LINEAR STATISTICAL CALIBRATION

by

Faqir Muhammad

A thesis Submitted for the

Degree of Doctor of Philosophy

in the Department of Statistics

University of Glasgow

October 1987

ProQuest Number: 13834280

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 13834280

Published by ProQuest LLC (2019). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Thesis
7756
copy 2

GLASGOW
UNIVERSITY
LIBRARY

ACKNOWLEDGEMENTS

First and foremost, I wish to thank my supervisor Mr.A.D.McLaren for his continuous guidance, encouragement and stimulating discussions.

I am also thankful to the head of the Statistics Department, Professor D.M. Titterington and all other members of the staff for providing friendly environment and statistical education.

I would like to thank the Government of Pakistan, Ministry of Education, Embassy of Pakistan and University of Agriculture Faisalabad for grant and study leave.

Finally I am thankful to my family for their patience and encouragement.

CONTENTS

SUMMARY

Much has been said about the classical and the inverse methods of calibration for the univariate and to some extent about the multivariate case also in the existing literature, see Brown(1982). We have explored the possibilities of using the best linear predictor both in the univariate and the multivariate situations.

First four chapters deal mainly with the univariate case, chapters five and six deal with the multivariate situation and chapter seven is devoted to the Bayesian version of best linear predictor.

First chapter introduces calibration and discusses different methods of calibration in the univariate situation. Chapter 2 gives a review of the calibration literature for classical, Bayesian and best linear predictor approaches with some comments. Chapter 3 deals with the derivation of the best linear predictor and approximates its unconditional mean squared error by Taylor's series. A simulation study is made to compare the approximated and the simulated values. Chapter 4 starts with the interval estimates and possible aims. Two situations with the known and unknown parameters are studied. Tail probabilities are calculated for different $P(t)$.

Chapter 5 introduces multivariate calibration and reviews the literature. Much attention is focussed on the case when there are q response variables and there is only one explanatory variable p i.e. general q and p = 1. Best linear predictor is derived and its mean squared error in canonical form is studied by simulation. In chapter 6 approximation to mean squared error is obtained by regressing simulated data and the interval estimates are studied.

Chapter 7 gives a Bayesian treatment of the best linear predictor both in the univariate and the multivariate case.

CHAPTER 1

UNIVARIATE CALIBRATION

1.1. INTRODUCTION

The word Calibration is being used in two different contexts in statistical literature

(i) in connection with regression;

(ii) in connection with probability forecasts.

We restrict ourselves to the calibration of first kind, for the second kind see Dawid (1985).

First kind is usually referred to as calibration, inverse regression, inverse prediction or very rarely discrimination as by Lieberman et al. (1967) because of sharing similar features with calibration problem. The only difference between discrimination problem and calibration problem lies in the fact that fixed variable is continuous in calibration while it is a finite set in discrimination.

Williams (1969a) emphasized the need to differentiate between two activities, both being called calibration in statistical literature, and categorized as under,

(a) absolute calibration;

(b) comparative calibration.

In absolute calibration non-standard measurement technique is calibrated against a standard measurement technique whereas in comparative calibration one instrument is calibrated against the other (or possibly others) with neither being standard. Both are conceptually different and lead to different issues in statistical modelling. For more details see Williams (1969a) and Rosenblatt and Spiegleman's discussion to Hunter and Lamboy (1981). We would concentrate on absolute calibration only.

Aitchison and Dunsmore (1975) used the terms, natural calibration and designed calibration to distinguish between the two types of regression experiments regarding the way in which the values of the fixed variable arise.

In natural calibration the values of fixed variable in the experiment occur naturally as if it can be assumed that future values would also arise in the same way as in the past. Thus the regression experiment provides some information about the pattern of observations. Range of defined values is not controlled which may have some effect on the accuracy of estimated calibration curve.

In case of controlled calibration the values of the fixed variable in regression experiment are at fixed prechosen levels preferably such that they give a reasonable cover to the range of values of controlled variable expected in future. This helps to improve the design of experiment.

Brown (1982) termed the calibration as random calibration when both the response and explanatory variables are random. This appears to agree with the idea of natural calibration discussed above. In our opinion this is a prediction problem rather than a calibration problem because both variables are random and the regression of either on the other, for given values, is reasonable.

1.2. DESCRIPTION OF THE PROBLEM

A calibration problem consists of

(a) A regression experiment comprising N pairs (t_i, x_i), $i = 1, 2, \ldots, N$ with t_i fixed and x_i independent random variable;

and

(b) A current or future situation involving a bivariate random variable (T,X) independent of regression experiment where instead of observing a pair (t,x), only the observation x has been made.

There can be two situations

(i) Only one value of X is observed;

(ii) More values X_1, X_2, X_3, ..., X_k (i.e. $k > 1$) are observed, with mean $\bar{X}_f$.

The problem is to estimate t on T corresponding to the observation(s) x on X, from the information provided by regression experiment and the current situation. Two types of estimation on t are required.

(i) Point estimation,

(ii) Interval estimation.

This problem of estimation of t is called calibration problem. It can be more clarified by the following two examples, as discussed by Aitchison and Dunsmore (1975). We will describe it as simple linear calibration if there is only one response variable X and one explanatory variable T and the regression in the experiment is linear. The "regression experiment" in calibration situations as (a) above is sometimes referred to as the "calibration experiment".

Example 1: Measuring Water Content Of Soil Specimens

Water content in agricultural soils is of interest for cultivation purposes. It can be measured by two methods, namely, (a) Laboratory method and (b) On-site method. On-site method is cheaper and quicker as compared to the laboratory method but less accurate. It is recommended that in future on-site method be used. Now the objective is to predict the observation by the laboratory method corresponding to the observation by the on-site method. Let the linear regression between the two methods hold, denoted by the following relation

$$x = \alpha + \beta t + e$$

where x denotes the observation by on-site method and t denotes the observation by laboratory method, e is a random error with zero mean and constant variance $\sigma^2_{x|t}$ (conditionally on t).

A regression experiment is performed to obtain N pairs of observations (t_i,x_i), i = 1, 2, ..., N. In the current situation an observation x is observed by the on-site method and the corresponding observation t by the laboratory method is to be estimated.

The data now consists of (N+1) pairs (t_1,x_1), (t_2,x_2), ..., (t_N,x_N); (.,x) where the dot in the last pair indicates unknown value t to be estimated, x is referred as current observation and the pairs (t_i,x_i), i = 1, 2, ..., N are observations from the regression experiment.

The problem is to make statements about the water content t by laboratory method using the information from the regression experiment and the water content observation x by the on-site method.

Example 2: Antibiotic Assay

Different concentrations of an antibiotic drug applied to an infected medium clear different circular areas and so the diameter of the cleared area may be used to help in estimating the concentration of an antibiotic. It has been observed that average diameter of the area cleared by a given concentration is a linear function of the log concentration level of the drug. Different log concentration levels t_i's of the drug are prepared and the corresponding diameters x_i's of areas cleared are noted. The pairs (t_i,x_i) , i = 1, 2, ..., N are related by the regression model

$$x = \alpha + \beta t + e$$

In the current situation a patient is under treatment and the clearance diameter has been measured. The problem is to infer about the patient's concentration level i.e. we want to estimate the unknown log concentration value t corresponding to the known diameter value x.

As our technique applies to the situations where X and T have linear relationship, Box and Tidwell (1962) and/or Box and Cox (1964)

transformations can be applied to get the desired relationship if necessary. These are demonstrated in chapter 4 on the data of antibiotic assay example taken from Aitchison and Dunsmore (1975).

1.3. ASSUMPTIONS FOR THE REGRESSION EXPERIMENT

Let the pairs (t_i,x_i), $i = 1, 2, \ldots, N$ be related by the linear relationship

$$x_i = \alpha + \beta t_i + e_i$$

which can be written as

$$x_i = \alpha^* + \beta(t_i-\bar{t}) + e_i$$

$$E(x|t_i) = \alpha^* + \beta(t_i-\bar{t})$$

$$VAR(x|t_i) = \sigma^2_{x|t}, \text{ for any } t_i$$

The least squares estimators of α^*, β and $\sigma^2_{x|t}$ based on N pairs of observations are denoted by $\hat{\alpha}^*$, $\hat{\beta}$ and $\hat{\sigma}^2_{x|t}$ where

$$\hat{\alpha}^* = \Sigma x_i/N$$

$$\hat{\beta} = S_{XT}/S_{TT}$$

$$\hat{\sigma}^2_{x|t} = \Sigma[x_i - \hat{\alpha}^* - \hat{\beta}(t_i-\bar{t})]^2/(N-2)$$

$S_{XT} = \Sigma(x_i - \bar{x})(t_i - \bar{t})$ and $S_{TT} = \Sigma(t_i - \bar{t})^2$.

Assuming $P(x|t)$ is normal, the estimators $\hat{\alpha}^*$ and $\hat{\beta}$ are distributed according to $N(\alpha^*, \sigma^2_{x|t}/N)$ and $N(\beta, \sigma^2_{x|t}/S_{TT})$ respectively. The unbiased estimator $\hat{\sigma}^2_{x|t}$ is $\sigma^2_{x|t}\ \chi^2/(N-2)$. It is well known that $\hat{\alpha}^*$, $\hat{\beta}$ and $\hat{\sigma}^2_{x|t}$ are mutually independent under the normality assumption.

1.4. ASSUMPTIONS FOR THE CURRENT SITUATION

In the current situation a bivariate random variable (T,X) is thought to be under consideration with distribution $P(t,x)$. The practical situation draws attention to the conditional distribution.

$$P(t|x) = \frac{P(x|t)\ P(t)}{\int P(x|t)\ P(t)\ dt}$$

From the regression experiment $P(x|t)$ is distributed as $N(\alpha+\beta t, \sigma^2_{x|t})$ and about T it is assumed that $E(T) = \mu$ and $VAR(T) = \sigma^2$ are known.

We assume that the regression experiment and the current situation share a common conditional distribution $P(x|t)$. In the regression experiment values of T are fixed while in the current situation T is a random variable, in our situation with known mean μ and variance σ^2.

1.5. THE BEST LINEAR PREDICTOR C+DX

To avoid modelling the distributional shape of $P(t)$ in the bivariate situation, we consider

$$E[T - (C + DX)]^2 \qquad \ldots(1.1)$$

which is minimised by

$$C = E(T) - DE(X)$$

$$D = COV(T,X)/VAR(X)$$

the minimum mean squared error (1.1) being

$$\sigma^2_{t|x} = (1 - \rho^2)VAR(T)$$

Note that $E(T|X)$ is not necessarily linear in X, and $\sigma^2_{t|x}$ is not necessarily a variance. If $E[T - m(X)]^2$ is minimised instead of (1.1) where $m(X)$ is any function of X, then the solution is $m(X) = E(T|X)$, with minimum mean squared error $E(VAR(T|X))$.

It should be clear that within the regression experiment C + DX has no particular role.

1.6. THE BEST LINEAR PREDICTOR $\alpha+\beta t$

In the regression experiment

$$E(X|t_i) = \alpha + \beta t_i$$

$$VAR(X|t_i) = \sigma^2_{x|t}$$

In the current situation which is bivariate, α and β minimise unconditional mean squared error

$$E[X - (\alpha + \beta T)]^2 \qquad \ldots\ldots(1.2)$$

because here also, by the assumption about $P(x|t)$ made in section 1.4,

$$E(X|T) = \alpha + \beta t$$

Arguing as in section 1.5, the mean squared error (1.2) is minimised by

$$\alpha = E(X) - \beta E(T)$$

$$\beta = COV(X,T)/VAR(T)$$

and

$$\sigma^2_{x|t} = (1- \rho^2)VAR(X)$$

where ρ is the correlation coefficient between X and T.

1.7. CALIBRATION METHODS

Two most commonly used methods to estimate t are

(a) classical estimator approach;

(b) inverse estimator approach.

These are the outcome of "inverse regression" and "direct regression". Controversy over the relative worth of the two estimators is not yet clearly resolved because of its philosophical nature. We propose another approach (compare Brown (1979a)).

(c) best linear predictor approach.

These are discussed briefly.

(a) Classical Estimator Approach

The usual regression model $P(x|t)$ is used to predict the value of t for an observed value of x. This is just the regression line in reverse using only $P(x|t)$.

Let the estimated least squares line from the regression experiment be

$$\hat{x} = \hat{\alpha} + \hat{\beta}t$$

then the corresponding equation for predicting t becomes

$$\check{t} = (x - \hat{\alpha})/\hat{\beta}$$

or

$$\check{t} = \bar{t} + S_{TT}/S_{XT}\ (x - \bar{x}) \qquad \ldots(1.3)$$

(b) Inverse Estimator Approach

The line for predicting t is fitted using data from regression experiment (x_i, t_i), $i = 1, 2, \ldots, N$, as if it were truly bivariate data, that is as if $t_1, t_2, \ldots, t_N$ were a random sample from the distribution $P(t)$. Thus the least squares estimated line for predicting t is

$$\tilde{t} = \bar{t} + S_{TX}/S_{XX}\ (x-\bar{x}) \qquad \ldots(1.4)$$

This is a "direct regression" assuming T a random variable whereas it is fixed in the regression experiment.

(c) Best Linear Predictor Approach

It is suggested that t should be estimated using the relation

$$t = C + DX$$

where C and D are functions of bivariate moments as in section 1.5 i.e.

$$C = E(T) - DE(X)$$

$$D = COV(T,X)/VAR(X)$$

Three mathematically independent functions α, β and $\sigma^2_{x|t}$ of the five bivariate moments as in section 1.6 can be estimated from the regression experiment and the moments $E(T) = \mu$ and $VAR(T) = \sigma^2$ are assumed known. So we have numerical values (estimated or exact) of five functions of the five moments. Solving these equations, estimates of the moments are obtained and thus the estimates of C, D and $\sigma^2_{t|x}$ are

$$\hat{C} = \frac{(-\hat{\alpha}/\hat{\beta} + \mu\hat{\tau}^2)}{(1 + \hat{\tau}^2)}$$

$$\hat{D} = \frac{\hat{\beta}^{-1}}{(1 + \hat{\tau}^2)}$$

$$\hat{\sigma}^2_{t|x} = (1 - \hat{\rho}^2)\sigma^2$$

where

$$\hat{\tau}^2 = \hat{\sigma}^2_{x|t}/(\sigma^2\hat{\beta}^2)$$
$$= (1 - \hat{\rho}^2)/\hat{\rho}^2$$

and

$$\hat{\rho}^2 = \hat{\beta}^2\sigma^2 \ (\hat{\beta}^2\sigma^2 + \hat{\sigma}^2_{x|t})^{-1} \qquad \ldots(1.5)$$

$\hat{C}$, $\hat{D}$ and $\hat{\sigma}^2_{t|x}$ are in terms of $\hat{\alpha}$, $\hat{\beta}$, $\hat{\sigma}^2_{x|t}$ from the regression experiment and $E(T) = \mu$, $VAR(T) = \sigma^2$ which are assumed to be known.

Thus the best linear predictor is

$$\hat{t} = \hat{C} + \hat{D}X$$
$$= \frac{(-\hat{\alpha}/\hat{\beta} + \mu\hat{\tau}^2)}{(1 + \hat{\tau}^2)} + \frac{\hat{\beta}^{-1}X}{(1 + \hat{\tau}^2)}$$
$$= \frac{(X - \hat{\alpha})}{\hat{\beta}(1 + \hat{\tau}^2)} + \frac{\hat{\tau}^2\mu}{(1 + \hat{\tau}^2)}$$
$$= \hat{\rho}^2[(X - \hat{\alpha})/\hat{\beta}] + (1 - \hat{\rho}^2)\ \mu \qquad \ldots(1.6)$$

So the best linear predictor is weighted average of the classical estimator and the μ.

If we put $\mu = \bar{t}$ in relation (1.6) then

$$\hat{t} = \bar{t} + \hat{\rho}^2/\hat{\beta}\ (x - \bar{x})$$

and also if we take $\sigma^2 = S_{TT}/(N-2)$, then $\hat{\rho}^2$ in relation (1.5) becomes $S^2_{XT}/(S_{TT}S_{XX})$, $\hat{\beta}$ is S_{XT}/S_{TT}; thus $\hat{t}$ is

$$\hat{t} = \bar{t} + S_{XT}/S_{XX}\ (x - \bar{x})$$

which is the inverse estimator (1.4).

It is of interest that the classical estimator and the inverse estimator are special cases of the best linear predictor. Best linear predictor gives classical estimator when $\hat{\rho}^2 = 1$ or equivalently $\hat{\sigma}^2_{x|t}/(\sigma^2\hat{\beta}^2) = 0$. If we omit the uninteresting

possibilities $\hat{\sigma}^2_{x|t} = 0$ (perfect fit of regression) or $\hat{\beta} = \infty$ (all $t_i = \bar{t}$), we conclude that the best linear predictor coincides with the classical estimator if and only if $\sigma^2 = \infty$. The inverse estimator is obtained with $\mu = \bar{t}$ and $\sigma^2 = S_{TT}/(N-2)$ as shown above.

Thus both the classical estimator and the inverse estimator rest on implicit assumptions about the distribution of T in future. $\sigma^2 = \infty$ has some theoretical appeal, in expressing the idea of complete ignorance about T, but an infinite variance is unrealistic in practice especially since it suggests the need to extrapolate the experimental regression.

The combination of values $\mu = \bar{t}$, $\sigma^2 = S_{TT}/(N-2)$ suggests that the inverse estimator will be satisfactory if the experimental design $\{t_i;\ i = 1, 2, \ldots, N\}$ agrees in first and second moments with the distribution of T in future. Any choice between the two estimators should depend therefore on σ^2, and perhaps on μ. Such choice is impossible without at least some consideration of the distribution $P(t)$.

1.8. SOURCES OF μ and σ^2

In practice μ and σ^2 are not known exactly. Sometimes may be assessed as follows

(a) an assumption implicit in any calibration technique is $t_* \leq T \leq t^*$. Otherwise the experimental regression has to be extrapolated. Bounds for μ and σ^2 can be deduced.

(b) Sometimes a random sample of T's (or more commonly of X's, Tallis (1969)) is available. Natural estimates $\hat{\mu}$ and $\hat{\sigma}^2$ result.

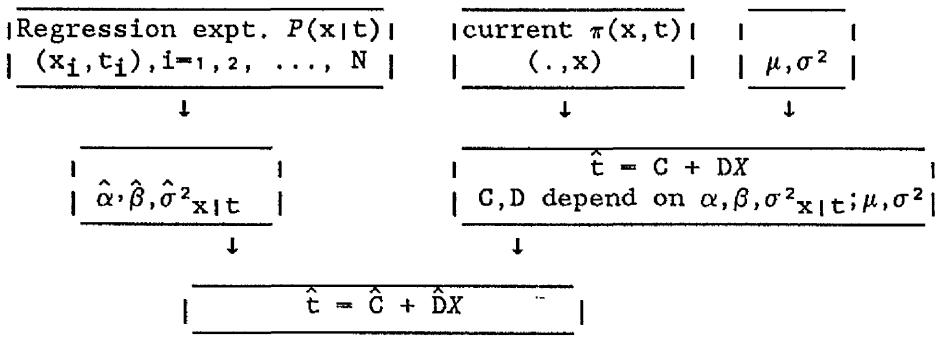
$$E(X) = EE(X|T) = \alpha + \beta\mu$$

$$VAR(X) = VAR(E(X|T)) + E(VAR(X|T))$$

$$= \beta^2\sigma^2 + \sigma^2_{x|t}$$

(c) In the absence of (b) μ and σ^2 may be regarded as the parameters of a subjective probability distribution.

Diagramatically the best linear predictor can be represented as under.



CHAPTER 2

REVIEW OF LITERATURE

The purpose of this part is to review the work done so far in the area of linear and non-linear univariate calibration. Estimation and optimal design aspects are described and some comments have been made wherever necessary.

2.1. CALIBRATION ESTIMATION

In general the theory of calibration can be classified under the following three types of models.

(a) Traditional approach;

(i) classical approach

(ii) inverse approach

(b) Bayesian approach;

(c) Best linear predictor approach.

These are elaborated separately in the following pages.

(a) TRADITIONAL APPROACH

As pointed out above traditional approach comprises the classical estimator and the inverse estimator. In classical approach $p(x|t)$ is assumed $N(\alpha + \beta t, \ \sigma^2_{x|t})$. The point estimate is

$$\check{t} = (x - \hat{\alpha})/\hat{\beta} = \bar{t} + S_{TT}/S_{XT} \ (x - \bar{x})$$

and an interval is available conditional on T i.e.

$$p(\text{Int. will contain } T|T=t) = 0.95$$

If β is not significantly different from zero at 5%, then the above interval for t will be either the entire real line or the complement of an interval in the ordinary sense Williams (1959) and Miller (1966).

In the inverse approach point estimate is derived using least squares, as if T was random in the experiment whereas it is fixed

i.e.

$$\tilde{t} = \bar{t} + S_{TX}/S_{XX} \ (x-\bar{x})$$

There is no theory of interval estimation associated with this estimator in a natural way. Both the classical and the inverse estimator agree when there is a perfect correlation between X and T.

Eisenhart (1939) was the first to discuss calibration problem who selected the classical estimator more appropriate as compared to the inverse estimator arguing "the fitting should be done in terms of deviations which actually represent error". This viewpoint was supported by Williams (1959).

Fieller-Creasy (1954) proposed methods for obtaining the confidence intervals for the ratio of two parameters (i.e. classical estimator) and Herson (1975) compared Fieller's method with that of so called Delta method (based on first order Taylor series expansion to approximate variance). He gave many rules of thumb for deciding when to approximate Fieller's interval by Delta interval. Mandel (1958) constructed simultaneous confidence intervals (for several T) to the problem of classical type and Miller (1966) gave another solution to this problem. Lieberman et al. (1967) gave a solution to this problem in terms of unlimited simultaneous discrimination intervals. They compared two methods for constructing these intervals, one based on Bonferroni inequality and the other on the idea of Lieberman and Miller (1963) and concluded that Bonferroni type interval is shorter.

Halperin (1961) considered the case when both variables are subject to error and derived confidence intervals under several different assumptions about the kind of information available. He gave the idea of uncertainty which was later on elaborated by Scheffé (1973). In (1970) he found that inverse estimator is superior in the sense of "closeness" for large samples if the values of explanatory

variable are restricted to a certain closed interval near $\bar{t}$; and is inferior elsewhere. He emphasized that the interval in which inverse estimator is superior is trivially small. Saw (1970) also gave this kind of views and found the use of inverse estimator unappealing.

Easterling (1969) considered point and interval estimates based on classical estimator and gave a procedure for obtaining exact confidence intervals by comparing them with Fieller's type (1954) and like Bonferroni intervals discussed by Lieberman et al. (1967).

Krutchkoff (1967) raised this problem of comparing these two estimators again on the basis of mean squared error. He simulated their mean squared errors and concluded (not quite correctly) that inverse estimator has a uniformly smaller mean squared error in the range of controlled variable in the experiment. He (1969a, 69b) claimed that inverse estimator is better for extrapolation for some cases while the classical is better for others. Using the Pitman closeness criterion he (1971) again concluded that inverse estimator is superior or equivalent to the classical estimator.

Williams (1969b) criticised the mean squared error as criterion in the problems of this kind and favoured the classical estimator even though it has infinite mean and mean squared error. We connect the latter with the implicit assumption $\sigma^2 = \infty$ discussed at the end of section 1.7c. He argued that inverse estimator is based on wrong regression. In (1969a) he considered some calibration situations and suggested some formulae to be used accordingly. He discussed the idea of Tallis (1969) of using supplementary information in the calibration situations and concluded that the use of additional information provides closer estimates.

Berkson (1969) showed that inverse estimator is inconsistent and mean squared error (asymptotic in N) for single x is smaller only for a limited range. For large N if t is estimated from the mean of

$k > 1$, then there is always a k_0 such that for all values of $k \geq k_0$, the mean squared error of classical estimator is smaller than the inverse estimator except at $t = \bar{t}$. We discuss the question of consistency in section 3.4.

Martinelle (1970) concluded that the inverse estimator has smaller mean squared error than the classical estimator for t near $\bar{t}$ but if more observations are made on the response variable i.e. ($k > 1$), the advantage of the inverse estimator is reduced.

Cox (1971) showed how the individual x values should be used for interval estimation in the cases where residual variance is constant, proportional to x or to x^2.

Shukla (1972) obtained asymptotic expressions for bias and mean squared error of both the classical and the inverse estimator. He concluded that if large number of observations are taken in the calibration experiment with small error and unknown t is estimated by large number of observations on x (i.e. large k) then it is unlikely that the inverse method will be advantageous over the classical method except in very trivial cases. However he recommended the use of inverse estimator for $k = 1$ and t close to $\bar{t}$ and classical estimator for large sample sizes N, k in the absence of any prior information about *T*. Again Shukla and Datta (1985) obtained exact expressions for the mean and mean squared error of the inverse estimator and compared them with the conditional classical estimator which they obtained from the classical estimator based on the test of hypothesis about the regression coefficient to overcome the difficulty of unbounded mean and variance.

Scheffé (1973) considered in detail the estimator of classical type for a polynomial in t. He used the idea of multiple comparisons and constructed the tables of interval estimates taking into account the intrinsic uncertainty in the estimation of regression parameters.

Odén (1973) found classical type simultaneous confidence intervals with large k. He gave a more precise form in probabilistic terms and good deal of improvement of Miller's method (1966).

Perng and Tong (1974) tackled the problem of classical type by proposing a two-stage sequential procedure for the construction of a fixed width confidence intervals for t, an unknown parameter. They showed that the limiting probabilities of "correct decision" are equal to a pre-assigned number p^*.

Minder and Whitney (1975) used the marginal likelihood methods to compare and make inferences about the unknown value t for both the methods. They found that a good number of cases considered by Krutchkoff (1967, 69a) give non-informative likelihood functions but cases which are common in practice tend to give likelihood functions which are informative and approximately normal in shape.

Schwartz (1975,76,77,78,79) considered different aspects i.e. non-linear calibration, calibration with non uniform variance, in practical problems from chemistry and made suggestions in some situations. Morris (1983) and Leary and Messick (1985) commented on practical calibration situations in chemistry. Makowski and Downing (1980) also solved a practical problem from chemistry by taking the relationship between x and t, both linear and quadratic. They constructed single and joint confidence intervals and compared them.

Naszódi (1978) proposed a modified form of the classical estimator, based on estimates of the first two moments of the estimator obtained from a Taylor's series expansion, which is practically unbiased, more efficient than the classical estimator and has advantage of consistency over the inverse estimator. He also discussed a mode of eliminating the error by experimental design.

Theobald and Mallinson (1978) considered the problem of estimating the calibration equations in both its structural and

functional relationship forms and showed that Barnett's (1969) structural relationship version of the problem is equivalent to a standard factor analysis model used by them. They discussed maximum likelihood estimators for certain constrained models and concluded that maximum likelihood method applied in Williams (1969a) is unworkable under some situations. Jansen (1980) raised the objection on the model proposed by them for not taking into account the goodness of fit test of the model. He suggested a model and compared results with them.

Brown (1978,79b) solved two practical situations. He used the method of generalised least squares and considered the error in both variables.

Trout and Swallow (1979) constructed uniform confidence bands of classical type for the simple inverse regression problem to provide joint confidence intervals for t in a specified range $t_a \leq t \leq t_b$. They compared relative efficiencies of their intervals with that of Scheffé's procedure (1973) and concluded that one has not to pay price in efficiency for the convenience of the uniform procedure.

Clark (1979,80) discussed practical aspects that arise while fitting smooth regression function to radio-carbon dates on tree rings data. He obtained estimates of the smooth function and calibration using biased estimators of the regression function. The cross-validation mean squared error has been proposed for selecting an appropriate regression estimator and its bias. He also proposed adjustment for intervals after theoretical calculations and simulation experiments.

Lundberg and DeMaré (1980) advocated that in applications with small measurement errors, simple approximate confidence intervals in calibration problems serve quite well when the relation of non-linear type in t is considered. They compared their results with simulation

results.

Lechner et al. (1980) discussed about pressure-volume calibration curve. They explained the appropriateness of applying splines to this curve and presented overview of the associated statistical uncertainties. In (1982) they implemented Scheffé's type calibration procedure on a pressure-volume example and compared results with a method as in Naszódi (1978).

Turiel et al. (1982) made simulation study about the linear calibration problem and inverse median estimation problem. They compared the classical estimator, the inverse estimator and the Naszódi (1978) estimator for small and large samples using the criteria of mean squared error, Pitman closeness and probability of over-estimation and suggested different estimators under different situations.

Grassia and Sundberg (1982) considered the statistical precision of class frequency estimates for populations of items. They took into account contributions of error from calibration, from sampling the population and from random mis-classification in the sorting of the sample.

Swallow and Trout (1983) presented the methodology for determining objectively the lower (or upper) limit associated with a simple linear regression i.e., the point below (or above) which a regression model fails. They gave methods for diagnosing, whether problems observed beyond the limit are due to increased variability or due to breakdown of the linear relationship, with multiple observations at some t value.

Schwenke and Milliken (1983) considered nonlinear models and picked the problem of classical type. They gave three techniques for obtaining confidence intervals for $\bar{t}$ based on asymptotic theory. They investigated small sample properties by a simulation study and

compared results.

Hochberg et al. (1983) proposed two new estimators for calibrating unknowns from dose-response curves in a system of quality controlled assays. The new estimators utilize the results of all other assays through the replications of the control samples in the system in contrast with classical estimator which only uses the results of one assay in which response of the unknown dose is measured. They compared results with an example.

Mandel (1984) dealt with the problem when both variables are subject to error and showed how the least squares formulae are modified in this situation. He explained this process in detail.

Knafl et al. (1984) solved a problem of classical type where Scheffé's (1973) procedure did not give good results because of the particular linear model assumed. They assumed a more general model and gave procedures for confidence intervals.

Spiegelman (1984) gave a method that divides the data into training and test groups. The test group is iteratively checked to see that a prechosen nominal confidence probability of coverage is met as in Scheffé (1973). It is shown that nominal probability level is still valid. In (1984) he gave a statistic for identifying influential observations in Scheffé's (1973) type calibration curve.

Oman (1984) analysed residuals in a calibration problem of classical type and proposed a statistic that is appropriate to specific situations and is similar to Cook's Distance. In (1985b) he gave an exact formula for the mean squared error of the inverse estimator in the linear calibration independent of Shukla and Datta (1985). He compared his results with simulation results of Krutchkoff (1967), asymptotic results of Berkson (1969) and Shukla (1972). These results were quite close to the simulation results but differed slightly from asymptotic results because of small N as one

can expect.

Mckeon and Chhikara (1985) compared the classical, inverse and Naszódi estimator from the point of view of regression estimation in sample surveys. They concluded that inverse estimator is more efficient than the classical estimator.

Carroll et al. (1985) and Carroll and Spiegelman (1986) discussed the cases where both variables are subject to error. They studied the effect of measurement errors in simple linear regression and emphasized that both criteria which define what is small measurement error, Draper and Smith for the first criterion and Scheffé (1973) and Mandel (1984) for the second criterion, are useful for point estimation and interval estimation respectively for calibration purposes.

Reilman and Gunst (1986) also discussed errors in both variables and contrasted maximum likelihood estimators of regression parameters with corresponding least squares estimators.

Lwin and Spiegelman (1986) took into account the error in the explanatory variable having a known finite bound. They gave an easily implementable accurate calibration curve procedure and produced conservative confidence intervals.

Among others Rothman (1968), Lindley (1972), Pepper (1973), Winslow (1976), Seber (1977), Dietrich and Marks (1979), Draper and Smith (1981), Sinclair (1982), Kurtz (1983), Branco (1985), Heldal and Spjøtvoll (1985), Schaffrin (1985), Lee and Yum (1985) and Currie (1985) have also considered the calibration problems with different situations.

(b) BAYESIAN APPROACH

Usually $P(x|t)$ is assumed $N(\alpha+\beta t, \sigma^2_{x|t})$ and in addition only prior distribution of T denoted by $\pi(t)$ or both $\pi(t)$ and

$\pi(\alpha, \beta, \sigma^2_{x|t})$ are assumed known. Special cases also include improper priors

$$\pi(\alpha, \beta, \sigma^2_{x|t})d\alpha\, d\beta\, d\sigma^2_{x|t} = d\alpha\, d\beta\, d\, \sigma^2_{x|t}/\sigma^2_{x|t}.$$

The required property of an interval estimate is

$$p(\text{Int. will contain } t|x;\ \text{expt.}) \approx 0.95$$

Dunsmore (1968) assumed that X and T have bivariate normal distribution in the experiment. He derived that posterior distribution $\pi(t_f|\text{expt.};\ x_f)$ is the Student's t distribution which gave inverse estimator $\tilde{t}$ as the conditional mean of $p(t_f|\text{expt.};\ x_f)$ where x_f and t_f denote the values in the future situation.

Hoadley (1970) considered the linear regression problem and showed that among a class of Bayes estimators, the inverse estimator $\tilde{t}$ is a particular solution for a certain informative prior. He showed that if $\pi(\alpha, \beta, \ln \sigma)$ is uniform and $\pi(t_f)$ is t distribution, then posterior of T_f is also Student's t distribution and $\tilde{t}$ is resulting Bayes estimator. His comment on Dunsmore's (1968) bivariate model is "the estimation of t is really a prediction problem (as opposed to a reverse prediction problem)".

Kalotay (1971) applied the structural technique of analysis to solve the linear calibration problem where there are k observations in the current situation to predict a single value of t. He derived a marginal structural distribution for t and compared results with Hoadley (1970).

Frazier (1974) determined the worth of Hoadley's intervals (1970) in practical situations when the distribution of true t is unknown which is usually the case. He made a simulation study and concluded that the confidence interval is always valid for t within the experimental range using end point design.

Aitchison and Dunsmore (1975) considered the problem of both the natural and designed calibration and derived the posterior

distribution of t_f. They assumed vague prior distributions for the parameters and a student's t for t_f having expectation $\bar{t}$ and variance $(1+1/N)S_{TT}/(N-5)$. The resulting calibrative distribution of t came out to be Student's t centred on the inverse estimator. We recall the characterization in section 1.7 of the inverse estimator as the best linear predictor when $\mu = \bar{t}$ and $\sigma^2 = S_{TT}/(N-2)$. Aitchison (1977) analysed a practical problem of system transfer and suggested some modifications.

Williford et al. (1979) derived a posterior distribution of t_f $\pi(t_f|\text{expt}; x_f)$. They assumed the prior distribution of t_f normal, X_i's as $N(\alpha+\beta t, \sigma^2_{x|t})$ and $\pi(\alpha, \beta, \sigma^2_{x|t}) \propto 1/\sigma^2_{x|t}$, which resulted in posterior distribution of t_f as approximately normal. They compared their confidence regions with that of Dunsmore (1968) and made a monte carlo study to compare their own posterior confidence intervals with Hoadley's posterior intervals (1970).

Hunter and Lamboy (1981) in their paper, which appeared with discussion, assumed locally uniform distributions for regression parameters and the response variable X. Since the observations of the regression experiment are independent of the future situation, they derived the distribution of T from the posterior densities of regression parameters and X. They obtained the posterior distribution of T approximately normal for unknown $\sigma^2_{x|t}$ using density of T along with the calibration line. Their posterior distribution of T is equivalent to the structural solution obtained by Kalotay (1971) and has infinite mean and variance suggesting classical estimator satisfactory against Hoadley (1970).

Davis and DeGroot (1982) considered all the four possible combinations of, "regression experiment" and "future situation", being controlled or random. They derived the posterior distribution of t_f assuming prior distributions for t and regression parameters.

They commented on these models as "since the value of t_f is unknown to the experimenter in the discussion, t_f must be regarded as stochastic (even though its value may have been "controlled" by someone else)".

Ansley and Wecker (1984) analysed the non-linear calibration problem. In their Bayesian analysis they derived the predictive distribution for the future observation x_f for given t_f conditional on the experiment in the first step and then they obtained the distribution $\pi(t_f|x_f)$ conditional on the experiment by assuming uniform prior distribution for t. They compared their results with the Lechner et al. (1982).

Bermúdez and Bernardo (1985) have done a Bayesian testing of a calibration procedure.

(c) Best Linear Predictor and related approaches

Usually $P(x|t)$ is assumed $N(\alpha+\beta t, \sigma^2_{x|t})$ and the first two moments of T i.e. $ET = \mu$ and $VAR(T) = \sigma^2$ are assumed known and thus the unconditional interval

$$p(\text{ Int. will contain t}) = 0.95$$

may seem appropriate. This formulation is discussed further in chapter 4.

Tallis (1969) considered the problem of obtaining a satisfactory estimate of a variable T from another variable X where X and T have joint distribution and experimental samples of X and (x_i, t_i) $i = 1, 2, \ldots, N$ have been observed. The estimate of t depends upon regression parameters and the first two moments of T. He used the information on X to get the first two moments of T and obtained the estimates of regression parameters from the pairs (x_i, t_i) $i = 1, 2, \ldots, N$. He illustrated this situation when X and T have bivariate normal distribution.

Brown (1979a) proposed the integrated mean squared error (IMSE)

as an optimising criterion for the linear inverse regression problem. He derived a predictor which depends upon regression parameters and first two moments of T. Like Tallis (1969) he suggested to replace the regression parameters with their estimates from regression experiment and mean and variance of T may be known "apriori". His estimator is inverse estimator when $\bar{t} = \mu$ and $VAR(T) = S_{TT}/(N-2)$. It is to be noted that replacing the regression parameters with their estimates introduces some uncertainty but he did not take into account this uncertainty.

Rao (1975, 81) considered a linear regression problem with a bivariate situation. In the later paper he assumed α and β known and proposed the best linear predictor as an estimate of T. Like others he also assumed first two moments of T i.e. μ and σ^2 known. This problem is mathematically much simpler with known α and β as compared to the problem when α and β are unknown.

Lwin and Maritz (1980) suggested the estimator $t = E(T|X=x_0)$ which has minimum mean squared error and is optimal when regression parameters are known. They assumed that the pairs (t_i, x_i) $1 = 1, 2, \ldots, N$, can be observed and the calibration curve exists. In practical situations regression parameters are not known and can be replaced by their estimates from regression experiment. They suggested to calculate the distribution of T from the t_i's in the data. In (1982) they used the total mean squared error as criterion and derived classical and inverse estimators from it. They compared the mean squared error of both the estimators in the asymptotic sense and showed that the inverse estimator is superior if the current t value is sampled from the same population as the previous t values.

Lwin (1985) took this problem again and discussed it from the point of view of Tallis (1969) and Lwin and Maritz (1980). He empasized the need of improvement in the estimation of parameters

where moments of T are obtained from the supplementary infomation on X or T and regression parameters are replaced by their estimates from the regression experiment.

Copas (1982) stressed the need to assume about the distribution of T in future. He explained how the predicted t value may vary depending upon the situation.

Muhammad and McLaren (1985) reporting some of results of the present thesis, assumed a bivariate random variable (X,T) in the current situation with $E(T) = \mu$ and $VAR(T) = \sigma^2$ known and replaced regression parameters by their estimates from the regression experiment to estimate the best linear predictor. They showed that unconditional mean squared error of the best linear predictor depends only on four invariants and splits into

(a) Intrinsic uncertainty in the bivariate situation;

and

(b) uncertainty due to estimation of regression parameters from the experiment.

They approximated mean squared error using Taylor's series and made a simulation study to check the accuracy of approximations.

2.2. OPTIMAL CALIBRATION DESIGNS

Very few material is available on this aspect of calibration problem. The few found only cover traditional approach and are reported here.

Ott and Myers (1968) were probabily the first to embark on optimal designs in calibration. They considered the classical estimator and used the criterion of minimizing integrated mean squared error. They concluded that for symmetrical designs with N even, N/2 observations should be taken at t = 1 and N/2 at t = -1. Similarly for odd N the optimal design is

$$t_1, t_2, t_3, \ldots, t_{(N-1)/2} = -1$$

$$t_{(N+1)/2} = 0$$

$$t_{(N+3)/2}, \ldots, t_N = 1$$

They also derived optimal designs using linear approximation for the particular cases when the true model is linear or quadratic.

Thomas and Myers (1973) looked into the designs for inverse estimator following the designs for classical estimator because of the long controversy between these estimators. They used the criterion of integrated mean squared error and developed designs for linear approximation when true model is linear or quadratic. The designs depend upon unknown parameters which are not stable. They showed that optimal designs exist and are near optimal and do not depend upon unknown model parameters. In their integrated mean squared error criterion, T is uniform over the range of the experiment but their use of the inverse estimator suggests $\mu = E(T) = \bar{t}$ and $VAR(T) = \sigma^2$ (Brown, 1979a). This may not be appropriate, depending on the design.

Andrews and Herzberg (1973) adopted sequential designs for the classical estimator in two stages. In the first stage they obtained the estimate from regression experiment and in second stage centred the design at this estimate. They proved that asymptotically parameter estimates approach their expected values with probability one. Perng and Tong (1977) also considered this problem by sequential procedure and made a monte carlo study. They reached the same conclusion as Andrews and Herzberg (1973) that the procedure is asymptotically optimal.

Ford (1976) considered the bivariate situation in the linear regression model and assumed T to be $N(\mu, \sigma^2)$. He reached the conclusion that end point design is D-optimal. He also considered designs for the subset of parameters.

Naszódi (1978) proposed a new estimator by modifying the

classical estimator based on an approximate bias correcting factor. He notes that use of the bias correcting formula might be complicated or inconvenient in some situations. An alternative procedure of reducing the bias is proposed based on experimental design. The end point design is the design which optimises the proposed criterion and the criterion is based on the integrated absolute bias over a specified range of T.

Spiegelman and Studden (1980) discussed designs in the context of Scheffé's approach and gave class of appropriate designs which depend upon location knots and slopes of the segments in linear splines.

Buonaccorsi (1986) used the criterion of minimizing asymptotic variance, average asymptotic variance and maximum asymptotic variance over the range of explanatory variable. His criterion of average asymptotic variance is close to Ott and Myers (1968). He has discussed the influence of designs on confidence regions.

CHAPTER 3

DERIVATIONS AND SIMULATIONS

3.1 DERIVATION OF THE BEST LINEAR PREDICTOR

As already discussed under section 1.2 we have a regression experiment and a current situation for the calibration problem; let us consider first the simple linear regression model

$$x = \alpha + \beta t + \epsilon$$

$$\text{or} \quad E(X|t) = \alpha + \beta t$$

$$\text{and} \quad VAR(X|T) = \sigma^2_{x|t} = VAR(X) - \{COV(X,T)\}^2/VAR(T)$$

The experiment provides least squares estimates of α, β, $\sigma^2_{x|t}$ which have the following meaning in the current situation

$$\alpha = E(X) - \beta E(T)$$

$$\beta = COV(X,T)/VAR(T)$$

$$\sigma^2_{x|t} = VAR(X) - \{COV(X,T)\}^2/VAR(T)$$

In the current situation consider a bivariate random variable (X,T), where only X has been observed and t is to be predicted by the best linear predictor $\hat{t} = \hat{C} + \hat{D}X$. Here C and D are required to be estimated which are defined by minimizing (1.1)

$$C = E(T) - DE(X)$$

$$D = COV(T,X)/VAR(X)$$

$$\sigma^2_{t|x} = VAR(T) - \{COV(T, X)\}^2/VAR(X)$$

To calculate three functions C, D, $\sigma^2_{t|x}$; five unknown moments EX, $COV(T, X)$, $VAR(X)$, $E(T)$ and $VAR(T)$ are needed. These are obtained from the regression experiment because $P(x|t)$ is the same in both the situations and $E(T) = \mu$ and $VAR(T) = \sigma^2$ are assumed known. So

$$E(X) = EE(X|T) = \alpha + \beta\mu$$

$$VAR(X) = \sigma^2_{x|t} + \beta^2\sigma^2$$

$$COV(T, X) = \beta\sigma^2$$

thus

$$D = \frac{\beta\sigma^2}{\beta^2\sigma^2 + \sigma^2_{x|t}} = \frac{\beta^{-1}}{1 + \tau^2} \qquad \ldots(3.1)$$

$$C = \mu - (\alpha + \beta\mu)\frac{\beta^{-1}}{1 + \tau^2} = \frac{(-\alpha/\beta + \mu\tau^2)}{1 + \tau^2} \qquad \ldots(3.2)$$

$$= \mu + \frac{-\alpha^*/\beta + (\bar{t} - \mu)}{1 + \tau^2} \qquad \ldots(3.3)$$

$$\sigma^2_{t|x} = \sigma^2 - \frac{\beta^2\sigma^4}{\beta^2\sigma^2 + \sigma^2_{x|t}}$$

$$= \sigma^2\{1 - (1 + \tau^2)^{-1}\} \qquad \ldots(3.4)$$

so

$$\hat{t} = \frac{1}{1 + \tau^2}\left[\frac{X - \alpha}{\beta}\right] + \frac{\tau^2}{1 + \tau^2}\mu \qquad \ldots(3.5)$$

where

$$\tau^2 = \frac{\sigma^2_{x|t}}{\beta^2\sigma^2}; \qquad \rho^2 = \frac{\beta^2\sigma^2}{\beta^2\sigma^2 + \sigma^2_{x|t}}$$

so

$$\hat{t} = \rho^2\,\check{t} + (1 - \rho^2)\,\mu \qquad \ldots(3.6)$$

It is clear that the best linear predictor is weighted average of classical estimator and μ.

In terms of regression estimates $\alpha = \hat{\alpha}$, $\beta = \hat{\beta}$ and $\sigma^2_{x|t} = \hat{\sigma}^2_{x|t}$; the best linear predictor would be

$$\hat{t} = \hat{\rho}^2\,\check{t} + (1 - \hat{\rho}^2)\,\mu \qquad \ldots(3.7)$$

where

$$\hat{\rho}^2 = \frac{\hat{\beta}^2\,\sigma^2}{\hat{\beta}^2\sigma^2 + \hat{\sigma}^2_{x|t}}$$

3.2 BIAS AND MEAN SQUARED ERROR OF THE BEST LINEAR PREDICTOR

These can be calculated under these two situations

(a) α, β and $\sigma^2_{x|t}$ known

(b) α, β and $\sigma^2_{x|t}$ unknown

3.2.1 BIAS

(a) α , β, $\sigma^2_{x|t}$ known

$$\begin{aligned} \text{bias} &= E[T - (C + DX)] \\ &= E(T) - C - DE(X) \\ &= E(T) - \mu + DE(X) - DE(X) \\ &= 0 \end{aligned}$$

so it is unbiased.

(b) α, β, $\sigma^2_{x|t}$ unknown

When α, β, $\sigma^2_{x|t}$ are unknown, these can be replaced by their estimates from the regression experiment, then

$$\begin{aligned} \text{bias} &= E[T - (\hat{C} + \hat{D}X)] \\ &= \mu - E(\hat{C}) - E\hat{D}E(X) \\ &= \mu - E(\hat{C}) - E\hat{D}(\alpha + \beta\mu) \end{aligned}$$

(substituting approximations to $E\hat{C}$ and $E\hat{D}$ from appendix A and simplifying)

$$\begin{aligned} &= \sigma^2/S_{TT}(\bar{t} - \mu)(1 - \rho^2)(2\rho^2 - 1) \\ &= \pm\, \sigma\left[\frac{(C_N)^{\frac{1}{2}}(1 - \rho^2)(2\rho^2 - 1)(B_N)^{\frac{1}{2}}}{(N - 2)}\right] \qquad \ldots(3.8) \\ &\neq 0, \quad \text{so biased} \end{aligned}$$

where

$$C_N = [\sigma^2(N - 2)]/S_{TT}$$

$$B_N = [(\bar{t} - \mu)^2(N - 2)]/S_{TT}$$

3.2.2 MEAN SQUARED ERROR

(a) α, β and $\sigma^2_{x|t}$ known

$$\begin{aligned} MSE &= E[\{T - (C + DX)\}^2] \\ &= E[\{(T - DX) - C\}^2] \\ &= VAR(T - DX) \\ &= (1 - \rho^2)\sigma^2 \end{aligned}$$

(b) α, β and $\sigma^2_{x|t}$ not known

When α, β and $\sigma^2_{x|t}$ are not known, these can be estimated and the values substituted would give the mean squared error as

$$\begin{aligned} MSE &= E[\{T - (\hat{C} + \hat{D}X)\}^2] \\ &= EE[\{T - (\hat{C} + \hat{D}X)\}^2 | \hat{C}, \hat{D}] \\ &= E[ET^2 - 2\hat{C}ET - 2\hat{D}ETX + \hat{C}^2 + 2\hat{C}\hat{D}EX + \hat{D}^2EX^2] \\ &= E[(\hat{C} - C)^2 + 2(EX)(\hat{C} - C)(\hat{D} - D) + (EX^2)(\hat{D} - D)^2 + (1 - \rho^2)\sigma^2] \end{aligned}$$

because the expression in square brackets is quadratic in $\hat{C}$, $\hat{D}$ whose minimum $(1 - \rho^2)\sigma^2$ is achieved when $\hat{C} = C$, $\hat{D} = D$. Note also that (T,X) and $(\hat{C},\hat{D})$ are independent. So

$$\begin{aligned} MSE &= E(\hat{C} - C)^2 + 2(EX)E(\hat{C} - C)(\hat{D} - D) + (EX^2)E(\hat{D} - D)^2 + (1 - \rho^2)\sigma^2] \\ &= (1 - \rho^2)\sigma^2[1 + Q_S] \qquad \ldots(3.9) \end{aligned}$$

where

$$Q_S = \frac{E(\hat{C} - C)^2 + 2(EX)E(\hat{C} - C)(\hat{D} - D) + (EX^2)E(\hat{D} - D)^2}{(1 - \rho^2)\sigma^2} \qquad \ldots(3.10)$$

so

$$\frac{MSE}{\sigma^2} = (1 - \rho^2)[1 + Q_S] \qquad \ldots(3.11)$$

Using the approximations to $VAR\hat{C}$, $VAR\hat{D}$ and $COV(\hat{C}, \hat{D})$ from (appendix A), as mean squared error = var + (bias)^2, and ignoring the (bias)^2 being of order $(1/N)^2$, we get the approximated value of Q_S. This is denoted by Q_A

$$Q_A = \frac{\rho^2}{N} + \frac{1}{N - 2}\left[2\rho^2(1 - \rho^2) + \frac{(1-2\rho^2)^2\sigma^2(N-2)}{S_{TT}} + \frac{\rho^2(\mu - \bar{t})^2(N-2)}{S_{TT}}\right]$$

or

$$Q_A = \frac{\rho^2}{N} + \frac{1}{N - 2}\left[2\rho^2(1 - \rho^2) + (1 - 2\rho^2)^2(C_N) + \rho^2(B_N)\right] \ldots(3.12)$$

$Q_S \approx Q_A$ and equation (3.9) can be written approximately as

$$MSE = (1 - \rho^2)\sigma^2[1 + Q_A] \qquad \ldots(3.13)$$

The expression (3.13) has two components

(1) $(1- \rho^2)\sigma^2$, the intrinsic uncertainty about t|x in a bivariate situation;

and

(2) $Q_A(1 - \rho^2)\sigma^2$, the uncertainty due to estimation of α, β and $\sigma^2_{x|t}$ in the experiment.

The main interest is in Q_A. It is quite clear from equation (3.12) that Q_A or Q_s or equivalently MSE/σ^2 is a function of the following four dimensionless quantities.

(1) N, size of the experiment.

(2) $C_N = (N - 2)\sigma^2/S_{TT}$, relative concentration of the experiment.

(3) $B_N = (N - 2)(\mu - \bar{t})^2/S_{TT}$, relative bias of the experiment.

(4) $\rho^2 = \beta^2\sigma^2/(\beta^2\sigma^2 + \sigma^2_{x|t})$, squared correlation coefficient.

The first three quantities are known and ρ^2 can be estimated from the regression experiment using known μ and σ^2. It should be clear that ρ^2 is not a function of the regression experiment alone as $\rho^2 = \beta^2\sigma^2(\beta^2\sigma^2 + \sigma^2_{x|t})^{-1}$ where σ^2 is not a parameter of the regression experiment.

Now we can prove the following theorem.

THEOREM 3.1 MSE/σ^2 DEPENDS ONLY ON THE ABOVE MENTIONED FOUR INVARIANTS

PROOF:

It can be proved through the following steps

Step 1:

MSE/σ^2 depends only on

α, β, $\sigma^2_{x|t}$; μ, σ^2(defining five bivariate moments)

and

N, $\bar{t}$, S_{TT} (defining the experiment)

Proof:

In the expression (3.11) EX, EX^2, ρ^2 and σ^2 are all (true rather than estimated) moments of the bivariate future distribution, so are functions of α, β, $\sigma^2_{x|t}$; μ and σ^2.

The best linear predictor is

$$\hat{t} = \hat{C} + \hat{D}X$$

where $\hat{C}$, $\hat{D}$ are functions of $\hat{\alpha}$, $\hat{\beta}$, $\hat{\sigma}^2_{x|t}$; μ, σ^2

Thus the distribution of $\hat{C}$, $\hat{D}$ depends on μ, σ^2 and parameters of the distribution of $\hat{\alpha}$, $\hat{\beta}$, $\hat{\sigma}^2_{x|t}$.

Under the normality assumption of $P(x_i|t_i)$, the following are independent:

$\hat{\sigma}^2_{x|t}$ is $\sigma^2_{x|t}\ \chi^2(N-2)/(N-2)$

$\hat{\beta}$ is $N(\beta,\ \sigma^2_{x|t}/S_{TT})$

$\hat{\alpha}^*$ is $N(\alpha^*,\ \sigma^2_{x|t}/N)$

where

$$\alpha^* = \alpha + \beta\bar{t}$$

It concludes that the distribution of $\hat{C}$, $\hat{D}$ depends on μ, σ^2 and α, β, $\sigma^2_{x|t}$, N, $\bar{t}$, S_{TT}.

Step 2:

MSE/σ^2 is not changed by changes of origin/scale of T and/or X

Proof

What happens if origin and scale of T are changed ?

Define $T' = f + gT$

$$\acute{\mu} = ET' = f + g\mu$$

$$\acute{\sigma}^2 = VAR(f + gT) = g^2\sigma^2$$

$$E(X|\acute{T}) = \alpha + \beta T$$

$$= \alpha + \beta[(T' - f)/g] = \acute{\alpha} + \acute{\beta}T'$$

where $\acute{\alpha} = \alpha - (f/g)\beta$, $\acute{\beta} = \beta/g$

$$(\sigma^2_{x|t})' = VAR(X|T') = VAR(X|T) = \sigma^2_{x|t}$$

$\acute{N} = N$; $\bar{T}' = f + g\,\bar{T}$; $S'_{TT} = g^2 S_{TT}$

What happens if also origin and scale of X are changed?

Define $X' = l + mX$

$$\mu'' = \acute{\mu} = f + g\mu$$

$$\sigma'' = \acute{\sigma} = g\sigma \qquad \text{as above}$$

$$E(X'|T') = l + mE(X|T')$$

$$= 1 + m\acute{\alpha} + m\acute{\beta}T'$$

i.e.

$$\alpha'' = 1 + m\acute{\alpha} = 1 + m[\alpha - (f/g)\beta]$$

$$\beta'' = m\acute{\beta} = (m/g)\beta$$

$$(\sigma^2_{X|t})'' = VAR(X'|T') = m^2\sigma^2_{X|t}$$

$$N'' = N;\ \bar{T}'' = f + g\bar{T};\ S'_{TT} = g^2 S_{TT}$$

We have

$$\hat{C} + \hat{D}X = \hat{\rho}^2\{(X - \hat{\alpha})/\hat{\beta}\} + (1 - \hat{\rho}^2)\mu$$

Note that $(\rho^2)'' = \rho^2$

Consider $E[\{T - (\hat{C} + \hat{D}X)\}^2]$

$$X'' - \hat{\alpha}'' = 1 + mX - [1 + m\{\hat{\alpha} - (f/g)\hat{\beta}\}]$$

$$= m(X - \hat{\alpha}) + m(f/g)\hat{\beta}$$

$$[(X'' - \hat{\alpha}'')/\hat{\beta}''] = g[(X - \hat{\alpha})/\hat{\beta}] + f$$

(using the well known fact that $\hat{\alpha}$ transforms in the same way as α, etc.)

$$(\hat{C} + \hat{D}X)'' = \hat{\rho}^2[g(X - \hat{\alpha})/\hat{\beta} + f] + (1 - \hat{\rho}^2)(f + g\mu)$$

$$= g\hat{\rho}^2[(X - \hat{\alpha})/\hat{\beta}] + g(1 - \hat{\rho}^2)\mu + f$$

$$T'' = f + gT$$

by last two lines

$$T'' - (\hat{C} + \hat{D}X)'' = g[T - (\hat{C} + \hat{D}X)]$$

So

$$E[\{T'' - (\hat{C} + \hat{D}X)\}^2] = MSE'' = g^2\,MSE$$

finally

$$MSE''/(\sigma^2)'' = MSE''/(g^2\sigma^2) = MSE/\sigma^2$$

Step 3:

MSE/σ^2 depends only on

$N,\ \rho,\ (B_N)^{\frac{1}{2}},\ C_N,\ \bar{t},\ S_{TT},\ \alpha,\ \beta$; a transformation of 8 variables listed in step 1

Step 4:

MSE/σ^2 depends only on the first four invariants given in step 3;

i.e. N, ρ, $(B_N)^{\frac{1}{2}}$ and C_N.

Proof

The four invariants are not affected by changes of scale and origin in X and T.

Consider two calibration situations or "systems" A and B with same values of N, ρ, $(B_N)^{\frac{1}{2}}$ and C_N.

We can transform system A to system A´ having $\bar{t} = 0$ (by origin of T), $S_{TT} = (N-2)$ by scale of T, $\alpha = 0$ (by origin of X), $\beta = 1$ (by scale of X).

MSE/σ^2 is same for A´ and for A; by step 2.

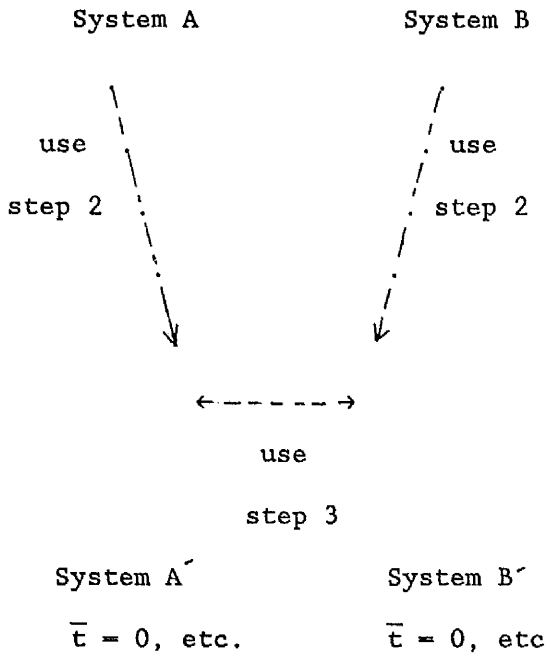
Similarly B can be transformed to a system B´ having $\bar{t} = 0$, etc; where MSE/σ^2 is same for B´ as for B.

Since N, ρ, $(B_N)^{\frac{1}{2}}$, C_N are unchanged by transformations, they have the same values for B´, B, A, A´. So B´, A´ agree in all 8 quantities listed in step 3. Hence MSE/σ^2 is same for B´as for A´.

Thus MSE/σ^2 for A is same as for B.

In other words MSE/σ^2 is a function of N, ρ, $(B_N)^{\frac{1}{2}}$, C_N.

Diagramatically



This completes the proof.

3.3 SIMULATIONS

By now it is quite clear that Q_S or Q_A is a function of four invariants, so to study Q we take the following set of values for these invariants.

$N = 10, 30, 50$

$C_N = 0.25, 1.0, 4.0$

$B_N = 0.0, 1.0, 4.0$

$\rho^2 = 0.7, 0.8, 0.9$

All the 81 (3^4) possible combinations of these four invariants are made and the approximated values of $E\hat{C}$, $E\hat{D}$, $VAR\hat{C}$, $VAR\hat{D}$, $COV(\hat{C}, \hat{D})$ and Q_A are calculated for these combinations using the formulae in appendix A for $k = 1$.

To check the accuracy of approximations, simulated estimates of the above quantities namely $\hat{E}\hat{C}$, $\hat{E}\hat{D}$, $\hat{VAR}\hat{C}$, $\hat{VAR}\hat{D}$, $\hat{COV}(\hat{C}, \hat{D})$ and Q_S for the same 81 combinations are needed. To obtain these $\hat{C}$ and $\hat{D}$ are required to be simulated which are not independent but

$\hat{C}$ is a function of μ, σ^2, $\hat{\alpha}^*$, $\hat{\beta}$, $\hat{\sigma}^2_{x|t}$, $\bar{t}$

and

$\hat{D}$ is a function of $\hat{\beta}$, $\hat{\sigma}^2_{x|t}$, σ^2

where

$\hat{\alpha}^*$ is $N(\alpha^*, \sigma^2_{x|t}/N)$

$\hat{\beta}$ is $N(\beta, \sigma^2_{x|t}/S_{TT})$

$\hat{\sigma}^2_{x|t}$ is $[\sigma^2_{x|t}/(N-2)]\chi^2(N-2)$

From the distribution theory $\hat{\alpha}^*$, $\hat{\beta}$ and $\hat{\sigma}^2_{x|t}$ are independent.

The values of $\alpha^* = 0$, $\beta = 1$, $\mu = 0$ and $\sigma^2 = 1$ are taken for convenience in simulations and $\bar{t}$, S_{TT} and $\sigma^2_{x|t}$ are calculated from the invariants. so

$$S_{TT} = (N-2)\sigma^2/C_N$$

$$\bar{t} = \mu + (B_N/C_N)^{\frac{1}{2}}\sigma$$

$$\sigma^2_{x|t} = [(1 - \rho^2)/\rho^2]\beta^2\sigma^2$$

10000 values of each $\hat{\alpha}^*$, $\hat{\beta}$ and $\hat{\sigma}^2_{x|t}$ are generated using NAG and thus values of $\hat{C}$ and $\hat{D}$ using these simulated values in formulae (3.3) and (3.1) respectively. From these 10000 $\hat{C}$ and $\hat{D}$ values, the following simulated estimates are calculated.

$$\hat{E}\hat{C} = \Sigma\hat{C}_i/10000 \qquad \hat{E}\hat{D} = \Sigma\hat{D}_i/10000$$

$$\hat{VAR}\hat{C} = \Sigma(\hat{C}_i - \hat{E}\hat{C})^2/10000 \qquad \hat{VAR}\hat{D} = \Sigma(\hat{D}_i - \hat{E}\hat{D})^2/10000$$

$$\text{Confidence interval(C.I)} = \hat{E}\hat{C} \pm 1.96[\hat{VAR}\hat{C}/10000]^{\frac{1}{2}}$$

$$\text{Confidence interval(C.I)} = \hat{E}\hat{D} \pm 1.96[\hat{VAR}\hat{D}/10000]^{\frac{1}{2}}$$

$$\hat{COV}(\hat{C}, \hat{D}) = \Sigma[(\hat{C}_i - \hat{E}\hat{C})(\hat{D}_i - \hat{E}\hat{D})]/10000.$$

Comparative study of the approximated and simulated values indicated that all the approximations are always good for N = 30, 50 and $\hat{COV}(\hat{C}, \hat{D})$ is good even for N = 10.

A detailed study is made for N = 10, at ρ^2= 0.7, 0.8, 0.9 separately, for the approximated values $\hat{EC}$, $\hat{ED}$, Q_A versus their simulated values by graphs. The labels in the graphs are according to the following table.

		B_N		
		0.0	1.0	4.0
C_N	0.25	A	B	C
	1.0	D	E	F
	4.0	G	H	I

The approximated $\hat{EC}$'s are plotted against the 95% confidence interval of simulated values, $\hat{E}\hat{C}$ for N=10 and ρ^2 =0.7, 0.8, 0.9,separately which indicate a reasonable approximation except at the point H and I, in figure 3.1, where B_N and C_N have high values. The approximations tended to be better with ρ^2= 0.9.

The approximated and simulated values of the bias, $(\hat{EC} - C)$ and $(\hat{E}\hat{C} - C)$, respectively are calculated where C is calculated by the formula (3.3) using $\alpha^* = 0$, $\beta = 1$, $\mu = 0$ and $\sigma^2 = 1$. The approximated bias values plotted against the interval estimates of simulated bias

values indicate the approximations to bias $\hat{C}$ are rather bad except where B_N is zero, figure 3.1. We also see in this figure that $\hat{E}\hat{C} - C$ is small as compared to $E\hat{C} - C$. Presumably true bias small compared to $E\hat{C} - C$. It is interesting that approximations to bias $\hat{C}$ are not used in approximating Q_A, so zero is better approximation than $E\hat{C} - C$ to true bias.

The $E\hat{D}$ values are plotted against 95% confidence interval of $\hat{E}\hat{D}$ for different values of ρ^2= 0.7, 0.8, 0.9 in figure 3.2 which show approximations are good except at G, H and I where high values of C_N are observed. The approximations are better for $\rho^2 = 0.9$.

The approximated bias $(E\hat{D} - D)$ and simulated bias $(\hat{E}\hat{D} - D)$ are calculated where D is calculated by relation 3.1. Their plots in figure 3.2 indicate that approximations are good except at points G, H, I; where C_N is high. The bias approximations are good for ρ^2= 0.9.

To compute the values of Q_A, expression (3.12) is used and to compute the values of Q_S expression (3.10). The values of Q_S are calculated in two ways.

(1) Q_S is calculated exactly, apart from the simulation error, and the terms $E(\hat{C} - C)^2$, $E(\hat{C} - C)(\hat{D} - D)$ and $E(\hat{D} - D)^2$ in (3.10) are replaced by

$$E(\hat{C} - C)^2 = \hat{VAR}\hat{C} + (\hat{E}\hat{C} - C)^2$$

$$E(\hat{D} - D)^2 = \hat{VAR}\hat{D} + (\hat{E}\hat{D} - D)^2$$

$$E(\hat{C} - C)(\hat{D} - D) = \hat{COV}(\hat{C}, \hat{D}) + (\hat{E}\hat{C} - C)(\hat{E}\hat{D} - D).$$

We will continue denoting this by Q_S

(2) Q_S is calculated by considering the simulated values as true values and ignoring the bias terms in the above expressions i.e.

$$E(\hat{C} - C)^2 = \hat{VAR}\hat{C} \text{ only etc.}$$

We will denote it by Q_{SO}.

The comparative study of Q_A and Q_S, table 1, indicates that the

values are in agreement except at high values of C_N. Because a confidence interval based on Q_s would be complicated, two independent simulations were carried out.

The values of Q_A plotted against the two independent sets of simulated Q_s values (including bias) individually for each value of ρ^2 =0.7, 0.8, 0.9, indicate a reasonable accuracy of approximations except for G, H, I, where C_N has high values,see figure 3.3.

Approximation error $Q_A - Q_s$ can be split into two components

(1) Omission of bias terms altogether i.e $(Q_s - Q_{so})$

(2) Approximating error of $VAR\hat{C}$, $VAR\hat{D}$, $COV(\hat{C}, \hat{D})$ i.e. $(Q_A - Q_{so})$.

Mathematically

$$Q_A - Q_s = (Q_A - Q_{so}) - (Q_s - Q_{so})$$

$$| Q_A - Q_{so}| >> | Q_s - Q_{so}|$$

$(Q_A - Q_{so})$ is plotted against $(Q_s - Q_{so})$ for ρ^2 = 0.7, 0.8, 0.9, figure 3.3, separately and it can be concluded that there is no point in adding approximated bias terms to improve Q_A values unless also added $O(1/N^2)$ terms in approximated variance and covariance terms.

Q_A is inversely proportional to N i.e. with the increase in the size of the experiment, Q_A tends to be small. It is quite obvious from table 3.1. It is an increasing function of C_N. A look at the table indicates that values of Q_A at N =30, ρ^2 = 0.7, B_N = 0.0, C_N = 0.25 and N = 50, ρ^2 =0.7, B_N = 0.0, C_N = 4.0 are 0.039762 and 0.036083 respectively. Similarly the values of Q_A at N = 30, ρ^2 = 0.8, B_N = 1.0, C_N = 0.25 and N = 50, ρ^2= 0.8, B_N = 1.0, C_N = 4.0 are 0.069881 and 0.069333 respectively which are approximately equal. This indicates the importance of C_N and suggests that C_N should be as small as possible or equivalently $S_{TT}/(N-2)$ should be large.

Q_A is also an increasing function of B_N. It is quite evident from table 3. 1. In the table the values of Q_A at N = 10, ρ^2 = 0.7, C_N = 1.0, B_N = 0.0 and N = 30, ρ^2 = 0.7, C_N = 1.0, B_N = 4.0 are

0.142500 and 0.144048. Similarly the values of Q_A at N = 10, $\rho^2 = 0.8$, $C_N = 1.0$, $B_N = 0.0$ and N = 30, $\rho^2 = 0.8$, $C_N = 1.0$, $B_N = 4.0$ are 0.165000 and 0.165238 respectively. These values are approximately equal by fixing ρ^2 and C_N and changing only the size of the experiment. This brings to light the inherent weakness in the planning of the experiment. It suggests that the experiment should be conducted in the right place i.e. with $\bar{t} = \mu$

Q_A is quadratic in ρ^2. To see how it depends on ρ^2, the derivative of Q_A with respect to ρ^2 is

$$\frac{\partial Q_A}{\partial \rho^2} = \frac{1}{N} + \frac{1}{N-2}\left[\, 2(1 - 2\rho^2) - 4C_N(1 - 2\rho^2) + B_N \,\right]$$

$\partial Q_A/\partial \rho^2 > 0$ always if both $C_N > 0.5$ and $\rho^2 > 0.5$ for all B_N and N; so Q is monotonically increasing function for these values of C_N and ρ^2. In most practical situations the range $0.5 < \rho^2 < 1$ is of particular interest. $\partial Q_A/\partial \rho^2 < 0$ for $\rho^2 = 1$, $C_N = 0.25$ and $B_N = 0.0$ for all N but the value of $\rho^2 = 1$ is not of interest in practical situations.

$\partial Q_A/\partial \rho^2 < 0$ for high values of C_N i.e. at $C_N = 4.0$ and low values of $\rho^2 (\rho^2 \leq 0.4)$. So Q_A is monotonically decreasing function for high values of C_N and low values of ρ^2, otherwise it is always monotonically increasing function.

Shukla and Datta (1985) and Oman (1985b) gave an exact formula for the conditional mean squared error of the inverse estimator as under;

$$\text{(Condit) MSE} = (t - \bar{t})^2\left[1 + \frac{\lambda}{2}\{(N-6)\Phi(N-1, \lambda) - (N-4)\Phi(N-3, \lambda)\}\right]$$

$$+\left[1 + \frac{1}{N}\right]\frac{S_{TT}}{2}\{(N-2)\Phi(N-1, \lambda) - (N-4)\Phi(N-3, \lambda)\} \ldots (3.14)$$

where $\lambda = \dfrac{\beta^2 S_{TT}}{\sigma^2_{x|t}}$ and $\Phi(J, \lambda) = E\left[\dfrac{1}{J + 2k}\right]$, $J > 0$

k is a Poisson random variable with parameter $\lambda/2$.

Unconditional mean squared error, which we are denoting in our work by MSE can be derived easily from (3.14) as under

$$(U)MSE = E(t - \bar{t})^2\left[1 + \frac{\lambda}{2}\{(N-6)\Phi(N-1, \lambda) - (N-4)\Phi(N-3, \lambda)\}\right]$$

$$+ \left[1 + \frac{1}{N}\right]\frac{S_{TT}}{2}\{(N-2)\Phi(N-1, \lambda) - (N-4)\Phi(N-3, \lambda)\} \quad ...(3.15)$$

Inverse estimator is a special case of our estimator with

$$\mu = \bar{t} \quad \text{and} \quad E(t - \bar{t})^2 = \frac{S_{TT}}{N - 2}$$

So exact MSE of our estimator for $B_N = 0.0$, $C_N = 1.0$ is

$$(U)MSE = \sigma^2\left[1 + \frac{\lambda}{2}\left[(N-6)\Phi(N-1, \lambda) - (N-4)\Phi(N-3, \lambda)\right]\right.$$

$$\left.+\left[1 + \frac{1}{N}\right]\left[\frac{N}{2} - 1\right]\{(N-2)\Phi(N-1, \lambda) - (N-4)\Phi(N-3, \lambda)\}\right] \quad ...(3.16)$$

According to formula (3.9)

$$MSE = (1 - \rho^2)\sigma^2(1 + Q_S) \quad ...(3.17)$$

λ in (3.16) and ρ^2 in (3.17) are related by $\lambda = \frac{(N-2)\rho^2}{(1 - \rho^2)}$

In our simulations inverse estimator corresponds to $B_N = 0.0$ and $C_N = 1.0$. To compare the results of (3.16) and (3.17), we took $\rho^2 = 0.7, 0.8, 0.9$ and $N = 9$ and computed these three cases. Results are as under,

	$MSE_{(theirs)}/\sigma^2$	$MSE_{(ours)}/\sigma^2$
$\rho^2 = 0.7$ $\lambda = 16.66$	0.351333	0.355470
$\rho^2 = 0.8$ $\lambda = 28$	0.242043	0.241784
$\rho^2 = 0.9$ $\lambda = 63$	0.123424	0.123471

It is quite clear that the exact results $MSE_{(theirs)}/\sigma^2$ and simulated results $MSE_{(ours)}/\sigma^2$ are quite close.

3.4 CALIBRATION BASED ON A SAMPLE MEAN $\bar{X}_f$ OF k OBSERVATIONS

If the k replicated measurements $X_1, X_2, \ldots, X_k$, with mean $\bar{X}_f$, corresponding to the unknown T value are used, then the best linear predictor based on the estimates, $\hat{\alpha}$, $\hat{\beta}$, $\hat{\sigma}^2_{x|t}$, would be

$$\hat{t} = \hat{C} + \hat{D}\bar{X}_f$$

where $$\hat{C} = E(T) - \hat{D}E(\bar{X}_f)$$

and $$\hat{D} = \frac{\hat{COV}(T, \bar{X}_f)}{\hat{VAR}(\bar{X}_f)}$$

$$= \frac{\hat{\beta}\sigma^2}{VAR(\bar{X}_f)}$$

$$= \frac{\hat{\beta}\sigma^2}{\hat{\beta}^2\sigma^2 + \frac{\hat{\sigma}^2_{x|t}}{k}}$$

$$= \frac{\hat{\beta}^{-1}}{1 + \frac{\hat{\tau}^2}{k}} \qquad \text{where } \hat{\tau}^2 = \frac{\hat{\sigma}^2_{x|t}}{\hat{\beta}^2\sigma^2}$$

and

$$\hat{C} = \mu - \frac{\hat{\beta}\sigma^2(\hat{\alpha} + \hat{\beta}\mu)}{\hat{\beta}^2\sigma^2 + \frac{\hat{\sigma}^2_{x|t}}{k}}$$

$$= \frac{-\frac{\hat{\alpha}}{\hat{\beta}} + \mu\frac{\hat{\tau}^2}{k}}{1 + \frac{\hat{\tau}^2}{k}}$$

and

$$\sigma^2_{t|\bar{x}_f} = VAR(T) - \frac{\{COV(\bar{X}_f, T)\}^2}{VAR(\bar{X}_f)}$$

$$= \frac{\hat{\tau}^2\sigma^2/k}{1 + \frac{\hat{\tau}^2}{k}}$$

So $$\hat{t} = \hat{C} + \hat{D}\bar{X}_f$$

$$= \frac{\mu\frac{\hat{\tau}^2}{k} - \frac{\hat{\alpha}}{\hat{\beta}}}{1 + \frac{\hat{\tau}^2}{k}} + \frac{\hat{\beta}^{-1}\bar{x}_f}{1 + \frac{\hat{\tau}^2}{k}} = \frac{\frac{\hat{\beta}\mu\hat{\tau}^2}{k} - \hat{\alpha} + \bar{X}_f}{\hat{\beta}\left[1 + \frac{\hat{\tau}^2}{k}\right]}$$

If $N \to \infty$; $\hat{\alpha} \to \alpha$; $\hat{\beta} \to \beta$, $\hat{\tau}^2 \to \tau^2$ and if also $k \to \infty$, then

$$\hat{t} = \frac{\bar{X}_f - \alpha}{\beta}$$

which is the classical estimator.

Like the classical estimator, the best linear predictor is consistent when N, $k \to \infty$.

The approximations to the $E\hat{C}$, $E\hat{D}$, $VAR\hat{C}$, $VAR\hat{D}$, $COV(\hat{C}, \hat{D})$, bias and mean squared error of the best linear predictor for this situation are obtained using Taylor's series and are given in appendix A.

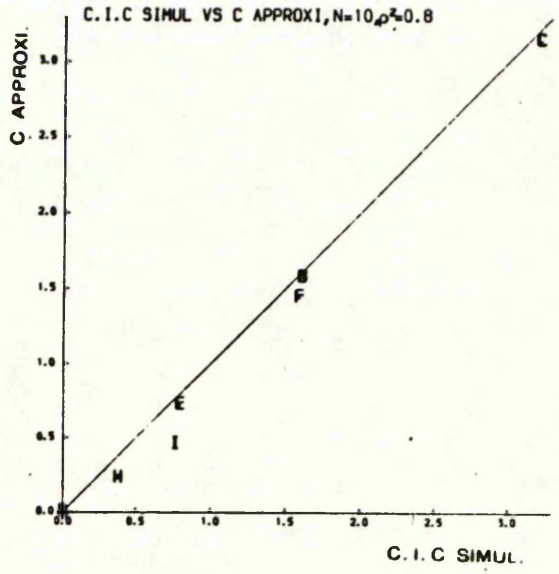
Berkson (1969) pointed out that the inverse estimator is inconsistent, and we showed in section 1.7c that the best linear predictor coincides with the inverse estimator if $\mu = \bar{t}$, $\sigma^2 = S_{TT/(N-2)}$ when $k = 1$. For general k, however, $\sigma^2 = k^{-1} S_{TT}/(N-2)$ would be necessary. The inconsistency of the inverse estimator when N, $k \to \infty$ may perhaps be traced to the fact that it does not correspond to any fixed combination of μ and σ^2.

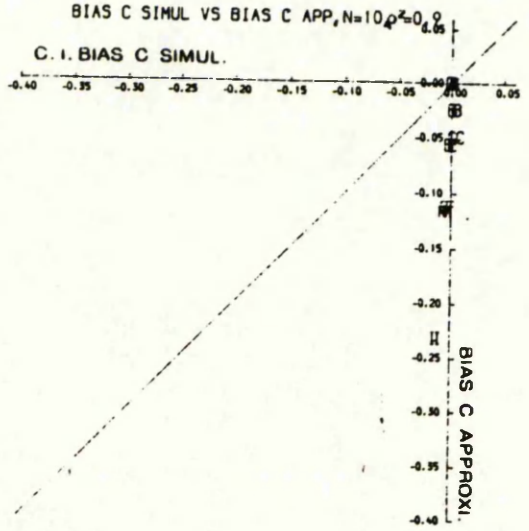
TABLE 3.1

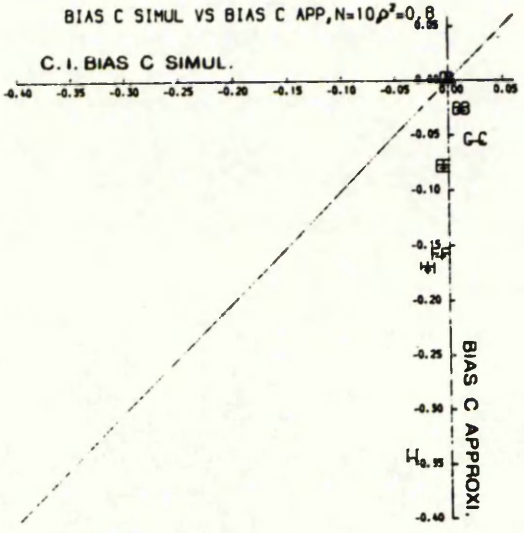
Approximated Q_A and simulated Q_s values for the 81 combinations.

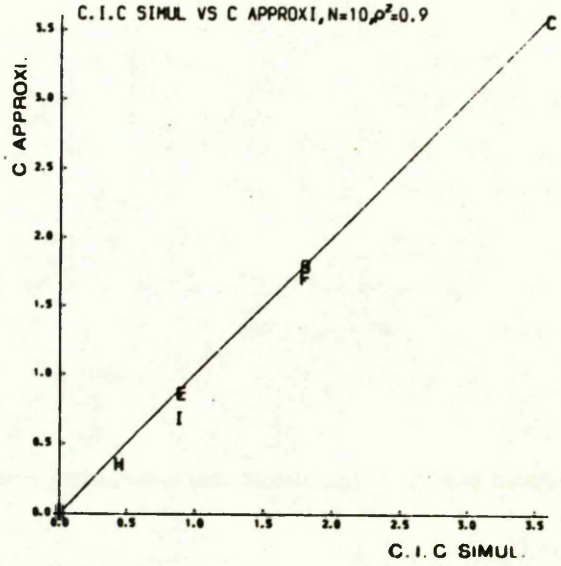
	Q_A			Q_s		
C_N	0.25	1.00	4.00	0.25	1.00	4.00
B_N			$\rho^2 = 0.7$	$N = 10$		
0.0	0.127500	0.142500	0.202500	0.133973	0.162922	0.343772
1.0	0.215000	0.230000	0.290000	0.231629	0.258394	0.414560
4.0	0.477500	0.492500	0.552500	0.520980	0.541941	0.622605
			$\rho^2 = 0.7$	$N = 30$		
0.0	0.039762	0.044048	0.061190	0.039907	0.045299	0.061460
1.0	0.064762	0.069048	0.086190	0.065927	0.071134	0.085475
4.0	0.139762	0.144048	0.161190	0.142287	0.147220	0.156220
			$\rho^2 = 0.7$	$N = 50$		
0.0	0.023583	0.026083	0.036083	0.023933	0.026764	0.036014
1.0	0.038167	0.040667	0.050667	0.038384	0.041207	0.050053
4.0	0.081917	0.084417	0.094417	0.081933	0.084722	0.092415
			$\rho^2 = 0.8$	$N = 10$		
0.0	0.131250	0.165000	0.300000	0.135583	0.186052	0.375541
1.0	0.231250	0.265000	0.400000	0.244554	0.301440	0.483354
4.0	0.531250	0.565000	0.700000	0.567710	0.644115	0.803068
			$\rho^2 = 0.8$	$N = 30$		
0.0	0.041310	0.050952	0.089524	0.041152	0.051864	0.091395
1.0	0.069881	0.079524	0.118095	0.070665	0.081792	0.122379
4.0	0.155595	0.165238	0.203810	0.157536	0.169960	0.213805
			$\rho^2 = 0.8$	$N = 50$		
0.0	0.024542	0.030167	0.052667	0.024706	0.030663	0.053485
1.0	0.041208	0.046833	0.069333	0.041197	0.047328	0.070650
4.0	0.091208	0.096833	0.119333	0.090843	0.097467	0.122241
			$\rho^2 = 0.9$	$N = 10$		
0.0	0.132500	0.192500	0.432500	0.135293	0.208648	0.554168
1.0	0.245000	0.305000	0.545000	0.254600	0.335410	0.711015
4.0	0.582500	0.642500	0.882500	0.608789	0.711931	1.178437
			$\rho^2 = 0.9$	$N = 30$		
0.0	0.042143	0.059286	0.127857	0.041677	0.059580	0.134697
1.0	0.074286	0.091429	0.160000	0.074682	0.093120	0.170378
4.0	0.170714	0.187857	0.256429	0.171870	0.191938	0.275592
			$\rho^2 = 0.9$	$N = 50$		
0.0	0.025083	0.035083	0.075083	0.025062	0.035249	0.077233
1.0	0.043833	0.053833	0.093833	0.043566	0.053962	0.096713
4.0	0.100083	0.110083	0.150083	0.099224	0.110216	0.155178

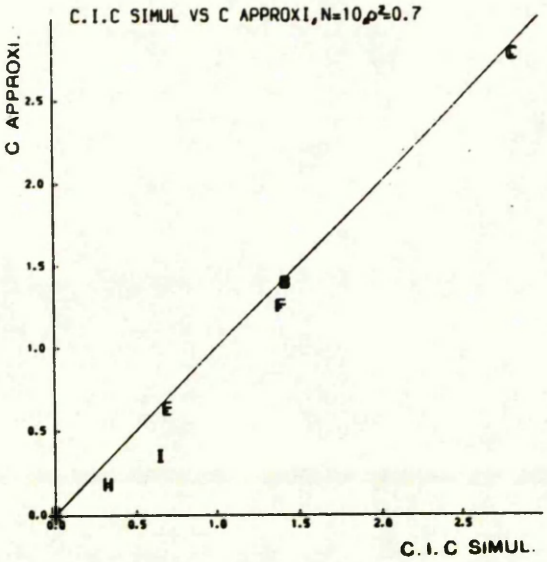
FIGURE 3.1


C.I.C SIMUL VS C APPROXI, N=10 ρ²=0.8
C APPROXI.
C.I.C SIMUL.


BIAS C SIMUL VS BIAS C APP, N=10 ρ²=0.8
C.I. BIAS C SIMUL.
BIAS C APPROXI.


BIAS C SIMUL VS BIAS C APP, N=10 ρ²=0.8
C.I. BIAS C SIMUL.
BIAS C APPROXI.


C.I.C SIMUL VS C APPROXI, N=10 ρ²=0.9
C APPROXI.
C.I.C SIMUL.


C.I.C SIMUL VS C APPROXI, N=10 ρ²=0.7
C APPROXI.
C.I.C SIMUL.

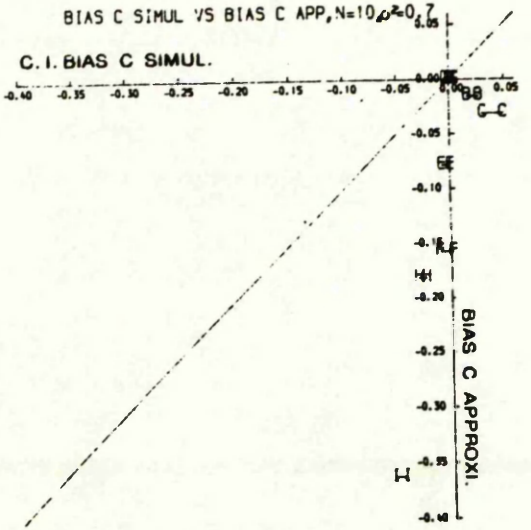
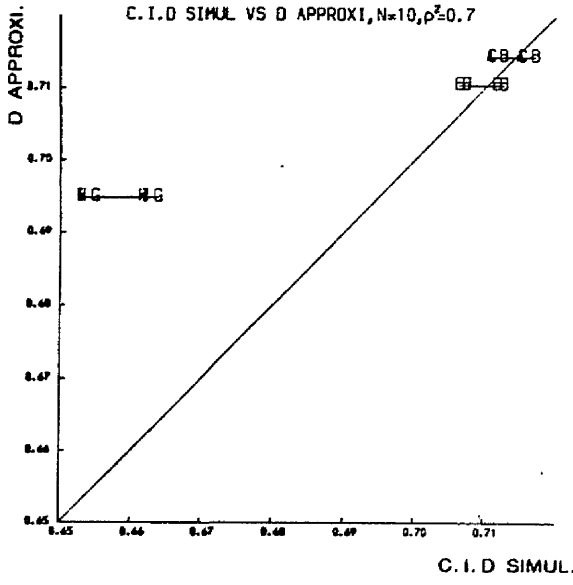
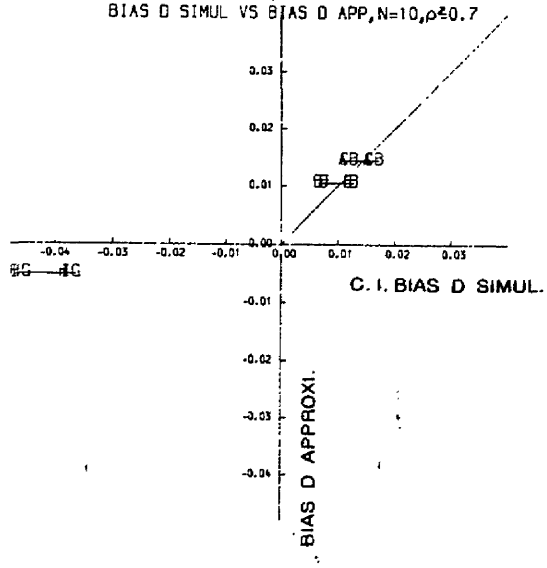
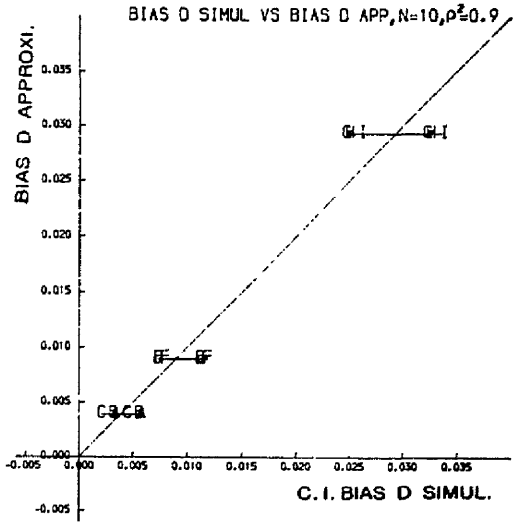
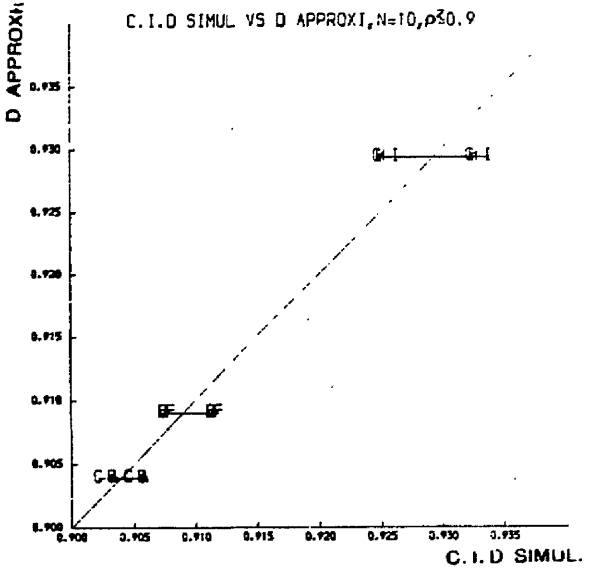

BIAS C SIMUL VS BIAS C APP, N=10 ρ²=0.7
C.I. BIAS C SIMUL.
BIAS C APPROXI.

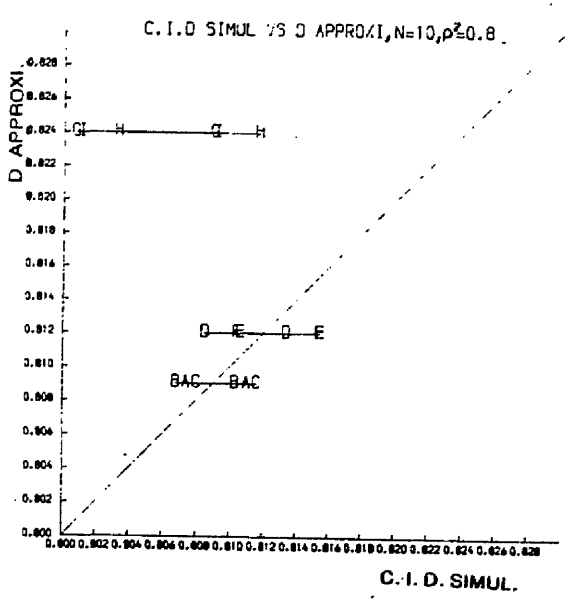
FIGURE 3.2


C.I.D SIMUL VS D APPROXI, N=10, ρ²=0.7
D APPROXI.
C.I. D SIMUL.


BIAS D SIMUL VS BIAS D APP, N=10, ρ²=0.7
C. I. BIAS D SIMUL.
BIAS D APPROXI.


BIAS D SIMUL VS BIAS D APP, N=10, ρ²=0.9
BIAS D APPROXI.
C. I. BIAS D SIMUL.


C.I.D SIMUL VS D APPROXI, N=10, ρ²0.9
D APPROXI.
C. I. D SIMUL.


C.I.D SIMUL VS D APPROXI, N=10, ρ²=0.8
D APPROXI.
C. I. D. SIMUL.

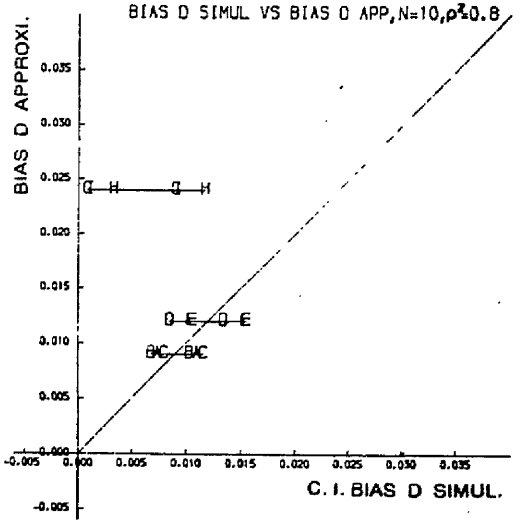
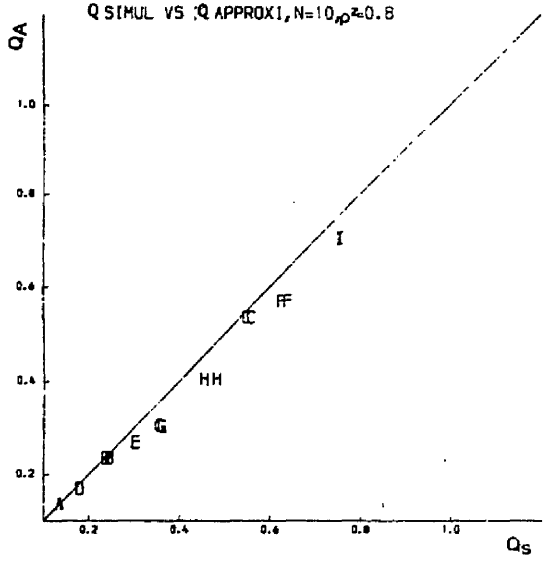
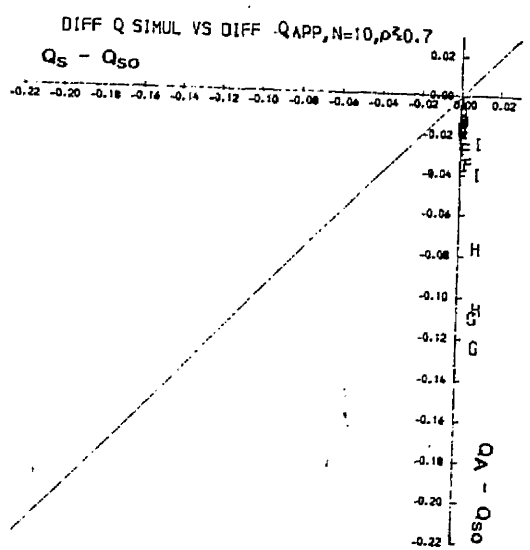
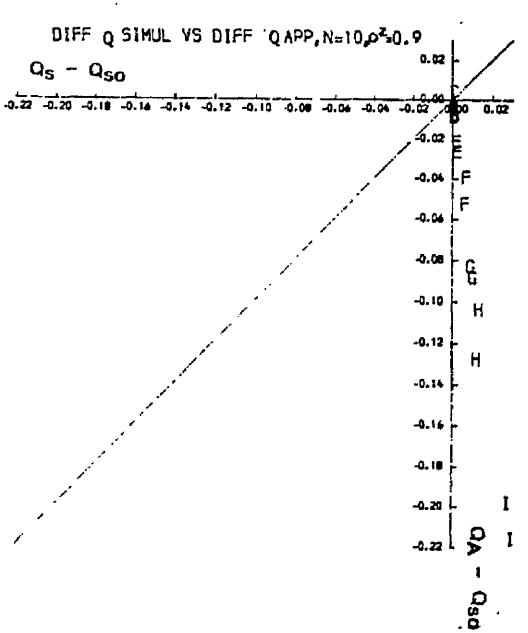
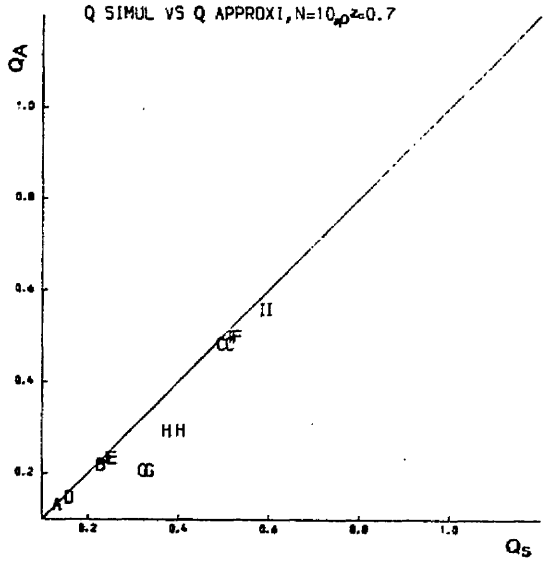

BIAS D SIMUL VS BIAS D APP, N=10, ρ²=0.8
BIAS D APPROXI.
C. I. BIAS D SIMUL.

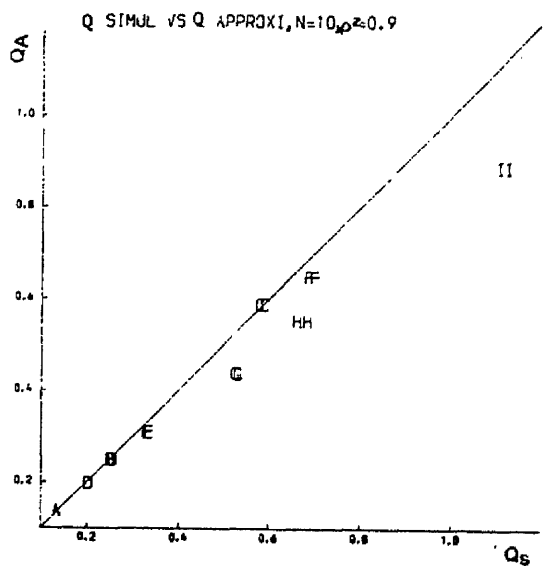
FIGURE 3.3

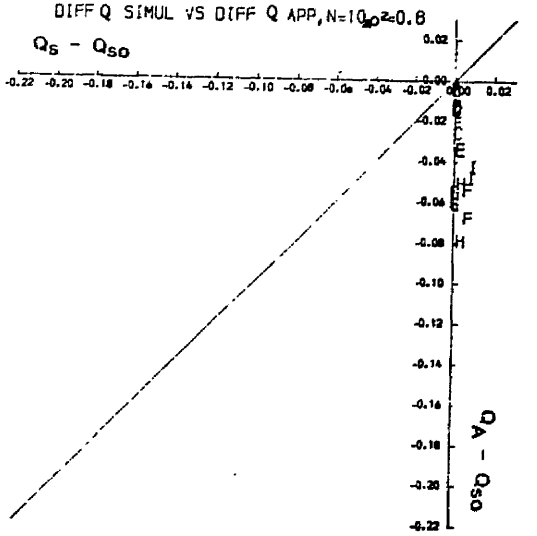

Q SIMUL VS Q APPROXI, N=10, ρ²=0.8
QA
QS


DIFF Q SIMUL VS DIFF Q APP, N=10, ρ²=0.7
Qs - Qso
QA - Qso


DIFF Q SIMUL VS DIFF Q APP, N=10, ρ²=0.9
Qs - Qso
QA - Qso


Q SIMUL VS Q APPROXI, N=10, ρ²=0.7
QA
QS


Q SIMUL VS Q APPROXI, N=10, ρ²=0.9
QA
QS


DIFF Q SIMUL VS DIFF Q APP, N=10, ρ²=0.8
Qs - Qso
QA - Qso

CHAPTER 4

INTERVAL ESTIMATION

Here we discuss possible interval estimates I for T with their justification and emphasize more on unconditional interval estimates using the best linear predictor $\hat{t}$ based on $\bar{X}_f$, μ, σ^2 and the regression parameters α, β, $\sigma^2_{\bar{x}|t}$. There can be two situations.

(a) α, β, $\sigma^2_{\bar{x}|t}$ known

(b) α, β, $\sigma^2_{\bar{x}|t}$ unknown and estimated by their estimates $\hat{\alpha}$, $\hat{\beta}$, $\hat{\sigma}^2_{\bar{x}|t}$.

$\hat{\sigma}^2_{\bar{x}|t}$ means $\hat{\sigma}^2_{x|t}/k$ and $\rho^2_{\bar{x}}$ means correlation coefficient between T and $\bar{X}_f$, where $\bar{X}_f$ is the mean of k observations in the current situation.

4.1. CASE (a) α, β, $\sigma^2_{\bar{x}|t}$ KNOWN.

We discuss this case mainly for the light it will throw on case (b) which occurs more commonly in practice.

4.1.1 POSSIBLE AIMS

There are three possible aims which are discussed as follows.

(a) Conditional confidence given $\bar{x}_f$, $P(T \in I|\bar{x}_f)$ should be 0.95.

If it were possible to calculate, it would be desirable because $\bar{x}_f$ is known; for instance if a specific distribution of $P(t)$ is assumed, $P(t|\bar{x}_f)$ can be deduced by Bayes formula. In particular if $P(t)$ is $N(\mu, \sigma^2)$, then $P(t|\bar{x}_f)$ is $N[C+D\bar{X}_f, (1-\rho^2_{\bar{x}})\sigma^2]$.

Thus an interval estimate

$$I^*: C+D\bar{X}_f \pm 1.96[(1-\rho^2_{\bar{x}})\sigma^2]^{\frac{1}{2}} \qquad \ldots(4.1)$$

would have conditional confidence 0.95 given $\bar{X}_f = \bar{x}_f$

However in our approach, no specific distribution is assumed for $P(t)$, only $E(T) = \mu$ and $VAR(T) = \sigma^2$ are assumed known. The conditional confidence of the interval I^* would depend upon the shape

of $P(t)$. Consider exponential shape $P(t) = e^{-t}$, $(t \geq 0)$. For $\bar{x}_f$ sufficiently negative, interval I^* would be entirely negative so $P(T \in I^* | \bar{x}_f) = 0$.

Thus from the point of view of $P(T \in I | \bar{x}_f)$, the interval I^* is appropriate when T is $N(\mu, \sigma^2)$ and is not appropriate when $P(t) = e^{-t}$, $(t \geq 0)$.

(b) Conditional confidence given T, $P(T \in I | T)$ should be 0.95.

It can be calculated (simulated) because it only involves $P(\bar{x}_f | t)$. It resembles classical approach to calibration problem. It is not suitable aim because for $P(t)$ specified exactly, $P(t | \bar{x}_f)$ can be deduced and conditional confidence given $\bar{x}_f$, $P(T \in I | \bar{x}_f)$ would be relevant. In our approach moments of $P(t)$ are known and this partial information about $P(t)$ makes conditional confidence given T not a suitable aim. Information about $P(t)$ is available and it is not appropriate to insist that $P(T \in I | T) = 0.95$ for all T. We show this by returning to the case $P(t)$ is $N(\mu, \sigma^2)$, for which I^* is the appropriate interval, and showing that conditional confidence given T of interval I^*

$$P[T \text{ within } C+D\bar{X}_f \pm 1.96[(1 - \rho^2_{\bar{X}})\sigma^2]^{\frac{1}{2}} | T] \qquad \ldots(4.2)$$

does depend on T.

For any $P(t)$, this probability is

$$P[(T-C)/D \text{ within } \bar{X}_f \pm (1.96/D)[(1 - \rho^2_{\bar{X}})\sigma^2]^{\frac{1}{2}} | T]$$

$$= P[\bar{X}_f \text{ within } (T-C)/D \pm (1.96/D)[(1 - \rho^2_{\bar{X}})\sigma^2]^{\frac{1}{2}} | T]$$

Using expressions for C and D

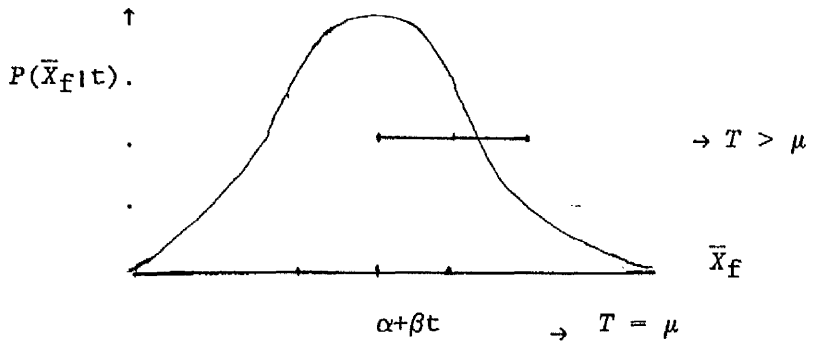
$$\frac{T - C}{D} = \beta T(1 + \tau^2_{\bar{X}}) + \alpha - \mu\beta\tau^2_{\bar{X}}$$

$$= \alpha + \beta T + \beta\tau^2_{\bar{X}}(T - \mu)$$

$$= P[\bar{X}_f \text{ within } \alpha+\beta T + \beta\tau^2_{\bar{X}}(T - \mu) \pm (1.96/D)[(1 - \rho^2_{\bar{X}})\sigma^2]^{\frac{1}{2}} | T]$$

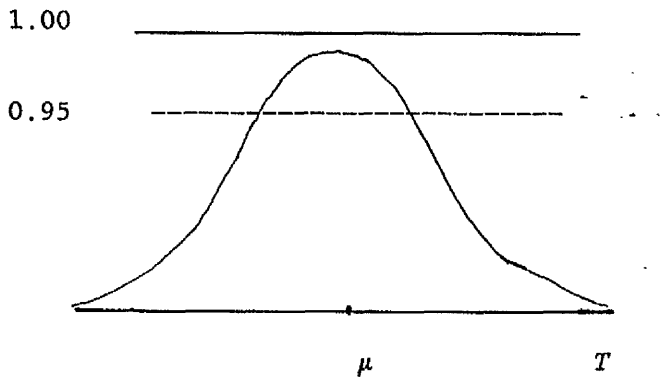
Note that $VAR(\bar{X}_f | T)$ does not depend upon T as $P(\bar{x}_f | t)$ is

$N(\alpha+\beta T,\ \sigma^2_{\bar{x}|t}/k)$ but width of interval for $\bar{x}_f$ unchanged, see the following diagram for the two intervals when $T = \mu$ and $T > \mu$.



so $P[\bar{X}_f$ within $\alpha+\beta t + \beta\tau^2_{\bar{x}}(T - \mu) \pm (1.96/D)[(1- \rho^2_{\bar{x}})\sigma^2]^{\frac{1}{2}}|T]$ depends on T.

$P(T \in I^*|T)$ is a decreasing function of $|T - \mu|$ as following.



The interval I^* in aim (a) above, was seen appropriate when $P(t)$ is $N(\mu,\ \sigma^2)$. Such information about $P(t)$ thus conflicts with the requirement that $P(T \in I|T) = 0.95$ for all T. It is strongly linked with the fact that

$$E(C+D\bar{X}_f|T) = \frac{T + \mu\tau^2_{\bar{x}}}{1 + \tau^2_{\bar{x}}} \neq T \quad \text{"biased"}$$

Unbiasedness (conditional on T) is not appropriate if there is information about $P(t)$. In our approach μ and σ^2 are assumed known and conditional confidence given T is still not appropriate.

(c) Unconditional confidence, $P(T \in I)$ should be 0.95.

We recall that aim (a), conditional confidence given $\bar{X}_f$ is desirable but not attainable with our limited assumptions. On the other hand aim (b), conditional confidence given T is not desirable (given μ and σ^2).

The unconditional confidence, $P(T \in I)$ of any interval is

$$P(T\epsilon I|\mu,\sigma^2;\ \alpha,\beta,\sigma^2_{\bar{X}|T}) = \int P(T \in I|T=t;\ \alpha,\beta,\sigma^2_{\bar{X}|t})P(T|\mu,\sigma^2)dt \quad ...(4.3)$$

Based on the best linear predictor $C+D\bar{X}_f$, the proposed interval is

$$\begin{aligned} & C+D\bar{X}_f \pm 1.96[(1-\rho^2_{\bar{X}})\sigma^2]^{\frac{1}{2}} \\ &= [(\bar{X}_f-\alpha)/\beta]\rho^2_{\bar{X}} + (1-\rho^2_{\bar{X}})\mu \pm 1.96[(1-\rho^2_{\bar{X}})\sigma^2]^{\frac{1}{2}} \\ &= \check{t}\,\rho^2_{\bar{X}} + (1-\rho^2_{\bar{X}})\mu \pm 1.96[(1-\rho^2_{\bar{X}})\sigma^2]^{\frac{1}{2}} \qquad ...(4.4) \end{aligned}$$

Now $E(C+D\bar{X}_f - T) = 0$, and

$$VAR(C+D\bar{X}_f - T) = E(C+D\bar{X}_f - T)^2 = (1-\rho^2_{\bar{X}})\sigma^2,$$

due to definition of the best linear predictor $\hat{t}$. We suggest that $P(T \in I)$ depends on $P(t)$ less strongly than does $P(T \in I|\bar{X}_f)$. So $P(T \in I) = 0.95$ is an aim easier to achieve approximately than ideal aim discussed in (a).

Note that for any interval, aim (c) would be implied by aim (a). Also aim (c) would be implied by the inappropriate aim (b).

By Chebychev's inequality

$$P[\,|C+D\bar{X}_f - T| > 1.96[(1-\rho_{\bar{X}}{}^2)\sigma^2]^{\frac{1}{2}}\,] \le (1/1.96)^2 = 0.26$$

The unconditional confidence (4.3) can be evaluated approximately for I^* for specified $P(t)$. Different choices of $P(t)$ are discussed below.

4.1.2 UNCONDITIONAL CONFIDENCE, AS A FUNCTION OF $P(t)$

Four different distributions, normal, exponential, uniform and triangular are considered as $P(t)$ in (4.3) and the tail probabilities are calculated.

(a) $P(t)$ Normal

From discussion on aims it can be concluded that Error probability is 0.025 for each tail for all normal distributions.

For $P(t)$ is $N(\mu, \sigma^2)$ ---> $P(T, \bar{X}_f)$ is bivariate normal

---> $P(T|\bar{X}_f)$ is $N[C+D\bar{X}_f, (1-\rho^2_{\bar{x}})\sigma^2]$

---> $P(T \in I^*|\bar{X}_f) = 0.95$ (all $\bar{X}_f$)

---> $P(T \in I^*) = 0.95$ because (a) ---> (c)

$P(T \in I^*)$ can be evaluated as under.

Error probability for upper tail

$$P(T>C+D\bar{X}_f+1.96[(1-\rho^2_{\bar{x}})\sigma^2]^{\frac{1}{2}}) = \int \left[P(T-(C+D\bar{X}_f) > 1.96[(1-\rho^2_{\bar{x}})\sigma^2]^{\frac{1}{2}}|T)\right]P(t)dt$$

referring to $C+D\bar{X}_f = \check{t}\rho^2_{\bar{x}} + (1-\rho^2_{\bar{x}})\mu$

$$E(\check{t}|T) = T; \qquad VAR(\check{t}|T) = \frac{\sigma^2_{\bar{x}|t}}{\beta^2}$$

so $E[T - (C+D\bar{X}_f)|T] = T - \rho^2_{\bar{x}}T - (1-\rho^2_{\bar{x}})\mu$

$$= (1-\rho^2_{\bar{x}})(T-\mu)$$

and $VAR[T-(C+D\bar{X}_f)|T] = (\rho^2_{\bar{x}})^2 \frac{\sigma^2_{\bar{x}|t}}{\beta^2}$

$$= (\rho^2_{\bar{x}})^2 \left[\frac{1-\rho^2_{\bar{x}}}{\rho^2_{\bar{x}}}\right]\sigma^2$$

$$= \rho^2_{\bar{x}}(1-\rho^2_{\bar{x}})\sigma^2$$

$$= \int P[N[(1-\rho^2_{\bar{x}})(T-\mu), \rho^2_{\bar{x}}(1-\rho^2_{\bar{x}})\sigma^2] > 1.96[(1-\rho^2_{\bar{x}})\sigma^2]^{\frac{1}{2}}P(t)dt$$

$$= \int \Psi\left[\frac{1.96\left[(1-\rho^2_{\bar{X}})\sigma^2\right]^{\frac{1}{2}} - (1-\rho^2_{\bar{X}})(T-\mu)}{\left[\rho^2_{\bar{X}}(1-\rho^2_{\bar{X}})\sigma^2\right]^{\frac{1}{2}}}\right] P(t)dt$$

$$= \int \Psi\left[\frac{1.96}{(\rho^2_{\bar{X}})^{\frac{1}{2}}} - \left[\frac{1-\rho_{\bar{X}}^2}{\rho^2_{\bar{X}}}\right]^{\frac{1}{2}}\left[\frac{t-\mu}{\sigma}\right]\right] P(t)dt \qquad \ldots(4.5)$$

$$= \int \Psi\left[\frac{1.96}{(\rho^2_{\bar{X}})^{\frac{1}{2}}} - \left[\frac{1-\rho^2_{\bar{X}}}{\rho^2_{\bar{X}}}\right]^{\frac{1}{2}} t^*\right] P(t^*)dt^* \qquad \ldots(4.6)$$

where $t^* = \left[\frac{t-\mu}{\sigma}\right]$

and $\Psi(Z) = P[\,N(0,1) > Z\,]$

Error probability for the lower tail

Similarly as above

$$P(T<C+D\bar{X}_f-1.96\left[(1-\rho^2_{\bar{X}})\sigma^2\right]^{\frac{1}{2}} = \int \Phi\left[\frac{-1.96}{(\rho^2_{\bar{X}})^{\frac{1}{2}}} - \left[\frac{1-\rho^2_{\bar{X}}}{\rho^2_{\bar{X}}}\right]^{\frac{1}{2}}\left[\frac{t-\mu}{\sigma}\right]\right] P(t)dt \ldots(4.7)$$

$$= \int \Phi\left[\frac{-1.96}{(\rho^2_{\bar{X}})^{\frac{1}{2}}} - \left[\frac{1-\rho^2_{\bar{X}}}{\rho^2_{\bar{X}}}\right]^{\frac{1}{2}} t^*\right] P(t^*)dt^* \ldots(4.8)$$

where $\Phi = 1 - \Psi(Z)$.

Considering $P(t)$ normal and $\rho^2_{\bar{X}} = 0.1(0.01)0.99$ in (4.5) and (4.7), Gauss-Hermite formula from NAG was used to solve the integral numerically and the results came out to be 0.025 for both upper and lower tail probabilities upto three decimal places. This confirms the accuracy of numerical integration.

(b) *P(t) Exponential*

The error probabilities when $P(t)$ is exponential i.e

$$P(t) = e^{-t}, \quad (t \geq 0)$$

can be calculated using (4.5) and (4.7) or alternatively the following procedure can be adopted.

Upper tail probability

$$P(T>C+D\bar{X}_f+1.96\left[(1-\rho^2_{\bar{X}})\sigma^2\right]^{\frac{1}{2}}) = P[\,T>C+D(\alpha+\beta T)+D\sigma_{\bar{X}|t}Z+1.96[(1-\rho_{\bar{X}}^2)\sigma^2]^{\frac{1}{2}}\,]$$

as $\bar{X}_f = \alpha + \beta T + \sigma_{\bar{X}|t}\, Z$

Z is N(0,1) and Z and T are independent

$$= P\left[T(1-D\beta)-D\sigma_{\bar{X}|t}Z > C+D\alpha+1.96[(1-\rho^2_{\bar{X}})\sigma^2]^{\frac{1}{2}}\right]$$

$$= P\left[T - \frac{D\sigma_{X|t}}{(1-D\beta)(k)^{\frac{1}{2}}} Z > \frac{C+D\alpha+1.96[(1-\rho^2_{\bar{X}})\sigma^2]^{\frac{1}{2}}}{(1- D\beta)}\right]$$

Note that $D\beta = \rho^2_{\bar{X}} < 1$

$$= P(T-aZ > f) = P(T > aZ+f)$$

where $a = D\,\frac{\sigma_{X|t}}{(1-D\beta)(k)^{\frac{1}{2}}} = \sigma\left[\frac{\rho^2_{\bar{X}}}{1-\rho^2_{\bar{X}}}\right]^{\frac{1}{2}} = \left[\frac{\rho^2_{\bar{X}}}{1-\rho^2_{\bar{X}}}\right]^{\frac{1}{2}}$ since $\sigma = 1$

and $f = \dfrac{C+D\alpha + 1.96[(1-\rho^2_{\bar{X}})\sigma^2]^{\frac{1}{2}}}{(1 - D\beta)}$

$$= \frac{[\mu(1 - \rho^2_{\bar{X}}) + 1.96[(1- \rho^2_{\bar{X}})\sigma^2]^{\frac{1}{2}}]}{(1 - \rho^2_{\bar{X}})}$$

$$= 1 + \frac{1.96}{(1 - \rho^2_{\bar{X}})^{\frac{1}{2}}} \qquad \text{since } \mu = 1 \text{ and } \sigma = 1$$

thus

$$= \int_{-\infty}^{\infty} P(T > aZ+f|Z)P(z)dz$$

$$= \int_{-\infty}^{-f/a} P(T > aZ+f|Z)P(z)dz + \int_{-f/a}^{\infty} P(T > aZ+f|Z)P(z)dz$$

$$= \int_{-\infty}^{-f/a} 1.\ P(z)dz + \int_{-f/a}^{\infty} e^{-aZ-f}\,P(z)dz$$

$$= P(Z < -f/a) + e^{-f + \frac{a^2}{2}} \int_{-f/a}^{\infty} (2\pi)^{-\frac{1}{2}}\, e^{-\frac{1}{2}(Z+a)^2} dz$$

$$= P(Z < -f/a) + e^{-f + \frac{a^2}{2}}\, \Psi(a - f/a)$$

$$= \Psi(f/a) + e^{-f + \frac{a^2}{2}}\, \Psi(a - f/a) \qquad \ldots(4.9)$$

where $\Psi(s) = P[N(0, 1) > s]$

similarly

Lower tail probability

$$P(T<C+D\bar{X}_{\bar{f}}-1.96[(1-\rho^2_{\bar{X}})\sigma^2]^{\frac{1}{2}}) = 1-[\Psi(f'/a)+e^{-f'+\frac{a^2}{2}}\,\Psi(a - f'/a)] \ldots(4.10)$$

where $f' = 1 - 1.96/(1- \rho^2_{\bar{X}})^{\frac{1}{2}}$.

Gauss-Laguerre formula from NAG is used to evaluate (4.5) and (4.7) for $P(t)$ to be exponential. Numerical integration results are produced in table (4.1). The results obtained by using (4.9) and (4.10) are the same as the results of (4.5) and (4.7).

From (4.5) it can be seen that error probability is a function of $\rho^2_{\bar{X}}$ for all shifted exponential distributions

$$P(t) = \theta e^{-\theta(t-t_0)}, \qquad t \geq t_0$$

If shape of $P(t)$ as exponential is fixed but μ and σ^2 change, then t^* is same but ρ^2 changes as illustrated in table 4.1.(b)

For the case of $P(t) = e^{-t}$, $(t \geq 0)$, results are given in table 4.1 when $\rho^2_{\bar{X}} = 0.1(0.01)0.99$. To find error probabilities for $P(t) = \theta e^{-\theta(t-t_0)}$, $(t \geq t_0)$ and given $\beta^2/\sigma^2_{\bar{X}|t}$, one must first calculate $\rho^2 = \beta^2\sigma^2(\beta^2\sigma^2 + \sigma^2_{\bar{X}|t})^{-1}$ as in table 4.1.(b). By relation (4.5) and (4.7) this determines the error probabilities for $P(t)$ of exponential shape, therefore table 4.1.(a) can be used and table 4.2 illustrates the results.

(C) <u>P(t) Uniform and Triangular</u>

For numerical integration of $P(t)$ uniform and triangular in (4.5) and (4.7) Gauss-Legendre formula is used with $\rho^2_{\bar{X}} = 0.1(0.01)0.99$. Results for lower and upper tail probabilities are given in table 4.3.(a) and 4.3.(b) respectively.

To find error probabilities for $P(t) = 1/(b-a)$, $a < x < b$ and given $\beta^2/\sigma^2_{\bar{X}|t}$, one needs to find $\rho^2 = \beta^2\sigma^2(\beta^2\sigma^2 + \sigma^2_{\bar{X}|t})^{-1}$ as in table 4.4.(a). By (4.5) and (4.7), this determines the error probabilities for $P(t)$ of uniform shape, therefore the table 4.4.(a) can be used and table 4.5.(a) illustrates the results.

Similar type of results are calculated for $P(t)$ to be triangular and are given in tables 4.3.(b), 4.4.(b) and 4.5.(b).

4.2. CASE (b) α, β, $\sigma^2_{\bar{x}|t}$ UNKNOWN

The unknown parameters can be estimated, then the interval for T would reflect uncertainty about α, β, $\sigma^2_{\bar{x}|t}$.

To study the error probability by simulation, following four pivotal functions are considered. Each function would lead to a slightly different interval for T using the approximation that the function is N(0, 1). This is further discussed under section 4.2.3.

4.2.1. PIVOTAL FUNCTIONS

The possible pivotal functions are

$$F_1 = \frac{T - (\hat{C}+\hat{D}\bar{X}_f)}{(\hat{MSE})^{\frac{1}{2}}}$$

$$F_2 = \frac{T - (\hat{C}+\hat{D}\bar{X}_f+\hat{B}_T)}{(\hat{MSE} - \hat{B}^2_T)^{\frac{1}{2}}}$$

$$F_3 = \frac{T - (\hat{C}+\hat{D}\bar{X}_f)}{(\hat{MSE} - \hat{B}^2_T)^{\frac{1}{2}}}$$

$$F_4 = \frac{T - (\hat{C}+\hat{D}\bar{X}_f+\hat{B}_T)}{(\hat{MSE})^{\frac{1}{2}}}$$

where B_T is the bias defined in appendix A.

It can be shown that the distribution of the above four functions depends mainly upon the four invariants already studied i.e. B_N, C_N, N and $\rho^2_{\bar{x}}$.

Consider the first two moments of F_1, for instance,

$$EF_1 \approx \frac{E[T - (\hat{C}+\hat{D}\bar{X}_f)]}{E(\hat{MSE})^{\frac{1}{2}}}$$

$$= \frac{B_T}{E(\hat{MSE})^{\frac{1}{2}}}$$

$$\approx \frac{B_T}{(MSE)^{\frac{1}{2}}}$$

while $$E(F_1^2) \approx \frac{MSE}{E(\hat{MSE})} \approx 1$$

Now approximately

$$B_T = E[T - (\hat{C}+\hat{D}\bar{X}_f)]$$

$$= \sigma\left[\frac{k}{N-2}(C_N)^{\frac{1}{2}}(1-\rho^2_{\bar{x}})(2\rho^2_{\bar{x}} - 1)(B_N)^{\frac{1}{2}}\right] \quad \ldots(4.11)$$

and

$$MSE = (1 - \rho^2_{\bar{x}})\,\sigma^2[1 + Q_A] \quad \ldots(4.12)$$

Thus the first two moments of F_1 depend approximately on

$$\frac{B_T/\sigma}{(MSE/\sigma^2)^{\frac{1}{2}}}$$

where B_T/σ and $(MSE/\sigma^2)^{\frac{1}{2}}$ both are functions of the same four invariants as shown in theorem 5.2 and theorem 3.1 respectively.

4.2.2. SIMULATIONS

F_1, F_2, F_3 and F_4 are calculated by assuming a specific $P(t)$ and generating T from it. $\bar{X}_f$ is simulated from $P(\bar{x}_f|t)$. To calculate $\hat{C}$, $\hat{D}$, $\hat{\rho}^2_{\bar{x}}$, $\hat{B}_T$ and $\hat{MSE}$, we take $k = 1$, $\alpha^* = 0$, $\beta = 1$, $\mu = 0$, $\sigma^2 = 1$.

S_{TT} and $\bar{t}$ are derived from B_N and C_N as in section 3.3 and

$$\sigma^2_{\bar{x}|t} = [(1 - \rho^2_{\bar{x}})k]/\rho^2_{\bar{x}}$$

$\hat{\alpha}^*$, $\hat{\beta}$ and $\hat{\sigma}^2_{x|t}$ are simulated from the usual normal theory and thus $\hat{C}$ and $\hat{D}$ as in section 3.3 so

$$\bar{X}_f = T - \bar{t} + (\sigma^2_{x|t}/k)^{\frac{1}{2}}\, N(0,1)$$

$$\hat{\rho}^2_{\bar{x}} = \hat{\beta}^2(\hat{\beta}^2 + \hat{\sigma}^2_{x|t}/k)^{-1}$$

$\hat{B}_T$ and $\hat{MSE}$ are calculated by the relations given in appendix A.

The same four invariants studied in section 3.3 each at three levels are taken i.e.

$B_N = 0.0,\ 1.0,\ 4.0,$

$C_N = 0.25,\ 1.0,\ 4.0,$

$N = 10,\ 30,\ 50$

$\rho^2_{\bar{x}} = 0.7,\ 0.8,\ 0.9.$

81 combinations are made of the values of these invariants and values of F_1, F_2, F_3 and F_4 are calculated by simulating $\hat{\alpha}^*$, $\hat{\beta}$, $\hat{\sigma}^2_{x|t}$ and the other quantities as mentioned above 10000 times for each combination. The point estimates of the lower tail probability $P_{iL} = P(F_i < -1.96)$ and upper tail probability $P_{iU} = P(F_i > 1.96)$ $i = 1, 2, \ldots, 4$ are calculated and also interval estimates for these tail probabilities.

The same 10000 values of T and of the N(0, 1) random variables defining $\bar{X}_f$, $\hat{\alpha}^*$, $\hat{\beta}$ and the χ^2 random variable defining $\hat{\sigma}^2_{x|t}$ are used for all the four pivotal functions, for all the 81 combinations. This reuse of the simulated values permits "paired data" rather than "two sample" comparisons particularly between pivotal functions.

4.2.3. UNCONDITIONAL CONFIDENCE, AS IT DEPENDS ON $P(t)$

Four distributions, normal, exponential, triangular and uniform as in section 4.1.2. are tried as $P(t)$ to simulate F_1, F_2, F_3, F_4 and all the four $P(t)$ are generated with location parameter zero and scale parameter one.

1000 values of each of the pivotal functions are used to see their distribution. Normal probability plots indicate that F_1, F_2, F_3 and F_4 are normal for $N = 30, 50$. For $N = 10$, they are normal most of the times but for high B_N, C_N where some outliers appear in tails. For normal probability plots when $N = 10$, $B_N = 4.0$, $C_N = 4.0$ and $\rho^2_{\bar{x}} = 0.7$ and for each $P(t)$, see figure 4.1.

Summary statistics table 4.6 and inspection of individual values of F_1, F_2, F_3 and F_4 indicate that $F_1 = F_3$ and $F_2 = F_4$ with difference in the third decimal place only quite a small number of times. It is because B^2_T, bias squared term, is small compared to MSE and correcting denominator makes little difference. They have approximately zero mean and unit variance which confirms the assumption made at the end of section 4.2.1.

As it is evident that $F_1 = F_3$ and $F_2 = F_4$, so we would deal with F_1 and F_2 only.

(a) *P(t) NORMAL*

Point Estimates

It is observed that $\hat{P}_{1L} \leq \hat{P}_{2L}$ whereas $\hat{P}_{1U} \geq \hat{P}_{2U}$ for all the 81 combinations.

For N = 10

$\hat{P}_{1L}$ ranges from 0.042 - 0.057

$\hat{P}_{2L}$ " " " 0.042 - 0.064

$\hat{P}_{1U}$ " " " 0.026 - 0.046

$\hat{P}_{2U}$ " " " 0.025 - 0.046

Sum of both tails; $\hat{P}_{1L} + \hat{P}_{1U} \leq \hat{P}_{2L} + \hat{P}_{2U}$ and individual sums are nearly 0.088, greater than 0.05.

For N = 30, 50; $\hat{P}_{1U} \approx \hat{P}_{2U}$ and $\hat{P}_{1L} \approx \hat{P}_{2L}$ and are between 0.021 - 0.030 and 0.029 - 0.035 respectively. $\hat{P}_{1L} + \hat{P}_{1U} \leq \hat{P}_{2L} + \hat{P}_{2U}$ and sums for N = 30, 50 are nearly 0.06 and 0.055 $\geq$ 0.05 respectively.

It is observed that there is more variation in F_2 as compared with F_1 and it has also been noted that there is not much change in tail probability because of the change in $\rho^2_{\bar{x}}$ from 0.7 to 0.9.

Interval Estimates

The intervals for the estimates P_{iL}, P_{iU}, i = 1, 2; are calculated by the formula

$$\hat{P}_i \pm 1.96 \left[\hat{P}_i(1-\hat{P}_i)/10000\right]$$

Their study indicates that for N = 10, $\hat{P}_{1L}$ and $\hat{P}_{2L}$ are significantly greater than 0.025 all the times and also $\hat{P}_{1U}$ and $\hat{P}_{2U}$.

For N = 30, 50; $\hat{P}_{1L}$, $\hat{P}_{2L}$ are significantly different from 0.025 but $\hat{P}_{1U}$, $\hat{P}_{2U}$ are almost non-significant all the times.

Contingency Table

81 contingency tables are made for the 81 combinations and for

each contingency table 3 conditional tests are made for comparing F_1 and F_2. Example follows which explains the procedure of tests in the tables.

Example for tests in contingency table

Consider the following contingency table

F_1 \ F_2	< -1.96	$-1.96 < F_2 < 1.96$	>1.96	TOTAL
< -1.96	417	2	0	419
$-1.96 < F_1 < 1.96$	8	9150	7	9165
> 1.96	0	1	415	416
TOTAL	425	9153	422	10000

(a) Comparison of $\hat{P}_{1L}$ and $\hat{P}_{2L}$

refer 8 to Bi(8+2, 1/2)

two sided P value = 0.1719 > 0.05 non-significant

(b) Comparison of $\hat{P}_{1U}$ and $\hat{P}_{2U}$

refer 1 to Bi(1+7, 1/2)

two sided P value = 0.070 > 0.05 non-significant

(c) Comparison of $P_{1U} + P_{1L}$, $P_{2U} + P_{2L}$

refer 8+7 to Bi(8+7+2+1, 1/2)

two sided P value = 0.0076 > 0.05 significant

(1) For N = 10, comparison of $\hat{P}_{1L}$, $\hat{P}_{2L}$ indicates that $\hat{P}_{1L} \leq \hat{P}_{2L}$ and is significant only when both B_N and C_N are high; otherwise non-significant.

For N = 30, 50; $\hat{P}_{1L} < \hat{P}_{2L}$ but usually non-significant.

(2) For N = 10; $\hat{P}_{1U} > \hat{P}_{2U}$ and is significant only when both B_N and C_N are high.

For N = 30, 50; $\hat{P}_{1U} > \hat{P}_{2U}$ and is non-significant.

(3) $\hat{P}_{1U} + \hat{P}_{1L} \leq \hat{P}_{2U} + \hat{P}_{2L}$ for N = 10, 30, 50.

For N =10; it is significant for both B_N and C_N high but for N = 30,

50 they are non-significant almost everywhere.

Inspection of the contingency tables indicates that confidence probability is almost 0.91 for N = 10 and for N = 30, 50, it is nearly 0.94 for both F_1 and F_2 in all the cases.

From the above results it is concluded that F_1 is better than F_2 for N = 10 and in case of N = 30, 50, they are almost equal. It is advisable to use F_1 always as it is simpler compared with F_2.

(b) *P(t) Exponential*

Point Estimates

N = 10

It is observed that

$\hat{P}_{1L} \leq \hat{P}_{2L}$

On the other hand

$\hat{P}_{1U} \geq \hat{P}_{2U}$

and

$\hat{P}_{1L} + \hat{P}_{1U} \leq \hat{P}_{2L} + \hat{P}_{2U}$

$\hat{P}_{1L}$ ranges between 0.030 - 0.056

$\hat{P}_{1U}$ ranges between 0.032 - 0.050

$\hat{P}_{2L}$ " " " 0.030 - 0.056

$\hat{P}_{2U}$ " " " 0.030 - 0.050

$\hat{P}_{1L} + \hat{P}_{1U} \leq \hat{P}_{2L} + \hat{P}_{2U}$ and is between 0.079 - 0.086

For α, β, $\sigma^2_{x|t}$ known and $\rho^2_{\bar{x}} = 0.7$, $P_L = 0.018$, $P_U = 0.032$ and $P_L + P_U = 0.050$.

N = 30, 50

There is less variation in $\hat{P}_{1L}$, $\hat{P}_{2L}$, $\hat{P}_{1U}$, $\hat{P}_{2U}$ and lower limits have tendency towards 0.018 and upper towards 0.032, the true values when α and β and $\sigma^2_{x|t}$ are known at $\rho^2_{\bar{x}} = 0.7$.

Interval Estimates

Lower tail probability 0.018 is not but upper tail probability 0.032 is mostly in the interval for N = 30, 50 . For N = 10 these

probabilities are not usually in the interval.

In the present situation both tails are being overestimated because of the uncertainty in the estimation of α, β, $\sigma^2_{x|t}$. Interval for t is also wider because of Q in the denominator, than when α, β, $\sigma^2_{x|t}$ are known.

Contingency Table

N = 10

$\hat{P}_{1L} < \hat{P}_{2L}$ and is significant only for both B_N and C_N high, otherwise non-significant.

$\hat{P}_{1U} > \hat{P}_{2U}$ and is significant for both B_N and C_N high only.

$\hat{P}_{1L} + \hat{P}_{1U} < \hat{P}_{2L} + \hat{P}_{2U}$ and is non-significant.

N = 30, 50

Same types of inequalities hold for both lower and upper tails as for N = 10 and are usually non-significant.

(c) <u>P(t) Uniform</u>

N = 10

$\hat{P}_{1L} \leq \hat{P}_{2L}$

$\hat{P}_{1U} \geq \hat{P}_{2U}$

$\hat{P}_{1L}$ ranges between 0.039 - 0.059

$\hat{P}_{2L}$ " " " 0.039 - 0.066

$\hat{P}_{1U}$ " " " 0.020 - 0.044

$\hat{P}_{2U}$ " " " 0.019 - 0.044

$\hat{P}_{1L} + \hat{P}_{1U} \leq \hat{P}_{2L} + \hat{P}_{2U}$ and ranges between 0.078 - 0.088.

When α, β, $\sigma^2_{x|t}$ are known, $\hat{P}_{1U} = \hat{P}_{2U} = \hat{P}_{1L} = \hat{P}_{2L} = 0.011$ for $\rho^2_{\bar{x}} = 0.7$.

From above it is clear that the lower and upper limits are always greater than 0.011. This is because of the uncertainty in the estimation of regression parameters from the regression experiment.

N = 30, 50

Point estimates $\hat{P}_{1L} \approx \hat{P}_{2L} \approx 0.030$ and $\hat{P}_{2U} \approx \hat{P}_{1U} \approx 0.030$ for

N = 30 and are about 0.022 for N = 50 most of the times and the variation decreases with the increase in N.

Interval Estimates

Lower and upper tail intervals never contain the value 0.011 either for F_1 or for F_2 in all the 81 cases.

Contingency Table

N = 10

In the contingency tables

$\hat{P}_{1L} \leq \hat{P}_{2L}$ and is significant only for both B_N and C_N high.

$\hat{P}_{1U} \geq \hat{P}_{2U}$ and is significant only when both B_N and C_N have high values.

$\hat{P}_{1L} + \hat{P}_{1U} < \hat{P}_{2U} + \hat{P}_{2L}$ and is non-significant all the times and confidence probability is 0.92 approximately.

N = 30, 50

Same types of inequalities hold as for N = 10 for lower and upper tail probabilities. Most of the times counts tend to be smaller but significance is still there for high B_N and C_N most of the times.

$\hat{P}_{1L} + \hat{P}_{1U}$ vs $\hat{P}_{2L} + \hat{P}_{2U}$ are mostly non-significant every where and confidence probability is 0.94 and 0.95 for N = 30 and 50 respectively.

(d) <u>P(t) Triangular</u>

Point Estimates

N =10

It is noted that

$\hat{P}_{1L} \leq \hat{P}_{2L}$ and

$\hat{P}_{1L}$ ranges from 0.038 - 0.054

$\hat{P}_{2L}$ " " " 0.038 - 0.060

$\hat{P}_{1U} \geq \hat{P}_{2U}$ and

$\hat{P}_{1U}$ ranges from 0.023 - 0.045

$\hat{P}_{2U}$ " " " 0.023 - 0.045

$\hat{P}_{1L} + \hat{P}_{1U} \leq \hat{P}_{2L} + \hat{P}_{2U}$ and range is 0.077 - 0.086

Upper and lower tail probabilities are always greater than 0.01, which is the value corresponding to $\rho^2 = 0.7$ when regression parameters are known, because of the uncertainty of estimation.

N = 30, 50

Same types of inequalities are observed for $\hat{P}_{iL}$, $\hat{P}_{iU}$, i = 1, 2 and the sum of lower and upper tail prababilities for both F_1 and F_2 is approximately 0.060 for N = 30 and 0.053 for N = 50.

Interval estimates

Both upper and lower tail confidence intervals never contain the value 0.01.

Contingency Table

N = 10

$\hat{P}_{1L} < \hat{P}_{2L}$ and is significant for high values of B_N and C_N at the same time.

$\hat{P}_{1U} > \hat{P}_{2U}$ and is significant when both B_N and C_N have high values.

$\hat{P}_{1L} + \hat{P}_{1U} < \hat{P}_{2L} + \hat{P}_{2U}$ and is non-significant.

The confidence probability is 0.92

N = 30 and 50

The same inequalities hold and are mostly non-significant. The confidence coefficients are 0.94 and 0.95 for N = 30 and 50 respectively for both F_1 and F_2.

It can be concluded from above that

(i) Error probabilities depend on N, B_N, C_N and to some extent on ρ^2.

(ii) Correcting for bias in F_1 (if any) leads to using F_2, F_3 or F_4 instead of F_1, but F_1 appears to be as good as any of these, in terms of error probabilities.

4.3. BOOTSTRAPPING

In the previous work it is assumed that conditionally on T, X is normal but in cases where we do not know the distribution of X, It remains to show how does it affect the mean squared error and ultimately confidence coefficient.

Bootstrapping is a resampling procedure to assess the accuracy of an estimator and is in fact computing power as a substitute for theoretical analysis.

Bootstrap algorithm is as follows

(1) We have pairs (x_i, t_i), $i = 1, 2, \ldots, N$ where X_i's are random and t_i's fixed. We call this regression experiment.

(2) Assign equal probabilities to each X_i for $i = 1, 2, \ldots, N$.

(3) Construct bootstrap sample $x^*_1, x^*_2, \ldots, x^*_N$ as follows

(a) Obtain $e_i = x_i - \hat{x}_i$

where $\hat{x}_i = \hat{\alpha} + \hat{\beta} t_i$

$\hat{\alpha}$ and $\hat{\beta}$ are the values of regression parameters estimated from the regression experiment.

(b) $x^*_i = \hat{\alpha} + \hat{\beta} t_i + e^*_i$

Where e^*_i is selected from $e_1, e_2, \ldots, e_N$ using sampling with replacement with the help of discrete uniform distribution between 1 and N from NAG.

(4) Calculate C^*, D^* from the bootstrap sample (x^*_i, t_i) $i = 1, 2, \ldots, N$ using μ and σ^2 calculated from the following relation with B_N and C_N known.

$$\mu = \bar{t} + [(B_N)\ S_{TT}/(N-2)]^{\frac{1}{2}}$$

$$\sigma^2 = [(C_N)(S_{TT})]/(N-2)$$

$\bar{t}$ and S_{TT} are calculated from the data.

(5) Repeat 3(b) and 4, B times and calculate

$$\hat{C}^* = \Sigma C^*_j/B \quad j = 1, 2, \ldots, B$$

$$\hat{D}^* = \Sigma D^*_j/B$$

$$\hat{VAR}(\hat{C}^*) = [\Sigma(C^*_j - \hat{C}^*)^2]/B$$

$$\hat{VAR}(\hat{D}^*) = [\Sigma(D^*_j - \hat{D}^*)^2]/B$$

$$\hat{COV}(\hat{C}^*, \hat{D}^*) = [\Sigma(C^*_j - \hat{C}^*)(D^*_j - \hat{D}^*)]/B$$

As B tends to infinity all these parameters tend to the true values.

(6) $\hat{MSE}^*$ is calculated by substituting these values in the expression (3.9) taking $EX = \hat{\alpha}^* + \hat{\beta}(\mu - \bar{t})$

$$EX^2 = \hat{\beta}^2\sigma^2 + \hat{\sigma}^2_{x|t} + (EX)^2$$

$$E(\hat{C} - C)^2 = VAR\hat{C} + (E\hat{C} - C)^2$$

and $$\hat{\rho}^2 = \hat{\beta}^2\sigma^2(\hat{\beta}^2\sigma^2 + \hat{\sigma}^2_{x|t})^{-1}$$

where α^*, β, $\sigma^2_{x|t}$, $\bar{t}$, C, D are estimated from the regression experiment.

4.3.1. EXAMPLES

Data of two examples from Aitchison and Dunsmore (1975), already discussed in chapter 1 have been taken to illustrate the bootstrap procedure in our calibration situation.

Example 1: Measuring water content of soil specimen

Plot of the data (x_i, t_i) i = 1, ..., 16 indicates linear relationship between X and T. This points out the suitability of our technique and there is no need to apply any transformation to make the relationship linear.

On the other hand $\hat{C}^*$, $\hat{D}^*$, $\hat{VARC}^*$, $\hat{VARD}^*$ and $\hat{COV}(C^*, D^*)$ are calculated to obtain the $\hat{MSE}^*$ using the formula (3.9) for the same 9 combinations. The results are given in table 4.7.

The bootstrap procedure is as under

(1) The data set (x_i, t_i) i = 1, 2, ..., 16.

(2) α^*, β are calculated by the usual formulae and $\sigma^2_{x|t}$ is calculated by RSS/N instead of RSS/(N-2), refer to 3(b).

(3) e_i's i = 1, 2, ..., 16 are calculated by

$$e_i = x_i - \hat{x}_i$$

where $\hat{x}_i = \hat{\alpha} + \hat{\beta}t_i$

(4) Bootstrapped $x^*_1, x^*_2, \ldots, x^*_{16}$ are obtained by the relation

$$x^*_i = \hat{\alpha} + \hat{\beta}t_i + e^*_i$$

where e^*_i is selected from the $e_1, e_2, \ldots, e_{16}$ with replacement using discrete uniform distribution between 1 and 16 from NAG.

(5) (x_i^*, t_i) $i = 1, 2, \ldots, 16$ are used to calculate $\hat{C}^*$, $\hat{D}^*$ and $\hat{MSE}^*$ by the relation given in the bootstrap algorithm.

Nine values of C, D and MSE are calculated for 9 different values of μ and σ^2 which arose as a result of three values of each B_N = 0.0, 1.0, 4.0 and C_N = 0.25, 1.0, 4.0, using our formulae (3.13).

On the other hand $\hat{C}^*$, $\hat{D}^*$, $\hat{VAR}C^*$, $\hat{VAR}D^*$ and $\hat{COV}(C^*, D^*)$ are calculated to obtain the $\hat{MSE}^*$ using the formula (3.9) for the same 9 combinations. The results are given in table 4.7.

A look at the table indicates that there is quite a good agreement between the two set of parameter values calculated.

Example 2: <u>Antibiotic Assay</u>

The data (d_i, c_i) $i = 1, 2, \ldots, 120$ as in Aitchison and Dunsmore (1975) shows a non-linear relationship between d_i(diameter cleared) and c_i(concentration). As our approach works for linear relationship, following criteria were applied to achieve the linearity and equality of variances.

(1) For the response variable Box-Cox (1964) transformation is applied as under. Let

$$W = \begin{cases} (d^{\lambda} - 1)/\lambda & \text{for } \lambda \neq 0 \\ \ln d & \text{for } \lambda = 0 \end{cases}$$

We used the data to estimate parameter λ as well as the regression parameters in the model to be fitted,

$$W = \alpha + \beta c + \epsilon \qquad \ldots(4.13)$$

and

$$L_{max}(\lambda) = -N/2(RSS/N) + (\lambda - 1)\sum_{1}^{N} \ln(d_i)$$

where RSS is residual sum of squares after fitting the model (4.13).

Several values of λ and $L_{max}(\lambda)$ are given below

λ	$L_{max}(\lambda)$
2.0	-119.625
2.5	-113.828
3.0	-109.964
3.5	-107.878
4.0	-107.415
4.5	-108.440
5.0	-110.806
5.5	-114.394

This gives $\lambda = 4$ as the appropriate transformation but d^4 against c gives unequal variances as

concentration	Mean	S.D
1	14364	7096
2	84208	20647
4	139539	25320
8	189912	31087
16	223066	36683
32	229635	44424

To get the equality of variances, d_i^2 was tried against c_i which gave the following values.

concentration	Mean	S.D
1	115.22	33.8
2	288.25	34.3
4	372.19	32.7
8	434.39	35.8
16	470.66	40.3
32	477.25	44.2

d_i^2 looks appropriate for equality of variances and then to get the linearity Box-Tidwell (1962) transformation is applied on c_i.

let

$$V = \begin{cases} c_i^A & A \neq 0 \\ \ln c_i & A = 0 \end{cases}$$

and the linear model is

$$d^2 = B_0 + B_1 V + \epsilon \qquad \ldots(4.14)$$

We define $Z = V \log V$ and the fitted model is

$$d^2 = \hat{B}_0 + \hat{B}_1 V + \hat{\gamma} Z \qquad \ldots(4.15)$$

The approximate estimate of A is obtained by the following relation starting usually with $A = 1$

$$\hat{A} = [(\hat{\gamma}/B_1)+1)](\text{current value of } \hat{A}) \qquad \ldots(4.16).$$

The procedure with $\hat{A}$ obtained from (4.16) is repeated until the decrease in residual sum of squares ($RSS_{A)}$ is small. This transformation gave the following results.

Value of $\hat{A}$	RSS_A
1.000	1135373
-1.55076	253984
-0.611635	173165
-0.807527	159560
-0.817510	159531
-0.740777	161299
-0.816020	159531
-0.817573	159531
-0.817574	159531
-0.817574	159531

Thus the linear relationship is $d^2 = \alpha + \beta/c$ with almost equal variances. So $x_i = d_i^2$ and $t_i = 1/c_i$.

Using the data $(d_i^2, 1/c)$ $i = 1, 2, \ldots, 120$, the same set of 9 parameters as in the previous example 1 are calculated by the same procedure. The results are given in table 4.7

It is evident from the table 4.7 that the values of the parameters obtained by bootstrapping are in good agreement with the values calculated using our formula obtained by Taylor's series approximations.

4.3.2. EFFECT OF NON-NORMALITY OF ERRORS

In the data of the above two examples errors are approximately normal. To show the effect of the non-normality of errors, data with shifted exponential errors $\sigma_{x|t}(v-1)$ where v is standard exponential

(e^{-v}) were simulated by generating errors from the exponential distribution using NAG.

C, D and MSE are calculated for both the examples by both the methods i.e. our method and bootstrapping. The results are produced in table 4.7 for example 1 and example 2.

Comparison of table mean squared errors indicates that the corresponding values are close by both the methods for all the 9 combinations of B_N and C_N although errors are strongly non-normal

This clearly indicates that our method works quite well whatever is the distribution of errors.

TABLE 4.1

(a) Upper and Lower tail probabilities for the exponential

$P(t) = e^{-t}$ ($t \geq 0$) i.e. with $\mu = \sigma^2 = 1$.

ρ^2	upper tail prob.	lower tail prob.	total
0.10	0.049	0.000	0.049
0.20	0.047	0.001	0.048
0.40	0.041	0.007	0.048
0.50	0.038	0.011	0.049
0.60	0.035	0.015	0.050
0.70	0.032	0.018	0.050
0.80	0.029	0.021	0.050
0.90	0.027	0.023	0.050
0.99	0.025	0.025	0.050

(b) The values of $\rho^2 = \beta^2\sigma^2/(\beta^2\sigma^2 + \sigma^2_{\bar{x}|t})$ calculated by fixing the values of regression parameters i.e $\beta^2/\sigma^2_{\bar{x}|t}$ and changing the scale of the exponential distribution.

$\beta^2/\sigma^2_{\bar{x}\|t}$ $P(t)$	0.67	1.0	2.33	4.0	9.0
$(1/8)e^{-t/8}$	0.977	0.984	0.993	0.996	0.998
$(1/4)e^{-t/4}$	0.91	0.94	0.97	0.98	0.99
$(1/2)e^{-t/2}$	0.72	0.80	0.90	0.94	0.97
e^{-t}	0.40	0.50	0.70	0.80	0.90
$2e^{-2t}$	0.14	0.20	0.37	0.50	0.69
$4e^{-4t}$	0.04	0.05	0.13	0.20	0.36
$8e^{-8t}$	0.01	0.02	0.04	0.06	0.12

TABLE 4.2

Upper (U) and Lower (L) tail probabilities for different scales of exponential distribution for fixed values of the parameters of the experiment i.e. $\beta^2/\sigma^2_{\bar{x}|t}$.

$\beta^2/\sigma^2_{\bar{x}\|t}$		0.67	1.0	2.33	4.0	9.0
$P(t)$						
$(1/8)e^{-t/8}$	U	0.025	0.025	0.025	0.025	0.025
	L	0.025	0.025	0.025	0.025	0.025
$(1/4)e^{-t/4}$	U	0.026	0.026	0.026	0.025	0.025
	L	0.024	0.024	0.024	0.025	0.025
$(1/2)e^{-t/2}$	U	0.031	0.029	0.026	0.026	0.025
	L	0.019	0.021	0.024	0.024	0.025
e^{-t}	U	0.041	0.038	0.032	0.029	0.026
	L	0.007	0.011	0.018	0.021	0.024
$2e^{-2t}$	U	0.048	0.046	0.042	0.038	0.032
	L	0.003	0.001	0.006	0.011	0.017
$4e^{-4t}$	U	---	---	0.048	0.046	0.042
	L	---	---	0.003	0.001	0.006
$8e^{-8t}$	U	---	---	---	---	0.048
	L	---	---	---	---	0.000

Missing values in the table correspond to very low values of $\rho^2 < 0.10$. (numerical integration not done).

TABLE 4.3

Lower and upper tail probabilities for $P(t)$ as uniform and triangular distributions.

(a) $P(t)$ Uniform $\quad P(t) = 1/(b - a) \quad a \leq t \leq b$

ρ^2	lower tail	upper tail	total
0.10	0.000	0.000	0.000
0.20	0.000	0.000	0.000
0.30	0.001	0.001	0.002
0.40	0.002	0.002	0.004
0.50	0.004	0.004	0.008
0.60	0.007	0.007	0.014
0.70	0.011	0.011	0.022
0.80	0.015	0.015	0.030
0.90	0.020	0.020	0.040
0.99	0.024	0.024	0.048

(b) $P(t)$ Triangular $\quad P(t) = \begin{cases} 4t & 0 < t \leq 1/2 \\ 4(1-t) & 1/2 < t \leq 1 \end{cases}$

0.10	0.000	0.000	0.000
0.20	0.000	0.000	0.000
0.30	0.000	0.000	0.000
0.40	0.001	0.001	0.002
0.50	0.003	0.003	0.006
0.60	0.006	0.006	0.012
0.70	0.010	0.010	0.020
0.80	0.014	0.014	0.028
0.90	0.019	0.019	0.028
0.99	0.024	0.024	0.048

TABLE 4.4

Values of $\rho^2 = \beta^2\sigma^2/(\beta^2\sigma^2 + \sigma^2_{\bar{x}|t})$ by fixing the values of $\beta^2/\sigma^2_{\bar{x}|t}$ and changing the scale of the uniform and symmetrical triangular distribution.

(a) *P(t) Uniform*

$\beta^2/\sigma^2_{\bar{x}\|t}$	0.67	1.00	1.50	2.33	4.00	9.00
$\sigma^2 = (b-a)^2/12$						
1/12	0.05	0.08	0.11	0.16	0.25	0.40
4/12	0.18	0.29	0.33	0.43	0.59	0.75
16/12	0.47	0.57	0.67	0.76	0.84	0.92
36/12	0.67	0.75	0.82	0.87	0.92	0.96
64/12	0.78	0.84	0.89	0.93	0.96	0.98
144/12	0.89	0.92	0.95	0.97	0.98	0.99

(b) *P(t) Triangular*

$\sigma^2 = (b-a)^2/24$						
1/24	0.03	0.04	0.06	0.09	0.14	0.27
4/24	0.10	0.14	0.20	0.28	0.40	0.60
16/24	0.31	0.40	0.50	0.61	0.73	0.86
36/24	0.50	0.60	0.69	0.78	0.86	0.93
64/24	0.64	0.73	0.80	0.86	0.91	0.96
144/24	0.80	0.86	0.90	0.93	0.96	0.98

TABLE 4.5

Probabilities (same for both tails) for different scales of uniform and triangular distribution for fixed values of the parameters of the regression experiment i.e $\beta^2/\sigma^2_{\bar{x}|t}$.

(a) P(t) Uniform

$\beta^2/\sigma^2_{\bar{x}\|t}$	0.67	1.00	1.50	2.33	4.00	9.00
$\sigma^2 = (b-a)^2/12$						
1/12	0.000	0.000	0.000	0.000	0.000	0.002
4/12	0.000	0.000	0.000	0.003	0.007	0.013
16/12	0.003	0.006	0.010	0.015	0.017	0.020
36/12	0.001	0.013	0.016	0.018	0.021	0.023
64/12	0.014	0.017	0.020	0.021	0.023	0.024
144/12	0.020	0.021	0.022	0.023	0.024	0.024

(b) P(t) Triangular

$\sigma^2 = (b-a)^2/24$

1/24	0.000	0.000	0.000	0.000	0.000	0.000
4/24	0.000	0.000	0.000	0.000	0.001	0.006
16/24	0.000	0.001	0.003	0.006	0.011	0.017
36/24	0.003	0.006	0.010	0.013	0.017	0.021
64/24	0.008	0.011	0.014	0.017	0.020	0.023
144/24	0.014	0.017	0.020	0.021	0.023	0.024

TABLE 4.6

Summary statistics for F_1, F_2, F_3 and F_4 based on a sample of 1000 simulated values of each with different $P(t)$ and $B_N = 4.0$, $C_N = 4.0$, $N = 10$, $\rho^2_{\bar{x}} = 0.7$.

(a) *P(t) Normal and Exponential*

	Normal				Exponential			
	F_1	F_2	F_3	F_4	F_1	F_2	F_3	F_4
Mean	-0.123	-0.154	-0.123	-0.155	-0.075	-0.101	-0.074	-0.103
Med	-0.039	-0.062	-0.040	-0.062	-0.066	-0.083	-0.067	-0.083
S.D	1.149	1.179	1.156	1.171	1.097	1.136	1.103	1.124
MIN	-8.363	-8.394	-8.366	-8.391	-4.675	-4.770	-4.690	-4.754
MAX	5.789	5.716	5.820	5.686	4.320	4.302	4.320	4.301

(b) *P(t) Uniform and Triangular*

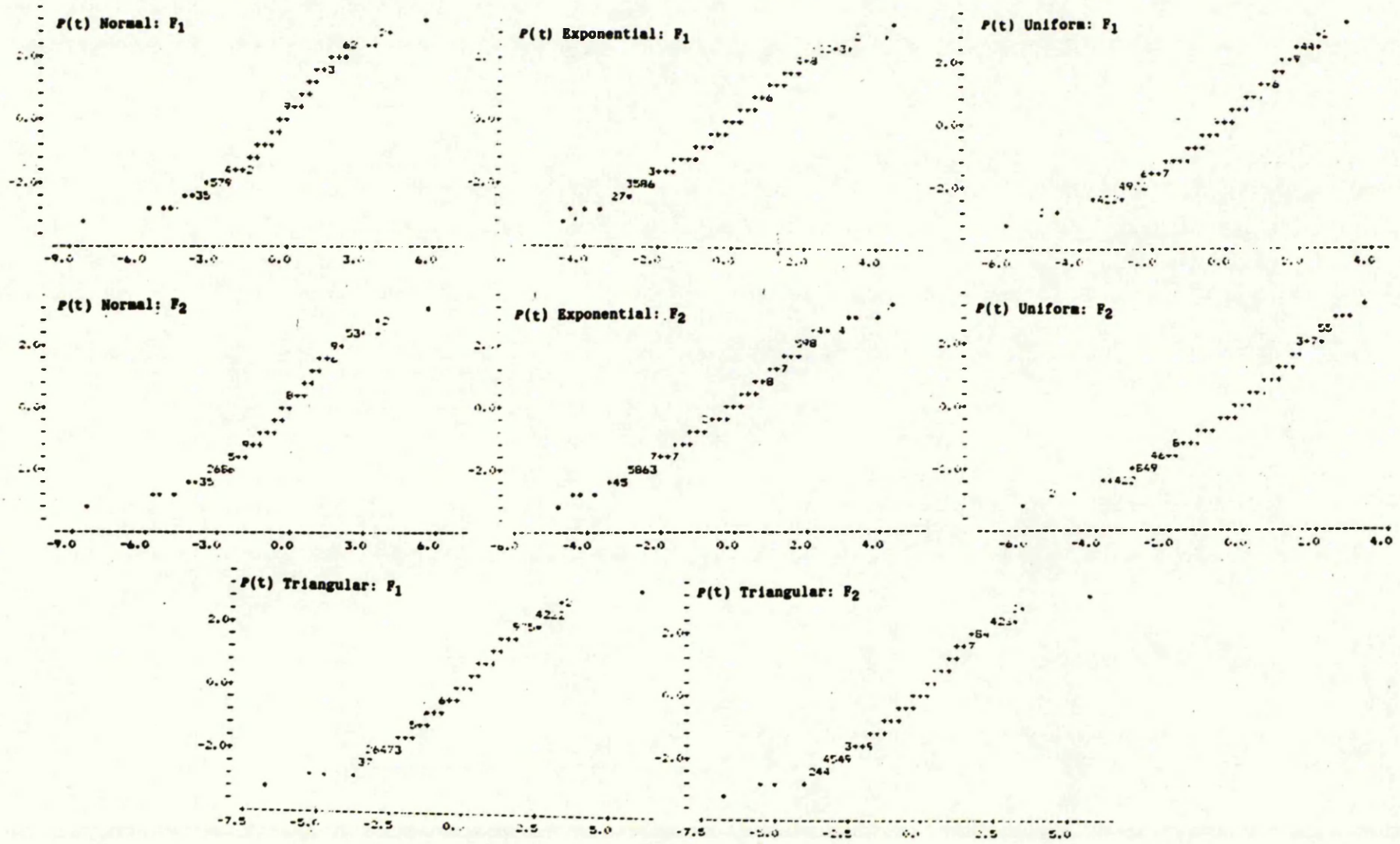
	Uniform				Triangular			
	F_1	F_2	F_3	F_4	F_1	F_2	F_3	F_4
Mean	-0.066	-0.093	-0.066	-0.094	-0.091	-0.116	-0.091	-0.117
Med	0.032	0.012	0.032	0.012	-0.015	-0.015	-0.015	-0.015
S.D	1.114	1.146	1.121	1.138	1.161	1.189	1.167	1.182
MIN	-5.979	-6.076	-5.997	-6.057	-6.688	-6.733	-6.678	-6.725
MAX	3.488	3.470	3.488	3.470	5.913	5.845	5.939	5.820

TABLE 4.7

MSE by our method and by bootstrapping for example 1 and example 2

		$P(x\mid t)$ empirical		$P(x\mid t)$ exponential	
MU	VART	$MSE_{(our)}$	$MSE_{(boot)}$	$MSE_{(our)}$	$MSE_{(boot)}$
(a) Example 1					
26.300	19.045	2.8896	2.8898	2.7792	2.7976
26.300	76.180	3.4363	3.3811	3.2903	3.2162
26.300	304.720	4.1999	4.1635	4.0162	3.9943
35.028	19.045	2.9637	3.0658	2.8526	2.9737
35.028	76.180	3.5723	3.5963	3.4214	3.4512
35.028	304.720	4.3583	4.3908	4.1680	4.1326
43.756	19.045	3.2601	3.5861	3.1463	3.4556
43.756	76.180	4.1164	4.2382	3.9461	4.0568
43.756	304.720	4.9921	5.0707	4.7751	4.7425
(b) Example 2					
0.3281	0.0291	0.0075	0.0075	0.0069	0.0069
0.3281	0.1164	0.0094	0.0093	0.0084	0.0084
0.3281	0.4658	0.0102	0.0102	0.0091	0.0091
0.6694	0.0291	0.0075	0.0076	0.0069	0.0066
0.6694	0.1164	0.0094	0.0094	0.0084	0.0084
0.6694	0.4658	0.0103	0.0103	0.0091	0.0092
1.0106	0.0291	0.0076	0.0077	0.0069	0.0071
1.0106	0.1164	0.0096	0.0097	0.0086	0.0087
1.0106	0.4658	0.0105	0.0105	0.0093	0.0093

FIGURE 4.1. Normal Probability Plots for F_1 and F_2 with different $P(t)$ and $B_N = 4.0$, $C_N = 4.0$, $N = 10$, $\rho^2 = 0.7$.



CHAPTER 5

MULTIVARIATE CALIBRATION

5.1. INTRODUCTION

Multivariate calibration is a topic of current interest and has many practical applications. The aim is to make inferences about a p-vector T from an observed q-vector X where $p \leq q$. The relationship between the response vector X and vector of explanatory variables T is determined from the data of the multivariate regression experiment (X_i, T_i), $i = 1, 2, \ldots, N$, where X_i and T_i are q×1 and p×1 vectors respectively. In this experiment the T_i's are fixed and X_i's are random.

Consider the multivariate linear regression model

$$X = T\beta^{\#} + \epsilon \qquad \ldots.(5.1)$$

where X is a N×q matrix of q response variables for each of N individuals. T is N× (p+1) matrix whose first column consists of 1's, the other columns listing p explanatory variables measured on N individuals. $\beta^{\#}$ is a (p+1)×q matrix of regression parameters and ϵ is a matrix of N×q random errors whose rows ϵ_i^T are independent and normally distributed with

$$E\epsilon_i = 0$$

and

$$E(\epsilon_i \epsilon_i^T) = \Gamma$$

The maximum likelihood estimates of $\beta^{\#}$ and Γ are

$$\hat{\beta}^{\#} = (T^T T)^{-1} T^T X$$

$$\hat{\Gamma} = \{X^T(I - T(T^T T)^{-1}T^T)X\}/N \qquad \ldots(5.2)$$

The unbiased estimate of Γ is

$$\hat{\Gamma} = \{X^T(I - T(T^T T)^{-1}T^T)X\}/(N-p-1) \qquad \ldots(5.3)$$

In the future situation a q×1 vector X is observed and p×1 vector T is to be predicted using the multivariate regression experiment.

Both the classical and inverse estimator have been studied by Brown (1982) along with the extension to the Lwin and Maritz (1980) approach.

If $\beta^{\#}$ is partitioned $\begin{bmatrix} \alpha^T \\ \cdots \\ \beta^T \end{bmatrix}$, where β is q×p,

the classical or maximum likelihood estimator for p×1 vector T is

$$\check{T} = \bar{T} + (\hat{\beta}^T \hat{\Gamma}^{-1} \hat{\beta})^{-1} \hat{\beta}^T \hat{\Gamma}^{-1} (X - \bar{X}) \qquad \ldots(5.4)$$

and the inverse estimator is

$$\tilde{T} = \bar{T} + S_{TX} S^{-1}{}_{XX} (X - \bar{X}) \qquad \ldots(5.5)$$

where S_{TX} is p×q matrix of sums of products corrected for the mean and S_{XX} is q×q matrix of sums of corrected squares and products.

We give the derivation of the best linear predictor in section 5.3 for general q and p = 1. It is interesting to note that the best linear predictor in this situation also gives classical estimator for $\sigma^2 = \infty$ and inverse estimator for $\mu = \bar{t}$ and $\sigma^2 = S_{TT}/(N-2)$ like the simple linear calibration problem discussed in chapter 1. The proof is given in appendix B.

For p = q = 1, the model (5.1) becomes the simple linear regression model and the whole calibration situation and the estimators become univariate estimators already discussed in section 1.2 and 1.7.

EXAMPLES

A very common practical example is to get the estimate of age based on different body measurements in the current situation Wood (1982), Oman and Wax (1984). The relationship between X and T is estimated from the data of the regression experiment (X,T) where X is a q-vector of response variables and T is a p-vector of explanatory variables, in this situation p is usually 1, so T is 1×1 i.e. (scalar).

Another example is by Brown (1982), analysed as one of his examples, on wheat quality data where p = 2 are the accurately determined measurements on moisture and protein content and q = 4 are the derived infrared reflectance measurements at different wavelengths. He estimated the moisture and protein content from the observations on infrared reflectance measurements. We will discuss this example in detail at the end of chapter 6. Some other examples are given in Williams (1959).

5.2. REVIEW

Most of the existing literature originates from early 1980's and is briefly described below.

Henschke (1980) constructed simultaneous confidence intervals for the multivariate linear calibration of classical type. Two of the three methods discussed are the extensions of methods first discussed by Miller (1966) based on Bonferroni inequality and the union intersection principle for univariate case. Third method developed is only applicable to univariate case.

Brown (1982) was the first to discuss the multivariate linear calibration problem in detail. He considered both the classical and the Bayesian approaches along with the multivariate extension of Lwin and Maritz (1980) and compared the results of the three approaches.

The results regarding the comparison of three approaches in table 3 of his paper are based on small samples of size five similar to simulations of size five so are not reliable. Large samples are needed for comparison particularly for Lwin and Maritz (1980).

In classical type calibration the interval estimates sometimes become empty or disjoint like the simple linear calibration problem.

He distinguished the calibration as controlled or natural depending upon T whether it is controlled or random respectively in

the experiment.

He has suggested that it is beneficial to treat the characteristics of the explanatory variable one at a time, forgetting the existence of the other p-1 variables.

Brown and Sundberg (1987) looked into the controlled calibration problem from the point of view of the profile likelihood function and compared confidence intervals with Brown (1982). When an inconsistency statistic is large in this method, there would be large regions for T.

Wood (1982) proposed an alternative method to overcome the difficulty of empty confidence intervals in the case of controlled calibration problem. He partitioned the quadratic form in T, obtained from the log-likelihood, into two parts, the first part showing the consistency of X with the model and second part a suitable expression for non-empty confidence intervals. The distribution of second part is quite complex and may be approximated by F distribution asymptotically. This method gave smaller confidence intervals compared with Brown (1982). His method can also be applied to non-linear models that can be approximated by linear models within approximate intervals.

Sjöström et al. (1983) described the use of partial least squares (PLS) in latent variables for multivariate calibration problems in analytical chemistry. They compared this method with principal component analysis combined with multiple regression and concluded that PLS approach has some obvious advantages over the traditional approach.

Oman and Wax (1984) solved a specific problem by applying Brown's classical approach (1982). They estimated the gestational age using femur length F and the biparietal diameter BPD individually and combining F and BPD. They discussed model choice in detail and also

tried quadratic regression making p = 2 instead of one.

Oman (1985a) also discussed the classical interval estimates following Brown (1982) and suggested some changes.

Naes (1985a) compared the classical and Bayesian approach (assuming the distribution of T only) in multivariate linear calibration. He took the risk function as criterion of comparison. He again (1985b, 86) considered calibration situations with error covariance matrix having linear factor structure and covariance adjustment respectively. He proposed new solution to this problem. Naes and Martens (1984), Naes et al. (1986) and Martens and Naes (1984) have described and clarified with examples the situations for multivariate calibration and applied different techniques to NIR (near infrared) instruments.

Fujikoshi and Nishii (1984) derived an asymptotic expansion up to order N^{-2} based on chi-squared percentiles for the distribution function of the statistic which is quadratic in T, proposed by Wood (1982) for confidence intervals. They (1986) obtained the asymptotic expression for bias and mean squared error of the classical estimator by expanding the estimate by Taylor's series. They used this information and the Akaike's criterion for selection of the best subset and compared them by applying on the wheat quality data analysed by Brown (1982). Nishii (1986) derived the cross validation criterion and obtained the asymptotic properties of it and the two criteria of Fujikoshi and Nishii (1986).

Spezzaferri (1985) used the Shannon information to derive the distribution of T given data and X. He adopted the Bayesian approach and solved a problem of choosing among k different calibration experiments associated with k different instruments considering equal and unequal costs for the instruments.

Sundberg (1985) compared classical and inverse estimators on the basis of mean squared error and derived the regions where inverse estimator has smaller mean squared error. This is the generalization of Berkson's (1969) work in the univariate case. Sundberg and Brown (1985) investigated unique natural extensions of the traditional solutions to estimation and prediction problem when there are more variables than observations.

5.3 DERIVATION OF BEST LINEAR PREDICTOR (p = 1 and general q)

Very often p = 1 is of interest in practical situations as we have seen above in the first example in section 5.1. It has also been suggested by Brown (1982) and Brown and Sundberg (1987) to consider the one explanatory variable at a time forgetting the existence of the other p-1 variables.

Consider the multivariate normal linear regression model with response q-vector X and an explanatory variable T

$$\underset{q\times 1}{X_i} = \underset{q\times 1}{\alpha} + T_i\underset{q\times 1}{\beta} + \underset{q\times 1}{\epsilon_i} \qquad i = 1,2, \ldots, N.$$

$$E(\epsilon_i) = 0$$

and $$E(\epsilon_i \epsilon_i^T) = \Gamma$$

but ϵ_i are independent for i = 1, 2, ..., N.

In the future situation (X,T), where q-vector X is observed and T is to be predicted; the joint distribution $P(x,t)$ is such that

$P(x|T=t)$ is same as above i.e. $N(\alpha + T\beta, \Gamma)$.

$P(t)$ is such that $E(T) = \mu$ and $VAR(T) = \sigma^2$ are known.

Future situation with p = 1 is described by the following parameters.

$$\underset{q\times 1}{\alpha}, \quad \underset{q\times 1}{\beta}, \quad \underset{q\times q}{\Gamma}, \quad \mu, \quad \sigma^2$$

These parameters define all the first and second order moments of future $P(x,t)$.

$$E(X_j) = EE(X_j|T) = E(\alpha_j + T\beta_j)$$
$$= \alpha_j + \mu\beta_j \qquad j = 1, 2, \ldots, q$$

$$E(T) = \mu$$

$$VAR(X_j) = VAR[E(X_j|T)] + E[VAR(X_j|T)]$$
$$= VAR(\alpha_j + \beta_j T) + \Gamma_{jj}$$
$$= \beta_j{}^2\sigma^2 + \Gamma_{jj}$$

$$COV(X_j, X_k) = COV[E(X_j|T), E(X_k|T)] + E[COV(X_j, X_k|T)]$$
$$= COV(\alpha_j+\beta_j T,\ \alpha_k+\beta_k T) + \Gamma_{jk}$$
$$= \beta_j\beta_k\sigma^2 + \Gamma_{jk}$$

$$COV(X) = \sigma^2\beta\beta^T + \Gamma \qquad COV(X) \text{ is a } q\times q \qquad \ldots(5.6)$$

$$VAR(T) = \sigma^2$$

$$COV(X_j, T) = COV(\alpha_j+\beta_j T+\epsilon_j, T) \qquad \epsilon_j \text{ is independent of } T$$
$$= \beta_j\ \sigma^2$$

Conversely α, β, Γ may be written in terms of the moments.

To derive the best linear predictor we minimize

$$E[T - (C + D^T X)]^2 \qquad \ldots(5.7)$$

Expression (5.7) is minimized by

$$C = E(T) - D^T E(X)$$
$$= \mu - D^T(\alpha + \mu\beta) \qquad \ldots(5.8)$$

For given D mean squared error is thus

$$VAR(T - D^T X) = VAR(T) - 2COV(T, D^T X) + VAR(D^T X)$$
$$= \sigma^2 - \underset{1\times q\ \ q\times 1}{2D^T COV(T, X)} + \underset{1\times q\ \ q\times q\ \ q\times 1}{D^T COV(X) D} \qquad \ldots(5.9)$$

differentiating this quadratic in D

$$0 = -2COV(T, X) + 2COV(X)D$$

Thus

$$D = \{COV(X)\}^{-1}\ COV(T, X)$$
$$= \sigma^2\{\Gamma + \sigma^2\beta\beta^T\}^{-1}\beta \qquad \ldots(5.10)$$

so best linear predictor $C + D^T X$ is

$$C + D^T X = E(T) + D^T[X - E(X)]$$

$$= \mu(1 - D^T\beta) + D^T(X - \alpha)$$

$$= \mu(1 - \rho^2) + (X - \alpha)^T\sigma^2\{\Gamma + \sigma^2\beta\beta^T\}^{-1}\beta \quad \ldots(5.11)$$

where $\rho^2 = D^T\beta$.

5.4. BIAS AND MEAN SQUARED ERROR

There are two situations

(1) α, β and Γ known

(2) α, β and Γ unknown

(a) BIAS

(1) α, β and Γ known

$$\text{bias} = E(T - (C + D^TX)$$

$$= \mu - C - D^TE(X)$$

$$= \mu - C - D^T(\alpha + \beta\mu)$$

$$= 0 \qquad \text{(using value of C from (5.8))}$$

(2) α, β and Γ unknown

$$\text{bias} = E[T - (\hat{C} + \hat{D}^TX)]$$

$$= E[T - (\hat{C} + \hat{D}_1X_1 + \hat{D}_2X_2 + \ldots + \hat{D}_qX_q]$$

$$= \mu - E(\hat{C}) - E(\hat{D}_1)E(X_1) + \ldots + E(\hat{D}_q)E(X_q)$$

$$= \mu - E(\hat{C}) - E(\hat{D})^T\{\alpha + \beta\mu\}$$

$$= \mu - E(\hat{C}) - D^T\{\alpha + \beta\mu\} - [E\hat{D} - D]^T\{\alpha + \beta\mu\}$$

$$= -[E(\hat{C}) - C] - [E\hat{D} - D]^T\{\alpha + \beta\mu\}$$

$$\neq 0 \qquad \ldots(5.12)$$

so the estimator is biased in this situation. We will discuss it further in theorem 5.2 where it is proved that bias/σ depends upon the same invariants as the MSE/σ^2.

(b) MEAN SQUARED ERROR

(1) α, β and Γ known

Substituting the value D from (5.10) in (5.9) we get

$$\text{MSE} = \sigma^2 - D^TCOV(T,X)$$

$$= \sigma^2[1 - \sigma^2\beta^T(\Gamma + \sigma^2\beta\beta^T)^{-1}\beta]$$

$$= \sigma^2(1 - \rho^2)$$

by definition of ρ^2.

(2) α, β and Γ unknown

$$\text{MSE} = E(T - \hat{C} - \hat{D}^T X)^2$$

$$= EE[(T - \hat{C} - \hat{D}^T X)^2 | \hat{C}, \hat{D}]$$

$E[(T - \hat{C} - \hat{D}^T X)^2 | \hat{C}, \hat{D}]$ is quadratic in $\hat{C}$, $\hat{D}$ and is minimised by $\hat{C} = C$ and $\hat{D} = D$ and its minimum is $\sigma^2(1 - \rho^2)$.

Thus

$$E[(T - \hat{C} - \hat{D}^T X)^2 | \hat{C}, \hat{D}] = \begin{bmatrix} \hat{C} - C \\ \hat{D} - D \end{bmatrix}^T M \begin{bmatrix} \hat{C} - C \\ \hat{D} - D \end{bmatrix} + \sigma^2(1 - \rho^2)$$

where M is a $(q+1)\times(q+1)$ symmetric matrix

$$\begin{bmatrix} 1 & EX_1 \ldots & EX_q \\ & EX_1^2 \ldots & EX_1X_q \\ \text{symmetric} & \ddots & \\ & & EX_q^2 \end{bmatrix}$$

Now

$$\text{MSE} = \sigma^2(1 - \rho^2) + E \text{ trace } MN$$

where

$$N = \begin{bmatrix} \hat{C} - C \\ \hat{D} - D \end{bmatrix} \begin{bmatrix} \hat{C} - C & \hat{D}^T - D^T \end{bmatrix}$$

Finally

$$\begin{aligned} \text{MSE} = {} & E(\hat{C} - C)^2 + 2(EX_1)E(\hat{C} - C)(\hat{D}_1 - D_1) + \ldots + 2(EX_q)E(\hat{C} - C)(\hat{D}_q - D_q) \\ & + E(X_1^2)E(\hat{D}_1 - D_1)^2 + \ldots + 2(EX_1X_q)E(\hat{D}_1 - D_1)(\hat{D}_q - D_q) \\ & + \ldots \\ & + \ldots \qquad + E(X_q^2)E(\hat{D}_q - D_q)^2 \\ & + \sigma^2(1 - \rho^2) \qquad \ldots(5.13) \end{aligned}$$

THEOREM 5.1. MSE/σ^2 DEPENDS ONLY UPON N, B_N, C_N, ρ and q when p=1

It would be proved in five steps.

STEP 1:

MSE/σ^2 depends upon

(i) the unconditional moments μ, σ^2, EX, $COV(X)$, $\beta\sigma^2$

and

(ii) the parameters of the distribution of $(\hat{C},\hat{D})$.

PROOF

Note that ρ^2 is merely a function of (i) so MSE/σ^2 depends only on

(i) μ, σ^2, α, β, Γ

and

(ii) the distribution of $\hat{\alpha}$, $\hat{\beta}$, $\hat{\Gamma}$

(Anderson (1984) theorem 8.2.2)

so MSE/σ^2 depends only on

μ, σ^2, α, β, Γ; $\bar{t}$, S_{TT}, N.

STEP 2:

MSE/σ^2 not changed by

(i) Changes of origin/scale of T

(ii) Changes of origin of X

(iii) $X \rightarrow HX$ where H is non-singular.

These may be proved as in step 2 of theorem 5.2.

STEP 3:

MSE/σ^2 depends only on

(q), N, ρ, $(\bar{t} - \mu)/(S_{TT}/(N - 2))^{\frac{1}{2}}$, $(N - 2)\sigma^2/S_{TT}$; $\bar{t}$, S_{TT}, α, β, Γ.

PROOF

This follows at once from step 1. Note that (q), N, ρ, $(\bar{t}-\mu)/(S_{TT}/(N - 2))^{\frac{1}{2}}$ and $(N - 2)\sigma^2/S_{TT}$ are invariants for

transformation (i), (ii) and (iii) of step 2.

STEP 4:

Consider two calibration situations or systems A and B which have the same values of q, N, ρ, $(B_N)^{\frac{1}{2}}$ and C_N.

By transformation of T, X of types mentioned in step 2, system A becomes system A' with

$$\bar{t} = 0,\ S_{TT} = N-2,\ \alpha = 0,\ \Gamma = I$$

$$\beta = \left[\frac{\rho}{[(N-2)\sigma^2/S_{TT}]^{\frac{1}{2}}(1-\rho^2)^{\frac{1}{2}}},\ 0,\ 0,\ \ldots,\ 0\right]^T \qquad \ldots(5.14)$$

Possibility of this transformation is shown in step 5.

By transformation of similar type, system B becomes system $\bar{B}$ with

$$\bar{t} = 0,\ S_{TT} = N-2,\ \alpha = 0,\ \Gamma = I$$

and β is same as above in relation (5.14).

By step 3, A' and B' agree in all quantities on which MSE/σ^2 depends.

By step 2,

$$(MSE/\sigma^2)_A = (MSE/\sigma^2)_{A'}$$

and

$$(MSE/\sigma^2)_B = (MSE/\sigma^2)_{B'}$$

thus

$$(MSE/\sigma^2)_A = (MSE/\sigma^2)_B.$$

In other words MSE/σ^2 depends only on q, N, ρ, (B_N), C_N.

STEP 5:

To show the possibility of $A \to A'$, by (i) ,(ii) and/or (iii) of step 2, where A' has $\bar{t} = 0$, $S_{TT} = N-2$, $\alpha = 0$, $\Gamma = I$ and β is as in (5.14) in step 4.

Firstly $\bar{t} = 0$ and $S_{TT} = N-2$ are ensured by choice of origin/scale of T. These values will not be disturbed by transformations of X, which are about to be described.

After a linear transformation $X \to HX$ has achieved simultaneously $\Gamma = I$ and the required β, a change of origin in X will ensure $\alpha = 0$, without disturbing Γ or β.

The transformation $X = HX$ can be done in the following stages.

(a) Linearly independent combinations $Y_2, Y_3, \ldots, Y_q$ are chosen with zero regression on T; i.e. $Y_j = m_j^T X$, where $m_j^T \beta = 0$, $(j = 2,3, \ldots, q)$.

(b) $Y_1 = m_1^T X$ is chosen to be uncorrelated (conditional on T) with $Y_2, \ldots, Y_q$. Thus $m_1^T \Gamma m_j = 0 \qquad (j = 2,3, \ldots, q)$.

(c) Write $Y = (Y_2, Y_3, \ldots, Y_q)^T$ and $COV(Y|T) = GG^T$ where G is $(q-1)\times(q-1)$ and non-singular.

Now $COV(G^{-1}Y|T) = G^{-1}GG^TG^{-1T} = I$.

The components of $Z = G^{-1}Y$ are uncorrelated (conditional on T) with Y_1, by (b). Change of scale of Y_1 is all that is needed to achieve $\Gamma = I$. Z has zero regression on T by (a).

(d) Consider

$$HX = \begin{bmatrix} \text{scaled version of } Y_1 \\ \cdots \\ Z \end{bmatrix}$$

If its regression vector is $(\beta^*, 0, 0, \ldots, 0)^T$ and its conditional covariance matrix is I then the identity

$$\rho^2 = \sigma^2 \beta^T (\Gamma + \sigma^2 \beta\beta^T)^{-1} \beta$$

shows that

$$\rho^2 = \sigma^2(\beta^*, 0, 0, \ldots, 0)\left\{I + \begin{bmatrix} \sigma^2\beta^{*2} & 0 & \ldots & 0 \\ 0 & 0 & & 0 \\ 0 & 0 & & 0 \end{bmatrix}\right\}^{-1} \begin{bmatrix} \beta^* \\ 0 \\ 0 \end{bmatrix}$$

$$= \frac{\sigma^2 \beta^{*2}}{1 + \sigma^2\beta^{*2}} \qquad \ldots(5.15)$$

thus

$$\beta^{*2} = \frac{\rho^2}{\sigma^2(1 - \rho^2)} \qquad \ldots(5.16)$$

Since system A´ has $S_{TT} = N-2$, we replace σ^2 by $(N-2)\sigma^2/S_{TT}$. Thus

$$\beta^{*2} = \frac{\rho^2}{[(N-2)\sigma^2/S_{TT}](1 - \rho^2)}$$

$$= \frac{\rho^2}{C_N(1 - \rho^2)} \qquad \ldots(5.17)$$

THEOREM 5.2. <u>σ^{-1}BIAS DEPENDS ONLY ON q, N, ρ, B_N and C_N</u>

PROOF:

From equation (5.12)

$$\text{Bias}/\sigma = -[[E\hat{C} - C] - [E\hat{D} - D]^T \{\alpha + \beta\mu\}]/\sigma \qquad \ldots(5.18)$$

Following proof that bias/σ depends upon q, N, ρ, B_N, C_N follows closely similar proof for MSE/σ^2. Only step 2 is slightly different.

STEP 1:

Same as above in theorem 5.1.

STEP 2:

Bias/σ is not changed by

(i) changes of origin/scale in T;

(ii) changes of origin in X;

(iii) $X \to HX$, where H is non-singular.

PROOF:

$$D = \sigma^2 \{\Gamma + \sigma^2\beta\beta^T\}^{-1} \beta$$

$$C = \mu - D^T(\alpha + \mu\beta)$$

(i) Change of origin/scale in T.

Consider T´, where $T = f + g\,T'$

$$\alpha + \beta T = \alpha' + \beta' T'$$

$$\alpha + \beta f + \beta g T' = \alpha' + \beta' T'$$

thus $\beta' = \beta g$ and $\alpha' = \alpha + \beta f$

also $\mu' = E(T') = E[(T-f)/g] = (\mu-f)/g$

$\sigma'^2 = VAR(T') = VAR[(T-f)/g] = \sigma^2/g^2$

Γ unchanged, i.e. $\Gamma = \Gamma'$.

Results above show $\sigma\beta$ unchanged, so $\{\Gamma + \sigma^2\beta\beta^T\}$ unchanged and

$$D' = \sigma'^2\{\Gamma' + \sigma'^2\beta'\beta'^T\}^{-1}\beta'$$

$$= \sigma' D/\sigma$$

$$= D/g$$

Similarly $E(\hat{D}') = E(\hat{D})/g$ so $[E(\hat{D}') - D'] = [E(\hat{D}) - D]/g$

Now $\alpha' + \beta'\mu' = \alpha + \beta f + g\beta[(\mu-f)/g]$

$$= \alpha + \mu\beta$$

so second term in bias/σ, namely $-[E(\hat{D})-D]^T\{\alpha + \beta\mu\}/\sigma$ is invariant.

Now

$$C' = \mu' - D'^T(\alpha' + \mu'\beta')$$

$$= [(\mu-f)/g] - [D^T/g][\alpha + \mu\beta) = (C - f)/g$$

similarly $\hat{C}' = (\hat{C} - f)/g$ and $E(\hat{C}') = [E(\hat{C}) - f]/g$

so first term in bias/σ, namely $-[E\hat{C} - C]/\sigma$ is also invariant.

Thus bias/σ invariant for changes of origin or scale in T, as required for (i).

(ii)

$$X' = m + X$$

μ, σ unchanged,

$$\alpha' + \beta' T = E(X'|T) = m + \alpha + \beta T$$

so

$$\alpha' = m + \alpha \quad \text{and} \quad \beta' = \beta$$

$$\Gamma' = \Gamma; \quad D' = D; \quad \hat{D}' = \hat{D};$$

$$C' = \mu - D^T(\alpha' + \mu\beta') = C - D^T m$$

similarly

$$\hat{C}' = \hat{C} - \hat{D}^T m \quad \text{and} \quad E\hat{C}' = E\hat{C} - E(\hat{D}^T)\, m$$

thus bias becomes

$$-[E\hat{C} - E\hat{D}^T m - (C - D^T m)] - [E\hat{D} - D]^T\{m + \alpha + \beta\mu\}$$

$$= -[E\hat{C} - C] - [E\hat{D} - D]^T\{\alpha + \beta\mu\}, \text{ i.e. not changed.}$$

so bias/σ is also invariant.

Note that $E\hat{C} - C$ is not invariant.

(iii)

$$X' = HX \text{ where H is q×q non-singular}$$

μ, σ unchanged.

$$\alpha' + \beta' T = E[X' | T] = H(\alpha + \beta T)$$

so $\alpha' = H\alpha$ and $\beta' = H\beta$

$$\Gamma' = COV(X' | T) = COV(HX | T) = H\,\Gamma\,H^T$$

thus

$$\{\Gamma' + \sigma^2\beta'\beta'^T\}^{-1} = H^{T-1}\{\Gamma + \sigma^2\beta\beta^T\}^{-1}\,H^{-1}$$

and

$$D' = H^{T-1}\,D$$

Similarly

$$\hat{D}' = H^{T-1}\,\hat{D}, \qquad E\hat{D}' = H^{T-1}\,E\hat{D},$$

and

$$[E\hat{D}' - D'] = H^{T-1}\,[E\hat{D} - D]$$

so second term in bias becomes

$$[E\hat{D} - D]^T\,H^{-1}\,H\{\alpha + \beta\mu\}, \text{ i.e. unchanged.}$$

$C = \mu - D^T(\alpha + \mu\beta)$ is also unchanged so $\hat{C}$, $E\hat{C}$, $(E\hat{C} - C)$ and first term in bias/σ are seen successively to be unchanged.

Thus bias/σ is invariant for non-singular transformations $X' = HX$ as required for (iii).

STEPS 3, 4 and 5:

Same as in theorem 5.1.

All the above steps complete the proof.

5.5. SIMULATION STUDY (p=1; general q)

We have proved in theorem 5.1 that MSE/σ^2 depends only on the same four invariants (as for the univariate case) and on q for any value of q and is also invariant under changes of origin/scale of T and $\boldsymbol{X}$. So it is enough to simulate the canonical system with $\bar{t} = 0$, $S_{TT} = N-2$, $\alpha = 0$, $\Gamma = I$, $\beta = (\beta^*, 0, 0, \ldots, 0)^T$ where

$$\beta^* = \rho/[C_N(1-\rho^2)]^{\frac{1}{2}} \qquad \ldots(5.19)$$

$$\mu = (B_N)^{\frac{1}{2}} \quad \text{and} \quad \sigma^2 = C_N$$

In the canonical form we have

$$EX_1 = \beta^*\mu$$

$$EX_2 = EX_3 = \ldots = EX_q = 0$$

$$EX_1^2 = \beta^{*2}\sigma^2 + 1 + (\beta^*\mu)^2$$

$$= \beta^{*2}(\sigma^2 + \mu^2) + 1$$

$$EX_2^2 = EX_3^2 = \ldots = EX_q^2 = 1$$

$$EX_1X_2 = EX_1X_3 = \ldots = EX_{q-1}X_q = 0.$$

Substituting above terms in the expression (5.13) we have

$$MSE = \sigma^2(1-\rho^2) + E(\hat{C}-C)^2 + 2\beta^*\mu E(\hat{C}-C)(\hat{D}_1-D_1) + \{\beta^{*2}(\sigma^2+\mu^2)+1\}E(\hat{D}_1-D_1)^2$$

$$+ (q-1)E(\hat{D}_2 - D_2)^2 \qquad \ldots(5.20)$$

Last term by symmetry of $X_2, X_3, \ldots, X_q$.

$$D = \left[\frac{\sigma^2\beta^*}{1+\sigma^2\beta^{*2}}, 0, \ldots, 0\right]^T \qquad \ldots(5.21)$$

(i.e. $D_2, D_3, \ldots, D_q = 0$)

$$C = \mu - D^T(\alpha + \mu\beta)$$

$$= \mu - \frac{\mu\sigma^2\beta^{*2}}{1+\sigma^2\beta^{*2}}$$

$$= \frac{\mu}{1+\sigma^2\beta^{*2}} \qquad \ldots(5.22)$$

and the best linear predictor is

$$C + D^TX = C + D_1X_1$$

$$= \frac{\mu + \sigma^2\beta^*X_1}{1+\sigma^2\beta^{*2}} \qquad \ldots(5.23)$$

To evaluate (5.20) we require to simulate $\hat{C}$, $\hat{D}$ and thus estimate $VAR(\hat{C})$, $VAR(\hat{D}_1)$, $VAR(\hat{D}_2)$, $COV(\hat{C},\hat{D}_1)$, $E\hat{C} - C$, $E\hat{D}_1 - D_1$, $E\hat{D}_2$ and

$$E(\hat{C} - C)^2 = VAR(\hat{C}) + (bias\hat{C})^2 \text{ etc.}$$

These can be simulated by simulating $\hat{\alpha}$, $\hat{\beta}$, $\hat{\Gamma}$ from their distributions.

Anderson (1984), theorem 8.2.2. states that if

$$\underset{q\times 1}{X_a} \text{ is } N\left(\underset{q\times p}{\beta}\ \underset{p\times 1}{t_a},\ \underset{q\times q}{\Gamma}\right) \qquad a = 1,2,\ldots,N$$

$\acute{p} = p + 1$ where p is the number of explanatory variables.

$$t_a = \begin{bmatrix} 1 \\ t_{1a} \\ \vdots \\ t_{pa} \end{bmatrix}$$

then

$$\hat{\beta} \text{ is } \underset{q\times \acute{p}}{N}(\beta, \ldots)$$

$COV(i^{th} \text{ and } j^{th} \text{ rows of } \hat{\beta})$ is $\gamma_{ij}\ A^{-1}$ ...(5.24)

where

$$\underset{p\times \acute{p}}{A} = \sum_a t_a t_a^T \quad \text{and} \quad \Gamma = \{\gamma_{ij}\}.$$

$N\hat{\Gamma}_{MLE}$ is $W(\Gamma, N-\acute{p})$, independent of $\hat{\beta}$.

In linear calibration when $p = 1$

$$\underset{q\times 1}{X_a} \text{ is } N\left(\begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \\ \vdots & \vdots \\ \alpha_q & \beta_q \end{bmatrix}\begin{bmatrix} 1 \\ t_a \end{bmatrix},\ \underset{q\times q}{\Gamma}\right) \quad a=1,2,\ldots,N$$

and in the canonical form

$$X_a \text{ is } N\left(\begin{bmatrix} \beta^* t_a \\ 0 \\ \vdots \\ 0 \end{bmatrix},\ I\right) \qquad a = 1, 2, \ldots, N$$

Thus $\Gamma = I$ and $A = \sum_a \begin{bmatrix} 1 & t_a \\ t_a & t^2_a \end{bmatrix} = \begin{bmatrix} N & 0 \\ 0 & S_{TT} \end{bmatrix}$ because $\bar{t} = 0$.

$\hat{\alpha}_i$, $\hat{\beta}_i$ ($i = 1, 2, \ldots, q$) are independent by (5.24) and also by (5.24) covariance matrix of $(\hat{\alpha}_i, \hat{\beta}_i)$ is $A^{-1} = \begin{bmatrix} N^{-1} & 0 \\ 0 & S_{TT}^{-1} \end{bmatrix}$

$\hat{\alpha}_i$ is N (0, 1/N) $\quad i = 1, 2, \ldots, q$

$\hat{\beta}_1$ is N (β^*, 1/(N-2))

$\hat{\beta}_i$ is N (0, 1/(N-2)) $\quad i = 2, \ldots, q$

Also independent

$N\,\hat{\Gamma}_{MLE}$ is W (I, N-2)

and unbiased $\hat{\Gamma}$ is $(N-2)^{-1}\, W$ (I, N-2).

Bartlett's decomposition of a Wishart matrix Anderson, (1984) corollary 7.2.1, was used to simulate $\hat{\Gamma}$.

$\hat{\alpha}_i$, $\hat{\beta}_i$, $\hat{\Gamma}$ were simulated 10000 times, and $\hat{C}$, $\hat{D}$ calculated from the following formulae.

$$\hat{C} = \mu - \hat{D}^T(\hat{\alpha} + \mu\hat{\beta})$$

$$\hat{D} = \sigma^2 \{\hat{\Gamma} + \sigma^2\hat{\beta}\hat{\beta}^T\}^{-1}\hat{\beta}$$

Natural estimates of $VAR(\hat{C})$, $VAR(\hat{D}_1)$, $VAR(\hat{D}_2)$, $COV(\hat{C},\hat{D}_1)$, $E\hat{C} - C$, $E\hat{D}_1 - D_1$ and $E\hat{D}_2$ were obtained from these simulations.

Now MSE/σ^2 can be calculated using (5.20) for any set of values of the invariants N, B_N, C_N, ρ^2 and q.

Q is defined as in relation (3.11)

$$MSE/\sigma^2 = (1 - \rho^2)(1 + Q_s)$$

so

$$Q_s = MSE/[\sigma^2(1 - \rho^2)] - 1$$

MSE/σ^2 is invariant so would be Q_s.

Q_s is calculated for the 81 combinations of the set of invariants made in section 3.3 by the procedure described above for q = 1, 2, 3 ,4, 8 and are given in table 7.2 where $\hat{\rho}^2$ should be read as ρ^2. The simulated values for q = 1 quite agreed with the simulated values of section 3.3. The simulated values of q = 2, 3, 4 and 8 are further discussed in the following chapter.

CHAPTER 6

APPROXIMATIONS AND INTERVAL ESTIMATES

Approximation to $Q_S(N-2)$ when $q > 1$ and $p = 1$ is obtained using simulated values and then this approximation is used to define an interval for T. Confidence probabilities are studied in section 6.5 using this approximation.

The procedure to obtain the approximation is described below.

6.1. APPROXIMATING Q(N-2) BASED ON REGRESSING SIMULATED VALUES

For any particular value of q, Q_S is a function of N, ρ^2, B_N, C_N i.e. $Q_S(N,\rho^2,B_N,C_N)$. An extensive study has already been made in chapter 3 for simple linear calibration problem (p=q=1) considering simulated values Q_S and approximated values Q_A, obtained with the help of Taylor's series.

Using Taylor's series we got a mathematical expression Q_A when $q = 1$.

$$Q_A = \rho^2/N + 1/(N-2)[2\rho^2(1-\rho^2) + (1-2\rho^2)^2C_N + \rho^2B_N] \quad ...(6.1)$$

This expression suggests following linear model for $q \geq 1$

$$Q_{A1} = 1/(N-2)[b_0 + b_1\rho^2 + b_2\rho^4 + (b_3 + b_4\rho^2 + b_5\rho^4)\, C_N + (b_6 + b_7\rho^2 + b_8\rho^4)\, B_N] \quad ...(6.2)$$

Simulated values Q_S corresponding to any set of values of the invariants for any $q \geq 1$ can be generated by the procedure described in section 5.5 and Q_{A1} in equation (6.2) can be replaced by those simulated values Q_S. Thus the following quadratic multiple regression model can be fitted.

$$Q_S(N-2) = b_0 + b_1\rho^2 + b_2\rho^4 + (b_3 + b_4\rho^2 + b_5\rho^4)C_N + (b_6 + b_7\rho^2 + b_8\rho^4)B_N + \text{error} \quad ...(6.3)$$

Where the regression coefficients b_i's ($i = 0, 1, 2, \ldots, 8$) may depend on N and q.

Here to increase the scope of study, sample space for the invariants ρ^2, B_N and C_N is increased and now five values of each instead of three values (as in section 3.3) are considered i.e. now ρ^2 = 0.3, 0.5, 0.7, 0.8, 0.9; B_N = 0.0, 1.0, 2.0, 3.0, 4.0; C_N = 0.25, 0.50, 1.0, 2.0, 4.0. The values of N are the same as before i.e. N = 10, 30, 50.

125 values of $Q_s(N-2)$ corresponding to 125 (5×5×5) combinations of ρ^2, B_N, C_N are calculated for each N = 10, 30, 50, making a total of 375 instead of 81 in the previous simulations.

Linear model (6.3) is fitted by ordinary least squares for q = 1, 2, 3, 4 and 8 using 125 values of Q_s for N = 10, 30, 50 separately. Estimates of partial regression coefficients along with other relevant statistics are given in table 6.1. R^2 is coefficient of determination, S is such that $(125-9)S^2$ = residual sum of squares, q is the number of response variables in the multivariate regression experiment.

Table 6.2 summarizes $Q_s(N-2)$, $Q_A(N-2)$ and $Q_R(N-2)$ where Q_R represents fitted values.

First we discuss the case of simple linear calibration.

6.2. RESULTS WHEN q = 1

Table 6.1(a) indicates that the values of R^2 are greater than 98% and S decreases with the increase in N; S = 0.2320 for N = 10; S = 0.0935 for N = 30; S = 0.0580 for N = 50. This along with graphs $Q_s(N-2)$ versus $Q_R(N-2)$ figure 6.1 shows that model fits to the situation very well. The $Q_A(N-2)$ values calculated from relation (6.1) are also plotted against $Q_s(N-2)$ in figure 6.1 to see how the approximation works. 95% interval estimates constructed for regression parameters b_i's (i = 0, 1, ..., 8) of expression (6.3) overlap for N = 30, 50 and also most of the times for N = 10. The approximation (6.3) suggested that the b_i's would depend only

slightly on N.

To compare mathematical expression (6.1) and multiple linear regression model (6.3), we define two quantities

$$S_A = E[Q_s(N-2) - Q_A(N-2)]^2$$

$$= \frac{1}{125} \sum_{1}^{125} [Q_s(N-2) - Q_A(N-2)]^2$$

and

$$S_R = E[Q_s(N-2) - Q_R(N-2)]^2$$

$$= \frac{1}{125-9} \sum_{1}^{125} [Q_s(N-2) - Q_R(N-2)]^2$$

$$= \hat{\sigma}^2$$

Q_A comes from mathematical approximation obtained by Taylor's series and Q_R is from regression model where coefficients of ρ^2, ρ^4, etc. in Q_R are functions of Q_1, Q_2, ..., Q_{125}.

S_A and S_R for each N are given as under

N	10	30	50
S_R	0.054	0.009	0.0034
S_A	0.445	0.045	0.0150
S_A/S_R	8.2	5.0	4.0

These results along with the graphs of $Q_s(N-2)$ versus $Q_A(N-2)$ figure 6.1 indicate that the approximations are getting better for large N.

It looks reasonable to pool the results of three regressions i.e. N = 10, 30, 50 because there is a reason to think that three functions are the same (mathematical approximation).

COMBINATION OF ESTIMATES

Let the linear model (6.3) be represented by

$$E(X) = A\theta$$

where $A_{125\times 9}$ is a matrix of ρ^2, B_N, C_N and $\theta_{9\times 1}$ is a parameter vector also

$$COV(X) = \sigma^2 I$$

then

$$\hat{\theta} = (A^TA)^{-1}A^TX$$

and $$COV(\hat{\theta}) = \sigma^2(A^TA)^{-1}$$

If we assume X_M (M = 1, 2, ..., 125) is $N(A\theta, \sigma^2 I)$ then

$$\text{Log lik} = \text{Const} - 125\log\sigma - 1/(2\sigma^2)[X - A\theta]^T[X - A\theta]$$

$$= \text{Const} - 125\log\sigma - 1/(2\sigma^2)[(\theta-\hat{\theta})^TA^TA(\theta-\hat{\theta})+RSS]$$

We have three independent sets of data for N = 10, 30, 50 with 125 observations in each. Let $\hat{\theta}_{10}$, $\hat{\theta}_{30}$, $\hat{\theta}_{50}$, be the estimates for N =10, 30, 50 with error variances σ^2_{10}, σ^2_{30}, σ^2_{50} respectively.

Combined log lik = const$-125\,\Sigma\log\sigma_i - 1/2\,\Sigma\,(1/\sigma^2_N)[\theta-\hat{\theta}_N]^TA^TA[\theta-\hat{\theta}_N]$

(A^TA is same each time because values of A are determined by ρ^2, B_N, C_N and is not diagonal; RSS_N absorbed in const.)

$$= \text{const} - 125\Sigma\log\sigma_i - 1/2(\theta-\hat{\theta})^T(1/\sigma^2_{10}+1/\sigma^2_{30}+1/\sigma^2_{50})(A^TA)(\theta-\hat{\theta})$$

for appropriate choice of $\hat{\theta}$ and const.

By comparing linear terms in θ

$$-2\hat{\theta}^TK\theta = -2[1/\sigma^2_{10}\hat{\theta}^T_{10}(A^TA)\theta+1/\sigma^2_{30}\hat{\theta}^T_{30}(A^TA)\theta+1/\sigma^2_{50}\hat{\theta}^T_{50}(A^TA)\theta]$$

where

$$K = [1/\sigma^2_{10}+ 1/\sigma^2_{30}+1/\sigma^2_{50}](A^TA)$$

coefficient vector $K\hat{\theta} = \Sigma\, 1/\sigma^2_N(A^TA)\hat{\theta}_N$

$$\hat{\theta} = \Sigma w_N\hat{\theta}_N$$

$$COV(\hat{\theta}) = \Sigma\, w_N^2 COV(\hat{\theta}_N)$$

where

$$w_N = \frac{\dfrac{1}{\sigma^2_N}}{\dfrac{1}{\sigma^2_{10}} + \dfrac{1}{\sigma^2_{30}} + \dfrac{1}{\sigma^2_{50}}}$$

and $\Sigma\, w_i = 1$

Using the above theory combined estimates for N = 10, 30, 50 are as follows

w.coefficient	w.S.D.	Interval Estimate		Coeffi. (b_i) in Q_A
-0.3001	0.0813	-0.4595	-0.1407	0
4.3100	0.2991	3.7238	4.8962	2.80*
-3.2438	0.2483	-3.7304	-2.7572	-2.00
1.5092	0.0288	1.4528	1.5656	1.00
-5.5310	0.1058	-5.7384	-5.3235	-4.00
5.1647	0.0878	4.9925	5.3370	4.00
-0.0150	0.0278	-0.0694	0.0394	0.00
0.8633	0.1021	0.6631	1.0634	1.00
0.2049	0.0847	0.0389	-0.3710	0.00

* *This value is for N = 10 and the values for any other N can be calculated by the relation 2+(N-2)/N.*

In the above results only some of the b_i's corresponding to $Q_A(N-2)$ i.e. coefficients derived from equation (6.1) lie in the pooled interval from regression. Presumably Q_R is a better approximation than Q_A.

6.3. RESULTS WHEN q = 2, 3, 4, 8

Linear model (6.3) is fitted to 125 values each of $Q_s(N-2)$ for q = 2, 3, 4, 8 and N = 10, 30, 50. Results for q = 2, 3, 4 are given in table 6.1(a) and results for q = 8 are given in table 6.1(b). It is clear from the table that R^2 is always very high i.e. for q = 2 it is 99.2% for N = 10, 99.8% for N = 30 and 99.9% for N = 50. S is very small i.e. for q = 2 it is 0.1641 for N = 10 ; 0.0738 for N = 30 and 0.0564 for N = 50. Similarly for q = 3, 4. For q = 8 R^2 is still high but S has increased.

Parameters for q = 8 and N = 10 are not reported here because of high estimation error as in multivariate regression experiments $N \geq p+q+1$ is required to avoid singularity of the error covariance matrix $\hat{\Gamma}$, Sundberg and Brown (1985).

The above statistics along with the graphs of $Q_s(N-2)$ versus $Q_R(N-2)$, given in figure 6.2 indicate that linear model (6.3) fits

reasonably well.

The nature of the fitted model is shown by contours of $Q_R(N-2)$ discussed in next section 6.4.

6.4. CONTOUR PLOTS

Contour plots are drawn for q = 1, 2, 3, 4 and 8 to see the behaviour of $Q_R(N-2)$. For each q they are of the following three types at N = 10, 30, 50.

(a) $Q_R(N-2)$ against C_N and B_N, fixing ρ^2.

(b) $Q_R(N-2)$ against ρ^2 and C_N, fixing B_N.

(c) $Q_R(N-2)$ against ρ^2 and B_N, fixing C_N.

All these contours are given from figure 6.3 to figure 6.62. There are four contours in each figure for the same value of N and either B_N or C_N or ρ^2 but for different q = 1, 2, 3, 4. These are fifteen in number of each type (a), (b), (c).

Contour key indicates different heights and is same for q = 1 and q = 2; and q = 3 and q = 4 is the same.

ρ^2 is denoted by A in the contours.

In figures 6.3 to 6.47, contour key starts at height 1.00 and goes up in steps of 0.75 for q = 1, 2 whereas for q = 3, 4; it starts at height 2.50 and goes up in steps of 0.75.

In figures 6.48 to 6.62 where q = 8, contour height starts at 7.00 and goes up in steps of 0.75.

The above types are now discussed.

(a)

A look on these contours indicates that Q(N-2) is linear in C_N and B_N. At low values of ρ^2 there is more distance between the contours for q = 1 as compared with q = 2, 3 and 4 which means that more B_N and C_N are required to go towards higher values of Q(N-2).

For $\rho^2 = 0.5$; $q = 1$; $N = 50$, contours are almost vertical indicating that all the increase is because of B_N. This connects with the fact that the mathematical approximation $Q_A(N-2)$ for $q = 1$ has zero coefficients of C_N, when $\rho^2 = 0.5$. Contours with $\rho^2 > 0.7$ are becoming more symmetrical.

One thing is quite clear that with the increase in q, height of the contours have increased.

(b)

To describe this type some cases are picked and discussed. Let us take the case with $N = 10$; $B_N = 4$; $q = 3$. For ρ^2 fixed and greater than 0.7, there is steady increase in $Q(N-2)$ with C_N. ρ^2 fixed and less than 0.6 there is a quadratic increase with C_N.

For $N = 10$; $B_N = 4$; $q = 1$. ρ^2 fixed, $Q(N-2)$ increases with C_N. C_N fixed and greater than 1.5, $Q(N-2)$ decreases with ρ^2, reaches a minimum, then increases as ρ^2 increases.

For $N = 10$; $B_N = 4$; $q = 4$. $Q(N-2)$ is not an increasing function of C_N. The same pattern is found when $N = 10$; $B_N = 1$; $q = 4$. For small N this behaviour may be because of estimation error or because of terms omitted from the model.

For $N = 10$; $B_N = 0.0$; $q = 1$, there is a plateau that indicates a slowly increasing function.

It is clear that $Q(N-2)$ has increased with the increase in q.

(c)

Some interesting cases are discussed for this type i.e. $Q(N-2)$ against ρ^2 and B_N, fixing values of C_N, N, q. For $N = 50$; $C_N = 0.25$; $q = 1$, $Q(N-2)$ increases with increase in ρ^2. Increase is slow at low levels of B_N but with increase in B_N and ρ^2, the lines are becoming straight which mean linear relationship.

For N = 10; C_N = 4; q = 1. ρ^2 fixed, Q(N-2) increases slowly with the increase in B_N. B_N fixed Q(N-2) decreases with ρ^2, reaches its minimum, then increases as ρ^2 increases.

Overall the shape of contours changes with change in N and q and Q(N-2) increases with increase in q.

For q = 8, cases only with N = 30, 50 are considered. Contours for N = 10 and q = 8 are not drawn because these are not reliable for the same reason as is given in section 6.3 for not reporting the results of parameter estimates.

Following observations have been made from the contours drawn for q = 8 and N = 30, 50.

(i) Contours Q(N-2) have linear relationship i.e. for type (a) above and are almost same but start from higher heights for N = 30 as compared with N = 50.

(ii) Contours of Q(N-2) versus ρ^2 and C_N look to depend more on C_N than ρ^2 and contours of Q(N-2) versus ρ^2 and B_N look linear for $C_N \leq 1$.

Overall the shape of the contours have changed with N and q.

The following message is obtained from the contours.

(i) Q(N-2) is an increasing function of q when ρ^2, B_N, C_N, N are fixed.

(ii) Q(N-2) is greater for small values of N as compared with large values of N i.e. Q(N-2) is greater for N = 10 than for N = 30 and 50 (ρ^2, B_N, C_N, q fixed).

6.5. INTERVAL ESTIMATES

We have discussed the interval estimates for q = 1 in section 4.2 where Q_A is used to obtain the tail probabilities. It is concluded there that the tail probabilities do not change very much with ρ^2 and the pivotal function F_1 gave better results compared with F_2, F_3 and F_4. So here only F_1 based on $\rho^2 = 0.7$ and 27 (3×3×3) combinations of the invariants B_N, C_N and N with the same values is simulated. Q_R is used instead of Q_A in the calculations of F_1 and the expression (4.12) becomes as under

$$\hat{MSE}_{reg} = (1-\hat{\rho}^2)\sigma^2[1 + \hat{Q}_R] \qquad \ldots(6.4)$$

where $\hat{Q}_R$ comes from the equation (6.2) substituting $\rho^2 = \hat{\rho}^2$ and

$$\hat{\rho}^2 = \hat{\beta}^T\hat{D}.$$

F_1 is defined as

$$F_1 = [T - (\hat{C}+\hat{D}^TX)]/(\hat{MSE}_{reg})^{\frac{1}{2}} \qquad \ldots(6.5)$$

The EF_1 and $E(F_1^2)$ are

$$\begin{aligned} EF_1 &\approx E[T - (\hat{C}+\hat{D}^TX)]/(MSE_{reg})^{\frac{1}{2}} \\ &= -[(E\hat{C}-C) - (E\hat{D}-D)^T\{\alpha + \beta\mu\}]/(MSE_{reg})^{\frac{1}{2}} \\ E(F_1^2) &\approx E[(T - (\hat{C}+\hat{D}^TX)]^2/E(\hat{MSE}_{reg}) \\ &= MSE/E(\hat{MSE}_{reg}) \\ &\approx 1 \end{aligned}$$

It follows from Theorem 5.1 and 5.2 that at least approximately both $E(F_1)$ and $VAR(F_1)$ depend only on the invariants q, N, ρ, B_N and C_N.

To simulate the upper tail probability $P(F_1>1.96)$ and lower tail probability $P(F_1< -1.96)$, it is required to obtain pivotal function F_1. To this end following are required to be simulated.

(i) T, X, which are simulated using canonical form as in section 5.5.

(ii) $\hat{\alpha}$, $\hat{\beta}$, $\hat{\Gamma}$ are generated from the distribution theory

as is described in section 5.5 to get the estimates of $\hat{C}$, $\hat{D}$ and $\hat{\rho}^2$.

Two cases are studied for T;

(a) T is generated from normal distribution with mean μ and variance σ^2.

(b) T is generated from exponential distribution with μ and σ^2 as location and scale parameters.

where

μ and σ^2 are calculated from the values of B_N and C_N.

β^* is calculated by the relation (5.19) and X's are generated from the standardized normal distribution. $\hat{Q}_R$ comes from the linear model (6.3) with regression coefficients b_i's ($i = 0, 1, \ldots, 8$) as are given in table 6.1(a) for $q = 2$, 3, and 4. For the cases $N = 10$, 30, 50; the corresponding b_i's of the table 6.1(a) were used at first. Then the cases $N = 30$, 50 were repeated with b_i's corresponding to $N = 10$, and slightly lower tail probabilities were usually found. So these latter b_i's were finally chosen for $q = 2$, 3 and 4. The coefficients used for $q = 1$ are the weighted coefficients given at the end of section 6.2 under the combination of estimates.

Three extreme cases are picked to see the distribution of F_1. Normal probability plots in figure 6.63 and summary statistics in table 6.3 are given for 1000 values of F_1 corresponding to normal and exponential distributions of T for $q = 1$, 2, 3, 4 and three combinations of the invariants.

It is clear from the table 6.3 that mean is approximately zero for the case (i) and (iii) where $B_N = 4.0$, $C_N = 0.25$ and $B_N = 0.0$, $C_N = 4.0$ respectively for $q = 1$, 2, 3 and 4. In case (ii) with $B_N = 4.0$, $C_N = 4.0$; means have increasing tendency with increasing value of q.

In all the cases standard deviation has increased for the higher values of q.

The tail probabilities for $P(t)$ to be normal and exponential are given in table 6.4 for N = 10, 30, 50 and q = 1, 2, 3 and 4.

For N = 10, the error probabilities are high and have increasing trend with the increasing q. For q = 1, the sum of lower and upper tail probabilities is around 0.085 for $P(t)$ to be normal and 0.080 for $P(t)$ to be exponential. This agrees with the results already obtained for q = 1 in section 4.2 case (b).

It should be made clear that using the procedure of canonical form as at the end of chapter 5, the upper and lower tail probabilities were same for $B_N = 0$ or $\bar{t} = \mu$ as with the procedure of chapter 4 case (b) but when $\bar{t} \neq \mu$ the values of the upper and lower tail probabilities exchanged. This happened because in the procedure of chapter 4 case (b) $\mu \leq \bar{t}$ always whereas it is opposite here i.e. $\bar{t} \leq \mu$. The change of sign of T exchanges the numerical values of the lower and upper tail probabilities and this is discussed as under.

Lower error prob. $= P(F_1 < -1.96)$

Upper error prob. $= P(F_1 > +1.96)$

Let $T' = -T$

$\bar{T}' = -\bar{T}$

$S_{T'T'} = S_{TT}$

$E(X|T) = \alpha + \beta T = \alpha - \beta T' = \alpha' + \beta' T$

$\alpha' = \alpha, \qquad \beta' = -\beta$

$\sigma^2_{x|T'} = \sigma^2_{x|T}$

$\mu' = -\mu \quad ; \quad \sigma'^2 = \sigma^2$

N, ρ^2, B_N, C_N are unchanged so MSE unchanged

$C' + D'X(\hat{T}') = -\hat{T} = -C - DX$

$C' = -C$

$$D' = -D$$

$$\text{Lower error prob.}' = P[F_1' < -1.96]$$

$$= P\left[\frac{T' - (\hat{C}' + \hat{D}'X)}{(\hat{MSE})^{\frac{1}{2}}} < -1.96\right]$$

$$= P\left[-\frac{T' - (\hat{C}' + \hat{D}'X)}{(\hat{MSE})^{\frac{1}{2}}} > 1.96\right]$$

$$= P\left[\frac{T - (\hat{C} + \hat{D}X)}{(\hat{MSE})^{\frac{1}{2}}} > 1.96\right]$$

$$= P[F_1 > 1.96] = \text{upper error prob.}$$

For N = 30 and 50, sum of error probabilities is very close and is always between 0.05 and 0.06.

It is observed from the table that the lower and the upper tail probabilities are near to each other for $B_N = 0.0$ and get apart for higher values of B_N.

6.6. EXAMPLE: WHEAT QUALITY DATA

The wheat quality data analysed by Brown (1982) consists of 21 samples of response variables 4-vector X and the 2-vector T of explanatory variables. X_1, X_2, X_3 and X_4 are the infrared reflectance measurements and T_1, T_2 denote the percentage of water and protein contents.

We are discussing only one explanatory variable at a time and that we select the protein percentage. The set of first 16 observations on X_1, X_2, X_3, X_4 and T is used as the regression experiment and the next set of 5 to test the predicted values. Thus $\mu \approx \bar{t}$ and $\sigma^2 \approx S_{TT}/15$. Hence approximately, $B_N = 0.0$ and $C_N = 1.0$.

To predict T, different subsets of the response variables can be used but we confine ourselves to the following subsets.

(i) X_2 only

(ii) X_1 and X_2

(iii) X_1, X_2 and X_3

(iv) X_1, X_2, X_3 and X_4.

The values of $\hat{C}$ and $\hat{D}$ are calculated for the above subsets, from the values of $\hat{\alpha}$, $\hat{\beta}$, $\hat{\Gamma}$ obtained from the first 16 observations and μ, σ^2 from $B_N = 0.0$ and $C_N = 1.0$. The point and interval estimates for protein percentage values T_{17}, T_{18}, T_{19}, T_{20}, T_{21} are calculated and are reported in table 6.5 along with the data.

It is clear from the table 6.5 that the values to be predicted are always in the 95% interval for all the four subsets of response variables. The interval estimate is getting shorter with the increase in q until $q = 3$ and it is almost the same for $q = 3$ and $q = 4$.

TABLE 6.1(a). Partial regression estimates along with their S.E.'s in brackets for different q and N. Also S and R^2.

N	q = 1	q = 2	q = 3	q = 4
b_0				
10	-0.5422(0.3915)	1.2306(0.2769)	3.8386(0.3229)	8.9605(0.4232)
30	-0.4081(0.1578)	-0.1111(0.1246)	0.3101(0.1894)	0.9699(0.3221)
50	-0.2441(0.0978)	0.1838(0.0952)	-0.0243(0.1244)	0.3425(0.2283)
b_1				
10	5.4520(1.4400)	3.015 (1.0180)	0.9560(1.1870)	-2.3600(1.5560)
30	4.6982(0.5804)	4.9255(0.4581)	5.2021(0.6964)	5.4810(1.1840)
50	4.0942(0.3597)	4.9598(0.3501)	5.6386(0.4574)	6.1038(0.8395)
b_2				
10	-4.5930(1.1950)	-2.0116(0.8452)	-0.4754(0.9856)	1.1730(1.2920)
30	-3.5677(0.4818)	-2.8512(0.3802)	-2.2225(0.5781)	-1.8138(0.9831)
50	-3.0357(0.2986)	-2.8091(0.2906)	-2.4855(0.3797)	-2.0672(0.6968)
b_3				
10	1.6681(0.1386)	0.9969(0.0980)	0.1286(0.1143)	-1.4228(0.1498)
30	1.6608(0.0559)	2.1445(0.0441)	2.6063(0.0670)	2.9120(0.1140)
50	1.4414(0.0346)	2.3670(0.0337)	3.0648(0.0440)	3.5284(0.0808)
b_4				
10	-5.4357(0.5096)	-1.4705(0.3604)	1.6764(0.4202)	5.0521(0.5509)
30	-5.9741(0.2054)	-5.2029(0.1621)	-4.5158(0.2465)	-3.8032(0.4192)
50	-5.3675(0.1273)	-5.9353(0.1239)	-5.9277(0.1619)	-5.3840(0.2971)
b_5				
10	5.1623(0.4230)	1.6352(0.2992)	-0.4441(0.3488)	-2.0651(0.4573)
30	5.4846(0.1705)	4.0896(0.1346)	2.8833(0.2046)	1.8909(0.3480)
50	5.0425(0.1057)	4.6463(0.1028)	3.8767(0.1344)	2.8491(0.2466)
b_6				
10	0.2302(0.1336)	0.6018(0.0945)	1.0525(0.1102)	1.6213(0.1445)
30	-0.0064(0.0539)	0.2276(0.0425)	0.4776(0.0646)	0.7367(0.1099)
50	-0.0337(0.0334)	0.1329(0.0325)	0.3067(0.0425)	0.5147(0.0779)
b_7				
10	-0.1597(0.4915)	-0.5000(0.3476)	-0.6582(0.4053)	-0.7725(0.5313)
30	0.8041(0.1981)	0.3100(0.1563)	-0.1125(0.2377)	-0.4330(0.4043)
50	0.9501(0.1228)	0.5582(0.1195)	-0.2529(0.1561)	-0.0652(0.2865)
b_8				
10	1.2038(0.4080)	1.2048(0.2885)	1.0786(0.3364)	0.8851(0.4410)
30	0.2833(0.1645)	0.5443(0.1298)	0.7398(0.1973)	0.8692(0.3356)
50	0.1124(0.1019)	0.3982(0.0992)	0.5442(0.1296)	0.6540(0.2379)
S				
10	0.2320	0.1641	0.1913	0.2508
30	0.0935	0.0738	0.1122	0.1908
50	0.0580	0.0564	0.0737	0.1353
R^2				
10	98.0%	99.2%	99.3%	99.2%
30	99.6%	99.8%	99.7%	99.4%
50	99.8%	99.9%	99.9%	99.7%

TABLE 6.1(b). Partial regression estimates with S.E's for q = 8

	N		
	10	30	50
b_0	---	6.9410(1.3190)	3.2410(1.0360)
b_1	---	1.5610(4.8490)	5.0630(3.8100)
b_2	---	2.7340(4.0250)	1.4990(3.1620)
b_3	---	2.0545(0.4667)	4.4555(0.3667)
b_4	---	3.4460(1.7160)	-1.2680(1.3480)
b_5	---	-4.3480(1.4250)	-2.2510(1.1190)
b_6	---	2.1522(0.4501)	1.5559(0.3536)
b_7	---	-1.4080(1.6550)	-1.0510(1.3000)
b_8	---	0.5670(1.3740)	0.6640(1.0790)
S	---	0.7814	0.6139
R^2	---	96.3%	98.0%

TABLE 6.2. Summary statistics of 125 values of $Q_S(N-2)$, $Q_R(N-2)$, $Q_A(N-2)$ for different N and q.

	Minimum			Median			Maximum		
N	10	30	50	10	30	50	10	30	50
q = 1									
$Q_S(N-2)$	0.865	0.745	0.795	2.819	2.514	2.449	9.427	7.717	7.749
$Q_R(N-2)$	0.806	0.768	0.782	2.782	2.489	2.461	8.720	7.616	7.411
$Q_A(N-2)$	0.700	0.740	0.748	2.400	2.467	2.480	7.060	7.180	7.204
q = 2									
$Q_S(N-2)$	2.104	1.415	1.361	4.972	4.029	3.971	11.064	8.953	9.040
$Q_R(N-2)$	2.130	1.348	1.303	4.917	3.965	3.920	10.817	8.900	8.990
q = 3									
$Q_S(N-2)$	4.025	2.167	1.958	7.935	5.557	5.348	14.784	10.688	10.468
$Q_R(N-2)$	4.231	2.048	1.852	7.907	5.576	5.361	14.760	10.604	12.882
q = 4									
$Q_S(N-2)$	7.845	3.070	2.677	12.496	7.173	6.626	20.015	13.204	13.158
$Q_R(N-2)$	7.986	2.936	2.530	12.394	7.130	6.703	20.164	12.956	12.882
q = 8									
$Q_S(N-2)$	---	8.325	6.099	---	15.100	12.816	---	26.557	26.310
$Q_R(N-2)$	---	8.330	5.863	---	15.743	12.719	---	25.579	25.587

TABLE 6.3. Summary statistics regarding F_1 based on 1000 values for some combinations of invariants, q and $P(t)$.

$P(t)$	Mean	S.D	Median	Minimum	Maximum
(i) $B_N = 4.0$, $C_N = 0.25$, N = 10					
q=1					
Normal	0.0702	1.1700	0.0389	-5.3230	4.4797
Exponential	0.0057	1.1291	-0.0140	-5.7439	4.0664
q=2					
Normal	-0.0084	1.1850	-0.0567	-5.2699	5.0908
Exponential	0.0531	1.2486	0.0323	-5.1363	5.6905
q=3					
Normal	0.2175	1.2977	0.1382	-4.3242	6.1449
Exponential	0.2347	1.3996	0.1779	-5.8286	10.0021
q=4					
Normal	0.1340	1.7104	0.1061	-9.7849	9.8618
Exponential	0.1998	1.4431	0.1625	-10.4561	6.3756
(ii) $B_N = 4.0$, $C_N = 4.0$, N = 10.					
q=1					
Normal	0.1327	1.1386	0.0314	-3.2836	5.9210
Exponential	0.0846	1.1332	-0.0141	-6.0603	8.3807
q=2					
Normal	0.1915	1.2272	0.0591	-5.8218	6.2986
Exponential	0.2845	1.2310	0.1297	-2.5214	6.5332
q=3					
Normal	0.5012	1.4123	0.3243	-4.0407	6.5045
Exponential	0.5530	1.6689	0.3151	-3.9202	20.1340
q=4					
Normal	0.6224	1.7630	0.4296	-7.0558	10.6921
Exponential	0.6704	1.8015	0.4627	-7.0706	17.9137
(iii) $B_N = 0.0$, C_N 4.0, N = 10.					
q=1					
Normal	0.0488	1.1372	0.0185	-5.9947	4.8935
Exponential	0.0082	1.1399	-0.0933	-5.4736	7.0026
q=2					
Normal	-0.0969	1.3277	-0.0543	-9.4903	5.1451
Exponential	-0.0144	1.2386	-0.0553	-3.6546	5.9887
q=3					
Normal	0.0442	1.4863	0.0291	-8.2817	6.1712
Exponential	0.0775	1.7132	-0.0461	-5.5343	19.2303
q=4					
Normal	-0.0130	1.7648	-0.0632	-7.5979	7.8095
Exponential	0.0278	1.9025	-0.0887	-10.2309	17.3228

TABLE 6.4. Lower(L) and Upper(U) 10^3 × error probabilities of F_1 for q = 1, 2, 3 and 4 for different $P(t)$ and N.

(a) P(t) Normal.

		q = 1			q = 2			q = 3			q = 4		
B_N		0.0	1.0	4.0	0.0	1.0	4.0	0.0	1.0	4.0	0.0	1.0	4.0
C_N													
						N = 10							
0.25	L	40	38	37	50	44	39	60	51	44	80	68	57
	U	42	47	49	46	54	57	62	70	76	76	91	103
1.0	L	39	36	32	53	40	31	63	46	34	84	61	45
	U	44	50	53	49	61	67	65	81	92	80	106	126
4.0	L	40	31	22	58	40	22	80	49	27	103	65	36
	U	46	52	56	55	71	83	78	104	124	100	141	170
						N = 30							
0.25	L	29	27	26	32	30	27	29	26	24	30	26	23
	U	29	29	32	30	32	33	29	32	31	29	32	33
1.0	L	30	27	24	34	29	24	28	25	20	31	24	19
	U	39	31	33	31	34	36	31	34	37	31	36	39
4.0	L	31	27	23	34	28	23	34	25	18	39	27	18
	U	30	34	36	33	38	42	34	42	48	38	47	55
						N = 50							
0.25	L	29	27	27	26	24	22	26	24	22	29	27	23
	U	29	29	29	27	28	29	28	28	29	22	24	24
1.0	L	29	27	25	24	23	21	27	24	20	31	26	21
	U	29	30	31	27	30	32	29	31	32	23	27	29
4.0	L	31	27	26	25	23	20	29	24	18	33	27	19
	U	29	32	34	28	33	36	32	36	39	28	33	38

(b) $P(t)$ Exponential

						N = 10							
0.25	L	34	31	33	45	38	37	54	47	42	74	61	54
	U	44	47	47	49	52	56	60	69	76	76	87	98
1.0	L	35	28	27	45	33	27	51	39	28	76	53	39
	U	45	48	49	51	58	63	64	80	93	82	102	121
4.0	L	31	21	17	41	25	15	57	30	15	88	49	24
	U	47	50	51	56	69	78	76	97	114	97	126	152
						N = 30							
0.25	L	26	24	24	26	24	22	26	23	20	23	19	18
	U	37	36	35	38	38	38	38	37	37	32	33	34
1.0	L	25	23	22	24	23	19	25	22	17	24	17	14
	U	37	36	38	38	40	41	39	41	41	34	38	39
4.0	L	24	21	18	23	18	13	25	17	12	27	16	10
	U	36	39	42	40	43	47	43	51	54	43	52	57
						N = 50							
0.25	L	23	23	22	23	22	21	22	19	17	20	18	16
	U	37	37	38	36	37	37	36	37	36	33	34	31
1.0	L	23	21	21	21	21	19	22	18	15	19	17	14
	U	36	38	39	36	38	38	36	38	40	34	36	37
4.0	L	23	20	18	22	18	15	21	17	12	19	15	11
	U	38	39	41	38	41	42	39	43	46	39	44	47

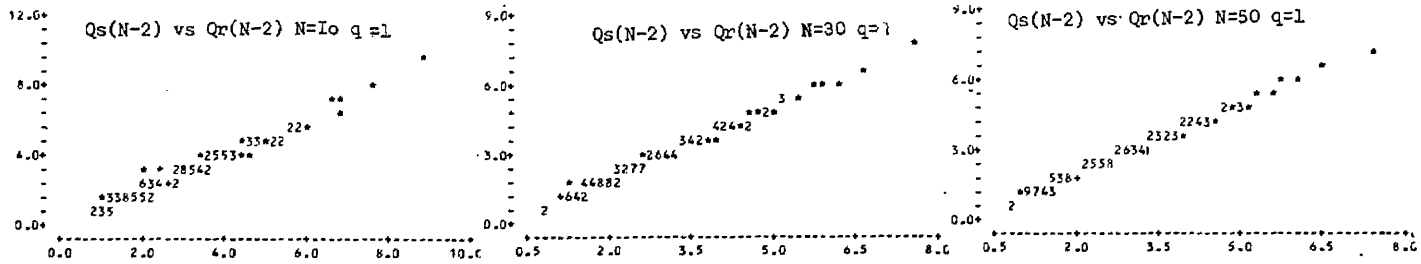
TABLE 6.5(a). Point and interval estimates for wheat quality data for different subsets of response variables.

	point estimate	interval estimate		$\hat{MSE}$
(i) X_2				
	9.4704	8.7515	10.1894	0.1345
	10.0894	9.3705	10.8083	--
	9.2641	8.5452	9.9830	--
	12.9780	12.2591	13.6969	--
	12.7717	12.0527	13.4906	--
(ii) X_1, X_2				
	9.1669	8.7369	9.5969	0.0481
	10.0420	9.6120	10.4720	--
	9.2325	8.8025	9.6625	--
	12.5773	12.1473	13.0073	--
	12.8811	12.4511	13.3112	--
(iii) X_1, X_2, X_3				
	9.1295	8.7509	9.5080	0.0373
	10.1736	9.7951	10.5522	--
	9.1149	8.7364	9.4935	--
	12.6602	12.2816	13.0387	--
	12.7719	12.3934	13.1505	--
(iv) X_1, X_2, X_3, X_4				
	9.2490	8.8729	9.6265	0.0368
	10.1817	9.8055	10.5579	--
	9.1522	8.7760	9.5284	--
	12.7134	12.3372	13.0895	--
	12.7666	12.3904	13.1428	--

(b) Wheat Quality Data.

Observation number	X_1	X_2	X_3	X_4	% protein
1	361	108	96	243	10.73
2	361	107	98	245	11.05
3	362	110	94	241	9.86
4	362	105	94	246	11.41
5	362	104	70	221	11.57
6	367	113	75	221	9.42
7	366	108	82	233	10.93
8	360	104	86	236	11.61
9	362	113	85	229	8.82
10	360	103	90	242	11.81
11	351	97	88	238	12.33
12	353	95	73	227	12.93
13	352	97	77	228	12.69
14	355	96	52	206	13.13
15	357	106	69	216	10.41
16	351	93	69	222	13.57
17	363	113	88	231	9.26
18	363	110	101	248	9.82
19	366	114	79	224	9.46
20	350	96	85	235	12.85
21	355	97	63	216	12.81

FIGURE 6.1


Qs(N-2) vs Qr(N-2) N=Io q =1
Qs(N-2) vs Qr(N-2) N=30 q=1
Qs(N-2) vs Qr(N-2) N=50 q=1

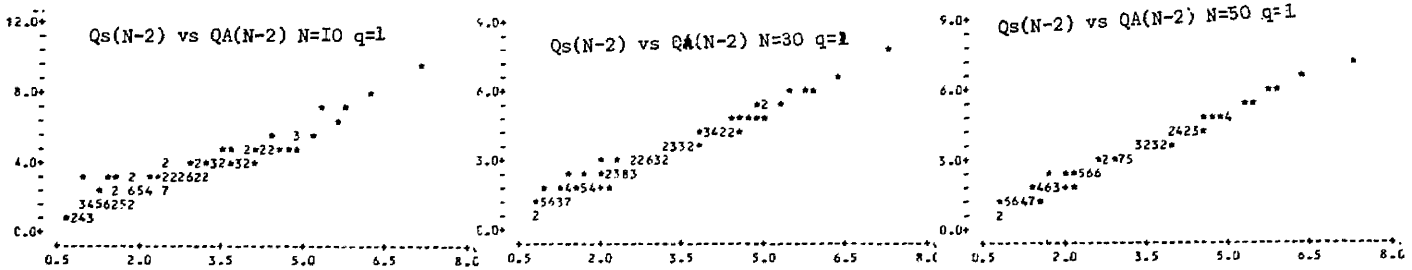
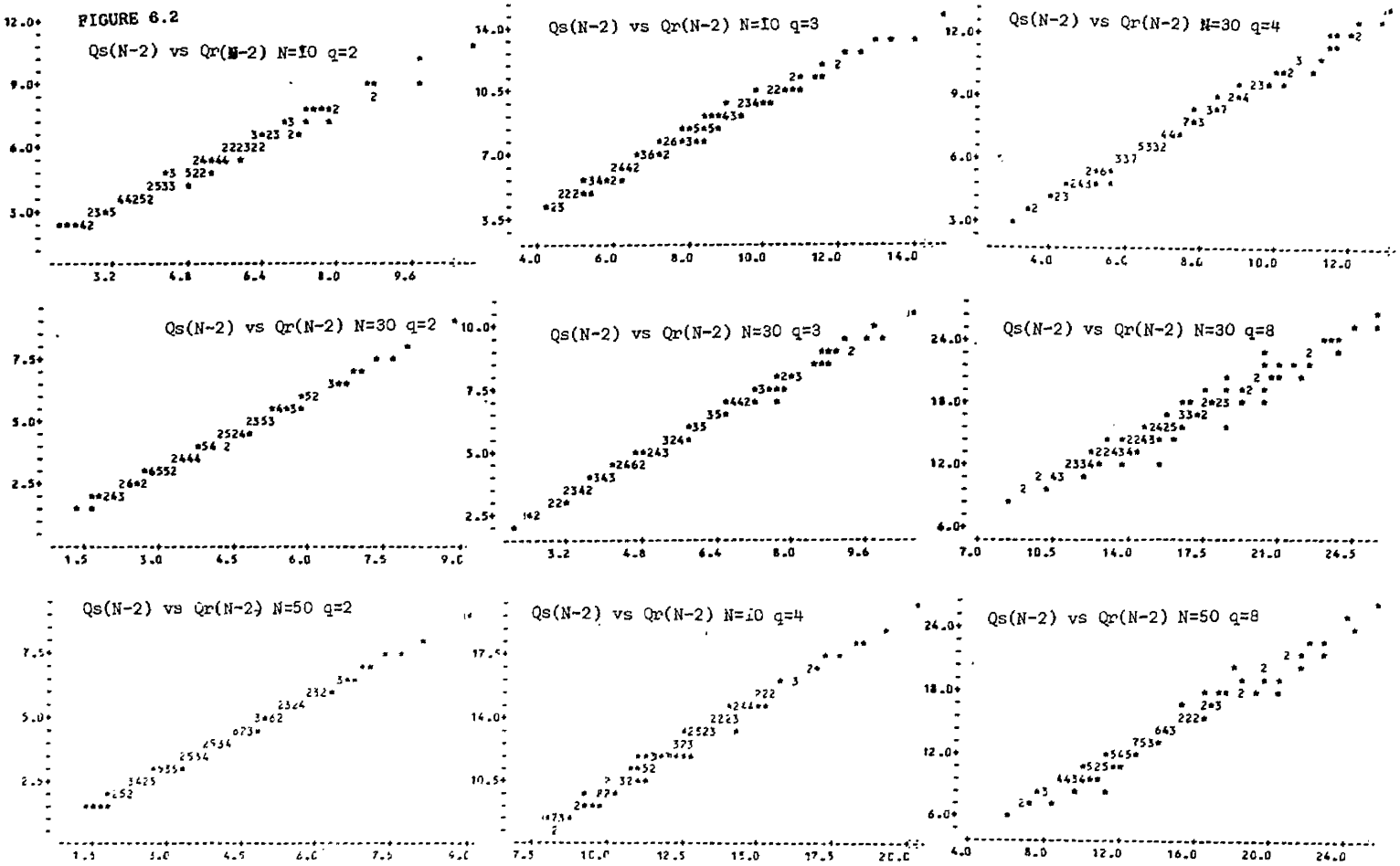
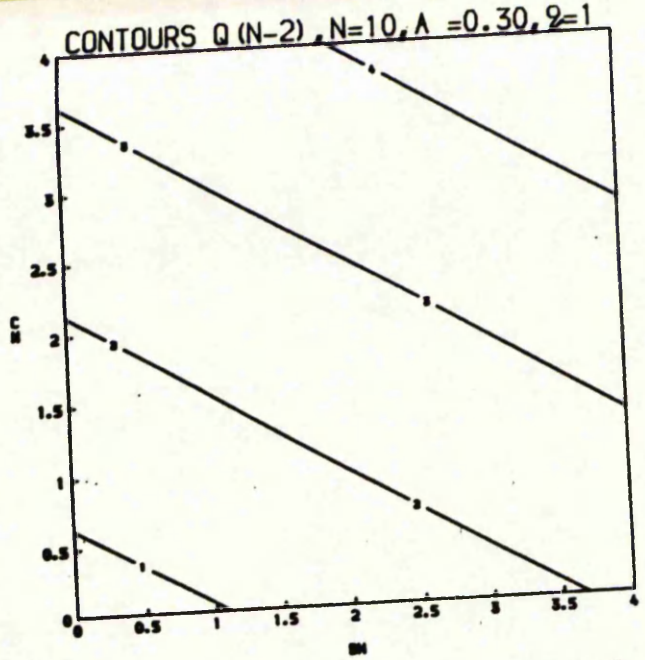
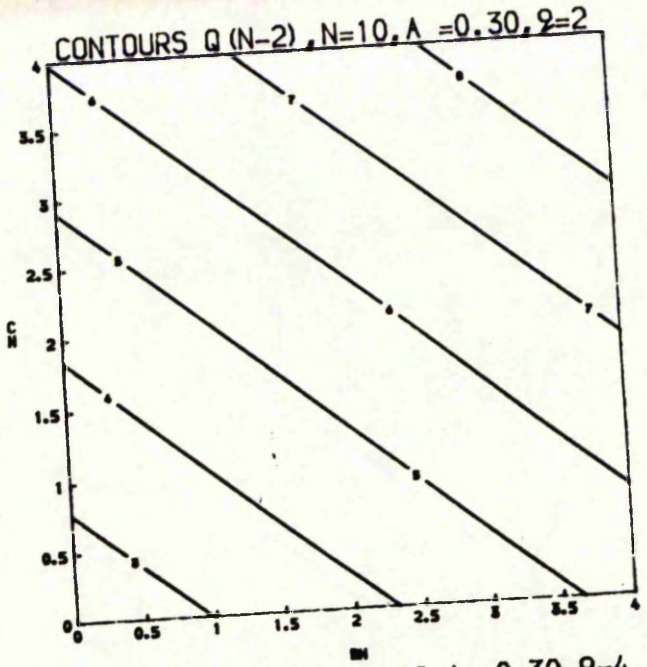

Qs(N-2) vs QA(N-2) N=IO q=1
Qs(N-2) vs QA(N-2) N=30 q=1
Qs(N-2) vs QA(N-2) N=50 q=1

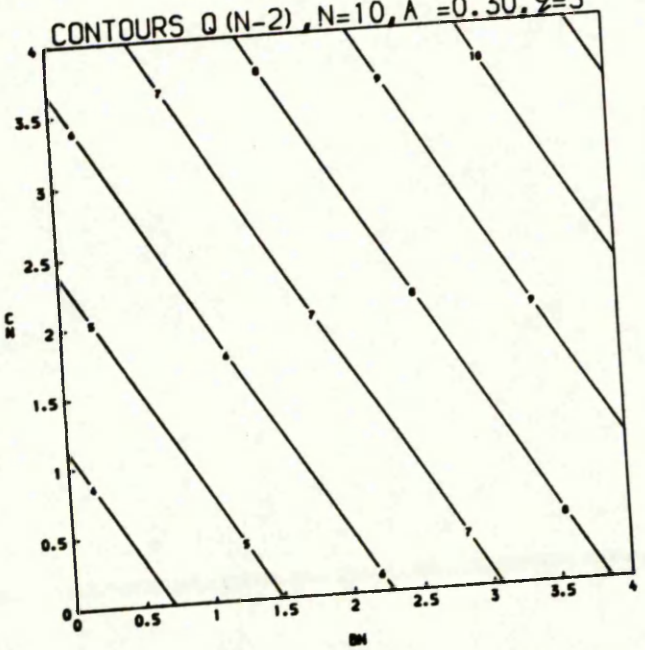
FIGURE 6.2

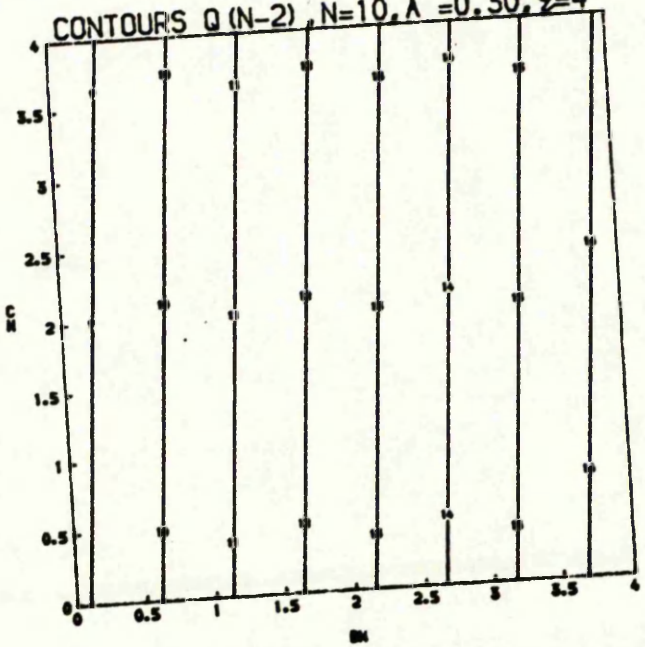

Qs(N-2) vs Qr(N-2) N=10 q=2
Qs(N-2) vs Qr(N-2) N=10 q=3
Qs(N-2) vs Qr(N-2) N=30 q=4
Qs(N-2) vs Qr(N-2) N=30 q=2
Qs(N-2) vs Qr(N-2) N=30 q=3
Qs(N-2) vs Qr(N-2) N=30 q=8
Qs(N-2) vs Qr(N-2) N=50 q=2
Qs(N-2) vs Qr(N-2) N=10 q=4
Qs(N-2) vs Qr(N-2) N=50 q=8





CONTOUR	KEY
1	1.000
2	1.750
3	2.500
4	3.250
5	4.000
6	4.750
7	5.500
8	6.250
9	7.000
10	7.750
11	8.500





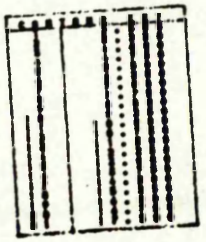


FIGURE 6.3

FIGURE 6.4

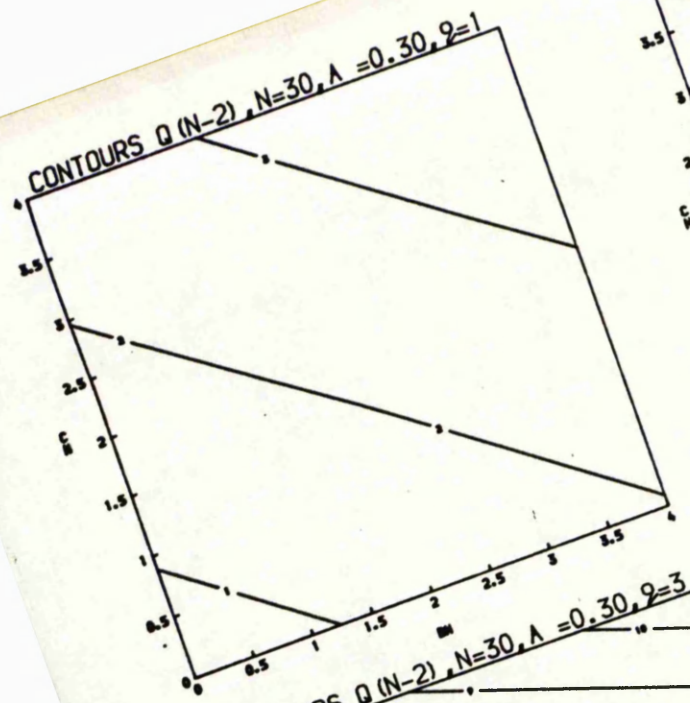
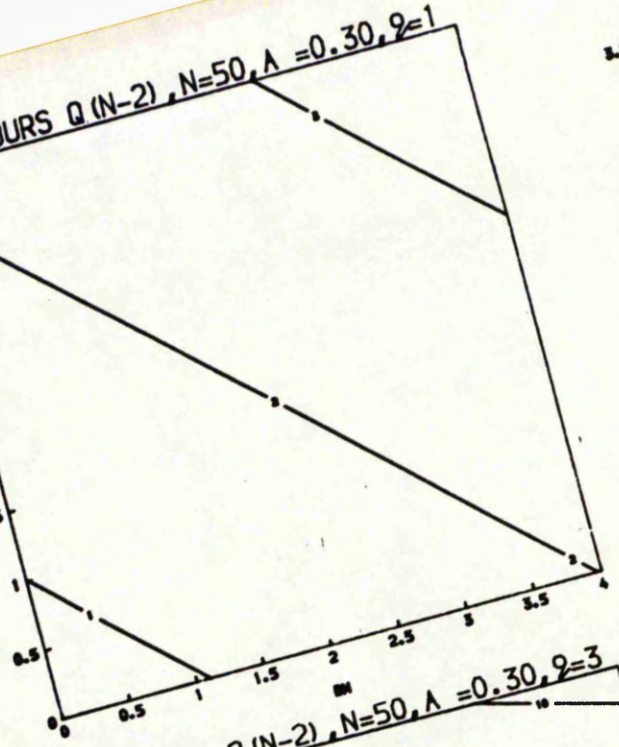
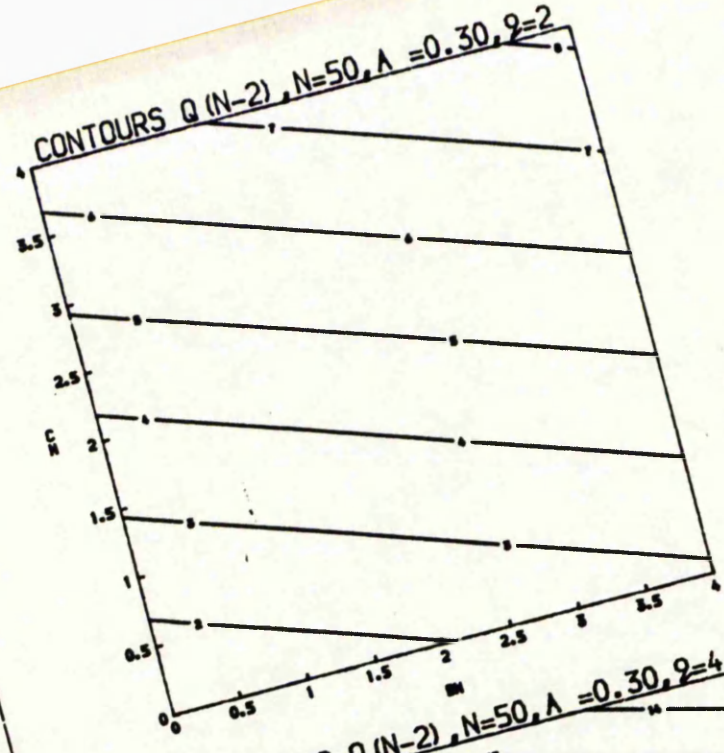
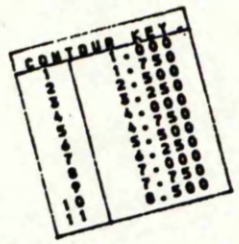
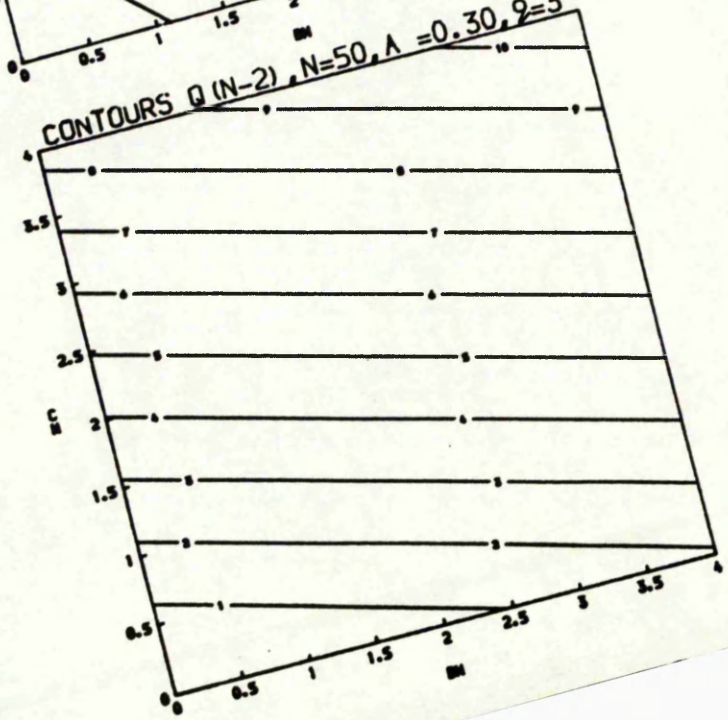

CONTOURS Q (N-2) , N=30, λ =0.30, ϱ=1
CONTOURS Q (N-2) , N=30, λ =0.30, ϱ=2
CONTOURS Q (N-2) , N=30, λ =0.30, ϱ=3
CONTOURS Q (N-2) , N=30, λ =0.30, ϱ=4
CONTOURS
FRACTION
IM

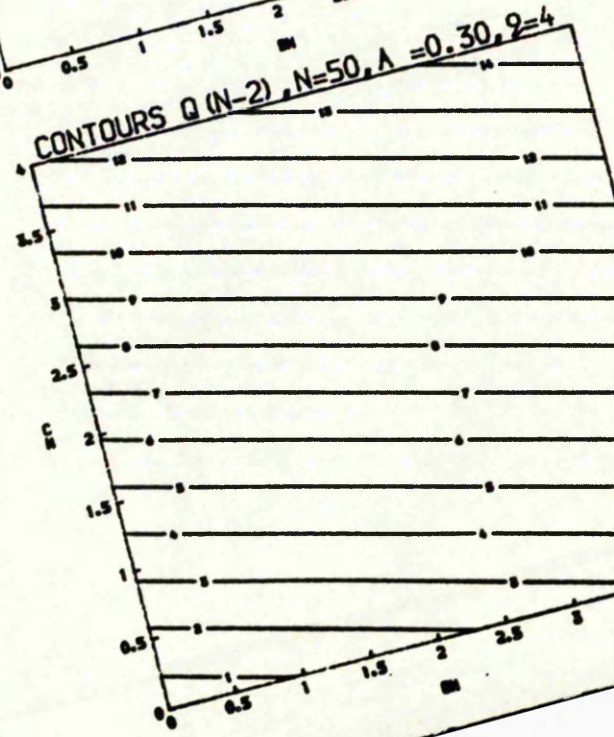
FIGURE 6.5

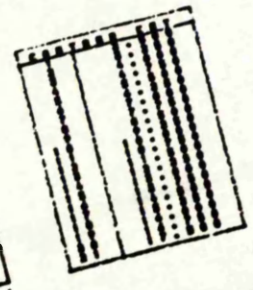

CONTOURS Q (N-2) , N=50, Λ =0.30, 9=1
CM
BM


CONTOURS Q (N-2) , N=50, Λ =0.30, 9=2
CM
BM


CONTOUR KEY


CONTOURS Q (N-2) , N=50, Λ =0.30, 9=3
CM
BM


CONTOURS Q (N-2) , N=50, Λ =0.30, 9=4
CM
BM



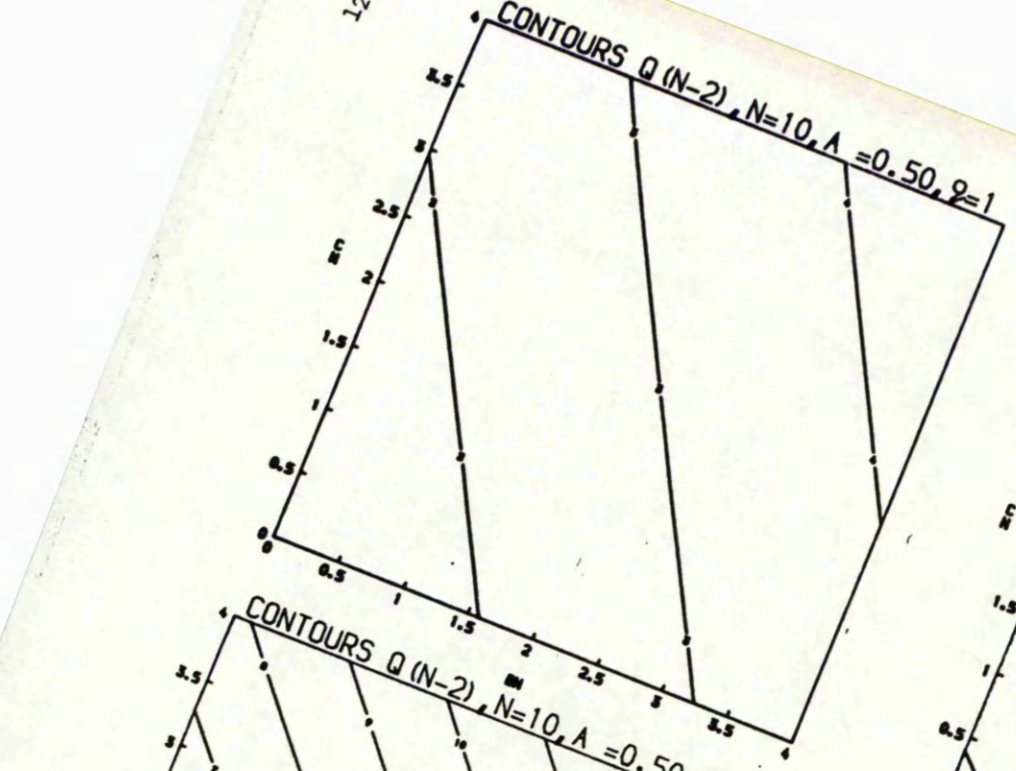
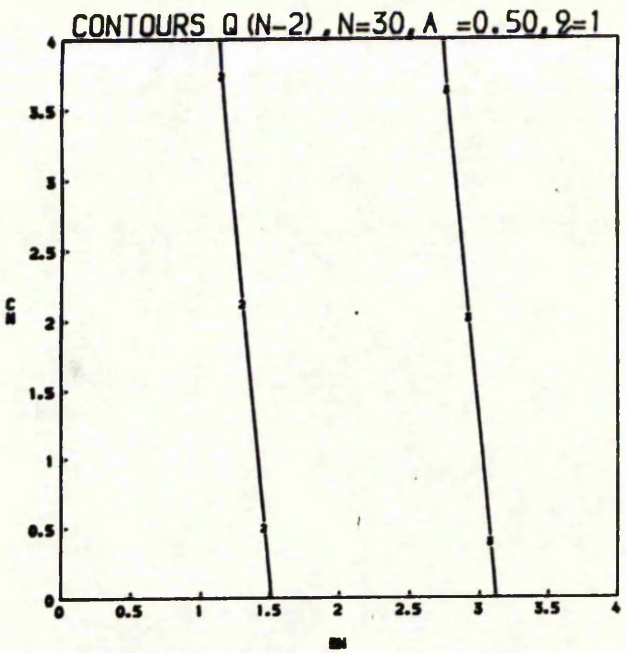

CONTOURS Q (N-2), N=10, A =0.50, Q=1
CONTOURS Q (N-2), N=10, A =0.50, Q=2
CONTOURS Q (N-2), N=10, A =0.50, Q=3
CONTOURS Q (N-2), N=10, A =0.50, Q=4
CONTOUR KEY

FIGURE 6.6


CONTOURS Q(N-2), N=30, A =0.50, Q=1
CM
BM

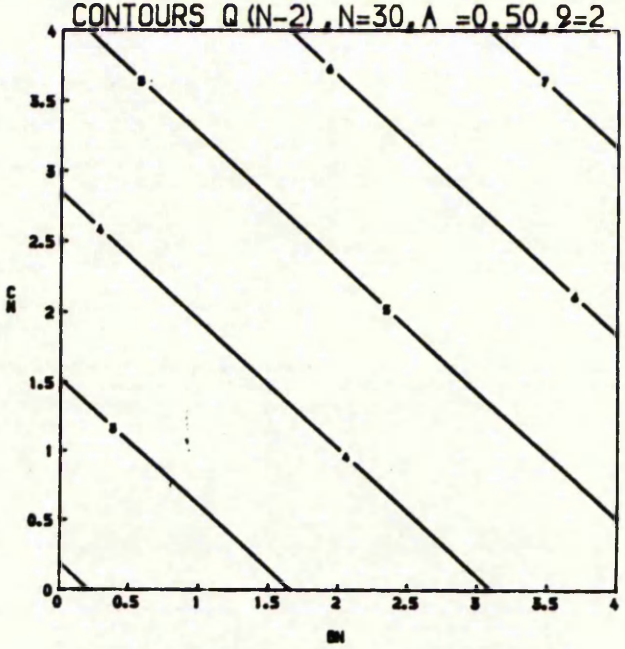
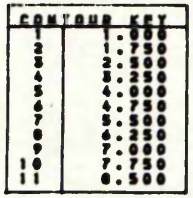
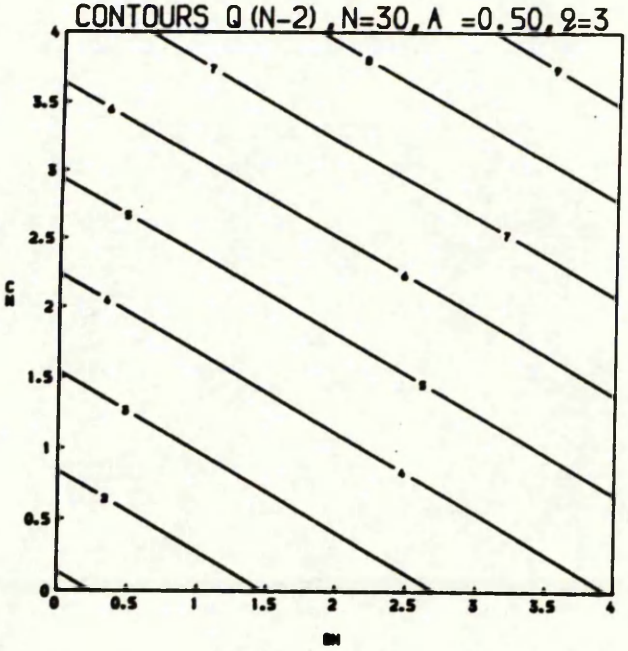
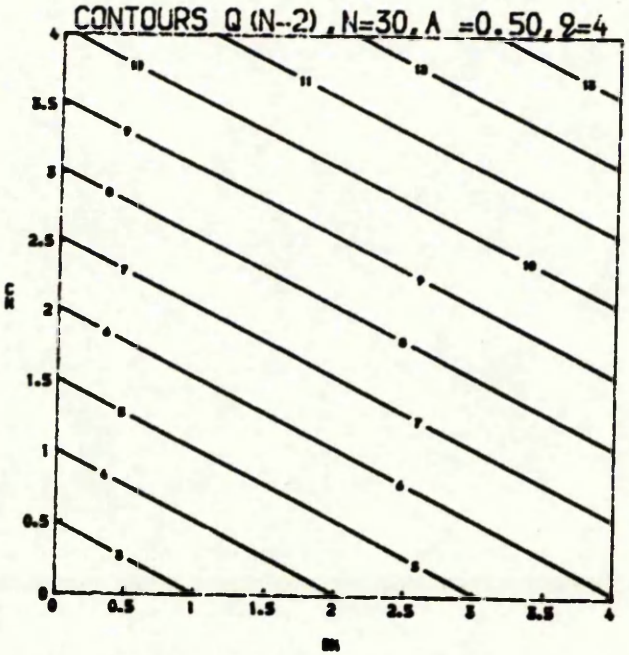
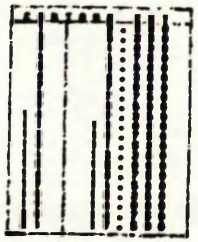

CONTOURS Q(N-2), N=30, A =0.50, Q=2
CM
BM

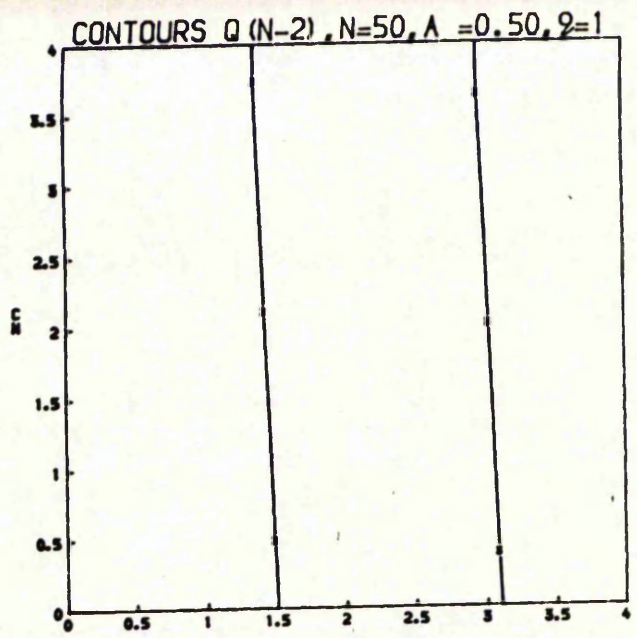
FIGURE 6.7


CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500


CONTOURS Q(N-2), N=30, A =0.50, Q=3
CM
BM


CONTOURS Q(N-2), N=30, A =0.50, Q=4
CM
BM




CONTOURS Q(N-2), N=50, A =0.50, Q=1
EM
EN

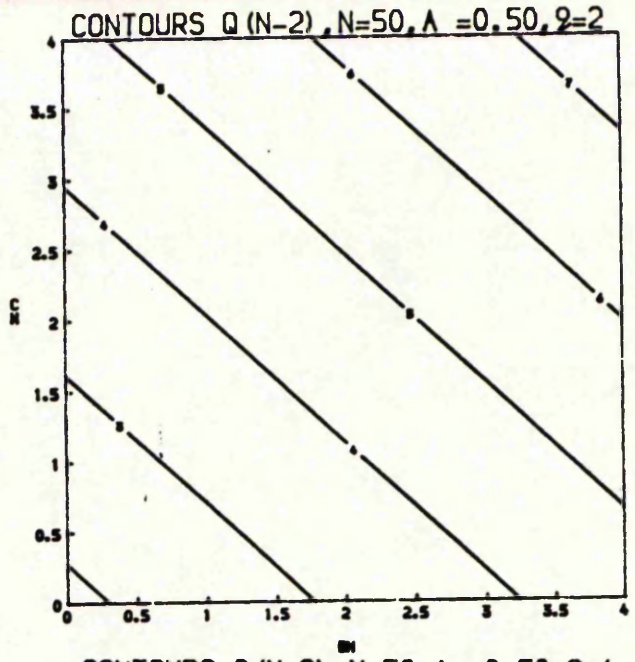
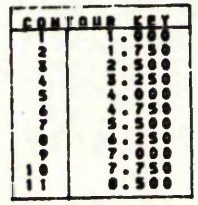
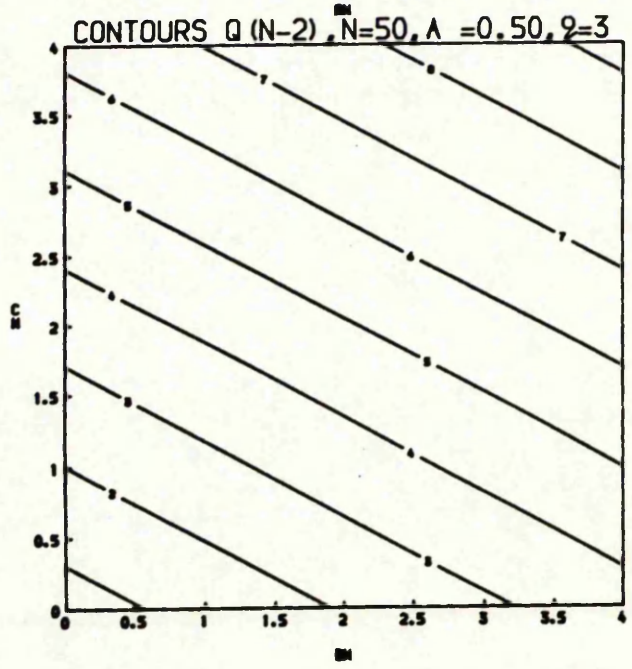
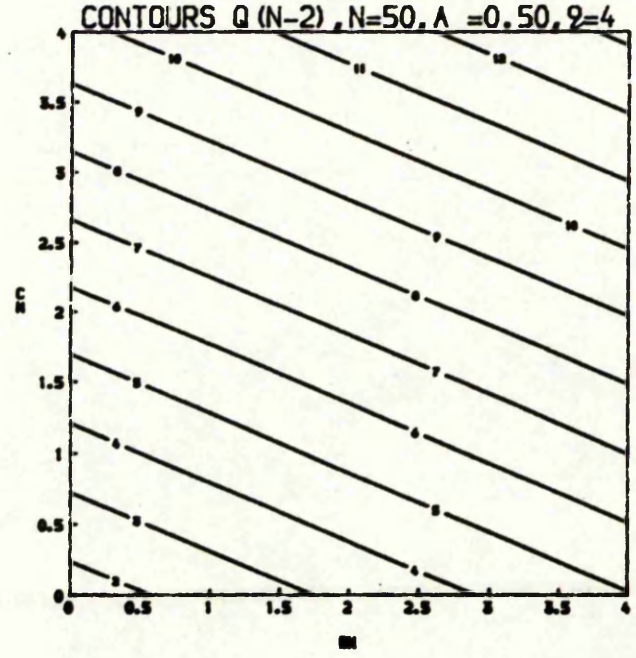
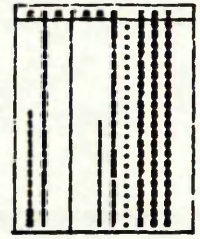

CONTOURS Q(N-2), N=50, A =0.50, Q=2
EM
EN

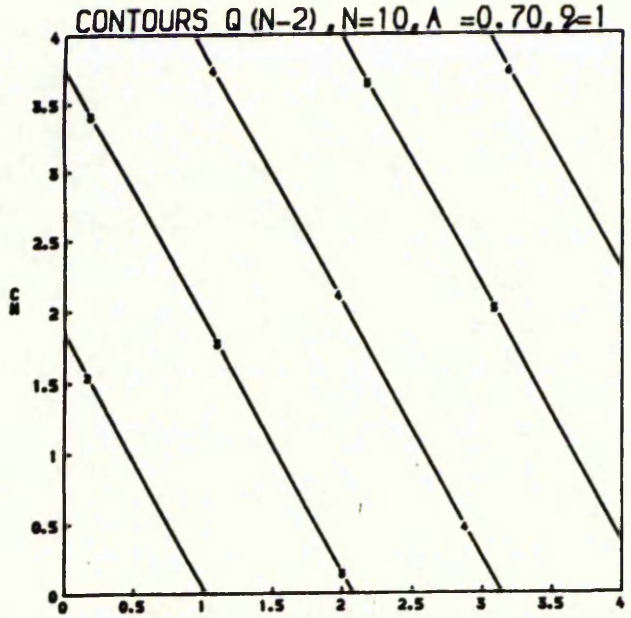
FIGURE 6.8


CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500


CONTOURS Q(N-2), N=50, A =0.50, Q=3
EM
EN


CONTOURS Q(N-2), N=50, A =0.50, Q=4
EM
EN




CONTOURS Q(N-2), N=10, A =0.70, Q=1
CN
BN

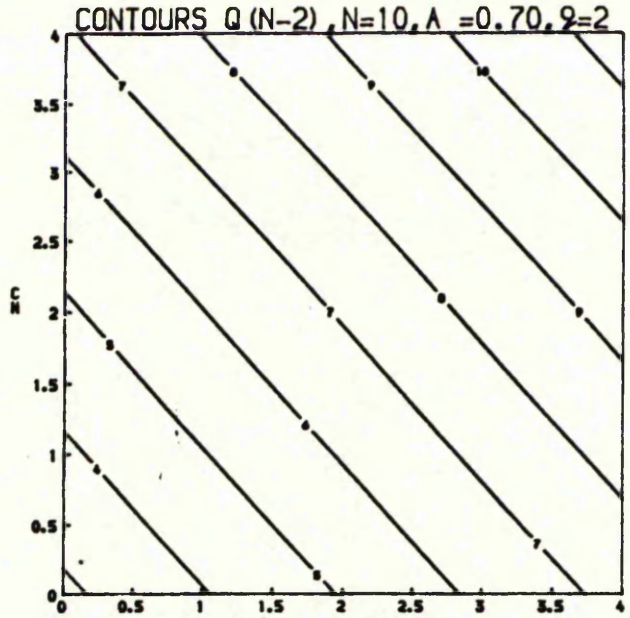
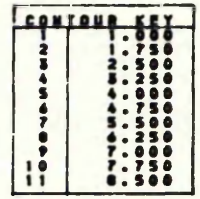
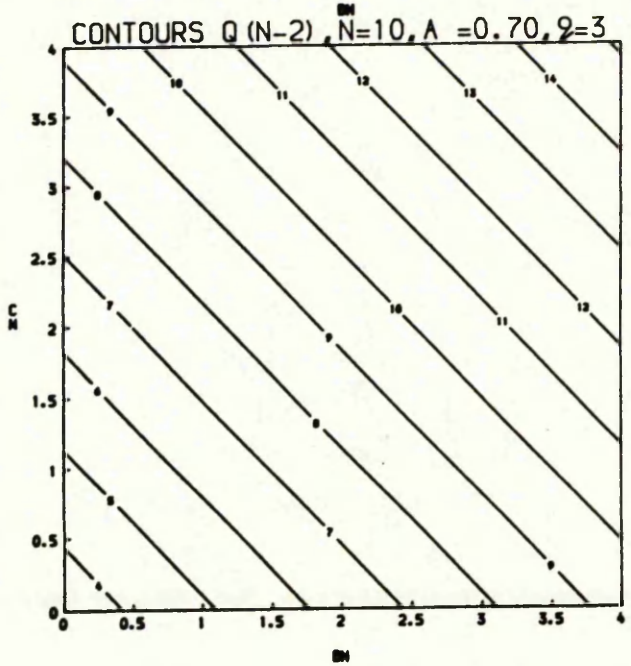
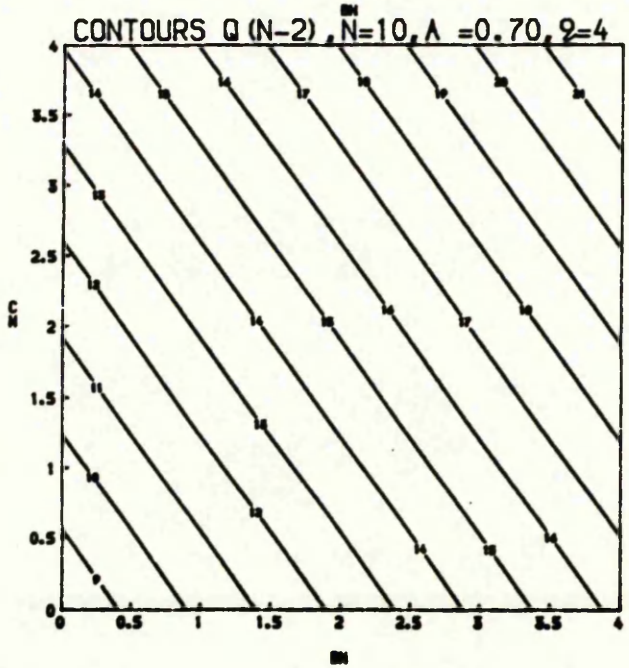
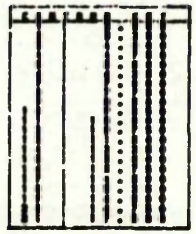

CONTOURS Q(N-2), N=10, A =0.70, Q=2
CN
BN

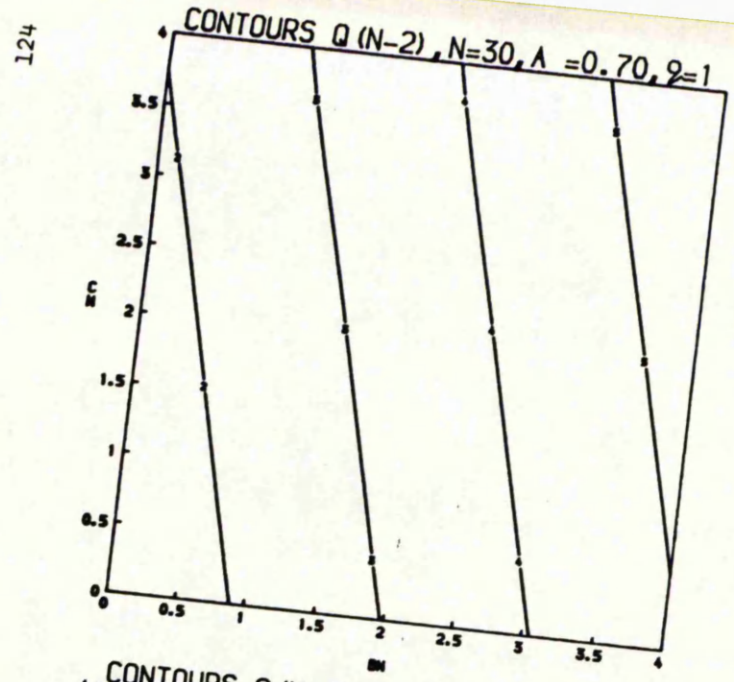
FIGURE 6.9

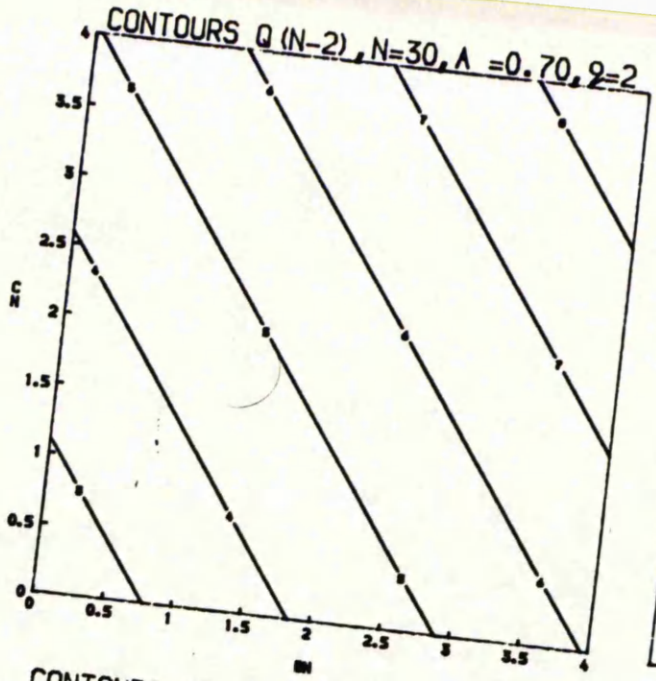

CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500


CONTOURS Q(N-2), N=10, A =0.70, Q=3
CN
BN


CONTOURS Q(N-2), N=10, A =0.70, Q=4
CN
BN

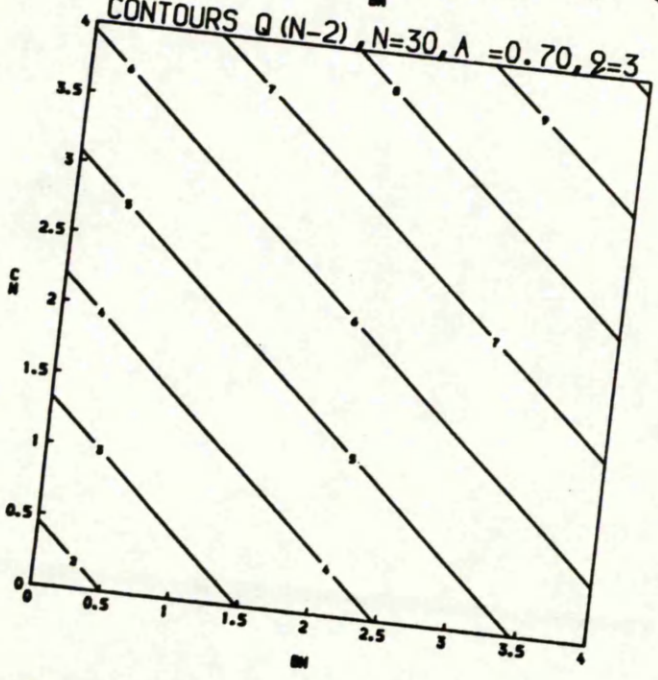


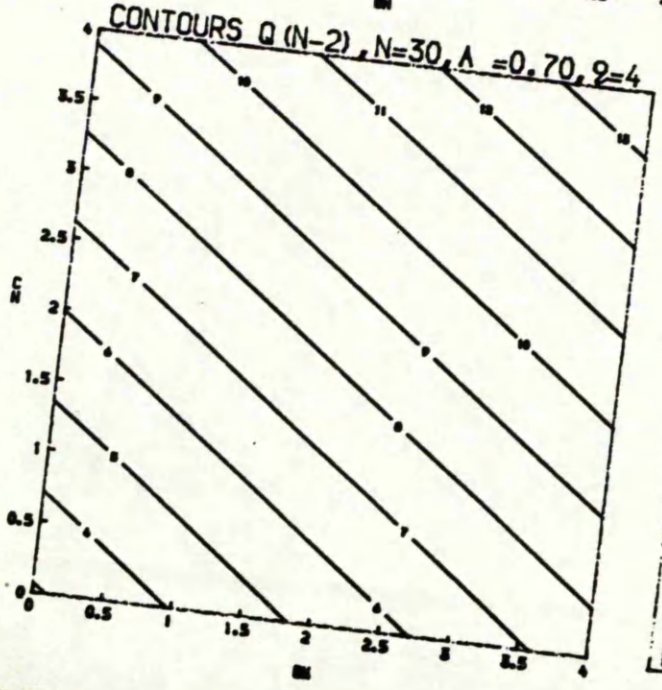

CONTOURS Q (N-2), N=30, Λ =0.70, Q=1
CM
BM

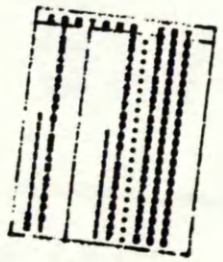

CONTOURS Q (N-2), N=30, Λ =0.70, Q=2
CM
BM

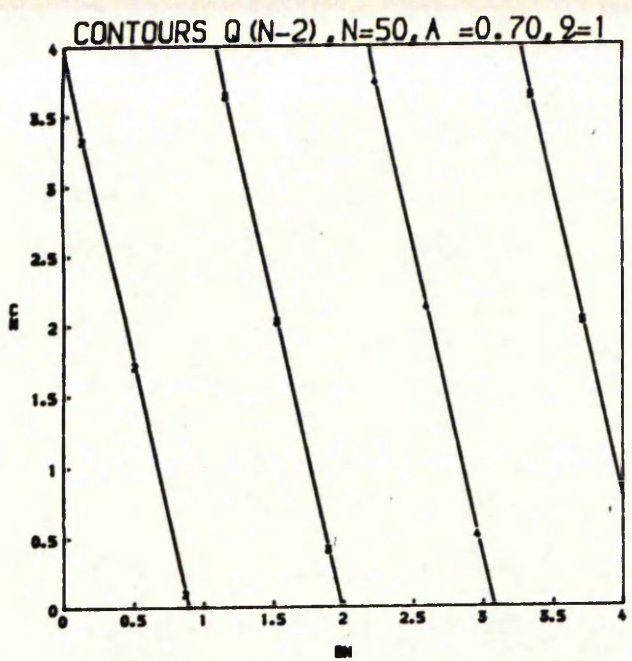

CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500

FIGURE 6.10


CONTOURS Q (N-2), N=30, Λ =0.70, Q=3
CM
BM


CONTOURS Q (N-2), N=30, Λ =0.70, Q=4
CM
BM




CONTOURS Q(N-2), N=50, A =0.70, Q=1
EN
EM

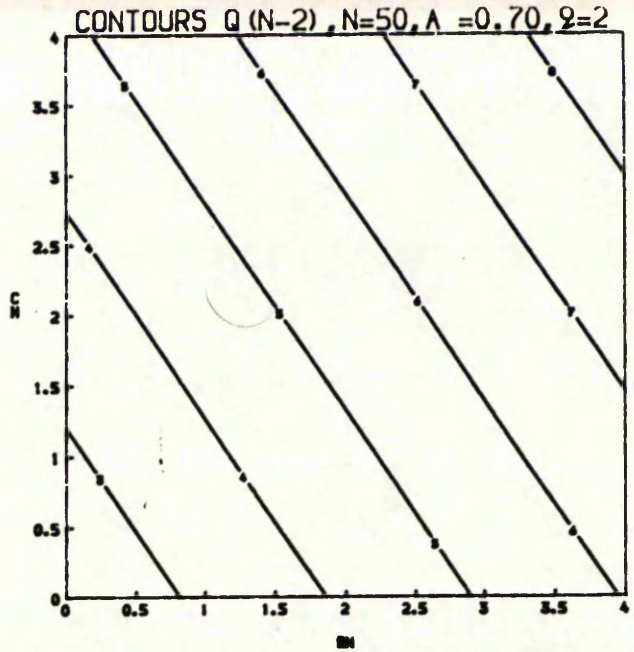
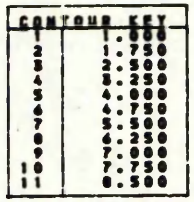
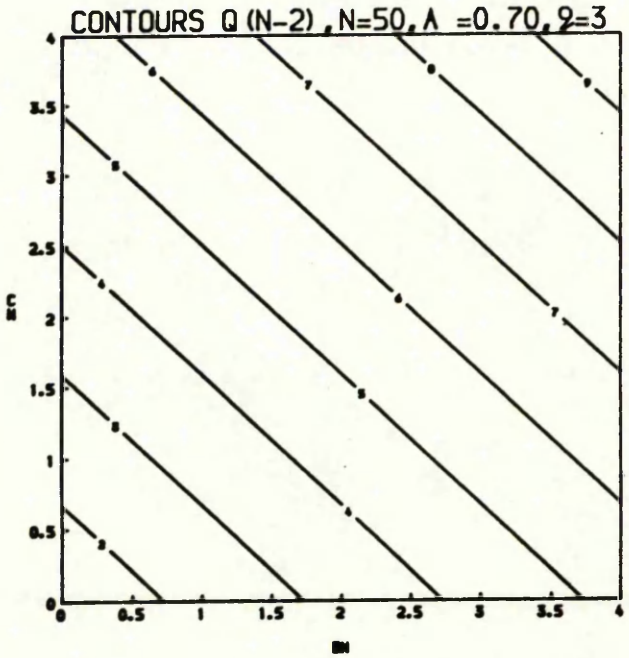
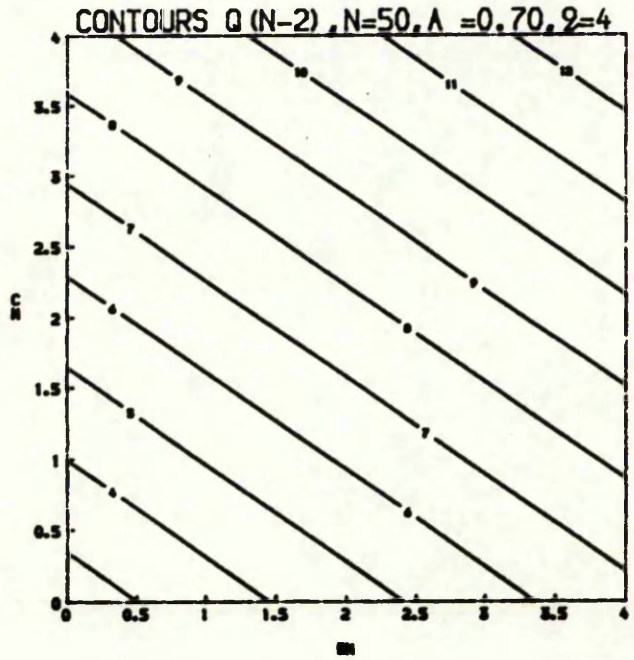
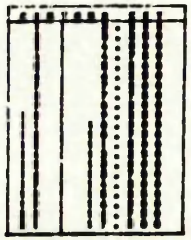

CONTOURS Q(N-2), N=50, A =0.70, Q=2
EN
EM

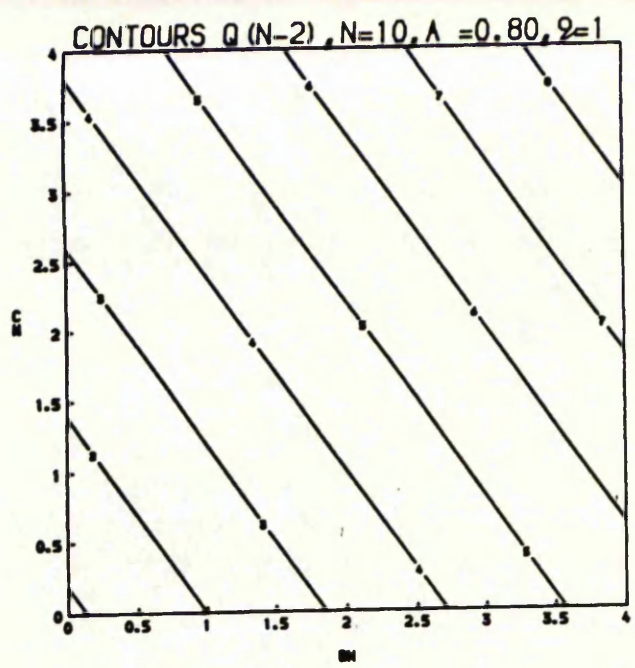
FIGURE 6.11


CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500


CONTOURS Q(N-2), N=50, A =0.70, Q=3
EN
EM


CONTOURS Q(N-2), N=50, A =0.70, Q=4
EN
EM





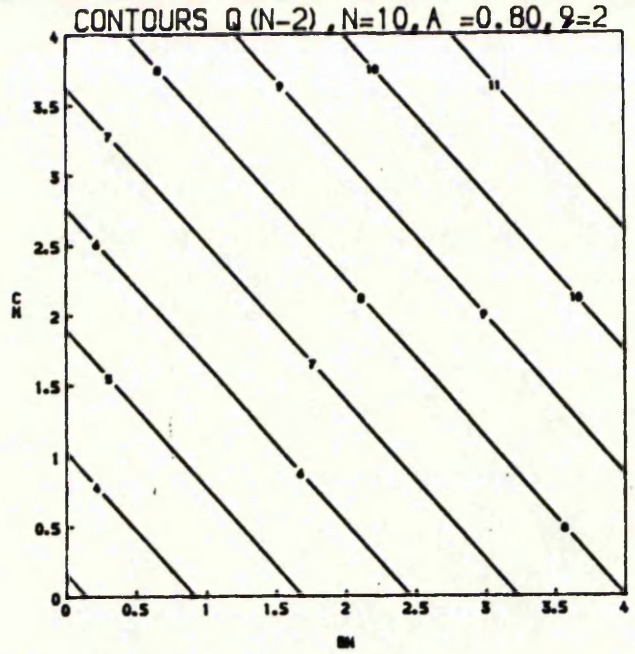
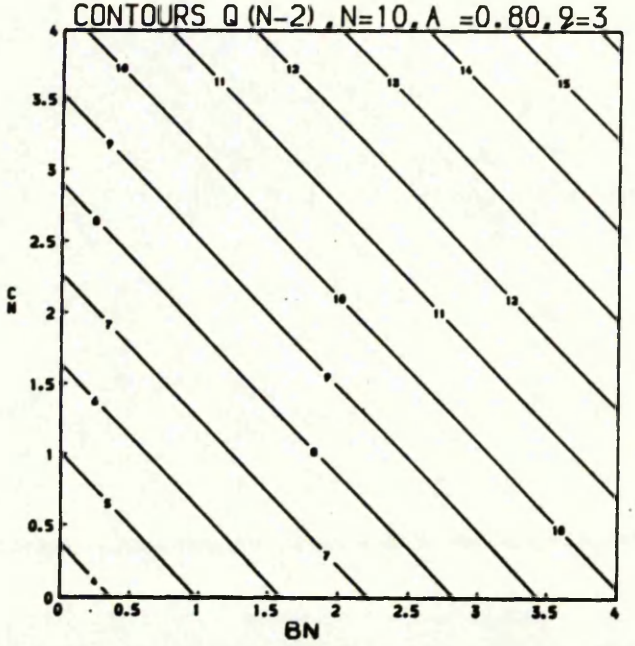
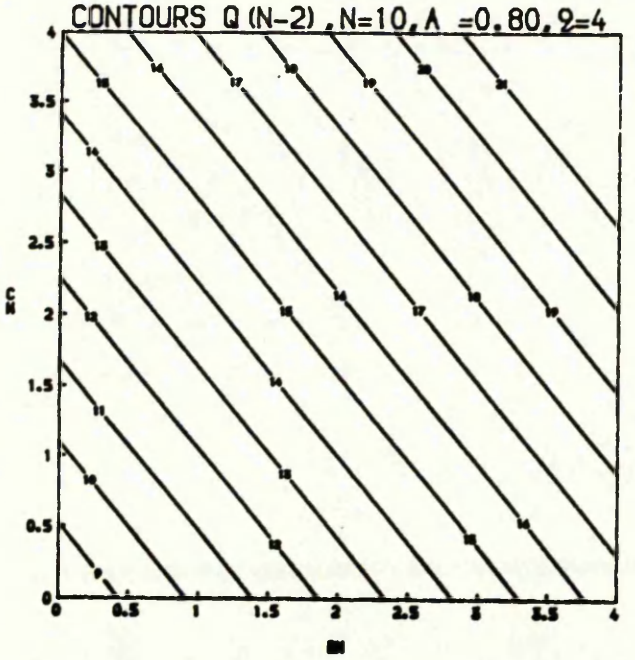
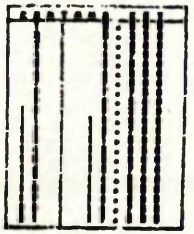


FIGURE 6.12

CONTOUR KEY	
1	1.000
2	1.750
3	2.500
4	3.250
5	4.000
6	4.750
7	5.500
8	6.250
9	7.000
10	7.750
11	8.500







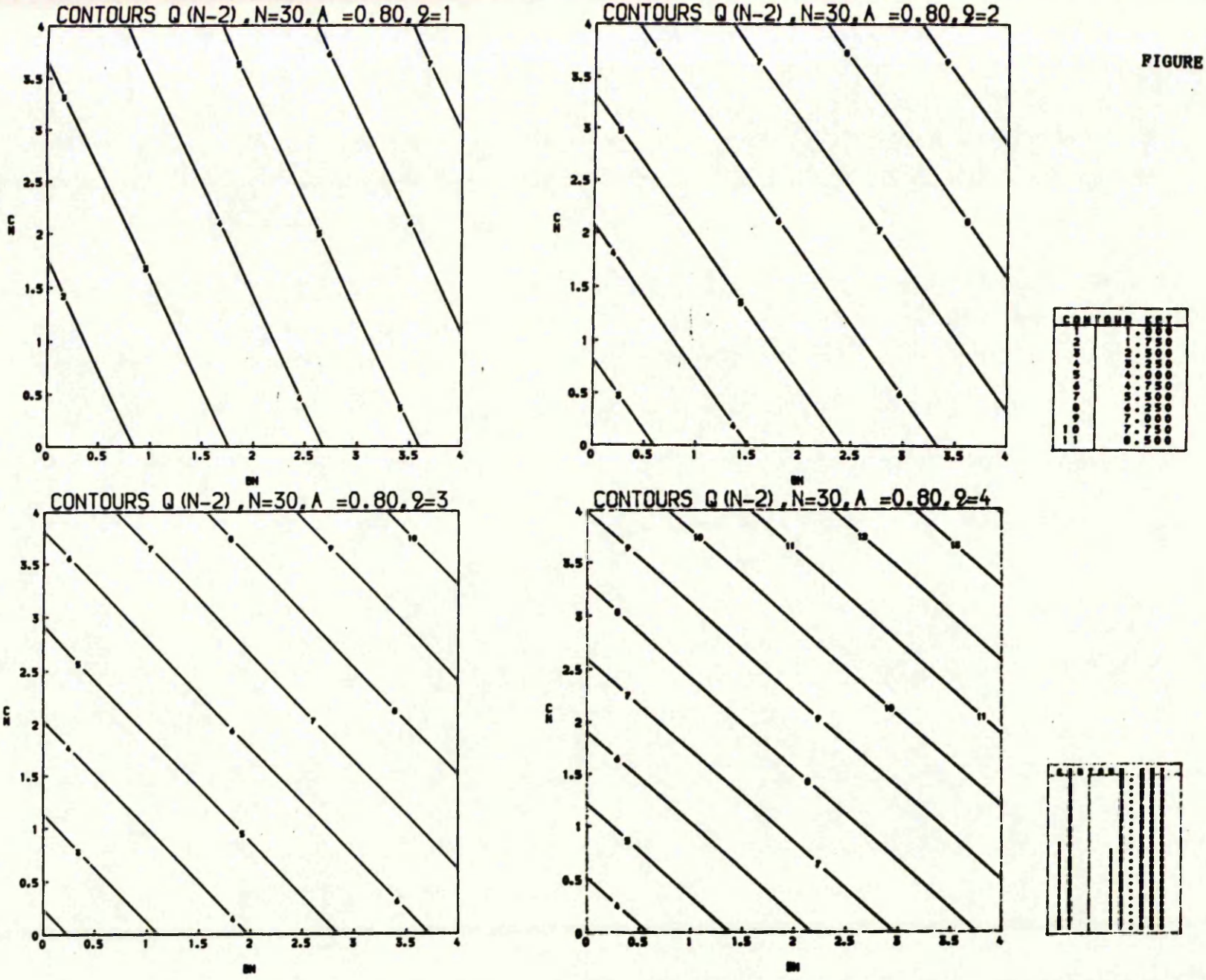
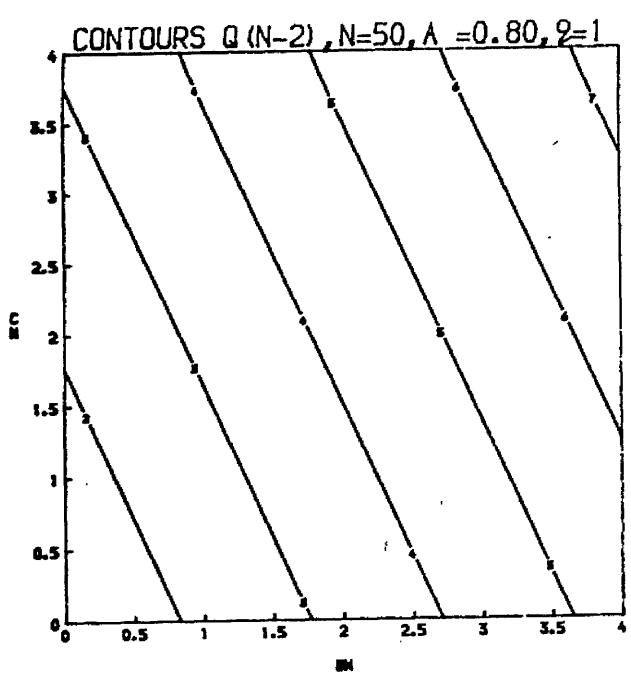

CONTOURS Q(N-2), N=30, A =0.80, Q=1
CONTOURS Q(N-2), N=30, A =0.80, Q=2
CONTOURS Q(N-2), N=30, A =0.80, Q=3
CONTOURS Q(N-2), N=30, A =0.80, Q=4
CN
BN
CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500

FIGURE 6.13



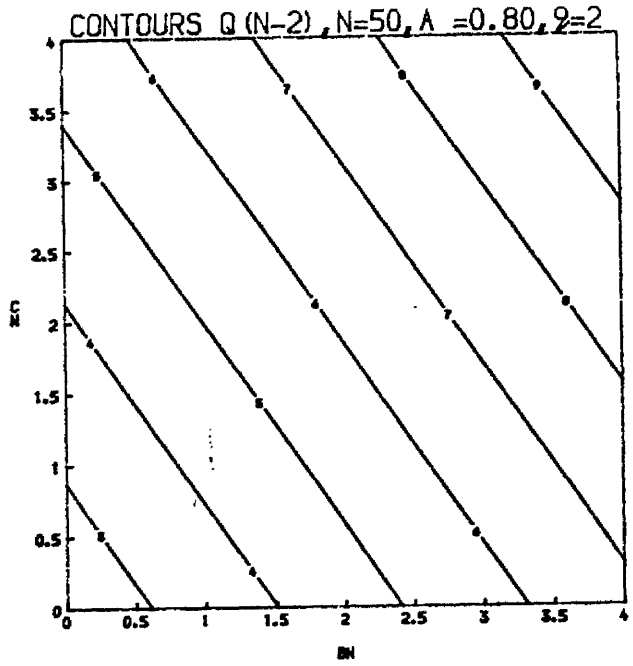
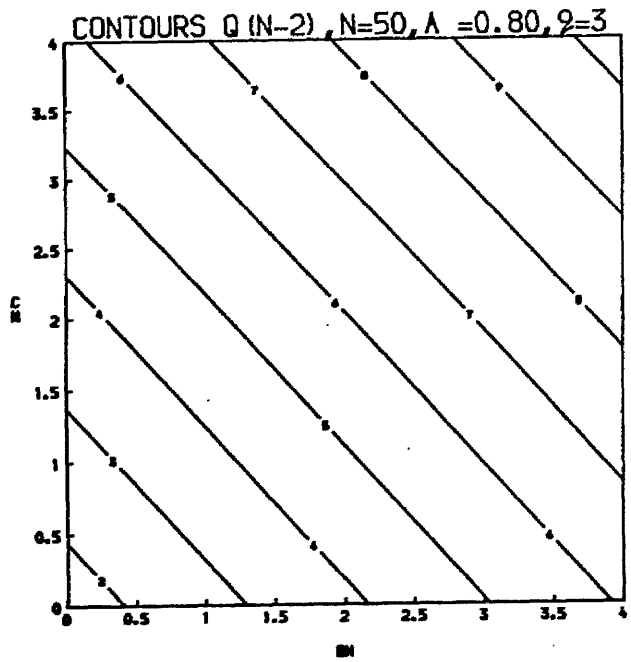
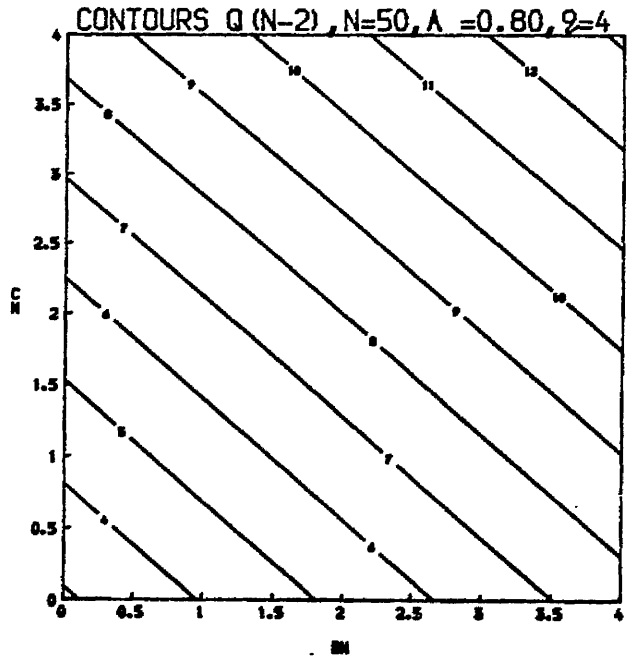
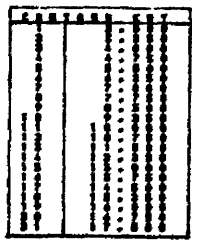


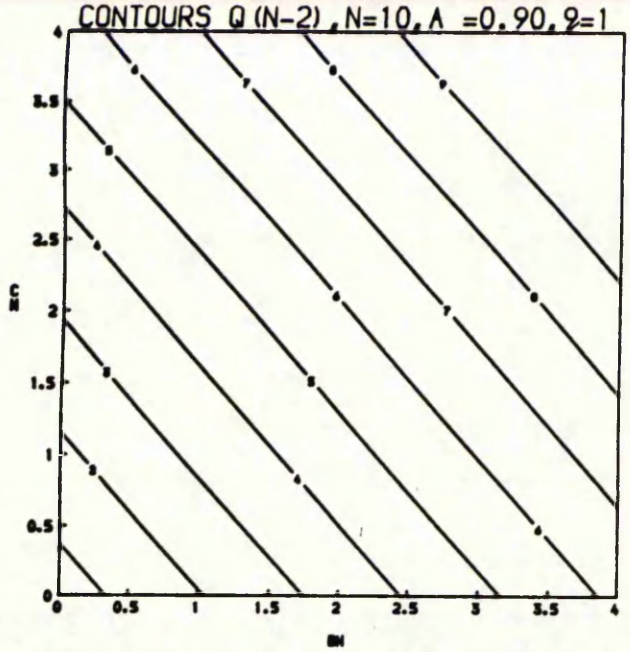
FIGURE 6.14

CONTOUR	KEY
1	1.000
2	1.750
3	2.500
4	3.250
5	4.000
6	4.750
7	5.500
8	6.250
9	7.000
10	7.750
11	8.500









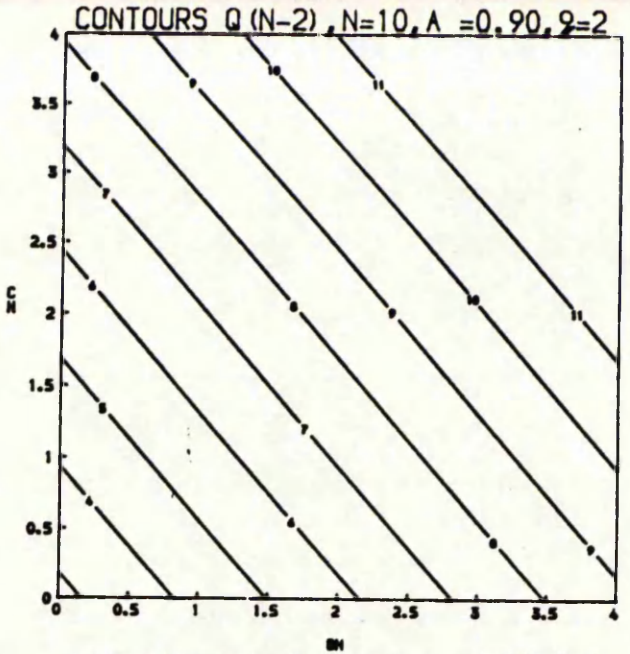
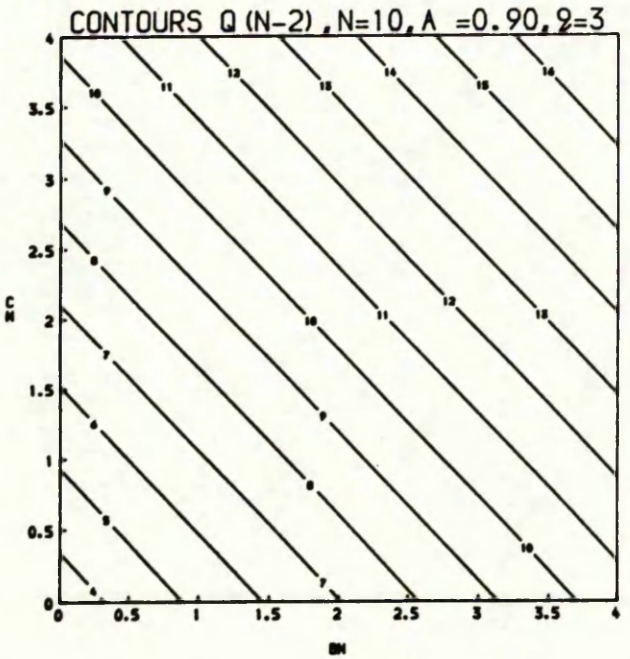
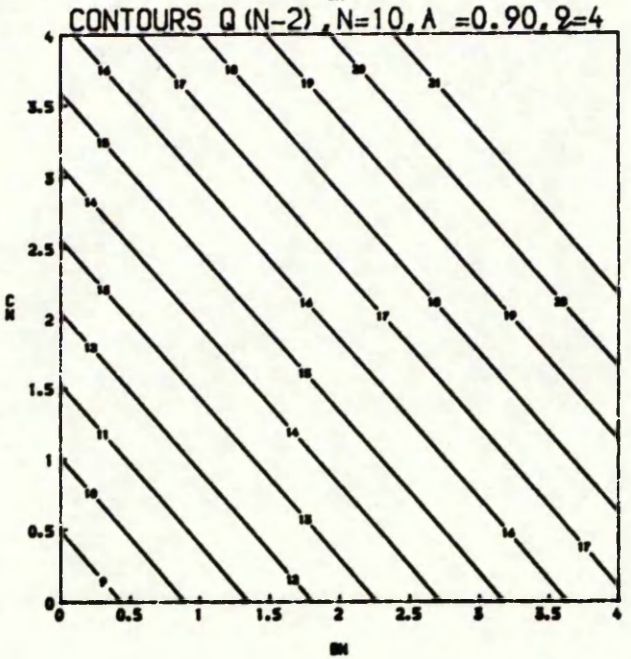
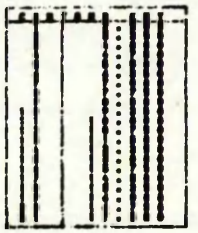


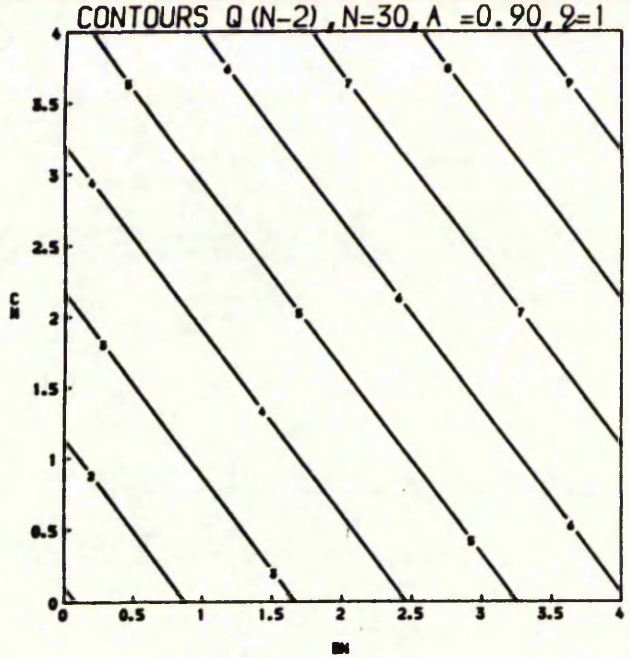
FIGURE 6.15

CONTOUR	KEY
1	1.000
2	1.750
3	2.500
4	3.250
5	4.000
6	4.750
7	5.500
8	6.250
9	7.000
10	7.750
11	8.500









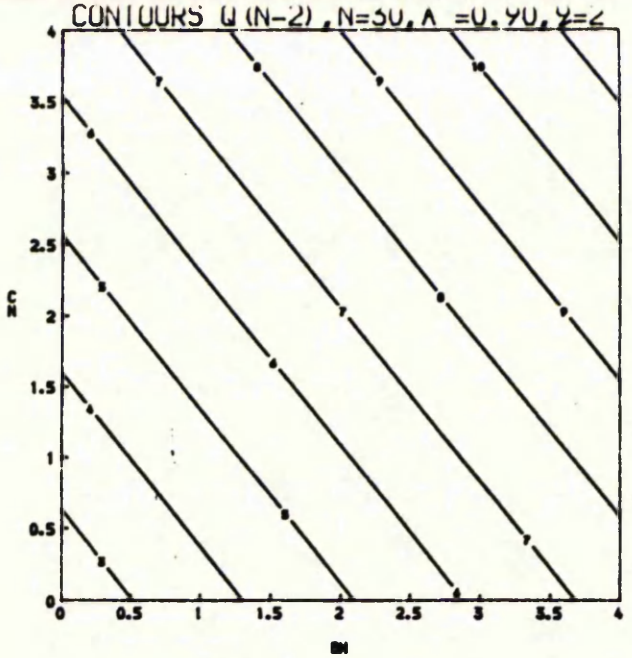
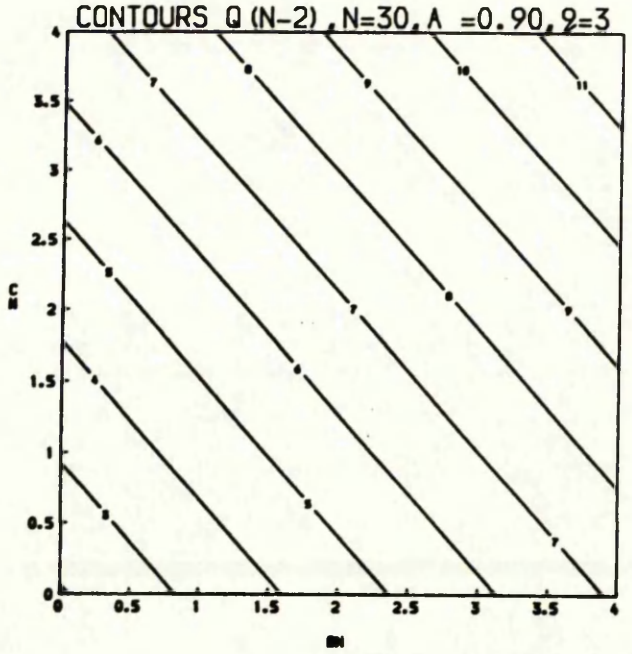
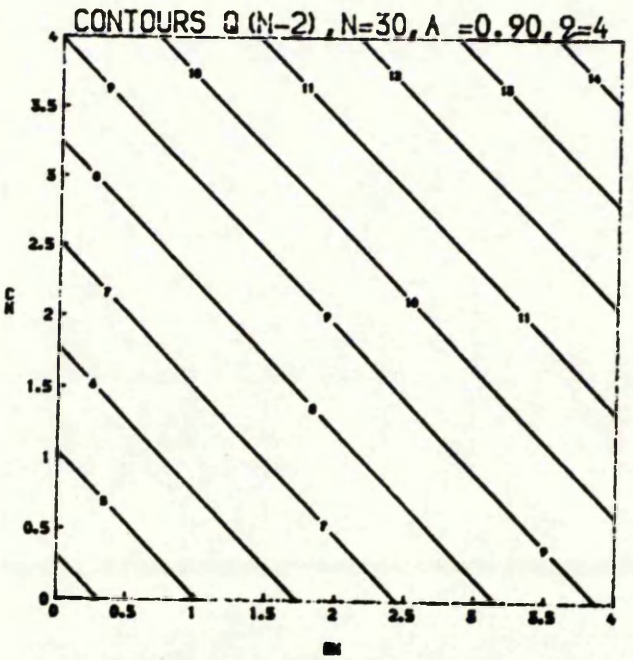
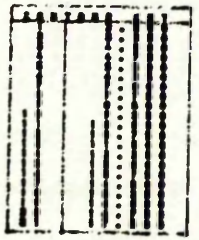


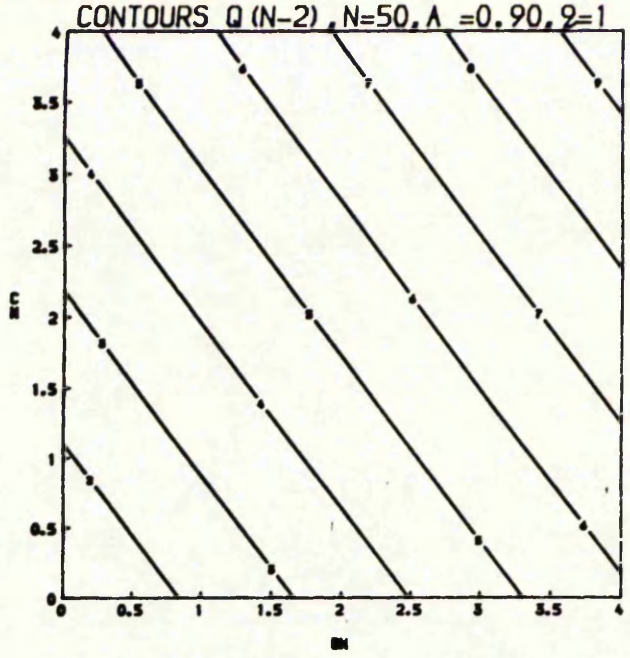
FIGURE 6.16

CONTOUR	KEY
1	1.000
2	1.750
3	2.500
4	3.250
5	4.000
6	4.750
7	5.500
8	6.250
9	7.000
10	7.750
11	8.500








CONTOURS Q(N-2), N=50, A =0.90, Q=1
BN
CN

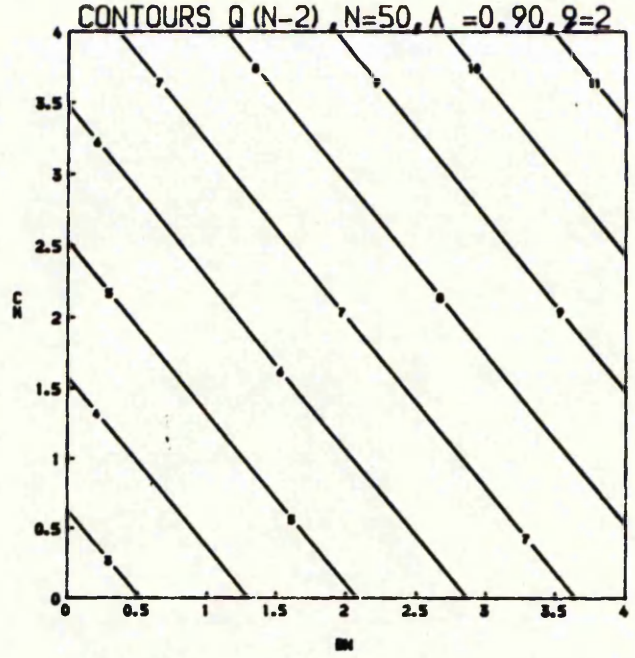
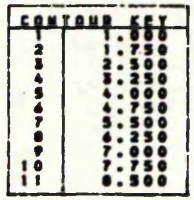
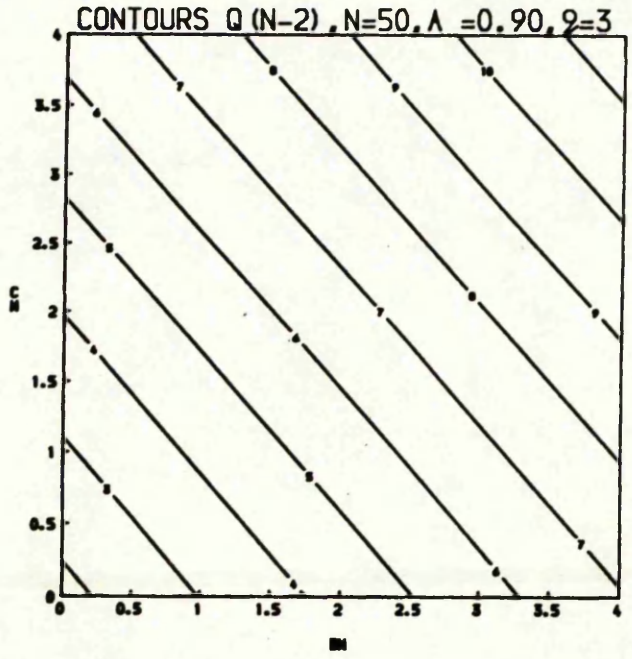
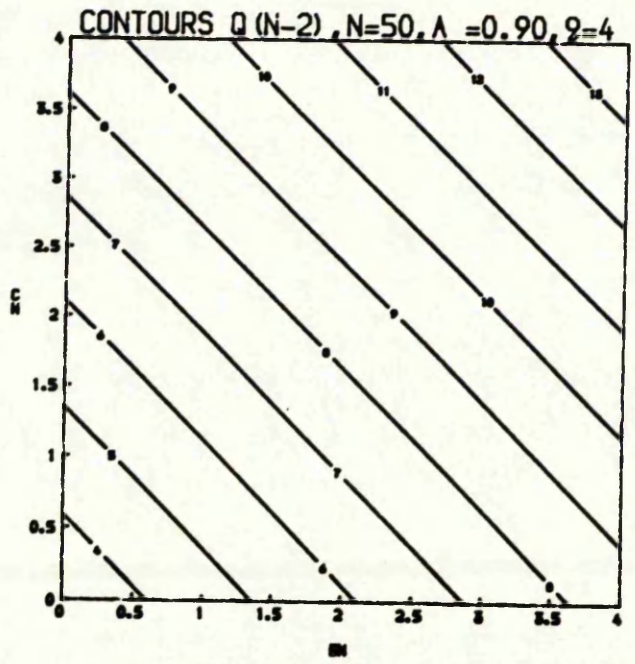
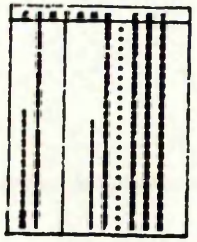

CONTOURS Q(N-2), N=50, A =0.90, Q=2
BN
CN

FIGURE 6.17


CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500


CONTOURS Q(N-2), N=50, A =0.90, Q=3
BN
CN


CONTOURS Q(N-2), N=50, A =0.90, Q=4
BN
CN



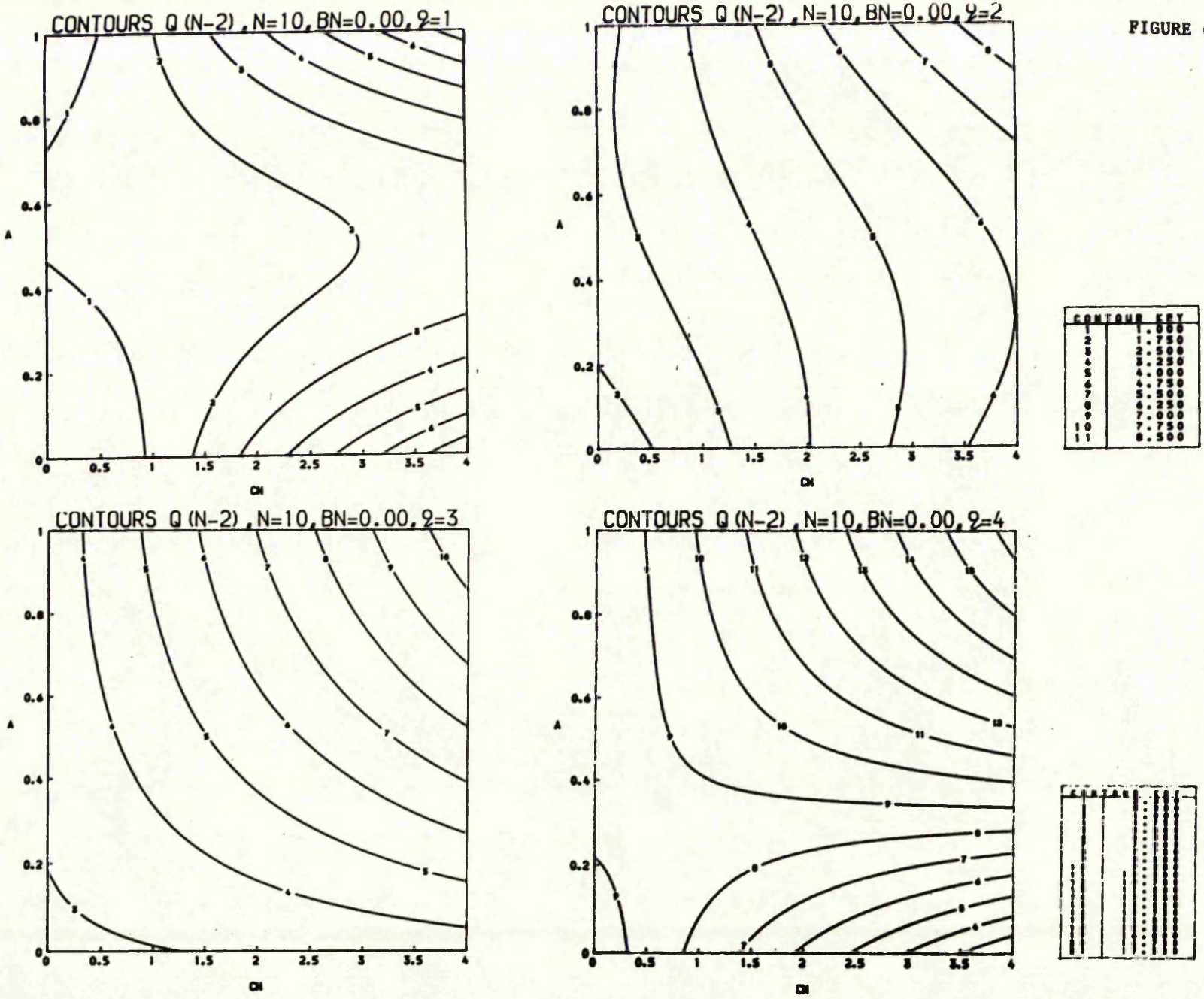

CONTOURS Q(N-2), N=10, BN=0.00, Q=1
CONTOURS Q(N-2), N=10, BN=0.00, Q=2
CONTOURS Q(N-2), N=10, BN=0.00, Q=3
CONTOURS Q(N-2), N=10, BN=0.00, Q=4
A
CN
CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500

FIGURE 6.18

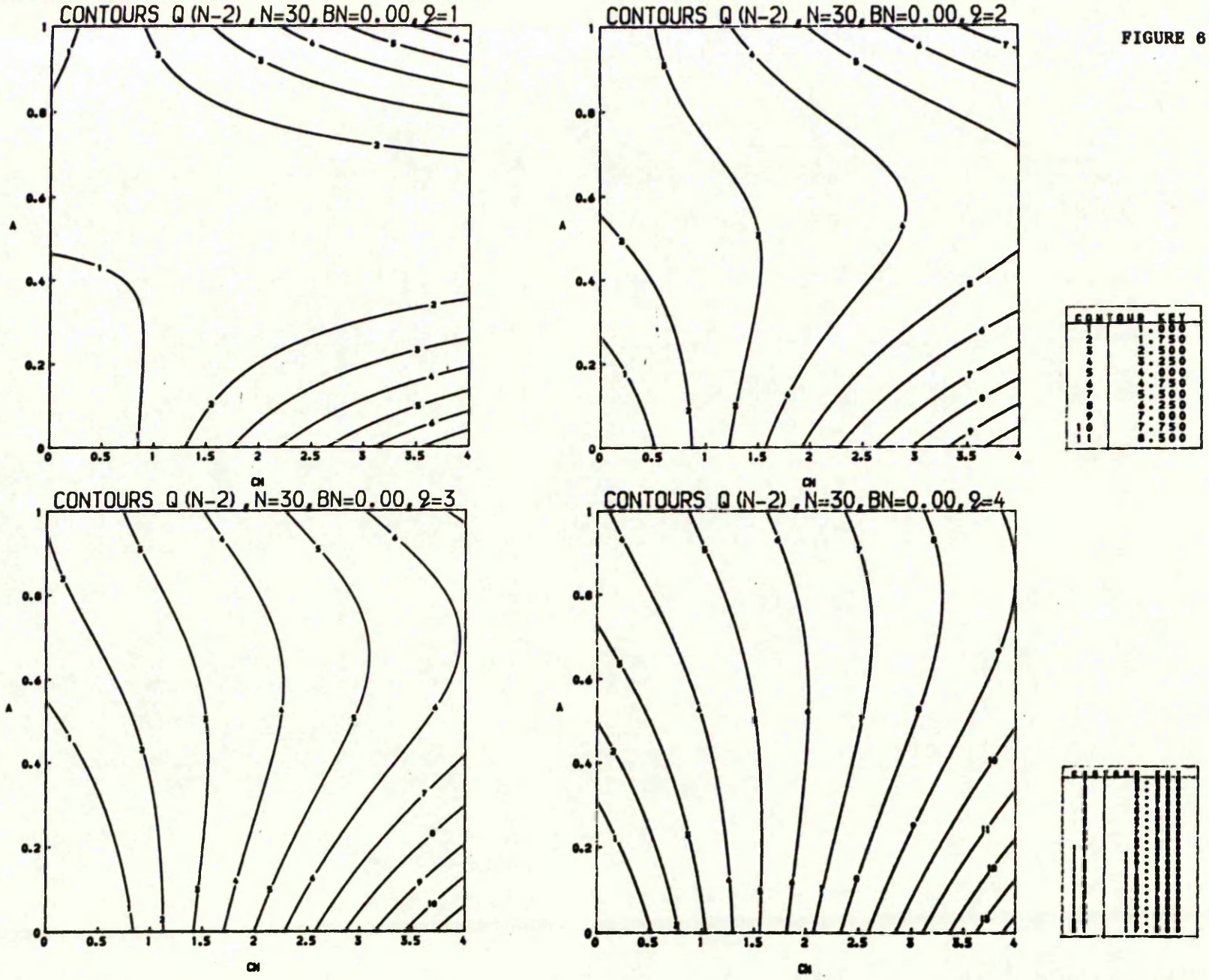
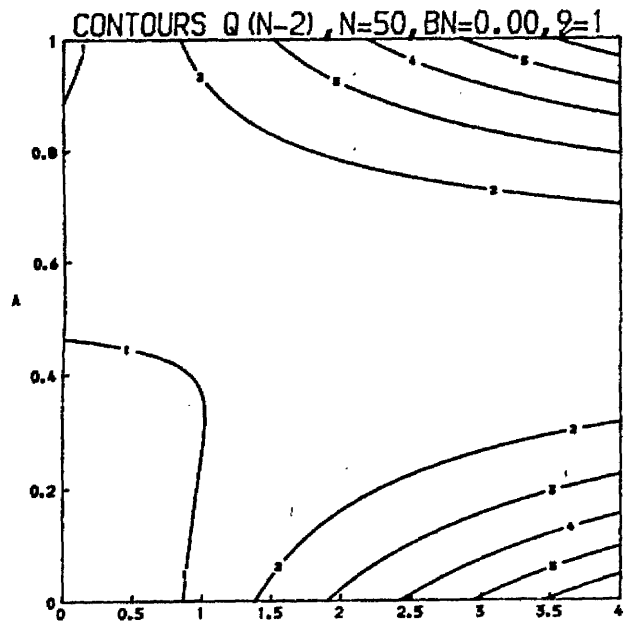
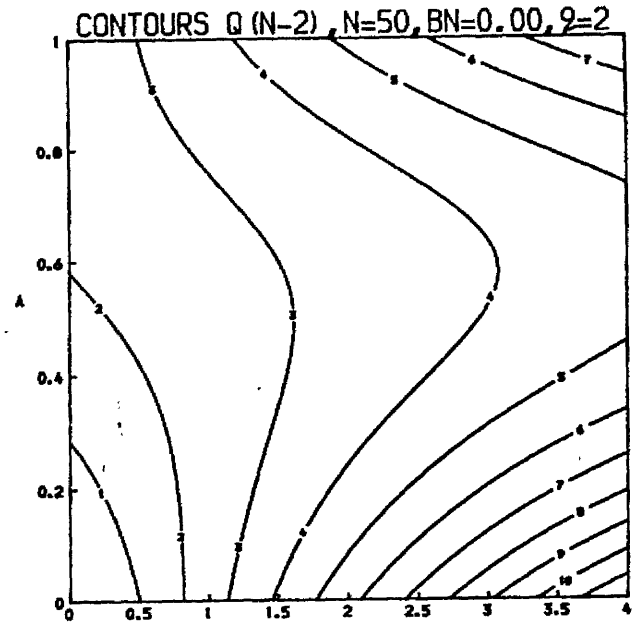
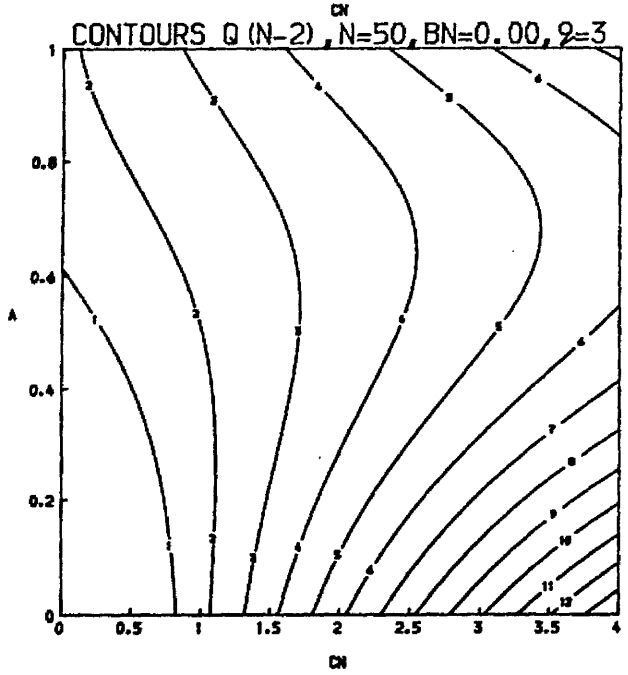
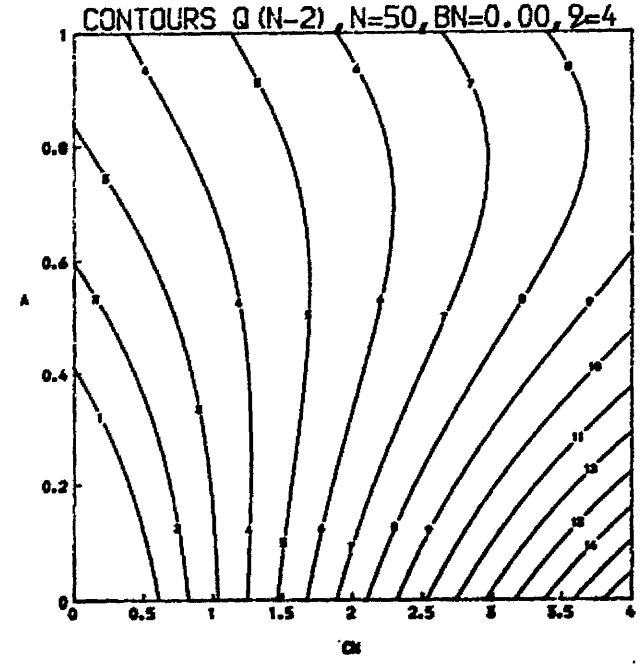

CONTOURS Q (N-2) , N=30, BN=0.00, Q=1
CONTOURS Q (N-2) , N=30, BN=0.00, Q=2
CONTOURS Q (N-2) , N=30, BN=0.00, Q=3
CONTOURS Q (N-2) , N=30, BN=0.00, Q=4
A
CN
CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500

FIGURE 6.19









CONTOUR	KEY
1	1.000
2	1.750
3	2.500
4	3.250
5	4.000
6	4.750
7	5.500
8	6.250
9	7.000
10	7.750
11	8.500

FIGURE 6.20

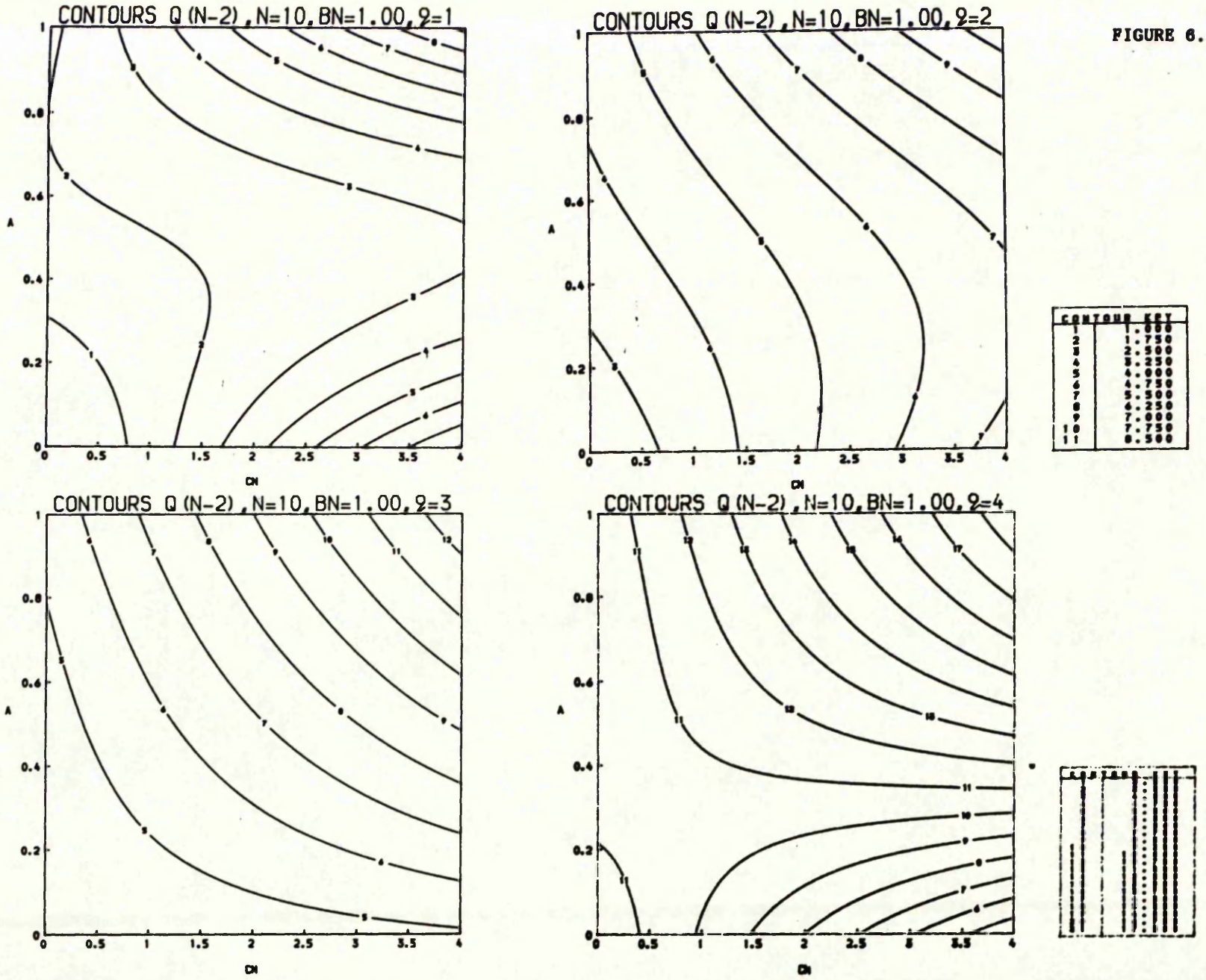

CONTOURS Q (N-2) , N=10, BN=1.00, Q=1
CONTOURS Q (N-2) , N=10, BN=1.00, Q=2
CONTOURS Q (N-2) , N=10, BN=1.00, Q=3
CONTOURS Q (N-2) , N=10, BN=1.00, Q=4
A
CN
CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500

FIGURE 6.21

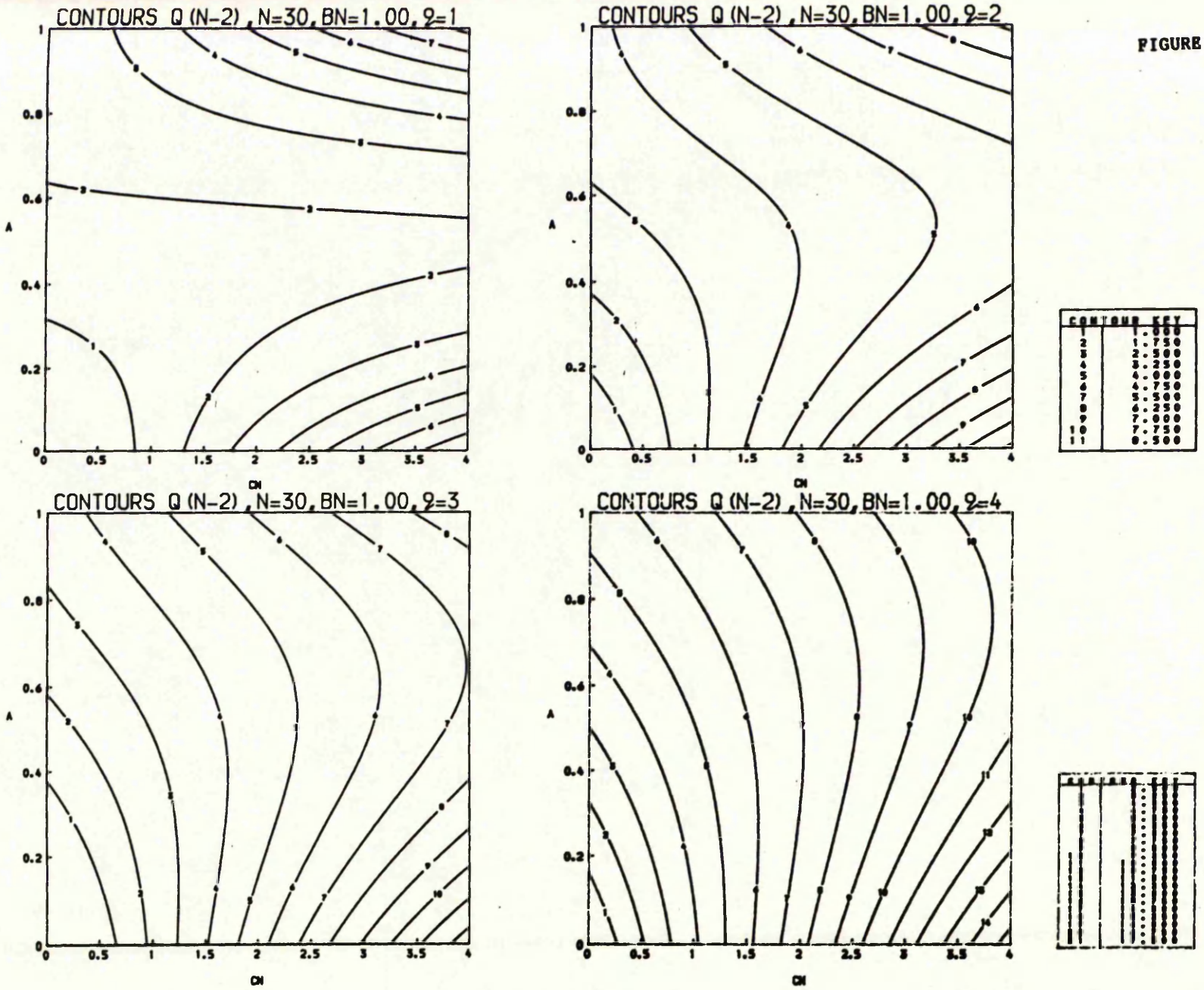
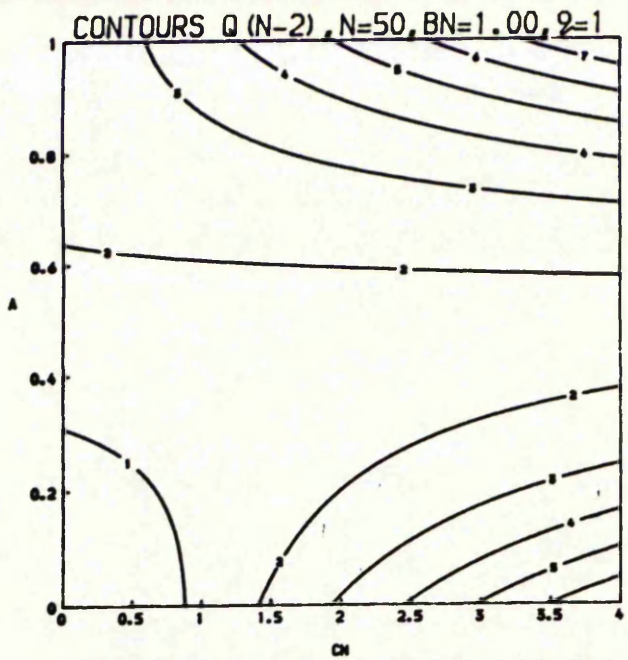

CONTOURS Q(N-2), N=30, BN=1.00, Q=1
CONTOURS Q(N-2), N=30, BN=1.00, Q=2
CONTOURS Q(N-2), N=30, BN=1.00, Q=3
CONTOURS Q(N-2), N=30, BN=1.00, Q=4
A
CN
CONTOUR KEY

FIGURE 6.22


CONTOURS Q (N-2) , N=50, BN=1.00, Q=1
A
CN

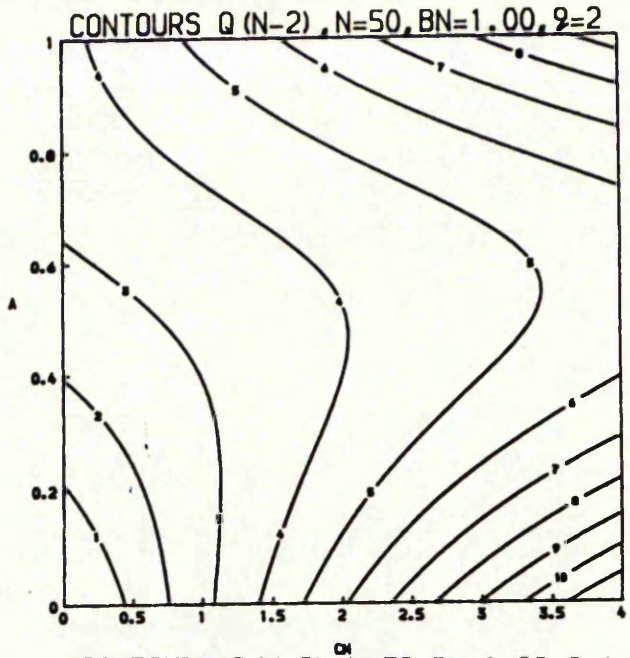
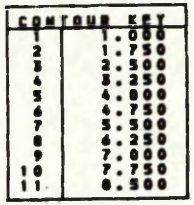
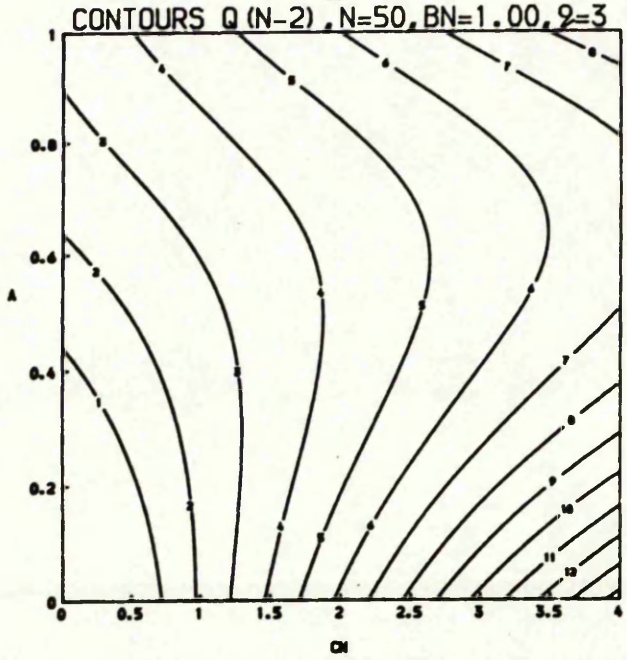
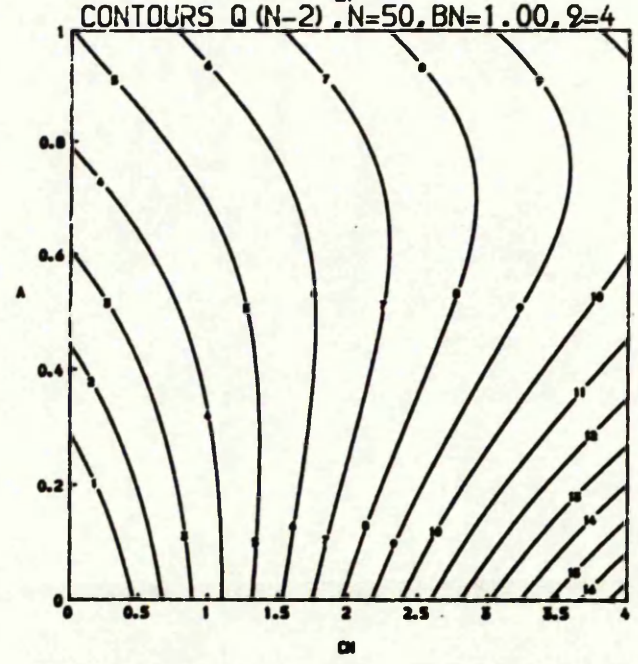
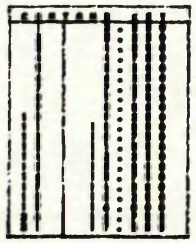

CONTOURS Q (N-2) , N=50, BN=1.00, Q=2
A
CN

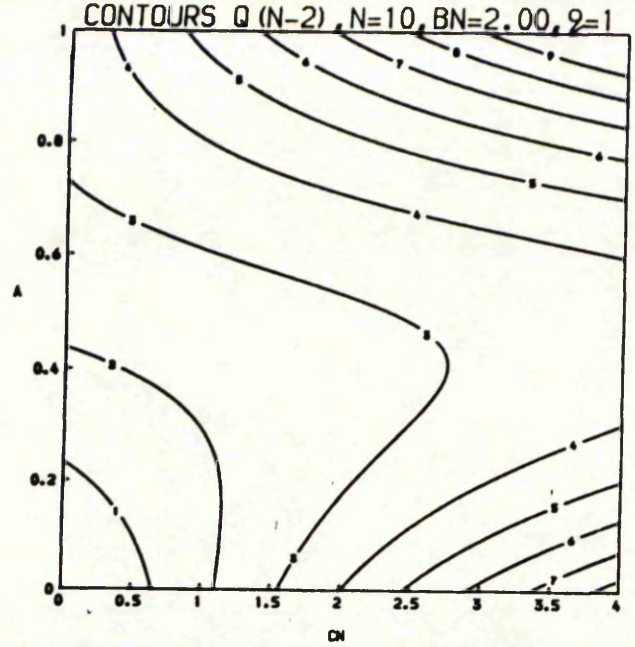
FIGURE 6.23


CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500


CONTOURS Q (N-2) , N=50, BN=1.00, Q=3
A
CN


CONTOURS Q (N-2) , N=50, BN=1.00, Q=4
A
CN




CONTOURS Q (N-2) , N=10, BN=2.00, Q=1
A
CN

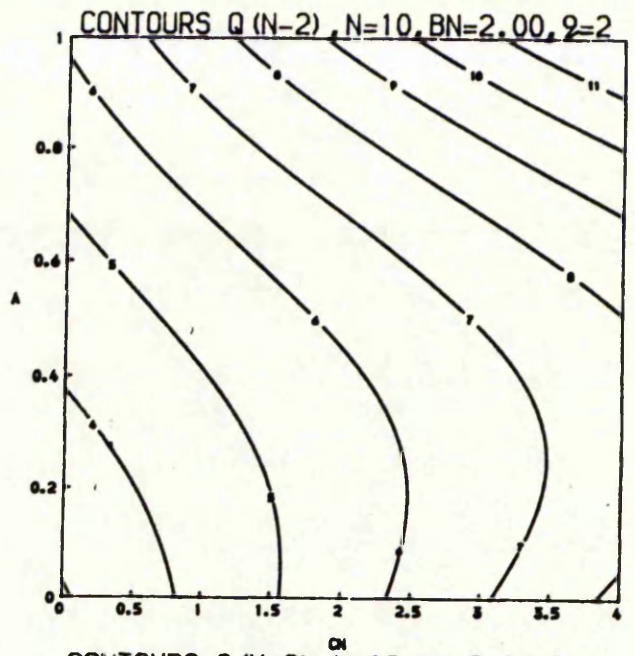
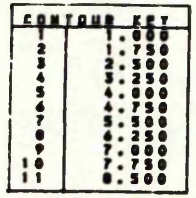
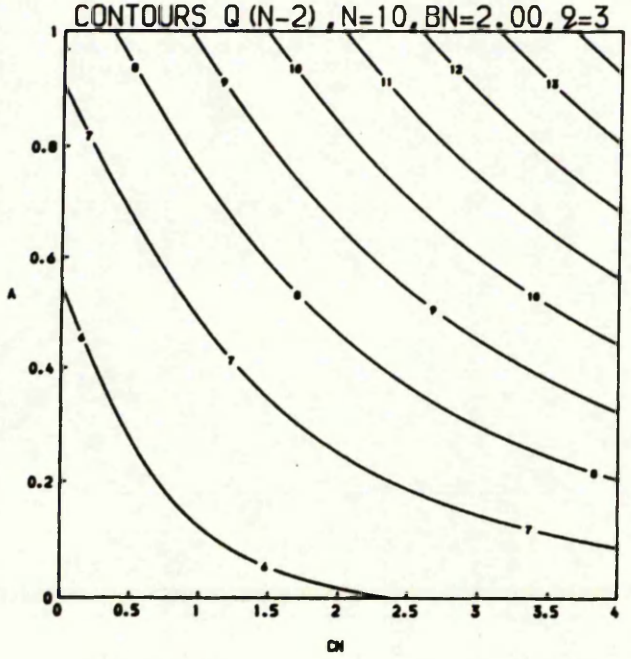
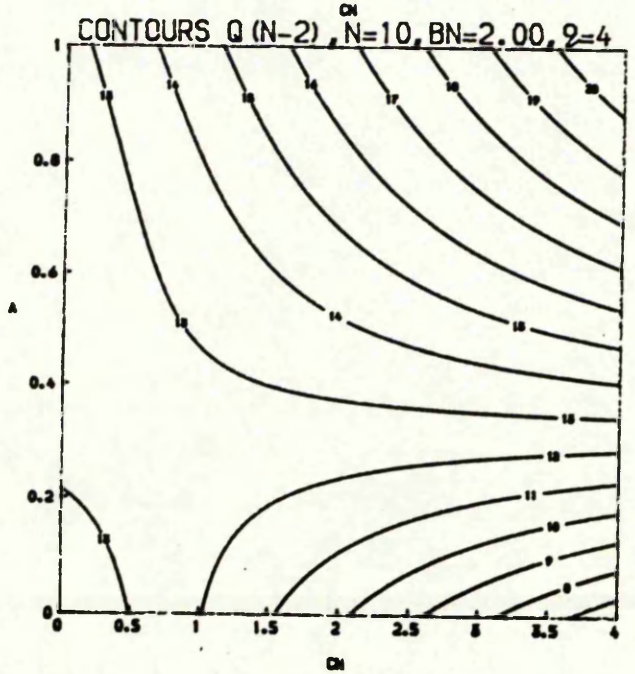
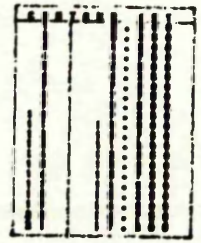

CONTOURS Q (N-2) , N=10, BN=2.00, Q=2
A
CN

FIGURE 6.24


CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500


CONTOURS Q (N-2) , N=10, BN=2.00, Q=3
A
CN


CONTOURS Q (N-2) , N=10, BN=2.00, Q=4
A
CN



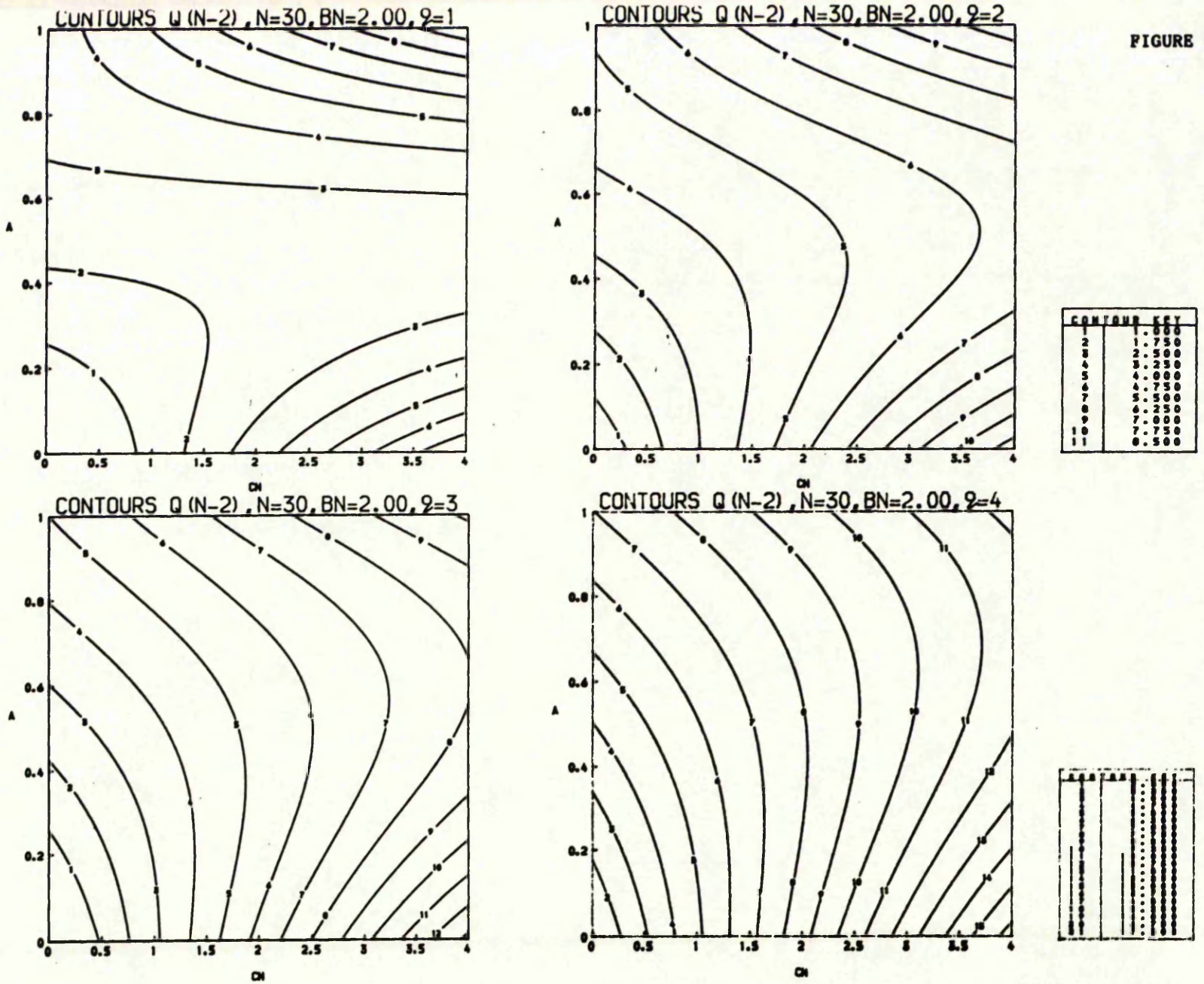

CONTOURS Q(N-2), N=30, BN=2.00, Q=1
CONTOURS Q(N-2), N=30, BN=2.00, Q=2
CONTOURS Q(N-2), N=30, BN=2.00, Q=3
CONTOURS Q(N-2), N=30, BN=2.00, Q=4
A
CN
CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500

FIGURE 6.25

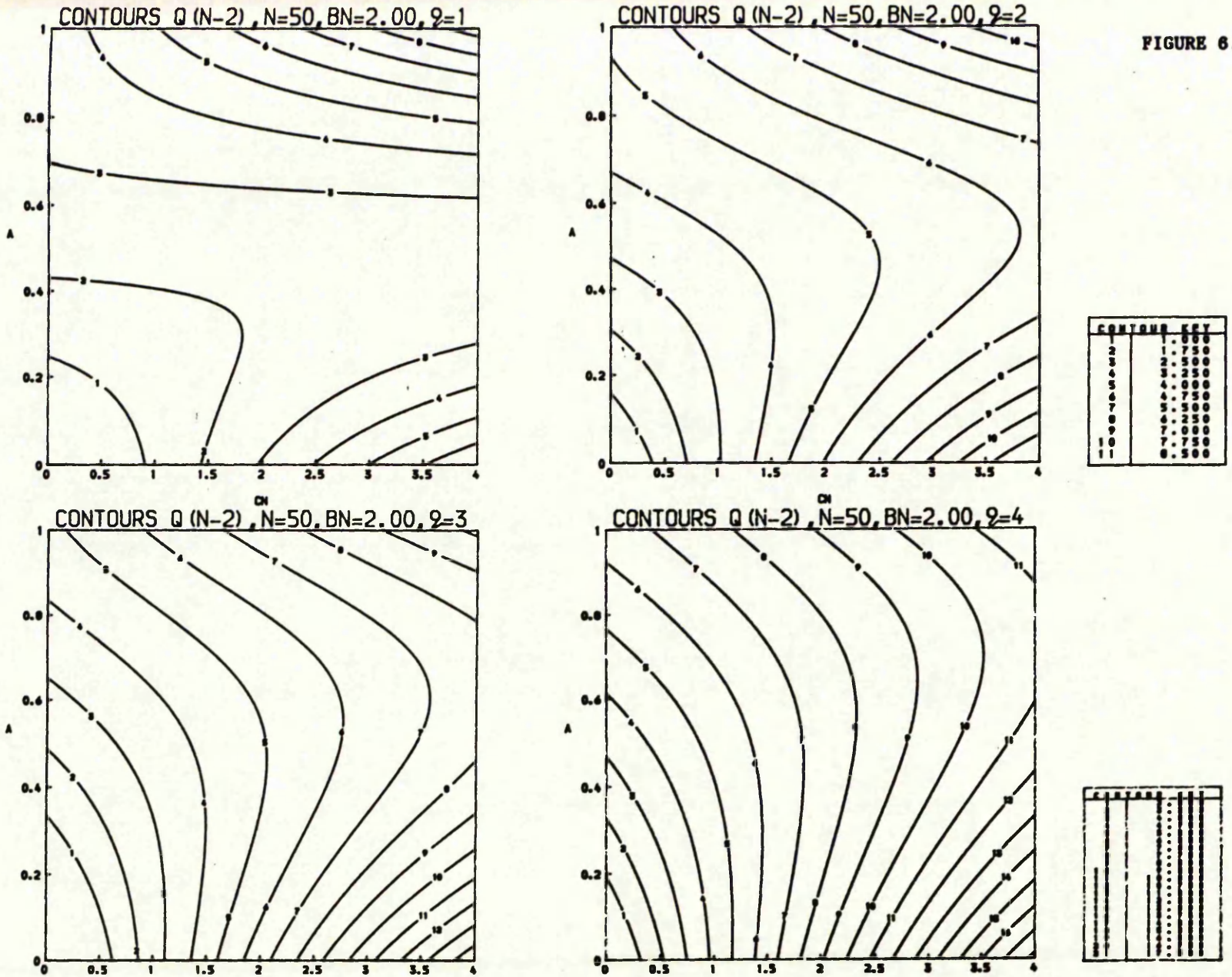
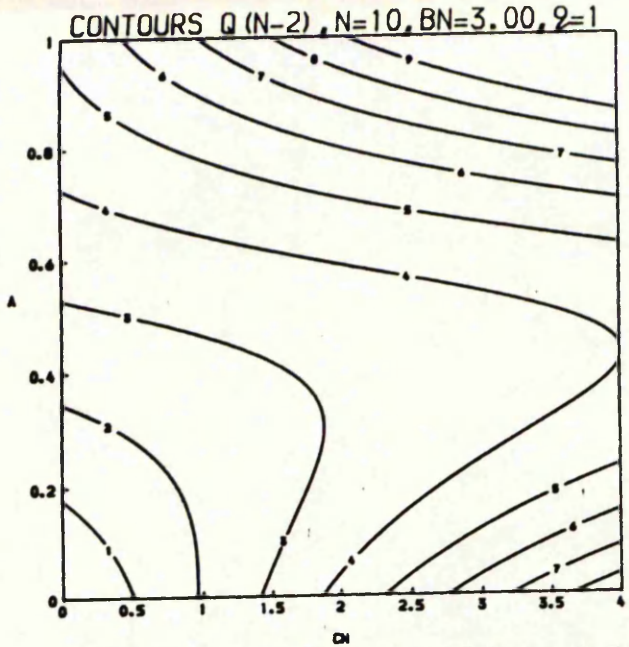

CONTOURS Q (N-2) , N=50, BN=2.00, 9=1
CONTOURS Q (N-2) , N=50, BN=2.00, 9=2
CONTOURS Q (N-2) , N=50, BN=2.00, 9=3
CONTOURS Q (N-2) , N=50, BN=2.00, 9=4
A
CN
CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500

FIGURE 6.26


CONTOURS Q(N-2), N=10, BN=3.00, Q=1
A
CN

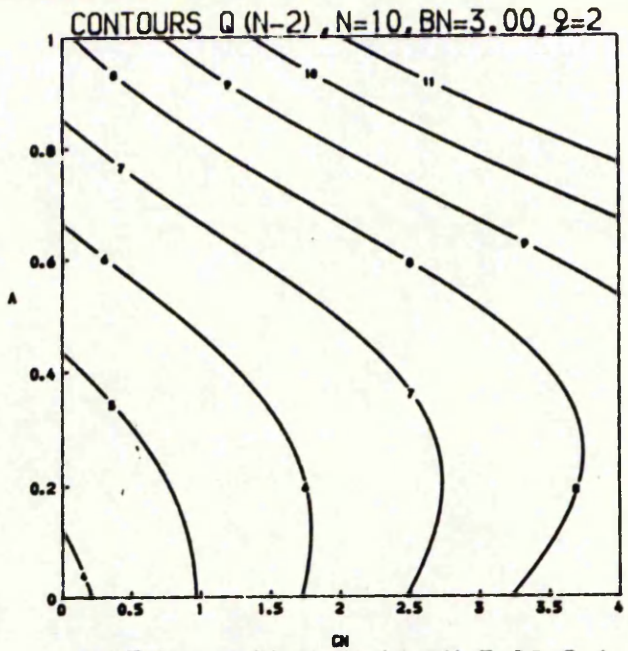
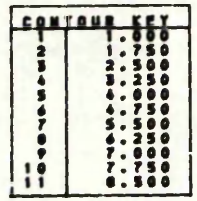
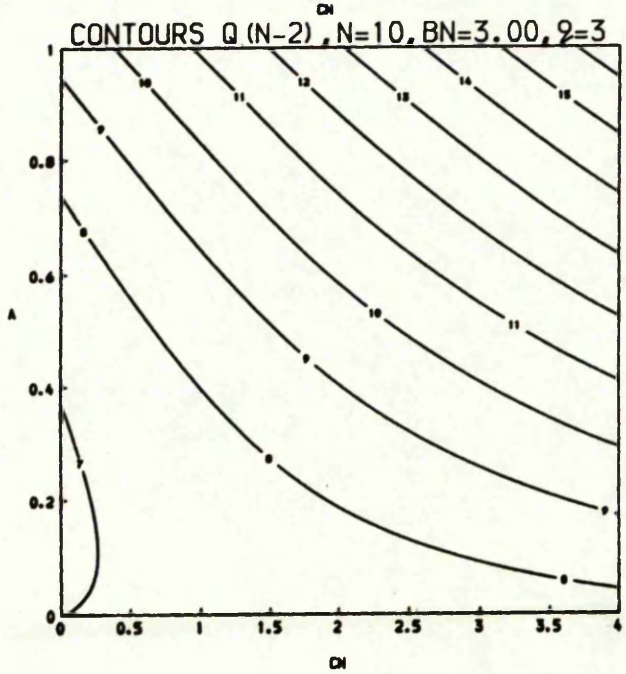
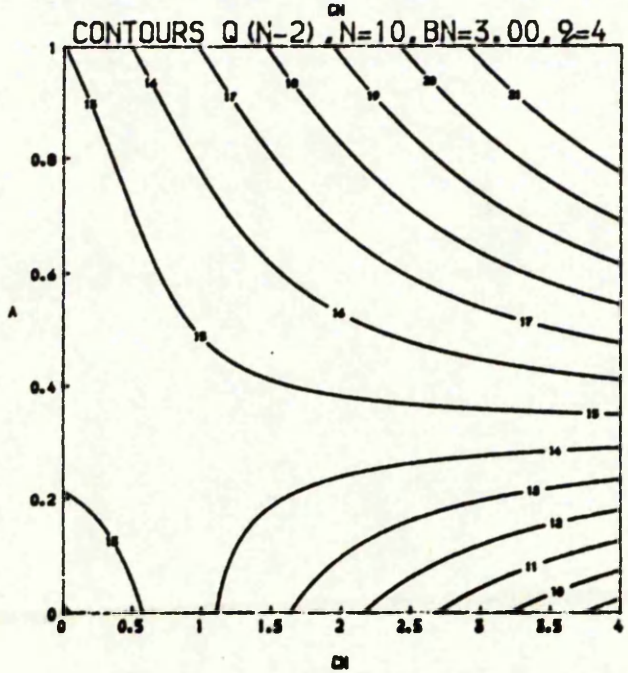
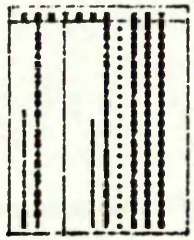

CONTOURS Q(N-2), N=10, BN=3.00, Q=2
A
CN

FIGURE 6.27


CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500


CONTOURS Q(N-2), N=10, BN=3.00, Q=3
A
CN


CONTOURS Q(N-2), N=10, BN=3.00, Q=4
A
CN



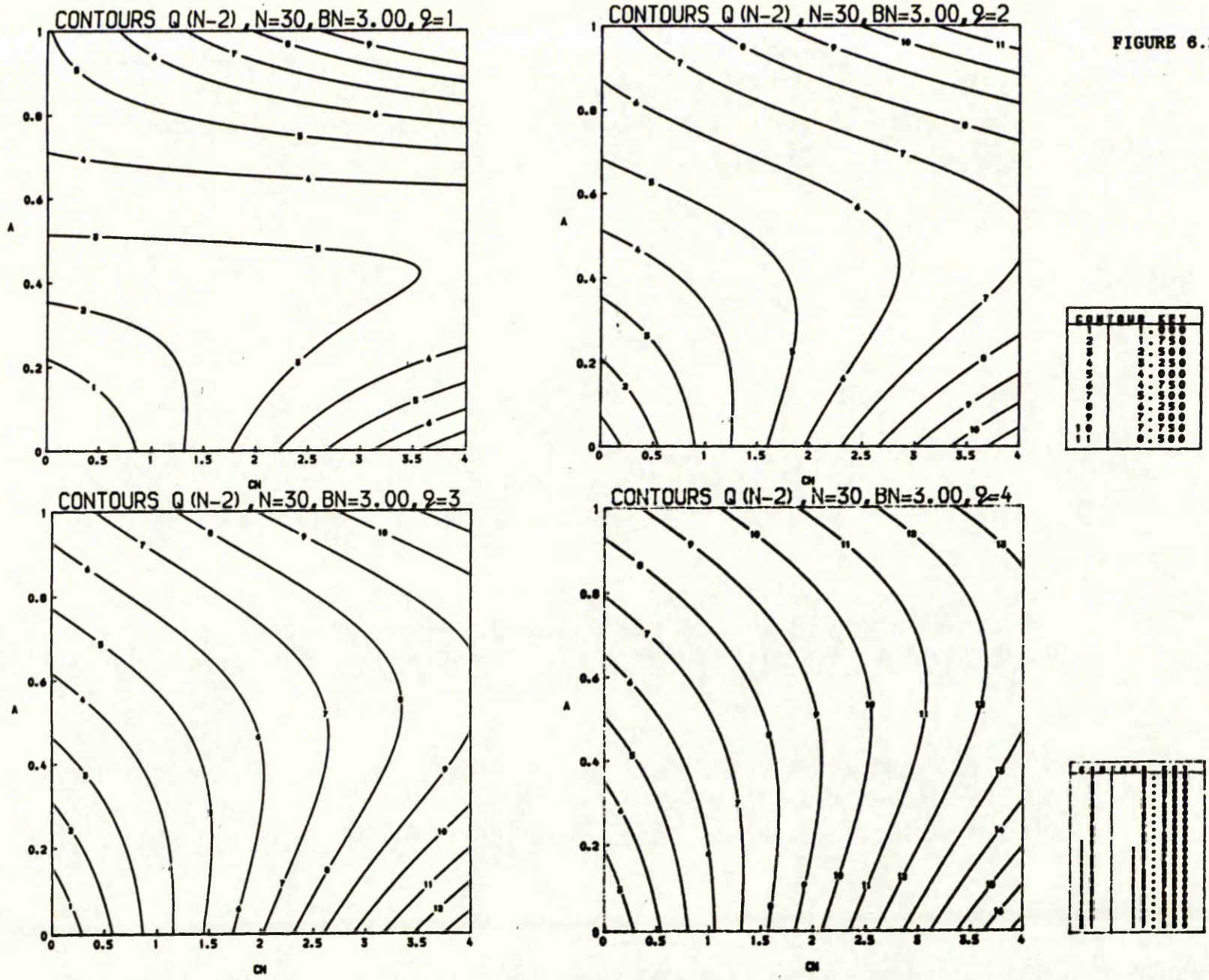

CONTOURS Q (N-2) , N=30, BN=3.00, 9=1
CONTOURS Q (N-2) , N=30, BN=3.00, 9=2
CONTOURS Q (N-2) , N=30, BN=3.00, 9=3
CONTOURS Q (N-2) , N=30, BN=3.00, 9=4
A
CH
CONTOUR KEY

FIGURE 6.28

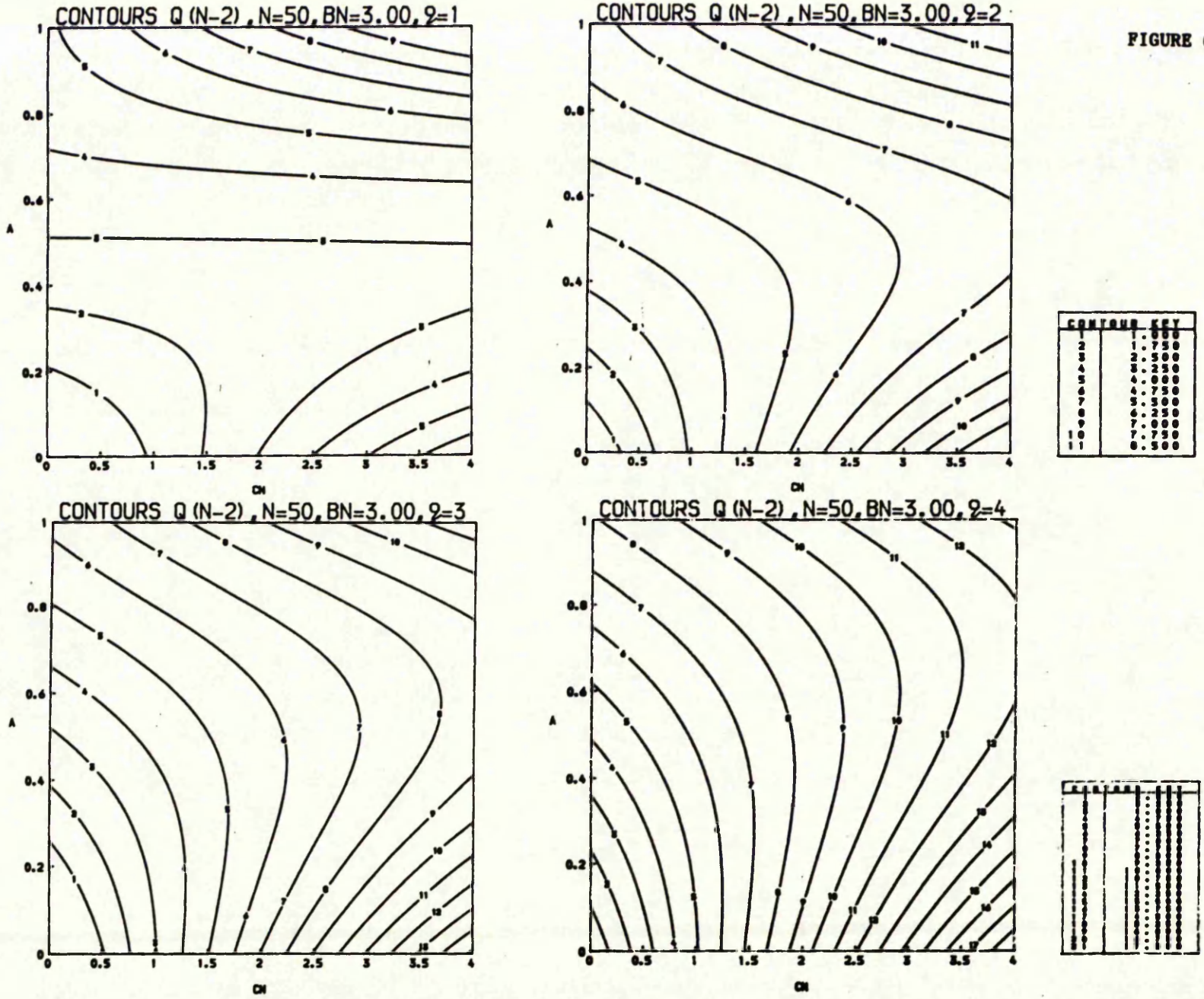
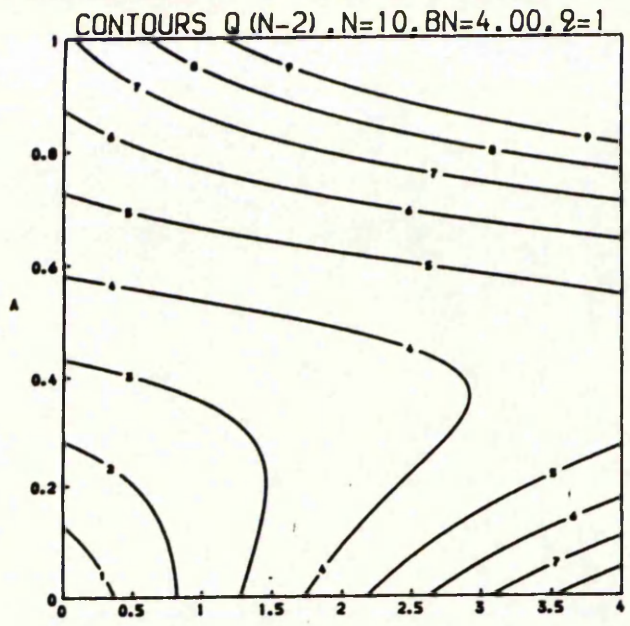
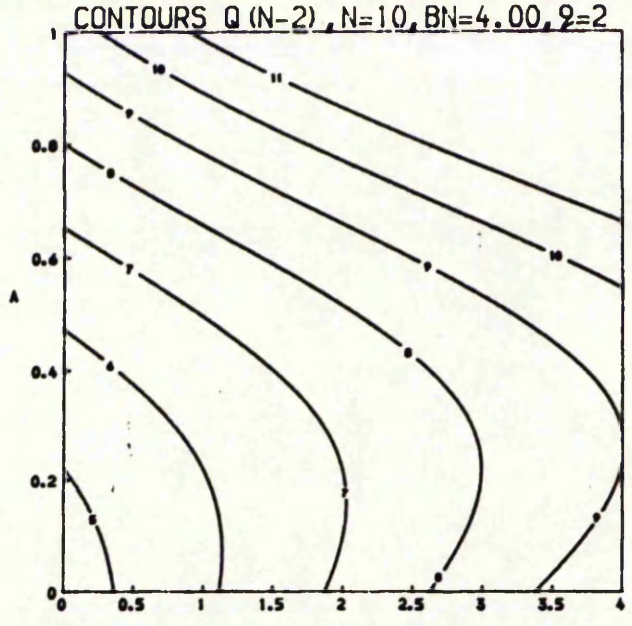
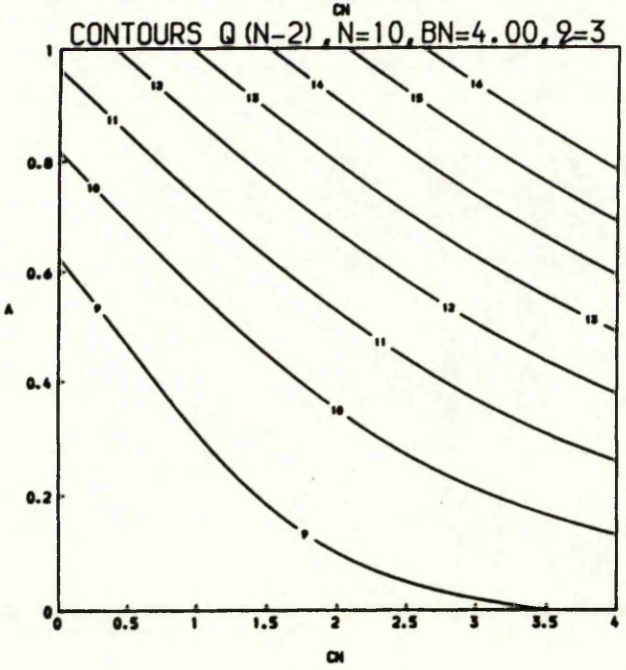
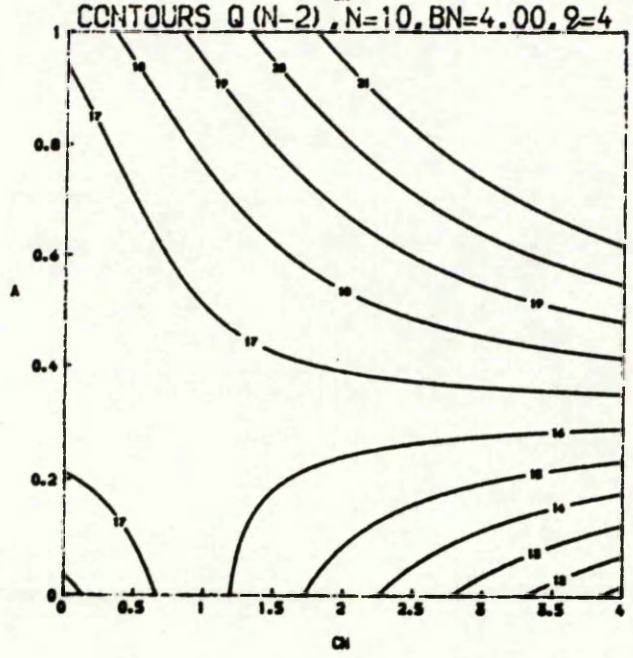

CONTOURS Q(N-2), N=50, BN=3.00, Q=1
CONTOURS Q(N-2), N=50, BN=3.00, Q=2
CONTOURS Q(N-2), N=50, BN=3.00, Q=3
CONTOURS Q(N-2), N=50, BN=3.00, Q=4
A
CN
CONTOUR KEY

FIGURE 6.29


CONTOURS Q(N-2), N=10, BN=4.00, Q=1
A
CN


CONTOURS Q(N-2), N=10, BN=4.00, Q=2
A
CN


CONTOURS Q(N-2), N=10, BN=4.00, Q=3
A
CN


CONTOURS Q(N-2), N=10, BN=4.00, Q=4
A
CN

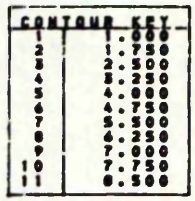
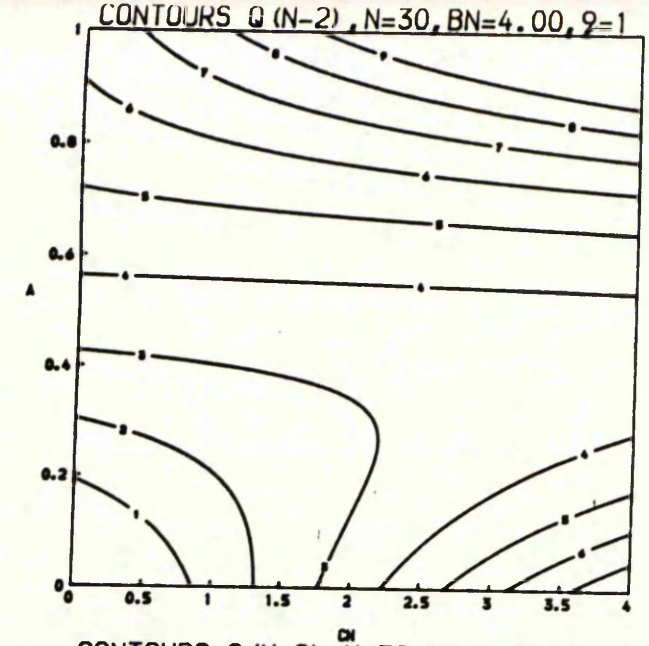

CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500

FIGURE 6.30


CONTOURS Q (N-2) , N=30, BN=4.00, 9=1
A
CN

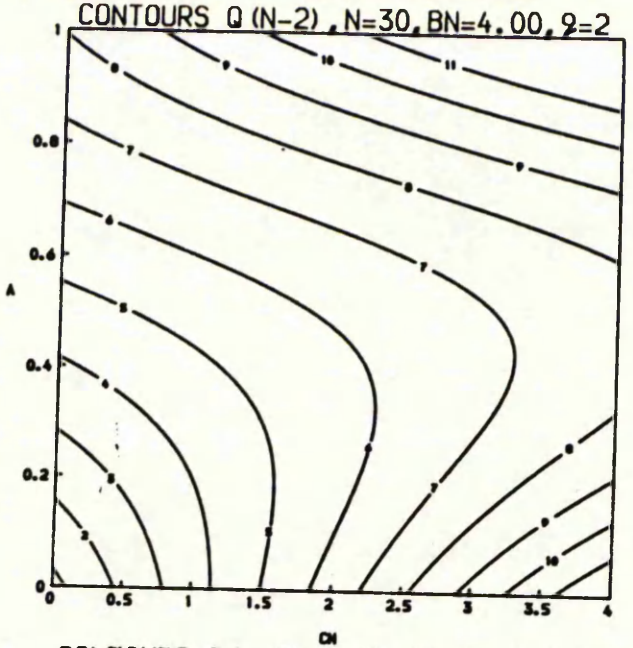
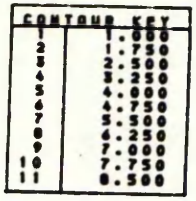
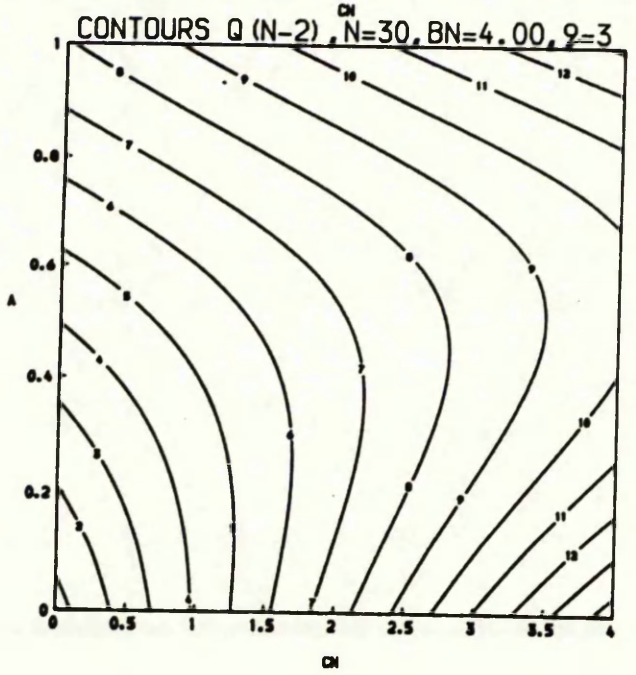
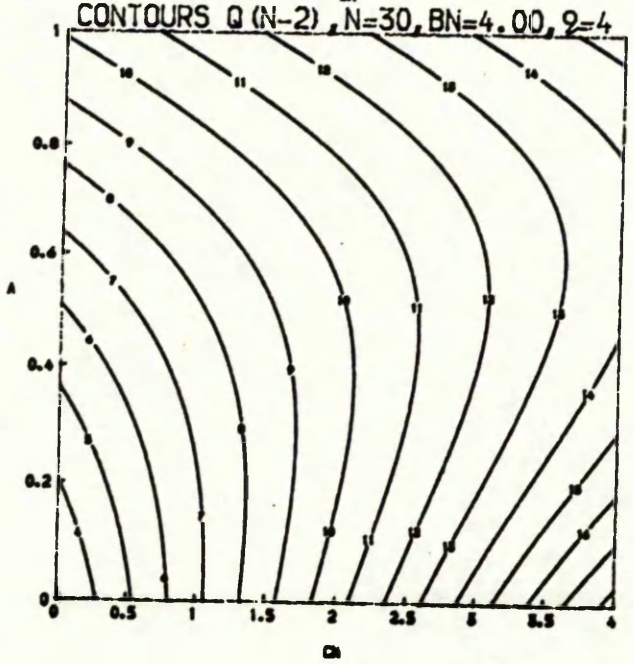
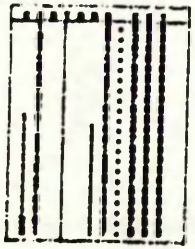

CONTOURS Q (N-2) , N=30, BN=4.00, 9=2
A
CN

FIGURE 6.31


CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500


CONTOURS Q (N-2) , N=30, BN=4.00, 9=3
A
CN


CONTOURS Q (N-2) , N=30, BN=4.00, 9=4
A
CN



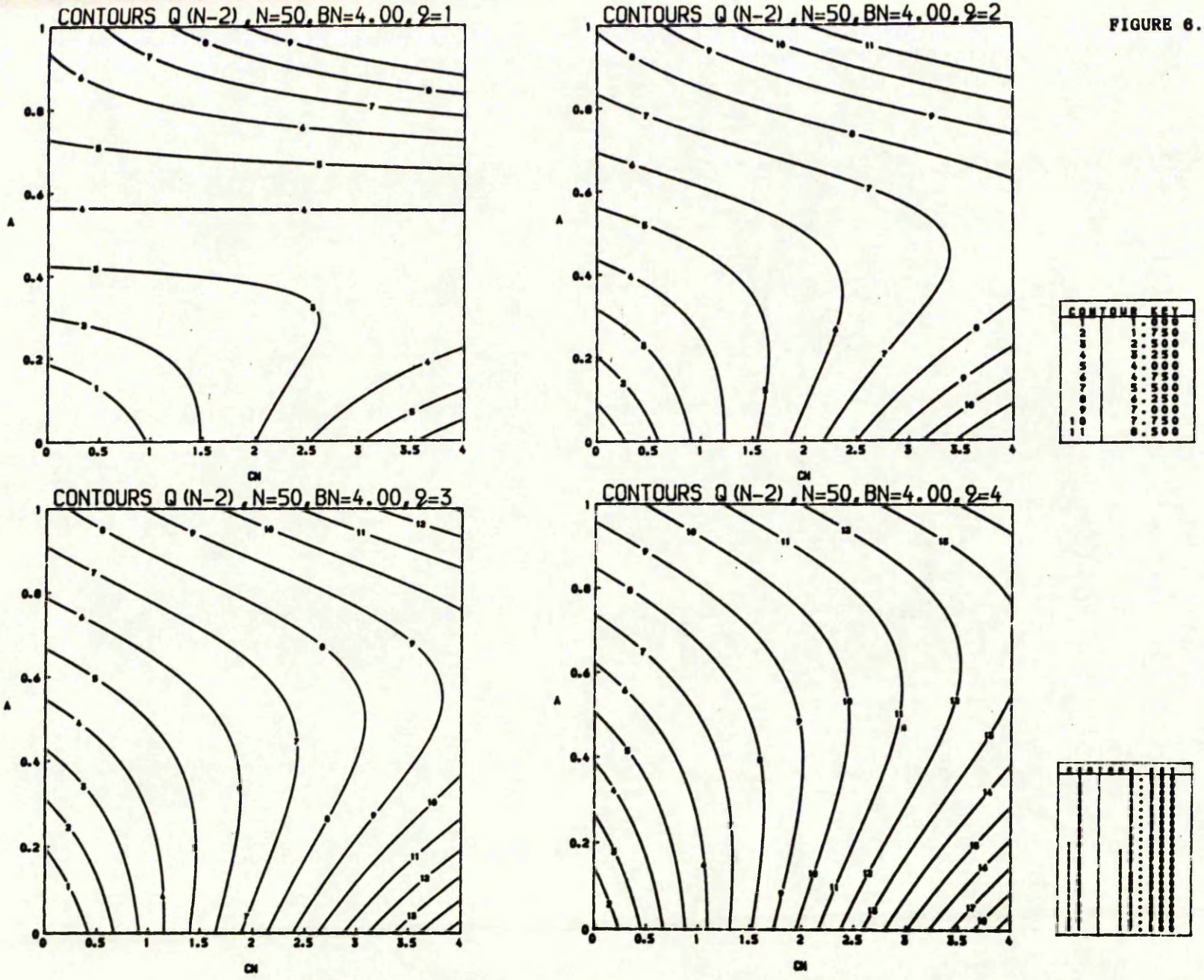
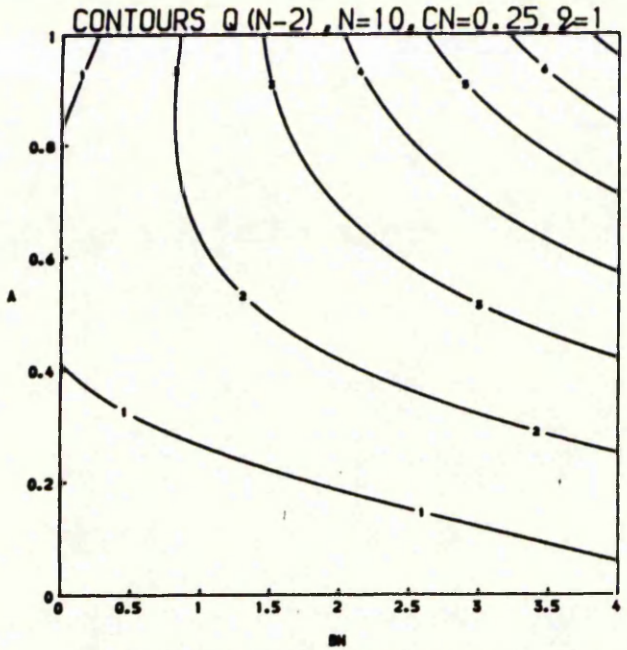

CONTOURS Q(N-2), N=50, BN=4.00, Ω=1
CONTOURS Q(N-2), N=50, BN=4.00, Ω=2
CONTOURS Q(N-2), N=50, BN=4.00, Ω=3
CONTOURS Q(N-2), N=50, BN=4.00, Ω=4
A
CN
CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500

FIGURE 6.32


CONTOURS Q(N-2), N=10, CN=0.25, Q=1
A
BN

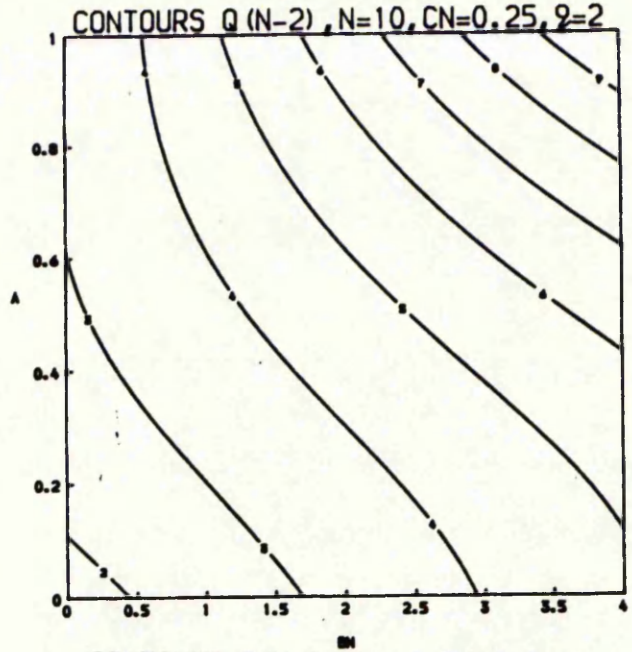
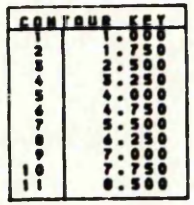
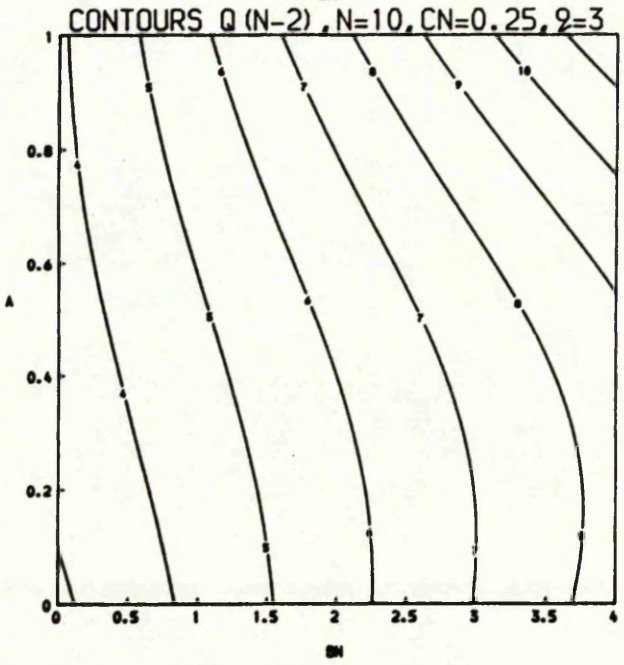
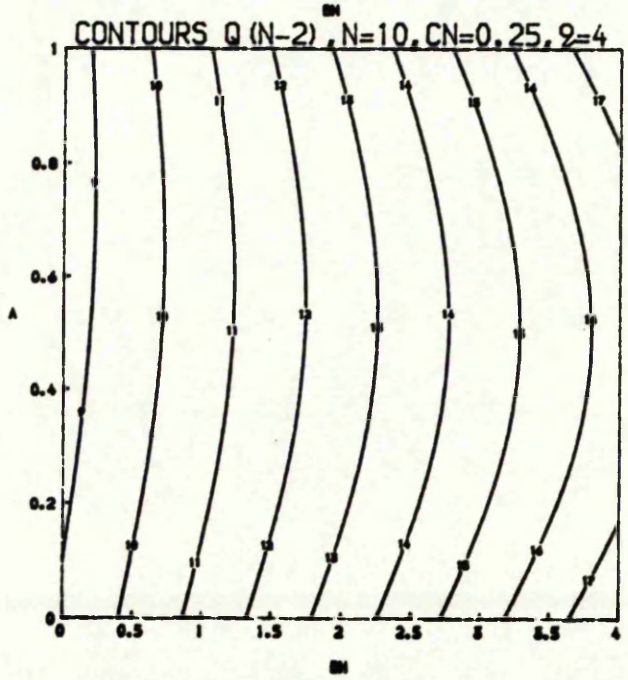
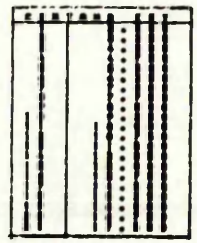

CONTOURS Q(N-2), N=10, CN=0.25, Q=2
A
BN

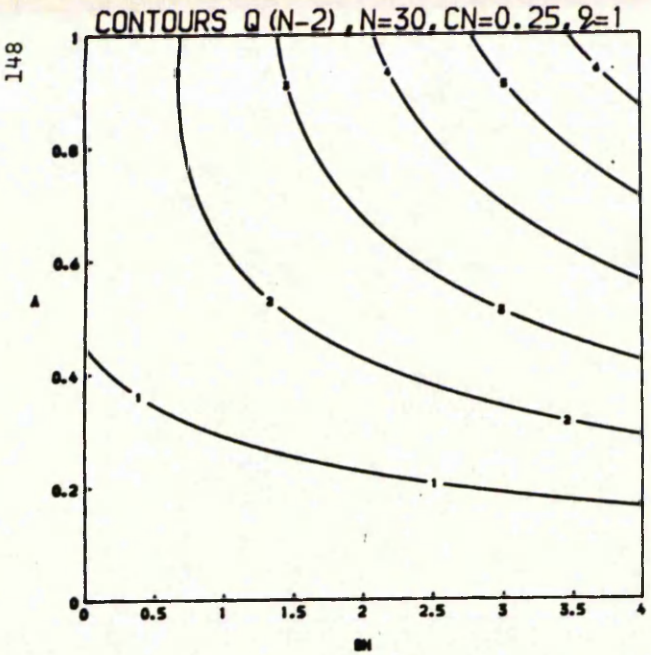
FIGURE 6.33


CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500


CONTOURS Q(N-2), N=10, CN=0.25, Q=3
A
BN


CONTOURS Q(N-2), N=10, CN=0.25, Q=4
A
BN




CONTOURS Q (N-2), N=30, CN=0.25, P=1
A
BM

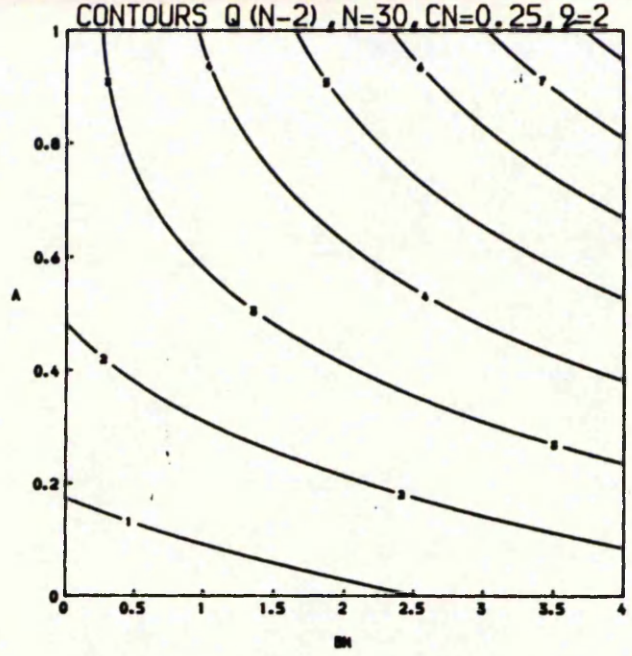
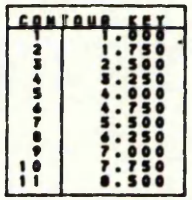
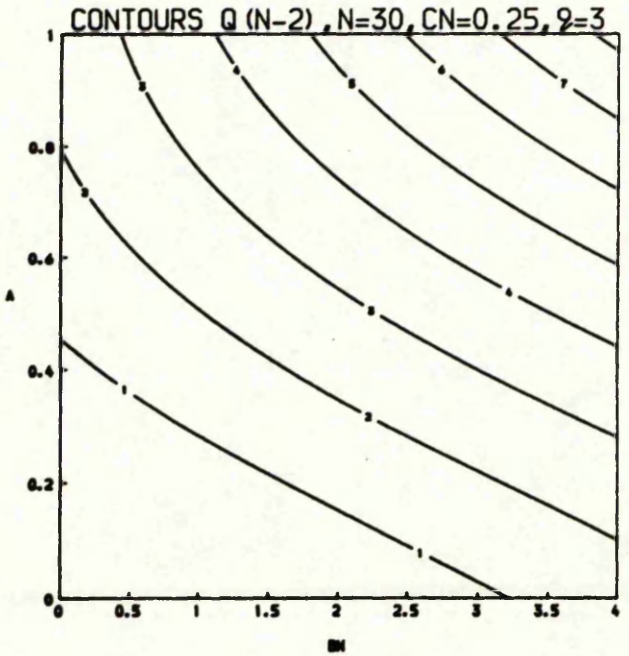
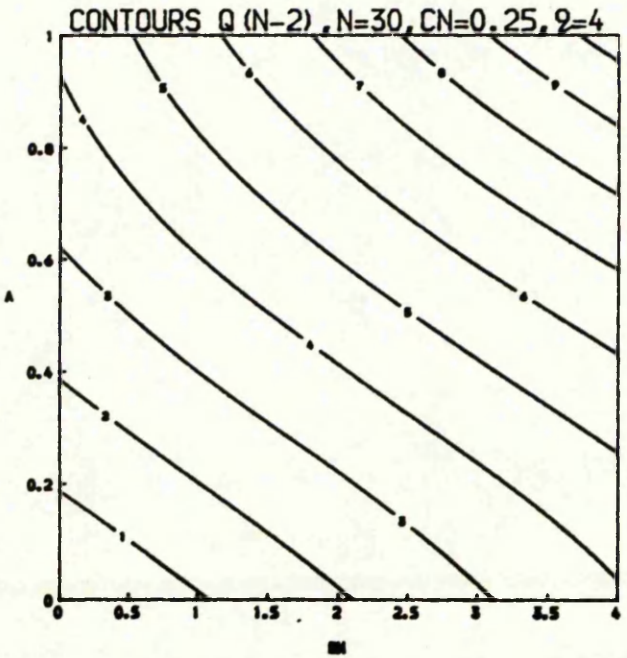
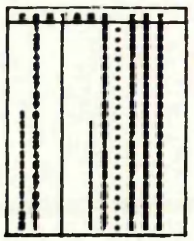

CONTOURS Q (N-2), N=30, CN=0.25, P=2
A
BM

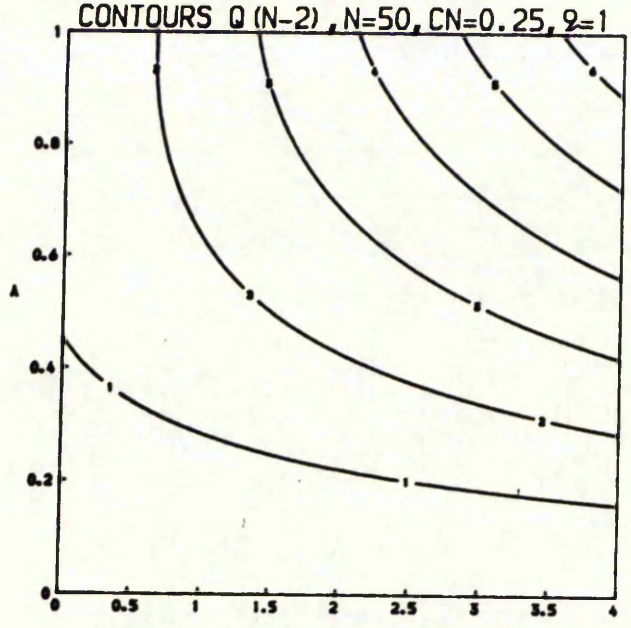
FIGURE 6.34


CONTOUR FEET
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500


CONTOURS Q (N-2), N=30, CN=0.25, P=3
A
BM


CONTOURS Q (N-2), N=30, CN=0.25, P=4
A
BM





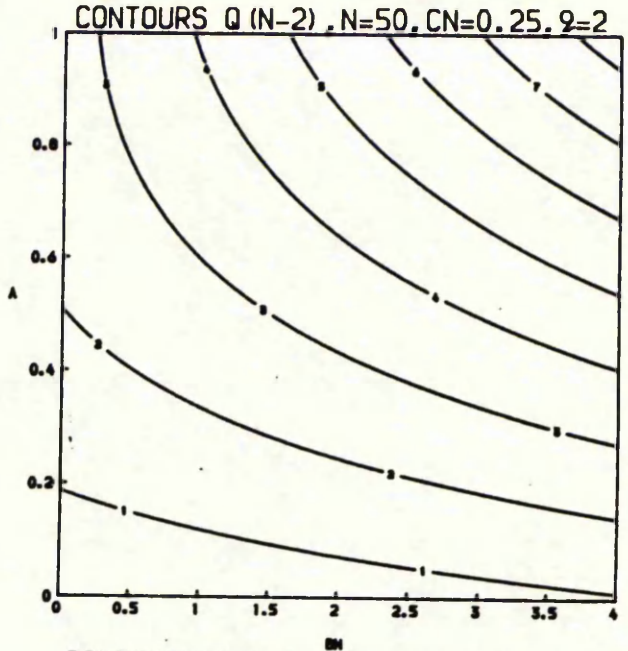
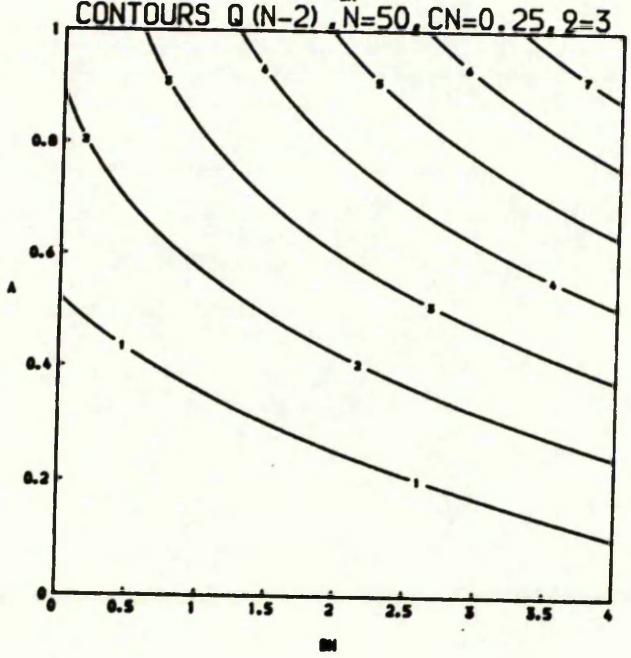
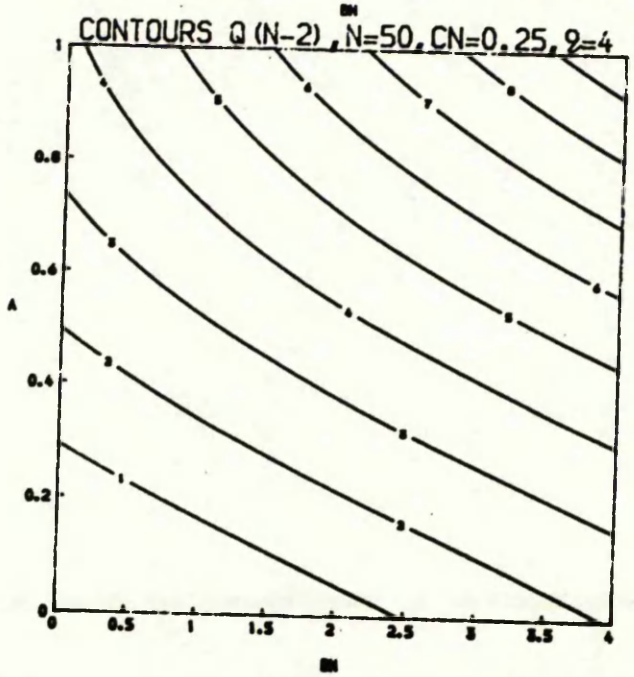
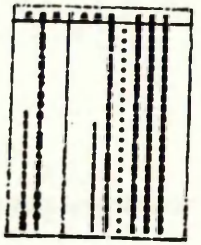


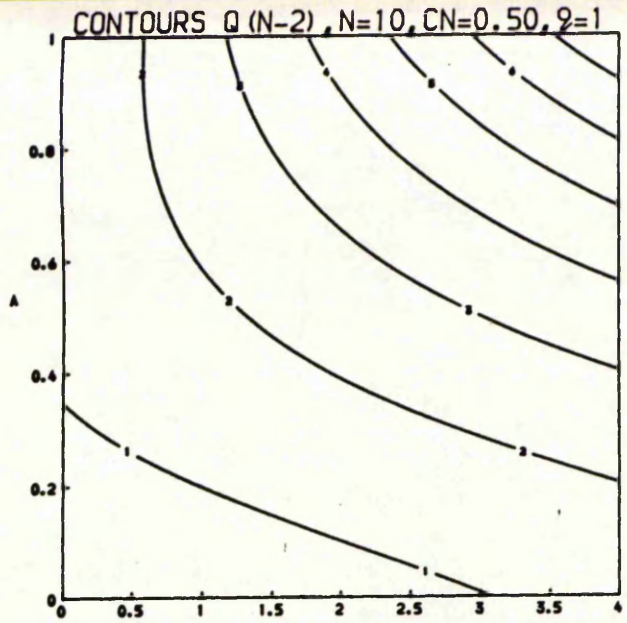
FIGURE 6.35

CONTOUR KEY	
1	1.000
2	1.750
3	2.500
4	3.250
5	4.000
6	4.750
7	5.500
8	6.250
9	7.000
10	7.750
11	8.500









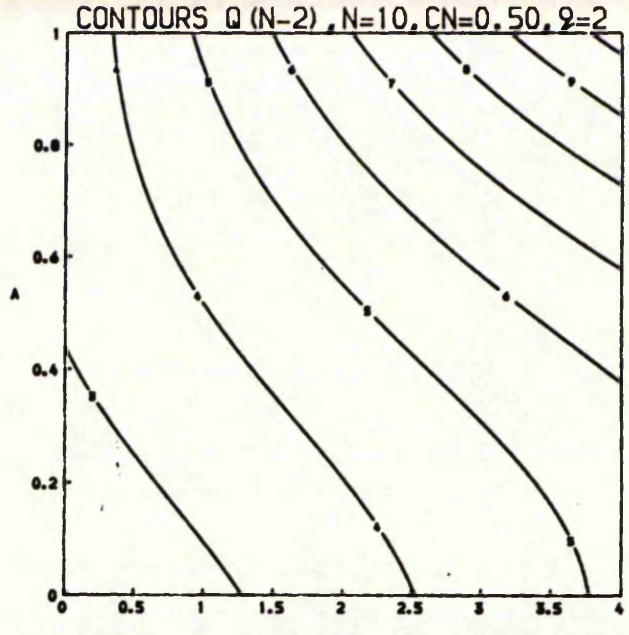
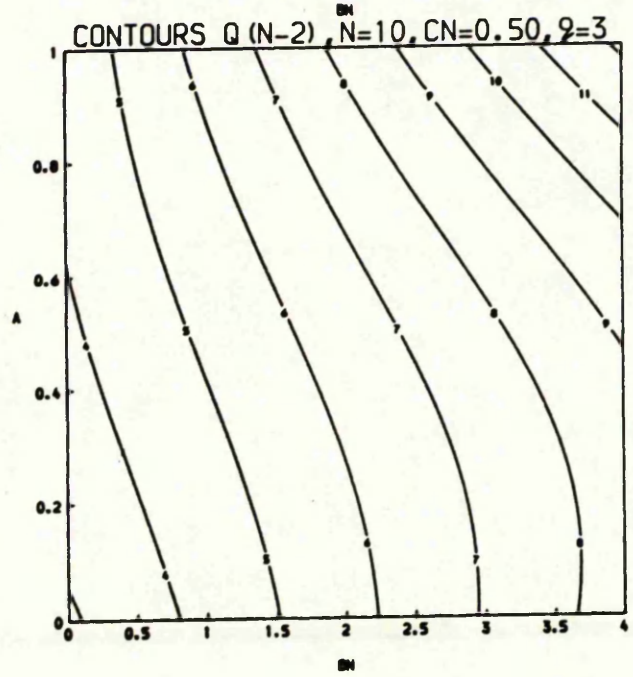
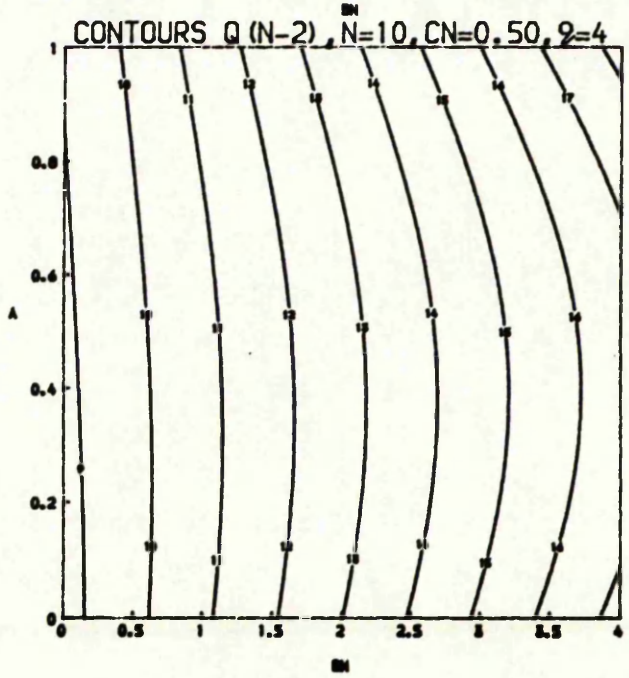
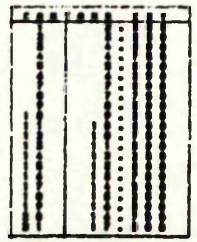


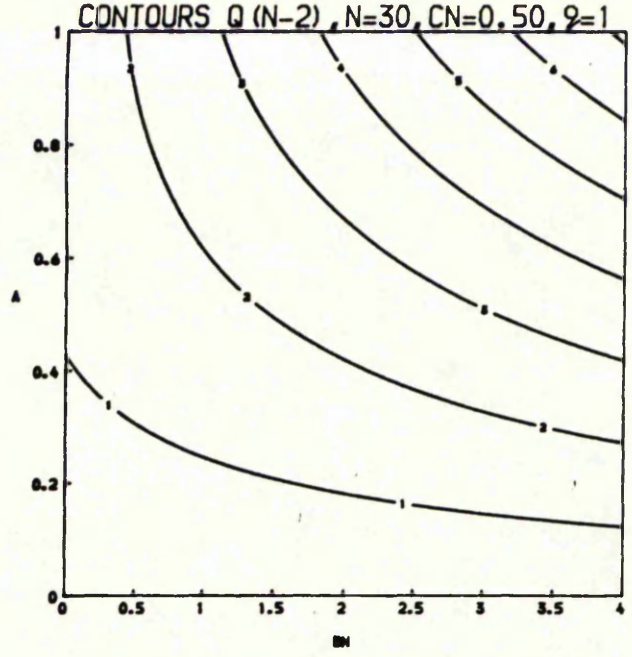
FIGURE 6.36

CONTOUR	KEY
1	1.000
2	1.750
3	2.500
4	3.250
5	4.000
6	4.750
7	5.500
8	6.250
9	7.000
10	7.750
11	8.500








CONTOURS Q(N-2), N=30, CN=0.50, 9=1
A
BN

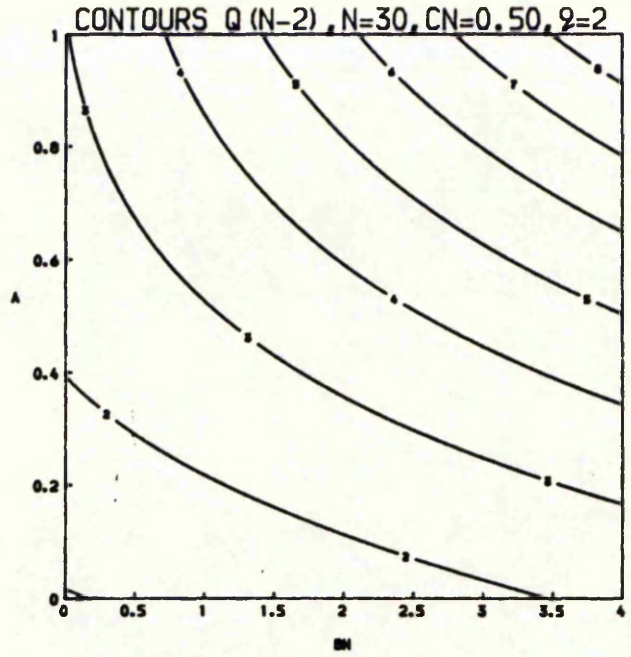
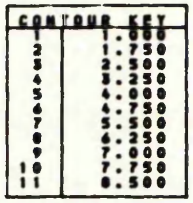
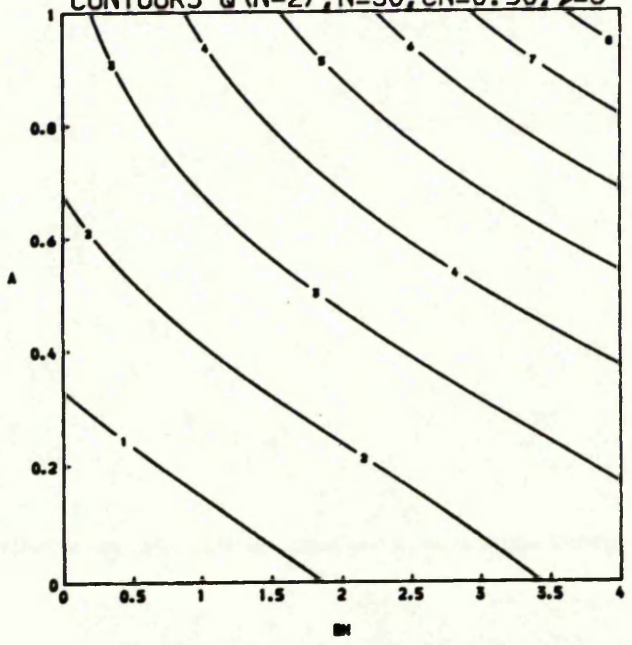
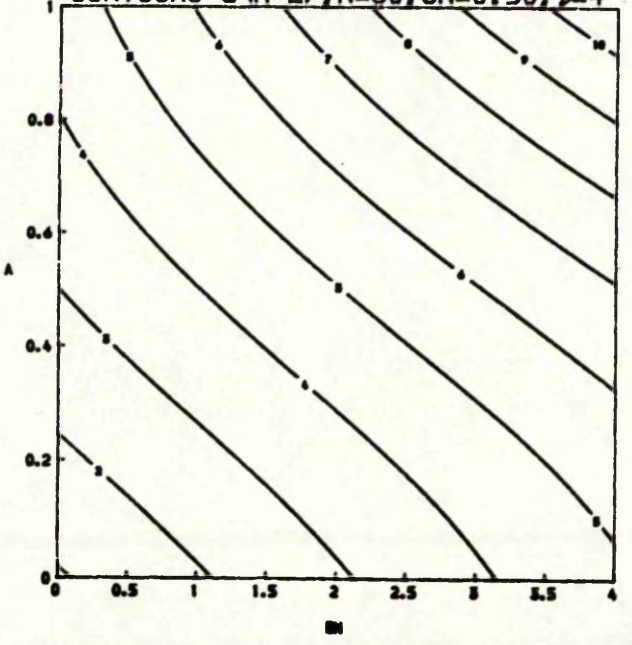
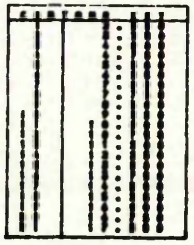

CONTOURS Q(N-2), N=30, CN=0.50, 9=2
A
BN

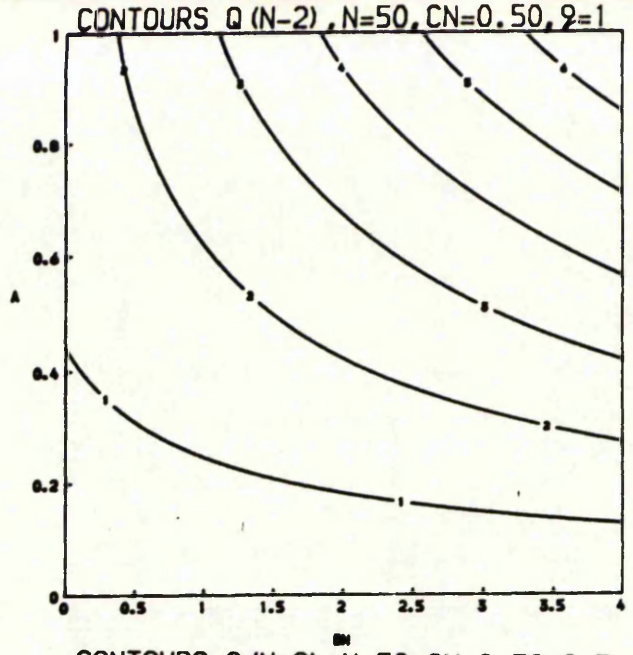
FIGURE 6.37


CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500


CONTOURS Q(N-2), N=30, CN=0.50, 9=3
A
BN


CONTOURS Q(N-2), N=30, CN=0.50, 9=4
A
BN




CONTOURS Q (N-2) , N=50, CN=0.50, Q=1
A
BN

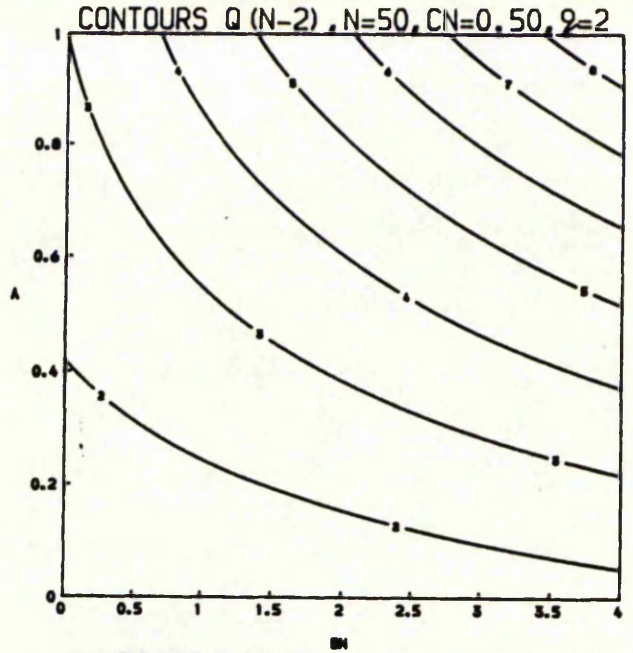
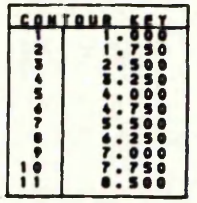
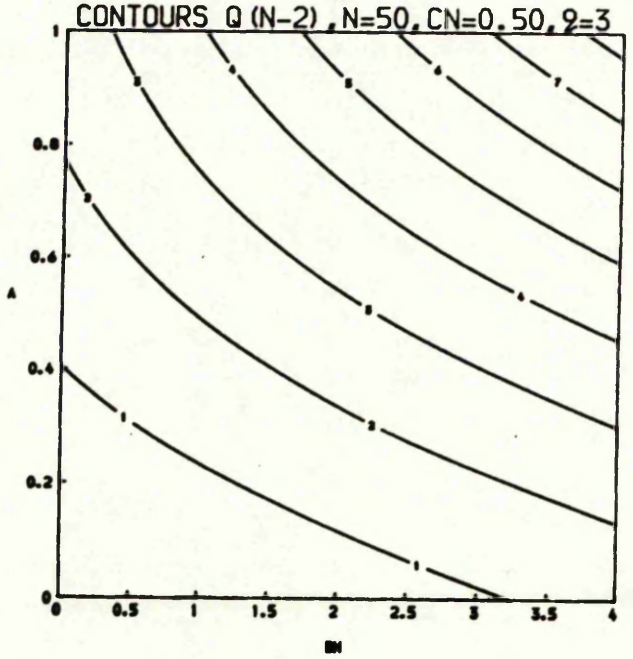
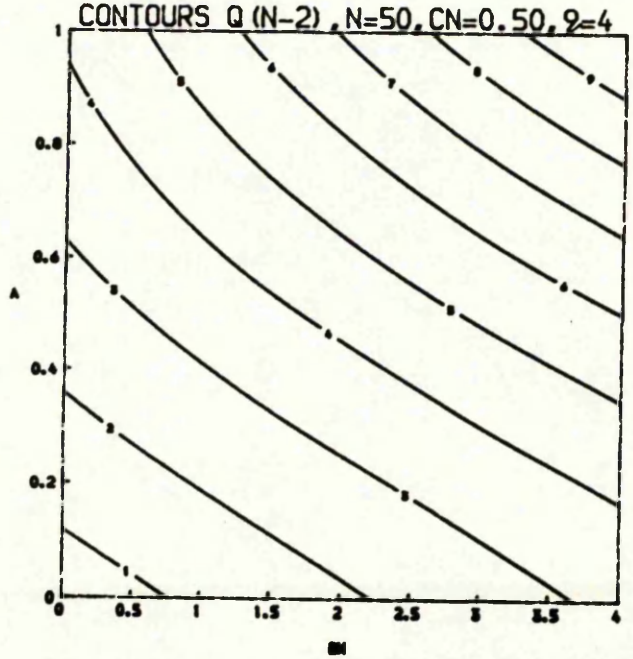
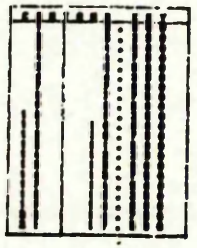

CONTOURS Q (N-2) , N=50, CN=0.50, Q=2
A
BN

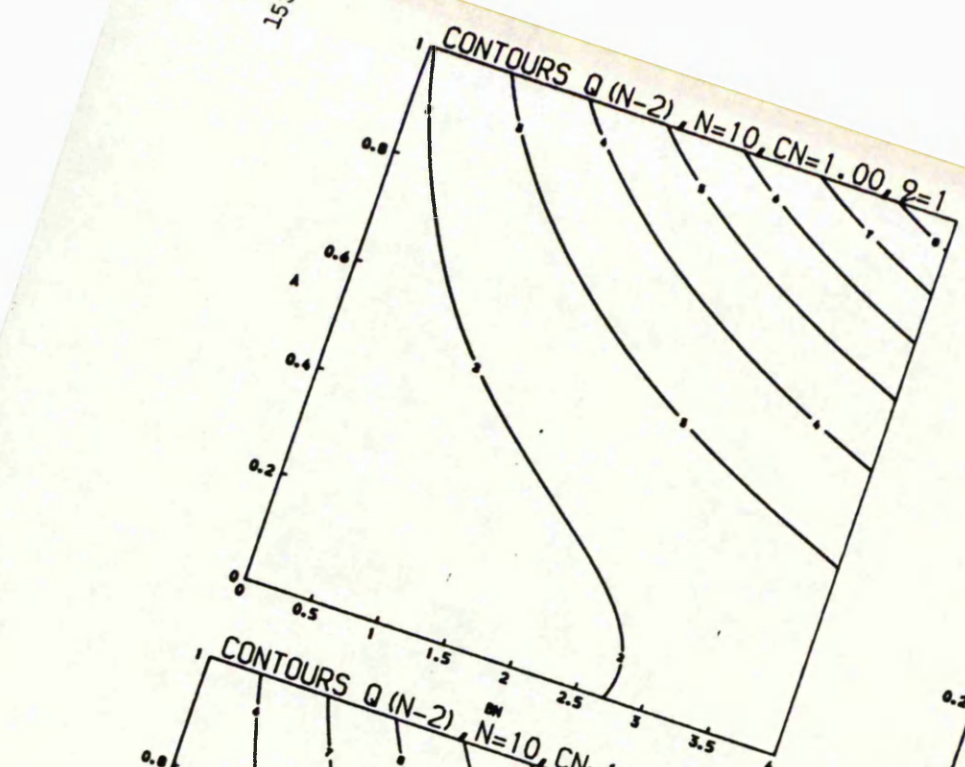
FIGURE 6.38


CONTOUR SET
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500


CONTOURS Q (N-2) , N=50, CN=0.50, Q=3
A
BN


CONTOURS Q (N-2) , N=50, CN=0.50, Q=4
A
BN




CONTOURS Q (N-2), N=10, CN=1.00, Q=1
BN

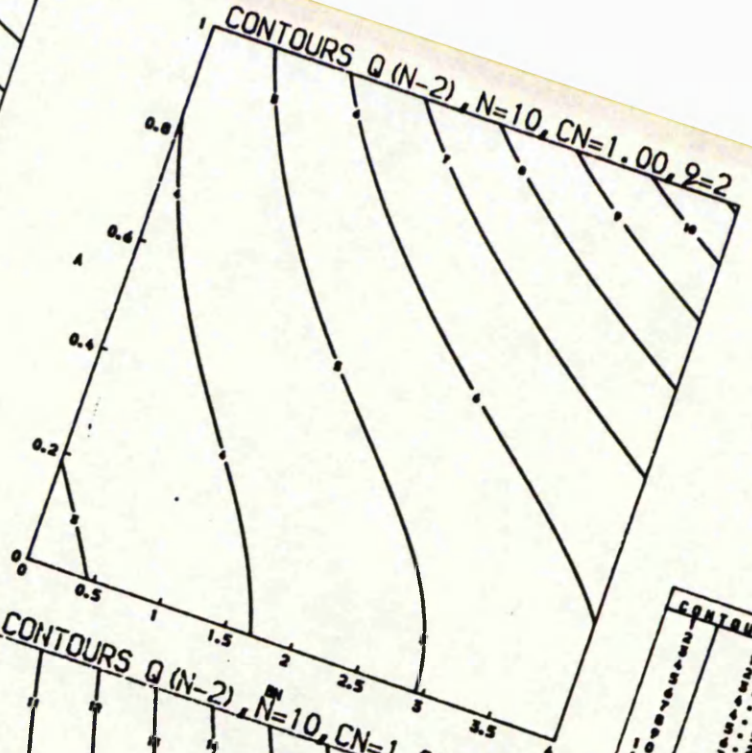
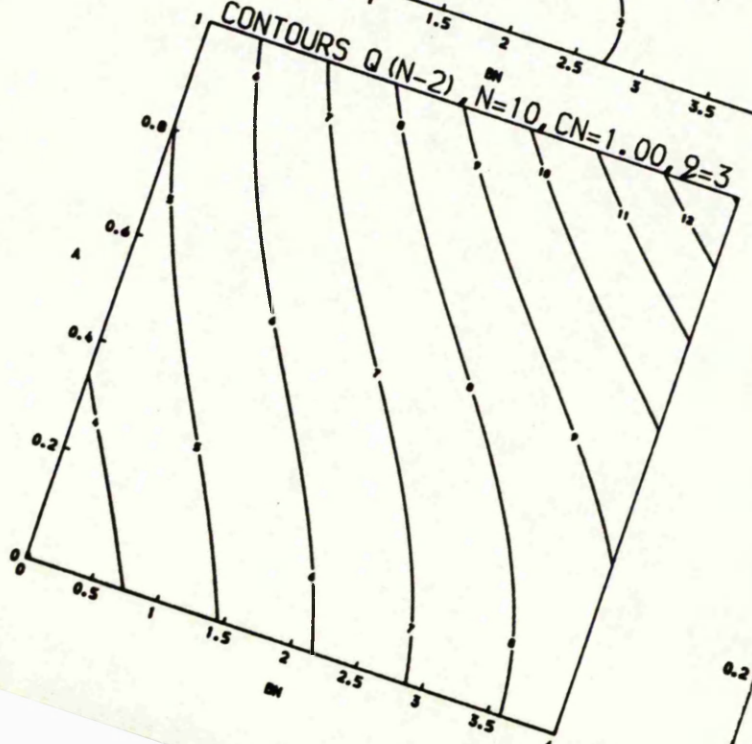
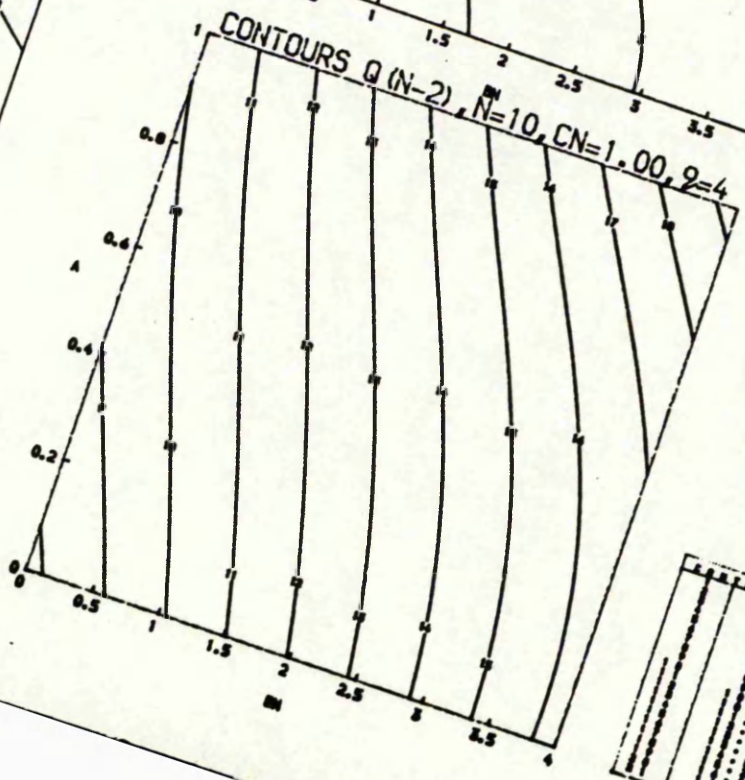
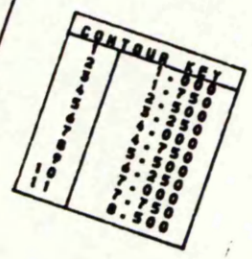
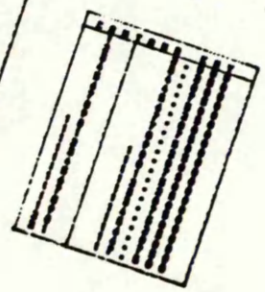

CONTOURS Q (N-2), N=10, CN=1.00, Q=2
BN

FIGURE 6.39


CONTOURS Q (N-2), N=10, CN=1.00, Q=3
BN


CONTOURS Q (N-2), N=10, CN=1.00, Q=4
BN


CONTOURS FEET



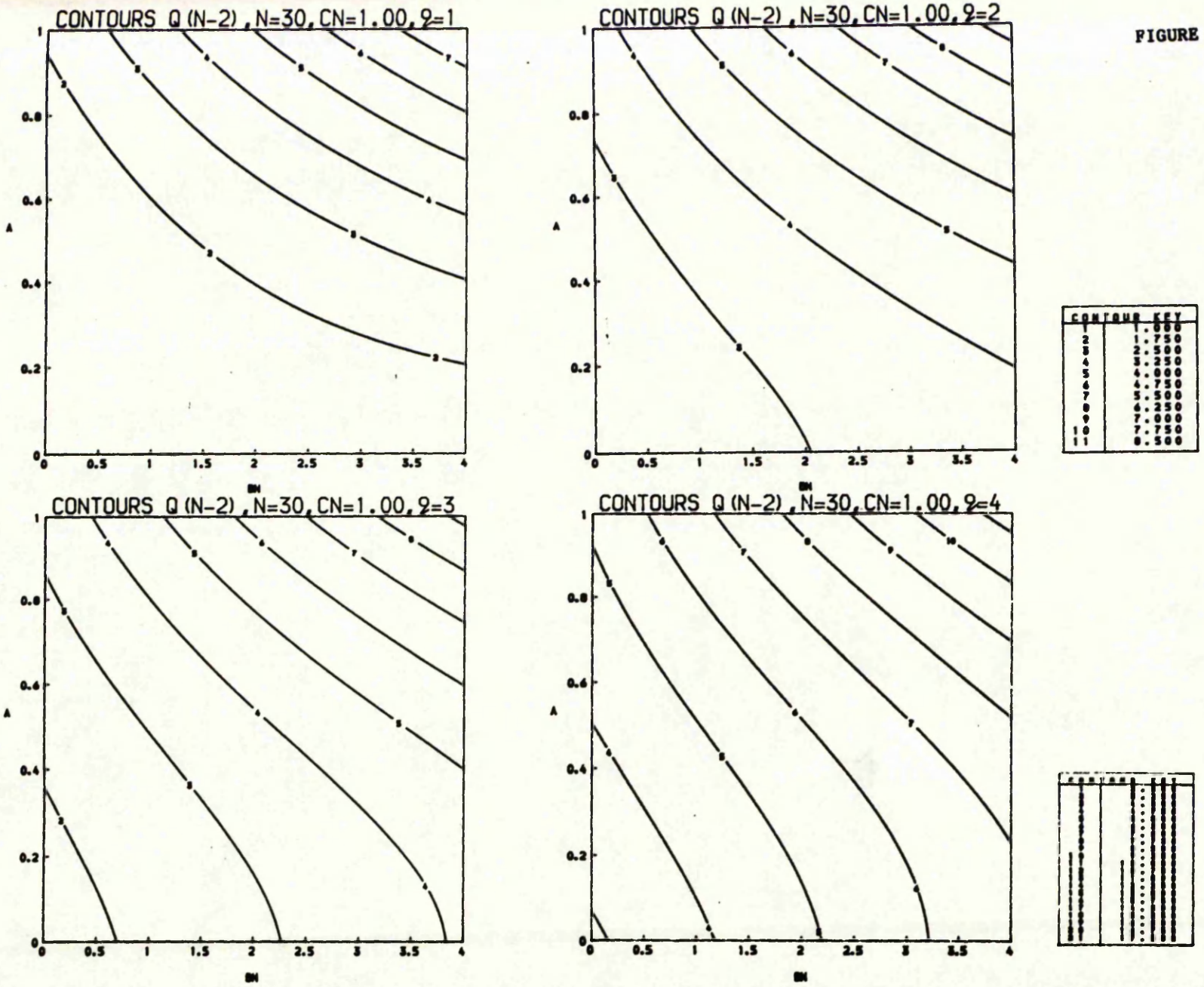

CONTOURS Q(N-2), N=30, CN=1.00, Q=1
CONTOURS Q(N-2), N=30, CN=1.00, Q=2
CONTOURS Q(N-2), N=30, CN=1.00, Q=3
CONTOURS Q(N-2), N=30, CN=1.00, Q=4
A
BN
CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500

FIGURE 6.40

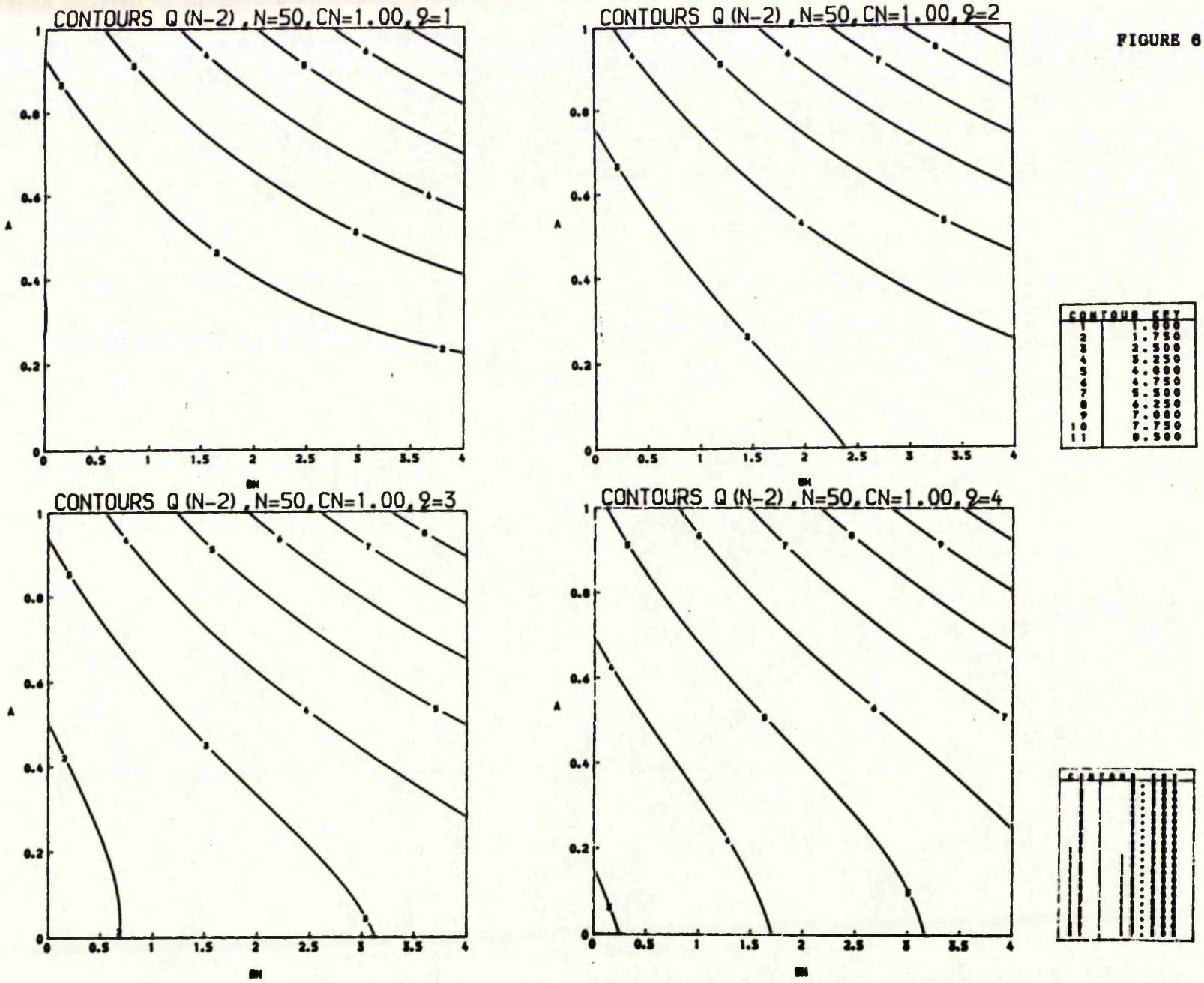
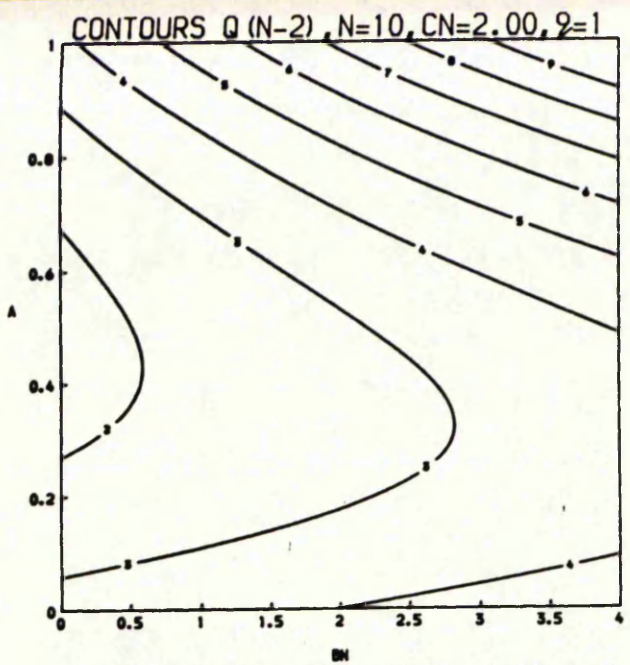

CONTOURS Q (N-2) , N=50, CN=1.00, Q=1
CONTOURS Q (N-2) , N=50, CN=1.00, Q=2
CONTOURS Q (N-2) , N=50, CN=1.00, Q=3
CONTOURS Q (N-2) , N=50, CN=1.00, Q=4
A
BN
CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500

FIGURE 6.41


CONTOURS Q(N-2), N=10, CN=2.00, Q=1
A
BN

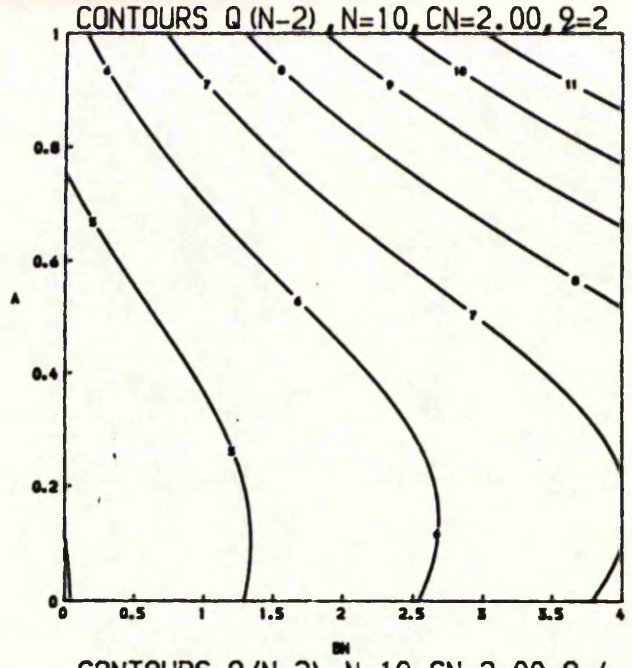
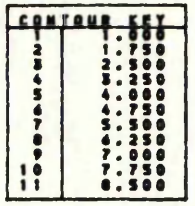
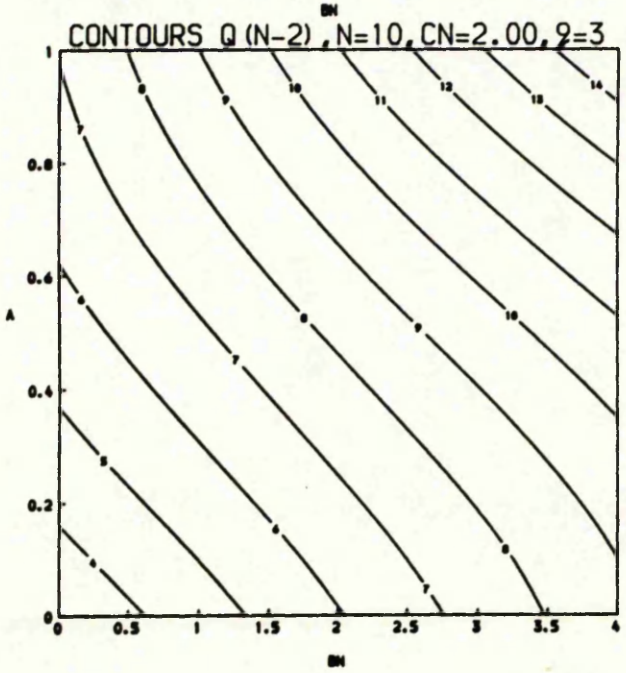
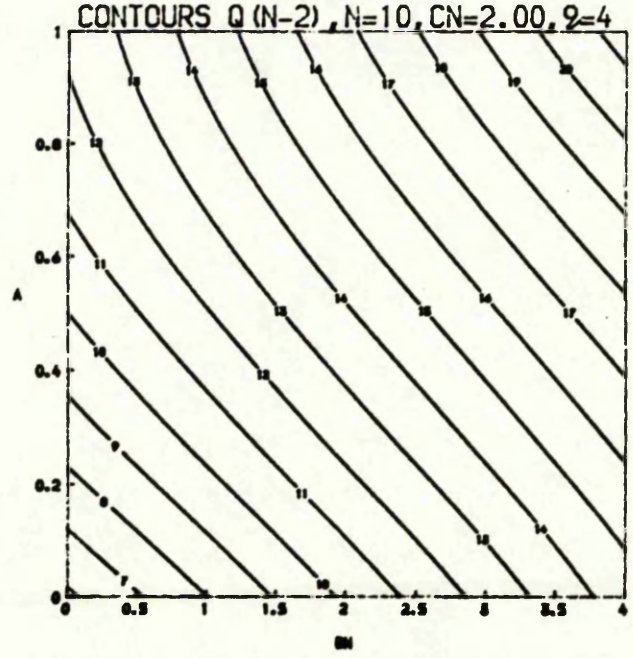
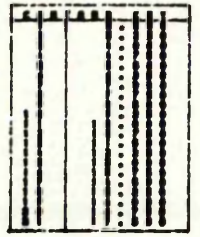

CONTOURS Q(N-2), N=10, CN=2.00, Q=2
A
BN

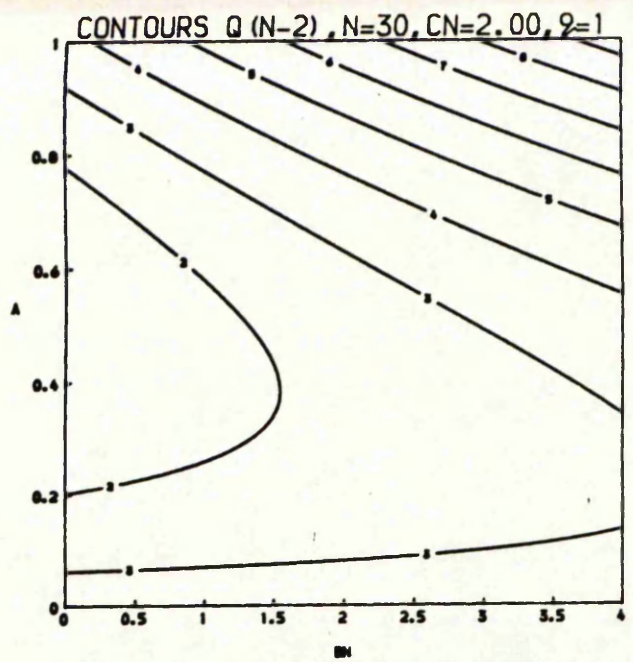
FIGURE 6.42


CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500


CONTOURS Q(N-2), N=10, CN=2.00, Q=3
A
BN


CONTOURS Q(N-2), N=10, CN=2.00, Q=4
A
BN





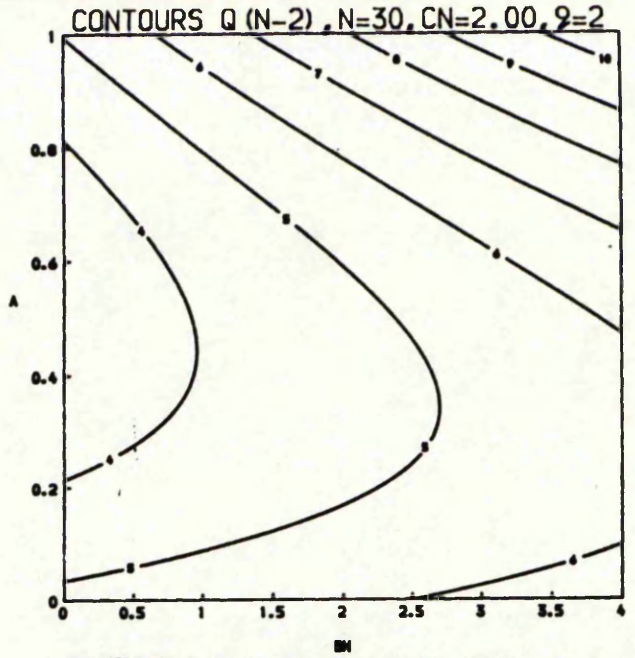
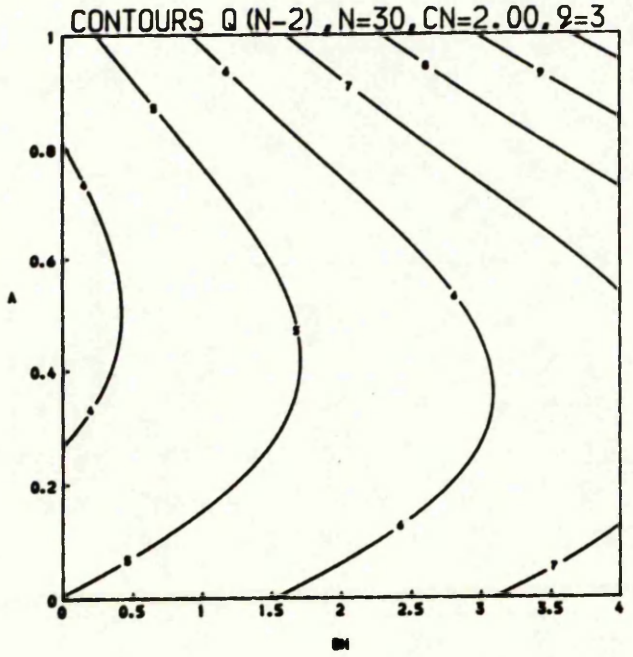
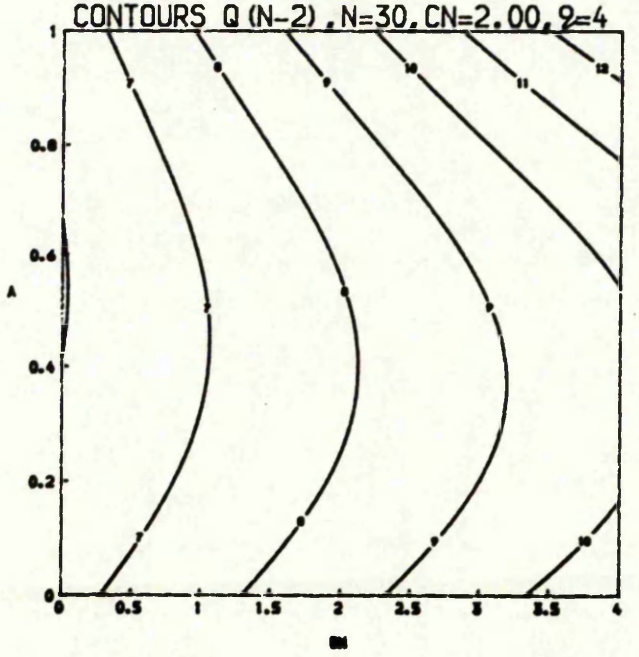
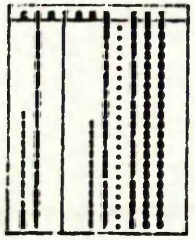


FIGURE 6.43

CONTOUR	KEY
1	1.000
2	1.750
3	2.500
4	3.250
5	4.000
6	4.750
7	5.500
8	6.250
9	7.000
10	7.750
11	8.500







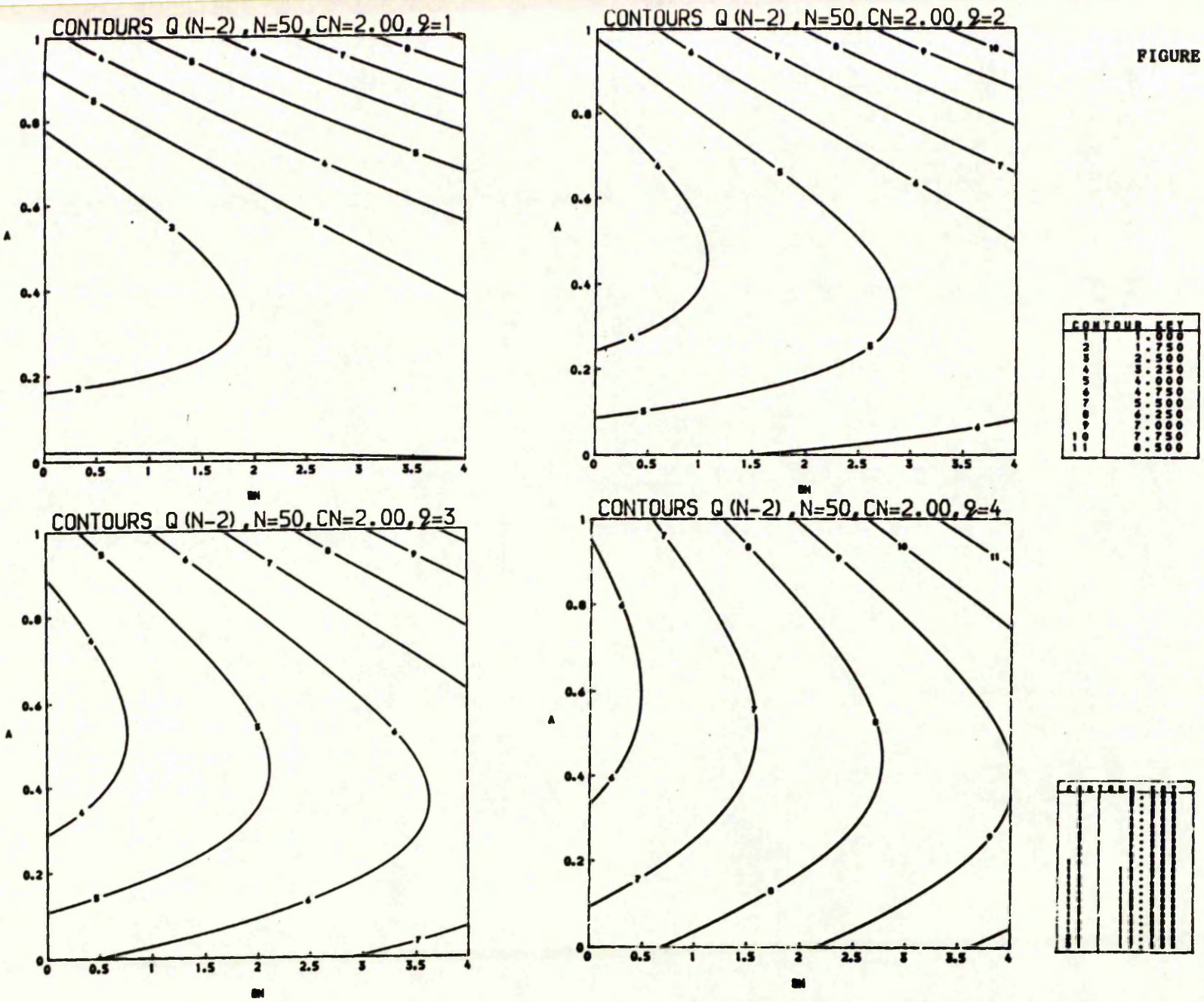
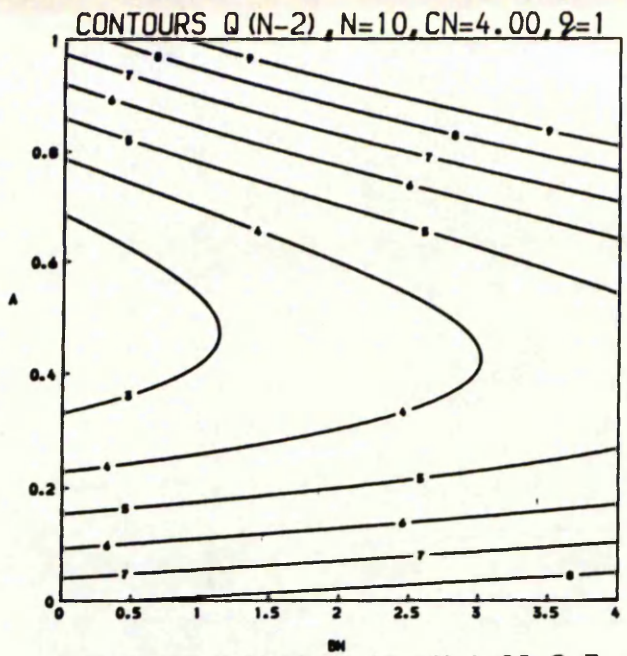

CONTOURS Q(N-2), N=50, CN=2.00, Q=1
CONTOURS Q(N-2), N=50, CN=2.00, Q=2
CONTOURS Q(N-2), N=50, CN=2.00, Q=3
CONTOURS Q(N-2), N=50, CN=2.00, Q=4
A
BN
CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500

FIGURE 6.44


CONTOURS Q (N-2) , N=10, CN=4.00, Q=1
A
BM

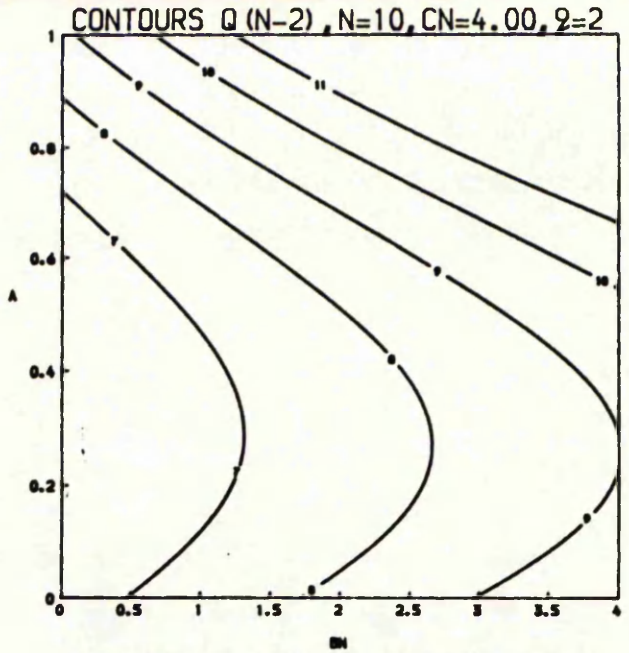
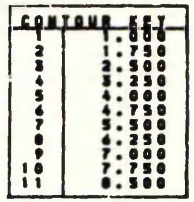
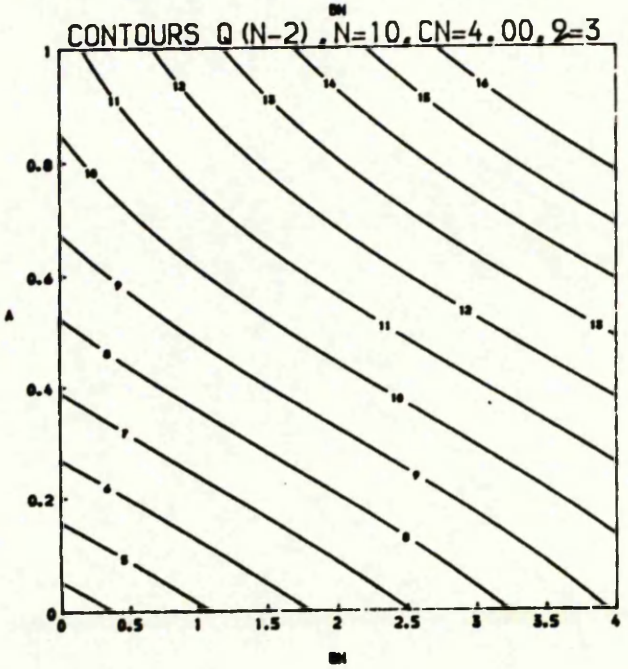
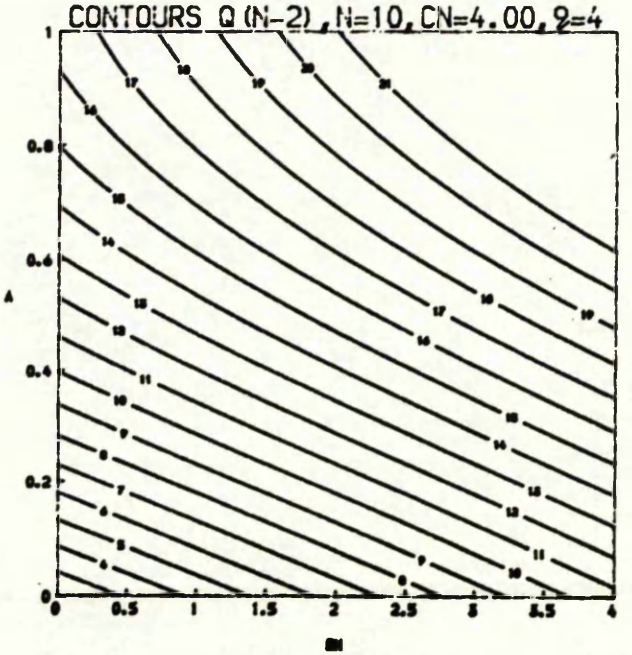
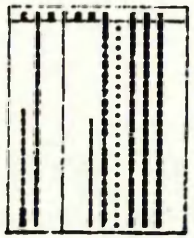

CONTOURS Q (N-2) , N=10, CN=4.00, Q=2
A
BM

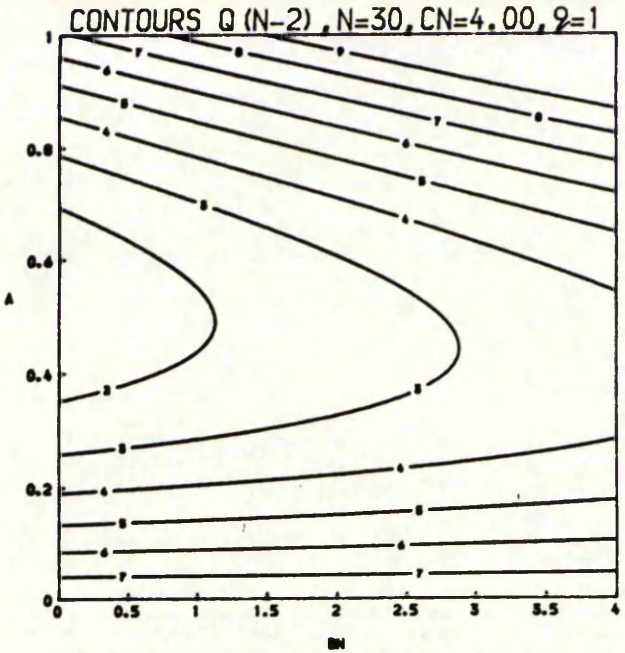
FIGURE 6.45


CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500


CONTOURS Q (N-2) , N=10, CN=4.00, Q=3
A
BM


CONTOURS Q (N-2) , N=10, CN=4.00, Q=4
A
BM





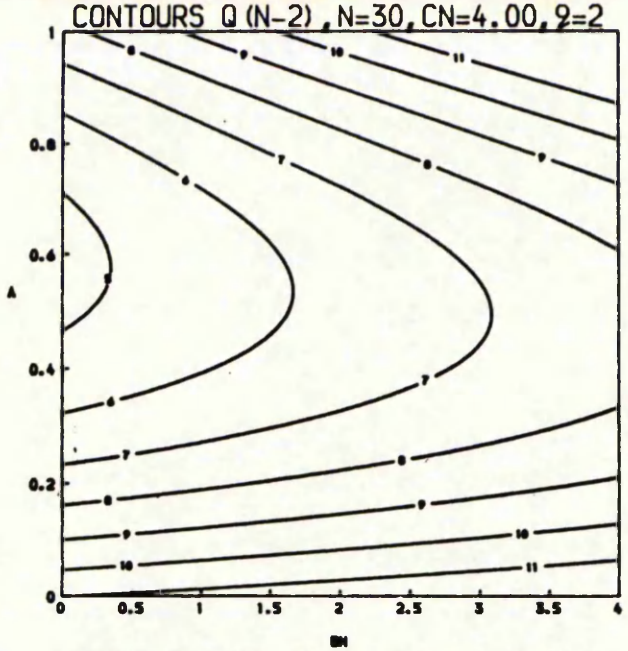
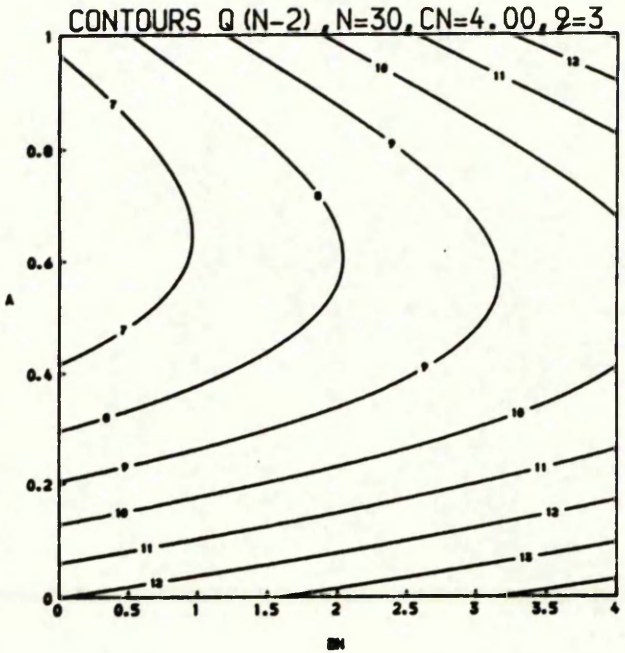
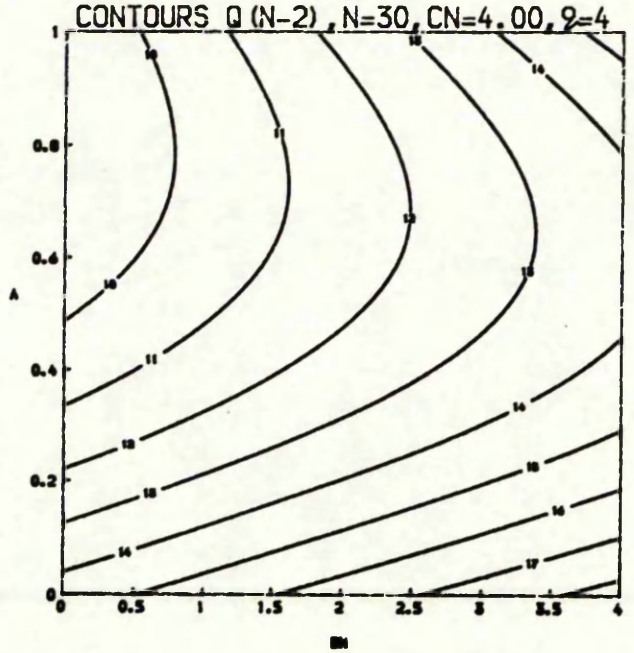
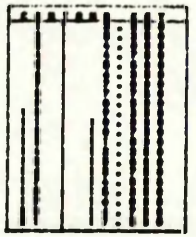


FIGURE 6.46

CONTOUR	KEY
1	1.000
2	1.750
3	2.500
4	3.250
5	4.000
6	4.750
7	5.500
8	6.250
9	7.000
10	7.750
11	8.500







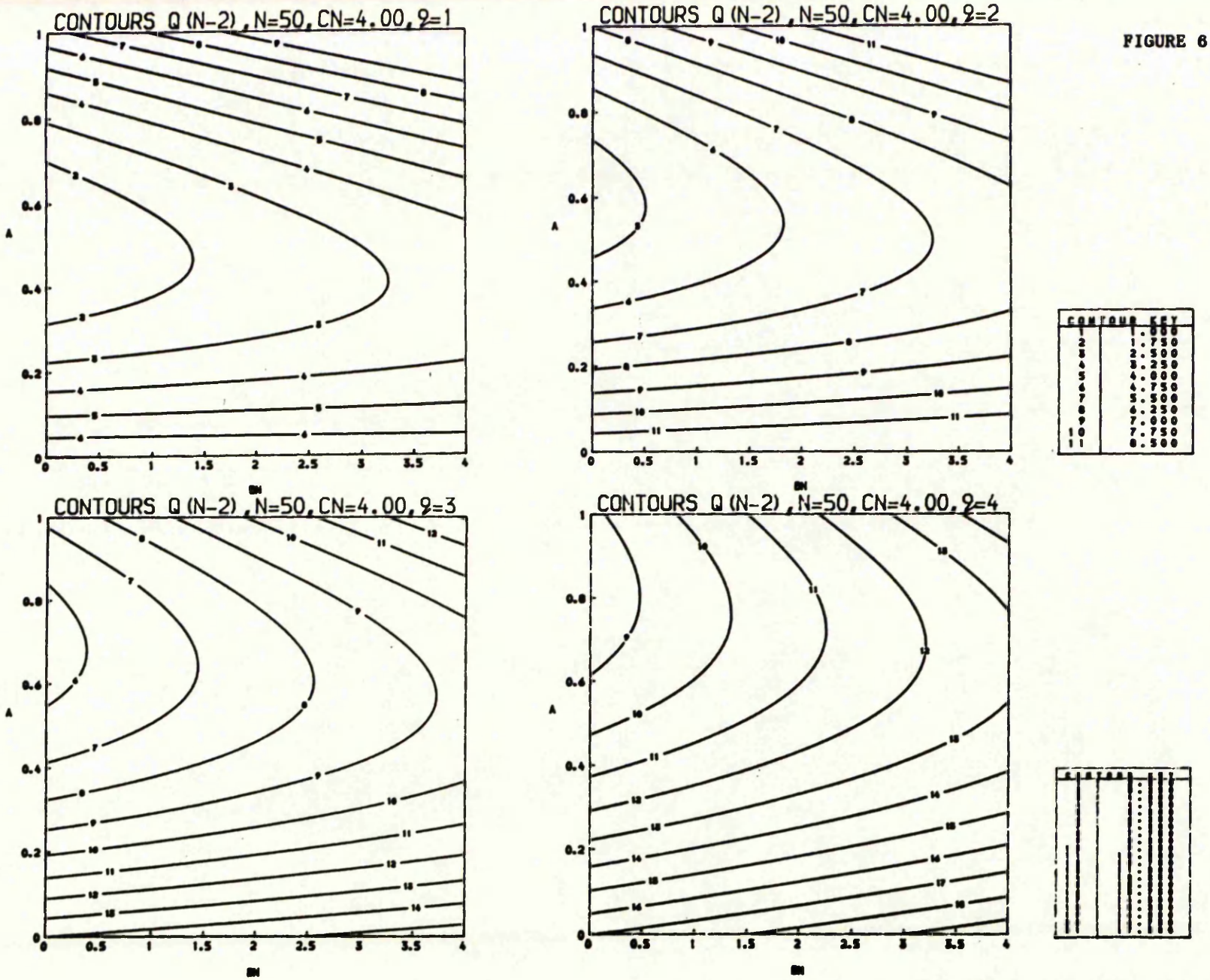
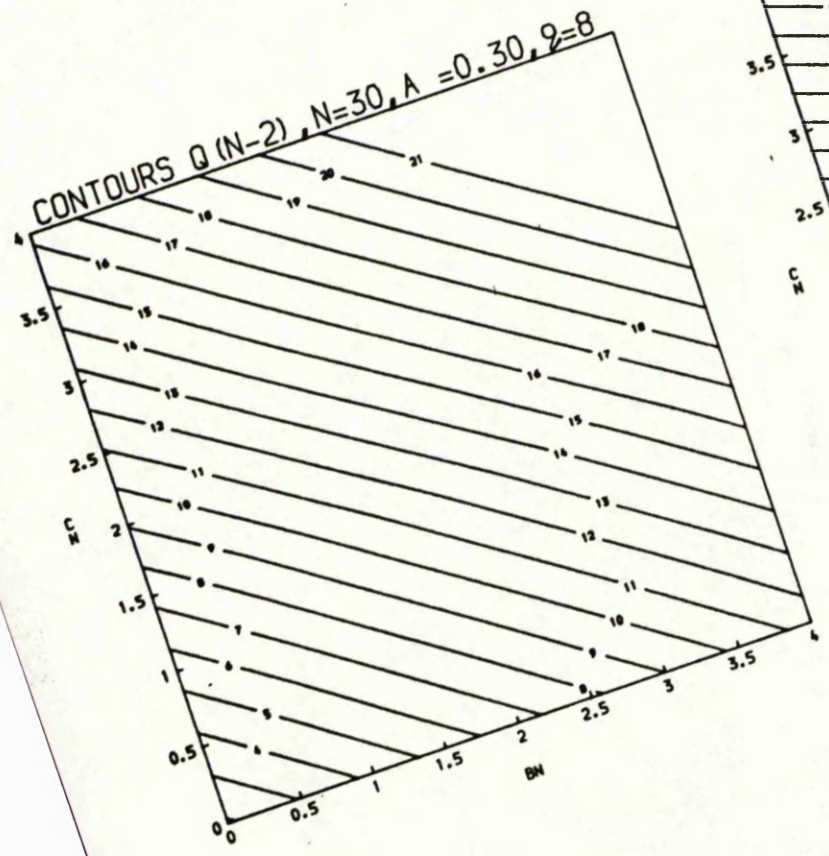
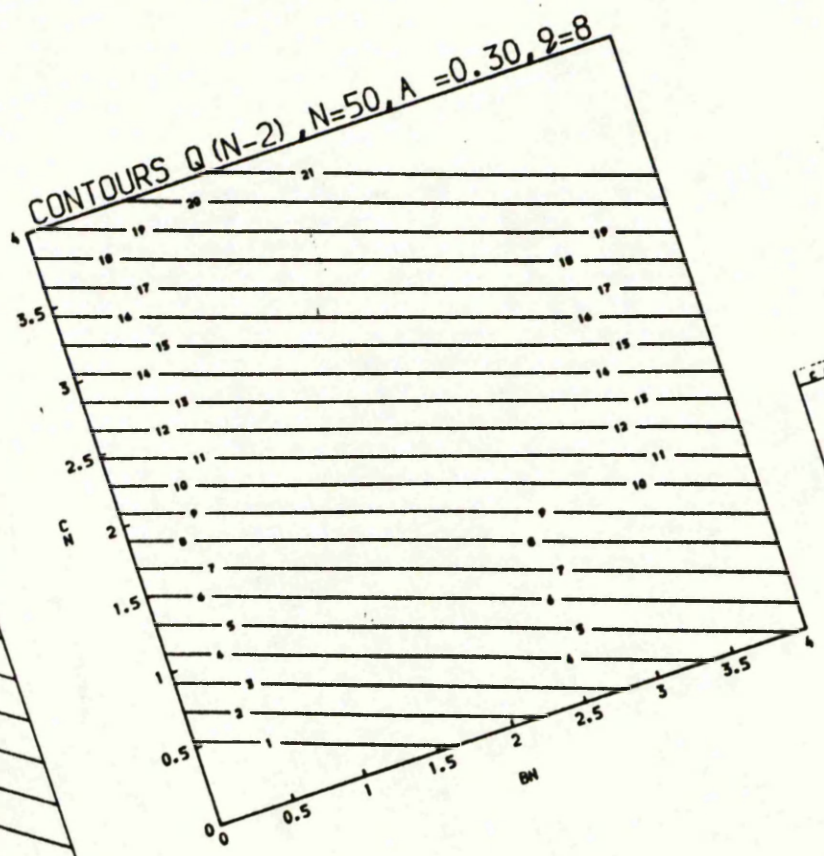

CONTOURS Q(N-2), N=50, CN=4.00, P=1
CONTOURS Q(N-2), N=50, CN=4.00, P=2
CONTOURS Q(N-2), N=50, CN=4.00, P=3
CONTOURS Q(N-2), N=50, CN=4.00, P=4
A
BN
CONTOUR KEY
1 1.000
2 1.750
3 2.500
4 3.250
5 4.000
6 4.750
7 5.500
8 6.250
9 7.000
10 7.750
11 8.500

FIGURE 6.47

FIGURE 6.48





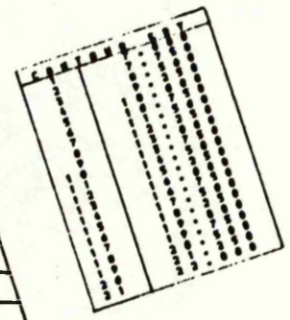
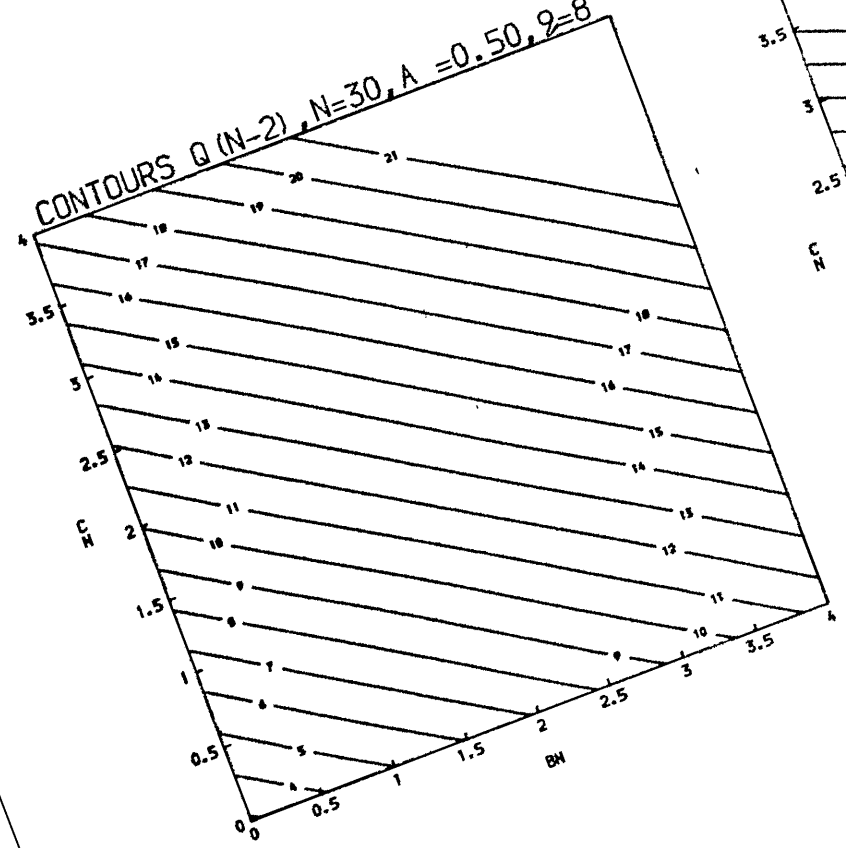
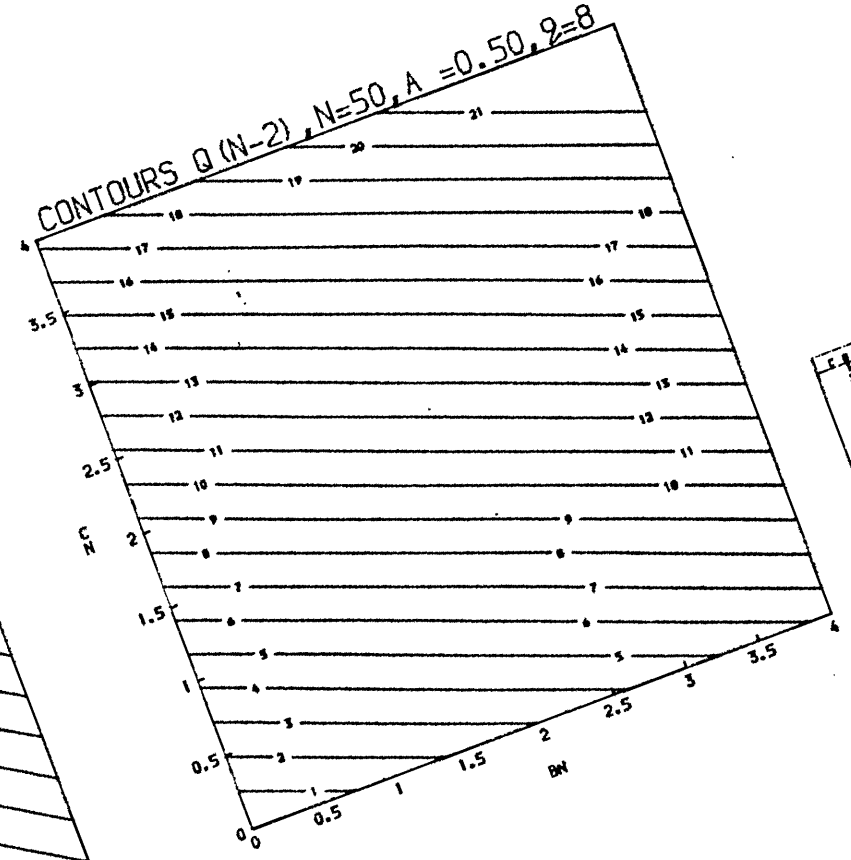


FIGURE 6.49





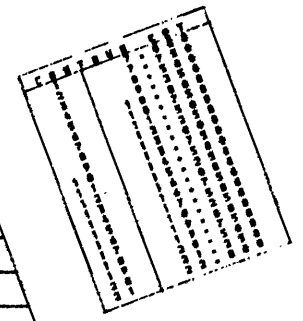


FIGURE 6.50

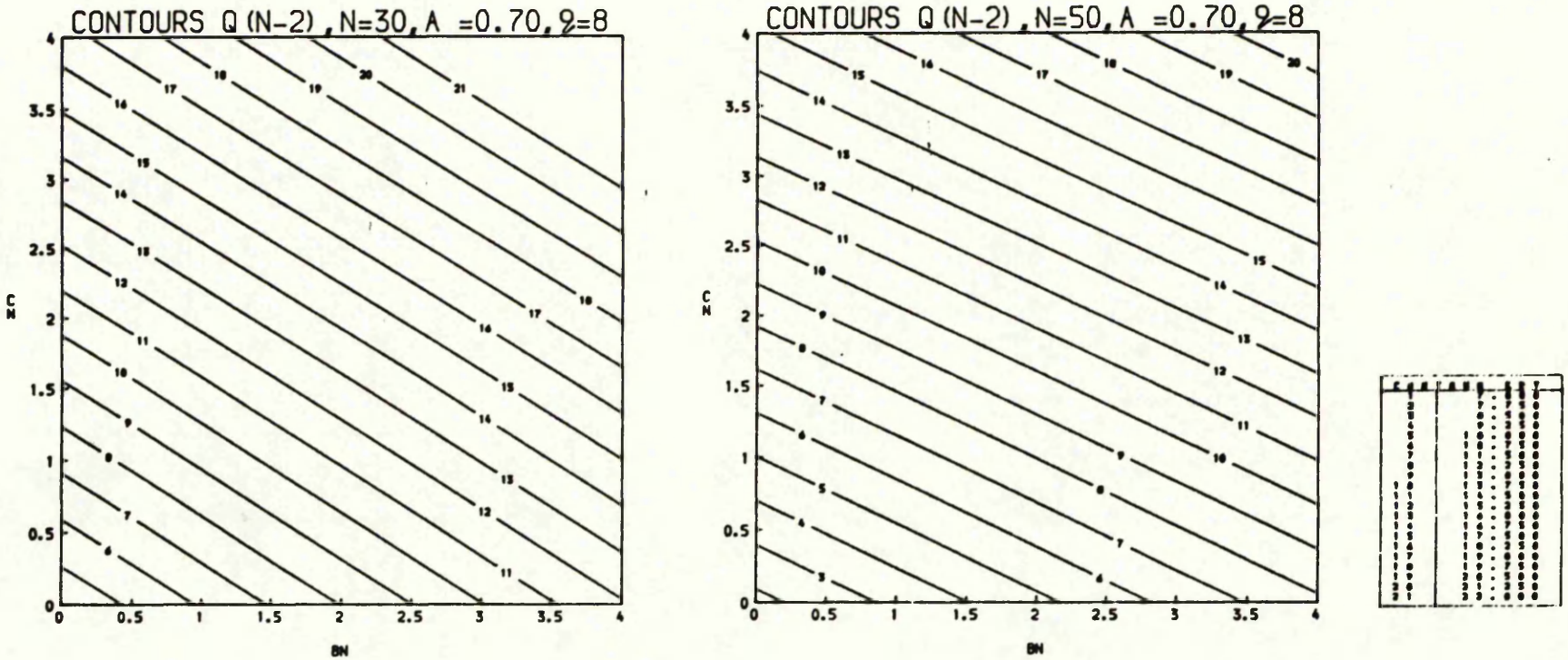
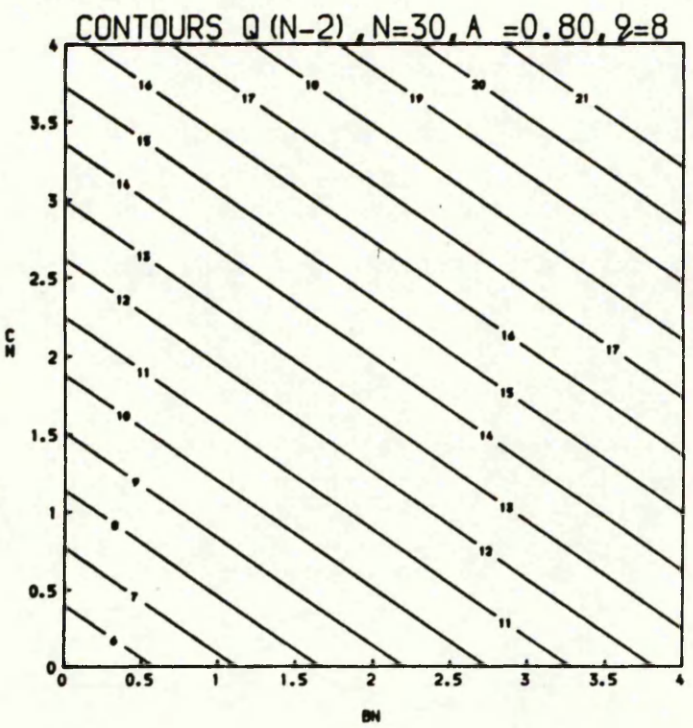
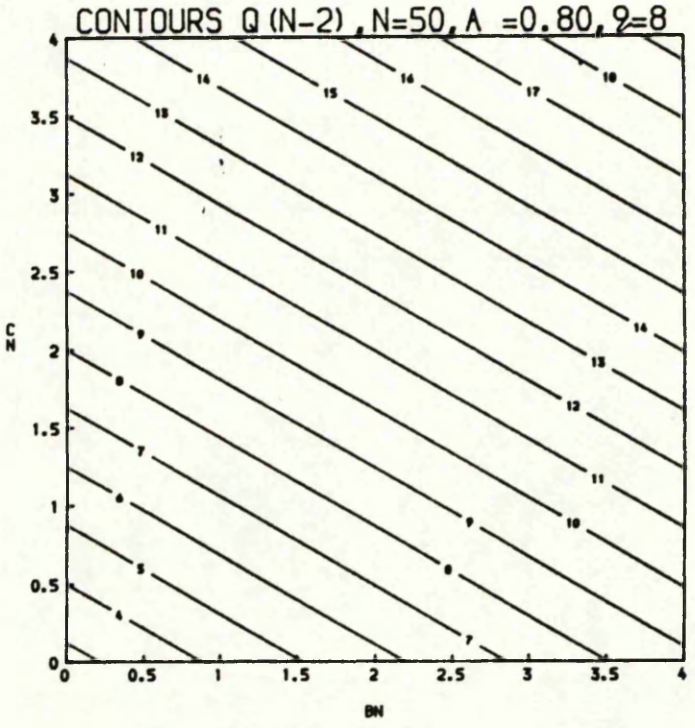
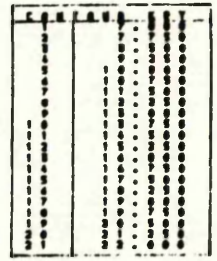

CONTOURS Q (N-2) , N=30, A =0.70, 9=8
CONTOURS Q (N-2) , N=50, A =0.70, 9=8
CN
BN
0
0.5
1
1.5
2
2.5
3
3.5
4

FIGURE 6.51


CONTOURS Q(N-2), N=30, A =0.80, ℓ=8
CN
BN
0
0.5
1
1.5
2
2.5
3
3.5
4
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21


CONTOURS Q(N-2), N=50, A =0.80, ℓ=8
CN
BN
0
0.5
1
1.5
2
2.5
3
3.5
4
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18



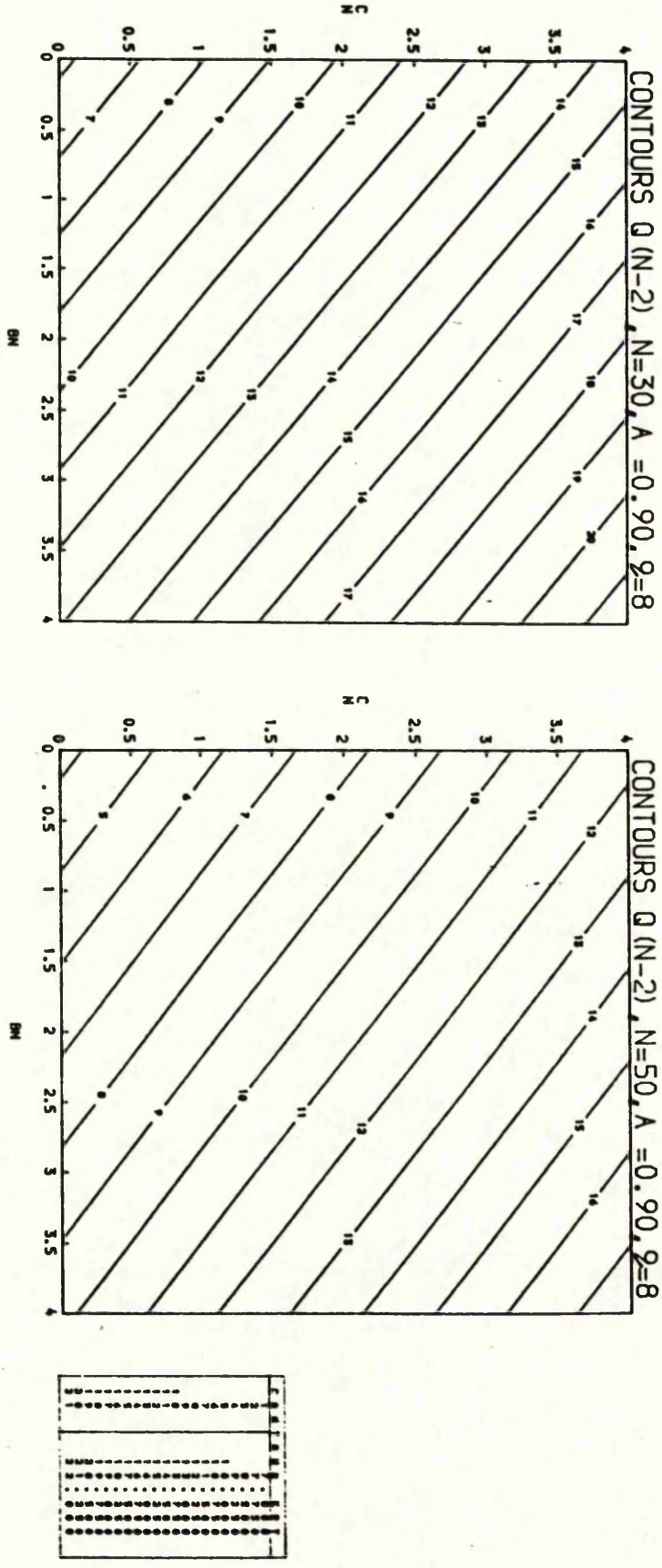


FIGURE 6.52

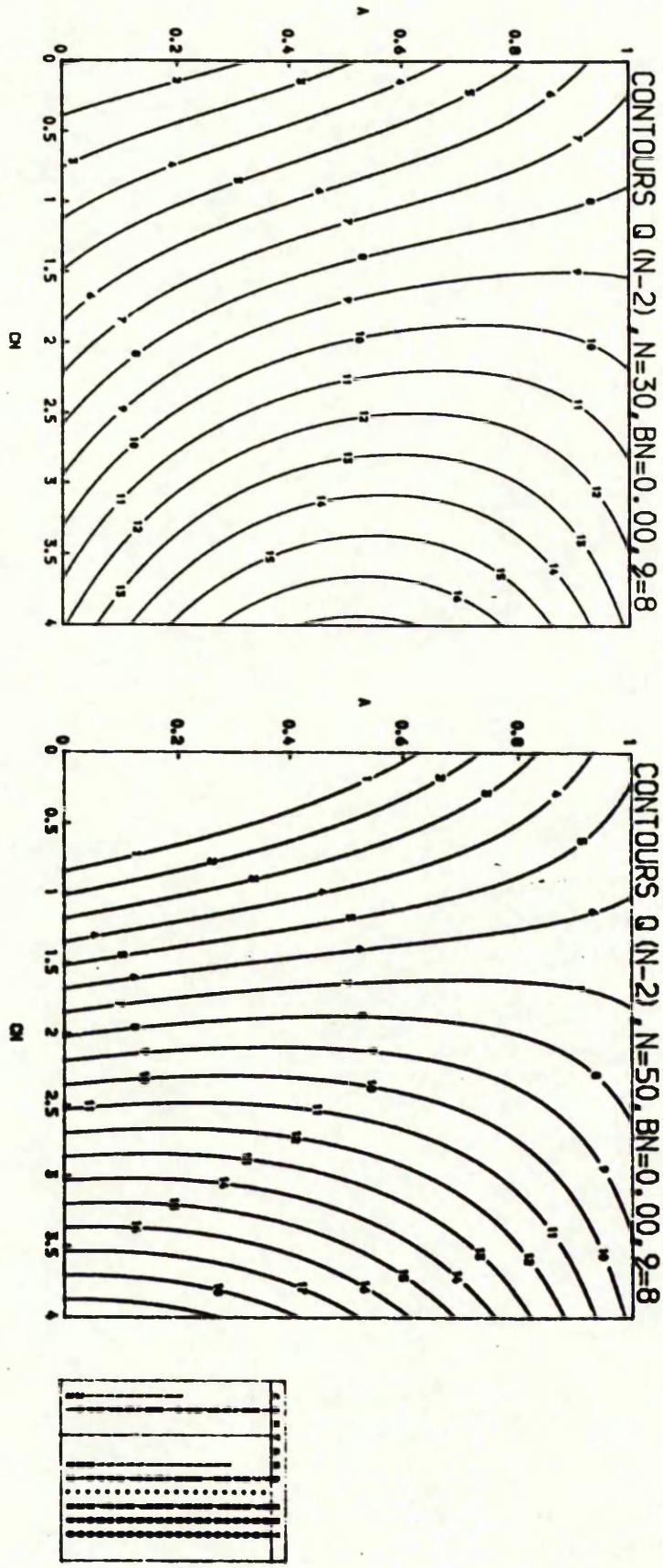


FIGURE 6.53

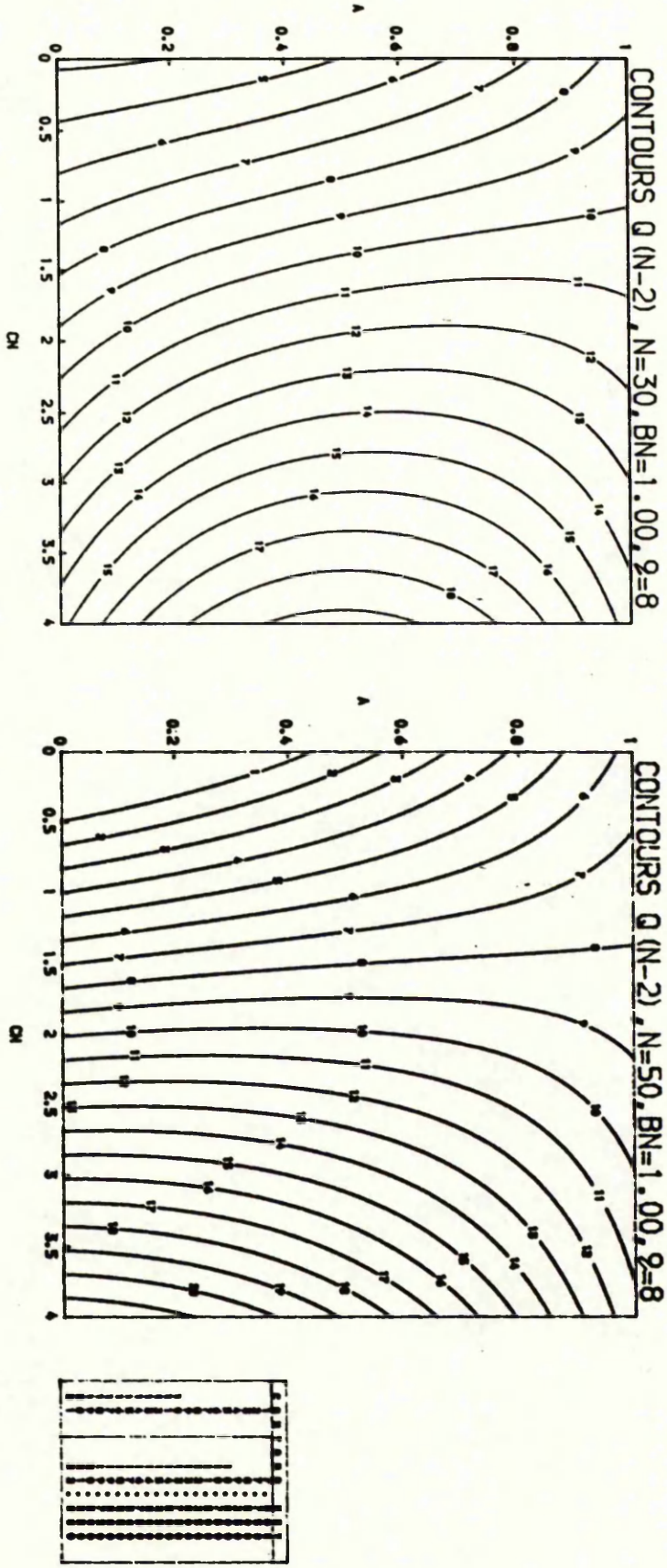


FIGURE 6.54

FIGURE 6.55

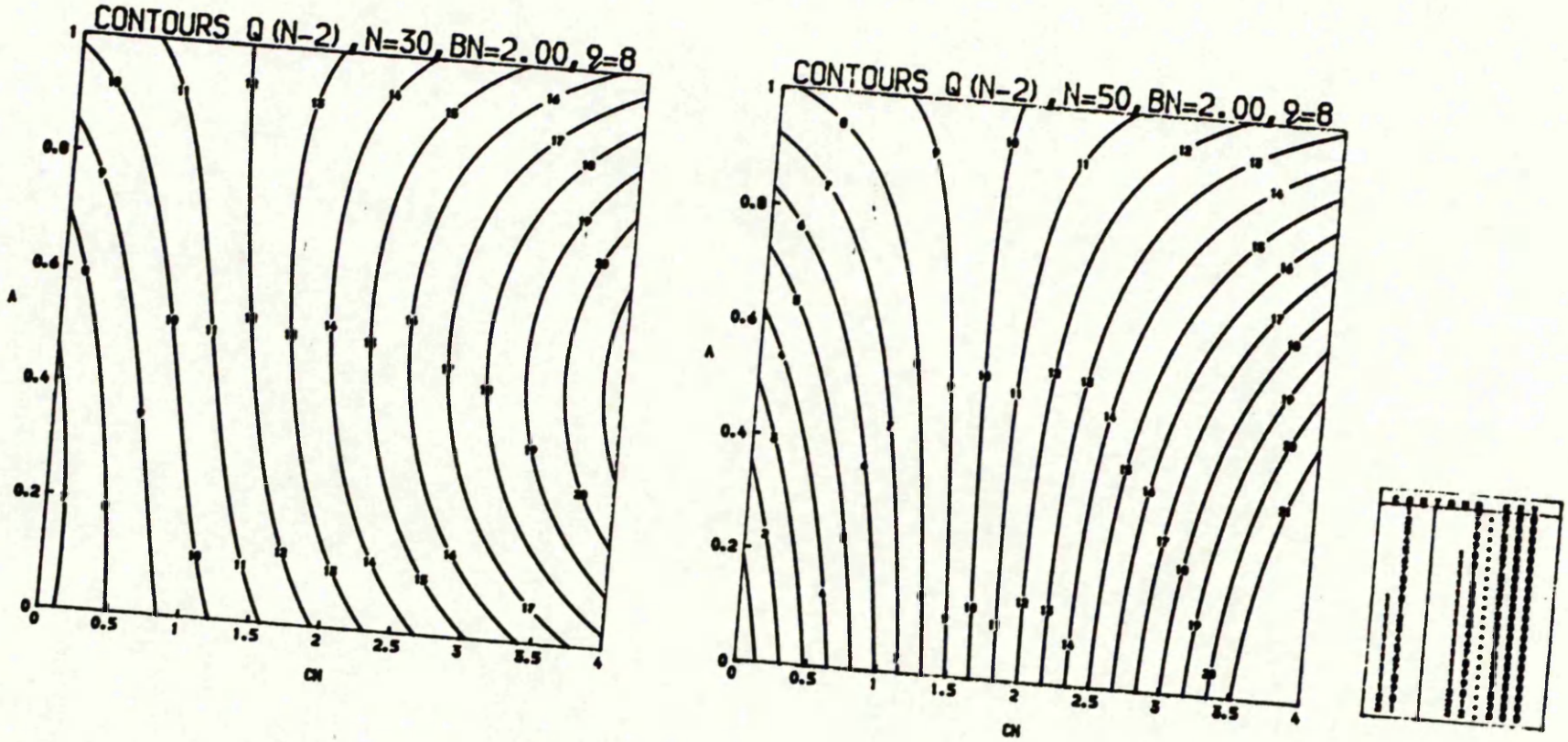

CONTOURS Q (N-2) , N=30, BN=2.00, 9=8
CONTOURS Q (N-2) , N=50, BN=2.00, 9=8
A
CN
0
0.2
0.4
0.6
0.8
1
0.5
1.5
2
2.5
3
3.5
4

FIGURE 6.56

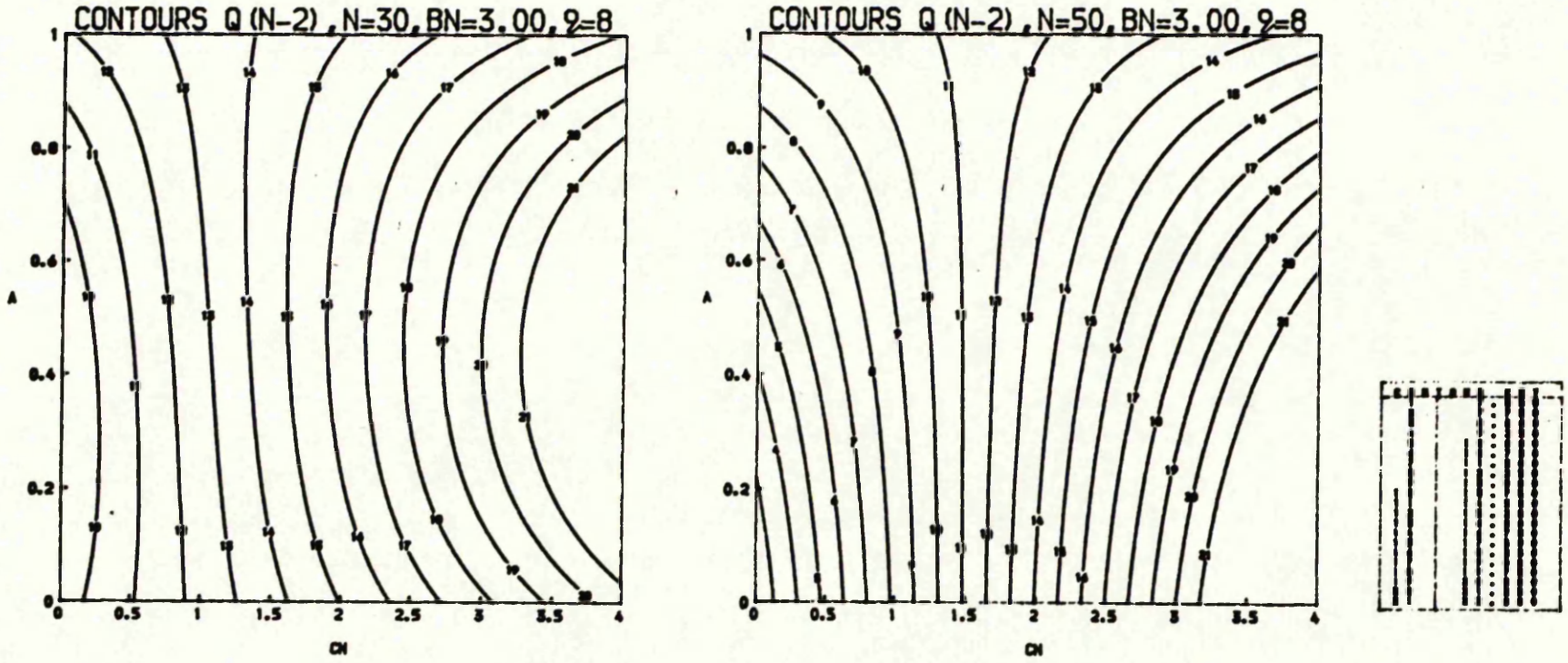

CONTOURS Q(N-2), N=30, BN=3.00, ϑ=8
CONTOURS Q(N-2), N=50, BN=3.00, ϑ=8
A
CN
0
0.2
0.4
0.6
0.8
1
0.5
1.5
2
2.5
3
3.5
4

FIGURE 6.57

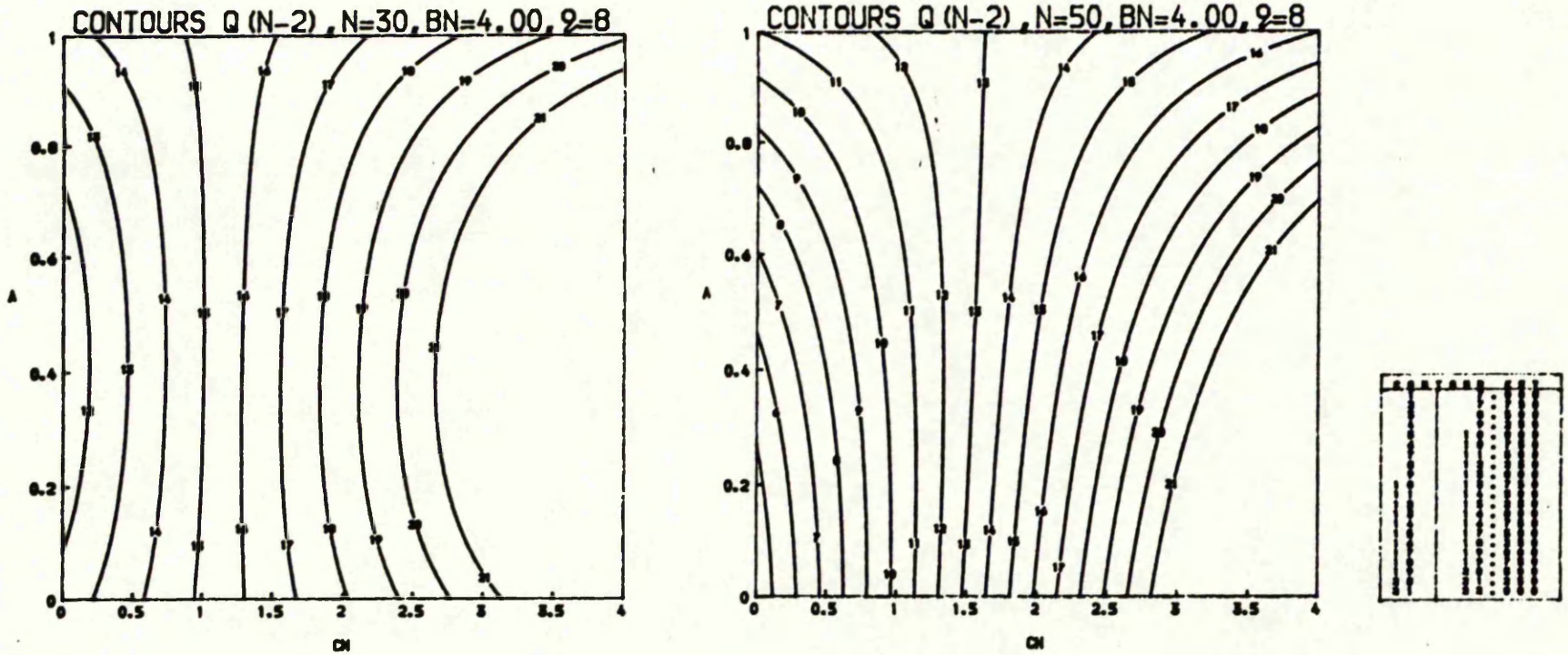
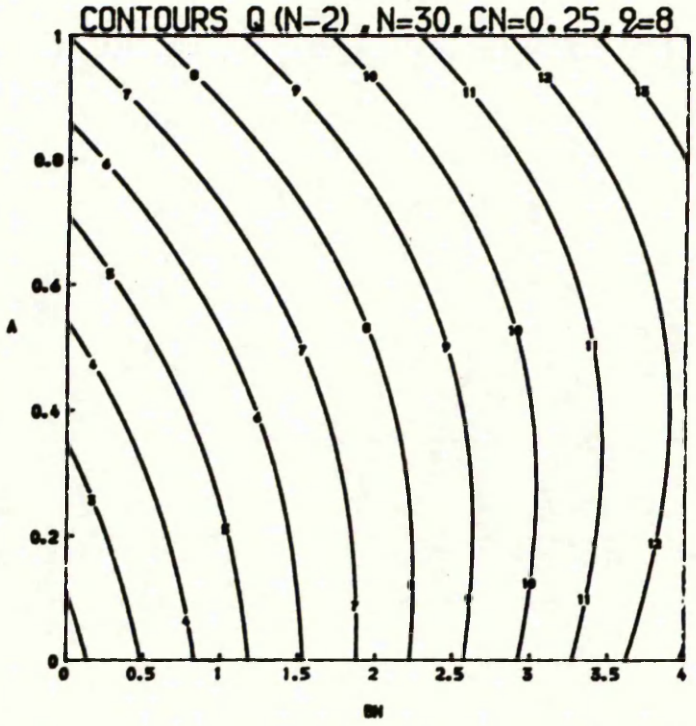
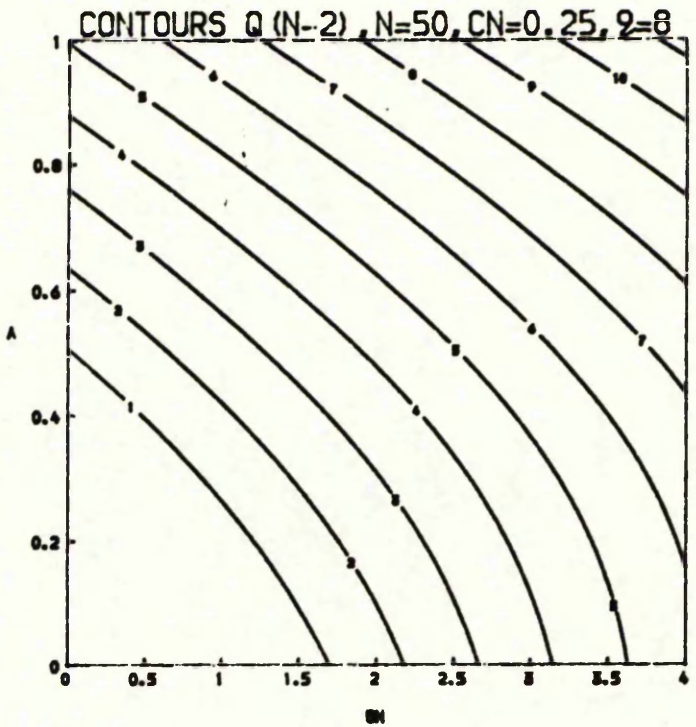

CONTOURS Q (N-2) , N=30, BN=4.00, 9=8
CONTOURS Q (N-2) , N=50, BN=4.00, 9=8
A
CN
0
0.2
0.4
0.6
0.8
1
0.5
1.5
2
2.5
3
3.5
4

FIGURE 6.58


CONTOURS Q(N-2), N=30, CN=0.25, ϑ=8
A
BN
0
0.2
0.4
0.6
0.8
1
0.5
1.5
2
2.5
3
3.5
4


CONTOURS Q(N-2), N=50, CN=0.25, ϑ=8
A
BN
0
0.2
0.4
0.6
0.8
1
0.5
1.5
2
2.5
3
3.5
4

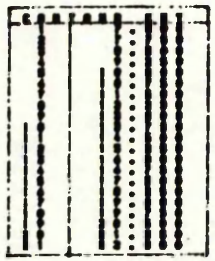


FIGURE 6.59

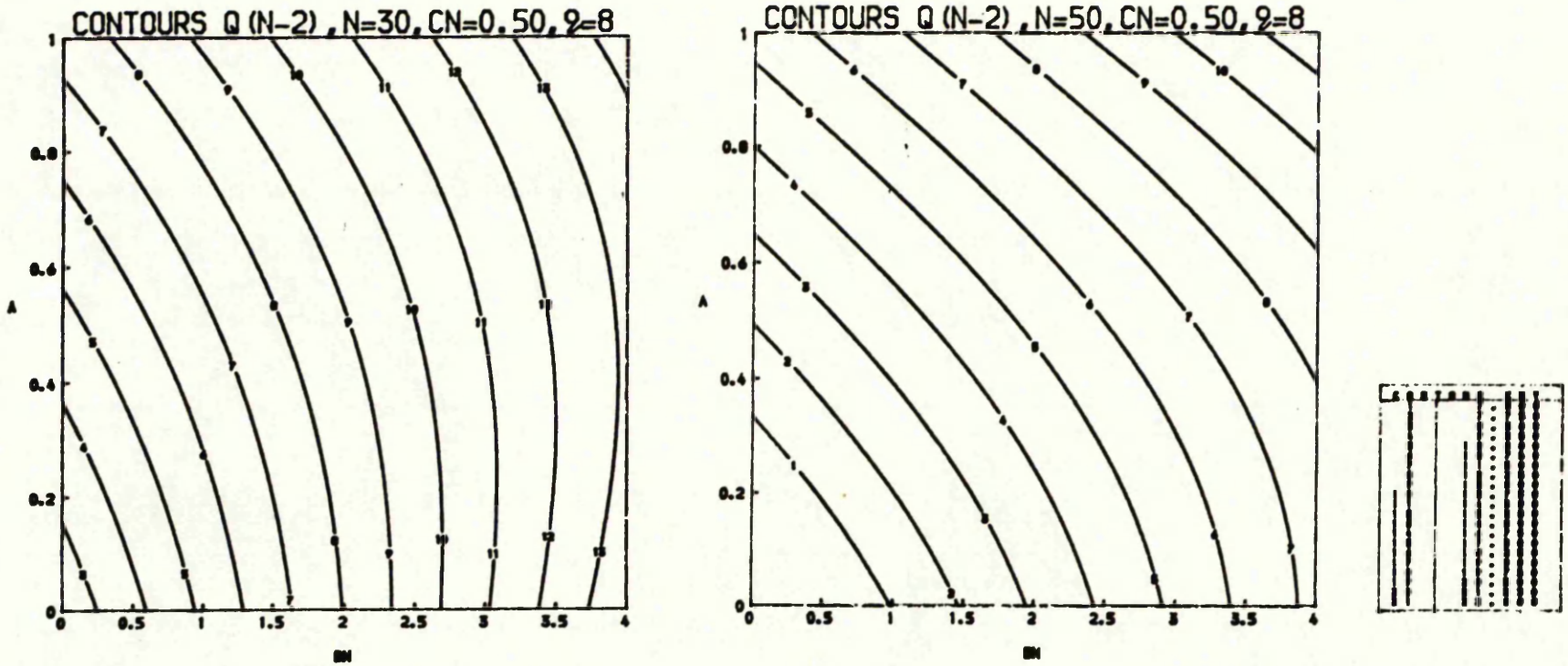

CONTOURS Q (N-2) , N=30, CN=0.50, 2=8
A
BN
0
0.2
0.4
0.8
1
0.5
1.5
2
2.5
3
3.5
4
CONTOURS Q (N-2) , N=50, CN=0.50, 2=8
A
BN

FIGURE 6.60

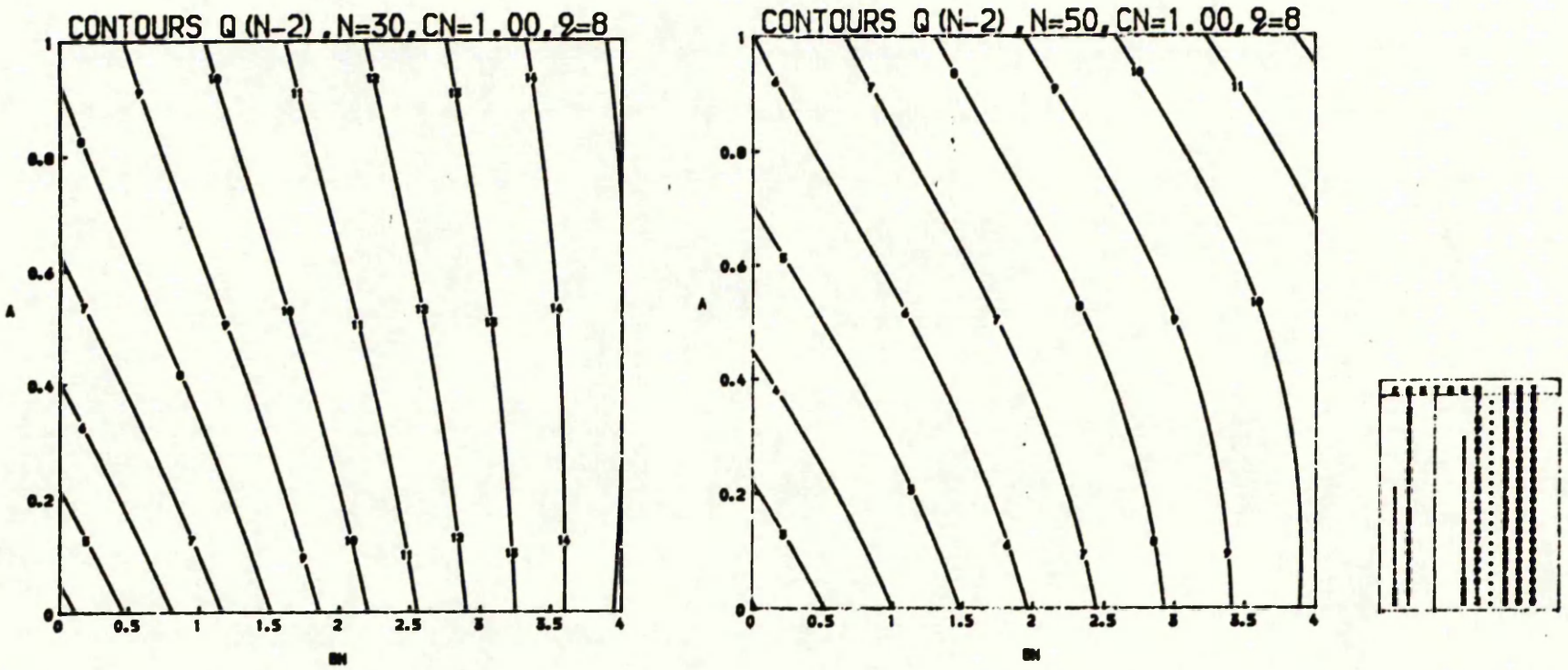
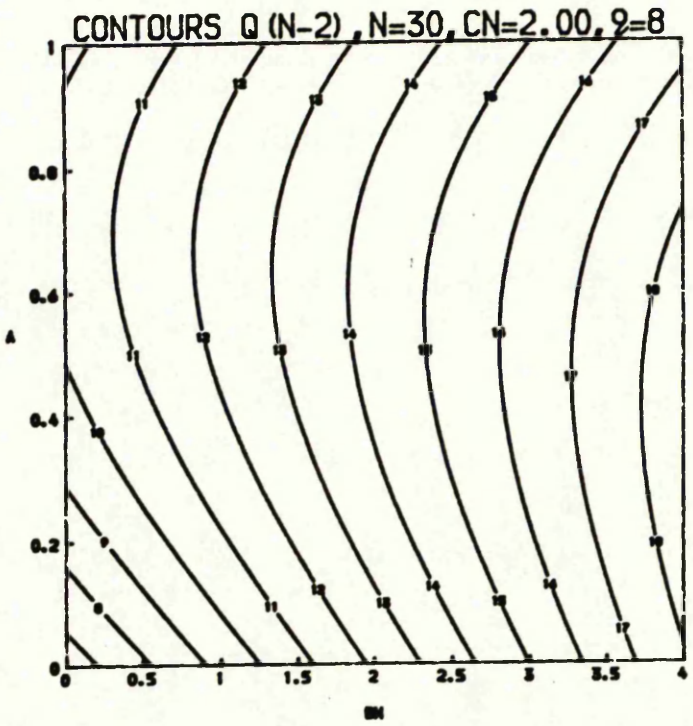
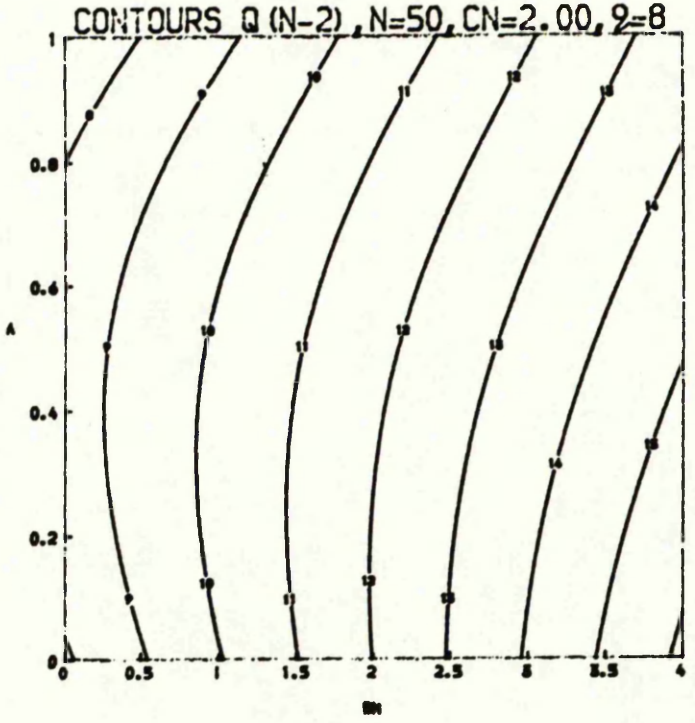

CONTOURS Q(N-2), N=30, CN=1.00, Q=8
CONTOURS Q(N-2), N=50, CN=1.00, Q=8
A
BN
0
0.2
0.4
0.6
0.8
1
0.5
1.5
2
2.5
3
3.5
4

FIGURE 6.61


CONTOURS Q(N-2), N=30, CN=2.00, 9=8
A
BN
0
0.2
0.4
0.6
0.8
1
0.5
1.5
2
2.5
3
3.5
4


CONTOURS Q(N-2), N=50, CN=2.00, 9=8
A
BN
0
0.2
0.4
0.6
0.8
1
0.5
1.5
2
2.5
3
3.5
4

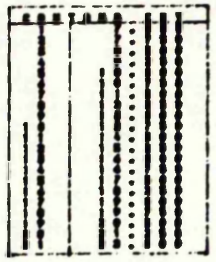
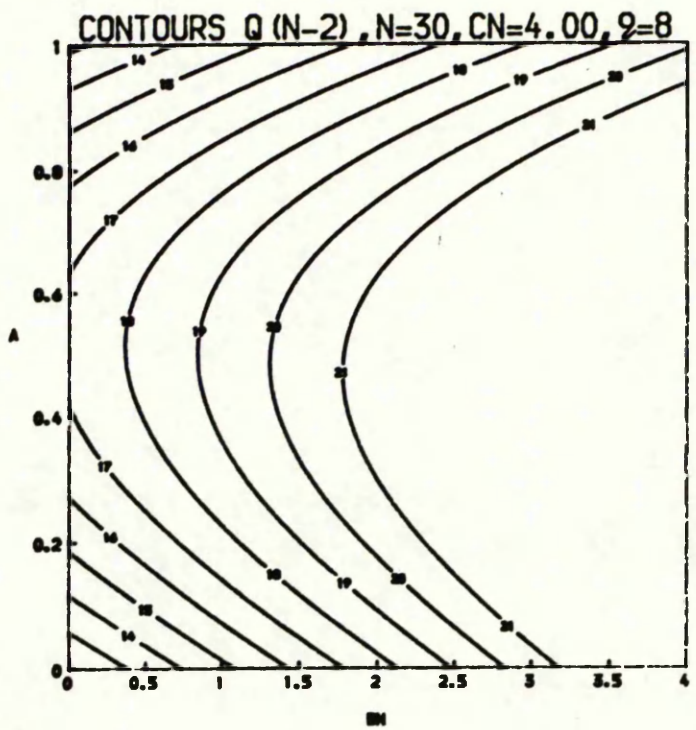
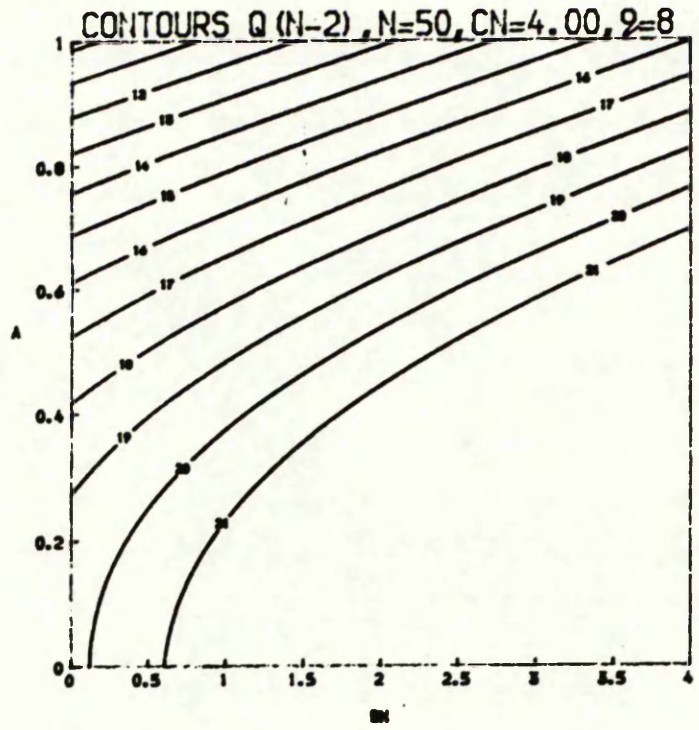


FIGURE 6.62


CONTOURS Q (N-2) , N=30, CN=4.00, 9=8
A
BN
0
0.2
0.4
0.6
0.8
1
0.5
1.5
2
2.5
3
3.5
4


CONTOURS Q (N-2) , N=50, CN=4.00, 9=8
A
BN
0
0.2
0.4
0.6
0.8
1
0.5
1.5
2
2.5
3
3.5
4

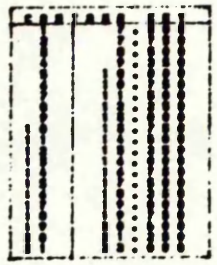
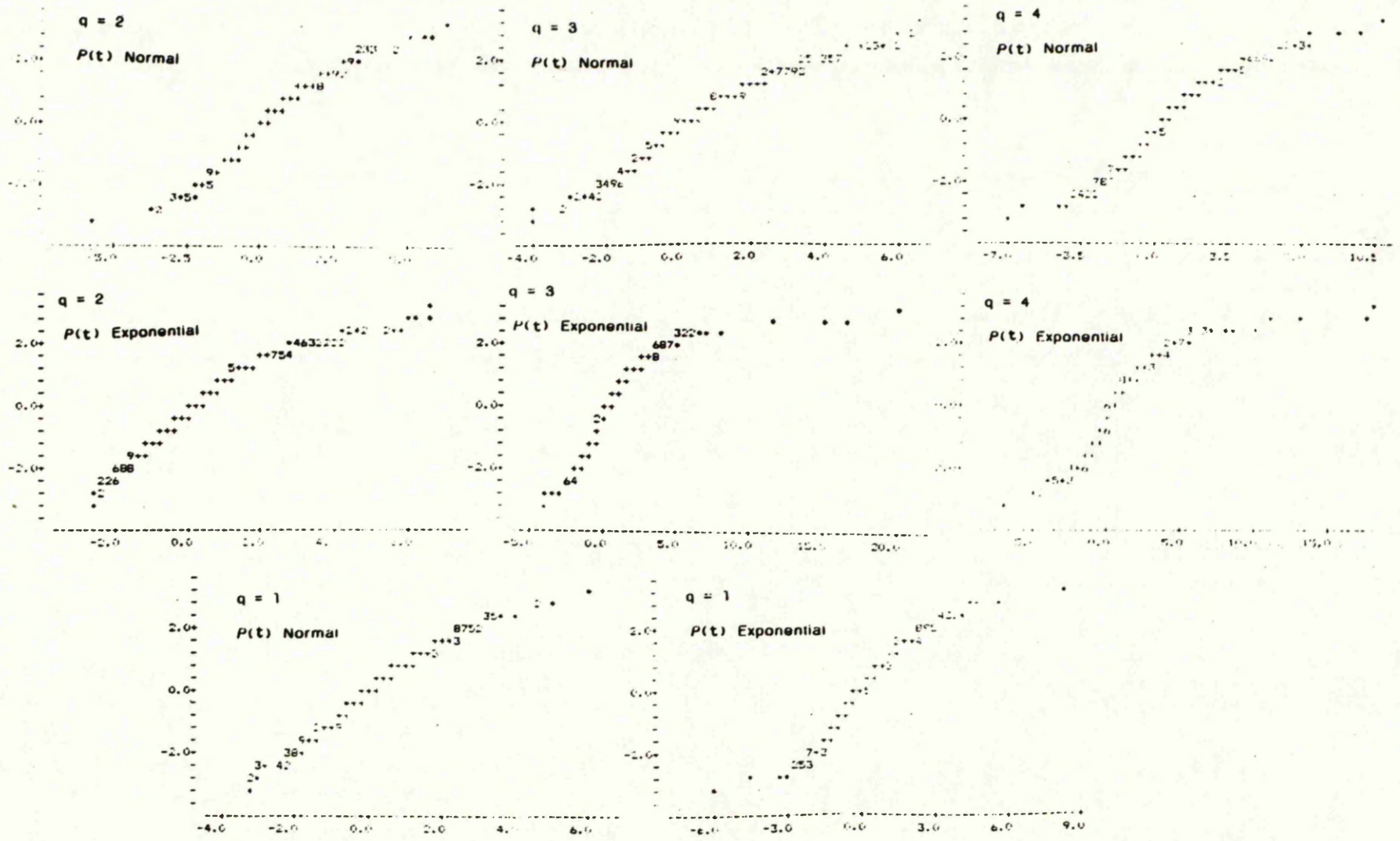


FIGURE 6.63. Normal Probability Plots of F_1 with different q and $P(t)$ for $B_N = 4.0$, $C_N = 4.0$, $N = 10$ and $\rho' = 0.7$



CHAPTER 7

BAYESIAN BEST LINEAR PREDICTOR

7.1. INTRODUCTION

Let $P(X|t;\theta)$ be the distribution of X in the regression experiment where $\theta = (\alpha,\beta,\Gamma)$ and prior distribution of θ denoted by $\pi(\theta)$ is required in the Bayesian analysis and also $\pi(t)$ to get the distribution of $\pi(t|\text{expt.};X)$ because by Bayes formula

$$\pi(t|\text{expt.};X) = \frac{\pi(t)\ \pi(X|t;\text{expt.})}{\int \pi(t)\ \pi(X|t;\text{expt.})\ dt}$$

Note that $\pi(t|\text{expt.}) = \pi(t)$ because experimental values tell us nothing about future T.

Now

$$\begin{aligned}\pi(X|t;\text{expt.}) &= \int \pi(X,\theta|t,\text{expt.})d\theta \\ &= \int \pi(X|\theta,t;\text{expt.})\ \pi(\theta|t;\text{expt.})d\theta \\ &= \int P(X|\theta;t)\ \pi(\theta|\text{expt.})d\theta\end{aligned}$$

$P(X|\theta;t)$ will be taken as $N_q(\alpha+\beta t;\ \Gamma)$, as in earlier chapters. For fuller discussion see Aitchison and Dunsmore (1975) chapter 10.

The best linear predictor of t using

$$\pi(t,X|\text{expt.}) = \pi(t)\ \pi(X|t;\text{expt.})$$

is denoted by

$$C_\pi + D_\pi^T X \qquad \ldots(7.1)$$

and its Bayesian mean squared error is defined by

$$MSE_\pi = E\left[\{T-(C_\pi+D_\pi^T X)\}^2|\text{expt.}\right] \qquad \ldots(7.2)$$

The advantage of using this linear predictor (Hartigan, 1969) instead of $E(T|\text{expt.};X)$ is that the former depends on $\pi(t)$ only through its moments μ, σ^2.

There are two situations under large N where $E(T|\text{expt};X)$ can be easily approximated.

(i) If N is very large then $\hat{\theta}_{ML} = \theta$;

$\pi(\theta|\text{expt.})$ is concentrated on point θ

$\pi(X|t;\text{expt.}) = P(X|\theta;t)$

thus

$$\pi(t|\text{expt.};X) = \frac{\pi(t)\ P(X|\theta;t)}{\int \pi(t)\ P(X|\theta;t)} \qquad \ldots(7.3)$$

If also $\pi(t)$ is $N(\mu,\sigma^2)$, then $\pi(t|\text{expt.};X)$ is

$$N[\,C+D^TX;\ (1-\rho^2)\sigma^2\,]$$

If $\pi(t)$ is not normal then $\pi(t|\text{expt.};X)$ is not normal.

Best linear predictor, using $\pi(t;X|\text{expt.}) = \pi(t)\ P(X|\theta;t)$ is

$$C_\pi + D_\pi{}^TX \approx \hat{C} + \hat{D}^TX$$

and

$$MSE_\pi \approx \hat{MSE}$$

(ii) If N is large but still some uncertainty about θ would lead to uncertainty about C and D.

When $\pi(\theta)$ has special form, formula for $\pi(X|t;\text{expt.})$, namely Student's t can be approximated by normal with mean vector and some covariance matrix. Thus $\pi(t;X|\text{expt.})$ is approximately normal.

Main idea below is not to use approximations to $E(T|\text{expt.};X)$ but instead to use relation (7.1).

We will discuss first p = q = 1 and then general q and p = 1.

7.2. BAYESIAN LINEAR PREDICTOR FOR q = p = 1

If $\pi(\alpha,\beta,\sigma^2{}_{x|t})$ is such that $\alpha,\beta,\log\sigma^2{}_{x|t}$ are independent uniform $(-\infty < \alpha,\beta,\log\sigma^2{}_{x|t} < \infty)$, i.e. "vague prior information", then:

$\pi(x|t;\text{expt.})$ is such that

$$\frac{x - (\hat{\alpha} + \hat{\beta}t)}{\hat{\sigma}_{x|t}[\,1+1/N\ +(t-\bar{t})^2/S_{TT}\,]^{\frac{1}{2}}} \quad \text{is } t_{N-2}.$$

then

$$x = \hat{\alpha}+\hat{\beta}t\ +\ \hat{\sigma}_{x|t}[\,1+1/N\ +(t-\bar{t})^2/S_{TT}\,]^{\frac{1}{2}}\ t_{N-2}$$

Consider $\pi(t,x|\text{expt.})$, then the Bayesian linear predictor of t would be

$$C_\pi + D_\pi X = t_\pi = E(T) + \frac{COV(X,T)}{VAR(X)}(x - E(X)) \qquad \ldots(7.4)$$

Where all the moments are conditional on "experiment" and

$$MSE_\pi = (1-\rho_\pi^2)VAR(T)$$
$$= (1-\rho_\pi^2)\sigma^2 \qquad \ldots(7.5)$$

where ρ_π^2 is given in expression (7.8).

We have assumed $E(T) = \mu$ and $VAR(T) = \sigma^2$ known and the values of $COV(X,T)$, $VAR(X)$ and $E(X)$ can be calculated using $\pi(X|t;\text{expt.})$.

$$E(X) = EE(X|T) = E(\hat{\alpha} + \hat{\beta}t)$$
$$= \hat{\alpha} + \hat{\beta}\,\mu$$

$$VAR(X) = VAR[E(X|T)] + E[VAR(X|T)]$$
$$= VAR(\hat{\alpha}+\hat{\beta}T) + E\hat{\sigma}^2_{x|t}[1+1/N+(t-\bar{t})^2/S_{TT}]VAR(t_{N-2})$$
$$= \hat{\beta}^2\sigma^2+(N-2)/(N-4)\hat{\sigma}^2_{x|t}[1+1/N+\{\sigma^2+(\mu-\bar{t})^2\}/S_{TT}]$$
$$[E(t-\bar{t})^2 = E(T-\mu+\mu-\bar{t})^2]$$

$$COV(X,T) = E(XT) - E(X)E(T)$$
$$= EE(XT|T) - E(X)E(T)$$
$$= E(\hat{\alpha}T + \hat{\beta}T^2) - (\hat{\alpha}+\hat{\beta}\mu)\mu$$
$$= \hat{\alpha}\mu + \hat{\beta}(\sigma^2+\mu^2) - (\hat{\alpha}+\hat{\beta}\mu)\mu$$
$$= \hat{\beta}\sigma^2$$

After substituting the values of these moments in (7.4) we get

$$t_\pi = \mu + \frac{COV(T,X)}{VAR(X)}(X - \hat{\alpha} - \hat{\beta}\mu)$$
$$= \mu + \frac{\hat{\beta}\sigma^2\hat{\beta}}{VAR(X)}[(X-\hat{\alpha})/\hat{\beta} - \mu]$$
$$= \mu + \frac{\hat{\beta}^2\sigma^2[(X-\hat{\alpha})/\hat{\beta} - \mu)]}{\hat{\beta}^2\sigma^2+[(N-2)/(N-4)]\hat{\sigma}^2_{x|t}[1+1/N+\{\sigma^2+(\mu-\bar{t})^2\}/S_{TT}]}$$
$$= \mu + \rho_\pi^2\left[\frac{X-\hat{\alpha}}{\hat{\beta}} - \mu\right]$$
$$= \rho_\pi^2\frac{(X-\hat{\alpha})}{\hat{\beta}} + (1-\rho_\pi^2)\,\mu \qquad \ldots(7.6)$$

and MSE_π (7.5) would be

$$MSE_\pi = \sigma^2\left[1 - \frac{\hat{\beta}^2\sigma^4}{\sigma^2\left[\hat{\beta}^2\sigma^2+((N-2)/(N-4))\hat{\sigma}^2_{x|t}\left[1+1/N+\{\sigma^2+(\mu-\bar{t})^2\}/S_{TT}\right]\right.}\right]$$

$$= \sigma^2\left[1 - \rho_\pi^2\right] \qquad \ldots(7.7)$$

where

$$\rho_\pi^2 = \frac{\hat{\beta}^2\sigma^2}{\hat{\beta}^2\sigma^2+((N-2)/(N-4))\hat{\sigma}^2_{x|t}\left[1+1/N+\{\sigma^2+(\mu-\bar{t})^2\}/S_{TT}\right]} \qquad \ldots(7.8)$$

When $N \to \infty$; $\bar{t}$ fixed and so $S_{TT} \to \infty$, then (7.6) would be

$$t_\pi = \mu + \frac{\hat{\beta}\sigma^2[x - (\hat{\alpha}+\hat{\beta}\mu)]}{\hat{\beta}^2\sigma^2 + \hat{\sigma}^2_{x|t}}$$

$$= (1-\hat{\rho}^2)\mu + \hat{\rho}^2\,(x-\hat{\alpha})/\hat{\beta}$$

$$= \hat{C} + \hat{D}X, \text{ which is the best linear predictor (3.7)}$$

and the mean squared error MSE_π (7.7) would be equal to $\hat{MSE}$ (3.11).

For $N < \infty$, $\rho_\pi^2 \neq \hat{\rho}^2$ and ρ_π^2 can be written in terms of the invariants already discussed i.e. N, B_N, C_N and $\hat{\rho}^2$ and after simple algebra it would be

$$\rho_\pi^2 = \frac{\hat{\rho}^2}{\hat{\rho}^2+(N-2)/(N-4)\{(1-\hat{\rho}^2)\}\left[1+ 1/N +B_N/(N-2)+ C_N/(N-2)\right]} \qquad \ldots(7.9)$$

So MSE_π (7.7) i.e. mean squared error (Bayesian) depends on the above four invariants N, B_N, C_N and $\hat{\rho}^2$.

It should be noted that both the Bayesian (7.6) and the best linear predictor (3.7) are weighted average of classical estimator and the μ. Since $\rho_\pi^2 < \hat{\rho}^2$, Bayesian gives less weight to classical estimator. In particular when $\mu = \bar{t}$; $\sigma^2 = S_{TT}/(N-2)$, $C_\pi + D_\pi X \neq$ inverse estimator.

The expression (7.7) can be written as under

$$MSE_\pi = \sigma^2(1 - \rho_\pi^2)$$

$$= \sigma^2(1 - \hat{\rho}^2)(1 + Q_\pi)$$

so

$$Q_\pi = \frac{1 - \rho_\pi^2}{1 - \hat{\rho}^2} - 1 \qquad \ldots(7.10)$$

As ρ_π^2 (7.9) is a function of the four invariants; N, B_N, C_N and

$\hat{\rho}^2$ so would be the Q_π.

81 values of the Q_π are calculated for the same combinations of the invariants as in chapter 3 to compare with the simulated values of Q_s (3.10). These values of Q_π are given in table 7.1 and simulated values Q_s are given in table 7.2.

7.3. BAYESIAN LINEAR PREDICTOR FOR GENERAL q and p = 1

Suppose (X,T) has a q+1 dimensional joint distribution which is known. Then we derived the best linear predictor (5.11) which is

$$\hat{t} = E(T) + D^T(X - EX)$$

where $D = \underset{\substack{\text{covariance}\\ \text{matrix}}}{\{COV(X)\}^{-1}} \quad \underset{\substack{\text{covariance}\\ \text{vector}}}{COV(T,X)}$

and has mean squared error

$$MSE = VAR(T) - D^T COV(T,X)$$

There are two cases here.

(a) CASE 1

Known joint distribution is frequency distribution (non-Bayesian).

$$P(x,t|\alpha,\beta,\Gamma;\mu;\sigma^2) = P(x|t;\alpha,\beta,\Gamma)\ P(t|\mu,\sigma^2)$$

then

$$D = \sigma^2\ \{\Gamma + \sigma^2\beta\beta^T\}^{-1}\ \beta$$

$$C = E(T) - D^T\ E(X)$$

and

$$MSE = \sigma^2[1 - \sigma^2\beta^T(\Gamma + \sigma^2\beta\beta^T)^{-1}\beta]$$

$$= \sigma^2(1 - \rho^2)$$

where $\rho^2 = \sigma^2\beta^T(\Gamma + \sigma^2\beta\beta^T)^{-1}\beta$

This case has already been discussed in chapter 5.

(b) CASE 2

Known joint distribution $\pi(X,t|\text{expt.})$ arises in Bayesian

analysis. This is a marginal distribution of

$$\pi(\Gamma, X, t \mid \text{expt.}) = \pi(X \mid \Gamma, t; \text{expt.})\ \pi(\Gamma \mid t; \text{expt})\ \pi(t \mid, \text{expt}; \mu, \sigma^2).$$

It is convenient to consider these conditional factors in finding moments of $\pi(x, t \mid \text{expt.})$

These moments will give the Bayesian linear predictor $C_\pi + D_\pi^T X$ as follows;

$$C_\pi + D_\pi^T X = \mu + [X - E(X)]^T \{COV(X \mid \text{expt.})\}^{-1}\ COV(T, X \mid \text{expt.}) \quad \ldots(7.11)$$

($\{COV(X \mid \text{expt.})\}^{-1}$: *posterior covariance matrix*; $COV(T, X \mid \text{expt.})$: *posterior covariance vector*)

Now we derive all the posterior moments from the standard prior stated after the relation (7.12) for the regression parameters, and assuming as always that $E(T) = \mu$ and $VAR(T) = \sigma^2$.

$\pi(X \mid \Gamma; t, \text{expt.})$ is such that

$$X = \hat{\alpha} + \hat{\beta} t + N[0,\ \Gamma\{1 + 1/N + (t - \bar{t})^2 / S_{TT}\}]$$

By $E(X)$ in equation (7.11) we mean

$$\begin{aligned} E(X \mid \text{expt.}) &= EE(X \mid t, \text{expt.}) = EEE(X \mid \Gamma, t, \text{expt.}) \\ &= EE(\hat{\alpha} + \hat{\beta} t) \\ &= \hat{\alpha} + \hat{\beta}\mu \end{aligned}$$

$$\begin{aligned} E[X(T - E(T)) \mid \text{expt.}] &= E(XT \mid \text{expt.}) - E(T) E(X \mid \text{expt.}) \\ &= \hat{\alpha}\mu + \hat{\beta}(\mu^2 + \sigma^2) - (\hat{\alpha} + \hat{\beta}\mu)\mu \\ &= \hat{\beta}\sigma^2 \end{aligned}$$

$$COV(X \mid \text{expt.}) = COV[E(X \mid \Gamma; t, \text{expt.})] + E[COV(X \mid \Gamma; t, \text{expt.})]$$

where both COV and E are posterior, i.e. conditioned on experiment.

$$\begin{aligned} &= COV(\hat{\alpha} + \hat{\beta} t) + E\{\Gamma[1 + 1/N + (t - \bar{t})^2 / S_{TT}]\} \\ &= \sigma^2 \hat{\beta}\hat{\beta}^T + E(\Gamma)\ E[1 + 1/N + (t - \bar{t})^2 / S_{TT}] \\ &= \sigma^2 \hat{\beta}\hat{\beta}^T + [(N-2)/(N-q-3)]\hat{\Gamma}[1 + 1/N + \{\sigma^2 + (\mu - \bar{t})^2\}/S_{TT}] \\ &= \sigma^2 \hat{\beta}\hat{\beta}^T + f\hat{\Gamma} \end{aligned}$$

where

$$\begin{aligned} f &= (N-2)/(N-q-3)[1 + 1/N + \{\sigma^2 + (\mu - \bar{t})^2\}/S_{TT}] \\ &= (N-2)/(N-q-3)[1 + 1/N + B_N/(N-2) + C_N/(N-2)] \quad \ldots(7.12) \end{aligned}$$

$$= f(N, B_N, C_N, q)$$

Following Press (1972), the prior densities for α, β and Γ are assumed as under,

$$\pi(\alpha, \beta) \propto \text{constant};$$

$$\pi(\Gamma) \propto 1/|\Gamma|^{(q+1)/2}.$$

The $E(\Gamma|\text{expt.};t) = [(N-2)/(N-q-3)]\ \hat{\Gamma}$ is obtained from Press (1972) theorem (8.6.3), where $\hat{\Gamma}$ is the usual unbiased estimate.

Using the above moments the Bayesian linear predictor (7.11) can be written as follows,

$$C_\pi + D_\pi{}^T X = \mu + [X - E(X)]^T [f\hat{\Gamma} + \sigma^2 \hat{\beta}\hat{\beta}^T]^{-1}\ \hat{\beta}\sigma^2 \qquad \ldots(7.13)$$

and

$$\begin{aligned} MSE_\pi &= \sigma^2 [1 - \sigma^2 \hat{\beta}^T \{f\hat{\Gamma} + \sigma^2 \hat{\beta}\hat{\beta}^T\}^{-1} \hat{\beta}] \\ &= \sigma^2 [1 - \rho_\pi{}^2] \end{aligned} \qquad \ldots(7.14)$$

$\rho_\pi{}^2$ has been defined as

$$\begin{aligned} \rho_\pi{}^2 &= \sigma^2 \hat{\beta}^T [f\hat{\Gamma} + \sigma^2 \hat{\beta}\hat{\beta}^T]^{-1}\ \hat{\beta} \\ &= b^T \{A + kbb^T\}^{-1} b \end{aligned}$$

where $b = \sigma\hat{\beta}$, $A = f\hat{\Gamma} + f\sigma^2 \hat{\beta}\hat{\beta}^T$ and $k = 1 - f$

By result following lemma (7.1), $\rho_\pi{}^2$ is

$$\begin{aligned} &= \frac{b^T A^{-1} b}{1 + kb^T A^{-1} b} \\ &= \frac{\sigma^2 \hat{\beta}^T f^{-1} (\hat{\Gamma} + \sigma^2 \hat{\beta}\hat{\beta}^T)^{-1} \hat{\beta}}{1 + (1-f)\sigma^2 \hat{\beta}^T f^{-1} (\hat{\Gamma} + \sigma^2 \hat{\beta}\hat{\beta}^T)^{-1} \hat{\beta}} \\ &= \frac{\hat{\rho}^2}{f + (1-f)\ \hat{\rho}^2} \end{aligned} \qquad \ldots(7.15)$$

LEMMA 7.1

If A is any non-singular square matrix, b a vector and k a scalar such that $A + kbb^T$ is non-singular, then

$$(A + kbb^T)^{-1} = \left[I - \frac{k}{1 + kb^T A^{-1} b} A^{-1} bb^T\right] A^{-1}$$

PROOF:

Premultiply by $(A + kbb^T)$ and postmultiply by A we obtain

$$A = [A + kbb^T][I - \frac{k}{1 + kb^TA^{-1}b} A^{-1}bb^T]$$

$$= A + gbb^T$$

where

$$g = k - \frac{k}{1 + kb^TA^{-1}b} - \frac{k^2}{1 + kb^TA^{-1}b} b^TA^{-1}b = 0.$$

Now we express $b^T\{A + kbb^T\}^{-1}b$ in terms of $b^TA^{-1}b$ and k.

$$b^T\{A + kbb^T\}^{-1}b = b^T[I - \frac{k}{1+kb^TA^{-1}b} A^{-1}bb^T]A^{-1}b$$

$$= b^TA^{-1}b - \frac{k}{1+kb^TA^{-1}b} (b^TA^{-1}b)^2$$

$$= \frac{b^TA^{-1}b}{1 + kb^TA^{-1}b}$$

By (7.14) $MSE_\pi = \sigma^2[1 - \rho_\pi^2]$

$$= \sigma^2[1 - \hat{\rho}^2][1 + Q_\pi]$$

Here $\hat{\rho}^2$ is the familiar non-Bayesian quantity. This is the definition of Q_π.

$$MSE_\pi/\sigma^2 = [1 - \hat{\rho}^2][1 + Q_\pi] \qquad \ldots(7.16)$$

and

$$\hat{MSE}/\sigma^2 = [1 - \hat{\rho}^2][1 + \hat{Q}_s] \qquad \ldots(7.17)$$

So comparing Q_π and $\hat{Q}_s$ is equivalent to comparing MSE_π/σ^2 and $\hat{MSE}/\sigma^2$.

$$Q_\pi = \frac{(1 - \rho_\pi^2)}{(1 - \hat{\rho}^2)} - 1$$

$$= \frac{f + (1 - f)\hat{\rho}^2 - \hat{\rho}^2}{(1- \hat{\rho}^2)[f +(1- f)\hat{\rho}^2]} - 1$$

$$= \frac{-(1- f)\hat{\rho}^2}{f + (1- f)\hat{\rho}^2}$$

$$= \frac{(f-1)\hat{\rho}^2}{f - (f - 1)\hat{\rho}^2} \qquad \ldots(7.18)$$

where f is as in (7.12).

Note that when N is large; MLE or unbiased estimators of Γ (as in case 1 above) are approximately equal to posterior expectation $E(\Gamma|\text{expt.})$. Also when N is large; $\bar{t}$ remaining fixed and $S_{TT} \to \infty$, then $f \approx 1$. Thus

$$C_\pi + D_\pi^T X = \hat{C} + \hat{D}^T X$$

and

$$MSE = MSE_\pi$$

To make a comparative study of Q_π and $\hat{Q}_S$, we calculate the 81 values of Q_π for the same values of the four invariants $(N, B_N, C_N, \hat{\rho}^2)$ as have already been used to calculate Q_S in section 5.5 for q = 1, 2, 3, 4 and 8. The values of Q_π are given in table 7.1 and the values of $\hat{Q}_S$ are given in table 7.2 for q = 1, 2, 3, 4.

One question the tables may help to answer is whether Q_π, for which there is a simple formula (7.18), is a good approximation to $\hat{Q}_S$, the non-Bayesian quantity which has had to be simulated.

In general Q_π and $\hat{Q}_S$ may be quite different as the tables 7.1 and 7.2 show, but we note that in the favourable case with $\hat{\rho}^2 = 0.9$, N = 50 (for instance);

$$Q_\pi - 0.057 < \hat{Q}_S < Q_\pi - 0.030$$

for all the 3x3x4 combinations of B_N, C_N and q.

The tabulated values are such that

(i) $\hat{Q}_S < Q_\pi$ for $\hat{\rho}^2 = 0.9$;

(ii) $\hat{Q}_S < Q_\pi$ for $\hat{\rho}^2 = 0.8$ and q = 1,2;

(iii) $\hat{Q}_S < Q_\pi$ for $\hat{\rho}^2 = 0.7$ and q = 1.

7.4 INTERVAL ESTIMATES

In section 7.3 we derived the Bayesian linear predictor $C_\pi + D_\pi^T X$ with mean squared error, conditional on the estimates $\hat{\alpha}$, $\hat{\beta}$, $\hat{\Gamma}$ from the experiment as,

$$MSE_\pi = \sigma^2(1 - \hat{\rho}^2)(1 + Q_\pi)$$

This leads us to propose

$$C_\pi + D_\pi^T X \pm 1.96\ (MSE_\pi)^{\frac{1}{2}}$$

as an interval for T, in the hope that its Bayesian confidence $\pi[\text{interval contains } T | \hat{\alpha}, \hat{\beta}, \hat{\Gamma}]$ is roughly 0.95. This confidence probability will depend on the shape of $\pi(t)$.

The inequalities $\hat{Q}_S < Q_\pi$ noted at the end of section 7.3 show that this Bayesian interval will often be wider than the interval

$$\hat{C} + \hat{D}^T X \pm 1.96\ (\hat{MSE})^{\frac{1}{2}}$$

proposed in chapter 6, since $\hat{MSE} = (1 - \hat{\rho}^2)(1 + \hat{Q}_S)$. However, this comparison of widths does not indicate which interval is to be preferred, for they aim to satisfy different criteria.

We end by recalling the different ways in which the linear predictors are defined:

(a) $\hat{C}$, $\hat{D}$ are estimates of C, D minimising

$$E[\{T - (C + D^T X)\}^2 | \alpha, \beta, \Gamma]$$

(b) C_π, D_π minimise

$$E[\{T - (C_\pi + D_\pi^T X)\}^2 | \hat{\alpha}, \hat{\beta}, \hat{\Gamma}]$$

so they are the functions of $\hat{\alpha}$, $\hat{\beta}$, $\hat{\Gamma}$ minimising

$$E[\{T - (C_\pi + D_\pi^T X)\}^2]$$

which can be written as

$$\int E[\{T - (C_\pi + D_\pi^T X)\}^2 | \alpha, \beta, \Gamma]\ \pi(\alpha, \beta, \Gamma)\, d\alpha d\beta d\Gamma$$

In (b) we ignore any difficulties due to the priors $\pi(\alpha, \beta, \Gamma)$ used in sections 7.2 and 7.3 being improper.

TABLE 7.1 81 values of Q_π for q = 1, 2, 3 and 4.

C_N	q = 1			q = 2			q = 3			q = 4		
	0.25	1.0	4.0	0.25	1.0	4.0	0.25	1.0	4.0	0.25	1.0	4.0
					$\hat{\rho}^2 = 0.7$ N = 10							
B_N												
0	0.309	0.373	0.592	0.456	0.522	0.744	0.641	0.707	0.928	0.880	0.944	1.155
1	0.393	0.452	0.655	0.543	0.602	0.806	0.728	0.788	0.988	0.965	1.022	1.212
4	0.608	0.655	0.818	0.760	0.806	0.967	0.943	0.988	1.143	1.170	1.212	1.353
					$\hat{\rho}^2 = 0.7$ N = 30							
0	0.083	0.101	0.173	0.112	0.130	0.203	0.142	0.161	0.234	0.174	0.194	0.268
1	0.107	0.126	0.196	0.137	0.155	0.226	0.167	0.186	0.258	0.200	0.219	0.291
4	0.179	0.196	0.261	0.209	0.226	0.292	0.240	0.258	0.325	0.274	0.291	0.359
					$\hat{\rho}^2 = 0.7$ N = 50							
0	0.048	0.059	0.101	0.064	0.075	0.118	0.080	0.091	0.134	0.097	0.108	0.152
1	0.062	0.073	0.115	0.078	0.089	0.132	0.095	0.106	0.148	0.112	0.123	0.166
4	0.105	0.115	0.156	0.121	0.132	0.172	0.138	0.148	0.189	0.155	0.166	0.207
					$\hat{\rho}^2 = 0.8$ N = 10							
0	0.369	0.450	0.739	0.558	0.644	0.951	0.806	0.899	1.222	1.150	1.248	1.581
1	0.476	0.552	0.825	0.672	0.753	1.041	0.929	1.015	1.315	1.279	1.368	1.674
4	0.761	0.825	1.059	0.974	1.041	1.283	1.246	1.315	1.561	1.605	1.674	1.917
					$\hat{\rho}^2 = 0.8$ N = 30							
0	0.096	0.117	0.203	0.130	0.152	0.239	0.166	0.188	0.277	0.204	0.227	0.318
1	0.125	0.146	0.230	0.159	0.181	0.267	0.196	0.218	0.306	0.235	0.258	0.347
4	0.209	0.230	0.310	0.246	0.267	0.349	0.284	0.306	0.389	0.325	0.347	0.432
					$\hat{\rho}^2 = 0.8$ N = 50							
0	0.055	0.068	0.117	0.073	0.086	0.137	0.093	0.106	0.157	0.112	0.125	0.177
1	0.072	0.084	0.134	0.091	0.103	0.153	0.110	0.123	0.173	0.130	0.143	0.194
4	0.122	0.134	0.182	0.141	0.153	0.202	0.161	0.173	0.223	0.181	0.194	0.244
					$\hat{\rho}^2 = 0.9$ N = 10							
0	0.435	0.536	0.916	0.674	0.788	1.215	1.009	1.140	1.623	1.510	1.663	2.216
1	0.569	0.667	1.035	0.826	0.935	1.347	1.182	1.308	1.771	1.713	1.857	2.382
4	0.946	1.035	1.373	1.248	1.347	1.718	1.661	1.771	2.182	2.258	2.382	2.836
					$\hat{\rho}^2 = 0.9$ N = 30							
0	0.109	0.134	0.234	0.148	0.174	0.277	0.190	0.217	0.323	0.236	0.263	0.373
1	0.143	0.168	0.266	0.183	0.209	0.311	0.226	0.253	0.358	0.273	0.300	0.408
4	0.242	0.266	0.363	0.285	0.311	0.410	0.332	0.358	0.460	0.382	0.408	0.514
					$\hat{\rho}^2 = 0.9$ N = 50							
0	0.062	0.077	0.134	0.083	0.098	0.156	0.105	0.120	0.180	0.128	0.143	0.204
1	0.082	0.096	0.153	0.103	0.118	0.176	0.125	0.140	0.199	0.148	0.164	0.224
4	0.139	0.153	0.209	0.161	0.176	0.233	0.185	0.199	0.258	0.209	0.224	0.283

TABLE 7.2 81 values of $\hat{Q}_s$ for q = 1, 2, 3 and 4.

	q = 1			q = 2			q = 3			q = 4		
C_N	0.25	1.0	4.0	0.25	1.0	4.0	0.25	1.0	4.0	0.25	1.0	4.0
					$\hat{\rho}^2 = 0.7$ N = 10							
B_N												
0	0.134	0.163	0.344	0.308	0.377	0.711	0.560	0.662	1.108	1.015	1.110	1.562
1	0.232	0.258	0.415	0.426	0.487	0.802	0.709	0.800	1.239	1.212	1.298	1.746
4	0.521	0.542	0.623	0.773	0.809	1.069	1.158	1.221	1.635	1.804	1.862	2.294
					$\hat{\rho}^2 = 0.7$ N = 30							
0	0.040	0.045	0.061	0.072	0.086	0.142	0.110	0.131	0.224	0.152	0.181	0.304
1	0.066	0.071	0.085	0.099	0.112	0.165	0.138	0.159	0.250	0.183	0.212	0.335
4	0.142	0.147	0.156	0.179	0.190	0.235	0.222	0.242	0.327	0.276	0.304	0.426
					$\hat{\rho}^2 = 0.7$ N = 50							
0	0.024	0.027	0.036	0.042	0.049	0.081	0.060	0.072	0.124	0.082	0.099	0.170
1	0.038	0.041	0.050	0.057	0.065	0.095	0.077	0.088	0.139	0.098	0.115	0.186
4	0.082	0.085	0.092	0.104	0.110	0.137	0.125	0.136	0.183	0.148	0.165	0.235
					$\hat{\rho}^2 = 0.8$ N = 10							
0	0.136	0.186	0.376	0.313	0.397	0.714	0.567	0.683	1.140	1.012	1.131	1.652
1	0.245	0.301	0.483	0.443	0.525	0.826	0.725	0.836	1.277	1.216	1.328	1.841
4	0.568	0.644	0.803	0.824	0.902	1.150	1.203	1.304	1.694	1.829	1.918	2.401
					$\hat{\rho}^2 = 0.8$ N = 30							
0	0.041	0.052	0.091	0.075	0.092	0.157	0.116	0.138	0.227	0.159	0.188	0.300
1	0.071	0.082	0.122	0.106	0.122	0.186	0.147	0.169	0.257	0.194	0.222	0.334
4	0.158	0.170	0.214	0.196	0.212	0.271	0.242	0.263	0.345	0.296	0.324	0.433
					$\hat{\rho}^2 = 0.8$ N = 50							
0	0.025	0.031	0.053	0.044	0.054	0.091	0.064	0.077	0.127	0.087	0.103	0.167
1	0.041	0.047	0.071	0.062	0.071	0.108	0.082	0.095	0.144	0.105	0.121	0.185
4	0.091	0.097	0.122	0.115	0.124	0.159	0.137	0.149	0.197	0.161	0.177	0.239
					$\hat{\rho}^2 = 0.9$ N = 10							
0	0.135	0.209	0.554	0.317	0.416	0.805	0.572	0.698	1.195	1.006	1.153	1.712
1	0.255	0.335	0.711	0.458	0.559	0.951	0.739	0.864	1.355	1.219	1.359	1.909
4	0.609	0.712	1.178	0.869	0.978	1.383	1.243	1.370	1.848	1.854	1.978	2.502
					$\hat{\rho}^2 = 0.9$ N = 30							
0	0.042	0.060	0.135	0.078	0.100	0.185	0.121	0.146	0.244	0.166	0.195	0.309
1	0.075	0.093	0.170	0.112	0.134	0.219	0.155	0.180	0.278	0.204	0.233	0.347
4	0.172	0.192	0.276	0.213	0.234	0.320	0.260	0.287	0.382	0.316	0.345	0.458
					$\hat{\rho}^2 = 0.9$ N = 50							
0	0.025	0.035	0.077	0.046	0.059	0.109	0.067	0.081	0.137	0.092	0.108	0.172
1	0.044	0.054	0.097	0.066	0.078	0.128	0.088	0.102	0.158	0.112	0.128	0.192
4	0.099	0.110	0.155	0.125	0.138	0.188	0.149	0.162	0.218	0.174	0.190	0.253

APPENDIX A

The asymptotic expressions $E\hat{C}$, $E\hat{D}$, $VAR(\hat{C})$, $VAR(\hat{D})$, $COV(\hat{C},\hat{D})$ are obtained by expanding the estimate by Taylor's series and these are used to calculate bias and mean squared error of the best linear predictor. The following results are based on a sample mean $\bar{x}_f$ of k values in the future situation and the results for a single value x can be achieved by putting k = 1.

A.1 Some facts

(1) $E(\bar{X}_f) = \alpha + \beta E(T)$

$= \alpha^* + \beta(\mu - \bar{t})$

(2) $VAR(\bar{X}_f) = E(VAR(\bar{X}_f|T)) + VAR(E(\bar{X}_f|T))$

(3) $\sigma^2_{\bar{x}|t} = (1 - \rho^2_{\bar{x}})VAR(\bar{X}_f)$

From (1), (2), (3) we can easily derive that

(a) $$(\bar{t} - \mu) = \frac{\alpha^*}{\beta} - \frac{E(\bar{X}_f)}{\beta}$$

(b) $$VAR(\bar{X}_f) = \frac{\sigma^2_{x|t}}{k} + \beta^2\sigma^2$$

(c) $$\frac{(E\bar{X}_f - \alpha^*)^2}{VAR(\bar{x}_f)} = \frac{\rho^2_{\bar{x}}(\mu - \bar{t})^2}{\sigma^2}$$

(4) $$\frac{1 - \rho^2}{k\rho^2} = \frac{1 - \rho^2_{\bar{x}}}{\rho^2_{\bar{x}}} = \frac{\sigma^2_{x|t}}{k\beta^2\sigma^2}$$

(5) $$\frac{D^2\sigma^2_{x|t}}{k} = \rho^2_{\bar{x}}(1 - \rho^2_{\bar{x}})\ \sigma^2$$

(6) $$\beta^2 = \frac{\rho^2_{\bar{x}}\ VAR(\bar{X}_f)}{\sigma^2}$$ (from (3) and (b) above)

(7) $$(1 - \rho^2_{\bar{x}})VAR(T) = \frac{\frac{\sigma^2_{x|t}}{k}}{\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}}$$

$$(8)\quad \left[\frac{-\beta^2 + \frac{\sigma^2_{x|t}}{\sigma^2}}{\beta^2 + \frac{\sigma^2_{x|t}}{\sigma^2}}\right]^2 = \left[1 - \frac{2}{1 + \frac{\sigma^2_{x|t}}{\beta^2\sigma^2}}\right]^2$$

A.2 *$E(\hat{C})$, $E(\hat{D})$, $VAR(\hat{C})$, $VAR(\hat{D})$, $COV(\hat{C},\hat{D})$*

$$\hat{C} = \frac{-\hat{\alpha}^*\hat{\beta} + (\bar{t}-\mu)\hat{\beta}^2}{\hat{\beta}^2 + \frac{\hat{\sigma}^2_{x|t}}{k\sigma^2}} + \mu$$

$$\sigma^2_{\hat{\sigma}^2_{x|t}} = VAR(\hat{\sigma}^2_{x|t})$$

$$\sigma^2_{\alpha^*} = VAR(\hat{\alpha}^*)$$

$$\sigma^2_{\beta} = VAR(\hat{\beta})$$

$E\hat{C} = E\ h(\hat{\beta},\ \hat{\sigma}^2_{x|t},\ \hat{\alpha}^*)$, where the three arguments are independent random variables with expectations $\underline{M} = (\beta,\ \sigma^2_{x|t},\ \alpha^*)$.

$$E\hat{C} \approx h(\beta,\ \sigma^2_{x|t},\ \alpha^*) + \frac{\sigma^2_{\beta}}{2}\ \frac{\partial^2 h}{\partial\beta^2}\Big|_{\underline{M}} + \frac{\sigma^2_{\hat{\sigma}^2_{x|t}}}{2}\ \frac{\partial^2 h}{\partial\sigma^2_{x|t}}\Big|_{\underline{M}} + \frac{\sigma^2_{\alpha^*}}{2}\ \frac{\partial^2 h}{\partial\alpha^{*2}}\Big|_{\underline{M}}$$

$$= \mu + \frac{-\alpha^*\beta + (\bar{t}-\mu)\beta^2}{\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}}\left[1 + \frac{2\sigma^4_{x|t}}{(N+k-3)k^2\sigma^4\left[\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\right]^2}\right]$$

$$+ \frac{\sigma^2_{x|t}}{\left[\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\right]^3}\left[\frac{\alpha^*\beta\left[-\beta^2 + \frac{3\sigma^2_{x|t}}{k\sigma^2}\right] + \frac{\sigma^2_{x|t}}{k\sigma^2}(\bar{t}-\mu)\left[\frac{\sigma^2_{x|t}}{k\sigma^2} - 3\beta^2\right]}{S_{TT}}\right]$$

or

$$= \mu + \left[-\frac{\alpha^*}{\beta} + (\bar{t}-\mu)\right]\rho^2_{\bar{x}}\left[1 + \frac{2(1-\rho^2_{\bar{x}})^2}{N+k-3}\right] + \frac{(1-\rho^2_{\bar{x}})}{S_{TT}}\ k\sigma^2\left[\frac{\alpha^*}{\beta}\ \rho^2_{\bar{x}}(3-4\rho^2_{\bar{x}})\right.$$

$$\left. + (1-\rho^2_{\bar{x}})(\bar{t}-\mu)(1-4\rho^2_{\bar{x}})\right]$$

$$\underline{VAR(\hat{C})} = VAR\ h(\hat{\beta},\ \hat{\sigma}^2_{x|t},\ \hat{\alpha}^*) \approx \left[\frac{\partial h}{\partial \alpha^*}\right]^2 \sigma^2_{\alpha^*} + \left[\frac{\partial h}{\partial \beta}\right]^2 \sigma^2_{\beta} + \left[\frac{\partial h}{\partial \sigma^2_{x|t}}\right]^2 \sigma^2_{\sigma^2_{x|t}}$$

$$= \frac{\sigma^2_{x|t}}{\left[\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\right]^2} \left[\frac{\beta^2}{N} + \frac{\left[\alpha^*\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}(2\beta(\bar{t}-\mu) - \alpha^*)\right]^2}{S_{TT}\left[\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\right]^2} + \frac{2\sigma^2_{x|t}}{N+k-3} \frac{\beta^2\left[\alpha^* - (\bar{t}-\mu)\beta\right]^2}{k^2\sigma^4\left[\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\right]^2}\right]$$

$$D = \frac{\beta^{-1}}{1 + \frac{\sigma^2_{x|t}}{k\sigma^2}}$$

$$\underline{N} = (\beta,\ \sigma^2_{x|t})$$

$$\hat{D} = g(\hat{\beta}\ ,\hat{\sigma}^2_{x|t})$$

$$\underline{E\hat{D}} = Eg(\hat{\beta},\ \hat{\sigma}^2_{x|t})$$

$$\approx g(\beta\ ,\ \sigma^2_{x|t}) + \frac{\sigma^2_{\beta}}{2}\frac{\partial^2 g}{\partial \beta^2}\Big|_{\underline{N}} + COV(\hat{\beta},\ \hat{\sigma}^2_{x|t})\frac{\partial^2 g}{\partial\beta\partial\sigma^2_{x|t}}\Big|_{\underline{N}} + 1/2\ \sigma^2_{\sigma^2_{x|t}}\frac{\partial^2 g}{\partial\ (\sigma^2_{x|t})^2}\Big|_{\underline{N}}$$

$$= \frac{\beta}{\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}}\left[1 - \frac{\sigma^2_{x|t}}{S_{TT}}\frac{\left[-\beta^2 + \frac{3\sigma^2_{x|t}}{k\sigma^2}\right]}{\left[\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\right]^2} + \frac{2\ \sigma^4_{x|t}}{(N+k-3)k\sigma^4\left[\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\right]^2}\right]$$

$$\underline{VAR(\hat{D})} = VAR\ g(\hat{\beta},\ \hat{\sigma}^2_{x|t}) \approx \left[\frac{\partial g}{\partial \beta}\right]^2 \sigma^2_{\beta} + \left[\frac{\partial g}{\partial \sigma^2_{x|t}}\right]^2 \sigma^2_{\sigma^2_{x|t}}$$

$$+ 2\left[\frac{\partial g}{\partial \beta}\right]\left[\frac{\partial g}{\partial \sigma^2_{x|t}}\right] COV(\hat{\beta},\ \hat{\sigma}^2_{x|t})$$

$$= \frac{\sigma^2_{x|t}}{\left[\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\right]^4}\left[\frac{\left[-\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\right]^2}{S_{TT}} + \frac{2\beta^2\sigma^2_{x|t}}{k^2\sigma^4(N+k-3)}\right]$$

$$\underline{COV(\hat{C},\ \hat{D})} \approx \sigma^2\beta\left[\frac{\partial g}{\partial \beta}\right]\left[\frac{\partial h}{\partial \beta}\right] + \sigma^2\sigma^2_{x|t}\left[\frac{\partial g}{\partial \sigma^2_{x|t}}\right]\left[\frac{\partial h}{\partial \sigma^2_{x|t}}\right]$$

$$\approx \frac{\sigma^2_{x|t}}{\left[\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\right]^4}\left[\frac{\left[-\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\right]\left[\alpha^*\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\{2\beta(\bar{t}-\mu)-\alpha^*\}\right]}{S_{TT}}\right.$$

$$\left. + \frac{2\sigma^2_{x|t}}{(N+k-3)}\frac{\beta}{k^2\sigma^4}\left[-\alpha^*\beta + (\bar{t}-\mu)\beta^2\right]\right]$$

A.3 Bias and Mean squared error of the best linear predictor

Bias $= B_T = E(T - (\hat{C} + \hat{D}\bar{X}_f)$

$= \mu - E\hat{C} - E\hat{D}(\alpha + \beta\mu)$

$= \mu - E\hat{C} - E\hat{D}[\alpha^* - \beta(\bar{t} - \mu)]$

substituting $E\hat{C}$, $E\hat{D}$ and simplifying we get

$$\approx (\bar{t} - \mu)\left[1 - \rho_{\bar{X}}^2\right]\left[2\rho_{\bar{X}}^2 - 1\right] \frac{k\sigma^2}{S_{TT}}$$

$$= \pm \sigma \left[\frac{k}{N-2} (C_N)^{\frac{1}{2}}(B_N)^{\frac{1}{2}}\left[1 - \rho_{\bar{X}}^2\right]\left[2\rho_{\bar{X}}^2 - 1\right]\right] \qquad \ldots(A.1)$$

Mean squared error

$$MSE = E[T - (\hat{C} + \hat{D}\bar{X}_f)]^2$$

$$= EE[(T - (\hat{C} + \hat{D}\bar{X}_f))^2 \mid \hat{C}, \hat{D}]$$

$$= E(\hat{C} - C)^2 + 2E(\bar{X}_f)E(\hat{C} - C)(\hat{D} - D) + (E\bar{X}_f^2)E(\hat{D} - D)^2$$

$$+(1 - \rho_{\bar{X}}^2)VAR(T)$$

plugging in values of $VAR(\hat{C})$, $VAR(\hat{D})$ and $COV(\hat{C},\hat{D})$, ignoring (bias)^2 being of order $(1/N)^2$

$$\approx \left[\frac{\sigma_{x|t}^2 \beta^2}{\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}}\right]^2 \left[\frac{1}{N} + \frac{2\sigma_{x|t}^2}{k^2\sigma^4(N+k-3)\left[\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\right]^2} E\{\bar{X}_f - E(\bar{X}_f)\}^2\right]$$

$$+ \frac{\sigma_{x|t}^2}{S_{TT}\left[\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\right]^4}\left[E\left\{\left[-\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\right]\bar{X}_f + \left[\alpha^*\beta^2 + \frac{\sigma^2_{x|t}}{k\sigma^2}\{-2E\bar{X}_f + \alpha^*\}\right]\right\}^2\right]$$

$$+ \left[1 - \rho_{\bar{X}}^2\right]VAR(T)$$

$$= \frac{\frac{\sigma^2_{x|t}}{\beta^2}}{1 + \frac{\sigma^2_{x|t}}{k\beta^2\sigma^2}} \left[\frac{1}{k} + \frac{1}{N\left[1 + \frac{\sigma^2_{x|t}}{k\beta^2\sigma^2}\right]} + \frac{2\sigma^2_{x|t}}{k\sigma^2(N+k-3)\beta^2\left[1 + \frac{\sigma^2_{x|t}}{k\beta^2\sigma^2}\right]^2} \right.$$

$$\left. + \frac{\sigma^2}{S_{TT}} \left[1 - \frac{2}{1 + \frac{\sigma^2_{x|t}}{k\beta^2\sigma^2}}\right]^2 + \frac{(\bar{t} - \mu)^2}{S_{TT}\left[1 + \frac{\sigma^2_{x|t}}{k\beta^2\sigma^2}\right]} \right]$$

$$= \sigma^2\left[1 - \rho^2_{\bar{x}}\right] \left[1 + \frac{k}{N\left[1 + \tau^2_{\bar{x}}\right]} + \frac{2\tau^2_{\bar{x}}}{(N+k-3)\left[1 + \tau^2_{\bar{x}}\right]^2} + \frac{k(C_N)}{N-2}\left[1 - \frac{2}{1 + \tau^2_{\bar{x}}}\right]^2 \right.$$

$$\left. + \frac{k(B_N)}{(N-2)\left[1 + \tau^2_{\bar{x}}\right]} \right] \quad \ldots(A.2)$$

where

C_N = Relative concentration of the experiment = $\frac{\sigma^2(N-2)}{S_{TT}}$

B_N = Relative bias of the experiment = $\frac{(\bar{t} - \mu)^2(N - 2)}{S_{TT}}$

and

$$\tau^2_{\bar{x}} = \frac{1 - \rho^2_{\bar{x}}}{\rho^2_{\bar{x}}} = \frac{\sigma^2_{x|t}}{k\beta^2\sigma^2}$$

Equation (A.2) can be written as

$$= \sigma^2\left[1 - \rho^2_{\bar{x}}\right]\left[1 + Q_A\right]$$

where Q_A is the quantity in the big brackets to the right of 1 in (A.2). Since $\rho^2_{\bar{x}} = 1/(1 + \tau^2_{\bar{x}})$, Q_A is a quadratic in $\rho^2_{\bar{x}}$, with coefficients depending on k, N, B_N, C_N.

APPENDIX B

Two results concerning the best linear predictor of section 5.3, for p = 1 and general q.

B.1

When $\sigma^2 \to \infty$, Best Linear Predictor → Classical Estimator.

Proof:

Using lemma 7.1 with $A = \hat{\Gamma}$, $k = \sigma^2$, $b = \hat{\beta}$,

$$\hat{D} = \sigma^2\hat{\Gamma}^{-1}\hat{\beta}/[1 + \sigma^2\hat{\beta}^T\hat{\Gamma}^{-1}\hat{\beta}].$$

Thus

$$\hat{\rho}^2 = \hat{D}^T\hat{\beta} \to 1 \quad \text{as} \quad \sigma^2 \to \infty.$$

Equation (5.11) implies

$$\begin{aligned}\hat{C} + \hat{D}^T X &= \mu\ (1 - \hat{\rho}^2) + \hat{D}^T(X - \hat{\alpha}) \\ &= \mu\ (1 - \hat{\rho}^2) + \hat{D}^T(X - \hat{\alpha}^* + \hat{\beta}\bar{t}) \\ &= \mu\ (1 - \hat{\rho}^2) + \hat{\rho}^2\bar{t} + \hat{D}^T(X - \bar{X})\end{aligned}$$

Since $\hat{D} \to \hat{\Gamma}^{-1}\hat{\beta}/[\hat{\beta}^T\hat{\Gamma}^{-1}\hat{\beta}]$ as $\sigma^2 \to \infty$,

$$\hat{C} + \hat{D}^T X \to \bar{t} + \hat{\beta}^T\hat{\Gamma}^{-1}(X - \bar{X})/[\hat{\beta}^T\hat{\Gamma}^{-1}\hat{\beta}]$$

which is the classical estimator (5.4) when p = 1.

B.2

When $\mu = \bar{t}$ and $\sigma^2 = S_{TT}/(N-2)$, the Best Linear Predictor for general q and p = 1 is the Inverse Estimator.

Proof:

We use the standard results from multivariate regression, derived from relation (5.2) when p = 1,

$$\hat{\beta} = S^T_{TX}/S_{TT}$$

$$(N-2)\hat{\Gamma} = S_{XX} - S^T_{TX}\ S_{TT}^{-1}\ S_{TX}$$

thus

$$\sigma^2\hat{\beta} = S^T_{TX}/(N-2) \quad \text{when} \quad \sigma^2 = S_{TT}/(N-2)$$

and

$$(N-2)[\hat{\Gamma} + \sigma^2\hat{\beta}\hat{\beta}^T] = (N-2)\hat{\Gamma} + S^T_{TX}\, S_{TX}/S_{TT} - S_{XX}$$

so

$$\begin{aligned}\hat{D} &= \sigma^2[\hat{\Gamma} + \sigma^2\hat{\beta}\hat{\beta}^T]^{-1}\,\hat{\beta} \quad \text{from (5.10)} \\ &= S_{TT}\, S_{XX}^{-1}\, S^T_{TX}/S_{TT} \\ &= S_{XX}^{-1}\, S^T_{TX}\end{aligned}$$

in agreement with (5.5).

Also when $\mu = \bar{t}$

$$\begin{aligned}\hat{C} &= \mu - \hat{D}^T(\hat{\alpha} + \mu\hat{\beta}) \qquad \text{from (5.8)} \\ &= \bar{t} - \hat{D}^T\hat{\alpha}^* \\ &= \bar{t} - \hat{D}^T\bar{X}\end{aligned}$$

in agreement with relation (5.5) when p = 1.

REFERENCES

Aitchison, J. (1977). A calibration problem in statistical diagnosis. Biometrika 64(3):461-472.

Aitchison, J. and Dunsmore, I.D. (1975). Statistical Prediction Analysis. Cambridge University Press, Cambridge.

Anderson, T.W. (1984). An Introduction to Multivariate Statistical Analysis. 2nd edition. John Wiley and Sons, Inc. New York. pp. 247, 291.

Andrews, D.F. and Herzberg, A.M. (1973). A simple method for constructing exact tests for sequentially designed experiments. Biometrika 60(3):489-497.

Ansley, C.F. and Wecker, W.E. (1984). A non-parametric Bayesian approach to the calibration problem. Proceed.Business Economics section. Amer.Statist.Assoc. pp. 96-101.

Barnett, V.D. (1969). Simultaneous pairwise linear structural relationships. Biometrics 25:129-142.

Berkson, J. (1969). Estimation of a linear function for a calibration line; consideration of a recent proposal. Technometrics 11(4):649-660.

Bermúdez, J.D. and Bernardo, J.M. (1985). Bayesian testing of a calibration procedure. Proceed.Intern.Statist.Institute. 45th session, contributed papers. 12-22 August, Amsterdam. pp. 119-120.

Box, G.E.P. and Cox, D.R. (1964). An analysis of transformations. J.Royal Statist.Soc. B, 26: 211-252.

Box, G.E.P. and Tidwell, P.W. (1962). Transformation of the independent variables. Technometrics 4(4):531-550.

Branco, J.A. (1985). A comparative calibration model for replicated observations. Proceed.Intern.Statist.institute. 45th session, contributed papers. 12-22 August, Amsterdam. pp. 101-102.

Brown, G.H. (1978). Calibration with an ultra-structural model. Appl.statist. 27(1):47-51.

Brown, G.H. (1979a). An optimization criterion for linear inverse estimation. Technometrics 21(4):575-579.

Brown, G.H. (1979b). A method for calibrating frequency distributions. Textile Research J. 49(2):101-104.

Brown, P.J. (1982). Multivariate calibration (with discussion). J.Royal Statist.Soc. B, 44(3):287-321.

Brown, P.J. and Sundberg, R. (1985). Multivariate calibration with more variables than observations. Institute of Actuarial Mathematics and Mathematical Statistics. Univ. of Stockholm, Sweden. Research Report No. 139.

Brown, P.J. and Sundberg, R. (1987). Confidence and conflict in multivariate calibration. J.Royal Statist.Soc. B, 49(1):46-57.

Buonaccorsi, J.P. (1986). Design considerations for calibration. Technometrics 28(2):149-155.

Carroll, R.J. and Spiegelman, C.H. (1986). The effect of ignoring small measurement errrors in precision instrument calibration. J.Quality Technology 18(3):170-173.

Carroll, R.J.; Spiegelman, C.H. and Waters, R. (1985). Calibration when both *X* and Y have error. Proceed.Intern.Statist.Institute. 45th session, contributed papers. 12-22 August, Amsterdam. pp. 559-560.

Clark, R.M. (1979). Calibration, cross-validation and carbon-14. *I*. J.Royal Statist.Soc. A, 142(1):47-62.

Clark, R.M. (1980). Calibration, cross-validation and carbon-14. *II*. J.Royal Statist.Soc. A, 143(2):177-194.

Copas, J.B. (1982). Discussion on paper by Brown. J.Royal.Statist. Soc. B, 44(3):312-313.

Cox, C.P. (1971). Interval estimation for *X*-predictions from linear

Y-on-X regression lines through the origin. J.Amer. Statist.Assoc. 66(366):749-751.

Creasy, M.A. (1954). Limits for the ratio of means. J.Royal Statist. Soc. B, 16:184-194.

Currie, L.A. (1985). The limitations of models and measurements as revealed through chemometric intercomparison. J.Research. National Bureau of Standards 90(6):409-422.

Davis, W.W. and DeGroot, M.H. (1982). A look at Bayesian prediction and calibration. Statistical Decision Theory and Related Topics III, Volume 1. Academic Press London. pp.271-289.

Dawid, A.P. (1985). Calibration based empirical probability. The Annals of Statistics 13(4):1251-1273.

Dietrich, F.J. and Marks, R.G. (1979). Analysis of factorial quantal response assay using inverse regression. Comm.Statist.Theory and Methods. A, 8(1):85-98.

Draper, N.R. and Smith, H. (1981). Applied Regression Analysis, 2nd edition, John Wiley and Sons, New York. pp. 47-51, 125.

Dunsmore, I.D. (1968). A Bayesian approach to calibration. J.Royal statist. Soc. B, 30:396-405.

Easterling, R.G. (1969). Discrimination intervals for percentiles in regression. J.Amer.Statist.Assoc. 64:1031-1041.

Eisenhart, C. (1939). The interpolation of certain regression methods and their use in biological and industrial research. Annals of Mathematical Statistics 10:162-184.

Fieller, E.C. (1954). Some problems in interval estimation. J.Royal Statist.Soc. B, 16:175-185.

Ford, I. (1976). Optimal static and sequential designs: A critical review. Ph.D Thesis, Deartment of statistics, University of Glasgow, Glasgow. pp. 44-46.

Frazier, L.T. (1974). An analysis of a Bayes inverse regression

method of confidence intervals in linear calibration. J.Statist.Comput.Simul. 3:99-103.

Fujikoshi, Y. and Nishii, R. (1984). On the distribution of a statistic in multivariate inverse regression analysis. Hiroshima Math.J. 14:215-225.

Fujikoshi, Y. and Nishii, R. (1986). Selection of variables in multivariate inverse regression problem. Hiroshima Math.J. 16:269-277.

Grassia, A. and Sundberg, R. (1982). Statistical precision in the calibration and use of sorting machines and other classifiers. Technometrics 24(2):117-121.

Halperin, M. (1961). Fitting of straight lines and prediction when both variables are subject to error. J.Amer.Statist.Assoc. 56:657-669.

Halperin, M. (1970). On inverse estimation in linear regression. Technometrics 12(4):727-736.

Hartigan, J.A. (1969). Linear Bayes methods. J.Royal Statist.Soc.B, 31:446-454.

Heldal, J. and Spjøtvoll, E. (1985). Calibration of categorical variables in registers - A study of the 2 × 2 case. Proceed.Intern.Statist.Institute. 45th session contributed papers, 12-22 August, Amsterdam. pp. 311-312.

Henschke, K. (1980). Simultaneous confidence procedures in multivariate calibration problems. Maths.Operationsforsch. Statist. Ser.Statistics 11(2):193-206.

Herson, J. (1975). Fieller's theorem vs. the Delta method for significance intervals for ratios. J.statist.Comput.Simul. 3:265-274.

Hoadley, B. (1970). A Bayesian look at inverse linear regression. J.Amer.Statist.Assoc. 65(329):356-369.

Hochberg, Y., Marom, I., Keret, R. and Peleg, S. (1983). On improved calibrations of unknowns in a system of quality controlled assays. Biometrics 39: 97-108.

Hunter, W.G. and Lamboy, W.F. (1981). A Bayesian analysis of the linear calibration problem (with discussion). Technometrics 23(4):323-350.

Jansen, A.A.M. (1980). Comparative calibration and congeneric measurements. Biometrics 36:729-734.

Kalotay, A.J. (1971). Structural solution to the linear calibration problem. Technometrics 13(4):761-769.

Knafl, G., Spiegelman, C.H., Sacks, J. and Yivisaker, D. (1984). Non-parametric calibration. Technometrics 26(3):233-241.

Krutchkoff, R.G. (1967). Classical and inverse regression methods of calibration. Technometrics 9(3):425-439.

Krutchkoff, R.G. (1969a). Classical and inverse regression methods of calibration in extrapolation. Technometrics 11(3):605-608.

Krutchkoff, R.G. (1969b). Letter to editor. Technometrics 12(2):433-434.

Krutchkoff, R.G. (1971). The calibration problem and closeness. J.Statist.Comput.Simul. 1:87-95.

Kurtz, D.A. (1983). The use of regression and statistical methods to establish calibration graphs in chromatography. Anal.Chim. 150:105-114.

Leary, J.J. and Messick, E.B. (1985). Constrained calibration curves: A novel application of Lagrange multipliers in analytical chemistry. Anal.Chem. 57:956-957.

Lee, S.H. and Yum, B.J. (1985). Choice of statistical calibration procedures when the standard measurement is also subject to error. J.Korean Statist. Soc. 14(2):63-75.

Lechner, J.A., Reeve, C.P. and Spiegelman, C.H. (1980). A new method

for assigning uncertainty in volume calibration. National Bureau of Standards. Washington, D.C., 108p. Report No. NBSIR-80-2151.

Lechner, J.A., Reeve, C.P. and Spiegelman, C.H. (1982). An implementation of the Scheffé approach to calibration using spline functions, illustrated by a pressure-volume calibration. Technometrics 24(3):229-234.

Lieberman, G.J., and Miller,JR. R.G. (1963). Simultaneous tolerance intervals in regression. Biometrika 50:155-168.

Lieberman, G.J., Miller, JR. R.G. and Hamilton, M.A. (1967). Unlimited simultaneous discrimination intervals in regression. Biometrika 54(1-2):133-145.

Lindley, D.V. (1972). Bayesian Statistics: A Review. Society for industrial and applied mathematics, Philadelphia.

Lundberg, E. and DeMaré, J. (1980). Interval estimates in the spectroscopy calibration problem. Scand.J.Statist. 7:40-42.

Lwin, T. (1985). Calibration with supplementary information. Proceed. Intern.Statist.Institute, 45th session contributed papers, 12-22 August, Amsterdam. pp. 65-66.

Lwin, T. and Maritz, J.S. (1980). A note on the problem of statistical calibration. Appl.Statist. 29(2):135-141.

Lwin, T. and Maritz, J.S. (1982). An analysis of the linear-calibration controversy from the perspective of compound estimation. Technometrics 24(3):235-242.

Lwin, T. and Spiegelman, C.H. (1986). Calibration with working standards. Appl.Statist. 35(3):256-261.

Makowski, G.G. and Downing, D. (1980). Confidence interval calibration of a standard curve. Technometrics 22(3):381-388.

Mandel, J. (1958). A note on confidence intervals in regression problems. Annals of Math.Statist. 29:903-907.

Mandel, J. (1984). Fitting straight lines when both variables are

subject to error. J.Quality Technology 16(1):1-14.

Martens, H. and Naes, T. (1984). Multivariate calibration. *I.* Trends in Analytical Chemistry 3(8):204-210.

Martinelle, S. (1970). On the choice of regression in linear calibration; comments on a paper by R.G.Krutchkoff. Technometrics 12(1): 157-161.

Mckeon, J.J. and Chhikara, R.S. (1985). Linear regression estimates in sample surveys under calibration. Proceed.Survey Research Methods. Amer.Statist.Assoc. pp 286-290.

Miller,JR. R.G, (1966). Simultaneous Statistical Inference. McGraw-Hill Book Co. New York.

Minder C.E. and Whitney, J.B. (1975). A likelihood analysis of the linear calibration problem. Technometrics 17(4):463-471.

Morris, A.G.G. (1983). Some comments on calibration procedures. Analyst 108:546-548.

Muhammad, F. and McLaren, A.D. (1985). An approach to linear calibration. Proceed.Intern.Statist.Institute. 45th session contributed papers, 12-22 August, Amsterdam. pp. 15-16.

Naes, T. (1985a). Comparison of approaches to multivariate linear calibration. Biometrical J. 27(3):265-275.

Naes, T. (1985b). Multivariate calibration when the error covariance matrix is structured. Technometrics 27(3):301-311.

Naes, T. (1986). Multivariate calibration using covariance adjustment. Biometrical J. 28:99-107.

Naes, T., Irgens, C. and Martens, H. (1986). Comparison of linear statistical methods for calibrations of NIR instruments. Appl.Statist 35(2):195-206.

Naes, T. and Martens, H. (1984). Mutivariate calibration. *II*. Trends in Analytical Chemistry 3(10):266-271.

Naszódi, L.L. (1978). Elimination of bias in the course of

calibration. Technometrics 20(2):201-205.

Nishii, R. (1986). Criteria for selection of response variables and the asymptotic properties in a multivariate calibration. Ann.Inst. Statist.Math. 38:319-329.

Odén, A. (1973). Simultaneous confidence intervals in inverse linear regression. Biometrika 60(2):339-343.

Oman, S.D. (1984). Analysing residuals in calibration problems. Technometrics 26(4):347-353.

Oman, S.D. (1985a). Confidence regions in multivariate calibration. Proceed.Intern.Statist.Institute, 45th session, contributed papers, 12-22 August, Amsterdam. pp. 617.

Oman, S.D. (1985b). An exact formula for mean squared error of the inverse estimator in linear calibration problem. J.Statist.Planning and Inference 11:189-196.

Oman, S.D. and Wax, Y. (1984). Estimating fetal age by ultra-sound measurements: An example of multivariate calibration. Biometrics 40:947-960.

Ott, R.L. and Myers, R.H. (1968). Optimal experimental design for estimating the independent variable in regression. Technometrics 10(4):811-823.

Pepper, M.P.G. (1973). A calibration of instruments with non-random errors. Technometrics 15(3):587-599.

Perng, S.K. and Tong, Y.L. (1974). A sequential solution to the inverse linear regression problem. The Annals of Statistics 2(3):535-539.

Perng, S.K. and Tong, Y.L. (1977). Optimal allocation of observations in inverse linear regression. The Annals of Statistics 5(1):191-196.

Press, S.J. (1972). Applied Multivariate Analysis. Holt, Rinehart and Winston, Inc. New York, pp. 232.

Rao, C.R. (1975). Some thoughts on regression and prediction. Sankhyā Ser. C, 87(2):102-120.

Rao, C.R. (1981). Some comments on the minimum mean square error as a criterion of estimation. Statistics and Related Topics. North-Holland publishing company, Holland. pp.123-143.

Reilman, M.A. and Gunst, R.F. (1986). Stochastic regression with errors in both variables. J.Quality Technology 18(3):162-169.

Rinco, S. and Chuiv, N.N. (1985). Multivariate calibration problem. Proceed.Intern.Statist.Institute, 45th session contributed papers, 12-22 August, Amsterdam. pp.95-96.

Rothman, D. (1968). Letter to editor. Technometrics 10(2):429-431.

Saw, J.G. (1970). Letter to editor. Technometrics 12(4):937.

Schaffrin, B. (1985). A calibration model for geodetic applications. Proceed.Intern.Statist.Institute, 45th session, contributed papers, 12-22 August, Amsterdam. pp. 523-525.

Scheffé, H. (1973). A statistical theory of calibration. The Annals of Statistics 1(1):1-37.

Schwartz, L.M. (1975). Random error propagation by monte-carlo simulation. Anal.Chem 47(6):963-964.

Schwartz, L.M. (1976). Non-linear calibration curves. Anal.Chem. 48(14):2287-2289.

Schwartz, L.M. (1977). Non-linear calibration. Anal.Chem. 49(13):2062-2068.

Schwartz, L.M. (1978). Statistical uncertainties of analyses by calibration of counting measurements. Anal.Chem. 50(7):980-984.

Schwartz, L.M. (1979). Calibration curves with non-uniform variance. Anal.Chem. 51(6):723-727.

Schwenke, J.R. and Milliken, G.A. (1983). On the calibration problem extended to non-linear models. Proceed.Biopharmaceutical section, Amer.Statist.Assoc. pp. 68-72.

Seber, G.A.F. (1977). Linear Regression Analysis. John Wiley and Sons, New York. pp. 187-191.

Shukla, G.K. (1972). On the problem of calibration. Technometrics 14(3):547-553.

Shukla, G.K. and Datta, P. (1985). Comparison of the inverse estimator with the classical estimator subject to a preliminary test in linear calibration. J.Statist.Planning and Inference 12(1):93-102.

Sinclair, S.D. (1982). A new approach to calibration intervals. Proceed. Intern. Assoc.for Statistical computing, Vienna. pp. 245-246.

Sjöström, M., Wold, S., Lindberg, W., Persson, J. and Martens, H. (1983). A multivariate calibration problem in analytical chemistry solved by partial least squares models in latent variables. Anal.Chim.Acta. 150:61-70.

Spezzaferri, F. (1985). A note on multivariate calibration experiments. Biometrics 41:267-272.

Spiegelmam, C.H. (1984). A new statistic for detecting influential observations in a Scheffé type calibration curve. Austral.J.Statist. 26(3):290-297.

Spiegelman, C.H. (1984). An iterative calibration curve procedure. J.Research, National Bureau of Standards, 89(2):187-192.

Spiegelman, C.H. and Studden, W.J. (1980). Design aspects of Scheffé calibration curve using linear splines. J.Research, National Bureau of Standards, 85(4):295-304.

Sundberg, R. (1985). When is the inverse regression estimator MSE-superior to the standard regression estimator in multivariate controlled calibration situations?. Statist.Prob.Letters 3:75-79.

Swallow, W.H. and Trout, J.R. (1983). Determination of limits for a

linear regression or calibration curve. J.Quality Technology 15(3):118-125.

Tallis, G.M. (1969). Note on a calibration problem. Biometrika 56(3):505-508.

Theobald, C.M. and Mallinson, J.R. (1978). Comparative calibration, linear structural relationships and congeneric measurements. Biometrics 34:39-45.

Thomas, M.A. and Myers, R.H. (1973). Optimal designs for the inverse regression method of calibration. Commu.Statist. 2(5):419-433.

Trout, J.R. and Swallow, W.H. (1979). Regular and inverse interval estimation of individual observations using uniform confidence bands. Technometrics 21(4):567-574.

Turiel, T.P., Hahn, G.H. and Tucker, W.T. (1982). New simulation results for the calibration and inverse median estimation problems. Commu.Statist.Simul.Comput. 11(6):672-713.

Williams, E.J. (1959). Regression Analysis. John Wiley and Sons Inc. New York.

Williams, E.J. (1969a). Regression methods in calibration problems. Bull.ISI. 43:17-28.

Williams, E.J. (1969b). A note on regression methods in calibration. Technometrics 11(1):189-192.

Williford, W.O., Carter, M.C. and Field, J.E. (1979). A further look at the Bayesian approach to calibration. J.Statist.Comput.Simul. 9:47-67.

Winslow, G.H. (1976). Some statistical aspects of the calibration and use of linear measuring systems. Nuclear Material Management 5:55-59.

Wood, J.T. (1982). Estimating the age of an animal: An application of multivariate calibration. Proceed.11th Intern.Biometric Conference. 6-11 september, Toulouse (France). pp. 117-121.

GLASGOW
UNIVERSITY
LIBRARY