	0.595	1.216	1.610	2.417	2.731	3.434	3.725
121	0.590	1.202	1.590	2.365	2.667	3.325	3.591
∞	0.584	1.190	1.569	2.318	2.607	3.222	3.472

Table A26 (cont.). Values of the coefficient $\beta^{LN}_{\mu,\sigma}(V_Y,n,P)$ in equation (II.45) for left-prediction in lognormal sequences with unknown mean and variance. Estimated coefficient of variation: $V_Y = 0.10$.

LEFT PREDICTION INTERVAL V_Y = 0.20

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	1.045	4.473	14.787	*******	*******	*******	*******
3	0.760	2.131	3.815	15.996	35.502	673.651	*******
4	0.676	1.712	2.723	7.028	10.766	35.188	67.561
5	0.634	1.538	2.330	5.068	6.955	15.278	22.401
6	0.609	1.441	2.128	4.241	5.534	10.282	13.643
7	0.593	1.383	2.006	3.801	4.808	8.185	10.317
8	0.579	1.339	1.924	3.518	4.372	7.028	8.596
9	0.571	1.309	1.864	3.328	4.081	6.385	7.576
10	0.563	1.284	1.820	3.187	3.872	5.833	6.891
11	0.558	1.264	1.784	3.081	3.718	5.471	6.406
12	0.553	1.250	1.756	3.000	3.598	5.222	6.056
13	0.548	1.238	1.734	2.935	3.501	5.017	5.782
14	0.546	1.226	1.714	2.878	3.423	4.851	5.561
15	0.543	1.218	1.698	2.834	3.357	4.718	5.385
16	0.540	1.210	1.684	2.795	3.302	4.607	5.238
17	0.538	1.204	1.672	2.760	3.254	4.511	5.112
18	0.536	1.196	1.660	2.730	3.215	4.427	5.004
19	0.534	1.192	1.651	2.706	3.178	4.359	4.917
20	0.533	1.188	1.643	2.684	3.148	4.298	4.839
21	0.531	1.181	1.635	2.662	3.119	4.241	4.766
31	0.524	1.157	1.588	2.545	2.950	3.928	4.365
41	0.518	1.143	1.565	2.483	2.870	3.776	4.175
61	0.515	1.132	1.542	2.430	2.794	3.641	4.006
121	0.510	1.117	1.521	2.372	2.719	3.506	3.837
∞	0.504	1.104	1.498	2.319	2.649	3.381	3.688

Table A26 (cont.). Values of the coefficient $\beta^{LN}_{\mu,\sigma}(V_Y,n,P)$ in equation (II.45) for left-prediction in lognormal sequences with unknown mean and variance. Estimated coefficient of variation: V_Y = 0.20.

LEFT PREDICTION INTERVAL V_Y= 0.40

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	0.895	5.439	31.623	*******	*******	*******	*******
3	0.613	2.104	4.403	35.896	138.785	*******	*******
4	0.532	1.612	2.855	10.193	19.215	136.635	443.110
5	0.494	1.417	2.350	6.444	10.041	33.328	62.669
6	0.470	1.311	2.101	5.065	7.273	17.911	27.811
7	0.455	1.248	1.954	4.381	5.997	12.725	18.003
8	0.442	1.201	1.857	3.961	5.275	10.193	13.680
9	0.434	1.169	1.787	3.686	4.812	8.889	11.363
10	0.427	1.143	1.736	3.486	4.490	7.826	9.909
11	0.422	1.121	1.695	3.338	4.256	7.158	8.930
12	0.418	1.106	1.662	3.226	4.078	6.712	8.248
13	0.414	1.094	1.637	3.137	3.936	6.355	7.730
14	0.411	1.081	1.614	3.061	3.822	6.070	7.322
15	0.408	1.072	1.596	3.001	3.727	5.845	7.004
16	0.406	1.065	1.580	2.949	3.648	5.661	6.742
17	0.404	1.058	1.567	2.904	3.581	5.501	6.520
18	0.402	1.050	1.554	2.864	3.525	5.364	6.333
19	0.401	1.046	1.543	2.831	3.474	5.254	6.183
20	0.400	1.041	1.534	2.803	3.431	5.156	6.048
21	0.398	1.035	1.525	2.774	3.390	5.065	5.925
31	0.391	1.009	1.473	2.622	3.158	4.576	5.264
41	0.386	0.995	1.447	2.542	3.050	4.343	4.960
61	0.383	0.983	1.422	2.475	2.948	4.142	4.695
121	0.378	0.968	1.398	2.401	2.849	3.943	4.437
∞	0.373	0.955	1.374	2.335	2.756	3.761	4.211

Table A26 (cont.). Values of the coefficient $\beta^{LN}_{\mu,\sigma}(V_Y,n,P)$ in equation (II.45) for left-prediction in lognormal sequences with unknown mean and variance. Estimated coefficient of variation: V_Y = 0.40.

LEFT PREDICTION INTERVAL V_Y = 0.05

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	1.875	3.175	5.572	49.272	353.883	*******	*******
3	1.423	1.752	2.159	4.045	5.669	15.851	29.246
4	1.285	1.457	1.649	2.357	2.825	4.635	5.957
5	1.217	1.330	1.451	1.857	2.097	2.896	3.387
6	1.174	1.258	1.346	1.622	1.777	2.252	2.521
7	1.148	1.214	1.281	1.490	1.600	1.930	2.107
8	1.126	1.181	1.237	1.402	1.488	1.735	1.864
9	1.111	1.158	1.204	1.342	1.411	1.620	1.708
10	1.099	1.139	1.180	1.296	1.355	1.517	1.597
11	1.089	1.124	1.160	1.261	1.312	1.447	1.516
12	1.082	1.113	1.145	1.234	1.279	1.398	1.456
13	1.074	1.104	1.132	1.213	1.252	1.357	1.407
14	1.069	1.095	1.121	1.194	1.229	1.323	1.368
15	1.064	1.088	1.112	1.179	1.211	1.295	1.335
16	1.060	1.082	1.105	1.166	1.195	1.272	1.308
17	1.056	1.077	1.098	1.154	1.181	1.252	1.285
18	1.053	1.072	1.092	1.144	1.170	1.234	1.264
19	1.050	1.068	1.086	1.135	1.159	1.219	1.248
20	1.049	1.065	1.082	1.128	1.150	1.206	1.233
21	1.045	1.060	1.077	1.120	1.141	1.194	1.219
31	1.033	1.041	1.051	1.080	1.092	1.125	1.140
41	1.023	1.030	1.038	1.058	1.067	1.091	1.101
61	1.017	1.021	1.025	1.040	1.044	1.060	1.067
121	1.009	1.010	1.013	1.019	1.022	1.029	1.032
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A27. Values of the penalty ratio $r^{LN}_{\mu,\sigma}(V_Y,n,P)$ in equation (II.47) for left-hand prediction in lognormal sequences with unknown mean and variance. V_Y is the estimated coefficient of variation and n is the sample size. For other values of V_Y, see continuation of the table.

LEFT PREDICTION INTERVAL V_Y= 0.10

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	1.936	3.437	6.679	186.673	*******	*******	*******
3	1.449	1.808	2.278	4.787	7.401	33.969	96.667
4	1.302	1.487	1.703	2.555	3.171	5.952	8.398
5	1.230	1.350	1.484	1.957	2.253	3.330	4.063
6	1.184	1.273	1.370	1.686	1.872	2.476	2.844
7	1.156	1.226	1.299	1.537	1.666	2.074	2.304
8	1.133	1.191	1.252	1.438	1.539	1.838	2.001
9	1.118	1.167	1.217	1.371	1.452	1.701	1.811
10	1.104	1.147	1.191	1.321	1.388	1.581	1.679
11	1.094	1.130	1.170	1.283	1.341	1.500	1.583
12	1.086	1.119	1.154	1.253	1.304	1.443	1.513
13	1.078	1.109	1.140	1.230	1.274	1.396	1.457
14	1.073	1.100	1.129	1.209	1.249	1.358	1.411
15	1.068	1.093	1.119	1.193	1.229	1.327	1.374
16	1.063	1.087	1.111	1.178	1.211	1.301	1.343
17	1.060	1.081	1.104	1.166	1.196	1.278	1.316
18	1.056	1.075	1.097	1.155	1.184	1.258	1.293
19	1.053	1.072	1.091	1.146	1.172	1.241	1.275
20	1.051	1.068	1.087	1.137	1.162	1.227	1.258
21	1.048	1.063	1.082	1.129	1.153	1.213	1.242
31	1.035	1.043	1.054	1.086	1.099	1.137	1.154
41	1.025	1.032	1.040	1.062	1.073	1.099	1.111
61	1.018	1.022	1.026	1.042	1.048	1.066	1.073
121	1.010	1.010	1.014	1.020	1.023	1.032	1.034
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A27 (cont.). Values of the penalty ratio $r_{\mu\sigma}^{LN}(V_Y,n,P)$ in equation (II.47) for left-prediction in lognormal sequences with unknown mean and variance. Estimated coefficient of variation: V_Y = 0.10.

LEFT PREDICTION INTERVAL V_Y= 0.20

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	2.071	4.050	9.869	*******	*******	*******	*******
3	1.506	1.930	2.546	6.899	13.404	199.265	*******
4	1.339	1.550	1.818	3.031	4.065	10.409	18.320
5	1.257	1.393	1.555	2.186	2.626	4.519	6.074
6	1.207	1.305	1.420	1.829	2.089	3.041	3.700
7	1.175	1.252	1.339	1.639	1.815	2.421	2.797
8	1.148	1.212	1.284	1.517	1.651	2.079	2.331
9	1.132	1.186	1.244	1.436	1.541	1.889	2.054
10	1.117	1.163	1.215	1.374	1.462	1.725	1.869
11	1.105	1.144	1.191	1.329	1.404	1.618	1.737
12	1.096	1.132	1.172	1.294	1.358	1.545	1.642
13	1.087	1.121	1.157	1.266	1.322	1.484	1.568
14	1.081	1.110	1.144	1.241	1.292	1.435	1.508
15	1.076	1.103	1.133	1.222	1.267	1.396	1.460
16	1.070	1.096	1.124	1.205	1.247	1.363	1.420
17	1.066	1.090	1.116	1.191	1.229	1.334	1.386
18	1.063	1.083	1.108	1.178	1.214	1.309	1.357
19	1.059	1.079	1.102	1.167	1.200	1.289	1.333
20	1.057	1.075	1.097	1.158	1.188	1.271	1.312
21	1.054	1.070	1.091	1.148	1.177	1.254	1.292
31	1.039	1.048	1.060	1.098	1.114	1.162	1.184
41	1.028	1.035	1.045	1.071	1.084	1.117	1.132
61	1.020	1.025	1.029	1.048	1.055	1.077	1.086
121	1.011	1.011	1.015	1.023	1.027	1.037	1.041
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A27 (cont.). Values of the penalty ratio $r^{LN}_{\mu,\sigma}(V_Y,n,P)$ in equation (II.47) for left-prediction in lognormal sequences with unknown mean and variance. Estimated coefficient of variation: V_Y = 0.20.

LEFT PREDICTION INTERVAL V_Y = 0.40

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	2.397	5.694	23.024	*******	*******	*******	*******
3	1.643	2.203	3.206	15.373	50.352	*******	*******
4	1.427	1.687	2.078	4.365	6.971	36.329	105.219
5	1.323	1.483	1.711	2.760	3.643	8.861	14.881
6	1.258	1.373	1.530	2.169	2.639	4.762	6.604
7	1.218	1.306	1.423	1.876	2.176	3.383	4.275
8	1.185	1.257	1.352	1.696	1.914	2.710	3.248
9	1.164	1.224	1.301	1.579	1.746	2.363	2.698
10	1.145	1.196	1.264	1.493	1.629	2.081	2.353
11	1.131	1.174	1.234	1.430	1.544	1.903	2.121
12	1.120	1.158	1.210	1.382	1.479	1.785	1.959
13	1.108	1.145	1.192	1.344	1.428	1.690	1.835
14	1.101	1.132	1.175	1.311	1.387	1.614	1.739
15	1.094	1.123	1.162	1.285	1.352	1.554	1.663
16	1.087	1.115	1.150	1.263	1.324	1.505	1.601
17	1.083	1.108	1.141	1.243	1.299	1.463	1.548
18	1.078	1.099	1.131	1.226	1.279	1.426	1.504
19	1.073	1.095	1.124	1.212	1.260	1.397	1.468
20	1.071	1.090	1.117	1.200	1.245	1.371	1.436
21	1.066	1.083	1.110	1.188	1.230	1.347	1.407
31	1.048	1.057	1.072	1.123	1.146	1.217	1.250
41	1.034	1.042	1.054	1.089	1.107	1.155	1.178
61	1.025	1.029	1.035	1.060	1.069	1.101	1.115
121	1.014	1.014	1.018	1.028	1.034	1.048	1.053
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A27 (cont.). Values of the penalty ratio $r^{LN}_{\mu,\sigma}(V_Y,n,P)$ in equation (II.47) for left-prediction in lognormal sequences with unknown mean and variance. Estimated coefficient of variation: V_Y = 0.40.

RIGHT PREDICTION INTERVAL V_Y = 0.05

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	1.153	3.291	6.122	16.337	18.664	19.051	19.051
3	0.899	1.986	2.974	6.322	8.318	13.806	15.978
4	0.819	1.687	2.368	4.287	5.321	8.297	9.808
5	0.780	1.556	2.119	3.554	4.262	6.205	7.177
6	0.756	1.480	1.984	3.183	3.744	5.206	5.917
7	0.740	1.434	1.899	2.967	3.442	4.648	5.215
8	0.727	1.399	1.842	2.820	3.245	4.287	4.766
9	0.719	1.375	1.799	2.717	3.106	4.065	4.463
10	0.712	1.354	1.767	2.639	3.003	3.861	4.241
11	0.706	1.338	1.741	2.579	2.924	3.719	4.073
12	0.702	1.326	1.720	2.532	2.862	3.618	3.945
13	0.697	1.316	1.704	2.494	2.811	3.532	3.841
14	0.694	1.307	1.689	2.461	2.769	3.460	3.755
15	0.692	1.299	1.677	2.434	2.733	3.402	3.685
16	0.689	1.293	1.666	2.411	2.703	3.353	3.625
17	0.687	1.288	1.658	2.390	2.676	3.309	3.572
18	0.685	1.282	1.649	2.372	2.654	3.270	3.526
19	0.683	1.278	1.642	2.357	2.634	3.238	3.489
20	0.682	1.275	1.636	2.343	2.617	3.210	3.455
21	0.680	1.269	1.630	2.330	2.600	3.183	3.423
31	0.673	1.249	1.594	2.257	2.503	3.031	3.242
41	0.668	1.237	1.577	2.218	2.456	2.954	3.151
61	0.664	1.227	1.559	2.184	2.410	2.885	3.069
121	0.659	1.215	1.543	2.146	2.365	2.813	2.985
∞	0.654	1.204	1.525	2.111	2.322	2.746	2.909

Table A28. Values of the coefficient $\beta^{LN}_{\mu,\sigma}(V_Y,n,P)$ in equation (II.46) for right-hand prediction in lognormal sequences with unknown mean and variance. V_Y is the estimated coefficient of variation of the lognormal population and n is the sample size. For other values of V_Y, see continuation of the table.

RIGHT PREDICTION INTERVAL V_Y= 0.10

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	1.086	2.883	4.911	8.911	9.092	9.096	9.096
3	0.857	1.813	2.630	5.039	6.210	8.405	8.858
4	0.785	1.556	2.134	3.642	4.377	6.199	6.955
5	0.749	1.442	1.926	3.087	3.624	4.965	5.565
6	0.727	1.376	1.811	2.797	3.234	4.299	4.778
7	0.713	1.335	1.739	2.625	3.000	3.905	4.304
8	0.701	1.304	1.689	2.506	2.846	3.642	3.990
9	0.693	1.283	1.652	2.423	2.736	3.477	3.771
10	0.687	1.265	1.625	2.358	2.654	3.323	3.608
11	0.682	1.250	1.603	2.309	2.590	3.215	3.483
12	0.677	1.240	1.585	2.270	2.540	3.137	3.387
13	0.673	1.231	1.570	2.239	2.499	3.071	3.308
14	0.671	1.222	1.558	2.212	2.464	3.015	3.242
15	0.668	1.216	1.547	2.190	2.435	2.969	3.188
16	0.666	1.211	1.538	2.170	2.411	2.931	3.142
17	0.664	1.206	1.530	2.153	2.389	2.896	3.102
18	0.662	1.200	1.523	2.138	2.371	2.866	3.066
19	0.661	1.197	1.517	2.125	2.354	2.841	3.037
20	0.660	1.194	1.512	2.114	2.340	2.818	3.011
21	0.658	1.189	1.506	2.103	2.327	2.797	2.986
31	0.651	1.171	1.475	2.042	2.247	2.676	2.843
41	0.646	1.161	1.460	2.009	2.208	2.614	2.772
61	0.643	1.152	1.445	1.981	2.170	2.558	2.707
121	0.639	1.141	1.430	1.949	2.132	2.501	2.639
∞	0.634	1.132	1.415	1.919	2.096	2.446	2.578

Table A28 (cont.). Values of the coefficient $\beta^{LN}_{\mu,\sigma}(V_Y,n,P)$ in equation (II.46) for right-prediction in lognormal sequences with unknown mean and variance. Estimated coefficient of variation: V_Y = 0.10.

RIGHT PREDICTION INTERVAL V_Y = 0.20

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	0.965	2.239	3.296	4.181	4.183	4.183	4.183
3	0.779	1.518	2.079	3.349	3.759	4.158	4.180
4	0.720	1.329	1.747	2.682	3.057	3.756	3.948
5	0.691	1.242	1.600	2.364	2.672	3.318	3.550
6	0.672	1.192	1.517	2.186	2.452	3.020	3.239
7	0.660	1.160	1.464	2.076	2.312	2.823	3.023
8	0.650	1.136	1.428	1.998	2.216	2.682	2.866
9	0.644	1.120	1.401	1.943	2.147	2.591	2.752
10	0.638	1.106	1.380	1.900	2.094	2.503	2.664
11	0.634	1.094	1.364	1.867	2.053	2.440	2.594
12	0.631	1.086	1.350	1.841	2.021	2.394	2.540
13	0.627	1.079	1.339	1.819	1.994	2.354	2.495
14	0.625	1.073	1.330	1.800	1.971	2.321	2.456
15	0.623	1.068	1.322	1.785	1.952	2.293	2.425
16	0.621	1.063	1.315	1.772	1.935	2.269	2.397
17	0.619	1.060	1.309	1.760	1.921	2.248	2.373
18	0.618	1.055	1.304	1.749	1.909	2.229	2.352
19	0.616	1.053	1.299	1.741	1.897	2.213	2.334
20	0.616	1.050	1.295	1.733	1.888	2.199	2.318
21	0.614	1.047	1.291	1.725	1.879	2.186	2.303
31	0.609	1.032	1.268	1.682	1.824	2.109	2.215
41	0.604	1.024	1.256	1.659	1.798	2.069	2.170
61	0.602	1.017	1.244	1.639	1.771	2.033	2.128
121	0.598	1.008	1.233	1.616	1.745	1.995	2.085
∞	0.594	1.001	1.222	1.595	1.720	1.959	2.045

Table A28 (cont.). Values of the coefficient $\beta^{LN}_{\mu,\sigma}(V_Y,n,P)$ in equation (II.46) for right-prediction in lognormal sequences with unknown mean and variance. Estimated coefficient of variation: V_Y = 0.20.

RIGHT PREDICTION INTERVAL V_Y = 0.40

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	0.771	1.434	1.745	1.832	1.832	1.832	1.832
3	0.649	1.097	1.368	1.755	1.811	1.832	1.832
4	0.608	0.992	1.215	1.591	1.694	1.811	1.825
5	0.588	0.941	1.140	1.482	1.588	1.749	1.787
6	0.575	0.911	1.096	1.412	1.514	1.685	1.734
7	0.567	0.892	1.068	1.366	1.462	1.633	1.686
8	0.560	0.878	1.048	1.332	1.425	1.591	1.645
9	0.555	0.868	1.032	1.308	1.396	1.562	1.612
10	0.551	0.859	1.021	1.288	1.374	1.532	1.585
11	0.548	0.852	1.012	1.272	1.357	1.510	1.563
12	0.546	0.847	1.004	1.260	1.342	1.493	1.545
13	0.543	0.843	0.998	1.250	1.330	1.478	1.529
14	0.542	0.839	0.992	1.241	1.320	1.466	1.516
15	0.540	0.836	0.988	1.233	1.311	1.455	1.504
16	0.539	0.833	0.984	1.227	1.304	1.446	1.494
17	0.538	0.831	0.981	1.221	1.297	1.437	1.485
18	0.537	0.828	0.977	1.216	1.292	1.430	1.477
19	0.536	0.827	0.975	1.211	1.287	1.423	1.471
20	0.536	0.825	0.972	1.208	1.282	1.418	1.465
21	0.535	0.823	0.970	1.204	1.278	1.412	1.459
31	0.531	0.814	0.956	1.182	1.252	1.380	1.424
41	0.527	0.809	0.950	1.170	1.239	1.363	1.406
61	0.525	0.804	0.943	1.160	1.227	1.348	1.389
121	0.523	0.799	0.936	1.148	1.214	1.331	1.370
∞	0.520	0.794	0.929	1.138	1.201	1.315	1.353

Table A28 (cont.). Values of the coefficient $\beta^{LN}_{\mu,\sigma}(V_Y,n,P)$ in equation (II.46) for right-prediction in lognormal sequences with unknown mean and variance. Estimated coefficient of variation: V_Y = 0.40.

RIGHT PREDICTION INTERVAL V_Y = 0.05

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	1.764	2.733	4.014	7.738	8.039	6.938	6.549
3	1.374	1.649	1.950	2.994	3.583	5.028	5.493
4	1.253	1.401	1.552	2.030	2.292	3.022	3.372
5	1.193	1.292	1.389	1.683	1.836	2.260	2.467
6	1.156	1.229	1.301	1.508	1.613	1.896	2.034
7	1.132	1.191	1.245	1.405	1.482	1.693	1.793
8	1.112	1.161	1.208	1.336	1.398	1.561	1.639
9	1.100	1.141	1.179	1.287	1.338	1.481	1.535
10	1.088	1.125	1.159	1.250	1.293	1.406	1.458
11	1.080	1.111	1.142	1.221	1.259	1.355	1.400
12	1.073	1.101	1.128	1.199	1.233	1.318	1.356
13	1.066	1.093	1.117	1.181	1.211	1.286	1.321
14	1.062	1.085	1.107	1.165	1.193	1.260	1.291
15	1.058	1.079	1.099	1.153	1.177	1.239	1.267
16	1.053	1.074	1.093	1.142	1.164	1.221	1.246
17	1.051	1.069	1.087	1.132	1.153	1.205	1.228
18	1.048	1.064	1.081	1.123	1.143	1.191	1.212
19	1.045	1.061	1.077	1.116	1.134	1.179	1.200
20	1.044	1.058	1.073	1.110	1.127	1.169	1.188
21	1.041	1.054	1.069	1.104	1.120	1.159	1.177
31	1.030	1.037	1.045	1.069	1.078	1.104	1.114
41	1.021	1.027	1.034	1.050	1.058	1.076	1.083
61	1.015	1.019	1.022	1.034	1.038	1.051	1.055
121	1.008	1.009	1.011	1.016	1.019	1.025	1.026
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A29. Values of the penalty ratio $r^{LN}_{\mu,\sigma}(V_Y,n,P)$ in equation (II.47) for right-hand prediction in lognormal sequences with unknown mean and variance. V_Y is the estimated coefficient of variation and n is the sample size. For other values of V_Y, see continuation of the table.

RIGHT PREDICTION INTERVAL V_Y = 0.10

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	1.714	2.547	3.471	4.643	4.338	3.719	3.529
3	1.352	1.602	1.859	2.626	2.963	3.436	3.436
4	1.239	1.375	1.509	1.897	2.088	2.534	2.698
5	1.183	1.274	1.361	1.609	1.729	2.030	2.159
6	1.147	1.216	1.280	1.457	1.543	1.758	1.853
7	1.125	1.180	1.229	1.367	1.432	1.596	1.670
8	1.106	1.152	1.194	1.306	1.358	1.489	1.548
9	1.094	1.134	1.168	1.262	1.305	1.422	1.463
10	1.084	1.118	1.148	1.229	1.266	1.359	1.399
11	1.076	1.105	1.133	1.203	1.236	1.314	1.351
12	1.069	1.096	1.120	1.183	1.212	1.283	1.314
13	1.062	1.088	1.110	1.167	1.192	1.255	1.283
14	1.058	1.080	1.101	1.152	1.176	1.233	1.258
15	1.055	1.075	1.093	1.141	1.162	1.214	1.237
16	1.051	1.070	1.087	1.131	1.150	1.198	1.219
17	1.048	1.066	1.082	1.122	1.140	1.184	1.203
18	1.045	1.061	1.076	1.114	1.131	1.172	1.190
19	1.043	1.058	1.072	1.107	1.123	1.162	1.178
20	1.041	1.055	1.068	1.102	1.117	1.152	1.168
21	1.039	1.051	1.065	1.096	1.110	1.144	1.158
31	1.028	1.035	1.043	1.064	1.072	1.094	1.103
41	1.020	1.026	1.032	1.047	1.053	1.069	1.075
61	1.015	1.018	1.021	1.032	1.035	1.046	1.050
121	1.008	1.008	1.011	1.015	1.017	1.022	1.024
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A29 (cont.). Values of the penalty ratio $r^{LN}_{\mu,\sigma}(V_Y,n,P)$ in equation (II.47) for right-prediction in lognormal sequences with unknown mean and variance. Estimated coefficient of variation: V_Y = 0.10

RIGHT PREDICTION INTERVAL V_Y= 0.20

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	1.625	2.237	2.698	2.621	2.432	2.136	2.045
3	1.312	1.517	1.702	2.099	2.185	2.123	2.044
4	1.213	1.328	1.430	1.681	1.777	1.917	1.930
5	1.163	1.241	1.310	1.482	1.553	1.694	1.736
6	1.131	1.190	1.242	1.370	1.425	1.542	1.584
7	1.111	1.159	1.199	1.301	1.344	1.441	1.478
8	1.095	1.135	1.169	1.253	1.288	1.369	1.401
9	1.084	1.119	1.147	1.218	1.248	1.323	1.346
10	1.075	1.105	1.130	1.191	1.217	1.278	1.302
11	1.068	1.093	1.116	1.170	1.194	1.246	1.268
12	1.062	1.085	1.105	1.154	1.175	1.222	1.242
13	1.056	1.078	1.096	1.140	1.159	1.202	1.220
14	1.052	1.072	1.089	1.129	1.146	1.185	1.201
15	1.049	1.067	1.082	1.119	1.134	1.171	1.185
16	1.045	1.062	1.077	1.111	1.125	1.158	1.172
17	1.043	1.059	1.072	1.103	1.117	1.148	1.160
18	1.041	1.054	1.067	1.097	1.110	1.138	1.150
19	1.038	1.052	1.063	1.091	1.103	1.130	1.141
20	1.037	1.049	1.060	1.086	1.097	1.123	1.133
21	1.035	1.046	1.057	1.081	1.092	1.116	1.126
31	1.025	1.031	1.038	1.055	1.060	1.077	1.083
41	1.018	1.023	1.028	1.040	1.045	1.056	1.061
61	1.013	1.016	1.019	1.027	1.030	1.038	1.041
121	1.007	1.008	1.010	1.013	1.015	1.019	1.019
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A29 (cont.). Values of the penalty ratio $r^{LN}_{\mu,\sigma}(V_Y,n,P)$ in equation (II.47) for right-prediction in lognormal sequences with unknown mean and variance. Estimated coefficient of variation: V_Y = 0.20.

RIGHT PREDICTION INTERVAL V_Y = 0.40

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
2	1.483	1.806	1.878	1.610	1.525	1.393	1.354
3	1.247	1.382	1.472	1.543	1.508	1.393	1.354
4	1.170	1.249	1.307	1.399	1.410	1.378	1.349
5	1.131	1.186	1.227	1.303	1.322	1.331	1.321
6	1.106	1.148	1.180	1.242	1.260	1.282	1.282
7	1.090	1.124	1.149	1.201	1.217	1.242	1.246
8	1.077	1.106	1.127	1.171	1.186	1.210	1.216
9	1.068	1.093	1.111	1.149	1.162	1.188	1.192
10	1.060	1.082	1.099	1.132	1.144	1.166	1.172
11	1.055	1.073	1.088	1.118	1.129	1.149	1.155
12	1.050	1.067	1.080	1.107	1.117	1.136	1.142
13	1.045	1.062	1.074	1.099	1.107	1.125	1.130
14	1.042	1.057	1.068	1.091	1.099	1.115	1.120
15	1.040	1.053	1.063	1.084	1.092	1.107	1.112
16	1.037	1.049	1.059	1.078	1.085	1.100	1.104
17	1.035	1.046	1.055	1.073	1.080	1.093	1.098
18	1.033	1.043	1.052	1.069	1.075	1.088	1.092
19	1.031	1.041	1.049	1.065	1.071	1.083	1.087
20	1.030	1.039	1.046	1.062	1.067	1.079	1.082
21	1.028	1.036	1.044	1.058	1.064	1.074	1.078
31	1.020	1.025	1.029	1.039	1.042	1.050	1.052
41	1.015	1.019	1.022	1.029	1.032	1.037	1.039
61	1.011	1.013	1.014	1.020	1.021	1.025	1.026
121	1.006	1.006	1.007	1.010	1.010	1.012	1.013
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A29 (cont.). Values of the penalty ratio $r^{LN}_{\mu,\sigma}(V_Y,n,P)$ in equation (II.47) for right-prediction in lognormal sequences with unknown mean and variance. Estimated coefficient of variation: V_Y = 0.40.

LEFT PREDICTION INTERVAL V_Y= 0.05

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.951	3.139	7.034	74.245	438.757	*******	*******
2	0.768	1.856	2.965	7.896	12.191	39.003	73.211
3	0.718	1.599	2.350	4.822	6.424	12.644	17.241
4	0.694	1.491	2.115	3.894	4.897	8.177	10.204
5	0.680	1.432	1.992	3.460	4.223	6.491	7.767
6	0.671	1.396	1.916	3.212	3.848	5.631	6.575
7	0.664	1.370	1.866	3.051	3.611	5.115	5.879
8	0.659	1.352	1.829	2.939	3.449	4.775	5.426
9	0.656	1.337	1.801	2.856	3.331	4.533	5.110
10	0.653	1.326	1.779	2.794	3.241	4.353	4.877
11	0.651	1.317	1.763	2.744	3.170	4.215	4.698
12	0.649	1.310	1.748	2.704	3.114	4.104	4.558
13	0.648	1.304	1.737	2.670	3.066	4.014	4.444
14	0.646	1.299	1.726	2.642	3.028	3.940	4.349
15	0.645	1.295	1.718	2.618	2.995	3.878	4.270
16	0.644	1.291	1.711	2.597	2.966	3.824	4.203
17	0.643	1.287	1.704	2.580	2.941	3.778	4.145
18	0.642	1.284	1.698	2.564	2.919	3.737	4.095
19	0.642	1.282	1.693	2.550	2.900	3.702	4.050
20	0.641	1.278	1.689	2.538	2.883	3.671	4.012
30	0.637	1.263	1.660	2.461	2.779	3.482	3.778
40	0.635	1.256	1.646	2.425	2.729	3.395	3.670
60	0.633	1.249	1.633	2.390	2.681	3.311	3.567
120	0.631	1.242	1.620	2.355	2.634	3.230	3.469
∞	0.628	1.235	1.606	2.321	2.590	3.153	3.377

Table A30. Values of the coefficient $\beta^{LN}_{\sigma Z}(V_Y,n,P)$ for left-hand prediction in lognormal sequences. In the associated normal sequence the mean is assumed known and the variance is unknown. V_Y is the estimated coefficient of variation of the lognormal population. For other values of V_Y, see continuation of the table.

LEFT PREDICTION INTERVAL V_Y = 0.10

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.904	3.208	7.895	207.264	*******	*******	*******
2	0.722	1.828	3.015	9.035	15.262	74.822	202.505
3	0.673	1.561	2.349	5.141	7.112	15.972	23.756
4	0.649	1.450	2.100	4.057	5.230	9.415	12.268
5	0.636	1.391	1.970	3.565	4.436	7.196	8.862
6	0.627	1.353	1.891	3.288	4.004	6.120	7.304
7	0.620	1.327	1.838	3.110	3.735	5.491	6.427
8	0.615	1.308	1.800	2.986	3.552	5.084	5.869
9	0.612	1.294	1.771	2.896	3.420	4.798	5.486
10	0.609	1.282	1.748	2.828	3.320	4.588	5.206
11	0.607	1.273	1.731	2.774	3.242	4.426	4.994
12	0.605	1.266	1.716	2.730	3.179	4.299	4.827
13	0.604	1.260	1.704	2.693	3.127	4.195	4.693
14	0.602	1.254	1.693	2.663	3.084	4.109	4.582
15	0.601	1.250	1.684	2.637	3.048	4.038	4.491
16	0.600	1.246	1.677	2.615	3.016	3.977	4.413
17	0.599	1.242	1.670	2.596	2.989	3.925	4.346
18	0.598	1.239	1.664	2.579	2.965	3.878	4.288
19	0.598	1.237	1.659	2.564	2.944	3.838	4.236
20	0.597	1.234	1.654	2.551	2.925	3.803	4.192
30	0.593	1.218	1.624	2.469	2.812	3.590	3.925
40	0.591	1.211	1.610	2.429	2.757	3.492	3.802
60	0.589	1.204	1.596	2.392	2.705	3.398	3.685
120	0.587	1.197	1.583	2.355	2.655	3.309	3.575
∞	0.584	1.190	1.569	2.318	2.607	3.222	3.472

Table A30 (cont.). Values of the coefficient $\beta^{LN}_{\sigma_Z}(V_Y,n,P)$ for left-prediction in lognormal sequences. In the associated normal sequence the mean is known and the variance is not. Estimated coefficient of variation of the lognormal population: $V_Y = 0.10$.

LEFT PREDICTION INTERVAL V_Y= 0.20

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.817	3.363	10.140	*******	*******	*******	*******
2	0.638	1.776	3.130	12.110	25.099	337.159	*******
3	0.590	1.490	2.354	5.899	8.859	26.818	48.849
4	0.567	1.374	2.074	4.432	6.025	12.796	18.385
5	0.554	1.311	1.930	3.804	4.932	8.994	11.804
6	0.546	1.272	1.844	3.460	4.364	7.322	9.167
7	0.539	1.245	1.787	3.244	4.019	6.398	7.787
8	0.534	1.226	1.745	3.096	3.788	5.819	6.948
9	0.531	1.211	1.714	2.988	3.624	5.423	6.389
10	0.529	1.199	1.689	2.908	3.500	5.136	5.991
11	0.526	1.190	1.671	2.843	3.405	4.919	5.693
12	0.524	1.182	1.655	2.792	3.328	4.750	5.463
13	0.523	1.176	1.642	2.750	3.265	4.613	5.280
14	0.521	1.171	1.630	2.714	3.213	4.500	5.129
15	0.520	1.166	1.621	2.684	3.169	4.408	5.006
16	0.519	1.162	1.613	2.658	3.132	4.328	4.901
17	0.518	1.158	1.606	2.636	3.098	4.261	4.812
18	0.518	1.155	1.599	2.616	3.070	4.201	4.735
19	0.518	1.153	1.594	2.599	3.045	4.150	4.667
20	0.517	1.149	1.589	2.584	3.022	4.105	4.609
30	0.513	1.134	1.557	2.490	2.888	3.836	4.261
40	0.511	1.126	1.542	2.445	2.824	3.713	4.104
60	0.509	1.119	1.528	2.402	2.763	3.596	3.956
120	0.507	1.112	1.513	2.360	2.704	3.486	3.817
∞	0.504	1.104	1.498	2.319	2.649	3.381	3.688

Table A30 (cont.). Values of the coefficient $\beta^{LN}_{\sigma Z}(V_Y,n,P)$ for left-prediction in lognormal sequences. In the associated normal sequence the mean is known and the variance is not. Estimated coefficient of variation of the lognormal population: $V_Y = 0.20$.

LEFT PREDICTION INTERVAL V_Y = 0.40

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.668	3.735	17.534	*******	*******	*******	*******
2	0.497	1.685	3.406	23.055	76.010	*******	*******
3	0.452	1.365	2.379	7.950	14.308	85.144	245.331
4	0.431	1.238	2.035	5.372	8.190	25.131	45.071
5	0.419	1.171	1.865	4.386	6.208	14.635	22.151
6	0.411	1.130	1.763	3.876	5.262	10.815	15.059
7	0.405	1.102	1.698	3.566	4.715	8.914	11.828
8	0.401	1.081	1.650	3.358	4.362	7.800	10.027
9	0.398	1.066	1.614	3.210	4.117	7.072	8.897
10	0.395	1.054	1.587	3.101	3.934	6.562	8.124
11	0.393	1.044	1.565	3.014	3.796	6.186	7.565
12	0.391	1.036	1.547	2.946	3.686	5.898	7.144
13	0.390	1.029	1.533	2.889	3.595	5.669	6.815
14	0.389	1.024	1.520	2.842	3.523	5.484	6.549
15	0.388	1.019	1.510	2.803	3.461	5.333	6.336
16	0.387	1.015	1.501	2.769	3.409	5.205	6.155
17	0.386	1.010	1.493	2.740	3.362	5.097	6.003
18	0.385	1.007	1.485	2.714	3.322	5.002	5.874
19	0.385	1.005	1.479	2.691	3.289	4.920	5.759
20	0.384	1.002	1.474	2.672	3.257	4.851	5.663
30	0.381	0.985	1.438	2.551	3.074	4.434	5.097
40	0.379	0.978	1.422	2.494	2.988	4.249	4.848
60	0.377	0.970	1.406	2.439	2.907	4.076	4.618
120	0.376	0.963	1.390	2.387	2.829	3.914	4.405
∞	0.373	0.955	1.374	2.335	2.756	3.761	4.211

Table A30 (cont.). Values of the coefficient $\beta^{LN}_{\sigma Z}(V_Y,n,P)$ for left-prediction in lognormal sequences. In the associated normal sequence the mean is known and the variance is not. Estimated coefficient of variation of the lognormal population: $V_Y = 0.40$.

LEFT PREDICTION INTERVAL V_Y = 0.05

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.515	2.542	4.379	31.987	169.422	*******	*******
2	1.223	1.503	1.846	3.402	4.707	12.371	21.681
3	1.143	1.295	1.463	2.078	2.481	4.010	5.106
4	1.105	1.207	1.317	1.678	1.891	2.594	3.022
5	1.083	1.160	1.240	1.491	1.631	2.059	2.300
6	1.069	1.130	1.193	1.384	1.486	1.786	1.947
7	1.058	1.110	1.162	1.314	1.394	1.622	1.741
8	1.050	1.095	1.139	1.266	1.332	1.514	1.607
9	1.045	1.083	1.121	1.231	1.286	1.438	1.513
10	1.041	1.074	1.108	1.204	1.251	1.381	1.444
11	1.036	1.067	1.097	1.182	1.224	1.337	1.391
12	1.033	1.061	1.088	1.165	1.202	1.302	1.350
13	1.031	1.056	1.081	1.150	1.184	1.273	1.316
14	1.028	1.052	1.075	1.138	1.169	1.250	1.288
15	1.027	1.048	1.070	1.128	1.156	1.230	1.265
16	1.025	1.045	1.065	1.119	1.145	1.213	1.245
17	1.023	1.042	1.061	1.112	1.136	1.198	1.227
18	1.022	1.039	1.057	1.105	1.127	1.185	1.213
19	1.022	1.038	1.054	1.099	1.120	1.174	1.199
20	1.020	1.035	1.052	1.093	1.113	1.164	1.188
30	1.014	1.023	1.033	1.060	1.073	1.105	1.119
40	1.011	1.017	1.025	1.045	1.054	1.077	1.087
60	1.008	1.011	1.017	1.029	1.035	1.050	1.056
120	1.005	1.006	1.008	1.015	1.017	1.025	1.027
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A31. Values of the penalty ratio $r_{\sigma_Z}^{LN}(V_Y,n,P)$ for left-hand prediction in lognormal sequences. In the associated normal sequence the mean is known but the variance is not. V_Y is the estimated coefficient of variation of the lognormal population. For other values of V_Y, see continuation of the table.

LEFT PREDICTION INTERVAL V_Y= 0.10

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.548	2.696	5.032	89.399	*******	*******	*******
2	1.236	1.537	1.922	3.897	5.855	23.220	58.329
3	1.151	1.313	1.497	2.217	2.728	4.957	6.843
4	1.111	1.219	1.338	1.750	2.007	2.922	3.534
5	1.088	1.169	1.256	1.538	1.702	2.233	2.553
6	1.073	1.137	1.205	1.418	1.536	1.899	2.104
7	1.061	1.115	1.172	1.341	1.433	1.704	1.851
8	1.053	1.100	1.147	1.288	1.363	1.578	1.690
9	1.048	1.088	1.129	1.249	1.312	1.489	1.580
10	1.043	1.078	1.114	1.220	1.274	1.424	1.500
11	1.038	1.070	1.103	1.196	1.244	1.374	1.438
12	1.035	1.064	1.094	1.177	1.220	1.334	1.391
13	1.033	1.059	1.086	1.162	1.200	1.302	1.352
14	1.030	1.055	1.079	1.149	1.183	1.275	1.320
15	1.028	1.051	1.074	1.137	1.169	1.253	1.294
16	1.026	1.048	1.069	1.128	1.157	1.234	1.271
17	1.025	1.044	1.065	1.120	1.147	1.218	1.252
18	1.023	1.041	1.061	1.112	1.137	1.204	1.235
19	1.023	1.040	1.057	1.106	1.129	1.191	1.220
20	1.022	1.037	1.054	1.100	1.122	1.180	1.208
30	1.015	1.024	1.035	1.065	1.079	1.114	1.130
40	1.012	1.018	1.026	1.048	1.058	1.084	1.095
60	1.008	1.012	1.018	1.032	1.038	1.055	1.062
120	1.005	1.006	1.009	1.016	1.018	1.027	1.030
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A31 (cont.). Values of the penalty ratio $r^{LN}_{\sigma_Z}(V_Y,n,P)$ for left-prediction in lognormal sequences. The associated normal sequence has known mean and unknown variance. Estimated coefficient of variation of the lognormal population: $V_Y = 0.10$.

LEFT PREDICTION INTERVAL V_Y= 0.20

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.620	3.045	6.767	963.371	*******	*******	*******
2	1.265	1.608	2.089	5.223	9.476	99.731	579.484
3	1.169	1.350	1.571	2.544	3.345	7.933	13.246
4	1.124	1.244	1.384	1.911	2.275	3.785	4.985
5	1.098	1.188	1.288	1.641	1.862	2.660	3.201
6	1.081	1.152	1.231	1.492	1.648	2.166	2.486
7	1.068	1.128	1.192	1.399	1.517	1.892	2.112
8	1.059	1.110	1.165	1.335	1.430	1.721	1.884
9	1.054	1.097	1.144	1.289	1.368	1.604	1.733
10	1.048	1.086	1.127	1.254	1.321	1.519	1.624
11	1.042	1.077	1.115	1.226	1.285	1.455	1.544
12	1.039	1.071	1.104	1.204	1.257	1.405	1.481
13	1.037	1.065	1.096	1.186	1.233	1.364	1.432
14	1.033	1.060	1.088	1.170	1.213	1.331	1.391
15	1.031	1.056	1.082	1.158	1.197	1.304	1.357
16	1.029	1.052	1.077	1.146	1.182	1.280	1.329
17	1.028	1.049	1.072	1.137	1.170	1.260	1.305
18	1.026	1.046	1.067	1.128	1.159	1.243	1.284
19	1.026	1.044	1.064	1.121	1.150	1.228	1.265
20	1.024	1.041	1.061	1.114	1.141	1.214	1.250
30	1.017	1.027	1.039	1.074	1.090	1.135	1.155
40	1.013	1.020	1.029	1.054	1.066	1.098	1.113
60	1.009	1.013	1.020	1.036	1.043	1.064	1.073
120	1.006	1.007	1.010	1.018	1.021	1.031	1.035
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A31 (cont.). Values of the penalty ratio $r_{\sigma_Z}^{LN}(V_Y,n,P)$ for left-prediction in lognormal sequences. The associated normal sequence has known mean and unknown variance. Estimated coefficient of variation of the lognormal population: $V_Y = 0.20$.

LEFT PREDICTION INTERVAL V_Y = 0.40

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.791	3.910	12.766	*******	*******	*******	*******
2	1.332	1.764	2.480	9.873	27.577	*******	*******
3	1.211	1.429	1.732	3.404	5.191	22.638	58.255
4	1.154	1.296	1.482	2.301	2.971	6.682	10.702
5	1.122	1.226	1.358	1.878	2.252	3.891	5.260
6	1.101	1.183	1.284	1.660	1.909	2.876	3.576
7	1.085	1.153	1.236	1.527	1.711	2.370	2.809
8	1.073	1.132	1.202	1.438	1.583	2.074	2.381
9	1.066	1.116	1.175	1.375	1.494	1.880	2.113
10	1.059	1.103	1.155	1.328	1.427	1.745	1.929
11	1.053	1.092	1.140	1.291	1.377	1.645	1.796
12	1.048	1.084	1.126	1.262	1.337	1.568	1.696
13	1.046	1.077	1.116	1.237	1.304	1.507	1.618
14	1.041	1.072	1.107	1.217	1.278	1.458	1.555
15	1.039	1.067	1.099	1.200	1.256	1.418	1.504
16	1.037	1.062	1.093	1.186	1.237	1.384	1.462
17	1.034	1.058	1.087	1.174	1.220	1.355	1.425
18	1.032	1.054	1.081	1.162	1.205	1.330	1.395
19	1.032	1.052	1.077	1.153	1.193	1.308	1.368
20	1.030	1.049	1.073	1.144	1.182	1.290	1.345
30	1.021	1.032	1.047	1.092	1.115	1.179	1.210
40	1.016	1.024	1.035	1.068	1.084	1.130	1.151
60	1.011	1.016	1.023	1.045	1.055	1.084	1.097
120	1.007	1.008	1.012	1.022	1.026	1.041	1.046
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A31 (cont.). Values of the penalty ratio $r_{\sigma_Z}^{LN}(V_Y,n,P)$ for left-prediction in lognormal sequences. The associated normal sequence has known mean and unknown variance. Estimated coefficient of variation of the lognormal population: $V_Y = 0.40$.

RIGHT PREDICTION INTERVAL V_Y = 0.05

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.951	2.736	5.172	15.170	18.260	19.051	19.051
2	0.784	1.735	2.607	5.616	7.463	12.815	15.127
3	0.737	1.519	2.134	3.886	4.840	7.628	9.075
4	0.715	1.427	1.946	3.272	3.934	5.755	6.676
5	0.703	1.376	1.846	2.968	3.495	4.877	5.551
6	0.694	1.345	1.784	2.789	3.241	4.383	4.923
7	0.688	1.323	1.743	2.671	3.076	4.070	4.529
8	0.683	1.307	1.712	2.587	2.960	3.856	4.260
9	0.680	1.294	1.689	2.525	2.876	3.700	4.067
10	0.678	1.284	1.671	2.478	2.810	3.582	3.921
11	0.675	1.276	1.657	2.440	2.759	3.490	3.807
12	0.673	1.270	1.645	2.409	2.717	3.416	3.716
13	0.672	1.265	1.635	2.383	2.682	3.355	3.642
14	0.670	1.260	1.626	2.362	2.653	3.304	3.579
15	0.669	1.257	1.619	2.343	2.629	3.261	3.527
16	0.669	1.253	1.613	2.327	2.608	3.224	3.482
17	0.668	1.250	1.608	2.314	2.589	3.193	3.443
18	0.667	1.247	1.603	2.302	2.572	3.164	3.410
19	0.667	1.245	1.598	2.291	2.558	3.139	3.379
20	0.666	1.243	1.595	2.281	2.545	3.118	3.353
30	0.662	1.229	1.571	2.222	2.466	2.984	3.193
40	0.660	1.223	1.559	2.193	2.428	2.922	3.117
60	0.658	1.217	1.548	2.165	2.392	2.861	3.044
120	0.657	1.211	1.536	2.138	2.356	2.803	2.975
∞	0.654	1.204	1.525	2.111	2.322	2.746	2.909

Table A32. Values of the coefficient $\beta^{LN}_{\sigma z}(V_Y,n,P)$ for right-hand prediction in lognormal sequences. In the associated normal sequence the mean is assumed known and the variance is unknown. V_Y is the estimated coefficient of variation of the lognormal population. For other values of V_y, see continuation of the table.

RIGHT PREDICTION INTERVAL V_Y= 0.10

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.904	2.438	4.275	8.717	9.080	9.096	9.096
2	0.753	1.597	2.332	4.578	5.733	8.120	8.709
3	0.710	1.409	1.939	3.342	4.042	5.828	6.603
4	0.690	1.329	1.779	2.868	3.378	4.671	5.262
5	0.678	1.284	1.693	2.626	3.042	4.068	4.534
6	0.671	1.256	1.640	2.481	2.843	3.712	4.101
7	0.665	1.237	1.604	2.385	2.712	3.480	3.819
8	0.661	1.222	1.578	2.316	2.619	3.319	3.622
9	0.658	1.211	1.558	2.265	2.551	3.200	3.478
10	0.656	1.203	1.542	2.226	2.498	3.110	3.368
11	0.653	1.196	1.530	2.194	2.456	3.038	3.282
12	0.651	1.190	1.519	2.169	2.423	2.980	3.213
13	0.650	1.185	1.511	2.148	2.394	2.933	3.155
14	0.649	1.181	1.503	2.129	2.370	2.893	3.107
15	0.648	1.178	1.497	2.114	2.350	2.859	3.067
16	0.647	1.175	1.492	2.101	2.333	2.830	3.032
17	0.646	1.172	1.487	2.090	2.317	2.805	3.002
18	0.645	1.170	1.483	2.079	2.304	2.782	2.976
19	0.645	1.168	1.479	2.070	2.292	2.762	2.952
20	0.645	1.166	1.476	2.062	2.281	2.745	2.931
30	0.641	1.154	1.455	2.012	2.216	2.639	2.805
40	0.640	1.148	1.445	1.988	2.185	2.588	2.745
60	0.638	1.143	1.435	1.965	2.154	2.539	2.687
120	0.636	1.137	1.425	1.942	2.125	2.492	2.631
∞	0.634	1.132	1.415	1.919	2.096	2.446	2.578

Table A32 (cont.). Values of the coefficient $\beta^{LN}_{\sigma_Z}(V_Y,n,P)$ for right-prediction in lognormal sequences. In the associated normal sequence the mean is known and the variance is not. Estimated coefficient of variation of the lognormal population: $V_Y = 0.10$.

RIGHT PREDICTION INTERVAL V_Y = 0.20

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.818	1.953	3.008	4.175	4.183	4.183	4.183
2	0.693	1.360	1.882	3.150	3.608	4.134	4.175
3	0.658	1.218	1.609	2.514	2.893	3.640	3.866
4	0.641	1.155	1.494	2.230	2.535	3.192	3.437
5	0.631	1.121	1.431	2.076	2.337	2.906	3.131
6	0.625	1.099	1.391	1.982	2.214	2.721	2.923
7	0.620	1.084	1.365	1.918	2.132	2.593	2.777
8	0.616	1.073	1.345	1.871	2.072	2.501	2.671
9	0.614	1.064	1.330	1.837	2.028	2.432	2.592
10	0.612	1.057	1.318	1.810	1.993	2.378	2.529
11	0.610	1.052	1.309	1.789	1.966	2.335	2.479
12	0.609	1.047	1.301	1.771	1.943	2.299	2.439
13	0.608	1.043	1.295	1.756	1.924	2.270	2.405
14	0.607	1.040	1.289	1.744	1.908	2.245	2.376
15	0.606	1.038	1.284	1.733	1.895	2.225	2.352
16	0.605	1.035	1.280	1.724	1.883	2.206	2.331
17	0.604	1.033	1.277	1.716	1.872	2.191	2.312
18	0.604	1.031	1.273	1.709	1.863	2.176	2.296
19	0.604	1.030	1.270	1.702	1.855	2.164	2.282
20	0.603	1.028	1.268	1.697	1.848	2.153	2.269
30	0.600	1.019	1.252	1.662	1.804	2.085	2.191
40	0.599	1.014	1.244	1.644	1.782	2.052	2.153
60	0.597	1.010	1.237	1.628	1.761	2.020	2.116
120	0.596	1.005	1.229	1.612	1.740	1.989	2.080
∞	0.594	1.001	1.222	1.595	1.720	1.959	2.045

Table A32 (cont.). Values of the coefficient $\beta^{LN}_{\sigma_Z}(V_Y,n,P)$ for right-prediction in lognormal sequences. In the associated normal sequence the mean is known and the variance is not. Estimated coefficient of variation of the lognormal population: $V_Y = 0.20$.

RIGHT PREDICTION INTERVAL V_Y = 0.40

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.675	1.312	1.682	1.832	1.832	1.832	1.832
2	0.590	1.009	1.280	1.715	1.795	1.831	1.832
3	0.565	0.927	1.145	1.536	1.653	1.798	1.820
4	0.553	0.890	1.084	1.430	1.543	1.724	1.770
5	0.546	0.869	1.049	1.367	1.472	1.656	1.711
6	0.542	0.855	1.027	1.325	1.424	1.603	1.660
7	0.539	0.846	1.012	1.296	1.390	1.563	1.620
8	0.536	0.839	1.001	1.274	1.365	1.531	1.588
9	0.535	0.833	0.992	1.258	1.345	1.507	1.562
10	0.533	0.829	0.986	1.245	1.330	1.487	1.541
11	0.532	0.826	0.980	1.235	1.318	1.471	1.524
12	0.531	0.823	0.976	1.226	1.308	1.458	1.509
13	0.530	0.821	0.972	1.219	1.299	1.446	1.497
14	0.529	0.819	0.969	1.213	1.292	1.436	1.487
15	0.528	0.817	0.966	1.208	1.285	1.428	1.478
16	0.528	0.816	0.964	1.203	1.280	1.421	1.470
17	0.527	0.814	0.962	1.199	1.275	1.414	1.463
18	0.527	0.813	0.960	1.195	1.271	1.408	1.456
19	0.527	0.812	0.958	1.192	1.267	1.403	1.451
20	0.526	0.811	0.957	1.190	1.263	1.399	1.446
30	0.524	0.805	0.947	1.172	1.242	1.370	1.414
40	0.523	0.802	0.943	1.163	1.232	1.356	1.399
60	0.522	0.799	0.938	1.154	1.221	1.342	1.383
120	0.521	0.797	0.934	1.146	1.211	1.328	1.368
∞	0.520	0.794	0.929	1.138	1.201	1.315	1.353

Table A32 (cont.). Values of the coefficient $\beta^{LN}_{\sigma_Z}(V_Y,n,P)$ for right-prediction in lognormal sequences. In the associated normal sequence the mean is known and the variance is not. Estimated coefficient of variation of the lognormal population: $V_Y = 0.40$.

RIGHT PREDICTION INTERVAL V_Y= 0.05

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.455	2.272	3.391	7.185	7.865	6.938	6.549
2	1.199	1.441	1.709	2.660	3.214	4.667	5.200
3	1.128	1.261	1.399	1.841	2.085	2.778	3.120
4	1.094	1.185	1.276	1.550	1.694	2.096	2.295
5	1.074	1.143	1.210	1.406	1.506	1.776	1.908
6	1.062	1.117	1.170	1.321	1.396	1.596	1.693
7	1.052	1.098	1.143	1.265	1.325	1.482	1.557
8	1.045	1.085	1.123	1.225	1.275	1.404	1.465
9	1.041	1.075	1.107	1.196	1.239	1.348	1.398
10	1.037	1.066	1.095	1.174	1.210	1.305	1.348
11	1.032	1.060	1.086	1.156	1.188	1.271	1.309
12	1.030	1.055	1.078	1.141	1.170	1.244	1.278
13	1.028	1.050	1.072	1.129	1.155	1.222	1.252
14	1.025	1.047	1.066	1.119	1.143	1.203	1.230
15	1.024	1.044	1.062	1.110	1.132	1.188	1.213
16	1.022	1.041	1.058	1.102	1.123	1.174	1.197
17	1.021	1.038	1.054	1.096	1.115	1.163	1.184
18	1.020	1.035	1.051	1.090	1.108	1.152	1.172
19	1.020	1.034	1.048	1.085	1.102	1.143	1.162
20	1.018	1.032	1.046	1.081	1.096	1.136	1.153
30	1.013	1.021	1.030	1.052	1.062	1.087	1.098
40	1.010	1.016	1.022	1.039	1.046	1.064	1.072
60	1.007	1.010	1.015	1.026	1.030	1.042	1.047
120	1.004	1.005	1.007	1.013	1.015	1.021	1.023
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A33. Values of the penalty ratio $r_{\sigma_Z}^{LN}(V_Y,n,P)$ for right-hand prediction in lognormal sequences. In the associated normal sequence the mean is known but the variance is not. V_Y is the estimated coefficient of variation of the lognormal population. For other values of V_Y, see continuation of the table.

RIGHT PREDICTION INTERVAL V_Y = 0.10

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.427	2.154	3.021	4.542	4.332	3.719	3.529
2	1.188	1.412	1.648	2.385	2.735	3.320	3.378
3	1.121	1.246	1.370	1.741	1.928	2.383	2.561
4	1.089	1.174	1.257	1.494	1.612	1.910	2.041
5	1.070	1.135	1.197	1.368	1.451	1.663	1.759
6	1.058	1.110	1.159	1.293	1.356	1.518	1.591
7	1.049	1.093	1.134	1.242	1.294	1.423	1.481
8	1.043	1.080	1.115	1.207	1.250	1.357	1.405
9	1.039	1.071	1.101	1.180	1.217	1.308	1.349
10	1.035	1.063	1.090	1.160	1.192	1.271	1.307
11	1.031	1.057	1.081	1.143	1.172	1.242	1.273
12	1.028	1.052	1.074	1.130	1.156	1.219	1.246
13	1.027	1.048	1.068	1.119	1.142	1.199	1.224
14	1.024	1.044	1.062	1.110	1.131	1.183	1.205
15	1.023	1.041	1.058	1.102	1.121	1.169	1.190
16	1.021	1.039	1.054	1.095	1.113	1.157	1.176
17	1.020	1.036	1.051	1.089	1.106	1.147	1.164
18	1.019	1.034	1.048	1.083	1.099	1.137	1.154
19	1.019	1.032	1.045	1.079	1.094	1.129	1.145
20	1.017	1.030	1.043	1.075	1.088	1.122	1.137
30	1.012	1.020	1.028	1.049	1.057	1.079	1.088
40	1.009	1.015	1.021	1.036	1.042	1.058	1.065
60	1.007	1.010	1.014	1.024	1.028	1.038	1.042
120	1.004	1.005	1.007	1.012	1.014	1.019	1.021
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A33 (cont.). Values of the penalty ratio $r^{LN}_{\sigma_Z}(V_Y,n,P)$ for right-prediction in lognormal sequences. The associated normal sequence has known mean and unknown variance. Estimated coefficient of variation of the lognormal population: $V_Y = 0.10$.

RIGHT PREDICTION INTERVAL V_Y= 0.20

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.378	1.952	2.462	2.617	2.432	2.136	2.045
2	1.168	1.358	1.541	1.975	2.098	2.110	2.041
3	1.108	1.216	1.317	1.576	1.682	1.859	1.890
4	1.080	1.154	1.223	1.398	1.474	1.630	1.681
5	1.063	1.120	1.171	1.302	1.359	1.484	1.531
6	1.052	1.098	1.139	1.242	1.287	1.389	1.429
7	1.044	1.083	1.117	1.202	1.239	1.324	1.358
8	1.038	1.072	1.101	1.173	1.205	1.277	1.306
9	1.035	1.063	1.089	1.151	1.179	1.241	1.267
10	1.031	1.056	1.079	1.135	1.159	1.214	1.237
11	1.027	1.051	1.071	1.121	1.143	1.192	1.212
12	1.025	1.046	1.065	1.110	1.130	1.174	1.192
13	1.024	1.043	1.060	1.101	1.118	1.159	1.176
14	1.022	1.039	1.055	1.093	1.109	1.146	1.162
15	1.020	1.037	1.051	1.086	1.101	1.136	1.150
16	1.019	1.034	1.048	1.080	1.095	1.126	1.140
17	1.018	1.032	1.045	1.076	1.088	1.118	1.131
18	1.017	1.030	1.042	1.071	1.083	1.111	1.123
19	1.017	1.029	1.040	1.067	1.079	1.105	1.116
20	1.016	1.027	1.038	1.064	1.074	1.099	1.110
30	1.011	1.018	1.025	1.042	1.048	1.064	1.071
40	1.008	1.013	1.019	1.031	1.036	1.048	1.052
60	1.006	1.009	1.012	1.020	1.024	1.031	1.034
120	1.004	1.004	1.006	1.010	1.012	1.016	1.017
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A33 (cont.). Values of the penalty ratio $r_{\sigma_Z}^{LN}(V_Y,n,P)$ for right-prediction in lognormal sequences. The associated normal sequence has known mean and unknown variance. Estimated coefficient of variation of the lognormal population: $V_Y = 0.20$.

RIGHT PREDICTION INTERVAL V_Y= 0.40

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.298	1.653	1.810	1.610	1.525	1.393	1.354
2	1.134	1.271	1.377	1.508	1.494	1.393	1.354
3	1.087	1.168	1.232	1.350	1.376	1.368	1.345
4	1.064	1.120	1.166	1.257	1.285	1.312	1.308
5	1.051	1.094	1.129	1.201	1.225	1.260	1.265
6	1.042	1.077	1.105	1.165	1.185	1.219	1.227
7	1.036	1.065	1.089	1.139	1.157	1.189	1.197
8	1.031	1.057	1.077	1.120	1.136	1.165	1.173
9	1.028	1.050	1.068	1.106	1.120	1.146	1.154
10	1.025	1.045	1.061	1.095	1.107	1.131	1.139
11	1.022	1.040	1.055	1.086	1.097	1.119	1.126
12	1.020	1.037	1.050	1.078	1.088	1.109	1.116
13	1.019	1.034	1.046	1.072	1.081	1.100	1.107
14	1.017	1.031	1.042	1.066	1.075	1.093	1.099
15	1.016	1.029	1.040	1.062	1.070	1.086	1.092
16	1.016	1.027	1.037	1.058	1.065	1.081	1.086
17	1.015	1.025	1.035	1.054	1.061	1.076	1.081
18	1.014	1.024	1.033	1.051	1.058	1.071	1.076
19	1.014	1.023	1.031	1.048	1.055	1.068	1.072
20	1.013	1.021	1.029	1.046	1.052	1.064	1.069
30	1.009	1.014	1.019	1.030	1.034	1.042	1.045
40	1.007	1.011	1.014	1.022	1.025	1.032	1.034
60	1.005	1.007	1.010	1.015	1.017	1.021	1.022
120	1.003	1.004	1.005	1.007	1.008	1.010	1.011
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A33 (cont.). Values of the penalty ratio $r^{LN}_{\sigma_Z}(V_Y,n,P)$ for right-prediction in lognormal sequences. The associated normal sequence has known mean and unknown variance. Estimated coefficient of variation of the lognormal population: $V_Y = 0.40$.

LEFT PREDICTION INTERVAL V_Y = 0.05

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.904	1.780	2.321	3.375	3.774	4.618	4.956
2	0.778	1.529	1.992	2.887	3.225	3.937	4.222
3	0.730	1.438	1.870	2.709	3.024	3.691	3.956
4	0.707	1.388	1.807	2.616	2.921	3.561	3.816
5	0.692	1.360	1.769	2.561	2.857	3.485	3.734
6	0.681	1.340	1.743	2.521	2.814	3.428	3.673
7	0.674	1.326	1.723	2.496	2.783	3.394	3.636
8	0.668	1.314	1.710	2.473	2.759	3.362	3.602
9	0.664	1.307	1.698	2.458	2.741	3.342	3.581
10	0.660	1.299	1.689	2.443	2.726	3.321	3.558
11	0.657	1.292	1.682	2.430	2.713	3.304	3.540
12	0.655	1.289	1.675	2.423	2.704	3.294	3.529
13	0.653	1.285	1.670	2.415	2.695	3.283	3.517
14	0.652	1.281	1.666	2.408	2.687	3.273	3.506
15	0.650	1.277	1.662	2.403	2.681	3.266	3.499
16	0.649	1.275	1.659	2.398	2.675	3.259	3.491
17	0.648	1.272	1.656	2.393	2.670	3.253	3.484
18	0.647	1.270	1.653	2.388	2.666	3.245	3.477
19	0.646	1.268	1.651	2.385	2.662	3.242	3.473
20	0.645	1.267	1.649	2.383	2.658	3.238	3.469
30	0.639	1.257	1.634	2.364	2.635	3.212	3.440
40	0.636	1.250	1.627	2.351	2.623	3.194	3.421
60	0.634	1.247	1.620	2.344	2.612	3.184	3.410
120	0.631	1.240	1.613	2.331	2.601	3.166	3.391
∞	0.628	1.235	1.606	2.321	2.590	3.153	3.377

Table A34. Values of the coefficient $\beta^{LN}_{\mu_Z}(V_Y,n,P)$ for left-hand prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. V_Y is the coefficient of variation of the lognormal population. For other values of V_Y, see continuation of the table.

LEFT PREDICTION INTERVAL V_Y = 0.10

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.857	1.749	2.318	3.469	3.920	4.898	5.301
2	0.732	1.489	1.970	2.930	3.302	4.107	4.435
3	0.685	1.396	1.842	2.736	3.081	3.825	4.128
4	0.662	1.346	1.777	2.635	2.967	3.679	3.968
5	0.647	1.317	1.737	2.575	2.897	3.593	3.874
6	0.637	1.296	1.710	2.532	2.850	3.530	3.806
7	0.630	1.282	1.690	2.506	2.816	3.491	3.763
8	0.624	1.270	1.676	2.481	2.790	3.455	3.725
9	0.620	1.263	1.664	2.465	2.770	3.433	3.701
10	0.616	1.254	1.654	2.449	2.755	3.409	3.675
11	0.613	1.247	1.647	2.435	2.740	3.391	3.655
12	0.611	1.244	1.640	2.427	2.730	3.379	3.642
13	0.609	1.240	1.635	2.419	2.720	3.367	3.629
14	0.607	1.236	1.631	2.411	2.712	3.356	3.617
15	0.606	1.233	1.626	2.406	2.705	3.348	3.609
16	0.605	1.231	1.623	2.401	2.699	3.341	3.600
17	0.604	1.228	1.620	2.395	2.693	3.333	3.591
18	0.603	1.226	1.617	2.390	2.689	3.325	3.584
19	0.602	1.223	1.615	2.387	2.685	3.321	3.580
20	0.601	1.222	1.612	2.385	2.680	3.317	3.575
30	0.595	1.212	1.598	2.364	2.656	3.288	3.542
40	0.592	1.205	1.590	2.350	2.643	3.268	3.521
60	0.590	1.202	1.583	2.342	2.631	3.257	3.509
120	0.587	1.195	1.576	2.329	2.618	3.237	3.488
∞	0.584	1.190	1.569	2.318	2.607	3.222	3.472

Table A34 (cont.). Values of the coefficient $\beta_{\mu Z}^{LN}(V_Y,n,P)$ for left-prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. Coefficient of variation of the lognormal population: $V_Y = 0.10$.

LEFT PREDICTION INTERVAL V_Y = 0.20

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.771	1.691	2.319	3.685	4.254	5.561	6.125
2	0.648	1.415	1.930	3.028	3.479	4.497	4.930
3	0.602	1.317	1.791	2.799	3.209	4.133	4.524
4	0.579	1.265	1.721	2.681	3.073	3.948	4.316
5	0.565	1.235	1.678	2.612	2.990	3.839	4.196
6	0.555	1.213	1.649	2.563	2.933	3.760	4.109
7	0.548	1.199	1.627	2.532	2.894	3.711	4.055
8	0.543	1.187	1.612	2.504	2.863	3.668	4.006
9	0.539	1.179	1.599	2.486	2.839	3.640	3.975
10	0.535	1.171	1.589	2.467	2.821	3.610	3.943
11	0.532	1.163	1.581	2.451	2.805	3.587	3.917
12	0.531	1.160	1.574	2.442	2.792	3.573	3.901
13	0.529	1.156	1.569	2.433	2.781	3.558	3.885
14	0.527	1.152	1.564	2.424	2.772	3.544	3.869
15	0.525	1.148	1.559	2.419	2.763	3.535	3.860
16	0.524	1.146	1.556	2.412	2.756	3.526	3.848
17	0.523	1.143	1.553	2.405	2.750	3.517	3.837
18	0.522	1.141	1.549	2.400	2.744	3.506	3.828
19	0.521	1.139	1.547	2.396	2.740	3.501	3.823
20	0.520	1.138	1.545	2.394	2.734	3.497	3.817
30	0.515	1.127	1.529	2.370	2.706	3.460	3.776
40	0.512	1.120	1.521	2.355	2.691	3.436	3.749
60	0.510	1.117	1.513	2.346	2.677	3.423	3.735
120	0.507	1.110	1.506	2.330	2.662	3.399	3.708
∞	0.504	1.104	1.498	2.319	2.649	3.381	3.688

Table A34 (cont.). Values of the coefficient $\beta^{LN}_{\mu Z}(V_Y,n,P)$ for left-prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. Coefficient of variation of the lognormal population: V_Y = 0.20.

LEFT PREDICTION INTERVAL V_Y = 0.40

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.623	1.588	2.335	4.207	5.086	7.322	8.381
2	0.506	1.282	1.865	3.265	3.903	5.478	6.205
3	0.463	1.177	1.703	2.955	3.517	4.894	5.523
4	0.442	1.122	1.622	2.799	3.326	4.605	5.186
5	0.429	1.090	1.573	2.709	3.212	4.439	4.994
6	0.420	1.068	1.541	2.645	3.135	4.319	4.856
7	0.414	1.054	1.516	2.605	3.082	4.246	4.771
8	0.408	1.040	1.499	2.569	3.041	4.181	4.695
9	0.405	1.032	1.485	2.546	3.009	4.140	4.648
10	0.401	1.024	1.474	2.522	2.985	4.096	4.598
11	0.399	1.016	1.465	2.502	2.962	4.062	4.559
12	0.397	1.013	1.457	2.490	2.946	4.041	4.534
13	0.395	1.008	1.451	2.479	2.931	4.019	4.509
14	0.394	1.004	1.446	2.467	2.918	3.999	4.485
15	0.392	1.001	1.441	2.461	2.907	3.985	4.470
16	0.391	0.998	1.437	2.452	2.898	3.972	4.454
17	0.390	0.995	1.433	2.444	2.889	3.958	4.437
18	0.389	0.993	1.430	2.437	2.882	3.943	4.422
19	0.389	0.991	1.427	2.432	2.876	3.936	4.415
20	0.388	0.990	1.425	2.429	2.869	3.929	4.405
30	0.383	0.979	1.407	2.400	2.831	3.876	4.343
40	0.380	0.971	1.398	2.380	2.811	3.841	4.303
60	0.378	0.968	1.390	2.369	2.794	3.822	4.282
120	0.376	0.961	1.382	2.350	2.774	3.787	4.242
∞	0.373	0.955	1.374	2.335	2.756	3.761	4.211

Table A34 (cont.). Values of the coefficient $\beta^{LN}_{\mu Z}(V_Y,n,P)$ for left-prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. Coefficient of variation of the lognormal population: V_Y = 0.40.

LEFT PREDICTION INTERVAL V_Y = 0.05

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.440	1.442	1.445	1.454	1.457	1.465	1.468
2	1.239	1.238	1.240	1.244	1.245	1.249	1.250
3	1.163	1.164	1.164	1.167	1.168	1.171	1.171
4	1.126	1.124	1.125	1.127	1.128	1.130	1.130
5	1.102	1.101	1.101	1.103	1.103	1.105	1.106
6	1.085	1.085	1.085	1.086	1.086	1.087	1.088
7	1.074	1.074	1.073	1.075	1.075	1.076	1.077
8	1.064	1.064	1.064	1.066	1.065	1.066	1.067
9	1.058	1.058	1.057	1.059	1.058	1.060	1.060
10	1.052	1.052	1.052	1.053	1.053	1.053	1.054
11	1.047	1.046	1.047	1.047	1.048	1.048	1.048
12	1.044	1.044	1.043	1.044	1.044	1.045	1.045
13	1.041	1.040	1.040	1.041	1.041	1.041	1.042
14	1.038	1.037	1.037	1.037	1.038	1.038	1.038
15	1.034	1.035	1.035	1.035	1.035	1.036	1.036
16	1.033	1.033	1.033	1.033	1.033	1.034	1.034
17	1.031	1.030	1.031	1.031	1.031	1.032	1.032
18	1.030	1.029	1.029	1.029	1.029	1.029	1.030
19	1.028	1.027	1.028	1.028	1.028	1.028	1.029
20	1.027	1.026	1.026	1.027	1.026	1.027	1.027
30	1.017	1.018	1.017	1.018	1.018	1.019	1.019
40	1.013	1.012	1.013	1.013	1.013	1.013	1.013
60	1.009	1.010	1.008	1.010	1.009	1.010	1.010
120	1.005	1.004	1.005	1.004	1.004	1.004	1.004
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A35. Values of the penalty ratio $r_{\mu_Z}^{LN}(V_Y,n,P)$ for left-hand prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. V_Y is the coefficient of variation of the lognormal population. For other values of V_Y, see continuation of the table.

LEFT PREDICTION INTERVAL V_Y= 0.10

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.468	1.470	1.478	1.496	1.504	1.520	1.527
2	1.253	1.252	1.256	1.264	1.267	1.274	1.277
3	1.173	1.173	1.174	1.180	1.182	1.187	1.189
4	1.133	1.131	1.133	1.136	1.138	1.142	1.143
5	1.108	1.107	1.107	1.111	1.112	1.115	1.116
6	1.089	1.089	1.090	1.092	1.093	1.095	1.096
7	1.078	1.078	1.077	1.081	1.080	1.083	1.084
8	1.068	1.068	1.068	1.070	1.070	1.072	1.073
9	1.061	1.061	1.061	1.063	1.063	1.065	1.066
10	1.055	1.055	1.054	1.056	1.057	1.058	1.059
11	1.050	1.048	1.050	1.050	1.051	1.052	1.053
12	1.046	1.046	1.046	1.047	1.047	1.049	1.049
13	1.043	1.042	1.042	1.043	1.044	1.045	1.045
14	1.040	1.039	1.039	1.040	1.040	1.041	1.042
15	1.036	1.036	1.037	1.038	1.038	1.039	1.040
16	1.035	1.035	1.035	1.035	1.036	1.037	1.037
17	1.033	1.032	1.033	1.033	1.033	1.034	1.034
18	1.031	1.030	1.031	1.031	1.031	1.032	1.032
19	1.030	1.029	1.029	1.030	1.030	1.031	1.031
20	1.028	1.028	1.028	1.029	1.028	1.029	1.030
30	1.018	1.019	1.018	1.020	1.019	1.020	1.020
40	1.013	1.013	1.014	1.014	1.014	1.014	1.014
60	1.010	1.010	1.009	1.010	1.009	1.011	1.011
120	1.005	1.004	1.005	1.004	1.004	1.005	1.005
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A35 (cont.). Values of the penalty ratio $r_{\mu Z}^{LN}(V_Y,n,P)$ for left-prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. Coefficient of variation of the lognormal population: $V_Y = 0.10$.

LEFT PREDICTION INTERVAL V_Y = 0.20

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.528	1.531	1.547	1.589	1.606	1.645	1.661
2	1.284	1.281	1.288	1.306	1.313	1.330	1.337
3	1.193	1.192	1.196	1.207	1.212	1.223	1.227
4	1.148	1.145	1.149	1.156	1.160	1.168	1.170
5	1.120	1.118	1.120	1.127	1.129	1.136	1.138
6	1.100	1.099	1.100	1.105	1.107	1.112	1.114
7	1.087	1.086	1.086	1.092	1.093	1.098	1.099
8	1.076	1.075	1.076	1.080	1.081	1.085	1.086
9	1.068	1.068	1.067	1.072	1.072	1.077	1.078
10	1.061	1.060	1.061	1.064	1.065	1.068	1.069
11	1.055	1.053	1.055	1.057	1.059	1.061	1.062
12	1.052	1.051	1.051	1.053	1.054	1.057	1.058
13	1.048	1.047	1.047	1.049	1.050	1.052	1.053
14	1.044	1.043	1.044	1.045	1.046	1.048	1.049
15	1.041	1.040	1.041	1.043	1.043	1.046	1.047
16	1.039	1.038	1.038	1.040	1.041	1.043	1.044
17	1.037	1.035	1.036	1.037	1.038	1.040	1.041
18	1.035	1.033	1.034	1.035	1.036	1.037	1.038
19	1.033	1.031	1.032	1.034	1.034	1.036	1.037
20	1.031	1.030	1.031	1.032	1.032	1.034	1.035
30	1.020	1.021	1.020	1.022	1.022	1.024	1.024
40	1.015	1.014	1.015	1.016	1.016	1.016	1.017
60	1.011	1.011	1.010	1.012	1.011	1.012	1.013
120	1.006	1.005	1.005	1.005	1.005	1.005	1.006
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A35 (cont.). Values of the penalty ratio $r^{LN}_{\mu_Z}(V_Y,n,P)$ for left-prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. Coefficient of variation of the lognormal population: $V_Y = 0.20$.

LEFT PREDICTION INTERVAL V_Y= 0.40

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.670	1.662	1.700	1.802	1.845	1.947	1.990
2	1.356	1.342	1.358	1.398	1.416	1.457	1.473
3	1.242	1.232	1.240	1.265	1.276	1.301	1.312
4	1.185	1.175	1.181	1.199	1.207	1.225	1.231
5	1.150	1.142	1.146	1.160	1.165	1.180	1.186
6	1.124	1.118	1.122	1.133	1.137	1.148	1.153
7	1.108	1.103	1.104	1.116	1.118	1.129	1.133
8	1.094	1.089	1.092	1.100	1.103	1.112	1.115
9	1.085	1.081	1.081	1.090	1.092	1.101	1.104
10	1.076	1.072	1.073	1.080	1.083	1.089	1.092
11	1.069	1.064	1.067	1.072	1.075	1.080	1.082
12	1.064	1.060	1.061	1.067	1.069	1.075	1.077
13	1.059	1.056	1.056	1.062	1.063	1.069	1.071
14	1.055	1.051	1.053	1.057	1.059	1.063	1.065
15	1.050	1.048	1.049	1.054	1.055	1.060	1.062
16	1.048	1.045	1.046	1.050	1.051	1.056	1.058
17	1.046	1.042	1.044	1.047	1.048	1.052	1.053
18	1.043	1.040	1.041	1.044	1.045	1.048	1.050
19	1.041	1.037	1.039	1.042	1.044	1.047	1.048
20	1.039	1.036	1.037	1.040	1.041	1.045	1.046
30	1.025	1.025	1.024	1.028	1.027	1.031	1.031
40	1.018	1.017	1.018	1.019	1.020	1.021	1.022
60	1.014	1.014	1.012	1.014	1.014	1.016	1.017
120	1.007	1.006	1.006	1.006	1.006	1.007	1.007
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A35 (cont.). Values of the penalty ratio $r_{\mu Z}^{LN}(V_Y,n,P)$ for left-prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. Coefficient of variation of the lognormal population: $V_Y = 0.40$.

RIGHT PREDICTION INTERVAL V_Y = 0.05

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.909	1.672	2.111	2.907	3.189	3.755	3.971
2	0.793	1.459	1.846	2.548	2.799	3.302	3.495
3	0.749	1.381	1.746	2.413	2.651	3.131	3.315
4	0.727	1.338	1.694	2.342	2.574	3.040	3.219
5	0.714	1.314	1.662	2.299	2.526	2.986	3.162
6	0.703	1.296	1.640	2.268	2.493	2.946	3.119
7	0.697	1.284	1.624	2.249	2.470	2.921	3.093
8	0.692	1.274	1.612	2.231	2.452	2.898	3.069
9	0.688	1.268	1.603	2.219	2.437	2.884	3.054
10	0.684	1.260	1.595	2.208	2.427	2.868	3.038
11	0.681	1.254	1.589	2.197	2.417	2.856	3.025
12	0.680	1.252	1.584	2.192	2.409	2.849	3.017
13	0.678	1.248	1.579	2.186	2.403	2.841	3.009
14	0.676	1.244	1.576	2.180	2.397	2.834	3.001
15	0.674	1.242	1.572	2.176	2.392	2.829	2.996
16	0.673	1.240	1.570	2.172	2.388	2.824	2.991
17	0.672	1.237	1.567	2.168	2.383	2.819	2.985
18	0.671	1.235	1.564	2.165	2.380	2.813	2.980
19	0.670	1.234	1.563	2.162	2.378	2.811	2.978
20	0.669	1.233	1.561	2.160	2.374	2.808	2.975
30	0.664	1.224	1.549	2.145	2.357	2.789	2.954
40	0.661	1.218	1.543	2.135	2.348	2.776	2.940
60	0.659	1.215	1.536	2.129	2.339	2.769	2.933
120	0.657	1.209	1.531	2.119	2.330	2.756	2.919
∞	0.654	1.204	1.525	2.111	2.322	2.746	2.909

Table A36. Values of the coefficient $\beta^{LN}_{\mu_Z}(V_Y,n,P)$ for right-hand prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. V_Y is the coefficient of variation of the lognormal population. For other values of V_Y, see continuation of the table.

RIGHT PREDICTION INTERVAL V_Y= 0.10

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.866	1.542	1.919	2.577	2.802	3.242	3.406
2	0.761	1.357	1.693	2.284	2.489	2.891	3.042
3	0.721	1.288	1.607	2.172	2.368	2.756	2.901
4	0.701	1.251	1.562	2.113	2.305	2.684	2.825
5	0.688	1.229	1.534	2.077	2.266	2.640	2.780
6	0.679	1.213	1.515	2.051	2.238	2.608	2.747
7	0.673	1.203	1.501	2.035	2.219	2.588	2.726
8	0.668	1.193	1.491	2.020	2.204	2.569	2.707
9	0.665	1.188	1.483	2.010	2.192	2.558	2.695
10	0.661	1.181	1.476	2.000	2.183	2.545	2.682
11	0.659	1.176	1.471	1.992	2.175	2.535	2.671
12	0.657	1.174	1.466	1.987	2.169	2.530	2.665
13	0.656	1.170	1.462	1.982	2.163	2.523	2.659
14	0.654	1.167	1.459	1.977	2.159	2.517	2.652
15	0.652	1.165	1.456	1.974	2.154	2.513	2.648
16	0.651	1.163	1.454	1.971	2.151	2.509	2.644
17	0.650	1.161	1.452	1.967	2.148	2.505	2.639
18	0.650	1.159	1.449	1.964	2.145	2.501	2.635
19	0.649	1.158	1.448	1.962	2.143	2.499	2.634
20	0.648	1.157	1.446	1.961	2.140	2.497	2.631
30	0.643	1.149	1.435	1.948	2.125	2.481	2.614
40	0.640	1.144	1.430	1.939	2.118	2.470	2.603
60	0.639	1.141	1.425	1.934	2.111	2.464	2.597
120	0.636	1.136	1.420	1.926	2.103	2.454	2.586
∞	0.634	1.132	1.415	1.919	2.096	2.446	2.578

Table A36 (cont.). Values of the coefficient $\beta_{\mu Z}^{LN}(V_Y,n,P)$ for right-prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. Coefficient of variation of the lognormal population: $V_Y = 0.10$.

RIGHT PREDICTION INTERVAL V_Y = 0.20

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.787	1.319	1.595	2.045	2.189	2.456	2.551
2	0.700	1.177	1.431	1.850	1.987	2.245	2.337
3	0.667	1.124	1.367	1.773	1.907	2.160	2.251
4	0.650	1.095	1.333	1.732	1.864	2.114	2.203
5	0.640	1.077	1.312	1.707	1.837	2.086	2.175
6	0.632	1.065	1.298	1.689	1.819	2.065	2.154
7	0.627	1.057	1.287	1.677	1.806	2.052	2.141
8	0.623	1.050	1.280	1.667	1.795	2.040	2.128
9	0.620	1.045	1.273	1.660	1.787	2.032	2.121
10	0.617	1.040	1.268	1.653	1.781	2.024	2.112
11	0.615	1.036	1.264	1.647	1.775	2.018	2.106
12	0.614	1.034	1.261	1.643	1.771	2.014	2.102
13	0.612	1.032	1.258	1.640	1.767	2.010	2.098
14	0.611	1.029	1.255	1.636	1.764	2.006	2.094
15	0.609	1.027	1.253	1.634	1.761	2.003	2.091
16	0.609	1.026	1.251	1.632	1.758	2.000	2.088
17	0.608	1.024	1.250	1.629	1.756	1.998	2.085
18	0.607	1.023	1.248	1.627	1.754	1.995	2.083
19	0.607	1.022	1.247	1.626	1.753	1.994	2.081
20	0.606	1.021	1.246	1.625	1.751	1.992	2.080
30	0.602	1.015	1.237	1.616	1.741	1.982	2.069
40	0.599	1.010	1.233	1.610	1.735	1.975	2.062
60	0.598	1.008	1.229	1.606	1.730	1.971	2.058
120	0.596	1.004	1.226	1.600	1.725	1.964	2.051
∞	0.594	1.001	1.222	1.595	1.720	1.959	2.045

Table A36 (cont.). Values of the coefficient $\beta^{LN}_{\mu Z}(V_Y,n,P)$ for right-prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. Coefficient of variation of the lognormal population: $V_Y = 0.20$.

RIGHT PREDICTION INTERVAL V_Y= 0.40

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	0.654	0.986	1.138	1.353	1.414	1.516	1.548
2	0.595	0.903	1.049	1.264	1.327	1.436	1.472
3	0.571	0.870	1.013	1.227	1.291	1.402	1.438
4	0.560	0.853	0.994	1.207	1.271	1.382	1.420
5	0.552	0.842	0.982	1.195	1.258	1.370	1.408
6	0.547	0.834	0.974	1.186	1.249	1.361	1.399
7	0.543	0.829	0.968	1.180	1.243	1.356	1.394
8	0.540	0.825	0.963	1.175	1.238	1.351	1.389
9	0.539	0.822	0.960	1.171	1.234	1.347	1.385
10	0.537	0.819	0.957	1.167	1.231	1.344	1.382
11	0.535	0.816	0.954	1.164	1.228	1.341	1.379
12	0.534	0.815	0.952	1.162	1.226	1.339	1.377
13	0.533	0.813	0.951	1.161	1.224	1.337	1.376
14	0.532	0.812	0.949	1.159	1.223	1.336	1.374
15	0.531	0.811	0.948	1.158	1.221	1.334	1.373
16	0.531	0.810	0.947	1.157	1.220	1.333	1.372
17	0.530	0.809	0.946	1.155	1.219	1.332	1.370
18	0.530	0.808	0.945	1.154	1.218	1.331	1.369
19	0.529	0.807	0.944	1.153	1.217	1.330	1.369
20	0.528	0.807	0.943	1.153	1.216	1.330	1.368
30	0.525	0.803	0.939	1.148	1.211	1.325	1.363
40	0.524	0.800	0.936	1.145	1.209	1.322	1.360
60	0.523	0.799	0.934	1.143	1.206	1.320	1.359
120	0.521	0.796	0.932	1.140	1.204	1.317	1.355
∞	0.520	0.794	0.929	1.138	1.201	1.315	1.353

Table A36 (cont.). Values of the coefficient $\beta^{LN}_{\mu Z}(V_Y,n,P)$ for right-prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. Coefficient of variation of the lognormal population: $V_Y = 0.40$.

RIGHT PREDICTION INTERVAL V_Y= 0.05

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.389	1.388	1.384	1.377	1.374	1.368	1.365
2	1.213	1.212	1.210	1.207	1.205	1.203	1.201
3	1.146	1.147	1.145	1.143	1.142	1.140	1.140
4	1.112	1.111	1.111	1.109	1.109	1.107	1.107
5	1.091	1.091	1.090	1.089	1.088	1.087	1.087
6	1.076	1.076	1.076	1.074	1.074	1.073	1.072
7	1.066	1.066	1.065	1.065	1.064	1.064	1.063
8	1.058	1.058	1.057	1.057	1.056	1.056	1.055
9	1.052	1.052	1.051	1.051	1.050	1.050	1.050
10	1.046	1.047	1.046	1.046	1.045	1.045	1.044
11	1.042	1.041	1.042	1.041	1.041	1.040	1.040
12	1.039	1.039	1.038	1.038	1.038	1.038	1.037
13	1.037	1.036	1.036	1.035	1.035	1.035	1.035
14	1.034	1.033	1.033	1.032	1.032	1.032	1.032
15	1.031	1.031	1.031	1.031	1.030	1.030	1.030
16	1.030	1.030	1.029	1.029	1.028	1.028	1.028
17	1.028	1.027	1.028	1.027	1.027	1.027	1.026
18	1.027	1.026	1.026	1.025	1.025	1.025	1.025
19	1.025	1.024	1.025	1.024	1.024	1.024	1.024
20	1.024	1.024	1.024	1.023	1.023	1.023	1.023
30	1.015	1.016	1.015	1.016	1.015	1.016	1.016
40	1.011	1.011	1.011	1.011	1.011	1.011	1.011
60	1.008	1.009	1.007	1.008	1.008	1.008	1.008
120	1.004	1.004	1.004	1.004	1.004	1.004	1.004
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A37. Values of the penalty ratio $r^{LN}_{\mu_Z}(V_Y,n,P)$ for right-hand prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. V_Y is the coefficient of variation of the lognormal population. For other values of V_Y, see continuation of the table.

RIGHT PREDICTION INTERVAL $V_Y = 0.10$

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.367	1.363	1.357	1.342	1.337	1.326	1.321
2	1.201	1.199	1.197	1.190	1.187	1.182	1.180
3	1.138	1.138	1.136	1.132	1.130	1.127	1.125
4	1.106	1.105	1.104	1.101	1.100	1.097	1.096
5	1.086	1.086	1.084	1.082	1.081	1.079	1.078
6	1.072	1.072	1.071	1.069	1.068	1.066	1.065
7	1.062	1.063	1.061	1.060	1.059	1.058	1.057
8	1.055	1.055	1.054	1.053	1.052	1.051	1.050
9	1.049	1.050	1.048	1.047	1.046	1.046	1.045
10	1.044	1.044	1.043	1.042	1.042	1.041	1.040
11	1.040	1.039	1.039	1.038	1.038	1.037	1.036
12	1.037	1.037	1.036	1.035	1.035	1.034	1.034
13	1.035	1.034	1.033	1.033	1.032	1.032	1.031
14	1.032	1.032	1.031	1.030	1.030	1.029	1.029
15	1.029	1.029	1.029	1.029	1.028	1.028	1.027
16	1.028	1.028	1.028	1.027	1.026	1.026	1.026
17	1.027	1.026	1.026	1.025	1.025	1.024	1.024
18	1.025	1.025	1.024	1.023	1.023	1.022	1.022
19	1.024	1.023	1.023	1.022	1.022	1.022	1.022
20	1.023	1.022	1.022	1.022	1.021	1.021	1.021
30	1.015	1.015	1.015	1.015	1.014	1.014	1.014
40	1.011	1.011	1.011	1.010	1.010	1.010	1.010
60	1.008	1.008	1.007	1.008	1.007	1.008	1.008
120	1.004	1.004	1.004	1.003	1.003	1.003	1.003
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A37 (cont.). Values of the penalty ratio $r^{LN}_{\mu z}(V_Y,n,P)$ for right-prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. Coefficient of variation of the lognormal population: $V_Y = 0.10$.

RIGHT PREDICTION INTERVAL $V_Y = 0.20$

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.325	1.317	1.306	1.282	1.272	1.254	1.247
2	1.179	1.176	1.171	1.160	1.155	1.146	1.142
3	1.123	1.123	1.119	1.112	1.109	1.103	1.100
4	1.095	1.094	1.091	1.086	1.084	1.079	1.077
5	1.077	1.077	1.074	1.070	1.068	1.065	1.063
6	1.064	1.064	1.063	1.059	1.057	1.054	1.053
7	1.056	1.056	1.054	1.052	1.050	1.048	1.047
8	1.049	1.049	1.048	1.045	1.044	1.042	1.041
9	1.044	1.044	1.042	1.041	1.039	1.038	1.037
10	1.039	1.039	1.038	1.036	1.035	1.033	1.033
11	1.036	1.035	1.035	1.032	1.032	1.030	1.030
12	1.033	1.033	1.032	1.030	1.029	1.028	1.028
13	1.031	1.031	1.030	1.028	1.027	1.026	1.026
14	1.029	1.028	1.028	1.026	1.025	1.024	1.024
15	1.026	1.026	1.026	1.025	1.024	1.023	1.022
16	1.025	1.025	1.024	1.023	1.022	1.021	1.021
17	1.024	1.023	1.023	1.021	1.021	1.020	1.019
18	1.023	1.022	1.022	1.020	1.020	1.019	1.018
19	1.022	1.021	1.021	1.019	1.019	1.018	1.018
20	1.020	1.020	1.020	1.019	1.018	1.017	1.017
30	1.013	1.014	1.013	1.013	1.012	1.012	1.012
40	1.010	1.009	1.010	1.009	1.009	1.008	1.008
60	1.007	1.008	1.006	1.007	1.006	1.006	1.006
120	1.004	1.003	1.003	1.003	1.003	1.003	1.003
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A37 (cont.). Values of the penalty ratio $r^{LN}_{\mu_Z}(V_Y,n,P)$ for right-prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. Coefficient of variation of the lognormal population: $V_Y = 0.20$.

RIGHT PREDICTION INTERVAL V_Y = 0.40

n \ P	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990	0.9995
1	1.257	1.242	1.224	1.189	1.177	1.153	1.144
2	1.143	1.137	1.129	1.111	1.105	1.092	1.088
3	1.099	1.096	1.091	1.079	1.075	1.066	1.063
4	1.077	1.074	1.070	1.061	1.058	1.052	1.049
5	1.062	1.061	1.057	1.050	1.047	1.043	1.041
6	1.052	1.051	1.048	1.042	1.040	1.036	1.034
7	1.045	1.045	1.041	1.037	1.035	1.031	1.030
8	1.040	1.039	1.037	1.032	1.031	1.027	1.026
9	1.036	1.035	1.033	1.029	1.027	1.025	1.024
10	1.032	1.031	1.029	1.026	1.025	1.022	1.021
11	1.029	1.028	1.027	1.024	1.023	1.020	1.019
12	1.027	1.026	1.025	1.022	1.021	1.019	1.018
13	1.025	1.024	1.023	1.020	1.019	1.017	1.017
14	1.023	1.022	1.021	1.019	1.018	1.016	1.015
15	1.021	1.021	1.020	1.018	1.017	1.015	1.015
16	1.020	1.020	1.019	1.017	1.016	1.014	1.014
17	1.019	1.019	1.018	1.016	1.015	1.013	1.013
18	1.018	1.018	1.017	1.015	1.014	1.012	1.012
19	1.017	1.017	1.016	1.014	1.013	1.012	1.012
20	1.016	1.016	1.015	1.013	1.013	1.011	1.011
30	1.011	1.011	1.010	1.009	1.008	1.008	1.008
40	1.008	1.008	1.007	1.007	1.006	1.006	1.005
60	1.006	1.006	1.005	1.005	1.004	1.004	1.004
120	1.003	1.003	1.003	1.002	1.002	1.002	1.002
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table A37 (cont.). Values of the penalty ratio $r_{\mu_Z}^{LN}(V_Y,n,P)$ for right-prediction in lognormal sequences. The associated normal sequences have unknown mean and known variance. Coefficient of variation of the lognormal population: $V_Y = 0.40$.

APPENDIX B

The tables in this appendix contain simple-prediction limits and associated penalty ratios for k-variate independent normal sequences with $k > 1$. For $k = 1$ see Appendix A. The values refer to ellipsoidal prediction regions centered on the mean. n denotes the available sample size and P is the desired probability content of the region.

For using the tables in Bayesian prediction with "diffuse" or conjugate prior see Paragraph II.2.4(a).

k=2	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
3	6.325	16.248	32.619	163.291	326.598	1632.992
4	3.354	5.809	8.441	19.268	27.318	61.207
5	2.701	4.181	5.529	9.930	12.624	21.799
6	2.415	3.552	4.500	7.246	8.756	13.366
7	2.254	3.220	3.984	6.034	7.087	10.090
8	2.151	3.015	3.674	5.355	6.179	8.419
9	2.078	2.876	3.468	4.924	5.613	7.422
10	2.025	2.776	3.322	4.627	5.228	6.765
11	1.984	2.700	3.212	4.410	4.950	6.303
12	1.952	2.640	3.127	4.244	4.740	5.961
13	1.925	2.592	3.059	4.115	4.576	5.696
14	1.903	2.553	3.003	4.010	4.445	5.487
15	1.884	2.519	2.957	3.924	4.337	5.318
16	1.868	2.492	2.918	3.851	4.247	5.179
17	1.855	2.467	2.884	3.790	4.171	5.061
18	1.843	2.446	2.855	3.737	4.105	4.960
19	1.832	2.428	2.968	3.691	4.049	4.875
20	1.823	2.412	2.807	3.651	3.999	4.799
21	1.815	2.397	2.787	3.615	3.955	4.734
22	1.807	2.384	2.769	3 583	3.917	4.674
32	1.759	2.303	2.658	3.389	3.680	4.323
42	1.736	2.263	2.604	3.297	3.568	4.161
62	1.712	2.224	2.551	3.207	3.460	4.004
122	1.688	2.184	2.499	3.120	3.356	3.858
∞	1.665	2.146	2.448	3.035	3.255	3.718

Table B1. Values of $\beta^{N}_{\underline{\mu},\underline{\Sigma}}(k,n,P)$ such that the prediction ellipsoid (II.92) has probability content P. Sampling is from a k-variate normal population with unknown mean vector $\underline{\mu}$ and covariance matrix $\underline{\Sigma}$. (For the case k=1 see Table A8 in Appendix A.)

k=3	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
4	9.605	24.554	49.262	246.550	493.179	2465.665
5	4.765	8.122	11.747	26.721	37.869	84.819
6	3.707	5.608	7.356	13.108	16.640	28.689
7	3.245	4.643	5.822	9.266	11.170	16.998
8	2.984	4.136	5.056	7.549	8.838	12.525
9	2.816	3.823	4.598	6.593	7.577	10.263
10	2.699	3.611	4.295	5.988	6.795	8.924
11	2.612	3.459	4.078	5.573	6.266	8.047
12	2.546	3.343	3.917	5.270	5.884	7.431
13	2.493	3.252	3.792	5.040	5.597	6.975
14	2.450	3.179	3.691	4.860	5.373	6.627
15	2.414	3.119	3.610	4.714	5.194	6.350
16	2.384	3.068	3.542	4.594	5.047	6.128
17	2.358	3.026	3.484	4.494	4.924	5.943
18	2.336	2.989	3.435	4.409	4.821	5.790
19	2.316	2.957	3.392	4.336	4.732	5.655
20	2.300	2.929	3.355	4.272	4.655	5.544
21	2.284	2.905	3.322	4.217	4.588	5.445
22	2.270	2.883	3.292	4.167	4.529	5.357
23	2.258	2.863	3.266	4.124	4.476	5.281
33	2.181	2.739	3.104	3.856	4.156	4.821
43	2.142	2.679	3.025	3.729	4.005	4.612
63	2.104	2.618	2.947	3.605	3.859	4.408
123	2.065	2.559	2.871	3.485	3.719	4.219
∞	2.027	2.500	2.796	3.368	3.583	4.032

Table B1 (cont.).

k=4	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
5	12.836	32.741	65.665	328.622	657.267	3286.329
6	6.141	10.384	14.985	34.028	48.214	107.969
7	4.675	6.989	9.130	16.202	20.551	35.405
8	4.032	5.687	7.093	11.217	13.504	20.514
9	3.669	5.003	6.076	9.001	10.518	14.869
10	3.434	4.582	5.470	7.770	8.910	12.028
11	3.271	4.296	5.068	6.994	7.915	10.352
12	3.149	4.089	4.782	6.461	7.243	9.260
13	3.055	3.933	4.568	6.073	6.760	8.493
14	2.981	3.810	4.402	5.779	6.396	7.928
15	2.920	3.711	4.270	5.548	6.113	7.497
16	2.870	3.630	4.161	5.362	5.886	7.153
17	2.828	3.562	4.071	5.209	5.700	6.876
18	2.791	3.504	3.994	5.081	5.545	6.648
19	2.759	3.454	3.929	4.972	5.415	6.456
20	2.732	3.411	3.873	4.879	5.303	6.293
21	2.708	3.373	3.823	4.798	5.206	6.153
22	2.686	3.340	3.780	4.727	5.121	6.033
23	2.667	3.309	3.741	4.664	5.046	5.923
24	2.649	3.283	3.706	4.608	4.979	5.833
34	2.540	3.115	3.491	4.266	4.576	5.265
44	2.486	3.032	3.385	4.103	4.386	5.007
64	2.431	2.951	3.282	3.945	4.202	4.759
124	2.375	2.869	3.180	3.792	4.026	4.523
∞	2.320	2.789	3.080	3.644	3.855	4.299

Table B1 (cont.).

k=5	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
6	16.039	40.860	81.933	410.009	820.111	4100.184
7	7.499	12.622	18.188	41.259	58.451	130.885
8	5.624	8.347	10.877	19.251	24.408	42.031
9	4.798	6.709	8.337	13.133	15.796	23.970
10	4.331	5.847	7.071	10.420	12.162	17.162
11	4.028	5.315	6.315	8.917	10.209	13.754
12	3.816	4.954	5.815	7.969	9.003	11.746
13	3.659	4.693	5.458	7.319	8.189	10.438
14	3.537	4.495	5.191	6.846	7.603	9.519
15	3.440	4.339	4.983	6.487	7.163	8.846
16	3.362	4.214	4.818	6.206	6.821	8.331
17	3.296	4.111	4.682	5.979	6.546	7.922
18	3.240	4.025	4.569	5.793	6.322	7.591
19	3.193	3.951	4.474	5.637	6.135	7.321
20	3.152	3.888	4.392	5.504	5.977	7.095
21	3.116	3.833	4.321	5.390	5.842	6.899
22	3.084	3.784	4.260	5.291	5.725	6.733
23	3.055	3.742	4.205	5.205	5.622	6.590
24	3.031	3.704	4.156	5.128	5.531	6.460
25	3.008	3.670	4.113	5.060	5.451	6.349
35	2.864	3.456	3.843	4.643	4.964	5.677
45	2.791	3.351	3.711	4.445	4.734	5.370
65	2.720	3.246	3.581	4.252	4.512	5.077
125	2.647	3.142	3.453	4.066	4.299	4.798
∞	2.574	3.039	3.327	3.884	4.093	4.528

Table B1 (cont.).

k=7	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
8	22.400	56.984	114.245	571.662	1143.488	5716.957
9	10.186	17.055	24.538	55.597	78.755	176.330
10	7.492	11.029	14.328	25.283	32.038	55.136
11	6.300	8.716	10.786	16.909	20.317	30.790
12	5.621	7.496	9.019	13.208	15.392	21.675
13	5.179	6.742	7.964	11.160	12.752	17.129
14	4.867	6.228	7.263	9.869	11.125	14.464
15	4.636	5.856	6.764	8.985	10.027	12.729
16	4.456	5.572	6.389	8.341	9.238	11.517
17	4.314	5.350	6.098	7.853	8.646	10.625
18	4.196	5.171	5.865	7.470	8.184	9.944
19	4.099	5.023	5.674	7.161	7.814	9.403
20	4.015	4.899	5.516	6.907	7.512	8.970
21	3.944	4.793	5.381	6.695	7.260	8.612
22	3.883	4.702	5.266	6.514	7.047	8.310
23	3.829	4.623	5.166	6.359	6.865	8.055
24	3.782	4.554	5.078	6.224	6.706	7.833
25	3.740	4.491	5.001	6.106	6.569	7.644
26	3.702	4.437	4.933	6.001	6.447	7.480
27	3.667	4.388	4.871	5.908	6.338	7.328
37	3.450	4.077	4.487	5.340	5.682	6.448
47	3.340	3.924	4.300	5.068	5.371	6.042
67	3.230	3.770	4.115	4.804	5.071	5.654
127	3.119	3.618	3.932	4.548	4.782	5.285
∞	3.006	3.467	3.751	4.298	4.503	4.928

Table B1 (cont.).

k=10	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
11	31.886	81.035	162.441	812.793	1625.479	8128.035
12	14.185	23.656	33.995	76.958	109.000	244.024
13	10.263	15.010	19.454	34.248	43.380	74.602
14	8.515	11.683	14.411	22.506	27.020	40.904
15	7.513	9.923	11.892	17.326	20.167	28.355
16	6.859	8.833	10.385	14.462	16.500	22.114
17	6.395	8.088	9.382	12.658	14.241	18.460
18	6.048	7.545	8.665	11.420	12.718	16.089
19	5.778	7.132	8.127	10.520	11.623	14.429
20	5.563	6.808	7.708	9.836	10.800	13.212
21	5.386	6.544	7.373	9.298	10.159	12.282
22	5.239	6.327	7.097	8.866	9.646	11.549
23	5.112	6.144	6.868	8.509	9.226	10.958
24	5.004	5.988	6.673	8.210	8.875	10.465
25	4.910	5.853	6.506	7.957	8.580	10.058
26	4.829	5.736	6.361	7.739	8.326	9.709
27	4.756	5.634	6.234	7.549	8.105	9.408
28	4.690	5.542	6.121	7.382	7.914	9.151
29	4.631	5.461	6.021	7.235	7.743	8.921
30	4.580	5.387	5.931	7.104	7.592	8.724
40	4.243	4.925	5.371	6.300	6.675	7.517
50	4.072	4.693	5.094	5.915	6.241	6.954
70	3.899	4.462	4.822	5.541	5.820	6.426
130	3.722	4.230	4.550	5.175	5.413	5.924
∞	3.543	3.999	4.279	4.818	5.019	5.441

Table B1 (cont.).

k=15	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
16	47.638	120.977	242.482	1213.251	2426.543	12133.211
17	20.815	34.605	49.685	112.400	159.183	356.346
18	14.841	21.600	27.944	49.102	62.175	106.914
19	12.166	16.580	20.402	31.762	38.107	57.640
20	10.622	13.921	16.627	24.122	28.050	39.379
21	9.607	12.263	14.362	19.898	22.673	30.328
22	8.885	11.128	12.852	17.235	19.361	25.033
23	8.343	10.298	11.769	15.407	17.126	21.601
24	7.920	9.666	10.956	14.076	15.521	19.208
25	7.578	9.166	10.321	13.063	14.312	17.447
26	7.297	8.759	9.811	12.268	13.370	16.098
27	7.063	8.423	9.392	11.625	12.614	15.038
28	6.861	8.139	9.041	11.095	11.996	14.178
29	6.687	7.898	8.743	10.652	11.481	13.474
30	6.537	7.687	8.488	10.273	11.044	12.885
31	6.405	7.505	8.263	9.948	10.669	12.369
32	6.286	7.344	8.069	9.666	10.344	11.935
33	6.182	7.200	7.896	9.416	10.059	11.567
34	6.088	7.072	7.740	9.196	9.808	11.227
35	6.003	6.956	7.601	9.000	9.584	10.936
45	5.450	6.223	6.732	7.792	8.222	9.183
55	5.167	5.854	6.300	7.211	7.572	8.373
75	4.877	5.482	5.867	6.640	6.942	7.599
135	4.580	5.106	5.436	6.082	6.328	6.849
∞	4.271	4.723	4.999	5.530	5.728	6.136

Table B1 (cont.).

k=20	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
21	63.363	160.848	322.378	1612.990	3226.309	16130.281
22	27.426	45.527	65.340	147.762	209.257	468.416
23	19.403	28.170	36.405	63.912	80.913	139.085
24	15.796	21.459	26.366	40.981	49.151	74.313
25	13.708	17.894	21.332	30.883	35.892	50.348
26	12.331	15.666	18.310	25.299	28.807	38.491
27	11.346	14.139	16.291	21.775	24.441	31.561
28	10.606	13.020	14.840	19.356	21.493	27.068
29	10.024	12.162	13.749	17.592	19.375	23.935
30	9.557	11.483	12.894	16.248	17.779	21.621
31	9.169	10.933	12.206	15.192	16.534	19.867
32	8.844	10.477	11.642	14.337	15.536	18.466
33	8.567	10.090	11.168	13.634	14.717	17.343
34	8.325	9.760	10.765	13.042	14.035	16.426
35	8.117	9.472	10.419	12.539	13.456	15.646
36	7.931	9.221	10.116	12.105	12.958	14.979
37	7.767	9.000	9.849	11.728	12.526	14.419
38	7.620	8.804	9.615	11.394	12.149	13.916
39	7.489	8.627	9.405	11.100	11.814	13.481
40	7.368	8.468	9.214	10.837	11.517	13.096
50	6.589	7.453	8.023	9.216	9.699	10.784
60	6.186	6.938	7.427	8.429	8.827	9.720
80	5.768	6.416	6.827	7.655	7.978	8.686
140	5.335	5.880	6.222	6.891	7.145	7.683
∞	4.881	5.331	5.605	6.129	6.325	6.738

Table B1 (cont.).

k=40	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
41	126.180	320.135	641.580	3210.006	6420.395	32100.687
42	53.826	89.146	127.853	288.982	409.218	916.027
43	37.598	54.376	70.175	123.021	155.703	267.634
44	30.259	40.901	50.142	77.747	93.201	140.817
45	25.981	33.704	40.078	57.820	67.146	94.083
46	23.147	29.197	34.012	46.792	53.222	70.987
47	21.104	26.087	29.944	39.825	44.640	57.532
48	19.561	23.799	27.019	35.033	38.839	48.783
49	18.339	22.043	24.803	31.531	34.661	42.685
50	17.352	20.645	23.065	28.856	31.509	38.202
51	16.530	19.504	21.661	26.750	29.046	34.766
52	15.834	18.550	20.502	25.041	27.066	32.054
53	15.241	17.747	19.527	23.629	25.440	29.862
54	14.726	17.051	18.695	22.444	24.081	28.046
55	14.271	16.447	17.976	21.429	22.927	26.529
56	13.867	15.920	17.351	20.552	21.932	25.207
57	13.508	15.453	16.796	19.790	21.070	24.095
58	13.189	15.033	16.304	19.113	20.309	23.124
59	12.895	14.657	15.861	18.516	19.639	22.259
60	12.632	14.314	15.467	17.981	19.040	21.519
70	10.886	12.115	12.931	14.647	15.347	16.926
80	9.959	10.975	11.637	13.004	13.552	14.777
100	8.974	9.787	10.308	11.360	11.772	12.675
160	7.916	8.541	8.929	9.697	9.989	10.608
∞	6.753	7.197	7.467	7.980	8.171	8.579

Table B1 (cont.).

k=2	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
2	5.477	14.071	28.249	141.414	282.843	1414.213
3	3.000	5.196	7.550	17.234	24.434	54.745
4	2.466	3.816	5.047	9.065	11.524	19.900
5	2.236	3.288	4.167	6.708	8.106	12.374
6	2.108	3.012	3.726	5.644	6.630	9.439
7	2.028	2.843	3.464	5.049	5.825	7.937
8	1.972	2.728	3.291	4.671	5.325	7.041
9	1.931	2.647	3.167	4.411	4.984	6.450
10	1.900	2.585	3.076	4.222	4.739	6.035
11	1.875	2.537	3.004	4.078	4.554	5.727
12	1.855	2.498	2.948	3.965	4.410	5.489
13	1.838	2.466	2.901	3.874	4.294	5.301
14	1.824	2.439	2.863	3.799	4.199	5.149
15	1.812	2.417	2.831	3.736	4.120	5.024
16	1.803	2.398	2.803	3.683	4.053	4.919
17	1.794	2.381	2.779	3.637	3.996	4.828
18	1.786	2.367	2.893	3.598	3.946	4.751
19	1.779	2.354	2.740	3.563	3.903	4.683
20	1.773	2.342	2.723	3.532	3.865	4.625
21	1.767	2.332	2.708	3.505	3.830	4.571
31	1.732	2.268	2.618	3.338	3.624	4.257
41	1.716	2.237	2.574	3.258	3.526	4.112
61	1.699	2.206	2.531	3.181	3.433	3.972
121	1.681	2.176	2.489	3.107	3.342	3.842
∞	1.665	2.146	2.448	3.035	3.255	3.718

Table B2. Values of $\beta_{\underline{\Sigma}}^{N}(k,n,P)$ such that the prediction ellipsoid (II.94) has probability content P. Sampling is from a k-variate normal population with known mean vector and unknown covariance matrix. (For the case k=1 see Table A7 in Appendix A.)

k=3	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
3	8.591	21.962	44.061	220.521	441.112	2205.357
4	4.349	7.414	10.723	24.393	34.569	77.429
5	3.432	5.192	6.811	12.136	15.406	26.561
6	3.035	4.343	5.446	8.667	10.448	15.900
7	2.813	3.899	4.767	7.117	8.332	11.808
8	2.671	3.627	4.362	6.255	7.188	9.737
9	2.573	3.443	4.095	5.709	6.479	8.509
10	2.501	3.311	3.905	5.335	5.999	7.705
11	2.446	3.212	3.764	5.063	5.654	7.139
12	2.402	3.134	3.654	4.857	5.394	6.722
13	2.367	3.071	3.566	4.695	5.191	6.402
14	2.337	3.020	3.495	4.565	5.029	6.148
15	2.313	2.977	3.436	4.457	4.896	5.945
16	2.292	2.941	3.386	4.368	4.786	5.776
17	2.273	2.910	3.343	4.292	4.692	5.635
18	2.257	2.883	3.306	4.226	4.612	5.511
19	2.244	2.859	3.274	4.170	4.543	5.410
20	2.232	2.838	3.246	4.120	4.483	5.320
21	2.220	2.819	3.220	4.076	4.429	5.240
22	2.211	2.802	3.197	4.037	4.382	5.170
32	2.149	2.699	3.058	3.799	4.094	4.750
42	2.118	2.648	2.990	3.686	3.959	4.560
62	2.088	2.598	2.924	3.576	3.829	4.373
122	2.057	2.549	2.859	3.471	3.703	4.202
∞	2.027	2.500	2.796	3.368	3.583	4.032

Table B2 (cont.).

k=4	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
4	11.717	29.889	59.944	299.989	600.000	2999.986
5	5.685	9.614	13.873	31.504	44.637	99.960
6	4.373	6.538	8.540	15.155	19.224	33.118
7	3.801	5.362	6.687	10.575	12.731	19.341
8	3.481	4.746	5.764	8.539	9.978	14.106
9	3.274	4.369	5.216	7.409	8.495	11.468
10	3.131	4.113	4.852	6.696	7.578	9.911
11	3.025	3.928	4.594	6.207	6.959	8.896
12	2.944	3.790	4.402	5.852	6.514	8.185
13	2.880	3.680	4.253	5.583	6.179	7.659
14	2.827	3.593	4.134	5.372	5.919	7.259
15	2.784	3.521	4.037	5.202	5.710	6.939
16	2.748	3.462	3.956	5.062	5.540	6.682
17	2.716	3.411	3.888	4.945	5.398	6.471
18	2.690	3.366	3.830	4.846	5.278	6.293
19	2.667	3.329	3.779	4.761	5.175	6.141
20	2.645	3.296	3.735	4.687	5.086	6.012
21	2.627	3.266	3.696	4.623	5.008	5.900
22	2.611	3.240	3.662	4.565	4.940	5.799
23	2.596	3.216	3.631	4.515	4.879	5.715
33	2.503	3.070	3.440	4.205	4.510	5.189
43	2.458	2.999	3.348	4.057	4.337	4.951
63	2.412	2.928	3.257	3.915	4.170	4.722
123	2.366	2.858	3.167	3.777	4.009	4.505
∞	2.320	2.789	3.080	3.644	3.855	4.299

Table B2 (cont.).

k=5	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
5	14.849	37.829	75.855	379.595	759.276	3796.043
6	7.014	11.807	17.013	38.594	54.676	122.432
7	5.303	7.870	10.255	18.150	23.012	39.627
8	4.552	6.365	7.909	12.459	14.985	22.740
9	4.130	5.575	6.742	9.935	11.596	16.363
10	3.857	5.089	6.046	8.537	9.774	13.169
11	3.667	4.759	5.586	7.656	8.650	11.286
12	3.526	4.522	5.259	7.053	7.891	10.059
13	3.417	4.342	5.015	6.614	7.346	9.196
14	3.331	4.202	4.825	6.281	6.936	8.565
15	3.261	4.088	4.674	6.020	6.617	8.082
16	3.203	3.995	4.550	5.810	6.362	7.698
17	3.154	3.917	4.447	5.638	6.153	7.389
18	3.113	3.851	4.361	5.494	5.980	7.135
19	3.076	3.794	4.286	5.372	5.833	6.924
20	3.044	3.745	4.222	5.266	5.707	6.741
21	3.016	3.701	4.166	5.175	5.599	6.585
22	2.991	3.663	4.117	5.095	5.503	6.451
23	2.970	3.629	4.072	5.024	5.420	6.330
24	2.950	3.598	4.033	4.962	5.345	6.226
34	2.824	3.407	3.789	4.578	4.895	5.598
44	2.761	3.314	3.671	4.396	4.682	5.312
64	2.699	3.222	3.554	4.220	4.478	5.039
124	2.636	3.130	3.440	4.050	4.282	4.779
∞	2.574	3.039	3.327	3.884	4.093	4.528

Table B2 (cont.).

k=7	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
7	21.119	53.725	107.711	538.968	1078.091	5390.000
8	9.663	16.179	23.278	52.744	74.713	167.282
9	7.144	10.516	13.661	24.106	30.547	52.570
10	6.032	8.345	10.327	16.189	19.452	29.480
11	5.401	7.202	8.665	12.689	14.788	20.825
12	4.991	6.497	7.674	10.754	12.288	16.506
13	4.702	6.017	7.016	9.535	10.747	13.974
14	4.489	5.670	6.549	8.699	9.708	12.325
15	4.323	5.406	6.198	8.092	8.962	11.173
16	4.192	5.200	5.926	7.632	8.402	10.326
17	4.084	5.034	5.708	7.270	7.965	9.679
18	3.995	4.896	5.531	6.980	7.617	9.165
19	3.919	4.781	5.383	6.741	7.331	8.754
20	3.854	4.683	5.257	6.541	7.093	8.414
21	3.798	4.599	5.151	6.371	6.892	8.127
22	3.749	4.526	5.057	6.225	6.720	7.885
23	3.706	4.462	4.976	6.098	6.571	7.675
24	3.667	4.404	4.904	5.987	6.441	7.496
25	3.633	4.354	4.841	5.889	6.326	7.340
26	3.601	4.309	4.783	5.802	6.224	7.196
36	3.405	4.023	4.428	5.269	5.606	6.363
46	3.305	3.883	4.255	5.015	5.315	5.978
66	3.206	3.742	4.085	4.768	5.034	5.612
126	3.107	3.604	3.917	4.530	4.763	5.264
∞	3.006	3.467	3.751	4.298	4.503	4.928

Table B2 (cont.).

k=10	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
10	30.529	77.585	155.525	778.190	1556.277	7782.023
11	13.628	22.728	32.662	73.939	104.723	234.450
12	9.889	14.464	18.747	33.002	41.802	71.889
13	8.226	11.287	13.922	21.743	26.104	39.517
14	7.275	9.608	11.514	16.776	19.527	27.455
15	6.654	8.569	10.075	14.030	16.008	21.453
16	6.215	7.860	9.118	12.301	13.840	17.940
17	5.887	7.344	8.433	11.115	12.379	15.660
18	5.632	6.951	7.921	10.254	11.329	14.064
19	5.429	6.644	7.522	9.598	10.540	12.894
20	5.262	6.393	7.204	9.084	9.925	12.000
21	5.123	6.188	6.941	8.671	9.434	11.295
22	5.005	6.015	6.723	8.330	9.032	10.727
23	4.903	5.867	6.538	8.044	8.696	10.254
24	4.815	5.740	6.380	7.803	8.413	9.863
25	4.738	5.629	6.242	7.594	8.170	9.528
26	4.670	5.532	6.121	7.413	7.959	9.238
27	4.609	5.446	6.015	7.254	7.776	8.992
28	4.554	5.369	5.920	7.114	7.613	8.771
29	4.506	5.300	5.835	6.988	7.469	8.583
39	4.191	4.864	5.305	6.223	6.593	7.424
49	4.032	4.647	5.044	5.857	6.179	6.885
69	3.871	4.431	4.787	5.502	5.779	6.380
129	3.708	4.214	4.532	5.155	5.392	5.902
∞	3.543	3.999	4.279	4.818	5.019	5.441

Table B2 (cont.).

k=15	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
15	46.216	117.365	235.242	1177.027	2354.092	11770.863
16	20.229	33.630	48.285	109.233	154.698	346.306
17	14.446	21.024	27.198	47.792	60.516	104.062
18	11.858	16.160	19.885	30.957	37.142	56.181
19	10.366	13.586	16.226	23.540	27.374	38.430
20	9.386	11.981	14.032	19.441	22.152	29.631
21	8.690	10.883	12.570	16.856	18.936	24.483
22	8.167	10.082	11.521	15.083	16.765	21.146
23	7.760	9.471	10.735	13.792	15.207	18.820
24	7.431	8.988	10.120	12.810	14.034	17.108
25	7.160	8.595	9.628	12.038	13.120	15.797
26	6.935	8.271	9.222	11.416	12.387	14.767
27	6.742	7.997	8.883	10.902	11.788	13.932
28	6.575	7.765	8.596	10.473	11.288	13.248
29	6.431	7.562	8.350	10.106	10.864	12.675
30	6.304	7.387	8.133	9.792	10.501	12.175
31	6.190	7.232	7.945	9.518	10.186	11.753
32	6.090	7.094	7.779	9.276	9.910	11.396
33	6.000	6.971	7.629	9.063	9.667	11.066
34	5.919	6.859	7.495	8.874	9.450	10.783
44	5.391	6.155	6.658	7.707	8.132	9.083
54	5.121	5.801	6.243	7.146	7.504	8.298
74	4.845	5.446	5.828	6.596	6.897	7.549
134	4.563	5.087	5.416	6.059	6.305	6.824
∞	4.271	4.723	4.999	5.530	5.728	6.136

Table B2 (cont.).

k=20	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
20	61.906	157.150	314.967	1575.907	3152.125	15759.391
21	26.823	44.526	63.904	144.514	204.657	458.120
22	18.995	27.577	35.639	62.566	79.209	136.157
23	15.477	21.025	25.833	40.153	48.158	72.811
24	13.441	17.546	20.918	30.283	35.195	49.370
25	12.100	15.373	17.968	24.826	28.268	37.771
26	11.141	13.884	15.997	21.383	24.000	30.992
27	10.421	12.794	14.582	19.019	21.120	26.597
28	9.855	11.958	13.518	17.296	19.049	23.532
29	9.402	11.296	12.684	15.984	17.490	21.270
30	9.024	10.761	12.014	14.953	16.273	19.554
31	8.709	10.317	11.465	14.118	15.299	18.184
32	8.440	9.940	11.003	13.432	14.499	17.086
33	8.205	9.620	10.610	12.854	13.833	16.190
34	8.003	9.339	10.273	12.364	13.268	15.427
35	7.823	9.096	9.979	11.941	12.781	14.775
36	7.664	8.880	9.718	11.572	12.360	14.228
37	7.521	8.690	9.491	11.247	11.992	13.737
38	7.395	8.518	9.287	10.960	11.665	13.312
39	7.277	8.365	9.101	10.704	11.375	12.935
49	6.524	7.379	7.944	9.125	9.603	10.677
59	6.135	6.881	7.365	8.360	8.754	9.640
79	5.733	6.376	6.785	7.608	7.928	8.633
139	5.316	5.859	6.199	6.866	7.120	7.656
∞	4.881	5.331	5.605	6.129	6.325	6.738

Table B2 (cont.).

k=40	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
40	124.669	316.301	633.896	3171.567	6343.465	31716.055
41	53.196	88.103	126.357	285.602	404.432	905.312
42	37.169	53.755	69.373	121.615	153.924	264.575
43	29.921	40.444	49.581	76.879	92.159	139.243
44	25.697	33.336	39.640	57.188	66.412	93.055
45	22.900	28.884	33.648	46.291	52.653	70.228
46	20.883	25.814	29.630	39.408	44.172	56.930
47	19.360	23.555	26.741	34.674	38.441	48.283
48	18.155	21.821	24.554	31.214	34.313	42.256
49	17.181	20.442	22.838	28.572	31.198	37.826
50	16.371	19.316	21.452	26.492	28.766	34.430
51	15.684	18.374	20.308	24.804	26.810	31.751
52	15.100	17.582	19.345	23.409	25.203	29.584
53	14.591	16.895	18.524	22.239	23.861	27.790
54	14.143	16.300	17.815	21.237	22.721	26.291
55	13.745	15.780	17.198	20.371	21.739	24.985
56	13.391	15.319	16.650	19.618	20.887	23.886
57	13.077	14.905	16.165	18.950	20.136	22.927
58	12.787	14.534	15.728	18.361	19.475	22.073
59	12.528	14.197	15.339	17.833	18.884	21.342
69	10.809	12.030	12.840	14.543	15.238	16.806
79	9.897	10.908	11.565	12.923	13.468	14.686
99	8.929	9.739	10.257	11.304	11.714	12.612
159	7.891	8.515	8.901	9.666	9.958	10.575
∞	6.753	7.197	7.467	7.980	8.171	8.579

Table B2 (cont.).

k=2	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
1	2.355	3.035	3.462	4.292	4.603	5.257
2	2.039	2.628	2.998	3.717	3.987	4.553
3	1.922	2.478	2.827	3.504	3.759	4.293
4	1.861	2.399	2.737	3.393	3.639	4.156
5	1.824	2.351	2.681	3.324	3.566	4.072
6	1.798	2.318	2.644	3.278	3.516	4.015
7	1.780	2.294	2.617	3.244	3.480	3.974
8	1.766	2.276	2.596	3.219	3.453	3.943
9	1.755	2.262	2.580	3.199	3.431	3.919
10	1.746	2.251	2.567	3.183	3.414	3.899
11	1.739	2.242	2.557	3.170	3.400	3.883
12	1.733	2.234	2.548	3.159	3.388	3.869
13	1.728	2.227	2.540	3.149	3.378	3.858
14	1.723	2.221	2.534	3.141	3.369	3.848
15	1.720	2.217	2.528	3.134	3.362	3.839
16	1.716	2.212	2.523	3.128	3.355	3.832
17	1.713	2.208	2.519	3.123	3.350	3.825
18	1.711	2.205	2.515	3.118	3.344	3.819
19	1.708	2.202	2.511	3.114	3.340	3.814
20	1.706	2.199	2.508	3.110	3.336	3.809
30	1.692	2.182	2.488	3.085	3.309	3.779
40	1.686	2.173	2.478	3.072	3.296	3.764
60	1.679	2.164	2.468	3.060	3.282	3.748
120	1.672	2.155	2.458	3.047	3.269	3.733
∞	1.665	2.146	2.448	3.035	3.255	3.718

Table B3. Values of $\beta_{\underline{\mu}}^{N}(k,n,P)$ such that the prediction ellipsoid (II.96) has probability content P. Sampling is from a k-variate normal population with unknown mean vector and known covariance matrix. (For the case k=1 see Table A6 in Appendix A.)

k=3	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
1	2.866	3.536	3.953	4.764	5.067	5.703
2	2.482	3.062	3.424	4.125	4.388	4.939
3	2.340	2.887	3.228	3.889	4.137	4.656
4	2.266	2.796	3.125	3.766	4.006	4.508
5	2.220	2.739	3.062	3.690	3.925	4.417
6	2.189	2.701	3.020	3.638	3.870	4.355
7	2.166	2.673	2.989	3.601	3.830	4.311
8	2.150	2.652	2.965	3.573	3.800	4.277
9	2.136	2.636	2.947	3.551	3.777	4.250
10	2.125	2.622	2.932	3.533	3.758	4.229
11	2.117	2.612	2.920	3.518	3.742	4.212
12	2.109	2.602	2.910	3.506	3.729	4.197
13	2.103	2.595	2.901	3.496	3.718	4.185
14	2.098	2.588	2.894	3.487	3.709	4.174
15	2.093	2.582	2.887	3.479	3.700	4.165
16	2.089	2.577	2.882	3.472	3.693	4.156
17	2.085	2.573	2.877	3.466	3.687	4.149
18	2.082	2.569	2.872	3.461	3.681	4.143
19	2.079	2.565	2.868	3.456	3.676	4.137
20	2.077	2.562	2.865	3.452	3.671	4.132
30	2.060	2.542	2.842	3.424	3.642	4.099
40	2.052	2.531	2.830	3.410	3.627	4.082
60	2.043	2.521	2.819	3.396	3.613	4.066
120	2.035	2.511	2.807	3.382	3.598	4.049
∞	2.027	2.500	2.796	3.368	3.583	4.032

Table B3 (cont.).

k=4	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
1	3.281	3.945	4.356	5.153	5.452	6.079
2	2.842	3.416	3.773	4.463	4.721	5.265
3	2.679	3.221	3.557	4.207	4.451	4.964
4	2.594	3.118	3.444	4.074	4.310	4.806
5	2.542	3.055	3.374	3.991	4.223	4.709
6	2.506	3.013	3.327	3.936	4.164	4.643
7	2.481	2.982	3.293	3.895	4.121	4.596
8	2.461	2.958	3.267	3.865	4.089	4.560
9	2.446	2.940	3.247	3.841	4.063	4.531
10	2.434	2.925	3.231	3.821	4.043	4.509
11	2.424	2.913	3.217	3.806	4.026	4.490
12	2.415	2.903	3.206	3.792	4.012	4.474
13	2.408	2.895	3.197	3.781	4.000	4.461
14	2.402	2.887	3.188	3.772	3.990	4.450
15	2.396	2.881	3.181	3.763	3.981	4.440
16	2.392	2.875	3.175	3.756	3.974	4.431
17	2.388	2.870	3.170	3.749	3.967	4.423
18	2.384	2.866	3.165	3.743	3.960	4.417
19	2.381	2.862	3.160	3.738	3.955	4.411
20	2.378	2.858	3.156	3.734	3.950	4.405
30	2.359	2.835	3.131	3.704	3.919	4.370
40	2.349	2.824	3.119	3.689	3.903	4.352
60	2.340	2.812	3.106	3.674	3.887	4.335
120	2.330	2.801	3.093	3.659	3.871	4.317
∞	2.320	2.789	3.080	3.644	3.855	4.299

Table B3 (cont.).

k=5	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
1	3.640	4.298	4.705	5.493	5.788	6.403
2	3.152	3.722	4.075	4.757	5.012	5.545
3	2.972	3.509	3.842	4.485	4.726	5.228
4	2.878	3.398	3.720	4.342	4.576	5.062
5	2.820	3.329	3.645	4.255	4.483	4.960
6	2.780	3.282	3.594	4.195	4.421	4.890
7	2.752	3.249	3.557	4.152	4.375	4.840
8	2.730	3.223	3.529	4.120	4.341	4.802
9	2.713	3.203	3.507	4.094	4.314	4.773
10	2.700	3.187	3.490	4.074	4.292	4.749
11	2.688	3.174	3.475	4.057	4.275	4.729
12	2.679	3.163	3.463	4.043	4.260	4.713
13	2.671	3.154	3.453	4.031	4.247	4.699
14	2.664	3.146	3.444	4.020	4.236	4.687
15	2.658	3.139	3.436	4.011	4.227	4.676
16	2.653	3.132	3.430	4.003	4.219	4.667
17	2.649	3.127	3.424	3.997	4.211	4.659
18	2.644	3.122	3.418	3.990	4.205	4.652
19	2.641	3.118	3.414	3.985	4.199	4.645
20	2.637	3.114	3.409	3.980	4.194	4.640
30	2.616	3.089	3.382	3.948	4.160	4.603
40	2.606	3.077	3.368	3.932	4.144	4.584
60	2.595	3.064	3.355	3.916	4.127	4.565
120	2.585	3.052	3.341	3.900	4.110	4.547
∞	2.574	3.039	3.327	3.884	4.093	4.528

Table B3 (cont.).

k=7	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
1	4.251	4.903	5.305	6.078	6.369	6.970
2	3.682	4.246	4.594	5.264	5.515	6.036
3	3.471	4.003	4.331	4.963	5.200	5.691
4	3.361	3.876	4.194	4.805	5.035	5.510
5	3.293	3.798	4.109	4.708	4.933	5.399
6	3.247	3.745	4.052	4.642	4.864	5.323
7	3.214	3.706	4.010	4.595	4.814	5.269
8	3.189	3.677	3.979	4.559	4.776	5.227
9	3.169	3.654	3.954	4.531	4.747	5.195
10	3.153	3.636	3.934	4.508	4.723	5.169
11	3.140	3.621	3.918	4.489	4.703	5.148
12	3.129	3.608	3.904	4.474	4.687	5.130
13	3.120	3.598	3.893	4.460	4.673	5.115
14	3.112	3.589	3.883	4.449	4.661	5.101
15	3.105	3.581	3.874	4.439	4.651	5.090
16	3.099	3.574	3.866	4.430	4.642	5.080
17	3.093	3.567	3.860	4.423	4.634	5.071
18	3.089	3.562	3.854	4.416	4.627	5.064
19	3.084	3.557	3.848	4.410	4.620	5.057
20	3.080	3.552	3.844	4.404	4.614	5.050
30	3.056	3.524	3.813	4.369	4.578	5.010
40	3.044	3.510	3.798	4.351	4.559	4.990
60	3.031	3.496	3.782	4.334	4.541	4.969
120	3.019	3.481	3.767	4.316	4.522	4.949
∞	3.006	3.467	3.751	4.298	4.503	4.928

Table B3 (cont.).

k=10	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
1	5.010	5.655	6.051	6.813	7.098	7.694
2	4.339	4.897	5.241	5.900	6.147	6.663
3	4.091	4.617	4.941	5.563	5.795	6.282
4	3.961	4.471	4.784	5.386	5.611	6.083
5	3.881	4.380	4.687	5.277	5.498	5.960
6	3.826	4.319	4.622	5.204	5.421	5.877
7	3.787	4.275	4.574	5.150	5.365	5.816
8	3.757	4.241	4.539	5.110	5.323	5.771
9	3.734	4.215	4.510	5.078	5.290	5.735
10	3.716	4.194	4.488	5.053	5.264	5.706
11	3.700	4.177	4.469	5.032	5.242	5.683
12	3.687	4.162	4.454	5.014	5.224	5.663
13	3.676	4.150	4.441	5.000	5.208	5.646
14	3.667	4.139	4.429	4.987	5.195	5.632
15	3.659	4.130	4.419	4.976	5.184	5.619
16	3.652	4.122	4.411	4.966	5.173	5.608
17	3.645	4.115	4.403	4.957	5.164	5.598
18	3.640	4.108	4.396	4.950	5.156	5.590
19	3.635	4.103	4.390	4.943	5.149	5.582
20	3.630	4.097	4.385	4.937	5.143	5.575
30	3.601	4.065	4.350	4.897	5.102	5.531
40	3.587	4.048	4.332	4.878	5.081	5.508
60	3.572	4.032	4.315	4.858	5.061	5.486
120	3.557	4.015	4.297	4.838	5.040	5.463
∞	3.543	3.999	4.279	4.818	5.019	5.441

Table B3 (cont.).

k=15	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
1	6.040	6.679	7.070	7.821	8.100	8.678
2	5.231	5.784	6.122	6.773	7.015	7.515
3	4.932	5.453	5.772	6.386	6.614	7.085
4	4.775	5.280	5.589	6.183	6.404	6.860
5	4.678	5.174	5.476	6.058	6.274	6.722
6	4.613	5.101	5.400	5.973	6.186	6.628
7	4.566	5.049	5.344	5.912	6.123	6.560
8	4.530	5.009	5.302	5.866	6.075	6.508
9	4.502	4.978	5.269	5.830	6.037	6.468
10	4.479	4.953	5.243	5.800	6.007	6.435
11	4.461	4.933	5.221	5.776	5.982	6.409
12	4.445	4.916	5.203	5.756	5.961	6.387
13	4.432	4.901	5.188	5.739	5.944	6.368
14	4.421	4.889	5.174	5.724	5.929	6.351
15	4.411	4.878	5.163	5.712	5.915	6.337
16	4.402	4.868	5.153	5.701	5.904	6.325
17	4.395	4.860	5.144	5.691	5.894	6.314
18	4.388	4.852	5.136	5.682	5.885	6.304
19	4.382	4.846	5.129	5.674	5.876	6.295
20	4.376	4.839	5.122	5.667	5.869	6.287
30	4.341	4.801	5.082	5.622	5.822	6.237
40	4.324	4.781	5.061	5.599	5.799	6.212
60	4.306	4.762	5.040	5.576	5.775	6.187
120	4.289	4.742	5.020	5.553	5.751	6.161
∞	4.271	4.723	4.999	5.530	5.728	6.136

Table B3 (cont.).

k=20	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
1	6.902	7.539	7.927	8.667	8.944	9.529
2	5.977	6.529	6.865	7.506	7.746	8.252
3	5.636	6.156	6.472	7.077	7.303	7.780
4	5.457	5.960	6.267	6.852	7.071	7.533
5	5.346	5.840	6.140	6.714	6.928	7.381
6	5.272	5.758	6.054	6.620	6.831	7.278
7	5.218	5.699	5.992	6.552	6.761	7.203
8	5.177	5.654	5.945	6.500	6.708	7.147
9	5.145	5.619	5.909	6.460	6.667	7.102
10	5.119	5.591	5.879	6.428	6.633	7.067
11	5.098	5.568	5.855	6.401	6.606	7.038
12	5.080	5.549	5.834	6.379	6.583	7.013
13	5.065	5.532	5.817	6.360	6.563	6.992
14	5.052	5.518	5.802	6.344	6.547	6.974
15	5.041	5.506	5.789	6.330	6.532	6.959
16	5.031	5.495	5.778	6.317	6.519	6.945
17	5.022	5.486	5.768	6.306	6.508	6.933
18	5.014	5.477	5.759	6.297	6.498	6.923
19	5.007	5.470	5.751	6.288	6.489	6.913
20	5.001	5.463	5.744	6.280	6.481	6.904
30	4.961	5.419	5.698	6.230	6.429	6.849
40	4.941	5.397	5.675	6.205	6.403	6.822
60	4.921	5.375	5.652	6.179	6.377	6.794
120	4.901	5.353	5.629	6.154	6.351	6.766
∞	4.881	5.331	5.605	6.129	6.325	6.738

Table B3 (cont.).

k=40	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
1	9.550	10.178	10.560	11.285	11.555	12.133
2	8.270	8.815	9.145	9.773	10.007	10.507
3	7.797	8.311	8.622	9.214	9.435	9.906
4	7.550	8.047	8.349	8.922	9.135	9.592
5	7.397	7.884	8.180	8.742	8.951	9.398
6	7.294	7.774	8.066	8.619	8.825	9.266
7	7.219	7.694	7.983	8.531	8.735	9.171
8	7.162	7.634	7.920	8.464	8.666	9.099
9	7.118	7.587	7.871	8.412	8.613	9.043
10	7.082	7.549	7.832	8.369	8.569	8.998
11	7.053	7.517	7.799	8.335	8.534	8.961
12	7.029	7.491	7.772	8.306	8.504	8.929
13	7.008	7.469	7.749	8.281	8.479	8.903
14	6.990	7.450	7.729	8.260	8.457	8.880
15	6.974	7.433	7.712	8.242	8.439	8.860
16	6.961	7.419	7.697	8.226	8.422	8.843
17	6.949	7.406	7.684	8.211	8.408	8.828
18	6.938	7.394	7.672	8.199	8.395	8.814
19	6.928	7.384	7.661	8.187	8.383	8.802
20	6.920	7.375	7.652	8.177	8.372	8.791
30	6.864	7.316	7.591	8.112	8.306	8.721
40	6.837	7.287	7.560	8.079	8.272	8.686
60	6.809	7.257	7.529	8.046	8.238	8.650
120	6.781	7.227	7.498	8.013	8.205	8.615
∞	6.753	7.197	7.467	7.980	8.171	8.579

Table B3 (cont.).

k=2	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
3	3.799	7.571	13.326	53.806	100.333	439.269
4	2.015	2.707	3.448	6.349	8.392	16.464
5	1.622	1.948	2.259	3.272	3.878	5.864
6	1.451	1.655	1.838	2.388	2.690	3.595
7	1.354	1.500	1.627	1.988	2.177	2.714
8	1.292	1.405	1.501	1.765	1.898	2.265
9	1.248	1.340	1.417	1.623	1.724	1.996
10	1.216	1.293	1.357	1.525	1.606	1.820
11	1.192	1.258	1.312	1.453	1.521	1.696
12	1.172	1.230	1.277	1.399	1.456	1.604
13	1.156	1.208	1.250	1.356	1.406	1.532
14	1.143	1.189	1.227	1.321	1.365	1.476
15	1.132	1.174	1.208	1.293	1.332	1.431
16	1.122	1.161	1.192	1.269	1.305	1.393
17	1.114	1.150	1.178	1.249	1.281	1.361
18	1.107	1.140	1.166	1.231	1.261	1.334
19	1.100	1.131	1.213	1.216	1.244	1.311
20	1.095	1.124	1.147	1.203	1.229	1.291
21	1.090	1.117	1.139	1.191	1.215	1.273
22	1.085	1.111	1.131	1.181	1.203	1.257
32	1.057	1.073	1.086	1.117	1.131	1.163
42	1.043	1.054	1.064	1.086	1.096	1.119
62	1.028	1.036	1.042	1.057	1.063	1.077
122	1.014	1.018	1.021	1.028	1.031	1.038
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B4. Values of the penalty ratio $r^{N}_{\underline{\mu},\underline{\Sigma}}(k,n,P)$ in equation (II.93). These values apply to the ellipsoidal $\underline{\mu},\underline{\Sigma}$ prediction region (II.92) and to sampling from a k-variate normal population with unknown mean vector and covariance matrix. (For the case k=1 see Table A3 in Appendix A.)

k=3	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
4	4.739	9.820	17.622	73.196	137.649	611.468
5	2.351	3.248	4.202	7.933	10.569	21.034
6	1.829	2.243	2.631	3.892	4.644	7.115
7	1.601	1.857	2.083	2.751	3.118	4.215
8	1.472	1.654	1.809	2.241	2.467	3.106
9	1.389	1.529	1.645	1.957	2.115	2.545
10	1.332	1.444	1.536	1.778	1.896	2.213
11	1.289	1.383	1.459	1.654	1.749	1.996
12	1.256	1.337	1.401	1.565	1.642	1.843
13	1.230	1.301	1.356	1.496	1.562	1.730
14	1.209	1.271	1.320	1.443	1.500	1.643
15	1.191	1.247	1.291	1.400	1.450	1.575
16	1.176	1.227	1.267	1.364	1.409	1.520
17	1.164	1.210	1.246	1.334	1.374	1.474
18	1.152	1.196	1.229	1.309	1.346	1.436
19	1.143	1.183	1.213	1.287	1.321	1.402
20	1.135	1.171	1.200	1.268	1.299	1.375
21	1.127	1.162	1.188	1.252	1.281	1.350
22	1.120	1.153	1.178	1.237	1.264	1.329
23	1.114	1.145	1.168	1.224	1.249	1.310
33	1.076	1.096	1.110	1.145	1.160	1.196
43	1.057	1.071	1.082	1.107	1.118	1.144
63	1.038	1.047	1.054	1.070	1.077	1.093
123	1.019	1.024	1.027	1.035	1.038	1.046
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B4 (cont.).

k=4	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
5	5.532	11.738	21.318	90.191	170.503	764.469
6	2.646	3.723	4.865	9.339	12.507	25.116
7	2.015	2.506	2.964	4.447	5.331	8.236
8	1.738	2.039	2.303	3.079	3.503	4.772
9	1.581	1.794	1.973	2.470	2.728	3.459
10	1.480	1.643	1.776	2.133	2.311	2.798
11	1.410	1.540	1.645	1.920	2.053	2.408
12	1.357	1.466	1.552	1.773	1.879	2.154
13	1.317	1.410	1.483	1.667	1.754	1.976
14	1.285	1.366	1.429	1.586	1.659	1.844
15	1.258	1.330	1.386	1.523	1.586	1.744
16	1.237	1.301	1.351	1.472	1.527	1.664
17	1.219	1.277	1.322	1.430	1.479	1.599
18	1.203	1.256	1.297	1.394	1.439	1.546
19	1.189	1.238	1.276	1.365	1.405	1.502
20	1.178	1.223	1.257	1.339	1.376	1.464
21	1.167	1.209	1.241	1.317	1.350	1.431
22	1.158	1.197	1.227	1.297	1.328	1.403
23	1.149	1.186	1.214	1.280	1.309	1.378
24	1.142	1.177	1.203	1.265	1.292	1.357
34	1.095	1.117	1.133	1.171	1.187	1.225
44	1.071	1.087	1.099	1.126	1.138	1.165
64	1.048	1.058	1.065	1.083	1.090	1.107
124	1.024	1.029	1.032	1.041	1.044	1.052
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B4 (cont.).

k=5	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
6	6.231	13.446	24.625	105.565	200.385	905.580
7	2.913	4.153	5.466	10.623	14.282	28.908
8	2.185	2.747	3.269	4.957	5.964	9.283
9	1.864	2.208	2.506	3.381	3.860	5.294
10	1.683	1.924	2.125	2.683	2.972	3.790
11	1.565	1.749	1.898	2.296	2.494	3.038
12	1.483	1.630	1.748	2.052	2.200	2.594
13	1.422	1.544	1.640	1.884	2.001	2.305
14	1.374	1.479	1.560	1.763	1.858	2.102
15	1.337	1.428	1.498	1.670	1.750	1.954
16	1.306	1.387	1.448	1.598	1.667	1.840
17	1.281	1.353	1.407	1.539	1.600	1.750
18	1.259	1.324	1.373	1.491	1.545	1.677
19	1.241	1.300	1.345	1.451	1.499	1.617
20	1.225	1.279	1.320	1.417	1.460	1.567
21	1.211	1.261	1.299	1.388	1.427	1.524
22	1.198	1.245	1.280	1.362	1.399	1.487
23	1.187	1.231	1.264	1.340	1.374	1.455
24	1.178	1.219	1.249	1.320	1.352	1.427
25	1.169	1.208	1.236	1.303	1.332	1.402
35	1.113	1.137	1.155	1.196	1.213	1.254
45	1.085	1.103	1.115	1.144	1.157	1.186
65	1.057	1.068	1.076	1.095	1.103	1.121
125	1.028	1.034	1.038	1.047	1.050	1.060
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B4 (cont.).

k=7	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
8	7.451	16.437	30.457	133.006	253.927	1159.982
9	3.388	4.919	6.542	12.936	17.489	35.778
10	2.492	3.181	3.820	5.882	7.114	11.187
11	2.096	2.514	2.876	3.934	4.512	6.247
12	1.870	2.162	2.405	3.073	3.418	4.398
13	1.723	1.945	2.123	2.596	2.832	3.475
14	1.619	1.797	1.936	2.296	2.470	2.935
15	1.542	1.689	1.803	2.090	2.227	2.583
16	1.482	1.607	1.703	1.941	2.051	2.337
17	1.435	1.543	1.626	1.827	1.920	2.156
18	1.396	1.492	1.563	1.738	1.817	2.018
19	1.363	1.449	1.513	1.666	1.735	1.908
20	1.336	1.413	1.470	1.607	1.668	1.820
21	1.312	1.383	1.435	1.558	1.612	1.747
22	1.292	1.356	1.404	1.516	1.565	1.686
23	1.274	1.334	1.377	1.479	1.524	1.634
24	1.258	1.314	1.354	1.448	1.489	1.589
25	1.244	1.295	1.333	1.421	1.459	1.551
26	1.232	1.280	1.315	1.396	1.432	1.518
27	1.220	1.266	1.299	1.375	1.407	1.487
37	1.148	1.176	1.196	1.242	1.262	1.308
47	1.111	1.132	1.146	1.179	1.193	1.226
67	1.074	1.088	1.097	1.118	1.126	1.147
127	1.037	1.044	1.048	1.058	1.062	1.072
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B4 (cont.).

k=10	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
11	9.001	20.265	37.962	168.711	323.867	1493.963
12	4.004	5.916	7.945	15.974	21.718	44.852
13	2.897	3.754	4.546	7.109	8.643	13.712
14	2.403	2.922	3.368	4.672	5.384	7.518
15	2.121	2.482	2.779	3.596	4.018	5.212
16	1.936	2.209	2.427	3.002	3.288	4.065
17	1.805	2.023	2.193	2.627	2.837	3.393
18	1.707	1.887	2.025	2.370	2.534	2.957
19	1.631	1.784	1.899	2.184	2.316	2.652
20	1.570	1.702	1.801	2.042	2.152	2.428
21	1.520	1.636	1.723	1.930	2.024	2.258
22	1.479	1.582	1.659	1.840	1.922	2.123
23	1.443	1.537	1.605	1.766	1.838	2.014
24	1.412	1.497	1.559	1.704	1.768	1.924
25	1.386	1.464	1.521	1.652	1.710	1.849
26	1.363	1.435	1.487	1.606	1.659	1.785
27	1.342	1.409	1.457	1.567	1.615	1.729
28	1.324	1.386	1.431	1.532	1.577	1.682
29	1.307	1.366	1.407	1.502	1.543	1.640
30	1.293	1.347	1.386	1.475	1.513	1.604
40	1.198	1.232	1.255	1.308	1.330	1.382
50	1.149	1.174	1.191	1.228	1.243	1.278
70	1.100	1.116	1.127	1.150	1.160	1.181
130	1.051	1.058	1.063	1.074	1.079	1.089
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B4 (cont.).

k=15	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
16	11.154	25.615	48.506	219.380	423.661	1977.395
17	4.874	7.327	9.939	20.324	27.792	58.075
18	3.475	4.574	5.590	8.879	10.855	17.424
19	2.849	3.511	4.081	5.743	6.653	9.394
20	2.487	2.948	3.326	4.362	4.897	6.418
21	2.249	2.597	2.873	3.598	3.959	4.943
22	2.080	2.356	2.571	3.116	3.380	4.080
23	1.953	2.181	2.354	2.786	2.990	3.520
24	1.855	2.047	2.192	2.545	2.710	3.130
25	1.774	1.941	2.065	2.362	2.499	2.843
26	1.709	1.855	1.963	2.218	2.334	2.624
27	1.654	1.783	1.879	2.102	2.202	2.451
28	1.607	1.723	1.808	2.006	2.094	2.311
29	1.566	1.672	1.749	1.926	2.004	2.196
30	1.531	1.628	1.698	1.858	1.928	2.100
31	1.500	1.589	1.653	1.799	1.863	2.016
32	1.472	1.555	1.614	1.748	1.806	1.945
33	1.447	1.525	1.579	1.703	1.756	1.885
34	1.425	1.497	1.548	1.663	1.712	1.830
35	1.406	1.473	1.521	1.627	1.673	1.782
45	1.276	1.318	1.347	1.409	1.436	1.497
55	1.210	1.239	1.260	1.304	1.322	1.365
75	1.142	1.161	1.174	1.201	1.212	1.238
135	1.072	1.081	1.087	1.100	1.105	1.116
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B4 (cont.).

k=20	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
21	12.983	30.172	57.513	263.190	510.124	2393.944
22	5.619	8.540	11.657	24.110	33.086	69.519
23	3.976	5.284	6.495	10.428	12.793	20.642
24	3.237	4.025	4.704	6.687	7.771	11.029
25	2.809	3.357	3.806	5.039	5.675	7.472
26	2.526	2.939	3.266	4.128	4.555	5.713
27	2.325	2.652	2.906	3.553	3.864	4.684
28	2.173	2.442	2.647	3.158	3.398	4.017
29	2.054	2.281	2.453	2.870	3.063	3.552
30	1.958	2.154	2.300	2.651	2.811	3.209
31	1.879	2.051	2.178	2.479	2.614	2.949
32	1.812	1.965	2.077	2.339	2.456	2.741
33	1.755	1.893	1.992	2.225	2.327	2.574
34	1.706	1.831	1.921	2.128	2.219	2.438
35	1.663	1.777	1.859	2.046	2.128	2.322
36	1.625	1.730	1.805	1.975	2.049	2.223
37	1.591	1.688	1.757	1.914	1.981	2.140
38	1.561	1.651	1.715	1.859	1.921	2.065
39	1.534	1.618	1.678	1.811	1.868	2.001
40	1.510	1.589	1.644	1.768	1.821	1.944
50	1.350	1.398	1.431	1.504	1.533	1.600
60	1.268	1.301	1.325	1.375	1.396	1.443
80	1.182	1.204	1.218	1.249	1.261	1.289
140	1.093	1.103	1.110	1.124	1.130	1.140
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B4 (cont.).

k=40	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
41	18.686	44.480	85.919	402.258	785.785	3741.756
42	7.971	12.386	17.122	36.213	50.084	106.775
43	5.568	7.555	9.398	15.416	19.056	31.196
44	4.481	5.683	6.715	9.743	11.407	16.414
45	3.847	4.683	5.367	7.246	8.218	10.967
46	3.428	4.057	4.555	5.864	6.514	8.275
47	3.125	3.625	4.010	4.991	5.463	6.706
48	2.897	3.307	3.618	4.390	4.753	5.686
49	2.716	3.063	3.322	3.951	4.242	4.976
50	2.570	2.869	3.089	3.616	3.856	4.453
51	2.448	2.710	2.901	3.352	3.555	4.052
52	2.345	2.577	2.746	3.138	3.313	3.736
53	2.257	2.466	2.615	2.961	3.114	3.481
54	2.181	2.369	2.504	2.813	2.947	3.269
55	2.113	2.285	2.407	2.685	2.806	3.092
56	2.054	2.212	2.324	2.575	2.684	2.938
57	2.000	2.147	2.249	2.480	2.579	2.809
58	1.953	2.089	2.183	2.395	2.486	2.695
59	1.910	2.036	2.124	2.320	2.404	2.595
60	1.871	1.989	2.071	2.253	2.330	2.508
70	1.612	1.683	1.732	1.835	1.878	1.973
80	1.475	1.525	1.558	1.630	1.659	1.722
100	1.329	1.360	1.380	1.424	1.441	1.477
160	1.172	1.187	1.196	1.215	1.223	1.236
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B4 (cont.).

k=2	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
2	3.290	6.556	11.540	46.598	86.891	380.418
3	1.802	2.421	3.084	5.679	7.506	14.726
4	1.481	1.778	2.062	2.987	3.540	5.353
5	1.343	1.532	1.702	2.210	2.490	3.329
6	1.266	1.403	1.522	1.860	2.037	2.539
7	1.218	1.325	1.415	1.664	1.790	2.135
8	1.184	1.271	1.344	1.539	1.636	1.894
9	1.160	1.233	1.294	1.454	1.531	1.735
10	1.141	1.204	1.256	1.391	1.456	1.623
11	1.126	1.182	1.227	1.344	1.399	1.541
12	1.114	1.164	1.204	1.307	1.355	1.477
13	1.104	1.149	1.185	1.277	1.319	1.426
14	1.096	1.137	1.170	1.252	1.290	1.385
15	1.089	1.126	1.156	1.231	1.266	1.351
16	1.083	1.117	1.145	1.214	1.245	1.323
17	1.077	1.109	1.135	1.199	1.228	1.299
18	1.073	1.103	1.182	1.185	1.212	1.278
19	1.068	1.097	1.119	1.174	1.199	1.260
20	1.065	1.091	1.112	1.164	1.187	1.244
21	1.061	1.086	1.106	1.155	1.177	1.230
31	1.040	1.057	1.069	1.100	1.113	1.145
41	1.031	1.042	1.052	1.074	1.083	1.106
61	1.020	1.028	1.034	1.048	1.055	1.069
121	1.010	1.014	1.017	1.024	1.027	1.034
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B5. Values of the penalty ratio $r_{\underline{\Sigma}}^{N}(k,n,P)$ in equation (II.95). These values apply to the ellipsoidal $\underline{\Sigma}$ prediction region (II.94) and to sampling from a k-variate normal population with known mean and unknown covariance matrix. (For the case k=1, see Table A2 in Appendix A.)

k=3	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
3	4.239	8.783	15.761	65.468	123.117	546.913
4	2.146	2.965	3.836	7.242	9.648	19.202
5	1.694	2.076	2.436	3.603	4.300	6.587
6	1.498	1.737	1.948	2.573	2.916	3.943
7	1.388	1.559	1.705	2.113	2.326	2.928
8	1.318	1.451	1.560	1.857	2.006	2.415
9	1.270	1.377	1.465	1.695	1.808	2.110
10	1.234	1.324	1.397	1.584	1.674	1.911
11	1.207	1.284	1.346	1.503	1.578	1.770
12	1.185	1.253	1.307	1.442	1.505	1.667
13	1.168	1.228	1.276	1.394	1.449	1.588
14	1.153	1.208	1.250	1.355	1.404	1.525
15	1.141	1.191	1.229	1.323	1.367	1.474
16	1.131	1.176	1.211	1.297	1.336	1.432
17	1.122	1.164	1.196	1.274	1.310	1.398
18	1.114	1.153	1.183	1.255	1.287	1.367
19	1.107	1.143	1.171	1.238	1.268	1.342
20	1.101	1.135	1.161	1.223	1.251	1.319
21	1.096	1.128	1.152	1.210	1.236	1.299
22	1.091	1.121	1.144	1.198	1.223	1.282
32	1.060	1.079	1.094	1.128	1.143	1.178
42	1.045	1.059	1.070	1.094	1.105	1.131
62	1.030	1.039	1.046	1.062	1.069	1.085
122	1.015	1.019	1.023	1.030	1.034	1.042
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B5 (cont.).

k=4	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
4	5.050	10.716	19.461	82.333	155.647	697.860
5	2.450	3.447	4.504	8.646	11.579	23.253
6	1.884	2.344	2.773	4.159	4.987	7.704
7	1.638	1.922	2.171	2.902	3.303	4.499
8	1.500	1.702	1.871	2.343	2.588	3.281
9	1.411	1.566	1.693	2.033	2.204	2.668
10	1.350	1.475	1.575	1.838	1.966	2.306
11	1.304	1.408	1.492	1.704	1.805	2.069
12	1.269	1.359	1.429	1.606	1.690	1.904
13	1.241	1.320	1.381	1.532	1.603	1.782
14	1.218	1.288	1.342	1.474	1.535	1.689
15	1.200	1.262	1.311	1.428	1.481	1.614
16	1.184	1.241	1.284	1.389	1.437	1.554
17	1.171	1.223	1.262	1.357	1.400	1.505
18	1.159	1.207	1.243	1.330	1.369	1.464
19	1.149	1.193	1.227	1.307	1.342	1.429
20	1.140	1.182	1.213	1.286	1.319	1.398
21	1.132	1.171	1.200	1.269	1.299	1.373
22	1.125	1.161	1.189	1.253	1.281	1.349
23	1.119	1.153	1.179	1.239	1.266	1.329
33	1.079	1.101	1.117	1.154	1.170	1.207
43	1.059	1.075	1.087	1.113	1.125	1.152
63	1.039	1.050	1.057	1.074	1.082	1.099
123	1.020	1.025	1.028	1.037	1.040	1.048
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B5 (cont.).

k=5	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
5	5.769	12.448	22.799	97.735	185.521	838.406
6	2.725	3.885	5.113	9.937	13.360	27.041
7	2.060	2.590	3.082	4.673	5.623	8.752
8	1.768	2.094	2.377	3.208	3.661	5.022
9	1.604	1.834	2.026	2.558	2.833	3.614
10	1.498	1.675	1.817	2.198	2.388	2.908
11	1.425	1.566	1.679	1.971	2.113	2.493
12	1.370	1.488	1.581	1.816	1.928	2.222
13	1.328	1.429	1.507	1.703	1.795	2.031
14	1.294	1.383	1.450	1.617	1.695	1.892
15	1.267	1.345	1.405	1.550	1.617	1.785
16	1.244	1.315	1.368	1.496	1.554	1.700
17	1.225	1.289	1.337	1.452	1.504	1.632
18	1.209	1.267	1.311	1.414	1.461	1.576
19	1.195	1.249	1.288	1.383	1.425	1.529
20	1.183	1.232	1.269	1.356	1.395	1.489
21	1.172	1.218	1.252	1.332	1.368	1.454
22	1.162	1.205	1.237	1.312	1.345	1.425
23	1.154	1.194	1.224	1.294	1.324	1.398
24	1.146	1.184	1.212	1.277	1.306	1.375
34	1.097	1.121	1.139	1.179	1.196	1.236
44	1.073	1.091	1.103	1.132	1.144	1.173
64	1.049	1.060	1.068	1.087	1.094	1.113
124	1.024	1.030	1.034	1.043	1.046	1.055
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B5 (cont.).

k=7	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
7	7.025	15.497	28.715	125.399	239.405	1093.642
8	3.215	4.667	6.206	12.272	16.591	33.942
9	2.376	3.033	3.642	5.609	6.783	10.667
10	2.006	2.407	2.753	3.767	4.320	5.981
11	1.797	2.077	2.310	2.952	3.284	4.225
12	1.660	1.874	2.046	2.502	2.729	3.349
13	1.564	1.736	1.871	2.218	2.387	2.835
14	1.493	1.635	1.746	2.024	2.156	2.501
15	1.438	1.559	1.652	1.883	1.990	2.267
16	1.394	1.500	1.580	1.776	1.866	2.095
17	1.359	1.452	1.522	1.692	1.769	1.964
18	1.329	1.412	1.474	1.624	1.691	1.860
19	1.304	1.379	1.435	1.568	1.628	1.776
20	1.282	1.351	1.402	1.522	1.575	1.707
21	1.263	1.326	1.373	1.482	1.530	1.649
22	1.247	1.305	1.348	1.448	1.492	1.600
23	1.233	1.287	1.326	1.419	1.459	1.557
24	1.220	1.270	1.307	1.393	1.430	1.521
25	1.209	1.256	1.290	1.370	1.405	1.489
26	1.198	1.243	1.275	1.350	1.382	1.460
36	1.133	1.161	1.180	1.226	1.245	1.291
46	1.099	1.120	1.134	1.167	1.180	1.213
66	1.067	1.080	1.089	1.109	1.118	1.139
126	1.033	1.040	1.044	1.054	1.058	1.068
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B5 (cont.).

k=10	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
10	8.618	19.402	36.346	161.528	310.079	1430.365
11	3.847	5.684	7.633	15.347	20.866	43.093
12	2.792	3.617	4.381	6.850	8.329	13.213
13	2.322	2.823	3.254	4.513	5.201	7.263
14	2.053	2.403	2.691	3.482	3.891	5.046
15	1.878	2.143	2.354	2.912	3.189	3.943
16	1.754	1.966	2.131	2.553	2.758	3.297
17	1.662	1.837	1.971	2.307	2.466	2.878
18	1.590	1.738	1.851	2.128	2.257	2.585
19	1.532	1.661	1.758	1.992	2.100	2.370
20	1.485	1.599	1.683	1.886	1.978	2.206
21	1.446	1.547	1.622	1.800	1.880	2.076
22	1.413	1.504	1.571	1.729	1.799	1.972
23	1.384	1.467	1.528	1.670	1.733	1.885
24	1.359	1.435	1.491	1.620	1.676	1.813
25	1.338	1.408	1.459	1.576	1.628	1.751
26	1.318	1.383	1.431	1.539	1.586	1.698
27	1.301	1.362	1.406	1.506	1.549	1.653
28	1.285	1.343	1.383	1.477	1.517	1.612
29	1.272	1.325	1.364	1.451	1.488	1.578
39	1.183	1.216	1.240	1.292	1.314	1.365
49	1.138	1.162	1.179	1.216	1.231	1.266
69	1.093	1.108	1.119	1.142	1.151	1.173
129	1.047	1.054	1.059	1.070	1.074	1.085
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B5 (cont.).

k=15	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
15	10.821	24.851	47.058	212.830	411.011	1918.342
16	4.736	7.121	9.659	19.751	27.009	56.439
17	3.382	4.452	5.441	8.642	10.566	16.959
18	2.776	3.422	3.978	5.598	6.485	9.156
19	2.427	2.877	3.246	4.257	4.779	6.263
20	2.198	2.537	2.807	3.515	3.868	4.829
21	2.035	2.304	2.514	3.048	3.306	3.990
22	1.912	2.135	2.305	2.727	2.927	3.446
23	1.817	2.005	2.147	2.494	2.655	3.067
24	1.740	1.903	2.024	2.316	2.450	2.788
25	1.677	1.820	1.926	2.177	2.291	2.574
26	1.624	1.751	1.845	2.064	2.163	2.407
27	1.579	1.693	1.777	1.971	2.058	2.270
28	1.540	1.644	1.720	1.894	1.971	2.159
29	1.506	1.601	1.670	1.827	1.897	2.066
30	1.476	1.564	1.627	1.771	1.833	1.984
31	1.449	1.531	1.589	1.721	1.778	1.915
32	1.426	1.502	1.556	1.677	1.730	1.857
33	1.405	1.476	1.526	1.639	1.688	1.803
34	1.386	1.452	1.499	1.605	1.650	1.757
44	1.262	1.303	1.332	1.394	1.420	1.480
54	1.199	1.228	1.249	1.292	1.310	1.352
74	1.134	1.153	1.166	1.193	1.204	1.230
134	1.068	1.077	1.083	1.096	1.101	1.112
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B5 (cont.).

k=20	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
20	12.684	29.478	56.190	257.139	498.395	2338.899
21	5.496	8.352	11.400	23.580	32.359	67.991
22	3.892	5.173	6.358	10.209	12.524	20.207
23	3.171	3.944	4.609	6.552	7.614	10.806
24	2.754	3.291	3.732	4.941	5.565	7.327
25	2.479	2.884	3.205	4.051	4.470	5.606
26	2.283	2.604	2.854	3.489	3.795	4.600
27	2.135	2.400	2.601	3.103	3.339	3.947
28	2.019	2.243	2.412	2.822	3.012	3.493
29	1.926	2.119	2.263	2.608	2.765	3.157
30	1.849	2.019	2.143	2.440	2.573	2.902
31	1.784	1.935	2.045	2.304	2.419	2.699
32	1.729	1.865	1.963	2.192	2.292	2.536
33	1.681	1.804	1.893	2.097	2.187	2.403
34	1.640	1.752	1.833	2.017	2.098	2.290
35	1.603	1.706	1.780	1.948	2.021	2.193
36	1.570	1.666	1.734	1.888	1.954	2.112
37	1.541	1.630	1.693	1.835	1.896	2.039
38	1.515	1.598	1.657	1.788	1.844	1.976
39	1.491	1.569	1.624	1.747	1.799	1.920
49	1.337	1.384	1.417	1.489	1.518	1.585
59	1.257	1.291	1.314	1.364	1.384	1.431
79	1.175	1.196	1.210	1.241	1.254	1.281
139	1.089	1.099	1.106	1.120	1.126	1.136
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B5 (cont.).

k=40	Probability content, P					
n	0.7500	0.9000	0.9500	0.9900	0.9950	0.9990
40	18.462	43.948	84.890	397.441	776.370	3696.922
41	7.878	12.241	16.922	35.790	49.498	105.526
42	5.504	7.469	9.290	15.240	18.839	30.840
43	4.431	5.619	6.640	9.634	11.279	16.231
44	3.805	4.632	5.309	7.166	8.128	10.847
45	3.391	4.013	4.506	5.801	6.444	8.186
46	3.092	3.587	3.968	4.938	5.406	6.636
47	2.867	3.273	3.581	4.345	4.705	5.628
48	2.689	3.032	3.288	3.912	4.200	4.926
49	2.544	2.840	3.058	3.580	3.818	4.409
50	2.424	2.684	2.873	3.320	3.521	4.013
51	2.323	2.553	2.720	3.108	3.281	3.701
52	2.236	2.443	2.591	2.934	3.085	3.448
53	2.161	2.347	2.481	2.787	2.920	3.239
54	2.094	2.265	2.386	2.661	2.781	3.065
55	2.035	2.193	2.303	2.553	2.661	2.912
56	1.983	2.128	2.230	2.458	2.556	2.784
57	1.936	2.071	2.165	2.375	2.464	2.672
58	1.894	2.019	2.106	2.301	2.383	2.573
59	1.855	1.973	2.054	2.235	2.311	2.488
69	1.601	1.671	1.719	1.822	1.865	1.959
79	1.466	1.516	1.549	1.619	1.648	1.712
99	1.322	1.353	1.374	1.417	1.434	1.470
159	1.169	1.183	1.192	1.211	1.219	1.233
∞	1.000	1.000	1.000	1.000	1.000	1.000

Table B5 (cont.).

APPENDIX C

Collected in this appendix are tables of simultaneous prediction limits and associated penalty ratios when sampling from independent and stationary univariate sequences. The distribution type of the sequence is known, but at least one distribution parameter is unknown. N denotes the number of future observations, n the available sample size and P the desired content of the prediction interval. For the use of these tables see Paragraph II.3.2.

n \ N	1	2	3	4	5	10	20
2	3.770	7.733	11.145	15.562	******	38.973	******
3	2.178	3.372	4.134	4.969	5.485	8.042	******
4	1.831	2.631 (2.48)	3.091 (2.87)	3.558 (3.15)	3.826 (3.36)	5.077 (4.02)	6.418 (4.64)
5	1.679	2.335 (2.24)	2.692 (2.57)	3.041 (2.80)	3.237 (2.98)	4.105 (3.52)	4.984 (4.04)
6	1.594	2.176 (2.11)	2.482 (2.40)	2.777 (2.61)	2.938 (2.76)	3.635 (3.25)	4.299 (3.71)
7	1.539	2.077 (2.02)	2.352 (2.29)	2.616 (2.48)	2.758 (2.63)	3.360 (3.08)	3.921 (3.50)
8	1.501	2.010 (1.96)	2.268 (2.22)	2.508 (2.40)	2.638 (2.54)	3.180 (2.96)	3.672 (3.36)
9	1.473	1.961 (1.92)	2.203 (2.17)	2.431 (2.34)	2.552 (2.47)	3.053 (2.87)	3.500 (3.26)
10	1.451	1.922 (1.88)	2.154 (2.13)	2.372 (2.29)	2.488 (2.42)	2.959 (2.81)	3.377 (3.18)
11	1.433	1.893 (1.86)	2.119 (2.09)	2.327 (2.26)	2.439 (2.38)	2.887 (2.76)	3.280 (3.11)
12	1.419	1.869 (1.84)	2.090 (2.07)	2.291 (2.23)	2.399 (2.35)	2.829 (2.71)	3.204 (3.06)
15	1.389	1.819 (1.79)	2.039 (2.01)	2.215 (2.16)	2.313 (2.28)	2.710 (2.62)	3.047 (2.95)
20	1.361	1.772 (1.75)	1.982 (1.96)	2.145 (2.11)	2.240 (2.21)	2.602 (2.54)	2.910 (2.85)
25	1.344	1.745 (1.72)	1.935 (1.93)	2.105 (2.07)	2.193 (2.18)	2.541 (2.49)	2.830 (2.79)
30	1.333	1.727 (1.71)	1.912 (1.91)	2.079 (2.05)	2.165 (2.15)	2.503 (2.46)	2.765 (2.75)
40	1.319	1.706 (1.69)	1.883 (1.89)	2.048 (2.02)	2.136 (2.12)	2.455 (2.42)	2.723 (2.70)
60	1.307	1.685 (1.67)	1.865 (1.86)	2.018 (1.99)	2.107 (2.09)	2.412 (2.38)	2.672 (2.65)
∞	1.282	1.645	1.812	1.960	2.036	2.326	2.559

Table C1. One-sided simultaneous prediction intervals of 0.90-content for normal sequences with unknown mean and variance. Approximate values of $\hat{\beta}^{N}_{\mu,\sigma}(0.90,n,N)$ from equation (II.130). In parenthesis are the exact values from Hahn (1970). n = available sample size; N = future sample size. Missing values are due to limitations in the available tables of the "Student's" t-distribution.

n \ N	1	2	3	4	5	10	20
2	38.973	77.964	******	155.936	******	389.847	779.696
3	8.042	11.460	******	16.269	******	25.781	36.486
4	5.077	6.530	7.524	8.333	8.989	11.420	14.449
		(6.30)	(7.07)	(7.63)	(8.07)	(9.43)	(10.76)
5	4.105	5.043	5.647	6.132	6.518	7.858	9.432
		(4.94)	(5.46)	(5.83)	(6.13)	(7.04)	(7.95)
6	3.635	4.355	4.812	5.155	5.433	6.365	7.419
		(4.30)	(4.70)	(4.99)	(5.22)	(5.94)	(6.64)
7	3.360	3.963	4.338	4.615	4.843	5.568	6.370
		(3.93)	(4.27)	(4.51)	(4.70)	(5.30)	(5.90)
8	3.180	3.711	4.033	4.273	4.462	5.075	5.736
		(3.69)	(3.99)	(4.20)	(4.37)	(4.90)	(5.43)
9	3.053	3.536	3.830	4.040	4.209	4.744	5.314
		(3.52)	(3.79)	(3.99)	(4.14)	(4.62)	(5.09)
10	2.959	3.409	3.674	3.870	4.024	4.507	5.014
		(3.39)	(3.65)	(3.83)	(3.97)	(4.41)	(4.85)
11	2.887	3.310	3.555	3.740	3.885	4.328	4.791
		(3.30)	(3.54)	(3.71)	(3.84)	(4.25)	(4.67)
12	2.829	3.233	3.471	3.640	3.769	4.189	4.618
		(3.22)	(3.45)	(3.61)	(3.74)	(4.13)	(4.52)
15	2.710	3.075	3.284	3.435	3.553	3.911	4.276
		(3.07)	(3.27)	(3.42)	(3.53)	(3.88)	(4.22)
20	2.602	2.932	3.117	3.252	3.354	3.667	3.979
		(2.93)	(3.11)	(3.24)	(3.34)	(3.65)	(3.95)
25	2.541	2.852	3.029	3.151	3.247	3.536	3.819
		(2.85)	(3.02)	(3.14)	(3.24)	(3.52)	(3.80)
30	2.503	2.802	2.968	3.088	3.161	3.452	3.719
		(2.80)	(2.97)	(3.08)	(3.17)	(3.44)	(3.71)
40	2.455	2.741	2.901	3.013	3.098	3.354	3.601
		(2.74)	(2.90)	(3.01)	(3.09)	(3.35)	(3.59)
60	2.412	2.684	2.833	2.941	3.025	3.262	3.492
		(2.68)	(2.83)	(2.94)	(3.02)	(3.26)	(3.49)
∞	2.326	2.576	2.713	2.807	2.878	3.090	3.291

Table C2. One-sided simultaneous prediction intervals of 0.99 content for normal sequences with unknown mean and variance. Approximate values of $\overline{\beta}^{N}_{\mu,\sigma}(0.99,n,N)$ from equation (II.130). In parenthesis are the exact values from Hahn (1970). n = available sample size; N = future sample size. Missing values are due to limitations in the available tables of the "Student's" t-distribution.

n \ N	1	2	3	4	5	10	20
2	389.847	779.696	******	******	******	3898.40	7796.70
3	25.781	36.486	******	******	******	81.637	115.461
4	11.420	14.449	******	******	******	24.820	31.305
5	7.858	9.432	******	******	******	14.274	17.028
6	6.365	7.419	8.096	8.614	9.041	10.456	12.074
7	5.568	6.370	6.874	7.259	7.574	8.574	9.709
8	5.075	5.736	6.152	6.459	6.693	7.488	8.363
9	4.744	5.314	5.660	5.935	6.145	6.788	7.505
10	4.507	5.014	5.323	5.559	5.726	6.303	6.916
11	4.328	4.791	5.071	5.275	5.434	5.943	6.487
12	4.189	4.618	4.866	5.058	5.204	5.673	6.163
15	3.911	4.276	4.495	4.565	4.699	5.164	5.539
20	3.667	3.979	4.160	4.293	4.406	4.714	5.018
25	3.536	3.819	3.982	4.100	4.191	4.487	4.746
30	3.452	3.719	3.898	4.003	4.066	4.320	4.580
40	3.354	3.601	3.749	3.848	3.927	4.161	4.385
60	3.262	3.492	3.623	3.719	3.788	4.002	4.209
∞	3.090	3.291	3.403	3.481	3.540	3.719	3.891

Table C3. One-sided simultaneous prediction intervals of 0.999-content for normal sequences with unknown mean and variance. Approximate values of $\overline{\beta}^{N}_{\mu,\sigma}(0.999,n,N)$ from equation (II.130). n = available sample size; N = future sample size. Missing values are due to limitations in the available tables of the "Student's" t-distribution.

n \ N	1	2	3	4	5	10	20
2	3898.40	7796.70	******	******	******	38984.9	77968.5
3	81.637	115.461	******	******	******	258.653	365.145
4	24.820	31.305	******	******	******	53.565	67.526
5	14.274	17.028	******	******	******	25.557	30.422
6	10.456	12.074	******	******	******	16.785	19.331
7	8.574	9.709	******	******	******	12.861	14.491
8	7.488	8.363	8.888	9.334	9.599	10.723	11.895
9	6.788	7.505	7.958	8.275	8.538	9.381	10.311
10	6.303	6.916	7.289	7.551	7.793	8.495	9.258
11	5.943	6.487	6.841	7.067	7.259	7.865	8.512
12	5.673	6.163	6.453	6.682	6.821	7.411	7.860
15	5.164	5.539	5.763	5.939	6.068	6.507	6.926
20	4.714	5.018	5.200	5.349	5.446	5.790	6.096
25	4.487	4.746	4.905	5.038	5.109	5.405	5.676
30	4.320	4.580	4.732	4.839	4.930	5.184	5.423
40	4.161	4.385	4.532	4.627	4.698	4.920	5.128
60	4.002	4.209	4.333	4.416	4.477	4.679	4.871
∞	3.719	3.891	3.988	4.056	4.107	4.265	4.417

Table C4. One-sided simultaneous prediction intervals of 0.9999-content for normal sequences with unknown mean and variance. Approximate values of $\overline{p}^{N}_{\mu,\sigma}(0.9999,n,N)$ from equation (II.130). n = available sample size; N = future sample size. Missing values are due to limitations in the available tables of the "Student's" t-distribution.

n \ N	1	2	3	4	5	10	20
2	16.755	(30.266)	******	(55.553)	******	(126.16)	(236.92)
3	3.457	(4.449)	******	(5.796)	******	(8.343)	(11.087)
4	2.183	2.446	2.606	2.718	2.804	3.052	3.270
5	1.765	1.918	2.013	2.077	2.130	2.278	2.416
6	1.563	1.669	1.732	1.778	1.814	1.922	2.018
7	1.445	1.526	1.574	1.607	1.633	1.715	1.793
8	1.367	1.432	1.471	1.496	1.518	1.586	1.650
9	1.313	1.366	1.397	1.421	1.438	1.495	1.547
10	1.272	1.316	1.345	1.364	1.379	1.427	1.474
11	1.241	1.281	1.305	1.322	1.334	1.375	1.419
12	1.216	1.250	1.272	1.286	1.300	1.337	1.373
15	1.165	1.192	1.205	1.218	1.227	1.256	1.282
20	1.119	1.137	1.146	1.154	1.161	1.181	1.200
25	1.092	1.106	1.113	1.119	1.126	1.139	1.155
30	1.076	1.087	1.095	1.097	1.101	1.113	1.127
40	1.055	1.064	1.069	1.072	1.074	1.084	1.091
60	1.037	1.040	1.043	1.047	1.049	1.055	1.060
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table C5. One-sided simultaneous prediction for normal sequences with unknown mean and variance. Values of the penalty ratio $\bar{r}^{N}_{\mu,\sigma}(0.99,n,N)$ in equation (II.128) (in parenthesis: approximate values). n = available sample size; N = future sample size.

n \ N	1	2	3	4	5	10	20
2	126.164	236.918	******	******	******	1048.24	2003.78
3	8.343	11.087	******	******	******	21.951	29.674
4	3.696	4.390	******	******	******	6.674	8.045
5	2.543	2.866	******	******	******	3.838	4.376
6	2.060	2.254	2.379	2.475	2.554	2.812	3.103
7	1.802	1.936	2.020	2.085	2.140	2.305	2.495
8	1.642	1.743	1.808	1.856	1.891	2.013	2.149
9	1.535	1.615	1.663	1.705	1.736	1.825	1.929
10	1.459	1.524	1.564	1.597	1.618	1.695	1.777
11	1.401	1.456	1.490	1.515	1.535	1.598	1.667
12	1.356	1.403	1.430	1.453	1.470	1.525	1.584
15	1.266	1.299	1.321	1.311	1.327	1.389	1.424
20	1.187	1.209	1.222	1.233	1.245	1.268	1.290
25	1.144	1.160	1.170	1.178	1.184	1.207	1.220
30	1.117	1.130	1.145	1.150	1.149	1.162	1.177
40	1.085	1.094	1.102	1.105	1.109	1.119	1.127
60	1.056	1.061	1.065	1.068	1.070	1.076	1.082
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table C6. One-sided simultaneous prediction for normal sequences with unknown mean and variance. Approximate values of the penalty ratio $\bar{r}^{N}_{\mu,\sigma}(0.999, n,N)$ in equation (II.128). n = available sample size; N = future sample size.

n \ N	1	2	3	4	5	10	20
2	1048.24	2003.78	******	******	******	9140.65	17651.9
3	21.951	29.674	******	******	******	60.645	82.668
4	6.674	8.045	******	******	******	12.559	15.288
5	3.838	4.376	******	******	******	5.992	6.887
6	2.812	3.103	******	******	******	3.936	4.376
7	2.305	2.495	******	******	******	3.015	3.281
8	2.013	2.149	2.229	2.301	2.337	2.514	2.693
9	1.825	1.929	1.995	2.040	2.079	2.200	2.334
10	1.695	1.777	1.828	1.862	1.897	1.992	2.096
11	1.598	1.667	1.715	1.742	1.767	1.844	1.927
12	1.525	1.584	1.618	1.647	1.661	1.738	1.779
15	1.389	1.424	1.445	1.464	1.477	1.526	1.568
20	1.268	1.290	1.304	1.319	1.326	1.358	1.380
25	1.207	1.220	1.230	1.242	1.244	1.267	1.285
30	1.162	1.177	1.187	1.193	1.200	1.215	1.228
40	1.119	1.127	1.136	1.141	1.144	1.154	1.161
60	1.076	1.082	1.087	1.089	1.090	1.097	1.103
∞	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table C7. One-sided simultaneous prediction for normal sequences with unknown mean and variance. Approximate values of the penalty ratio $\bar{r}^{N}_{\mu,\sigma}(0.9999,n,N)$ in equation (II.128). n = available sample size; N = future sample size.

n \ N	1	2	3	4	5	10	20
1	31.821	63.657	******	127.32	******	318.31	636.62
2	6.965	9.925	******	14.089	******	22.327	31.598
3	4.541	5.841	6.730	7.453	8.040	10.214	12.924
4	3.747	4.604	5.155	5.598	5.950	7.173	8.610
5	3.365	4.032 (4.00)	4.455 (4.39)	4.773 (4.67)	5.030 (4.90)	5.893 (5.61)	6.869
6	3.143	3.707 (3.68)	4.058 (4.01)	4.317 (4.25)	4.530 (4.44)	5.208 (5.03)	5.959
7	2.998	3.499 (3.48)	3.802 (3.77)	4.029 (3.98)	4.207 (4.15)	4.785 (4.67)	5.408
8	2.896	3.355 (3.34)	3.633 (3.61)	3.833 (3.80)	3.993 (3.95)	4.501 (4.42)	5.041
9	2.821	3.250 (3.24)	3.503 (3.49)	3.690 (3.66)	3.837 (3.80)	4.297 (4.23)	4.781
10	2.764	3.169 (3.16)	3.404 (3.39)	3.581 (3.56)	3.720 (3.69)	4.144 (4.09)	4.587
11	2.718	3.106 (3.10)	3.335 (3.32)	3.497 (3.48)	3.621 (3.60)	4.025 (3.98)	4.437
14	2.624	2.977 (2.97)	3.180 (3.17)	3.326 (3.31)	3.440 (3.42)	3.787 (3.76)	4.140
19	2.539	2.861 (2.86)	3.042 (3.04)	3.174 (3.16)	3.273 (3.26)	3.579 (3.56)	3.883
24	2.492	2.797 (2.79)	2.965 (2.96)	3.090 (3.08)	3.184 (3.17)	3.467 (3.46)	3.745
29	2.462	2.756 (2.75)	2.922 (2.92)	3.038 (3.03)	3.130 (3.12)	3.396 (3.39)	3.659
39	2.426	2.707	2.865	2.976	3.060	3.313	3.557
59	2.391	2.662	2.810	2.917	3.000	3.235	3.463
∞	2.326	2.576	2.713	2.807	2.878	3.090	3.291

Table C8. One-sdied simultaneous prediction intervals of 0.99-content for normal sequences with known mean and unknown variance. Approximate values of $\beta_Q^N(0.99,n,N)$ from equation (II.132a). In parenthesis are exact values from Krishnaiah and Armitage (1966), equations (II.131).

n \ N	1	2	3	4	5	10	20
1	318.31	636.62	******	******	******	3183.00	6366.20
2	22.33	31.60	******	******	******	70.70	99.99
3	10.21	12.92	******	******	******	22.20	28.00
4	7.17	8.61	******	******	******	13.03	15.54
5	5.89	6.87	7.49	7.98	8.37	9.68	11.18
6	5.21	5.96	6.43	6.79	7.09	8.02	9.08
7	4.78	5.41	5.80	6.09	6.31	7.06	7.89
8	4.50	5.04	5.37	5.63	5.83	6.44	7.12
9	4.30	4.78	5.07	5.30	5.46	6.01	6.59
10	4.14	4.59	4.85	5.05	5.20	5.69	6.21
11	4.02	4.44	4.68	4.86	5.00	5.45	5.92
14	3.79	4.14	4.27	4.42	4.55	5.00	5.36
19	3.58	3.88	4.06	4.19	4.30	4.60	4.90
24	3.47	3.74	3.90	4.02	4.11	4.40	4.65
29	3.40	3.66	3.84	3.94	4.00	4.25	4.51
39	3.31	3.56	3.70	3.80	3.88	4.11	4.33
59	3.23	3.46	3.59	3.69	3.76	3.97	4.17
∞	3.09	3.29	3.40	3.48	3.54	3.72	3.89

Table C9. One-sided simultaneous prediction intervals of 0.999-content for normal sequences with known mean and unknown variance. Approximate values of $\beta_{\sigma}^{N}(0.999,n,N)$ from equation (II.132a).

n \ N	1	2	3	4	5	10	20
1	3183.0	6366.2	******	******	******	31831.	63662.
2	70.70	99.99	******	******	******	224.00	316.23
3	22.20	28.00	******	******	******	47.91	60.40
4	13.03	15.54	******	******	******	23.33	27.77
5	9.68	11.18	******	******	******	15.54	17.90
6	8.02	9.08	******	******	******	12.03	13.56
7	7.06	7.89	8.38	8.80	9.05	10.11	11.22
8	6.44	7.12	7.55	7.85	8.10	8.90	9.78
9	6.01	6.59	6.95	7.20	7.43	8.10	8.83
10	5.69	6.21	6.55	6.77	6.95	7.53	8.15
11	5.45	5.92	6.20	6.42	6.55	7.12	7.65
14	5.00	5.36	5.58	5.75	5.88	6.30	6.71
19	4.60	4.90	5.07	5.22	5.31	5.65	5.95
24	4.40	4.65	4.81	4.94	5.01	5.30	5.57
29	4.25	4.51	4.65	4.76	4.85	5.10	5.34
39	4.11	4.33	4.48	4.57	4.64	4.86	5.06
59	3.97	4.17	4.30	4.38	4.44	4.64	4.83
∞	3.72	3.89	3.99	4.06	4.11	4.27	4.42

Table C10. One-sided simultaneous prediction intervals of 0.9999-content for normal sequences with known mean and unknown variance. Approximate values of $\beta_{\sigma}^{N}(0.9999,n,N)$ from equation (II.132a).

n \ N	1	2	3	4	5	10	20
1	13.68	(24.71)	******	(45.36)	******	(103.01)	(193.44)
2	2.99	(3.85)	******	(5.02)	******	(7.23)	(9.60)
3	1.95	(2.27)	(2.48)	(2.66)	(2.79)	(3.31)	(3.93)
4	1.61	(1.79)	(1.90)	(1.99)	(2.07)	(2.32)	(2.62)
5	1.45	1.55	1.62	1.66	1.70	1.82	(2.09)
6	1.35	1.43	1.48	1.51	1.54	1.63	(1.81)
7	1.29	1.35	1.39	1.42	1.44	1.51	(1.64)
8	1.25	1.30	1.33	1.35	1.37	1.43	(1.53)
9	1.21	1.26	1.29	1.30	1.32	1.37	(1.45)
10	1.19	1.23	1.25	1.27	1.28	1.32	(1.39)
11	1.17	1.20	1.22	1.24	1.25	1.29	(1.35)
14	1.13	1.15	1.17	1.18	1.19	1.22	(1.26)
19	1.09	1.11	1.12	1.13	1.13	1.15	(1.18)
24	1.07	1.08	1.09	1.10	1.10	1.12	(1.14)
29	1.06	1.07	1.08	1.08	1.08	1.10	(1.11)
39	1.04	(1.05)	(1.06)	(1.06)	(1.06)	(1.07)	(1.08)
59	1.03	(1.03)	(1.04)	(1.04)	(1.04)	(1.05)	(1.05)
∞	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table C11. One-sided simultaneous prediction intervals of 0.99-content for normal sequences with known mean and unknown variance. Values of $r_{\sigma}^{N}(0.99,n,N)$. In parenthesis are approximate values from using equation (II.132a).

n \ N	1	2	3	4	5	10	20
1	103.01	193.44	******	******	******	855.88	1636.1
2	7.23	9.60	******	******	******	19.01	25.70
3	3.31	3.93	******	******	******	5.97	7.20
4	2.32	2.62	******	******	******	3.50	3.99
5	1.91	2.09	2.20	2.29	2.36	2.60	2.87
6	1.69	1.81	1.89	1.95	2.00	2.16	2.33
7	1.55	1.64	1.70	1.75	1.78	1.90	2.03
8	1.46	1.53	1.58	1.62	1.65	1.73	1.83
9	1.39	1.45	1.49	1.52	1.54	1.62	1.69
10	1.34	1.39	1.43	1.45	1.47	1.53	1.60
11	1.30	1.35	1.37	1.40	1.41	1.47	1.52
14	1.23	1.26	1.26	1.27	1.29	1.34	1.38
19	1.16	1.18	1.19	1.20	1.21	1.24	1.26
24	1.12	1.14	1.15	1.15	1.16	1.18	1.20
29	1.10	1.11	1.13	1.13	1.13	1.14	1.16
39	1.07	1.08	1.09	1.09	1.10	1.11	1.11
59	1.05	1.05	1.06	1.06	1.06	1.07	1.07
∞	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table C12. One-sided simultaneous prediction intervals of 0.999-content for normal sequences with known mean and unknown variance. Approximate values of $f_{\sigma}^{N}(0.999,n,N)$.

n \ N	1	2	3	4	5	10	20
1	855.88	1636.13	******	******	******	7463.3	14412.9
2	19.01	25.70	******	******	******	52.52	71.59
3	5.97	7.20	******	******	******	11.23	13.67
4	3.50	3.99	******	******	******	5.47	6.29
5	2.60	2.87	******	******	******	3.64	4.05
6	2.16	2.33	******	******	******	2.82	3.07
7	1.90	2.03	2.10	2.17	2.20	2.37	2.54
8	1.73	1.83	1.89	1.94	1.97	2.09	2.21
9	1.62	1.69	1.74	1.78	1.81	1.90	2.00
10	1.53	1.60	1.64	1.67	1.69	1.77	1.85
11	1.47	1.52	1.55	1.58	1.60	1.67	1.73
14	1.34	1.38	1.40	1.42	1.43	1.48	1.52
19	1.24	1.26	1.27	1.29	1.29	1.32	1.35
24	1.18	1.20	1.21	1.22	1.22	1.24	1.26
29	1.14	1.16	1.17	1.17	1.18	1.20	1.21
39	1.11	1.11	1.12	1.13	1.13	1.14	1.15
59	1.07	1.07	1.08	1.08	1.08	1.09	1.09
∞	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table C13. One-sided simultaneous prediction intervals of 0.9999-content for normal sequences with known mean and unknown variance. Approximate values of $r_{\sigma}^{N}(0.9999, n, N)$.

N \ n	1	2	3	4	7	9	∞
1	0.961	0.979	0.985	0.987	0.990	0.991	0.9938
2	0.931 (0.924)	0.960 (0.959)	0.970	0.975	0.981	0.982	0.9876
3	0.906 (0.889)	0.943 (0.939)	0.957 (0.955)	0.962	0.971	0.974	0.9815
4	0.884 (0.854)	0.928 (0.920)	0.944 (0.940)	0.951 (0.950)	0.962	0.965	0.9754
5	0.866 (0.821)	0.914 (0.901)	0.931 (0.926)	0.939 (0.938)	0.953	0.957	0.9693
6	0.850 (0.790)	0.901 (0.882)	0.920 (0.912)	0.928 (0.926)	0.944 (0.943)	0.948	0.9633
7	0.834 (0.759)	0.888 (0.864)	0.909 (0.898)	0.918 (0.915)	0.936 (0.934)	0.940	0.9574
8	0.821 (0.730)	0.876 (0.846)	0.898 (0.885)	0.909 (0.903)	0.927 (0.925)	0.932 (0.931)	0.9514
9	0.808 (0.702)	0.864 (0.829)	0.887 (0.871)	0.900 (0.892)	0.918 (0.916)	0.924 (0.923)	0.9455
10	0.797 (0.675)	0.853 (0.812)	0.877 (0.858)	0.891 (0.880)	0.910 (0.907)	0.916 (0.915)	0.9397
11	0.786 (0.649)	0.843 (0.795)	0.868 (0.845)	0.881 (0.869)	0.902 (0.899)	0.909 (0.907)	0.9338
12	0.776 (0.624)	0.833 (0.779)	0.859 (0.832)	0.872 (0.858)	0.894 (0.890)	0.902 (0.899)	0.9281

Table C14. Probability content P of the interval (II.133a) for $\beta_{\mu}^{N}(\cdot) = 2.50$ and for several values of n and N. In parenthesis, approximate values from equation (II.134). (Where approximate values are missing they coincide with the exact ones up to the precision of the table.)

N \ n	1	2	3	4	7	9	∞
1	0.9933	0.9978	0.9988	0.9991	0.9995	0.9995	0.99977
2	0.9873 (0.9867)	0.9957	0.9976	0.9983	0.9989	0.9991	0.99954
3	0.9829 (0.9802)	0.9936	0.9963	0.9974	0.9983	0.9987	0.99931
4	0.9768 (0.9736)	0.9915	0.9951	0.9965	0.9978	0.9982	0.99908
5	0.9723 (0.9671)	0.9895 (0.9893)	0.9940 (0.9939)	0.9957	0.9973	0.9978	0.99885
6	0.9680 (0.9607)	0.9877 (0.9872)	0.9928 (0.9927)	0.9948	0.9968	0.9973	0.99862
7	0.9640 (0.9543)	0.9858 (0.9851)	0.9917 (0.9915)	0.9939	0.9963	0.9969	0.99839
8	0.9602 (0.9479)	0.9839 (0.9830)	0.9905 (0.9903)	0.9931	0.9958	0.9964	0.99816
9	0.9564 (0.9416)	0.9820 (0.9809)	0.9894 (0.9891)	0.9922 (0.9921)	0.9953	0.9960	0.99793
10	0.9528 (0.9354)	0.9802 (0.9788)	0.9882 (0.9879)	0.9914 (0.9913)	0.9948 (0.9947)	0.9956 (0.9955)	0.99770
11	0.9497 (0.9291)	0.9786 (0.9767)	0.9872 (0.9867)	0.9907 (0.9905)	0.9943 (0.9942)	0.9952 (0.9951)	0.99747
12	0.9467 (0.9229)	0.9769 (0.9746)	0.9861 (0.9855)	0.9898 (0.9896)	0.9938 (0.9936)	0.9947 (0.9946)	0.99724

Table C15. Probability content P of the interval (II.133a) for $\beta_{\mu}^{N}(\cdot) = 3.50$ and for several values of n and N. In parenthesis, approximate values from equation (II.134). (Where approximate values are missing they coincide with the exact ones up to the precision of the table.)

P = 0.990 A/B = 1.0

N \ n	10	20	30	40	60
1	1.438	1.168	1.099	1.068	1.040
2	1.408	1.158	1.094	1.064	1.038
3	1.394	1.153	1.091	1.063	1.037
4	1.385	1.150	1.089	1.062	1.037
5	1.378	1.148	1.088	1.061	1.036
6	1.373	1.146	1.087	1.060	1.036
7	1.369	1.145	1.087	1.060	1.036
8	1.365	1.143	1.086	1.060	1.036
9	1.362	1.142	1.085	1.059	1.035
10	1.360	1.142	1.085	1.059	1.035
15	1.350	1.138	1.083	1.058	1.035
20	1.344	1.136	1.082	1.057	1.034
25	1.340	1.135	1.081	1.057	1.034
30	1.336	1.134	1.080	1.056	1.034
40	1.331	1.132	1.080	1.056	1.034
60	1.324	1.130	1.078	1.055	1.033

P = 0.990 A/B = 2.0

N \ n	10	20	30	40	60
1	1.372	1.142	1.084	1.057	1.034
2	1.352	1.136	1.081	1.056	1.033
3	1.343	1.133	1.079	1.055	1.032
4	1.336	1.131	1.078	1.054	1.032
5	1.332	1.130	1.077	1.054	1.032
6	1.328	1.129	1.077	1.053	1.032
7	1.325	1.128	1.076	1.053	1.032
8	1.323	1.127	1.076	1.053	1.031
9	1.321	1.126	1.076	1.052	1.031
10	1.319	1.126	1.075	1.052	1.031
15	1.313	1.124	1.074	1.052	1.031
20	1.308	1.122	1.073	1.051	1.031
25	1.305	1.121	1.073	1.051	1.031
30	1.303	1.120	1.072	1.051	1.031
40	1.299	1.119	1.072	1.050	1.030
60	1.294	1.118	1.071	1.050	1.030

Table C16. Left-hand simultaneous prediction interval when sampling is from the Extreme Type I distribution (II.136) with a and b unknown. Each table collects penalty factors $r^{E1}_{a,b}(P,n,N)$ for fixed P and a/b ratios. Other values of P and a/b are considered in the continuation of the table.

P = 0.990 A/B = 3.0

N \ n	10	20	30	40	60
1	1.323	1.124	1.073	1.050	1.029
2	1.310	1.120	1.071	1.049	1.029
3	1.303	1.118	1.070	1.048	1.029
4	1.299	1.117	1.070	1.048	1.028
5	1.296	1.116	1.069	1.048	1.028
6	1.293	1.115	1.069	1.048	1.028
7	1.291	1.114	1.068	1.047	1.028
8	1.290	1.114	1.068	1.047	1.028
9	1.288	1.113	1.068	1.047	1.028
10	1.287	1.113	1.068	1.047	1.028
15	1.282	1.112	1.067	1.047	1.028
20	1.279	1.111	1.066	1.046	1.028
25	1.277	1.110	1.066	1.046	1.028
30	1.275	1.109	1.066	1.046	1.028
40	1.272	1.109	1.065	1.046	1.028
60	1.269	1.108	1.065	1.046	1.028

P = 0.990 A/B = 4.0

N \ n	10	20	30	40	60
1	1.285	1.109	1.065	1.044	1.026
2	1.277	1.107	1.064	1.044	1.026
3	1.272	1.106	1.063	1.043	1.026
4	1.269	1.105	1.063	1.043	1.026
5	1.267	1.104	1.062	1.043	1.026
6	1.265	1.104	1.062	1.043	1.026
7	1.264	1.103	1.062	1.043	1.026
8	1.263	1.103	1.062	1.043	1.026
9	1.261	1.103	1.062	1.043	1.026
10	1.261	1.103	1.061	1.043	1.026
15	1.257	1.102	1.061	1.042	1.026
20	1.255	1.101	1.061	1.042	1.025
25	1.253	1.101	1.061	1.042	1.025
30	1.252	1.100	1.060	1.042	1.025
40	1.250	1.100	1.060	1.042	1.025
60	1.248	1.099	1.060	1.042	1.025

Table C16 (cont.).

P = 0.990 A/B = 5.0

N \ n	10	20	30	40	60
1	1.256	1.098	1.058	1.040	1.023
2	1.250	1.097	1.057	1.039	1.023
3	1.247	1.096	1.057	1.039	1.023
4	1.245	1.095	1.057	1.039	1.023
5	1.243	1.095	1.057	1.039	1.023
6	1.242	1.095	1.057	1.039	1.023
7	1.241	1.094	1.057	1.039	1.023
8	1.240	1.094	1.056	1.039	1.023
9	1.239	1.094	1.056	1.039	1.023
10	1.239	1.094	1.056	1.039	1.023
15	1.236	1.093	1.056	1.039	1.023
20	1.235	1.093	1.056	1.039	1.023
25	1.234	1.093	1.056	1.039	1.023
30	1.233	1.093	1.056	1.039	1.023
40	1.231	1.092	1.056	1.039	1.024
60	1.229	1.092	1.055	1.039	1.024

P = 0.990 A/B = 6.0

N \ n	10	20	30	40	60
1	1.231	1.089	1.053	1.036	1.021
2	1.228	1.088	1.052	1.036	1.021
3	1.226	1.088	1.052	1.036	1.021
4	1.224	1.087	1.052	1.036	1.021
5	1.223	1.087	1.052	1.036	1.021
6	1.222	1.087	1.052	1.036	1.021
7	1.222	1.087	1.052	1.036	1.021
8	1.221	1.087	1.052	1.036	1.022
9	1.221	1.087	1.052	1.036	1.022
10	1.220	1.087	1.052	1.036	1.022
15	1.219	1.086	1.052	1.036	1.022
20	1.218	1.086	1.052	1.036	1.022
25	1.217	1.086	1.052	1.036	1.022
30	1.216	1.086	1.052	1.036	1.022
40	1.215	1.086	1.052	1.036	1.022
60	1.214	1.086	1.052	1.036	1.022

Table C16 (cont.).

P = 0.990 A/B = 7.0

N \ n	10	20	30	40	60
1	1.212	1.081	1.048	1.033	1.019
2	1.209	1.081	1.048	1.033	1.019
3	1.208	1.081	1.048	1.033	1.020
4	1.207	1.081	1.048	1.033	1.020
5	1.206	1.081	1.048	1.033	1.020
6	1.206	1.081	1.048	1.033	1.020
7	1.205	1.081	1.048	1.033	1.020
8	1.205	1.081	1.048	1.033	1.020
9	1.205	1.080	1.048	1.033	1.020
10	1.204	1.080	1.048	1.033	1.020
15	1.203	1.080	1.048	1.034	1.020
20	1.203	1.080	1.048	1.034	1.020
25	1.202	1.080	1.048	1.034	1.020
30	1.202	1.080	1.048	1.034	1.020
40	1.201	1.080	1.048	1.034	1.020
60	1.200	1.080	1.048	1.034	1.021

P = 0.990 A/B = 8.0

N \ n	10	20	30	40	60
1	1.195	1.075	1.044	1.030	1.018
2	1.193	1.075	1.044	1.030	1.018
3	1.193	1.075	1.045	1.031	1.018
4	1.192	1.075	1.045	1.031	1.018
5	1.192	1.075	1.045	1.031	1.018
6	1.191	1.075	1.045	1.031	1.018
7	1.191	1.075	1.045	1.031	1.019
8	1.191	1.075	1.045	1.031	1.019
9	1.191	1.075	1.045	1.031	1.019
10	1.191	1.075	1.045	1.031	1.019
15	1.190	1.075	1.045	1.031	1.019
20	1.190	1.075	1.045	1.031	1.019
25	1.189	1.075	1.045	1.032	1.019
30	1.189	1.075	1.045	1.032	1.019
40	1.189	1.075	1.045	1.032	1.019
60	1.188	1.075	1.045	1.032	1.019

Table C16 (cont.).

P = 0.990 A/B = 9.0

N \ n	10	20	30	40	60
1	1.180	1.069	1.041	1.028	1.016
2	1.180	1.070	1.041	1.028	1.017
3	1.180	1.070	1.042	1.029	1.017
4	1.179	1.070	1.042	1.029	1.017
5	1.179	1.070	1.042	1.029	1.017
6	1.179	1.070	1.042	1.029	1.017
7	1.179	1.070	1.042	1.029	1.017
8	1.179	1.070	1.042	1.029	1.017
9	1.179	1.070	1.042	1.029	1.017
10	1.179	1.070	1.042	1.029	1.018
15	1.178	1.071	1.042	1.029	1.018
20	1.178	1.071	1.042	1.030	1.018
25	1.178	1.071	1.043	1.030	1.018
30	1.178	1.071	1.043	1.030	1.018
40	1.178	1.071	1.043	1.030	1.018
60	1.178	1.071	1.043	1.030	1.018

P = 0.990 A/B = 10.0

N \ n	10	20	30	40	60
1	1.168	1.064	1.038	1.026	1.015
2	1.168	1.065	1.039	1.026	1.016
3	1.168	1.065	1.039	1.027	1.016
4	1.168	1.066	1.039	1.027	1.016
5	1.168	1.066	1.039	1.027	1.016
6	1.168	1.066	1.039	1.027	1.016
7	1.168	1.066	1.039	1.027	1.016
8	1.168	1.066	1.040	1.027	1.016
9	1.168	1.066	1.040	1.027	1.016
10	1.168	1.066	1.040	1.028	1.016
15	1.168	1.066	1.040	1.028	1.017
20	1.168	1.067	1.040	1.028	1.017
25	1.168	1.067	1.040	1.028	1.017
30	1.168	1.067	1.040	1.028	1.017
40	1.168	1.067	1.040	1.028	1.017
60	1.168	1.067	1.041	1.028	1.017

Table C16 (cont.).

P = 0.995 A/B = 1.0

N \ n	20	30	40	60
1	1.219	1.137	1.096	1.059
2	1.205	1.128	1.090	1.056
3	1.198	1.124	1.087	1.054
4	1.194	1.121	1.085	1.053
5	1.191	1.119	1.084	1.052
6	1.188	1.118	1.083	1.051
7	1.186	1.116	1.082	1.051
8	1.184	1.115	1.081	1.051
9	1.183	1.114	1.081	1.050
10	1.181	1.114	1.080	1.050
15	1.177	1.111	1.078	1.049
20	1.174	1.109	1.077	1.048
25	1.171	1.107	1.076	1.047
30	1.170	1.106	1.075	1.047
40	1.167	1.105	1.074	1.046
60	1.164	1.103	1.073	1.046

P = 0.995 A/B = 2.0

N \ n	20	30	40	60
1	1.189	1.118	1.083	1.051
2	1.180	1.112	1.079	1.049
3	1.175	1.109	1.077	1.048
4	1.172	1.107	1.075	1.047
5	1.169	1.106	1.074	1.046
6	1.167	1.105	1.074	1.046
7	1.166	1.104	1.073	1.045
8	1.165	1.103	1.073	1.045
9	1.163	1.102	1.072	1.045
10	1.163	1.102	1.072	1.045
15	1.159	1.100	1.070	1.044
20	1.157	1.098	1.069	1.043
25	1.155	1.097	1.069	1.043
30	1.154	1.096	1.068	1.043
40	1.152	1.095	1.067	1.042
60	1.149	1.094	1.066	1.042

Table C16 (cont.).

P = 0.995 A/B = 3.0

N \ n	20	30	40	60
1	1.167	1.104	1.073	1.045
2	1.160	1.100	1.070	1.043
3	1.156	1.098	1.068	1.042
4	1.154	1.096	1.068	1.042
5	1.152	1.095	1.067	1.042
6	1.151	1.094	1.066	1.041
7	1.150	1.094	1.066	1.041
8	1.149	1.093	1.066	1.041
9	1.148	1.093	1.065	1.041
10	1.147	1.092	1.065	1.040
15	1.145	1.091	1.064	1.040
20	1.143	1.090	1.063	1.039
25	1.142	1.089	1.063	1.039
30	1.141	1.088	1.062	1.039
40	1.139	1.087	1.062	1.039
60	1.137	1.086	1.061	1.038

P = 0.995 A/B = 4.0

N \ n	20	30	40	60
1	1.149	1.093	1.065	1.040
2	1.144	1.090	1.063	1.039
3	1.141	1.088	1.062	1.038
4	1.139	1.087	1.061	1.038
5	1.138	1.086	1.061	1.038
6	1.137	1.086	1.060	1.038
7	1.136	1.085	1.060	1.037
8	1.136	1.085	1.060	1.037
9	1.135	1.085	1.060	1.037
10	1.135	1.084	1.059	1.037
15	1.133	1.083	1.059	1.037
20	1.131	1.082	1.058	1.036
25	1.130	1.082	1.058	1.036
30	1.130	1.081	1.057	1.036
40	1.128	1.081	1.057	1.036
60	1.127	1.080	1.056	1.035

Table C16 (cont.).

P = 0.995 A/B = 5.0

N \ n	20	30	40	60
1	1.134	1.084	1.058	1.036
2	1.131	1.082	1.057	1.035
3	1.129	1.081	1.056	1.035
4	1.127	1.080	1.056	1.035
5	1.127	1.079	1.056	1.035
6	1.126	1.079	1.055	1.034
7	1.125	1.078	1.055	1.034
8	1.125	1.078	1.055	1.034
9	1.124	1.078	1.055	1.034
10	1.124	1.078	1.055	1.034
15	1.122	1.077	1.054	1.034
20	1.121	1.076	1.054	1.034
25	1.121	1.076	1.053	1.033
30	1.120	1.075	1.053	1.033
40	1.119	1.075	1.053	1.033
60	1.118	1.074	1.052	1.033

P = 0.995 A/B = 6.0

N \ n	20	30	40	60
1	1.122	1.076	1.053	1.033
2	1.120	1.075	1.052	1.032
3	1.118	1.074	1.052	1.032
4	1.117	1.074	1.052	1.032
5	1.117	1.073	1.051	1.032
6	1.116	1.073	1.051	1.032
7	1.116	1.072	1.051	1.032
8	1.115	1.072	1.051	1.032
9	1.115	1.072	1.051	1.032
10	1.115	1.072	1.051	1.032
15	1.114	1.071	1.050	1.031
20	1.113	1.071	1.050	1.031
25	1.112	1.070	1.050	1.031
30	1.112	1.070	1.050	1.031
40	1.111	1.070	1.049	1.031
60	1.110	1.069	1.049	1.031

Table C16 (cont.).

P = 0.995 A/B = 7.0

N \ n	20	30	40	60
1	1.112	1.070	1.049	1.030
2	1.111	1.069	1.048	1.030
3	1.110	1.069	1.048	1.030
4	1.109	1.068	1.048	1.030
5	1.108	1.068	1.048	1.030
6	1.108	1.068	1.048	1.030
7	1.108	1.067	1.047	1.029
8	1.107	1.067	1.047	1.029
9	1.107	1.067	1.047	1.029
10	1.107	1.067	1.047	1.029
15	1.106	1.066	1.047	1.029
20	1.106	1.066	1.047	1.029
25	1.105	1.066	1.047	1.029
30	1.105	1.066	1.046	1.029
40	1.104	1.065	1.046	1.029
60	1.104	1.065	1.046	1.029

P = 0.995 A/B = 8.0

N \ n	20	30	40	60
1	1.104	1.065	1.045	1.028
2	1.103	1.064	1.045	1.028
3	1.102	1.064	1.045	1.028
4	1.101	1.063	1.045	1.028
5	1.101	1.063	1.044	1.028
6	1.101	1.063	1.044	1.028
7	1.101	1.063	1.044	1.028
8	1.100	1.063	1.044	1.028
9	1.100	1.063	1.044	1.028
10	1.100	1.063	1.044	1.027
15	1.099	1.062	1.044	1.027
20	1.099	1.062	1.044	1.027
25	1.099	1.062	1.044	1.027
30	1.099	1.062	1.044	1.027
40	1.098	1.062	1.044	1.027
60	1.098	1.061	1.043	1.027

Table C16 (cont.).

P = 0.995 A/B = 9.0

N \ n	20	30	40	60
1	1.097	1.060	1.042	1.026
2	1.096	1.060	1.042	1.026
3	1.095	1.060	1.042	1.026
4	1.095	1.059	1.042	1.026
5	1.095	1.059	1.042	1.026
6	1.095	1.059	1.042	1.026
7	1.094	1.059	1.042	1.026
8	1.094	1.059	1.042	1.026
9	1.094	1.059	1.041	1.026
10	1.094	1.059	1.041	1.026
15	1.094	1.059	1.041	1.026
20	1.093	1.059	1.041	1.026
25	1.093	1.058	1.041	1.026
30	1.093	1.058	1.041	1.026
40	1.093	1.058	1.041	1.026
60	1.092	1.058	1.041	1.026

P = 0.995 A/B = 10.0

N \ n	20	30	40	60
1	1.090	1.056	1.039	1.024
2	1.090	1.056	1.039	1.024
3	1.089	1.056	1.039	1.024
4	1.089	1.056	1.039	1.024
5	1.089	1.056	1.039	1.024
6	1.089	1.056	1.039	1.024
7	1.089	1.056	1.039	1.024
8	1.089	1.056	1.039	1.024
9	1.089	1.056	1.039	1.024
10	1.089	1.056	1.039	1.024
15	1.088	1.055	1.039	1.024
20	1.088	1.055	1.039	1.024
25	1.088	1.055	1.039	1.024
30	1.088	1.055	1.039	1.024
40	1.088	1.055	1.039	1.024
60	1.088	1.055	1.039	1.024

Table C16 (cont.).

	$\Gamma(k,\lambda)$					LN_{μ_Z}	LN_{σ_Z}		$LN_{\mu,\sigma}$	
n	k=1 (EX)	k=2	k=3	k=4	k=5	V_Y=0.25	$\hat{V}_Y$=0.1	$\hat{V}_Y$=0.3	$\hat{V}_Y$=0.1	$\hat{V}_Y$=0.3
1	46.06	8.25	4.80	3.75	3.16	1.66	*****	*****	*****	*****
2	5.88	2.85	2.27	2.02	1.87	1.34	5.86	16.00	*****	*****
3	3.15	*****	*****	*****	*****	1.23	2.73	4.15	7.40	25.66
4	2.34	*****	*****	*****	*****	1.17	2.01	2.60	3.17	5.30
5	1.96	1.53	1.41	1.34	1.30	1.14	1.70	2.05	2.25	3.09
6	1.75	*****	*****	*****	*****	1.12	1.54	1.77	1.87	2.35
7	1.61	*****	*****	*****	*****	1.10	1.43	1.61	1.67	1.99
8	1.52	*****	*****	*****	*****	1.09	1.36	1.50	1.54	1.78
9	1.45	*****	*****	*****	*****	1.08	1.31	1.43	1.45	1.64
10	1.39	1.24	1.19	1.16	1.14	1.07	1.27	1.37	1.39	1.54
15	1.25	*****	*****	*****	*****	1.05	1.17	1.23	1.23	1.31
20	1.18	1.10	*****	*****	*****	1.03	1.12	1.16	1.16	1.22
25	1.14	*****	*****	*****	*****	*****	*****	*****	*****	*****
30	1.12	*****	*****	*****	*****	1.02	1.08	1.10	1.10	1.13
40	1.09	*****	*****	*****	*****	1.02	1.06	1.08	1.07	1.10
50	1.07	*****	*****	*****	*****	*****	*****	*****	*****	*****
60	*****	*****	*****	*****	*****	1.01	1.04	1.05	1.05	1.06

Table C17. Penalty factors for left-hand simultaneous intervals of P=0.99-content for the next N=2 observations. Comparison is between Gamma sequences with given shape factor k=1(1)5 and lognormal sequences with known or estimated coefficient of variation V_Y. For lognormal sequences the associated normal population has unknown mean (first lognormal column), unknown variance (following two columns) or both mean and variance unknown (last two columns). Other combinations of P and N are considered in the continuation of the table.

	$\Gamma(k,\lambda)$					LN_{μ_Z}	LN_{σ_Z}		$LN_{\mu,\sigma}$	
n	k=1 (EX)	k=2	k=3	k=4	k=5	V_Y=0.25	$\hat{V}_Y$=0.1	$\hat{V}_Y$=0.3	$\hat{V}_Y$=0.1	$\hat{V}_Y$=0.3
1	168.93	14.39	6.86	5.08	3.96	1.71	*****	*****	*****	*****
2	10.20	3.70	2.69	2.29	2.08	1.36	23.22	*****	*****	*****
3	4.40	*****	*****	*****	*****	1.24	4.96	13.26	33.97	*****
4	2.96	*****	*****	*****	*****	1.18	2.92	5.00	5.95	19.21
5	2.35	1.69	1.50	1.41	1.36	1.15	2.23	3.21	3.33	6.29
6	2.03	*****	*****	*****	*****	1.12	1.90	2.49	2.48	3.79
7	1.82	*****	*****	*****	*****	1.11	1.70	2.11	2.07	2.85
8	1.69	*****	*****	*****	*****	1.09	1.58	1.89	1.84	2.37
9	1.59	*****	*****	*****	*****	1.08	1.49	1.74	1.69	2.09
10	1.52	1.30	1.23	1.19	1.17	1.07	1.42	1.63	1.58	1.89
15	1.32	*****	*****	*****	*****	1.05	1.25	1.36	1.33	1.47
20	1.23	1.14	*****	*****	*****	1.04	1.18	1.25	1.23	1.32
25	1.18	*****	*****	*****	*****	*****	*****	*****	*****	*****
30	1.15	*****	*****	*****	*****	1.03	1.11	1.16	1.14	1.19
40	1.11	*****	*****	*****	*****	1.02	1.08	1.11	1.10	1.14
50	1.08	*****	*****	*****	*****	*****	*****	*****	*****	*****
60	*****	*****	*****	*****	*****	1.01	1.06	1.07	1.07	1.09

Table C17 (cont.). P = 0.99; N = 10.

	← $\Gamma(k,\lambda)$ →					← LN_{μ_Z} →	← LN_{σ_Z} →		← $LN_{\mu,\sigma}$ →	
n	k=1 (EX)	k=2	k=3	k=4	k=5	V_Y=0.25	$\hat{V}_Y$=0.1	$\hat{V}_Y$=0.3	$\hat{V}_Y$=0.1	$\hat{V}_Y$=0.3
1	302.69	18.73	8.06	5.20	4.34	1.74	*****	*****	*****	*****
2	13.10	4.15	2.89	2.42	2.17	1.37	58.33	*****	*****	*****
3	5.12	*****	*****	*****	*****	1.25	6.84	27.40	96.66	*****
4	3.29	*****	*****	*****	*****	1.19	3.53	7.25	8.40	43.25
5	2.56	1.75	1.54	1.45	1.39	1.15	2.55	4.08	4.06	9.42
6	2.17	*****	*****	*****	*****	1.12	2.10	2.97	2.84	4.92
7	1.93	*****	*****	*****	*****	1.11	1.85	2.43	2.30	3.45
8	1.77	*****	*****	*****	*****	1.09	1.69	2.11	2.00	2.74
9	1.66	*****	*****	*****	*****	1.08	1.58	1.91	1.81	2.35
10	1.57	1.33	1.25	1.20	1.18	1.08	1.50	1.77	1.68	2.09
15	1.35	*****	*****	*****	*****	1.05	1.29	1.43	1.37	1.56
20	1.25	1.15	*****	*****	*****	1.04	1.21	1.30	1.26	1.37
25	1.19	*****	*****	*****	*****	*****	*****	*****	*****	*****
30	1.16	*****	*****	*****	*****	1.03	1.13	1.18	1.15	1.22
40	1.12	*****	*****	*****	*****	1.02	1.10	1.13	1.11	1.15
50	1.09	*****	*****	*****	*****	*****	*****	*****	*****	*****
60	*****	*****	*****	*****	*****	1.01	1.06	1.08	1.07	1.10

Table C17 (cont.). P = 0.999; N = 2.

	← Γ(k,λ) →					← LN_{μ_Z} →	← LN_{σ_Z} →		← $LN_{\mu,\sigma}$ →	
n	k=1 (EX)	k=2	k=3	k=4	k=5	V_Y=0.25	$\hat{V}_Y$=0.1	$\hat{V}_Y$=0.3	$\hat{V}_Y$=0.1	$\hat{V}_Y$=0.3
1	*****	35.24	11.39	7.14	5.43	1.78	*****	*****	*****	*****
2	23.99	5.50	3.43	2.76	2.42	1.39	*****	*****	*****	*****
3	7.39	*****	*****	*****	*****	1.26	18.36	*****	*****	*****
4	4.26	*****	*****	*****	*****	1.19	6.00	23.14	24.51	*****
5	3.11	1.92	1.65	1.52	1.44	1.15	3.65	6.32	7.09	33.58
6	2.54	*****	*****	*****	*****	1.13	2.75	4.90	4.13	10.58
7	2.20	*****	*****	*****	*****	1.11	2.29	3.57	3.03	5.86
8	1.99	*****	*****	*****	*****	1.10	2.02	2.88	2.49	4.12
9	1.83	*****	*****	*****	*****	1.09	1.84	2.48	2.17	3.25
10	1.72	1.39	1.29	1.23	1.21	1.08	1.71	2.21	1.96	2.75
15	1.43	*****	*****	*****	*****	1.05	1.44	1.70	1.51	1.81
20	1.30	1.17	*****	*****	*****	1.04	1.30	1.45	1.34	1.52
25	1.23	*****	*****	*****	*****	1.03	1.23	1.34	1.26	1.39
30	1.19	*****	*****	*****	*****	1.03	1.18	1.26	1.20	1.30
40	1.14	*****	*****	*****	*****	1.02	1.13	1.19	1.15	1.21
50	1.11	*****	*****	*****	*****	*****	*****	*****	*****	*****
60	*****	*****	*****	*****	*****	1.01	1.08	1.12	1.09	1.13

Table C17 (cont.). P = 0.999; N = 10.