INTERVAL ESTIMATION OF FUNCTIONS OF BERNOULLI PARAMETERS WITH RELIABILITY AND BIOMEDICAL APPLICATIONS*

by

Dar-Shong Hwang

Technical Report No. 152

March 1971

University of Minnesota
Minneapolis, Minnesota

*Supported by National Science Foundation Grant GP-24183.

Dar-Shong Hwang
320 words

ABSTRACT

Interval Estimation of Functions of Bernoulli Parameters With Reliability and Biomedical Applications

The reliability of series, parallel and other systems is expressible in terms of polynomials in Bernoulli parameters, p_i. For products of two or more parameters the Lehmann-Scheffe theory of exponential families applies when the components are sampled by the inverse binomial rule. For example if X, Y are the number of successes before r_1, r_2 failures on two populations of components, then the conditional distribution of X given $Y - X$ depends only on p_1p_2, the reliability of a series system. This conditional distribution is shown to be related to hypergeometric functions. In the most general case considered, a conditional distribution is given which depends only on the quotient $\theta = p_1p_2\cdots p_{k_1}/p_1'p_2'\cdots p_{k_2}'$ where the p's are parameters of $k_1 + k_2$ Bernoulli populations.

A second technique is "multiple stage compounding." It is known that if Y is Poisson, $EY = \lambda$, and if X given $Y = y$ is binomial (y, p), then X is Poisson, $EX = \lambda p$. This fact is exploited by taking for example a binomial observation X_1 from population 1, where the sample size is a Poisson variate having known mean λ, and then taking a second binomial observation X_2 from population 2, where the number of trials is the observed value of X_1. Then p_1p_2 can be estimated from the value of X_2, which is Poisson with mean λp_1p_2. Similar methods can be used to estimate any rational function of Bernoulli parameters.

Lieberman and Ross have described a technique for estimating reliability when time failure has an exponential distribution. We show that their technique applies also in our Bernoulli model. The product of Bernoulli parameters is the type of function that can be estimated by this technique, so that the reliability of series and parallel systems can be treated. The method depends on properties of independent geometric random variables X and Y; for example, the fact that $\min(X, Y)$ and $|X - Y|$ are independently distributed.

The results described in the first paragraph are also shown to apply to biomedical problems including estimation of the difference of two bacterial densities, comparison of two Yule's birth processes, and comparison of two effectiveness indices.

Approved
R. J. Buehler
April 23, 1971

Acknowledgements

My sincere gratitude goes to my advisor, Professor Robert J. Buehler, for his excellent guidance and many valuable suggestions. It has been both a privilege and a pleasure to work under his direction.

I am thankful to the National Science Foundation for its financial support of this research, under Grant No. GP-24183.

Finally my thanks are also due to Mr. Gerald DuChaine Jr. for his skillful typing of both the rough and final drafts.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION AND NOTATIONAL CONVENTIONS

1.1 Introduction and Summary.

If π_1, π_2 are two Bernoulli populations with parameters p_1, p_2 respectively, various functions of p_1 and p_2 like the difference $p_1 - p_2$, the product $p_1 p_2$, the ratio p_1/p_2 are relevant in various real-life problems, and it is of practical importance to be able to estimate these functions. In recent years, statisticians and reliability engineers have been interested in estimating the reliability of a complex system (a piece of equipment, an aircraft, a missile, etc.) based on tests of its components. If components operate independently and each component assumes only two states of performance, success or failure, then we have a so-called "attribute failure model" and the system reliability can be expressed as a function of Bernoulli parameters. Interval estimation of such functions is the main objective in our investigation.

Suppose π_i, $i = 1, 2, \ldots, k$, are k independent components of a system with component reliabilities p_i. There are two basic reliability functions, namely, the reliability of a series system R_1:

$$R_1 = p_1 p_2 \cdots p_k$$

and that of a parallel system R_2:

$$R_2 = 1 - q_1 q_2 \cdots q_k, \quad (q_i = 1 - p_i).$$

Many authors have written about constructing confidence limits for

R_1 and R_2 (or special cases) using various approximations. Among them Buehler (1957) and Harris (1968) use Poisson approximation to estimate R_2, Madansky (1965) uses a chi-square distribution approximation to estimate R_1 and R_2 while Rosenblatt (1963) employs normal approximation. All of these approaches assume ordinary binomial sampling.

In Chapter II, we assume inverse binomial sampling and find that if (using notation defined below in Section 1.2)

$$X_i \sim NB(r_i, 1-p_i),\ i = 1,2,\ldots, k_1$$

and

$$Y_j \sim NB(s_j, 1-p'_j),\ j = 1,2,\ldots, k_2$$

are the $k_1 + k_2$ resulting negative binomial r.v.'s, then the conditional distribution of X_1 given the observed differences $X_i - X_1 = w_i$, $i = 2,3,\ldots, k_1$, and the observed sums $Y_j + X_1 = v_j$, $j = 1,2,\ldots, k_2$, is a distribution depending on the parametric function $\theta \equiv p_1p_2\cdots p_{k_1}/p'_1p'_2\cdots p'_{k_2}$ only.

In Chapter III, we apply this result and treat the special case $k_2 = 0$ and $\theta = R_1$. The resulting distribution is found to be expressible in terms of generalized hypergeometric functions, and exact confidence limits for the product of any number of Bernoulli parameters are constructed. With simple relabeling, the method also gives confidence limits for R_2, the reliability of a parallel system. By randomization, our confidence limits can be made to be "uniformly most accurate unbiased" in Lehmann's sense. Under usual assumptions about approximating a negative binomial r.v. by

a Poisson r.v., our exact solution for a parallel system tends to Harris' (1968) Poisson approximation, as one would expect.

A general class of systems called "monotonic" or "coherent" structures which includes parallel and series systems as special cases have been discussed by various authors. The reliability of a system in this class is a polynomial in its component reliabilities. In Chapter IV we introduce a multiple stage compounding technique which considerably augments the class of functions one can estimate. The class includes any polynomial, and more generally, any rational function of Bernoulli parameters. The basic idea is that in drawing binomial samples from a Bernoulli population, the sample sizes should be determined by the outcomes of a known Poisson distribution. The technique provides interval estimates of every system reliability in the attribute failure model. By randomization, many of them can be made to be "uniformly most accurate" in a certain class. The technique is easily extended to cover the case of a multistate failure model. It also provides a non-parametric solution for estimating the system reliability in the variable failure model, irrespective of the underlying time to failure distribution used (Section 6.1).

In Chapter V we discuss a method of estimating the reliability of a series or a parallel system using "left-over information." The method is a counterpart of Lieberman and Ross' (1970) solution for the exponential time-to-failure model.

In Chapter VI we use the relationship between the exponential distribution and the Poisson process to estimate the reliability

of a series system when the individual component has an exponential time-to-failure model (Section 6.2). We also discuss applications of the results obtained in Chapter II to some biomedical problems including estimation of the difference of two bacterial densities, comparison of two Yule's birth processes and comparison of two effectiveness indices.

1.2 Notational Conventions.

Throughout the thesis we shall adopt the following notational conventions. Except when they are encountered for the first time in the sequel, we shall not define them again.

π_i, π_j' : π_i means the i^{th} Bernoulli population. When there are two sets of Bernoulli populations, π_j' will denote the j^{th} Bernoulli population in the second set.

p_i : the probability of success for the Bernoulli population π_i. $0 < p_i < 1$.

p_i' : the probability of success for the Bernoulli population π_j'. $0 < p_i' < 1$.

q_i : equal to $1 - p_i$; i.e., probability of failure for the population π_i.

q_j' : equal to $1 - p_j'$; i.e., probability of failure for the population π_j'.

r.v. : Random variable.

CDF : Cumulative distribution function.

PMF : Probability mass function.

PGF : Probability generating function.

IBS : Inverse binomial sampling.

RHS : Right hand side.

LHS : Left hand side.

MSCP : Multiple stage compound Poisson (see Section 4.2.1).

$X \sim NB(r, q)$: means X is a negative binomial random variable with parameters r and q; more precisely, it means

(1.1) $P\{X = x\} = \binom{r+x-1}{x} q^r p^x$, $x = 0,1,2,\ldots$ where $p = 1 - q$.

$X \sim GD(p)$: means X is a geometric random variable with parameter p.

(1.2) $P\{X = x\} = pq^x$, $x = 0,1,2,\ldots$ where $q = 1 - p$.

$X \sim B(n, p)$: means X is a binomial random variable with parameters n and p; and

(1.3) $P\{X = x\} = \binom{n}{x} p^x q^{n-x}$, $x = 0,1,2,\ldots, n$.

$X \sim P_0(\lambda)$: means X is a Poisson random variable with parameter λ.

(1.4) $P\{X = x\} = e^{-\lambda}\lambda^x/x!$, $x = 0,1,2,\ldots$.

$T \sim Exp(\lambda)$: means T is an exponential r.v. with parameter $\lambda > 0$.

(1.5) $f_T(t) = \lambda e^{-\lambda t}$, $t > 0$.

$X \sim Gamma(n,\lambda)$: means X is a gamma r.v. with parameters $n > 0$, $\lambda > 0$.

(1.6) $f_X(x) = \frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda x}$, $x > 0$.

$X_n \xrightarrow{d} X$: means the sequence of r.v.'s $X_1, X_2, \ldots, X_n, \ldots$ converges in distribution to the r.v. X.

$F(a, b, c; z)$: the hypergeometric function.

$$(1.7) \qquad F(a, b, c; z) = \{1 + \frac{a \cdot b}{1 \cdot c} z + \frac{a(a+1)b(b+1)}{1 \cdot 2 \cdot c(c+1)} z^2 + \frac{a(a+1)(a+2)b(b+1)(b+2)}{1 \cdot 2 \cdot 3 \cdot c(c+1)(c+2)} z^3 + \ldots \}$$

See Whittaker and Watson (1915).

CHAPTER II

INFERENCE ABOUT THE QUOTIENT OF PRODUCTS OF BERNOULLI PARAMETERS

Let $\pi_1, \pi_2, \ldots, \pi_{k_1}$ and $\pi'_1, \pi'_2, \ldots, \pi'_{k_2}$ be $k_1 + k_2$ independent Bernoulli populations with parameters $p_1, p_2, \ldots, p_{k_1}$ and $p'_1, p'_2, \ldots, p'_{k_2}$ respectively. In this chapter we describe how to apply inverse binomial sampling to make statistical inferences about the parametric function

$$\theta \equiv \frac{p_1 p_2 \cdots p_{k_1}}{p'_1 p'_2 \cdots p'_{k_2}} . \tag{2.1}$$

The product $p_1 p_2$ and the quotient p_1/p_2 of two Bernoulli parameters are special cases of practical interest. Our aim is to get rid of nuisance parameters by devising suitable sampling rules so that it is possible to make exact probability statements about θ and obtain statistical tests and confidence limits for θ having certain optimum properties.

2.1 Design of Experiments.

The sampling method we use is inverse binomial sampling in which the number of successes (or failures) of the experiment on each population is fixed, while the sample size is random.

This type of sampling plan apparently was first used by J.B.S. Haldane (1945) in estimating abnormal blood corpuscle proportions. Later it was used for example in bacteria counting (Sandelius, 1950), animal population estimation (Chapman, 1952), and reliability testing for exponential populations (Nadler, 1960).

The IBS plan as applied to our present problem will be conducted in the following way:

For each of the Bernoulli populations π_i, $i = 1,2,\ldots, k_1$, we continue to perform independent, identical trials sequentially until r_i^{th} failure ($i = 1,2,\ldots, k_1$) is encountered where each r_i is any positive integer fixed in advance of the experiment. Then the sampling is terminated. Let X_i denote the number of successes encountered prior to the occurrence of the r_i^{th} failure. Then it is well-known that $X_i \sim NB(r_i, q_i)$, i.e., X_i has a negative binomial distribution with parameters r_i and q_i. The probability law is given by (1.1).

Similarly, for each of the Bernoulli populations $\pi_1', \pi_2', \ldots, \pi_{k_2}'$, we do inverse binomial sampling until s_j^{th} failure ($j = 1,2,\ldots, k_2$) is encountered, where each s_j is any positive integer fixed in advance of the experiment. Let Y_j denote the number of successes encountered prior to the occurrence of the s_j^{th} failure. Then $Y_j \sim NB(s_j, q_j')$. Thus we have $k_1 + k_2$ independent negative binomial random variables $X_1, X_2,\ldots, X_{k_1}; Y_1, Y_2,\ldots, Y_{k_2}$.

2.2 Derivation of a Probability Distribution Involving Only θ.

Let $X = (X_1,\ldots, X_{k_1})$, $Y = (Y_1,\ldots, Y_{k_2})$, $x = (x_1,\ldots, x_{k_1})$, $y = (y_1,\ldots, y_{k_2})$. Then

$$(2.2) \qquad P\{X = x, Y = y\} = \prod_{i=1}^{k_1} \prod_{j=1}^{k_2} \binom{r_i+x_i-1}{x_i} \binom{s_j+y_j-1}{y_j} p_i^{x_i} q_i^{r_i} p_j'^{y_j} q_j'^{s_j}$$

where $x_i = 0,1,2,\ldots$, $y_j = 0,1,2,\ldots$. We assume without loss of generality $k_1 > 0$. Let random variables W_i and V_j be defined by

(2.3) $W_1 = X_1$

$W_i = X_i - X_1$, for $i = 2,3,\ldots, k_1$

$V_j = Y_j + X_1$, for $j = 1,2,3,\ldots, k_2$.

Let $U = (W_2, W_3,\ldots, W_{k_1}, V_1, V_2,\ldots, V_{k_2})$ and $u = (w_2, w_3,\ldots, w_{k_1}, v_1, v_2,\ldots, v_{k_2})$. Then

(2.4) $P(w_1 = w_1, U = u) = C(w_1, u)\theta^{w_1}$

where θ is defined by (2.1) and

(2.5) $$c(w_1, u) = \binom{r_1+w_1-1}{w_1} q_1^{r_1} p_1^{w_1} \prod_{i=2}^{k_1} \prod_{j=1}^{k_2} \binom{r_i+w_i+w_1-1}{w_i+w_1} \cdot \binom{s_j+v_j-w_1-1}{v_j-w_1} q_i^{r_i} p_i^{w_i} q_j'^{w_j} p_j'^{v_j}$$

and the range of the variables is

$w_1 = 0,1,2,\ldots$

$w_i = -w_1, -w_1+1, -w_1+2,\ldots$

$v_j = w_1, w_1+1, w_1+2,\ldots$.

The new parameter θ is the desired parametric function we are interested in, other parameters in (2.4) are nuisance parameters of which we wish to eliminate. To get rid of nuisance parameters, we find the conditional distribution of w_1 given $U = u$:

(2.6) $P\{W_1 = w_1 | U = u\} = b(w_1, u)\theta^{w_1} / \sum_t b(t, u)\theta^t$

where

$$(2.7)\qquad b(w_1, u) = \binom{r_1+w_1-1}{w_1} \prod_{i=2}^{k_1} \prod_{j=1}^{k_2} \binom{r_i+w_i+w_1-1}{w_i+w_1} \binom{s_j+v_j-w_1-1}{v_j-w_1}$$

and where w_1 is any non-negative integer satisfying $\max\{0, \max_{2\le i\le k_1}(-w_i)\} \le w_1 \le \min_{1\le j\le k_2} v_j$. The values of t in the summation in the denominator are the same as the values of w_1.

The reasons for the restriction of the values of the conditioned random variable W_1 to the double inequality can be seen briefly as follows: In (2.6), once the first conditioning random variable W_2 is given to be a definite value (an integer), w_2, say, then noting the definitions of W_2 and W_1, we know the conditioned random variable W_1 can take only those non-negative integral values greater or equal to $\max(0, -w_2)$. Likewise, once $W_i = w_i$ for some $i = 3,4,\ldots, k_1$ is given, W_1 can only assume the integral values $\ge \max(0, -w_i)$. But $W_1 = w_1$ is conditioned by the conditions $W_2 = w_2$, $W_3 = W_3,\ldots, W_{k_1} = w_{k_1}$ <u>simultaneously</u>; therefore the possible values of $W_1 = w_1$ must satisfy

$$\begin{aligned} &\max(0, -w_2) \le w_1,\\ &\max(0, -w_3) \le w_1,\\ &\qquad\vdots\\ &\max(0, -w_{k_1}) \le w_1 \end{aligned}$$

simultaneously. This implies w_1 must satisfy $\max_{2\le i\le k_1}\{\max(0, -w_i)\} \le w_1$, which can be reduced to $\max\{0, \max_{2\le i\le k_1}(-w_i)\} \le w_1$.

Similar observations lead to the other inequality

$w_1 \le \min_{1\le j\le k_2} v_j$.

We note the conditional distribution (2.6) is easily verified to belong to the one parameter exponential family. It depends only on the parameter θ, not on the individual (p_i)'s and (p'_j)'s. Thus statistical inferences about θ can be made from this conditional distribution. The disappearance of all nuisance parameters is also assured by Lehmann and Scheffe's remarkable theorem about exponential families (see, for example, Lehmann (1959), Chapter 2, Lemma 8). We note the joint distribution of the $k_1 + k_2$ negative binomial random variables in (2.2) can be rewritten as

$$(2.8) \qquad P\{X = x, Y = y\} = A(p, p')h(x, y) \cdot \exp\{x_1 \log \theta + \sum_{i=2}^{k_1} (x_i - x_i)\log p_i + \sum_{j=1}^{k_2} (y_i + x_1)\log p'_j\},$$

where

$$A(p, p') = \prod_{i=1}^{k_1} q_i^{r_i} \prod_{j=1}^{k_2} q'^{s_j}_j,$$

and

$$h(x, y) = \prod_{i=1}^{k_1} \binom{r_i + x_i - 1}{x_i} \prod_{j=1}^{k_2} \binom{s_j + y_j - 1}{y_j},$$

which is seen to belong to the (multi-parameter) exponential family by straightforward identification. By Lemma 8, p. 52 of Lehmann (1959), the conditional distribution of X_1 given $X_i - X_1 = w_i$, $i = 2,3,\ldots, k_1$ and $Y_j + X_1 = v_j$, $j = 1,2,\ldots, k_2$ again belongs to the (one parameter) exponential family and depends on the parameter $\log p_1 p_2 \ldots p_{k_1} / p'_1 p'_2 \ldots p'_{k_2}$ only, not on the nuisance parameters $\log p_i$, $i = 2,3,\ldots, k_1$ and $\log p'_j$, $j = 1,2,\ldots, k_2$. Making the transformation (2.3), and noting the logarithm function is one to

one, it is equivalent to say the conditional distribution of W_1 given $U = u$ again belongs to the (one parameter) exponential family and depends on the desired new parameter θ only.

Before going to the discussion of statistical inferences about θ, it should be pointed out that the above conditional distribution technique in getting rid of nuisance parameters can be generalized to include more general parametric functions: θ' = quotient or ratio of two product terms; each product term is a mixture of (p_i)'s, (q_i)'s, (p_j')'s and (q_i')'s.

2.3 Statistical Inferences About θ.

As shown in the previous section, the probability distribution (2.6) depends on the single parameter θ only. Therefore, as usual, one-side and two-side conditional tests about the hypothetical values of θ can be obtained by observing W_1 given the observed values of $W_2,\ldots, W_{k_1}$ and $V_1,\ldots, V_{k_2}$. Since the joint distribution of the original random variables is seen in (2.8) to belong to the exponential family, we can apply Lehmann and Scheffe's theorem about uniformly most powerful unbiased tests for the multiparameter exponential family (see Theorem 3, p. 134 of Lehmann (1959)) and write down the tests of hypothesis which might be of interest about θ(θ_0, θ_1, θ_2 below being specified values):

Null Hypothesis	Alternative Hypothesis
H_1: $\theta \leq \theta_0$	k_1: $\theta > \theta_0$
H_1': $\theta = \theta_0$	k_1': $\theta > \theta_0$
H_2: $\theta < \theta_1$ or $\theta \geq \theta_2$	k_2: $\theta_1 < \theta < \theta_2$
H_3: $\theta_1 \leq \theta \leq \theta_2$	k_3: $\theta < \theta_1$ or $\theta > \theta_2$
H_4: $\theta = \theta_0$	k_4: $\theta \neq \theta_0$.

For example, let α be the size of the test, $0 < \alpha < 1$, and suppose the observed value of W_1 is k and that of U is u in (2.6). Then to test H_1: $\theta \leq \theta_0$ against the alternatives k_1: $\theta > \theta_0$, we reject H_1 if $k \geq c(u)$, where the critical value $c(u) \geq 0$ is, under $\theta = \theta_0$, the smallest number satisfying

$$(2.9) \qquad \sum_{w_1 \geq c(u)} P_{\theta_0}(W_1 = w_1 | U = u) \leq \alpha \ .$$

Since we are dealing with discrete distributions, for most values of θ_0 under null hypothesis, the size of our test is actually less than α. To obtain tests with size exactly equal to α, randomization at the critical value $c(u)$ can be employed. In that case, the test would have the property of being uniformly most powerful unbiased.

One and two-sided confidence intervals can be constructed as follows:

Let $0 < \alpha < 1$. Upon observing $W_1 = k$, the upper confidence limit $\theta_2(k, u)$ (having confidence level $1 - \alpha$) conditioned on the observed $U = u$ is

$$(2.10) \qquad \theta_2(k, u) = \sup\{\theta: \sum_{w_1 \leq k} P_\theta\{W_1 = w_1 | U = u\} \geq \alpha\}$$

and thus

$$P\{\theta < \theta_2(k, u) | U = u\} \geq 1 - \alpha.$$

The corresponding $1 - \alpha$ lower confidence limit for θ is

$$(2.11) \qquad \theta_1(k, u) = \inf\{\theta: \sum_{w_1 \geq k} P_\theta\{W_1 = w_1 | U = u\} \geq \alpha\}$$

and thus

$$P\{\theta \geq \theta_1(k, u) | U = u\} \geq 1 - \alpha.$$

The problem created by the discreteness of the distribution mentioned in hypothesis testing has similar bearing in interval estimation here. We note by proper randomization, the confidence limits $\theta_1(k, u)$ and $\theta_2(k, u)$ can be made to have the property of being uniformly most accurate unbiased (see Lehmann (1959), Chapter 4).

CHAPTER III

APPLICATION TO RELIABILITY PROBLEMS

We come now to the problems which originally motivated the writing of this thesis. One area in the statistical reliability theory which has received much attention in recent years is the following: A system (for example, electronic equipment, aircraft, missle, etc.) is built up from several components. Some systems will function successfully if and only if every component of it operates successfully. This type is called a "series system." If all the components serve the same purpose and the successful operation of anyone of them guarantees the proper functioning of the system itself, then we have a "parallel system." For example, a manual brake and an automatic brake in a braking system form a parallel system. More complex systems may have both series-connected and parallel-connected components. In any case, one is interested in assessing the reliability, defined as the probability of successful operation of the entire system. It may turn out that testing the system as a whole is quite expensive and sometimes not even feasible, while testing the individual component is cheaper, feasible and more informative. Clearly it would be desirable to make inferences about system reliability based on testing data of individual components. This method of estimating system reliability is also useful for exploratory system design and will be our primary concern in this chapter.

It is assumed that by analysis of the system structure, the system reliability R can be expressed as a function of the individual

reliability of the components. Data obtained from testing individual components may be available in one of two forms: "attribute data" or "variable data."

3.1 Attribute Failure Model.

Component data in the form of "success" or "failure" are called attribute data and are analysed in terms of Bernoulli populations. When a component from population π_i, $i = 1,2,\ldots, k$ is tested, it either succeeds or fails, with probabilities p_i and q_i, where $q_i + p_i = 1$, $0 < p_i < 1$, and the component outcomes are mutually independent. This is often called the "attribute failure model." Assuming there are k components in a system, it is obvious that in this model, the system failure probability is $q_1q_2\ldots q_k$ for the parallel system and $1 - p_1p_2\ldots p_k$ for the series system, while the system reliability is $\theta \equiv 1 - q_1q_2\ldots q_k$ and $\theta \equiv p_1p_2\ldots p_k$ respectively. In any case, the problem reduces to constructing confidence limits for the product of k Bernoulli parameters. For a system which is comprised of components connected in both parallel and series, a more complicated expression for the system reliability is needed.

Starting with Buehler (1957), much has been written about the problem by various writers from academic, industry and military establishments. As far as I am aware, every writer treats the problem from the viewpoint of fixed sample size approach, sometimes called positive binomial sampling, which leads to binomial distributions, and great difficulties--both theoretical and practical--are encountered, so that no one seems to have made exact confidence limit statements

about the θ's. Instead various large sample approximations were introduced, notably the Poisson approximations of Buehler (1957), Harris (1968) and others, the chi-square distribution approximations of Madansky (1965), Myhre and Saunders (1965), the normal approximations through maximum likelihood estimators of Rosenblatt (1963), DeCicco (1960), and Thomas (1960). For rather complete literature surveys in this area, see Rosenblatt (1963) and Mann (1970).

In view of the difficulties encountered in the fixed sample size approach, it is interesting to see what one can accomplish by adopting random sample size approach. It turns out that if we adopt the inverse sampling plan as described in Chapter II, we find the product function $\theta = p_1 p_2 \cdots p_k$ is just a special case of the more general parametric function

$$\theta = \frac{p_1 p_2 \cdots p_{k_1}}{p_1' p_2' \cdots p_{k_2}'} \tag{3.1}$$

treated in Chapter II.

3.2 Statistical Inferences About Reliability of Series and Parallel Systems.

If we set $k_1 = k$ and $k_2 = 0$ in (3.1), then the parameter θ becomes $\theta^* = p_1 p_2 \cdots p_k$ which is the reliability expression for the series system of k components in our present context. Thus results obtained in Chapter II provide tests and confidence limits for θ^*. As before, the observed random variable, X_i, $i = 1,2,\ldots, k$ would be the number of successes occurring prior to the r_i^{th} failure. If we are dealing with a parallel system, we simply relabeled the

random variables X_i, $i = 1,2,\ldots, k$ to be the number of failures (instead of successes) encountered prior to the occurrence of the r_i^{th} success in each of k Bernoulli populations, the resulting conditional distribution in Equation (2.6) will depend on the parameter $\theta = q_1 q_2 \ldots q_k$, and this can be used to construct confidence limits for the parametric function $1 - q_1 q_2 \ldots q_k$, the reliability of a parallel system.

The simple way by which we can handle both the series system and the parallel system is not always shared by procedures used by other writers. For example, the Poisson approximation approach can handle parallel system but not series system. In reliability problems, failure probabilities q_i, $i = 1,2,\ldots, k$ are small. Therefore, by taking sample size large enough, one can approximate an original binomial random variable by a Poisson random variable, thus making inference about $1 - q_1 q_2 \ldots q_k$, the reliability of a parallel system. But in a series system, the reliability is $p_1 p_2 \ldots p_k$, where $p_i = 1 - q_i$ is close to 1. Therefore, Poisson approximation does not provide a satisfactory approximation.

We have seen that mathematically speaking, the difference between analysis for a parallel system and a series system under our IBS plan amounts only to a simple relabeling of the random variables involved. Therefore, we will briefly summarize the results for the series system of k (≥ 2) components only:

As in Chapter II, let $W_1 = X_1$, $W_i = X_i - X_1$, $i = 2,3,\ldots, k$. If we define

$$U = (W_2, W_3, \ldots, W_k),$$

and

$$u = (w_2, w_3, \ldots, w_k),$$

then the conditional distribution of W_1 given $U = u$ (from (2.6)) is

$$(3.2) \qquad P_{\theta^*}\{W_1 = w_1 | U = u\} = b'(w_1, u)(\theta^*)^{w_1} / \sum_t b'(t, u)(\theta^*)^t,$$

where

$$(3.3) \qquad b'(w_1, u) = \binom{r_1+w_1-1}{w_1} \prod_{i=2}^{k} \binom{r_i+w_i+w_1-1}{w_i+w_1},$$

$\theta^* = P_1 P_2 \ldots P_k$ and w_1 is any non-negative integer satisfying

$$(3.4) \qquad \max\{0, \max_{2 \le i \le k} (-w_i)\} \le w_1 < \infty.$$

The values of t in the summation in the denominator run through all possible values of w_1 in (3.4).

The distribution (3.2) can incidentally be expressed in terms of generalized hypogeometric functions. We do this for $k = 2$ in Section 3.3.

Having found the distribution (3.2) depending on the desired parameter θ^* only, it is a routine job to conduct tests of hypothesis and constructions of confidence limits for θ^*. For example, let $0 < \alpha < 1$. Upon observing $W_1 = k$, the $1 - \alpha$ upper confidence limit $\theta_2^*(k, u)$ conditioned on the observed $U = u$ is:

$$(3.5) \qquad \theta_2^*(k, u) = \sup\{\theta^*: \sum_{w_1 \le k} P_{\theta^*}\{W_1 = w_1 | U = u\} \ge \alpha\}$$

and thus

$$P\{\theta^* < \theta_2^*(k, u) | U = u\} \geq 1 - \alpha.$$

The corresponding $1 - \alpha$ lower confidence limit for θ^* is

$$(3.6) \qquad \theta_1^*(k, u) = \inf\{\theta^*: \sum_{w_1 \geq k} P_{\theta^*}\{W_1 = w_1 | U = u\} \geq \alpha\}$$

and thus

$$P\{\theta^* \geq \theta_1^*(k, u) | U = u\} \geq 1 - \alpha .$$

Remarks about using randomization to obtain uniformly most accurate unbiased confidence limits mentioned at the end of Chapter II are also applicable here.

3.3 Estimating the Product of Two Bernoulli Parameters.

In this section we discuss some specific properties of the particular case of θ^* by taking $k = 2$; that is, θ^* becomes $\lambda \equiv p_1 p_2$, the product of two Bernoulli parameters. Physically this represents the reliability of a series system of two components π_1, π_2 (the method to be developed is easily adapted to treat the reliability of a parallel system, as commented upon before). It is highly desirable to find some convenient, known function which can be used to express the conditional distributions and facilitate computation. It turns out that the hypergeometric function is involved. Some interesting special cases will be mentioned and finally approximations appropriate for parallel systems will be given.

3.3.1 Probability distributions depending on $\lambda \equiv p_1 p_2$.

The definition of the hypergeometric function $F(a, b, c; z)$ given in (1.7) involves three parameters a, b and $c > 0$. The series is absolutely convergent when $|z| < 1$. For our purposes,

a, b, c will be positive integers; z will be a real number with $|z| < 1$. By straightforward rearrangement in each term of $F(a, b, c; z)$, it is seen that the function can be expressed as

$$(3.7) \qquad F(a, b, c; z) = \sum_{n=0}^{\infty} f(a, b, c; z, n),$$

where

$$(3.8) \qquad f(a, b, c; z, n) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+n)\Gamma(b+n)}{\Gamma(c+n)} \frac{z^n}{n!}.$$

Here $\Gamma(x)$, $x > 0$, is the familiar gamma function. Define $F_x(a, b, c; z)$ as the partial sum:

$$(3.9) \qquad F_x(a, b, c; z) = \sum_{n=0}^{x} f(a, b, c; z, n).$$

Then $F_\infty(a, b, c; z) = F(a, b, c; z)$.

We shall see presently that the conditional distribution (3.2) can be expressed in terms of $F_x(a, b, c; z)$ and $F(a, b, c; z)$. When $k = 2$, (3.2) becomes (note w_2 is replaced by w):

$$(3.10) \qquad P_\lambda\{X_1 = x_1 | X_2 - X_1 = w\} = b'(x_1, w)\lambda^{x_1} / \sum_{t=\max(0,-w)}^{\infty} b'(t, w)\lambda^t,$$

where

$$b'(x_1, w) = \binom{r_1+x_1-1}{x_1}\binom{r_2+w+x_1-1}{w+x_1},$$

and $\text{Max}(0, -w) \leq x_1 < \infty$.

Case 1: If $w \geq 0$, then (3.10) becomes

$$(3.11) \qquad P_\lambda\{X_1 = x_1 | X_2 - X_1 = w\} = b'(x_1, w)\lambda^{x_1} / \sum_{t=0}^{\infty} b'(t, w)\lambda^t, \quad x_1 = 0,1,2,\ldots.$$

Each term of the sum in the denominator can be written as

$$(3.12) \qquad b'(t, w)\lambda^t = \frac{\Gamma(r_2+w)}{\Gamma(r_2)\Gamma(w+1)} f(r_1, r_2+w, w+1; \lambda, t)$$

where $f(a, b, c; z, n)$ is defined in (3.8). The numerator of (3.11) is just one of the individual terms of the sum in the denominator and therefore it can be expressed as the RHS of (3.12) with t replaced by x_1. Thus using (3.7), (3.11) becomes

$$(3.13) \qquad P_\lambda\{X_1 = x_1 | X_2 - X_1 = w\} = \frac{f(r_1, r_2+w, w+1; \lambda, x_1)}{F(r_1, r_2+w, w+1; \lambda)} , \quad x_1 = 0,1,2,....$$

Introducing (3.9), the CDF of the conditional distribution (3.13) is given at non-negative integer values of x by

$$(3.14) \qquad P_\lambda\{X_1 \leq x | X_2 - X_1 = w\} = \frac{F_x(r_1, r_2+w, w+1; \lambda)}{F(r_1, r_2+w, w+1; \lambda)} ,$$

as we promised to show.

Case 2: If $w < 0$, then by defining $Y = X_1 + w$, the distribution of the translated r.v. Y can be found in a way similar to Case 1, and by translating back to the original r.v. X_1, we have

$$(3.15) \qquad P_\lambda\{X_1 = x_1 | X_2 - X_1 = w\} = \frac{f(r_2, r_1-w, 1-w; \lambda, x_1+w)}{F(r_2, r_1-w, 1-w; \lambda)} ,$$

$$x_1 = -w, -w+1, -w+2,....$$

The CDF form of (3.15) is:

$$(3.16) \qquad P_\lambda\{X_1 \leq x | X_2 - X_1 = w\} = \frac{F_{x+w}(r_2, r_1-w, 1-w; \lambda)}{F(r_2, r_1-w, 1-w; \lambda)} ,$$

$$x = -w, -w+1, -w+2,....$$

Combining (3.14) and (3.16), the cumulative distribution function of X_1 given $X_2 - X_1 = w$, $w = 0, \pm 1, \pm 2, \ldots$, is:

$$(3.17)\quad P\{X_1 \leq x \mid X_2 - X_1 = w\} = \begin{cases} F_x(r_1, r_2+w, 1+w; \lambda)/F(r_1, r_2+w, 1+w; \lambda), & \text{if } w \geq 0 \\ F_{x+w}(r_2, r_1-w, 1-w; \lambda)/F(r_2, r_1-w, 1-w; \lambda), & \text{if } w < 0 \end{cases}$$

where x is an integer ≥ 0 if $w \geq 0$; and an integer $\geq -w$ if $w < 0$.

By analogy with the incomplete gamma function and incomplete beta function, it is appropriate to call the distribution (3.17) the "incomplete hypergeometric function."

As a corollary to the conditional distributions (3.13), and (3.15) we can incidentally find the probability distribution of the difference of two negative binomial random variables.

If $x_i \sim NB(r_i, q_i)$, $p_i = 1 - q_i$, $i = 1,2$, and if x_1 and x_2 are independent, then

$$(3.18)\quad P\{X_2 - X_1 = w\} = c(r_1, r_2) p_1^{\alpha} p_2^{\beta} F(r_1+\alpha, r_2+\beta, 1 + |w|; p_1 p_2)$$

where

$$\alpha = \max(0, -w), \quad \beta = \max(0, w),$$

$$c(r_1, r_2) = \binom{r_1+\alpha-1}{\alpha} \binom{r_2+\beta-1}{\beta} q_1^{r_1} q_2^{r_2},$$

$$w = 0, \pm 1, \pm 2, \pm 3, \ldots.$$

In particular,

$$P\{X_2 - X_1 = w\} = \begin{cases} \binom{r_2+w-1}{w} q_1^{r_1} q_2^{r_2} p_2^{w} F(r_1, r_2+w, 1+w; p_1 p_2), \\ \qquad \text{for } w = 0,1,2,3,\ldots \\ \binom{r_1-w-1}{-w} q_1^{r_1} q_2^{r_2} p_1^{-w} F(r_2, r_1-w, 1-w; p_1 p_2), \\ \qquad \text{for } -w = 0,1,2,3,\ldots. \end{cases}$$

Using this result, the incomplete hypergeometric function (3.17) can be written in a single expression:

$$(3.19)\qquad P\{X_1 \le x | X_2 - X_1 = w\} = \frac{F_{x+\alpha}(r_1+\alpha,\ r_2+\beta,\ 1 + |w|;\ \lambda)}{F(r_1+\alpha,\ r_2+\beta,\ 1 + |w|;\ \lambda)}$$

$$x = \alpha,\ \alpha+1,\ \alpha+2,\ldots$$

$$w = 0,\ \pm 1,\ \pm 2,\ldots$$

and the corresponding PMF (3.13) and (3.15) can be put in a combined form:

$$(3.20)\qquad P\{X_1 = x_1 | X_2 - X_1 = w\} = \frac{f(r_1+\alpha,\ r_2+\beta,\ 1 + |w|;\ \lambda,\ x_1-\alpha)}{F(r_1+\alpha,\ r_2+\beta,\ 1 + |w|;\ \lambda)},$$

$$x_1 = \alpha,\ \alpha+1,\ \alpha+2,\ldots$$

$$w = 0,\ \pm 1,\ \pm 2,\ldots.$$

The probability generating function for (3.19) is immediately found to be the ratio of two hypergeometric functions:

$$(3.21)\qquad \eta(t) = F(r_1+\alpha,\ r_2+\beta,\ 1 + |w|;\ \lambda t)/F(r_1+\alpha,\ r_2+\beta,\ 1 + |w|;\ \lambda).$$

We see $\eta(t)$ is convergent for $|\lambda t| <$ | or $|t| < 1/p_1p_2$. Noting the simple formula $\frac{d}{dz} F(a, b, c; z) = \frac{ab}{c} F(a+1, b+1, c+1; z)$, the mean of this distribution is found by straightforward computation:

$$(3.22)\qquad E(X_1 | X_2 - X_1 = w) = \frac{\lambda(r_1+\alpha)(r_2+\beta)}{1 + |w|} \frac{F(r_1+\alpha+1,\ r_2+\beta+1,\ 2 + |w|;\ \lambda)}{F(r_1+\alpha,\ r_2+\beta,\ 1 + |w|;\ \lambda)}.$$

Similarly, the variance can also be obtained easily, but the algebraic expression involved is somewhat complicated and will not be given here.

3.3.2 Some special cases.

Some special cases of the CDF (3.19) or its corresponding probability mass function (3.20) are interesting. The case with either $r_1 = 1$ or $r_2 = 1$ reduces to the negative binomial distribution or "truncated negative binomial distribution," depending on the sign of the value of the conditioning r.v. Let us first consider

Case A: $r_1 = r_2 = 1$.

In this case we test both components π_1, π_2 until each of them encounters first failure. The PMF (3.20) becomes:

$$(3.23) \qquad P\{X_1 = x_1 | X_2 - X_1 = w\} = (1-p_1p_2)(p_1p_2)^{x_1}, \; x_1 = 0,1,2,\ldots,$$

for $w \geq 0$. This is a geometric distribution with parameter $1-p_1p_2$. Note the interesting fact that the conditioning value w disappears--the distribution is the same for all values of $w = 0,1,2,\ldots$. The corresponding CDF is

$$P\{X_1 \leq x | X_2 - X_1 = w\} = 1 - (p_1p_2)^{x+1}, \; x = 0,1,2,\ldots.$$

For later use, we note further that since $w \geq 0$, therefore $\min(X_1, X_2) = X_1$, and from (3.23), we have

$$(3.24) \qquad P\{\min(X_1, X_2) = x | X_2 - X_1 = w\} = (1-p_1p_2)(p_1p_2)^{x}, \; x = 0,1,2,\ldots,$$

a formula which also holds for $w < 0$. Since this does not depend on w, the unconditional distribution is the same:

$$\min(X_1, X_2) \sim GD(1-p_1p_2).$$

This result has been noted previously by Ferguson (1965). The signifi-

cance of the r.v. $\min(X_1, X_2)$ is that in our series system of two components, if we use IBS to test the "system" as a whole until first system failure, the r.v. X describing number of system successes prior to first system failure is exactly the r.v. $\text{Min}(X_1, X_2)$. For

$$\begin{aligned}\{X = x\} &= \{\text{The series system first fails at trial number } x + 1\} \\ &= \{\text{Either component } \pi_1 \text{ first fails at trial number } x + 1 \text{ or component } \pi_2 \text{ fails at trial number } x + 1\} \\ &= \{\text{Either } X_1 = x \text{ or } X_2 = x\} \\ &= \{\text{Min}(X_1, X_2) = x\}.\end{aligned}$$

Another way to look at the same matter is to find the distribution of X directly. In a series system of two independent components, the probability of (system) success in a trial is p_1p_2, while probability of failure is $1\text{-}p_1p_2$, it follows $X \sim GD(1\text{-}p_1p_2)$, the same as that of $\min(X_1, X_2)$, as we claimed.

The above remark about $\min(X_1, X_2)$ is closely related to another point of practical interest: In using the conditional distribution $P\{X_1 = x_1 | X_2 = x_1 = w\}$ to analyze experiment outcome, we have assigned the r.v. X_1 to component π_1 and X_2 to component π_2. The assignment is of course arbitrary. But it is important to keep the <u>same</u> assignment of r.v.'s to components in the subsequent experiments (if there are any), otherwise the analysis may lead to "testing the system," instead of "testing the components," and possibly lose information also. To see this, suppose when $X_1 \leq X_2$

is observed, we associate X_1 to component π_1, and X_2 to π_2; but if $X_1 > X_2$, we associate X_1 to π_2 and X_2 to π_1. A little reflexion shows that the conditioned r.v. X_1 in $P\{X_1 = x_1 | X_2 - X_1 = w\}$ actually becomes $\min(X_1, X_2)$ in this practice. But $\min(X_1, X_2)$ has already been shown to be equivalent to the r.v. X in "testing the system." The possible loss of information is now seen from the fact that the r.v. $\min(X_1, X_2)$ utilizes only the smaller of the two observations X_1, X_2, not both of them.

Next let us consider

Case B: $r_2 = 1$, r_1 = an arbitrary positive integer.

In this case we test component π_1 until r_1 failures, and test component π_2 until first failure. The PMF (3.20) becomes:

Case B1: If $w \geq 0$,

$$(3.25) \qquad P\{X_1 = x_1 | X_2 - X_1 = w\} = \binom{r_1+x_1-1}{x_1}(1-\lambda)^{r_1}\lambda^{x_1}, \quad x_1 = 0,1,2,\ldots.$$

This is a $NB(r_1, 1-\lambda)$ distribution and is free of the observed value w of the conditioning r.v. W. Note this is exactly the distribution one would obtain if he "tests the system" until r_1 (systems) failures. The CDF is obtained either from (3.19) or directly from (3.25):

$$(3.26) \qquad P\{X_1 \leq x | X_2 - X_1 = w\} = \sum_{x_1=0}^{x} \binom{r_1+x_1-1}{x_1}(1-\lambda)^{r_1}\lambda^{x_1}$$

$$= I_{1-\lambda}(r_1, x+1), \quad x = 0,1,2,\ldots,$$

where

$$I_x(m, n) = \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} \int_0^x t^{m-1}(1-t)^{n-1}dt$$

is the incomplete beta function which has been extensively tabulated by K. Pearson (1932). Note that (3.26) is also free of w.

Case B2: If $w < 0$, then (3.20) becomes

$$P\{X_1 = x_1 | X_2 - X_1 = w\} \tag{3.27}$$

$$= \binom{r_1+x_1-1}{x_1}(1-\lambda)^{r_1}\lambda^{x_1} \Big/ \sum_{t=-w}^{\infty}\binom{r_1+t-1}{t}(1-\lambda)^{r_1}\lambda^{t} \tag{3.28}$$

$$= \frac{\binom{r_1+x_1-1}{x_1}(1-\lambda)^{r_1}\lambda^{x_1}}{1 - I_{1-\lambda}(r_1,-w)}, \quad x_1 = -w, -w+1, -w+2,\ldots, \tag{3.29}$$

which is not free of w. Note in (3.28) the numerator is an individual term of $NB(r_1, 1-\lambda)$ with the first $(-w-1)$ terms missing, while the denominator is an adjusting factor to sum the remaining probabilities to unity. For this reason, (3.29) can be called a "truncated negative binomial distribution" truncated at $-w$. Its CDF is obtained from (3.19) or directly from (3.29)

$$P\{X_1 \le x | X_2 - X_1 = w\} = \frac{\sum_{x_1=-x}^{x}\binom{r_1+x_1-1}{x_1}(1-\lambda)^{r_1}\lambda^{x_1}}{1 - I_{1-\lambda}(r_1, -w)} \tag{3.30}$$

$$= \frac{I_{1-\lambda}(r_1, x+1) - I_{1-\lambda}(r_1,-w)}{1 - I_{1-\lambda}(r_1, -w)},$$

$$x = -w, -w+1, -w+2,\ldots.$$

Again Pearson's tables can be used.

Finally consider

Case C: $r_1 = 1$, r_2 = an arbitrary positive integer.

This case is similar to Case B just discussed. The cases $w \ge 0$ and $w < 0$ correspond to the cases $w < 0$ and $w \ge 0$ in Case B respectively.

3.4 Approximation for Parallel Systems.

Recall that the conditional distribution given by (3.20) is appropriate for a series system of two components. In this section we obtain a corresponding conditional distribution for a parallel system. Then we develop a Bessel function approximation for this distribution. Generalization to any number of components follows. Finally, a different type of approximation involving the known distribution of the product of two gamma r.v.'s is indicated.

If we relabel the r.v.'s X_1 and X_2 in (3.20) to represent numbers of failures prior to the r_1^{th} and r_2^{th} success respectively, then we obtain a conditional distribution appropriate for a parallel system of two components:

$$(3.31) \qquad P\{X_1 = x_1 | X_2 - X_1 = w\} = \frac{f(r_1+\alpha,\ r_2+\beta,\ 1 + |w|;\ \mu,\ x_1-\alpha)}{F(r_1+\alpha,\ r_2+\beta,\ 1 + |w|;\ \mu)} ,$$

where $\mu = q_1 q_2$,

α, β are defined immediately after Equation (3.18),

$x_1 = \alpha, \alpha+1, \alpha+2, \ldots$

$w = 0, \pm 1, \pm 2, \ldots$.

Note that $1 - \mu$ is the system reliability for this parallel system case, and (3.31) can be used to construct confidence limits for μ, and hence for $1 - \mu$. Now to derive an approximating distribution, we note the failure probabilities q_1 and q_2 are usually very small, and therefore r_1, r_2 must be comparatively very large in order to observe a few failures. With this in mind, the assumptions in Theorem 3.1 below are probably appropriate.

Theorem 3.1 (Bessel function approximation).

Let $X_i \sim NB(r_i, p_i)$, $i = 1,2$, be two independent negative binomial r.v.'s. If $r_i \to \infty$, $q_i = 1 - p_i \to 0$ $i = 1,2$, in such a way that $(r_1+\alpha)(r_2+\beta)q_1q_2 = \lambda$ remains fixed, where λ is a positive constant, then for each x_1, w, (3.31) has the limiting distribution:

$$(3.32) \qquad P\{X_1 = x_1 | X_2 - X_1 = w\} \to \frac{\lambda^{x_1+w/2}}{x_1!(x_1+w)!I_{|w|}(2\sqrt{\lambda})}$$

where

$$(3.33) \qquad I_v(x) = (\frac{x}{2})^v \sum_{k=0}^{\infty} \frac{(x^2/4)^k}{k!\Gamma(v+k+1)}$$

is the modified Bessel function of order v. For our purpose v will be a non-negative integers and $x > 0$. The proof is given in Appendix A. It essentially involves checking that limit of infinite sums is the infinite sum of limits using uniform convergence and Stirling's formula.

We note that if $X \sim P_0(\lambda)$ and $Y \sim P_0(\mu)$ are two independent Poisson r.v.'s with parameters λ and μ, then it follows from Harris (1968) that we have the probability mass function:

$$P\{Y = y | X - Y = w\} = \frac{\rho^{y+(w/2)}}{y!(y+w)!I_{|w|}(2\sqrt{\rho})},$$

where

$$\rho = \lambda\mu$$

$$y = \alpha, \alpha+1, \alpha+2, \ldots, \alpha = \max(0, -w);$$

$$w = 0, \pm 1, \pm 2, \ldots,$$

and this is exactly of the same form as the limiting distribution (3.32). This is somewhat expected as one can see from the mode of convergence of Theorem 3.1.

More generally, if there are $k > 2$ components in a parallel system, we observe

$$X_i \sim NB(r_i, p_i),\ i = 1,2,\ldots, k.$$

Define the r.v.'s W_i, $i = 1,2,\ldots, k$ and the vectors U and u as in (3.2). We obtain the conditional distribution of W_1 given $U = u$ by relabeling (3.2):

$$P\{W_1 = w_1 | U = u\} = b(w_1, u)\theta^{w_1} / \sum_t b'(t, u)\theta^t \tag{3.34}$$

where $b'(x, y)$ is defined in (3.3),

$$\theta = q_1 q_2 \cdots q_k,$$

and w_1 satisfies the double inequalities (3.4). Note $1 - \theta$ is the system reliability in this case.

Corollary 3.1.

Let $X_i \sim NB(r_i, p_i)$, $i = 1,2,\ldots, k$ be k independent r.v.'s. If $r_i \to \infty$, $q_i = 1 - p_i \to 0$, $i = 1,2,\ldots, k$ in such a way that $\prod_{i=1}^{k} r_i q_i = \rho$ remains fixed, where ρ is a positive constant, then for each w_1 and $u = (w_2,\ldots, w_k)$, (3.34) has the limiting value

$$h(w_1, \rho) / \sum_t h(t, \rho) \tag{3.35}$$

where

$$\rho = \prod_{i=1}^{k} r_i q_i, \text{ and } h(t, \rho) = \frac{\rho^t}{t! \prod_{i=2}^{k} (t+w_i)!}.$$

Another approximation for a parallel system is to observe that if $X \sim NB(r, p)$, then the r.v. Y', defined as $Y' = 2p(r+X)$ converges to $\chi^2(2r)$ in distribution as $p \to 0$. (See Chapman (1952)). It is also true $Y = 2pX \xrightarrow{d} \chi^2(2r)$. Thus if $X_i \sim NB(r_i, p_i)$, $i = 1,2$, then

$$Y_i \equiv 2p_iX_i \xrightarrow{d} U_i, \quad i = 1,2,$$

where $U_i \sim \chi^2(2r_i)$, $i = 1,2$. Now U_1, U_2 are two independent chi-square r.v.'s and the density $f_U(u)$ of the product $U \equiv U_1U_2$ follows from Malik (1968)

$$(3.36) \qquad f_U(u) = \frac{uK_{r_2-r_1}(\sqrt{u})}{r_2!(r_1-1)!2^{r_1+r_2-1}}, \quad u > 0$$

where $K_r(x)$ is the modified Bessel function of the second kind of order r. It follows $V \equiv X_1X_2 = Y_1Y_2/4p_1p_2$ has approximately the density function

$$(3.37) \qquad f_V(v) = \frac{2v(p_1p_2)^{r_2}}{r_2!(r_1-1)!} \frac{K_{r_2-r_1}(2\sqrt{p_1p_2v})}{(\sqrt{p_1p_2v})^{r_1-r_2}}, \quad v > 0.$$

This can be used to construct approximate confidence limits for p_1p_2 if both p_1 and p_2 are small, for any values of r_1 and r_2 not necessarily large. Unlike previous solutions, the confidence limits constructed here are unconditional.

3.5 Other Parametric Functions and Other Systems.

So far our discussion has been centered on creating probability distributions depending on the parametric function θ in (2.1) and its particular cases which are relevant to a purely series or purely parallel systems of the attribute failure model. It should be pointed

out that in the expression for θ in (2.1), if we replace some or all of (p_i)'s and (p_j')'s by (q_i)'s and/or (q_j')'s, and let θ' represent the new parametric function, then a distribution depending on θ' only can also be obtained and statistical inference about θ' carried out. For instance, $\theta' = q_1p_2q_3p_4/p_1'p_2'q_3'$. The rule is for each failure probability q_i (or q_i') in θ', we sample the corresponding Bernoulli population π_i (or π_i') until r_i^{th} (or s_i^{th}) success; while for each p_j (or p_j') in θ', we sample until r_j^{th} (or s_j^{th}) failure. Then a distribution similar to (2.6) can be obtained which depends on the desired θ'.

3.5.1 A parallel-series system.

We now discuss a way to provide confidence limits for the reliability of a system having both parallel and series elements. Suppose there are three independent, dissimilar components π_1, π_2 and π_3 in a given system. Four different ways have been found to connect them (P_s stands for system reliability in each configuration).

a. All components are in series, and $P_s = p_1p_2p_3$.

b. All components are in parallel, and $P_s = 1 - q_1q_2q_3$.

c. Two in parallel form a subsystem and the third component is connect to it in series. Here $P_s = (1-q_1q_2)p_3$.

d. Two in series form a subsystem and the third component is connected in parallel to it. Here $P_s = 1 - (1-p_1p_2)q_3$.

We have already given confidence limits for cases a. and b. Case c. and d. can be called "Parallel-series systems." We shall now find a distribution depending on the P_s in Case c. Case d. can be treated in a similar way.

Lemma 3.1. (Mixture of a negative binomial distribution and a binomial distribution.)

Suppose $X \sim NB(r, p_1)$ and $(Y|X = n) \sim B(n, p_2)$, then $Y \sim NB(r, p_1/(1-q_1q_2))$. In particular, if $p_1 = p_2 = p$, then $Y \sim NB(r, 1/(1+q))$, where $q = 1 - p$.

Proof:

The probability generating function (PGF) of X is:

$$\varphi_x(t) = (\frac{p_1}{1-q_1t})^r.$$

When $X = 1$, the PGF of $(Y|X = 1)$ is:

$$\varphi_{Y|X=1}(t) = (q_2 + p_2t).$$

By a theorem for the PGF of a random sum (see, Feller(1968),p. 287), the PGF of Y is:

$$\varphi_y(t) = \varphi_x(\varphi_{Y|X=1}(t)) = (\frac{p_1}{1-q_1(q_2+p_2t)})^r = (\frac{p_1/(1-q_1q_2)}{1-(q_1p_2/(1-q_1q_2))t})^r$$

which we recognize to be the PGF of $Y \sim NB(r, p_1/(1-q_1q_2))$.

To find confidence limits for P_s in Case c., we perform IBS on component π_1 until r_1^{th} success. If X represents the outcome, then $X \sim NB(r, p_1)$. Suppose n is the observed value of X, we perform ordinary binomial sampling on component π_2 until n observations are obtained. Let $(Y|X = n)$ represent the outcome. Then $(Y|X = n) \sim B(n, p_2)$, and $Y \sim NB(r, p_1/(1-q_1q_2))$. Now perform IBS on components π_1, π_2 and π_3 until first success, first failure, and r_3^{th} failure, respectively. Let X_1, X_2, and X_3 represent the outcomes. Then $X_1 \sim NB(1, p_1)$, $X_2 \sim NB(1, q_2)$ and

$X_3 \sim NB(r_3, q_3)$. Thus we have four independent negative binomial r.v.'s Y, X_1, X_2 and X_3. Let $W_1 = X_3$, $W_2 = X_1 - X_3$, $W_3 = X_2 - X_3$, $V = Y + X_3$, as in (2.3). By (2.6) we have

$$(3.38) \qquad P\{X_3 = x \mid X_1 - X = w_2,\ X_2 - X_3 = W_3,\ Y + X_3 = v\}$$

$$= b(x, u)\theta^x / \sum_t b(t, u)\theta^t, \text{ where}$$

$$b(x, u) = \binom{r_3+x-1}{x}\binom{r+v-x-1}{v-x},\ \theta = (1-q_1q_2)p_3$$

and x is any non-negative integer satisfying $\max(0, \max_{2\le i\le 3}(-w_i)) \le x \le v$. The values of t in the summation in the denominator are the same as the above-mentioned values of x.

For Case d., we need a distribution depending on system failure probability $P_f \equiv 1 - P_s = (1-p_1p_2)q_3$, and this can be found by interchanging the labels for success and failure in Case c.

3.5.2 A series system with some identical components.

If we have a purely series system consisting of i_j components of type j, $j = 1,2,\ldots, k$ and i_j being any positive integer, the system reliability is

$$(3.39) \qquad P_s = p_1^{i_1} p_2^{i_2} \cdots p_k^{i_k}$$

where p_j is the reliability of each type j component, $j = 1,2,\ldots, k$. This is, of course, a special case of (3.2) and IBS on each of the $N \equiv \sum_{j=1}^{k} i_j$ components is required in order to obtain a distribution like (3.2) which depends on P_s in (3.39). Another way to accomplish this purpose is to do IBS on each type of component, thereby reducing the number of samples from N to k. Let us illustrate this by

taking a series system of three components π_1, π_2, π_2', two of them (π_2 and π_2') are identical. Then $P_s = p_1 p_2^2$. We perform IBS on π_i until r_i^{th} failure and obtain the r.v.'s $X_i \sim NB(r_i, q_i)$, $i = 1, 2$. By Lehmann and Scheffe's Theorem mentioned in Section 2.2, it is easy to see the conditional distribution:

$$(3.40) \qquad P\{X_1 = x | X_2 - 2X_1 = w\}$$

depends on the system reliability P_s only. Although this method cuts down sampling efforts, the distribution becomes "more discrete" in the sense that probabilities of (3.40) are spread among fewer sample points, and randomization becomes more important.

CHAPTER IV

COMPOUND PROBABILITY DISTRIBUTIONS AND THEIR APPLICATIONS

Let the r.v. X have the distribution function $F_X(x; \theta)$ for each given value of the parameter θ. Suppose now that θ itself is a r.v. Y having distribution function $F_Y(y)$. Then the absolute (or marginal) distribution of X is sometimes called a "compound distribution," and Y is called the "compounder." In the present reliability problems, the usefulness of the compounding technique depends on its ability to produce desired reliability expressions in the compound distribution. We have already used this technique in Lemma 3.1 where the compounder is $X \sim NB(r, p_1)$ and the compound distribution (of Y) is found to be negative binomial. The technique produces the probability distribution (3.38) which depends on the desired reliability expression $(1-q_1q_2)p_3$ for a parallel-series system. In this chapter we exploit the properties of the compound Poisson distribution and find confidence limits for a large class of functions of Bernoulli parameters. A special case of these functions is the reliability expression of any monotonic system consisting of independent components in the attribute failure model. It is conjectured that the technique is also applicable to the non-monotonic systems. As we shall see later, the basic feature of our approach is to require that the number of observations taken from any Bernoulli population be a r.v. following a known Poisson distribution. The distribution theoretic facts which we use are given for example by Feller (1968), p. 301, and have been used previously by Paulson (1967) in a binomial selection problem. In practice, this can be carried out using the computer-generated uniform random variates or the published tables of the Poisson distribution such as General Electric Co. (1962) and a

table of random numbers, such as Rand Corporation (1955).

4.1 Compound Poisson Distributions and System Reliability.

Let $Y \sim P_0(\lambda)$, where λ is any positive number fixed in advance of the binomial sampling. Suppose $(X|Y = n) \sim B(n, p)$. Then it is well known that

$$(4.1) \qquad X \sim P_0(\lambda p).$$

X is called a compound Poisson r.v. In the practical sampling situation[§] if $n = 0$, we take $X = 0$. Now if $Y \sim P_0(\lambda_i)$ and

$$(4.2) \qquad (X_i|Y_i = n_i) \sim B(n_i, p_i), \; i = 1,2,\ldots, k,$$

then

$$X_i \sim P_0(\lambda_i p_i), \; i = 1,2,\ldots, k.$$

The X_i will be independent when the Y_i are. The conditional distribution of X_1 given that $X_2 - X_1 = w_2, \ldots, X_k - X_1 = w_k$ is found (see Harris (1968)) to depend on the product of the k Poisson parameters $\theta \equiv (\lambda_1 p_1)(\lambda_2 p_2)\ldots(\lambda_k p_k) = \lambda(p_1 p_2 \ldots p_k)$, where $\lambda = \prod_{i=1}^{k} \lambda_i$. For example, when $k = 2$, we have (see the discussion following Theorem 3.1 of Section 3.4):

$$P\{X_1 = x|X_2 - X_1 = w\} = \frac{\mu^{x+(w/2)}}{x!(x+w)!I_{|w|}(2\sqrt{\mu})},$$

where

[§]In the practical situations the case $Y = 0$ is very unlikely, because the sample size usually desired is moderate, say 30 or more. If we choose λ to be 30, then $P\{Y = 0\}$ is negligible.

$$\mu = \lambda_1\lambda_2 p_1 p_2$$

$$w = 0, \pm 1, \pm 2, \pm 3, \ldots$$

$$x = \alpha, \alpha+1, \alpha+2, \ldots$$

$$\alpha = \max(0, -w)$$

and $I_{|w|}$ is the Bessel function defined in (3.33).

Since λ_i is preassigned, $i = 1, 2$, we can use this conditional distribution to construct confidence limits for $\mu' \equiv p_1 p_2$. More generally, confidence limits for the reliability $\theta' \equiv p_1 p_2 \cdots p_k$ of a series system can be found.

To find confidence limits for the reliability $\theta'' = 1 - q_1 q_2 \cdots q_k$ of a parallel system, no additional sampling is required. We simply let $Z_i = n_i - X_i$ denote the number of failures observed in the i^{th} sample. Then $Z_i \sim P_0(\lambda_i q_i)$ and the same technique is seen to apply.

4.2 Multiple-Stage Compound Poisson Distributions and System Reliability.

In the previous section, each of the k binomial distributions was compounded individually with a Poisson distribution to give k compound Poisson distributions. We shall loosely call this type of compounding procedure "horizontal compounding." Instead of this procedure, if we perform the compounding process successively on the k binomial distributions as shown below, a single "multiple-stage" compounding Poisson distribution results and interesting facts relevant to the estimation of the system reliability emerge.

4.2.1 Multiple-stage compound Poisson distributions and the estimation of various functions of Bernoulli parameters.

Let $Y \sim P_0(\lambda)$, where $\lambda > 0$ is preassigned. Suppose

(4.3) $(X_1|Y = n_1) \sim B(n_1, p_1)$.

then, as before:

(4.4) $X_1 \sim P_0(\lambda p_1)$.

Now suppose

$$(X_2|X_1 = n_2) \sim B(n_2, p_2).$$

Then, by the same argument,

$$X_2 \sim P_0(\lambda p_1 p_2).$$

Repeating this process, and letting k be any positive integer, suppose

$$X_{k-1} \sim P_0(\lambda p_1 p_2 \cdots p_{k-1})$$

and

$$(X_k|X_{k-1} = n_k) \sim B(n_k, p_k).$$

Then

(4.5) $X_k \sim P_0(\lambda\theta)$

where $\theta = p_1 p_2 \cdots p_k$. The distribution of X_1 can be called the first stage or single stage compound Poisson distribution, while that of $X_2, X_3, \ldots, X_k$ can be called a second stage, a third stage,..., k^{th} stage or a MSCP (Multiple-Stage Compound Poisson) distribution.

Note that whenever a new observation of Y is generated (as described in Section 4.1), we can go through the k-stage compounding process described above and obtain a new observation of X_k.[§] Thus

[§]If the observed n_i is zero for the i^{th} stage, $i < k$, we take $x_k = 0$.

a sequence of observations $X_{k_1}, X_{k_2}, \ldots, X_{k_n}, \ldots$ on the Poisson r.v. X_k can be obtained and the problem of constructing confidence limits for $\lambda\theta$ or equivalently for θ (because λ is known) reduces to the elementary problem of finding confidence limits for the parameter of a Poisson distribution.

We shall call the type of compounding procedure for obtaining a MSCP r.v. "vertical compounding."

The parameter θ is the reliability of a series system. For the reliability of a parallel system $\theta'' = 1 - q_1 q_2 \cdots q_k$, we observe the number of failures in the compounding process and obtain at the k^{th} stage:

$$(4.6) \qquad Z_k \sim P_0(\lambda q_1 q_2 \cdots q_k)$$

which can be used to construct confidence limits for θ''.

As in the familiar case of constructing confidence limits for the parameter of a Poisson distribution, the lower or upper confidence limit for θ and θ'' constructed from (4.5) and (4.6) can be made to be uniformly most accurate by randomization.

From the k Bernoulli populations we have created MSCP r.v.'s X_k and Z_k whose parameters involve the product of Bernoulli parameters. More generally, let

$$(4.7) \qquad f(p, q) \equiv \prod_{i=1}^{k} \prod_{j=1}^{k} p_i^{\delta_i} q_j^{\gamma_j}$$

where $\delta_i = 0$ or 1, $\gamma_j = 0$ or 1, $i, j = 1,2,\ldots, k$;

$p = (p_1,\ldots, p_k)$, $q = (q_1,\ldots, q_k)$

be any product of the p_i's and q_j's. It is clear that by the same vertical compounding technique we can create

a sum X of MSCP r.v.'s so that the parameter of the Poisson r.v. X involves the sum of the products $f(p, q)$ of Bernoulli parameters (using the fact that the sums of k independent Poisson r.v.'s is a Poisson r.v. whose parameter is the sum of the individual parameters). Furthermore, we show that given a real polynomial $A(p, q)$ of any arbitrary degree in the (parameter) variables p_i and q_j of p and q, we can create a Poisson r.v. w whose parameter is $A(p, q) - C$, where C is the known constant term of $A(p, q)$. To see this, we know $A(p, q)$ can be expressed as

$$A(p, q) = \sum_{v=1}^{m} \lambda_v f_v(p, q) + C \tag{4.8}$$

where λ_v's are arbitrary real coefficients, m is a positive integer, and each f has the form (4.7) except that the exponents δ_i, γ_i can be $0,1,2,\ldots$. For each product term $\lambda_v f_v(p, q)$, we note its sign is determined by the sign of λ_v. If $\lambda_v > 0$ for all v, then we can, as before, create a Poisson r.v. X whose parameter is $\sum_v \lambda_v f_v(p, q) = A(p, q) - C$ (we treat $p_i^{\ell_i}$ as:

$$p_i^{\ell_i} = p_i p_i \cdots p_i, \ \ell_i\text{-fold}$$

and the same for $q_j^{m_j}$). If $\lambda_v < 0$ for some v, then there is no way we can create a MSCP r.v. whose parameter is the negative term $\lambda_v f_v(p, q)$, and the above procedure fails. One way to overcome this difficulty is to observe that we can express $\lambda_v f_v(p, q)$ as a sum of several positive product terms and the constant λ_v. More precisely, let $\lambda_v' = -\lambda_v > 0$. Then as we show by example in the next paragraph,

$$(4.10) \qquad \lambda_v f_v(p, q) = \lambda_v' \sum_i f_{v_i}(p, q) + \lambda_v$$

where $f_{v_i}(p, q) > 0$ is of the same form as $f_v(p, q)$, only the powers ℓ_i, m_j of the variables p_i's and q_j's may be less than those of $f_v(p, q)$.

The rule calls for expressing one of the q_i's (or p_i's) in terms of $(1-p_i)$ (or $1-q_i$) in each step, so that the original negative term $\lambda_v f_v(p, q)$ are split into two new product terms, one positive, one negative. Observe that the new negative product term has one less factor in its product now. The process is continued as long as there are still negative product terms. Example: let $\lambda = -1$,

$$-p_1q_2p_3q_4 = -(1-q_1)q_2p_3q_4 = q_1q_2p_3q_4 - (1-p_2)p_3q_4$$

$$= q_1q_2p_3q_4 + p_2p_3q_4 - (1-q_3)q_4$$

$$= q_1q_2p_3q_4 + p_2p_3q_4 + q_3q_4 - (1-p_4)$$

$$= q_1q_2p_3q_4 + p_2p_3q_4 + q_3q_4 + p_4 - 1.$$

More examples are given in P_s' of Table 4.1.

Substituting (4.10) into (4.8), and doing the same conversion for each negative term in (4.8), we obtain at the end that $A(p, q)$ can be expressed as a sum of positive product terms plus a constant C', where C' is a known constant obtained from combining C with the negative constants λ_v's in the conversion process. And the problem recuces to the case where $\lambda_v > 0$ for all v mentioned above.

One final extension is the following:

If $A_1(p, q)$, $A_2(p, q)$ are two polynomial of the form (4.8), then $R(p, q) = A_1(p, q)/A_2(p, q)$ is a rational function of the Bernoulli parameters p, q. Let $i = 1,2$. Since we can create a Poisson r.v. W_i with

$$(4.11) \qquad W_i \sim P_0(A_i(p, q) - C_i)$$

where C_i is a constant, it is clear that we can arrange W_1, W_2 to be independent, and find that the conditional distribution of W_1 given the difference $W_2 + W_1 = w$, depends on the ratio $[A_1(p, q) - C_1]/[A_2(p, q) - C_2]$ only. Since in (4.11), C_i is a known constant, thus if $Y \sim P_0(C_i)$, then $Y + W_i \sim P_0(A_i(p, q))$. Thus we can construct confidence limits for this rational function of the Bernoulli parameters.

4.2.2 Applications to the monotonic structures and other systems. Sampling considerations.

One immediate application of the above result is that we can construct confidence limits for the reliability of a large class of systems called monotonic structures or systems when their components are independent (Mine (1959), Barlow and Proschan (1965)). Birnbaum, Esary and Saunders (1961) call them coherent structures. The class includes series, parallel, "fail-safe," and r-out-of-k systems. A monotonic structure assumes only two states of performance, either complete success or complete failure; so does each of its components. Thus it is natural to treat them using Bernoulli r.v.'s. In the following discussion, the meaning and definitions of the terms like

structure function φ, cut, path as well as the results in (4.12) below are given in Chapter 7 of Barlow and Proschan (1965). Our purpose here is to show that the reliability of a monotonic structure can be expressed in terms of a polynomial in the Bernoulli parameters of the component populations.

Let $x = (x_1, x_2, \ldots, x_k)$ where x_i is the indicator function indicating the success or failure of component π_i, $i = 1,2,\ldots, k$. Correspondingly, let $X = (X_1, X_2, \ldots, X_k)$, where X_i is the Bernoulli r.v. associated with x_i. The reliability of a monotonic structure with structure function $\varphi(x)$ is given by $E\varphi(X)$. It is known (see, for example, Barlow and Proschan (1965)) that a monotonic structure can be represented in terms of paths A_j, $j = 1,2,\ldots, r$ of the structure and

$$(4.12) \qquad \varphi(x) = 1 - \prod_{j=1}^{r} [1 - \alpha_j(x)]$$

where $\alpha_j(x)$ is the binary function associated with the path A_j, and

$$(4.13) \qquad \alpha_j(x) = \prod_{i \in A_j} x_i, \quad j = 1,2,\ldots, r.$$

Thus

$$\varphi(x) = 1 - \prod_{j=1}^{r} [1 - \prod_{i \in A_j}^{r} x_i]$$

and

$$(4.14) \qquad E\varphi(X) = 1 - E(1 - \prod_{i \in A_1} X_i)(1 - \prod_{i \in A_2} X_i)\ldots(1 - \prod_{i \in A_r} X_i).$$

Since $X_1, \ldots, X_k$ are independent Bernoulli r.v.'s, and $EX_i = p_i$, $i = 1,2,\ldots, k$, it is clear that after multiplying the r factors under the expectation sign and then take expectation, we can express each resulting term in the form:

$$(4.15) \qquad EX_1^{\ell_1} \cdots X_k^{\ell_k} = p_1^{\delta_1} \cdots p_k^{\delta_k}$$

where $\delta_i = 0$ if $\ell_i = 0$ and $\delta_i = 1$ if $\ell_i = 1,2,\ldots$. Thus $E\varphi(X)$ is expressed as a polynomial in the Bernoulli parameters p_i, $i = 1,2,\ldots, k$, and we can construct confidence limits for it as shown in the previous section. It should be pointed out that the structure function $\varphi(x)$ can be represented in terms of cuts of the system. In that case the system reliability $E\varphi(X)$ can be expressed in terms of a polynomial in which each individual term is in the form:

$$q_1^{\gamma_1} q_2^{\gamma_2} \cdots q_k^{\gamma_k}$$

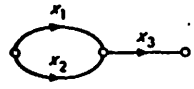
where γ_i, as in (4.7), is either 0 or 1, $i = 1,2,\ldots, k$.

For non-monotonic systems consisting of independent components in the attribute failure model, the system reliabilities can also be expressed in terms of a polynomial in the Bernoulli parameters (of the components), so that the technique developed in the previous section should be applicable.

If the reliability P_s of a system (monotonic or not) can be expressed in terms of a polynomial in its component parameters p, q, then

$$(4.16) \qquad P_s = \sum_{v=1}^{m} \lambda_v f_v(p, q) + C$$

as in (4.8). Without loss of generality, assume $\lambda_v f_v(p, q) > 0$ for each v. In (4.16), a component parameter p_i (or q_i) may appear in more than one product term $\lambda_v f_v(p, q)$. Therefore, to apply the vertical compounding technique, more than one binomial sample from this component is needed. For example, in a parallel-series system (Fig. 4.0) with three components π_i, $i = 1,2,3$, we have:



$$P_s = (1-q_1q_2)p_3 = p_3p_1 + p_3p_2q_1 .$$

Figure 4.0. A parallel-series system.

Here p_3 appears in both product terms of P_s. If by vertical compounding, we create

$$W_1 \sim P_0(\lambda p_3 p_1)$$

and

$$W_2 \sim P_0(\lambda p_3 p_2 q_1)$$

to obtain

$$W \equiv w_1 + w_2 \sim P_0(\lambda P_s),$$

we would need two binomial samples from component π_3 in the process of obtaining W (it is seen that two binomial samples from π_1 are also needed). We now describe a method to reduce multiple samples from the same component and show that in two special classes of P_s, $\mathfrak{A}$ and $\mathfrak{B}$, one binomial sample from each component is sufficient to estimate P_s.

Let π_i, $i = 1,2,\ldots, k$ be the k independent components (similar or dissimilar) of a system with parameters p or q (defined

immediately after (4.7)). For $j = 1,2,\ldots, k$, define a recursive formula as follows:

$$(4.17) \qquad P_s^{(j)} = \delta_j'(1-u_j) + \delta_j u_j P_s^{(j-1)}$$

with

$$P_s^{(0)} = 1$$

where

$$\delta_j = 0 \text{ or } 1, \; j = 1,2,\ldots, k$$
$$\delta_j' = 0 \text{ or } 1, \; j = 1,2,\ldots, k$$

and $(u_1, u_2,\ldots, u_j)$ is either a permutation of $(p_1, p_2,\ldots, p_j)$ or of $(q_1, q_2,\ldots, q_j)$. For example:

$$P_s^{(1)} = \delta_1'(1-u_1) + \delta_1 u_1$$

$$P_s^{(2)} = [\delta_2'(1-u_2) + \delta_2 u_2(\delta_1'(1-u_1) + \delta_1 u_1)]$$

$$P_s^{(3)} = \{\delta_3'(1-u_3) + \delta_3 u_3[\delta_2'(1-u_2) + \delta_2 u_2(\delta_1'(1-u_1) + \delta_1 u_1)]\}.$$

Note the way $P_s^{(2)}$ is built up from $P_s^{(1)}$, and $P_s^{(3)}$ is built up from $P_s^{(2)}$. For the parallel-series system mentioned above, P_s can be written as

$$P_s = P_s^{(3)} = \{0(1-p_3) + p_3[(1-q_1) + q_1(0(1-p_2) + p_2)]\}$$

or

$$P_s = P_s^{(3)} = p_3[p_1 + q_1(p_2)]$$

and this belongs to the class $\mathcal{G}$ to be specified below.

We now specify the classes $\mathcal{G}$ and $\mathcal{B}$ in terms of (4.17). Let k_i, $i = 1,2,\ldots, \ell$, and ℓ be positive integers. Suppose a system consists of $k = \sum_{i=1}^{\ell} k_i$ components. Then for $\ell \geq 1$ define

$$\mathcal{G} = \{P_s: P_s \text{ can be expressed as } P_s = C \pm \sum_{i=1}^{\ell} P_s^{(k_i)}\},$$

and for $\ell \geq 2$ define

$$\mathcal{B} = \{P_s: P_s \text{ can be expressed as } P_s = C \pm \prod_{i=1}^{\ell} P_s^{(k_i)}\}$$

where $C = 0$ or 1. For each $P_s \in \mathcal{G}$, we shall create a single MSCP r.v. W to estimate P_s; for $P_s \in \mathcal{B}$ we shall create $\ell > 1$ MSCP r.v.'s $w_1, w_2, \ldots, w_\ell$, and use the conditional r.v. $(w_1 = z | w_2 - w_1 = z_2, \ldots, w_\ell - w_1 = z_\ell)$ to estimate P_s. The procedures are best illustrated by considering examples. In the next section we consider a system which consists of four components π_i, $i = 1,2,3,4$, with corresponding component reliabilities p_i, $i = 1,2,3,4$. We also consider other well-known systems in Section 4.2.4.

4.2.3 Reliability estimation for a system of four components.

Ten possible ways have been found to connect four components (see Shooman (1968), p. 131). The ten system configurations and their reliabilities P_s are listed in Table 4.1. The configurations are easy to understand. For example, the first system is a purely series system, while the tenth system is a purely parallel system. P_s in each case is obtained by inspection. We also rewrite P_s in the alternate form P_s' of which the eight systems (1, 2, 3, 4, 5, 8, 9 and 10) are in the desired form $C \pm P_s^{(k)}$. Therefore they belong to the class $\mathcal{G}$ (with $\ell = 1$), and the remaining two systems (6 and 7)

involve product forms. They belong to the class $\mathcal{B}$ (with $\ell = 2$ in both systems). We shall first illustrate the procedure for G using system 1 in Table 4.1. (This is for illustration purposes only; actually for system 1, it is simpler to use the form $P_s = p_1p_2p_3p_4$, as in deriving (4.5)). We have

$$(4.18) \qquad P_s' = 1 - \{q_4 + p_4[q_3 + p_3(q_2 + p_2(q_1)]\}.$$

First multiply out inside the braces in (4.18) (the original order of terms should not be disturbed) to obtain another form of P_s'. Call it P_s^*:

$$(4.19) \qquad P_s^* = 1 - \{q_4 + p_4q_3 + p_4p_3q_2 + p_4p_3p_2q_1\}.$$

For each term within the braces of P_s^*, we wish to find a corresponding (compound) Poisson r.v. whose parameter is this particular term multiplied by a known constant λ. These Poisson r.v.'s should be independent so that the sum of them is again a Poisson r.v. whose parameter is $\lambda\ (P^* \pm C)$. In (4.19), we need to create four independent Poisson r.v.'s X_i, $i = 1,2,3,4$, with

$$(4.20) \qquad \begin{aligned} X_4 &\sim P_0(\lambda q_4) \\ X_3 &\sim P_0(\lambda p_4 q_3) \\ X_2 &\sim P_0(\lambda p_4 p_3 q_2) \\ X_1 &\sim P_0(\lambda p_4 p_3 p_2 q_1) \end{aligned}$$

so that $X \equiv X_1 + X_2 + X_3 + X_4 \sim P_0(\lambda(1-P_s^*))$. We recall in a first

Table 4.1. Reliability Configuration for Four Independent, Dissimilar Components.

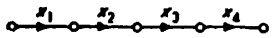
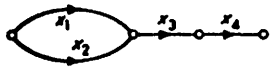
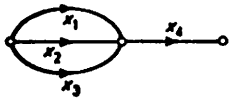
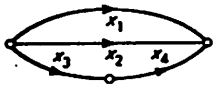
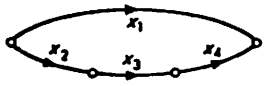
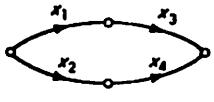
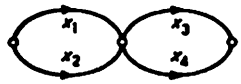
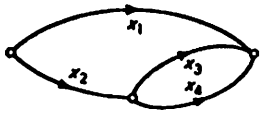
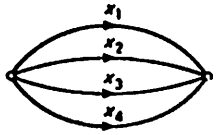
System Configuration	P_s	P'_s (Alternate Form of P_s)
1. 	$p_1p_2p_3p_4$	$1+\{q_4+p_4[q_3+p_3(q_2+p_2(q_1))]\}$
2. 	$p_3p_4(1-q_1q_2)$	$p_4p_3[p_1+ q_1(p_3)]$
3. 	$p_4(1-q_1q_2q_3)$	$p_4\{p_3+q_3[p_2+q_2(p_1)]\}$
4. 	$1-q_1q_2(1-p_3p_4)$	$1-\{q_1q_2[q_3+p_3(q_4)]\}$
5. 	$1-q_1(-p_2p_3p_4)$ $= \{p_1+q_1(p_2p_3p_4)\}$	$1-\{q_1[q_3+p_3(q_4+p_4(q_2))]\}$
6. 	$1-(1-p_1p_3)(1-p_2p_4)$	$1-\{q_3+p_3(q_1)\}\{q_4+p_4(q_2)\}$

Table 4.1 (cont.).

7.		$(1-q_1q_2)(1-q_3q_4)$	$\{p_2+q_2(p_1)\}\{p_4+q_4(p_3)\}$
8.		$p_1\{1-q_2(1-p_3p_4)\}$	$1-\{q_1+p_1[q_2(q_3+p_3(q_4))]\}$
9.		$1-q_1\{1-p_2(1-q_3q_4)\}$	$p_1+q_1\{p_2[p_3+q_3(p_4)]\}$
10.		$1-q_1q_2q_3q_4$	$p_4+q_4\{p_3+q_3[p_2+q_2(p_1)]\}$

stage compound Poisson distribution, the number of successes and failures are independent r.v.'s (see Feller (1968), p. 301). Clearly this also holds true in any higher stage compound Poisson distribution. The implication of this unexpected result is that in our compounding process, both the information about the number of successes and the number of failures can be utilized. The form of P_s in the classes $\mathfrak{A}$ and $\mathfrak{B}$ are designed to systematically exploit this property.

Let $Y \sim P_0(\lambda)$ where λ is preassigned. The order of taking binomial samples is indicated by the order of appearance of the component parameters p_i or q_i in the form P_s'. In our example we should take one sample from each of the four components in the order π_4, π_3, π_2 and π_1. The following four steps explain the procedure:

Step 1. Compound component π_4 with Y.[§] Let $X_4 = x_4$ and $X_4' = x_4'$ be the numbers of failures and successes in $y = x_4 + x_4'$ trials, where y is the observed value of Y. We have

$$X_4 \sim P_0(\lambda q_4)$$

$$X_4' \sim P_0(\lambda p_4)$$

and X_4, X_4' are independent. X_4 can be used to estimate the first parametric term of P_s', i.e., q_4.

Step 2. Compound component π_3 with X_4', using the observed value $X_4' = x_4'$ in Step 1. Let $X_3 = x_3$ and $X_3' = x_3'$ be the number of failures and successes in $x_4' = x_3 + x_3'$ trials. We have:

[§]This is a shortened expression to mean that we do binomial sampling on the component π_4 with sample size determined by an observation of a Poisson r.v. Y. This is true for π_i, i=1,2,3,4.

$$X_3 \sim P_0(\lambda p_4 q_3)$$

$$X_3' \sim P_0(\lambda p_4 p_3)$$

and X_3, X_3' are independent. X_3 can be used to estimate the second parametric term of P_s', i.e., $p_4 q_3$.

Step 3. Compound component π_2 with X_3', using the observed value $X_3' = x_3'$ in Step 2. Let $X_2 = x_2$, $X_2' = x_2'$ be the numbers of failures and successes in $x_3' = x_2 + x_2'$ trials. We have:

$$X_2 \sim P_0(\lambda p_4 p_3 q_2)$$

$$X_2' \sim P_0(\lambda p_4 p_3 p_2)$$

and X_2, X_2' are independent. X_2 can be used to estimate the third parametric term of P_s', i.e., $p_4 p_3 q_2$.

Step 4. Compound component π_1 with X_2', using the observed value $X_2' = x_2'$ in Step 3. Let $X_1 = x_1$, $X_1' = x_1'$ be the numbers of failures and successes in $X_2' = x_1 + x_1'$ trials. We have

$$X_1 \sim P_0(\lambda p_4 p_3 p_2 q_1)$$

$$X_1' \sim P_0(\lambda p_4 p_3 p_2 p_1)$$

and X_1, X_1' are independent. We need only X_1 to estimate the last parametric term of P_s', i.e., $p_4 p_3 p_2 q_1$. This finishes the sampling process.

In looking back, we see in each step we did a compounding which requires a binomial sample from the relevant component, and in total we draw four binomial samples, one for each of the four components

as we desired. The relationship between the r.v.'s we have created can be depicted as in Figure 4.1. From the independence of X_j and X'_j for each j it is seen that X_1, X_2, X_3, X_4 are jointly independent. Therefore

$$(4.21) \qquad X \equiv X_4 + X_3 + X_2 + X_1 \sim P_0(\lambda(q_4 + p_4q_3 + p_4p_3q_2 + p_4p_3p_2q_1))$$

or $X \sim P_0(\lambda(1-P'_s))$ as desired.

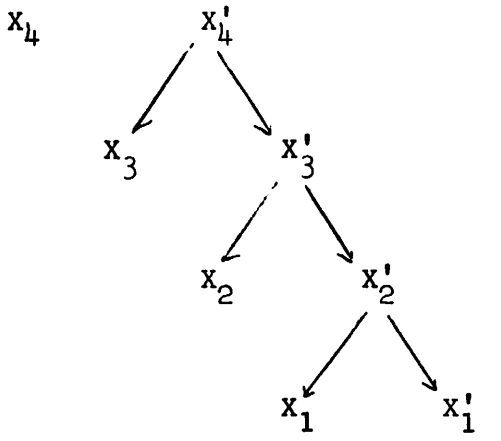


Figure 4.1. Relationship between the r.v.'s created.

If $P_s \in \mathcal{R}$, then P_s can be written in the form:

$$(4.22) \qquad P_s = C \pm \prod_{i=1}^{\ell} P_s^{(k_i)}, \ \ell > 1, \ C \text{ is known.}$$

For each fixed i, we take k_i binomial samples and create a MSCP r.v. W_i with

$$W_i \sim P_0(\lambda_i P_s^{(k_i)})$$

as in the case for $\mathcal{G}$ with $\ell = 1$. $\lambda_i > 0$ is a known constant.

The r.v.'s $W_1, \ldots, W_\ell$ are independent. Then by the result derived in Section 4.1 the conditional distribution:

$$P\{W_1 = z | W_2 - W_1 = z_2, \ldots, W_\ell - W_1 = z_\ell\}$$

depends on the parameter $\lambda \prod_{i=1}^{\ell} P_s^{(k_i)}$, where $\lambda = \prod_{i=1}^{\ell} \lambda_i$, as desired. We illustrate this case using System 6 in Table 4.1. We have

$$P_s = P_s' = 1 - \{q_3 + p_3(q_1)\}\{q_4 + p_4(q_2)\}.$$

This is in the form (4.22) with $C = 1$, $\ell = 2$ and $k_1 = k_2 = 2$

$$P_s^{(k_1)} = q_3 + p_3(q_1)$$

$$P_s^{(k_2)} = q_4 + p_4(q_2).$$

As in the case for class $\mathcal{C}$ with $\ell = 1$, we create:

$$W_1 \sim P_0(\lambda_1(q_3 + p_3 q_1)),$$

using one binomial sample from each of the two components π_3, π_1. Similarly, we create:

$$W_2 \sim P_0(\lambda_2(q_4 + p_4 q_2))$$

using one binomial sample from each of the two components π_4, π_2. By generating independent observations from $Y_i \sim P_0(\lambda_i)$, $i = 1,2$, W_1 and W_2 are seen to be independent. Then by the result in Section 4.1, the conditional distribution:

$$P\{W_1 = z | W_2 - W_1 = z_2\}$$

depends on the desired parameter $\lambda(q_3 + p_3 q_1)(q_4 + p_4 q_2)$, where $\lambda = \lambda_1 \lambda_2$.

4.2.4 Other well-known systems.

System 6 is a special case of the so-called "series-parallel system." Its general form has ℓ series subsystems connected in parallel. Each series subsystem has k_i components connected in series (Fig. 4.2). Let $p_{i1}, p_{i2}, \ldots, p_{ik_i}$ be the component reliabilities of the i^{th} subsystem, $i = 1,2,\ldots, \ell$. Then

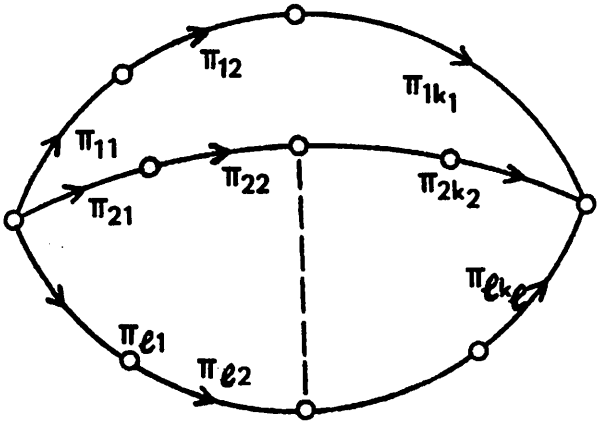


Figure 4.2. General "series-parallel system." The system reliability is given in (4.23).

$$(4.23) \qquad P_s = 1 - \prod_{i=1}^{\ell} (1 - p_{i1}p_{i2}\cdots p_{ik_i}).$$

We see $(1 - p_{i1}p_{i2}\cdots p_{ik_i})$ can be written in the form (4.17). (Compare the P_s and P_s' for system 1 in Table 4.1). Therefore, P_s of the general "series-parallel system" belongs to the class $\mathcal{B}$.

Similarly, system 7 is a special case of the "parallel-series system" whose general form is shown in Figure 4.3.

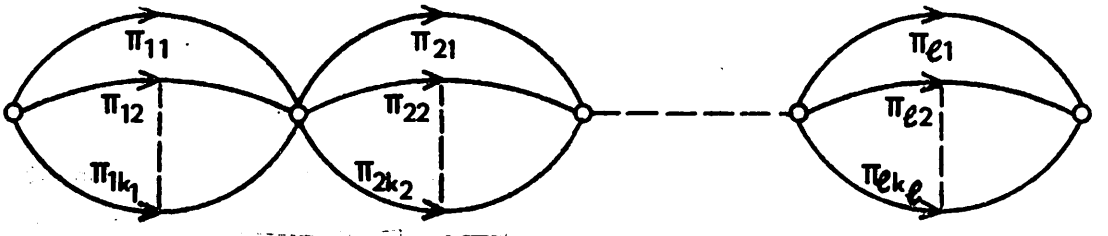


Figure 4.3. General "parallel-series system." The system reliability is given in (4.24).

Here

$$(4.24) \qquad P_s = \prod_{i=1}^{\ell} (1 - q_{i1}\ q_{i2} \cdots\ q_{ik_i}).$$

When compared to (4.23), the P_s of the general "parallel-series system" is seen to belong to $\mathcal{B}$.

Another well-known system is the "r-out-of-k system" which consists of k components and will function if any r out of k components function, where $1 < r < k$. (The cases $r = 1$ and $r = k$ correspond to a purely parallel and a purely series system, respectively, which have already been discussed and will be ruled out here.) The system reliability P_s in this case is somewhat complicated even with independent components assumption and small number of k. As mentioned in the previous section, this type of system is a special case of the monotonic system. Therefore the compounding technique can be used to estimate its P_s. It turns out that the P_s belongs to neither class $\mathcal{C}$ nor class $\mathcal{B}$. Therefore more than one binomial sample is needed from some of the components unless in the special case where the k components are identical. In this latter case:

$$(4.25) \qquad P_s = \sum_{i=r}^{k} \binom{k}{i} p^i (1-p)^{k-i}$$

where p is the probability of success of an individual component. We see (4.25) can be written in terms of the incomplete Beta function:

$$(4.26) \qquad P_s = r\binom{k}{r} \int_0^p t^{r-1}(1-t)^{k-r} dt$$

which is readily seen to be a strictly increasing function of p,

and confidence limits for P_s are obtained from confidence limits for p by simple transformation, as indicated in (4.26). When the components are not identical, then in the case $k = 3$, the only possible value of r is $r = 2$. In this 2-out-of-3 system, we have

$$P_s = p_1[p_2 + q_2(p_3)] + q_1[p_2(p_3)]$$

and we need five binomial samples in using the compounding technique (one for π_1, two for each of π_2, π_3). When $k = 4$, r can be 2 or 3. In the 3-out-of-4 system:

$$P_s = p_1\{p_2(p_3 + q_3(p_4)) + q_2(p_3)(p_4)\} + q_1\{p_2(p_3(p_4))\}$$

and we need 9 binomial samples (one for π_1, 2 for π_2, 3 for π_3, 3 for π_4).

4.3 Generalization to Multistate Failure Models.

So far we have been concerned with the attribute failure model in which both the system and components are capable of assuming one of two states of performance in a trial--success or failure, and binomial r.v.'s describe them well. In many cases, a component may assume one of three or more states of performance in each trial, and multinomial r.v.'s are needed to describe them. We know in N Bernoulli trials, if $N \sim P_0(\lambda)$, then the numbers of successes and failures are independent. There is a corresponding result for the multinomial distribution: if $E_1, E_2, \ldots, E_r$ are the r possible outcomes in a trial with $P(E_i) = p_i$, $i = 1,2,\ldots, r$, $\sum_{i=1}^{r} p_i = 1$. Let X_i = number of occurrence of the outcome E_i in N trials, $i = 1,2,\ldots, r$; we know the distribution of $(X_1, X_2, \ldots, X_r)$ is multinomial. Now suppose the trial number N is a r.v. with

$$N \sim P_0(\lambda)$$

then the absolute distribution of X_i is

$$X_i \sim P_0(\lambda p_i),\ i = 1,2,\ldots,\ r$$

and $X_1, X_2,\ldots, X_r$ are independent (Feller (1968), p. 301). We can apply the vertical and horizontal compounding technique to this case and establish confidence limits for a system whose components involve multiple states of performance. We illustrate the idea by an example. In Shooman (1968) p. 144-145, we see a semiconductor diode assumes three mutually exclusive states in a trial: it may operate properly, it may fail because of open-circuit or it may fail because of short circuit. Thus if outcome x_g represents good, outcome x_s shorted and outcome x_o opened, with $P(x_g) = p_1$, $P(x_s) = p_2$, and $P(x_o) = p_3$, then

$$P_1 + P_2 + P_3 = 1.$$

The reliability P_s of a single diode is:

$$P_s = p_1 = 1 - p_2 - p_3.$$

Now if two identical diodes are connected in series, then the system fails if either diode opens or if both short. The paths are: $x_{1g}x_{2g}$, $x_{1g}x_{2s}$, $x_{2g}x_{1s}$ and the system reliability is:

$$P_s = p_1^2 + 2p_1p_2.$$

To find confidence limits for this P_s, we perform N trials on

the given diode where $N \sim P_0(\lambda)$, $\lambda > 0$ is known. Let X_1, X_2, X_3 represent the numbers of occurrences of good, shorted and opened in the N trials. Then

$$X_1 \sim P_0(\lambda p_1)$$

$$X_2 \sim P_0(\lambda p_2)$$

$$X_3 \sim P_0(\lambda p_3)$$

and X_1, X_2, X_3 are independent. Given the observed values $X_1 = x_1$, $X_2 = x_2$, $X_3 = x_3$, if we perform x_1 trials on the diode, and let Y_1, Y_2, Y_3 represent the numbers of good, shorted and opened in $x_1 = y_1 + y_2 + y_3$ trials, then

$$Y_1 \sim P_0(\lambda p_1^2)$$

$$Y_2 \sim P_0(\lambda p_1 p_2)$$

$$Y_3 \sim P_0(\lambda p_1 p_3)$$

and Y_1, Y_2, Y_3 are independent. Now perform x_2 trials on the diode, and let Z_1, Z_2, Z_3 be the numbers of occurrence of each state. Then

$$Z_1 \sim P_0(\lambda p_2 p_1).$$

We see

$$Y_1 + Y_2 + Z_1 \sim P_0(\lambda(p_1^2 + 2p_1 p_2))$$

which can be used to construct confidence limits for P_s.

4.4 Approximations.

The vertical compounding technique discussed in the previous sections requires that at the first stage of compounding, the size of the binomial sample be determined by a known Poisson distribution. We now discuss a way to eliminate this requirement and provide approximate solutions for estimating certain functions of Bernoulli parameters.

At the beginning of Section 4.2.1, we considered the problem of estimating the function $\theta = p_1 p_2 \cdots p_k$ using the MSCP r.v. X_k, where

$$X_k \sim P_o(\lambda\theta).$$

Recall that in obtaining an observation of X_k, we start the sampling process by generating a value n_1 from $Y \sim P_o(\lambda)$. Let us now discard the r.v. Y and concentrate on the given k Bernoulli populations π_i, with parameter p_i, $i = 1,2,\ldots, k$. It might happen that one of the p_i's, say p_1, is very small. If we obtain a large binomial sample with sample size n_1 from the corresponding population π_1, and let V_1 denote the number of successes, then

$$V_1 \sim B(n_1, p_1).$$

It is well-known that V_1 can be approximated by the Poisson r.v. U_1 where

$$U_1 \sim P_o(n_1 p_1).$$

Now use the r.v. U_1 to play the role of the Poisson r.v. Y in

generating the sample size for the next binomial sample from population π_2. If $V_1 = v_1$ is the observed value of V_1, we then take $U_1 = v_1$. Let X_2 be the binomial r.v. describing the outcome of sampling from π_2 given the sample size v_1. Then $(X_2|U_1 = v_1) \sim B(v_1, p_2)$ and $X_2 \sim P_0(n_1p_1p_2)$, as in Section 4.2.1. Continuing in the same fashion as in obtaining a MSCP r.v., at the k^{th} stage we have a MSCP r.v. X_k where:

$$X_k \sim P_0(n_1p_1p_2 \cdots p_k)$$

and confidence limits for θ can be established as before.

Note that in this new approach we eliminate the requirement of generating the binomial sample size for π_1 at the first stage of sampling, and instead of exact confidence limits, we obtain approximate confidence limits. Clearly the procedure will work for estimating the following two classes of functions: (i) Any product of Bernoulli parameters as long as there is one (or more) parameter whose value is very small. For example, the reliability of a parallel system $\theta = 1 - q_1q_2 \cdots q_k$, where the failure probability q_i, $i = 1,2,\ldots, k$ are usually very small. (ii) Any sum of products of Bernoulli parameters in which each product term has one or more small parameters. For example, the P_s' of the system 4 and 5 in Table 4.1.

It happens that the procedure also works for a series system. In this case $P_s = p_1p_2 \cdots p_k$. We shall estimate $P_s' = 1 - P_s$. Now for $k = 2, 3$, P_s' can be expressed as follows:

$$k = 2, \quad P_s' = 1 - p_1 p_2 = q_1 + q_2 p_1$$

$$k = 3, \quad P_s' = 1 - p_1 p_2 p_3 = q_2 + q_3 p_2 + p_3 p_2 q_1 \ .$$

Both of these alternative forms belong to class (ii) above. Therefore, we can use the new procedure to estimate $1 - p_1 p_2$ and $1 - p_1 p_2 p_3$. The general k components case follows easily.

CHAPTER V

SOME PROPERTIES OF INDEPENDENT GEOMETRIC RANDOM VARIABLES AND THEIR APPLICATION TO RELIABILITY PROBLEMS

Lieberman and Ross (1970) have described a technique for estimating reliability when time to failure has an exponential distribution. In this section we show that their technique applies also in our Bernoulli model.

Recall that in Section 3.3.2, we discussed a series system consisting of two components π_1, π_2. If we do IBS on π_i until first failure, i = 1,2, and let X_1, X_2 represent the outcome, then:

$$Y \equiv \text{Min}(X_1, X_2) \sim GD(1-p_1p_2).$$

We have stated that Y utilizes only the smaller of the two observations X_1, X_2. Thus it loses "information," and is not the best r.v. to use in estimating the system reliability $P_s \equiv p_1p_2$. By examining the matter further, it is seen the information lost is $\max(X_1, X_2)$. We now propose a way to estimating p_1p_2 by using Y and the "leftover information." The procedure can be extended to cover systems with k components where the system reliability is $p_1p_2 \cdots p_k$, k being a positive integer. First we need to establish some facts about the geometric r.v.'s.

5.1 Some Properties of Independent Geometric r.v.'s.

Ferguson (1965) has proved that if $X_i \sim GD(1-p_i)$, i = 1,2, are independent, then the r.v.'s $Y \equiv \text{Min}(X_1, X_2)$ and $V = X_2 - X_1$ are independent. This fact will be used later. We now prove results of related interest.

Theorem 5.1.

Let $X_i \sim GD(1-p_i)$, $i = 1,2$, be two independent geometric r.v.'s. Then:

$$(5.1) \qquad P\{|X_2 - X_1| = x | X_2 \geq X_1\} = (1-p_2)p_2^x, \quad x = 0,1,2,\ldots$$

and

$$(5.2) \qquad P\{|X_2 - X_1| = x | X_1 \geq X_2\} = (1-p_1)p_1^x, \quad x = 0,1,2,\ldots.$$

(5.1) means when $X_2 \geq X_1$, the distribution of the difference $|X_2 - X_1|$ is geometric with the same parameter as that of the larger of X_2, X_1. The interesting fact is that p_1, which is the parameter of the smaller of X_2, X_1, is not involved in the distribution. Similar comments apply to (5.2).

Proof:

For $x = 0,1,2,\ldots$

$$P\{|X_2 - X_1| = x | X_2 \geq X_1\} = \frac{P\{|X_2 - X_1| = x,\ X_2 - X_1 \geq 0\}}{P\{X_2 - X_1 \geq 0\}}$$

$$= \frac{P\{X_2 - X_1 = x\}}{\sum_{y=0}^{\infty} P\{X_2 - X_1 = y\}} .$$

It follows from (3.18) that for $w = 0,1,2,\ldots$

$$P\{X_2 - X_1 = w\} = q_1 q_2 p_2^w/(1-p_1p_2).$$

Therefore (5.1) is true. The proof of (5.2) is similar.

The events $\{X_1 \leq X_2\} = \bigcup_{i=0}^{\infty} \{X_2 - X_1 = i\}$ and the event $\{\text{Min}(X_1, X_2) = w\}$ are independent. This follows easily from Ferguson's result mentioned above. Therefore

$$(5.3) \qquad P\{\text{Min}(X_1, X_2) = w | X_1 \leq X_2\} = (1-p_1p_2)(p_1p_2)^w, \quad w = 0,1,2,\ldots.$$

We now can easily show:

Theorem 5.2.

Let $X_i \sim GD(1-p_i)$, $i = 1,2$, be independent. If $X_1 \leq X_2$, then $\text{Min}(X_1, X_2)$ and $|X_2 - X_1|$ are independent.

Theorems 5.1 and 5.2 will be useful for estimating p_1p_2, as we will show later. To estimate $p_1p_2 \ldots p_k$, $k > 2$, we need more general results as presented in Theorems 5.3 and 5.4.

Theorem 5.3.

Let $X_i \sim GD(1-p_i)$, $i = 1,2,\ldots, k$, be independent geometric r.v.'s. Define $W_i \equiv |X_{i+1} - X_i|$, $i = 1,2,\ldots, k-1$. Then

$$(5.4) \qquad P\{W_i = w_i | X_1 \leq X_2 \leq X_3 \leq \ldots \leq X_k\} = (1 - \prod_{j=i+1}^{k} p_j)(\prod_{j=i+1}^{k} p_j)^{w_i},$$

$$w_i = 0,1,2,\ldots$$

$$i = 1,2,\ldots, k-1.$$

Again the conditioned r.v. is geometric. Note the disappearance of the parameters $p_1, p_2,\ldots, p_i$ of the smallest i r.v.'s $X_1, X_2,\ldots, X_i$ in the condition $X_1 \leq X_2 \leq \ldots \leq X_k$. Example: if $k = 4$, then when $X_1 \leq X_2 \leq X_3 \leq X_4$, we have:

$$W_1 \equiv |X_2 - X_1| \sim GD(1-p_2p_3p_4),$$
$$W_2 \equiv |X_3 - X_2| \sim GD(1-p_3p_4),$$
$$W_3 \equiv |X_4 - X_3| \sim GD(1-p_4).$$

The proof starts with the observation that

$$P\{|X_{i+1} - X_i| = w_i | X_1 \le X_2 \le \dots \le X_k\}$$

$$= \frac{P\{X_{j+1}-X_j \ge 0, j=1,2,\dots,i-1,i+1,\dots,k-1, X_{i+1}-X_i \ge 0, |X_{i+1}-X_i| = w_i\}}{P\{X_{j+1}-X_j \ge 0, j=1,2,\dots,k-1\}}$$

$$= \frac{P\{X_{j+1}-X_j \ge 0, j=1,2,\dots,i-1,k+1,\dots,k-1, X_{i+1}-X_i = w_i\}}{P\{X_{j+1}-X_j \ge 0, j=1,2,\dots,k-1\}} .$$

The denominator can be written as:

$$(5.5) \qquad \sum_{\ell_{k-1}=0}^{\infty} \sum_{\ell_{k-2}=0}^{\infty} \dots \sum_{\ell_1=0}^{\infty} P\{X_{j+1} - X_j = \ell_j,\ j = 1,2,\dots,\ k-1\}$$

$$= \sum_{\ell_{k-1}=0}^{\infty} \sum_{\ell_{k-2}=0}^{\infty} \dots \sum_{\ell_1=0}^{\infty} \sum_{m=0}^{\infty} P\{X_1 = m,\ X_{j+1} = m + \sum_{i=1}^{j} \ell_i,\ j=1,2,3,\dots,\ k-1\}$$

$$= \left(\prod_{i=1}^{k} q_i\right) \prod_{i=1}^{k} \left[1 - \prod_{j=i}^{k} p_j\right]^{-1}.$$

Applying similar arguments to the numerator, we obtain the result (5.4).

Having found the distributions of W_i, $i = 1,2,\dots,\ k-1$ we shall next show that $W_0, W_1, W_2, \dots, W_k$ are jointly independent, where $W_0 = \text{Min}(X_1, X_2, \dots, X_k)$. First we find that

$$(5.6) \qquad W_0 \sim GD(1-p_1p_2\cdots p_k).$$

This is easily seen from the fact:

$$Y_1 \equiv \text{Min}(X_1, X_2) \sim GD(1-p_1p_2)$$

$$Y_2 \equiv \text{Min}(X_1, X_2, X_3) = \text{Min}(Y_1, X_3) \sim GD(1-p_1p_2p_3)$$

$$Y_3 \equiv \text{Min}(X_1, X_3, X_4) = \text{Min}(Y_2, X_4) \sim GD(1-p_1p_2p_3p_4).$$

Then obviously:

$$Y_{k-1} \equiv W_0 = \text{Min}(X_1, X_2, \ldots, X_k) \sim GD(1-p_1p_2\cdots p_k).$$

Corresponding to (5.3), it is not difficult to show that the events $\{W_0 = w_0\}$ and $\{X_1 \le X_2 \le \ldots \le X_k\}$ are independent. Therefore

$$P\{W_0 = w_0 | X_1 \le X_2 \le \ldots \le X_k\} = (1-p_1p_2\cdots p_k)(p_1p_2\cdots p_k)^{w_0}, \quad w_0 = 0,1,2,\ldots. \tag{5.7}$$

We can now prove a generalized version of Theorem 5.2.

Theorem 5.4.

Let $X_i \sim GD(1-p_i)$, $i = 1,2,\ldots, k$, be independent. Let $W_i = |X_{i+1} - X_i|$, $i=1,2,\ldots, k-1$, and let $W_0 = \text{Min}(X_1, X_2, \ldots, X_k)$. Given that $X_1 \le X_2 \le \ldots \le X_k$, the r.v.'s $W_0, W_1, \ldots, W_{k-1}$ are independent.

Proof:

We have to show that for each vector $(w_0, w_1, \ldots, w_{k-1})$

$$P\{W_0 = w_0, W_1 = w_1, \ldots, W_{k-1} = w_{k-1} | X_1 \le X_2 \le \ldots \le X_k\} = \prod_{i=0}^{k-1} P\{W_i = w_i | X_1 \le X_2 \le \ldots \le X_k\}.$$

The right-hand side is easily found from (5.7) and Theorem (5.4) to be:

$$\prod_{i=0}^{k-1} (1 - \prod_{j=i+1}^{k} p_j)(\prod_{j=i+1}^{k} p_j)^{w_i}. \tag{5.8}$$

The left-hand side can be written as:

$$(5.9)\qquad P\{X_1 = w_0,\ X_2 - X_1 = w_1, \ldots,\ X_k - X_{k-1} = w_{k-1}\} \ /\ P\{X_1 \le X_2 \le \ldots \le X_k\}$$

$$= P\{X_1 = w_0,\ X_2 = w_0 + w_1, \ldots, X_k = \sum_{i=0}^{k-1} w_i\} \ /\ P\{X_1 \le X_2 \le \ldots \le X_k\},$$

using (5.5) for the denominator here it is easily seen (5.9) is equal to (5.8).

Let us define

$$V \equiv \text{Max}(X_i,\ i = 1,2,\ldots,\ k)$$

and

$$R \equiv V - W_0 = \text{Max}(X_i,\ i = 1,2,\ldots,\ k) - \text{Min}(X_i,\ i = 1,2,\ldots,\ k).$$

We note that if the X_i's have identical distributions. Then R becomes the sample range from a geometric distribution. It is clear from Theorem 5.4 that W_0 and R are independent when $X_1 \le X_2 \le \ldots \le X_k$, because $R = \sum_{i=1}^{k-1} W_i$.

By repeatedly applying Theorems 5.3 and 5.4, it is interesting to visualize a hierarchy of independent geometric r.v.'s. For example if $k = 4$ then when $X_1 \le X_2 \le X_3 \le X_4$, we have Figure 5.1.

$$\underbrace{X_1 \le X_2 \le X_3 \le X_4}$$

$$W_0 \qquad \underbrace{W_{(1)} \le W_{(2)} \le W_{(3)}}$$

$$U_0 \qquad \underbrace{U_{(1)} \le U_{(2)}}$$

$$V_0 \qquad V_1$$

Figure 5.1. Hierarchy of independent geometric r.v.'s.

(1) In the first row are the four original geometric r.v.'s which have been ordered.

(2) Given that $X_1 \leq X_2 \leq X_3 \leq X_4$, we create (by Theorems 5.3 and 5.4) four independent geometric r.v.'s. W_0, W_1, W_2, W_3. $W_{(1)}$, $W_{(2)}$, $W_{(3)}$ in the second row are the ordered values of W_1, W_2, W_3.

(3) Given that $W_{(1)} \leq W_{(2)} \leq W_{(3)}$, we create (again by Theorems 5.3 and 5.4) three independent geometric r.v.'s U_i, $i = 0,1,2$ where

$$U_0 = \text{Min}(W_i,\ i = 1,2,3)$$

$$U_i = |W_{(i+1)} - W_{(i)}|, i = 1,2$$

and $U_{(1)}$, $U_{(2)}$ in the third row are the ordered values of U_1, U_2.

(4) Similarly given that $U_{(1)} \leq U_{(2)}$, we create two independent geometric r.v.'s V_i, $i = 0,1$ where

$$V_0 = \text{Min}(U_i,\ i = 1,2)$$

$$V_1 = |U_{(2)} - U_{(1)}| \ .$$

With $k = 4$, there are $4 + 3 + 2 = 9$ geometric r.v.'s created in three "stages" as shown in Figure 5.1; the r.v.'s within each stage are independent. In general, for any k original independent geometric r.v.'s, we can generate $k + (k-1) + \ldots + 2 = \frac{k(k+1)}{2} - 1$ new geometric r.v.'s in $(k-1)$ stages.

5.2 Application to the Estimation of System Reliability. Two Components Case.

We now apply the results obtained in the previous section to the estimation of system reliability P_s. First let $P_s = p_1 p_2$, the reliability of a series system with two components π_1, π_2.

The sampling procedure is the inverse binomial sampling used in Chapters 2 and 3. For component π_i, $i = 1,2$ we observe X_i: the number of successes prior to the first failure for the i^{th} component,

$$X_i \sim GD(1-p_i).$$

Let X_{1i}, $i = 1,2,\ldots, p_1$ be a set of n_1 observations for X_1. (For convenience, we use the same symbol for a r.v. and its observed value.) Similarly, let X_{2j}, $j = 1,2,\ldots, n_2$ be a set of n_2 observations for X_2. To do the analysis, define U and V as:

$$U = \text{Min}(X_1, X_2)$$

$$V \equiv \text{Max}(X_1, X_2) - \text{Min}(X_1, X_2) = |X_2 - X_1|.$$

By Theorems 5.1 and 5.2

Case A. If $X_1 \geq X_2$, then U and V are independent, and

$$U \sim GD(1-p_1p_2)$$

$$V \sim GD(1-p_1).$$

Case B. If $X_1 \leq X_2$, then U and V are independent, and

$$U \sim GD(1-p_1p_2)$$

$$V \sim GD(1-p_2).$$

Thus, in Case A, U can be used as an observation to estimate p_1p_2, while the "leftover" information V can be combined with another observation for X_2, say X_2', to form $U' \equiv \text{Min}(V, X_2')$, which is $GD(1-p_1p_2)$, and U' becomes another observation for estimating p_1p_2.

Similarly, in Case B, U is an observation for estimating p_1p_2, while the "leftover information" V can be combined with another observation for X_1, say X_1', to form $U' \equiv \text{Min}(V, X_1')$ which is $GD(1-p_1p_2)$ and U' becomes another observation for estimating p_1p_2. We see that in any case, in order to utilize the observed value of V, we pick the next observation from the component whose observed value is $\text{Min}(X_1, X_2)$, corresponding to the fact that $\text{Min}(X_1, X_2)$ has been used to estimate p_1p_2. It is best to illustrate the procedure by an example.

Suppose we have five observations for X_1: $X_{11} = 2$, $X_{12} = 3$, $X_{13} = 6$, $X_{14} = 5$, $X_{15} = 1$ and four observations for X_2: $X_{21} = 8$, $X_{22} = 1$, $X_{23} = 0$, $X_{24} = 4$. Then $U_1 = \text{Min}(2, 8) = 2 \sim GD(1-p_1p_2)$, $V_1 = |2-8| = 6 \sim GD(1-p_2)$. Since $2 \leq 8$, we have Case B. Therefore we should combine $V_1 = 6$ with the next observation for X_1 which is 3. Treating V_1 as X_2, we have: $U_2 = \text{Min}(3, 6) = 3 \sim GD(1-p_1p_2)$, $V_2 = |3-6| = 3 \sim GD(1-p_2)$. Here $3 \leq 6$, and this is Case B; we should combine $V_2 = 3$ with the next observation for X_1, which is 6. Treating V_2 as X_2, we have: $U_3 = \text{Min}(6, 3) = 3 \sim GD(1-p_1p_2)$, $V_3 = |6-3| = 3 \sim GD(1-p_1)$. Here $6 \geq 3$, and this is Case A. Therefore we combine $V_3 = 3$ with the next observation for X_2, which is 1. Treating V_3 as X_1, we have: $U_4 = \text{Min}(3, 1) = 1 \sim GD(1-p_1p_2)$, $V_4 = |3-1| = 2 \sim GD(1-p_1)$. Here $3 \geq 1$, using Case A, we combine $V_4 = 2$ with the next observation from X_2, which is 2. Treating V_4 as X_1, we have: $U_5 = \text{Min}(2, 2) = 2 \sim GD(1-p_1p_2)$, $V_5 = |2-2| = 0 \sim GD(1-p_1)$. Note $\text{Max}(X_1, X_2) = \text{Min}(X_1, X_2) = 2$, so that V_5

can be $GD(1-p_1)$ or $GD(1-p_2)$, (either one can be used). We use $V_5 \sim GD(1-p_1)$, i.e., we are treating this case as Case A. Thus we should combine $V_5 = 0$ with the next observation for X_2, which is 4.

Continue in the same way. We obtain:

$$U_6 = 0,\ V_6 = 4,\ U_7 = 4,\ V_7 = 1.$$

When V_7 is obtained, the next observation should be from X_2, but there is not any observation from X_2 available. Therefore the process is terminated there.

In looking back, we have "manufactured" 7 observations for $U \sim GD(1-p_1p_2)$, i.e., 2, 3, 3, 1, 2, 0, 4, out of the $5 + 4 = 9$ original observations for X_1 and X_2; we see that

$$Y \equiv 2 + 3 + 3 + 1 + 2 + 0 + 4 = 15 \tag{5.10}$$

is an observation from $NB(7, 1-p_1p_2)$, and confidence limits for p_1p_2 can be easily constructed. The simplicity of the calculation involved appears to be an advantage of this procedure. Note there is one observation (the value 1 from X_1) whose information is not used. In general, if we have n_1 and n_2 observations for X_1 and X_2 respectively, we can manufacture n observations for U, where $\text{Min}(n_1, n_2) - 1 \leq n \leq n_1 + n_2 - 1$. The difference $(n_1 + n_2 - 1) - n$ is the number of observations wasted. Note also that in the example given, we worked with the original observations in the order: 2, 3, 6, 5, 1 from X_1 and 8, 1, 0, 4 from X_2. If we alter the order, say, to 1, 2, 3, 6, 5 from X_1, and let 8, 1, 0, 4 from X_2 remain the

same, we see U_1 becomes $\text{Min}(1, 8) = 1$ instead of $\text{Min}(2, 8) = 2$ as given in the example. Thus the same set of original observations may lead to different values for Y in (5.10), and also to different confidence limits. The simple alternative sampling procedure described below provides one way to avoid these two problems. The analysis of data remains the same; only the way observations are taken is different.

Instead of fixing sample sizes n_1 and n_2 separately as we did before, let the combined sample size $N(= n_1 + n_2)$ be fixed. Then take one observation each of X_1 and X_2, say X_{11}, X_{21}, and do exactly the same analysis as we did in obtaining U_1 and V_1 before. If $X_{11} \geq X_{21}$, take one observation X_{12} of X_1; if $X_{11} \leq X_{21}$, take one observation X_{22} of X_2, and do the same analysis as before in utilizing the leftover information. Continue in this fashion until we obtain N observations of both X_1 and X_2 (the sample size n_i for X_i, $i = 1,2$, will not be known until N is reached). In this way we shall always obtain $N - 1$ observations for U, out of the N original observations. The difference of 1 represents the left over information V_{N-1} from the last stage analysis.

It is clear that if we observe Y_i, the number of failures prior to the first success, then

$$Y_i \sim GD(1-q_i), \; i = 1,2,$$

and the same procedure can be used to construct confidence limits for the reliability of a parallel system $P_s = 1 - q_1 q_2$.

5.3 Application to the Estimation of System Reliability. Three or More Components.

To generalize the procedure to a series or parallel system with $k \geq 3$ components, we apply Theorems 5.3 and 5.4 in addition to Theorems 5.1 and 5.2. We shall illustrate the $k = 3$ components case. Here $P_s = p_1p_2p_3$. For component π_i, $i = 1,2,3$, as before, we observe X_i: the number of successes prior to the first failure for the i^{th} component, $X_i \sim GD(1-p_i)$. Let X_{ij}, $j = 1,2,\ldots, n_i$ be a sample of n_i observations from X_i, $i = 1,2,3$. To analyse the data, let:

$$Y_{(i)} = \text{the } i^{th} \text{ largest of } X_1, X_2, X_3.$$

$$p_{(i)} = 1 - q_{(i)} = \text{the parameter of } X_{(i)}.$$

We have $X_{(1)} \leq X_{(2)} \leq X_{(3)}$. Define

$$U = X_{(1)}$$

$$V = X_{(2)} - X_{(1)}$$

$$W = X_{(3)} - X_{(2)}.$$

By Theorems 5.3 and 5.4, conditioned on the observed ordering, U, V, W are independent and:

$$U \sim GD(1-p_1p_2p_3)$$

$$V \sim GD(1-p_{(2)}p_{(3)})$$

$$W \sim GD(1-p_{(3)}).$$

U can be used as an observation to estimate $p_1p_2p_3$, while V and

W are the "leftover information." It is clear if we combine V with a new observation from $X_{(1)}$, say $X'_{(1)}$ to form $U' \equiv \text{Min}(V, X'_{(1)}) \sim GD(1-p_1p_2p_3)$, then U' becomes another observation for estimating $p_1p_2p_3$. By Theorem 5.1, the leftover information $V' = |V - X'_{(1)}|$ is

Case A. $V' \sim GD(1-p_{(2)}p_{(3)})$ if $V \geq X'_{(1)}$, or

Case B. $V' \sim GD(1-p_{(1)})$ if $V \leq X'_{(1)}$.

In Case A, we should take another observation $X''_{(1)}$ of $X_{(1)}$ so that $\text{Min}(V', X''_{(1)}) \sim GD(1-p_1p_2p_3)$. In Case B, we combine W, which has not been used so far, and V', so that $Y_1 \equiv \text{Min}(W, V') \sim GD(1-p_{(1)}p_{(3)})$ and go on to pick one observation $X'_{(2)}$ from $X_{(2)}$, so that $\text{Min}(Y_1, X'_{(2)}) \sim GD(1-p_1p_2p_3)$. A numerical example follows: let $n_1 = 4$, $n_2 = 3 = n_3$, and

For component π_1, $X_{11} = 1$, $X_{12} = 7$, $X_{13} = 1$, $X_{14} = 5$

For component π_2, $X_{21} = 2$, $X_{22} = 8$, $X_{23} = 4$

For component π_3, $X_{31} = 6$, $X_{32} = 9$, $X_{33} = 3$.

Table 5.2 demonstrates the procedure. The notation $Z \sim p_i$ means the quantity Z is from $GD(1-p_i)$. There are four columns in the table, the symbol * placed between the column "Observations" and the column "Leftover A," means the corresponding r.v.'s in these two columns are independent. A similar remark applied to the symbol * between corresponding quantities in "Leftover B" and "Leftover C." Under columns "Leftover A" and "Leftover B," we find the instruction "next observation from X_i," indicating which component should provide

the next observation. This instruction is dictated by the component parameter p_i which is lacking in the distribution of the quantity immediately above the instruction. For example, in the stage where we observe $X_{22} = 8 \sim p_2$, the quantity placed immediately above the instruction "next observation from X_3" is $2 = \text{Min}(4, 2) \sim p_1p_2$. Here p_1p_2 lacks p_3, so the instruction says that the next observation should be from component π_3. And, in the next stage we combine a new observation $X_{32} = 9$ from component π_3 with the leftover information in the previous stage: $2 = \text{Min}(4, 2) \sim p_1p_2$. The utilization of previous leftover information in Table 5.2 is indicated by the arrows.

Note the procedure terminated at the stage which created U_7, because at that stage, the instruction says we should take an observation from X_2, but we have used up all the available observations from X_2. The leftover information in every stage except the last one is utilized and we "manufactured" 7 observations for U where

$$U \sim GD(1-p_1p_2p_3)$$

out of $4 + 3 + 3 = 10$ original observations. The sum of these 7 values $(1 + 1 + 4 + 2 + 1 + 1 + 4 = 14)$ is considered to be from $NB(7, 1-p_1p_2p_3)$, and confidence limits for $p_1p_2p_3$ can be constructed. Note, upon termination, there is one observation, namely $X_{33} = 3$, not used. This problem of possible waste of observations can be handled using a sampling procedure analogous to the alternative sampling procedure discussed in $k = 2$ case. With obvious modification, the procedure is applicable to a parallel system.

Table 5.1. Illustration of the Procedure for $k \equiv 3$.

Observations	Leftover A	Leftover B	Leftover C
Observe $X_{11}=1$, $X_{21}=2$, $X_{31}=6$			
$U_1=1=\text{Min}(1,2,6)\sim p_1p_2p_3$ *		$1=\lvert 2-1\rvert\sim p_2p_3$ Next observation from X_1	* $4=\lvert 6-2\rvert\sim p_3$
Observe $X_{12}=7\sim p_1$			
$U_2=1=\text{Min}(7,1)\sim p_1p_2p_3$ *	$6=\lvert 1-7\rvert\sim p_1$,	$4=\text{Min}(6,4)\sim p_1p_3$ Next observation from X_2	* $2=\lvert 4-6\rvert\sim p_1$
Observe $X_{22}=8\sim p_2$			
$U_3=4=\text{Min}(8,4)\sim p_1p_2p_3$ *	$4=\lvert 8-4\rvert\sim p_2$,	$2=\text{Min}(4,2)\sim p_1p_2$ Next observation from X_3	* $2=\lvert 4-2\rvert\sim p_2$
Observe $X_{32}=9\sim p_3$			
$U_4=2=\text{Min}(9,2)\sim p_1p_2p_3$ *	$7=\lvert 9-2\rvert\sim p_3$,	$2=\text{Min}(7,2)\sim p_2p_3$ Next observation from X_1	* $5=\lvert 7-2\rvert\sim p_3$
Observe $X_{13}=1\sim p_1$			
$U_5=1=\text{Min}(1,2)\sim p_1p_2p_3$ *	$1=\lvert 1-2\rvert\sim p_2p_3$,		$5=\lvert 7-2\rvert\sim p_3$
Observe $X_{14}=5\sim p_1$			
$U_6=1=\text{Min}(5,1)\sim p_1p_2p_3$ *	$4=\lvert 5-1\rvert\sim p_1$,	$4=\text{Min}(4,5)\sim p_1p_3$ Next observation from X_2	* $1=\lvert 4-5\rvert\sim p_3$
Observe $X_{23}=4\sim p_2$			
$U_7=4=\text{Min}(4,4)\sim p_1p_2p_3$ *	$0=\lvert 4-4\rvert\sim p_1$,	$0=\text{Min}(0,1)\sim p_1p_3$ Next observation from X_2	* $1=\lvert 0-1\rvert\sim p_3$

CHAPTER VI

OTHER RELIABILITY APPLICATIONS AND APPLICATION TO BIOMEDICAL PROBLEMS

So far our main concern is about the attribute failure model and its generalization--the multistate failure model. In this chapter we shall briefly discuss the "variable failure model" using results of previous chapters. After that we shall mention some application of our previous results to biomedical problems.

6.1 A Non-Parametric Method for the "Variable Failure Model."

The so-called "variable failure model" involves a time factor and describes the behavior of "time to failure" using various continuous distributions. Examples are the exponential, the gamma, the Weibull, and the extreme value distributions. In these cases, one may speak of reliability of a component as probability of successful operation of that component at least to a given fixed time t_0, sometimes called the "mission time." If $F_T(t)$ is the CDF of the r.v. T which stands for "time to failure" of the component involved, then $R \equiv 1 - F_T(t_0) = P\{T > t_0\}$ is the reliability of the component.

Now suppose a system is made up of $k \geq 2$ components, each one of them having a time to failure distribution $F_{T_i}(t_i)$, with mission time a_i, $i = 1,2,\ldots, k$. Then the component reliability is $1 - F_{T_i}(a_i)$ for component $i = 1,2,\ldots, k$. As in the case of the attribute failure model, one practical problem is to estimate system reliability when it can be expressed as a function of the

component reliabilities $1-F_{T_i}(a_i)$, $i = 1,2,\ldots, k$. The solution is not easy. However a non-parametric type solution is readily available. For each r.v. T_i, define a Bernoulli r.v. X_i as:

$$X_i = \begin{cases} 1 & \text{if } T_i \leq a_i \\ 0 & \text{if } T_i > a_i \end{cases}.$$

Thus we obtain k Bernoulli r.v.'s X_i, $i = 1,2,\ldots, k$ with parameter $p_i = 1 - F_{T_i}(a_i)$, and techniques derived in the previous chapters for the attribute failure model are applicable in constructing confidence limits for the system reliability. The next section discusses a more efficient method to estimate the reliability of a series system consisting of k independent components having exponential time to failure distributions.

6.2 Confidence Limits for an Exponential-Failure-Time Series System and for the Difference of Poisson Parameters.

In the special case in which a system consists of $k \geq 2$ components in series, and the time to failure for each component follows the exponential law with parameter λ_i, $i = 1,2,\ldots, k$, confidence limits for the system reliability R have been found using failure times data of individual component (see Lentner and Buehler (1963), El Mawaziny (1965)). The following alternative sampling scheme provides a simple solution for this problem involving much less computation.

For component 1, we test one item from it until it fails, then test another item from it until it fails (this can be called replacement testing). Continue in this fashion until a predetermined total

testing time t_0 is reached. Then stop testing component 1. Let T_1 be the time elapsed until the first failure, and for $j \geq 2$, let T_j be the time elapsed from the $(j-1)^{th}$ failure to the j^{th} failure. It is clear that $\{T_j\}$ is a sequence of i.i.d. r.v.'s with the common exponential distribution

$$T \sim \text{Exp}(\lambda_1).$$

The T_j's are called inter-arrival times, and it has been shown (see, for example, Parzen (1962), p. 174) that if the inter-arrival times $\{T_j\}$ are exponentially distributed with parameter λ_1, then the renewal counting process $\{N_1(t), t \geq 0\}$ is a Poisson process with intensity λ_1. In other words, for any time $t > 0$, the number of failures $N_1(t)$ occur during the period from 0 to t is a Poisson r.v. with

$$N_1(t) \sim P_0(\lambda_1 t).$$

In particular,

$$N_1(t_0) \sim P_0(\lambda_1 t_0)$$

where $t = t_0$ is the predetermined total testing time.

Similarly, for each i, $i = 2,3,\ldots, k$, if we use the same testing procedure until the same total testing time t_0 is reached, then

$$N_i(t_0) \sim P_0(\lambda_i t_0), \; i = 2,3,\ldots, k$$

where $N_i(t_0)$ = number of failures occur for component i during the period from 0 to t_0. Let

$$N(t_0) = N_1(t_0) + N_2(t_0) + \ldots + N_k(t_0).$$

Clearly

$$N(t_0) \sim P_0((\lambda_1 + \lambda_2 + \ldots + \lambda_k)t_0).$$

Since t_0 is preassigned, construction of confidence limits for $\theta \equiv \lambda_1 + \lambda_2 + \ldots + \lambda_k$, (and hence for system reliability $R = e^{-\theta t_0}$) reduces to the elementary problem of finding confidence limits for the parameter for a Poisson distribution.

The same technique can be applied when the inter-arrival time Y is distributed in such a way that a known function $X = g(Y)$ has the exponential distribution. For example, let $X = g(Y) = Y^{\beta}$ so that time to failure Y has the Weibull CDF

$$F(y; \beta, \theta) = 1 - \exp(-y^{\beta}/\theta),$$

where β is known, and θ is unknown. Since $X = Y^{\beta}$ is exponentially distributed, the Poisson technique can be used.

We have appealed to the Poisson process for estimating the sum of the parameters of the exponential distributions. It is interesting to mention here that one can appeal to the exponential distribution for estimating the difference of parameters of two Poisson processes. (Birnbaum (1954) treated this estimation problem using Chebyshev's inequality, but the solution does not seem to be satisfactory.) Let $\{X(t), t > 0\}$, $\{Y(t), t > 0\}$ be two independent Poisson processes with parameters or intensities λ_1, λ_2 respectively. Suppose we are interested in estimating the difference of the two intensities $\lambda_1 - \lambda_2$. It is well-known that the inter-arrival times of a Poisson

process are independent, identical exponential distributions with the same parameter as the Poisson process. It follows that the sum of the n inter-arrival times $X_1, X_2, \ldots, X_n$ from the process $\{X(t), t > 0\}$ has a gamma distribution with parameters n and λ_1, n being a positive integer. Thus

$$X \equiv X_1 + X_2 + \ldots + X_n \sim \text{Gamma } (n, \lambda_1).$$

Similarly, if $Y_1, Y_2, \ldots$ are inter-arrival times from the process $\{Y(t), t > 0\}$, then

$$Y \equiv Y_1 + Y_2 + \ldots + Y_m \sim \text{Gamma } (m, \lambda_2).$$

Since it has been shown that the conditional density of the r.v. X given the sum $X + Y$ depends on $\lambda_1 - \lambda_2$ only, confidence limits for $\lambda_1 - \lambda_2$ can be constructed. (See Lentner and Buehler (1963).)

The rest of the chapter will discuss applications to some biomedical problems.

6.3 Confidence Limits for the Difference of Two Bacterial Densities.

$p_1 - p_2$.

In the dilution method in bacteriology, a bacterial culture is diluted in a certain volume of water, from which a number of samples of fixed size v are taken and tested for the presence or absence of bacteria. The result of each sample is a Bernoulli variable with probability of success (at least one occurrence of bacteria)

$$p = 1 - e^{-\lambda v}$$

and probability of failure $q = e^{-\lambda v}$ where λ is the density of bacteria in the volume v. Procedures for estimating λ have been

discussed by various authors (see Cochran (1950)). If one is interested in using the dilution method to find the difference of two bacterial densities, i.e., $\lambda_1 - \lambda_2$, the following method is applicable. We continue taking samples of fixed size v from each of two kinds of diluted water with the same volume until r_1^{th} and r_2^{th} successes are observed. For the first kind of bacteria, let

X = number of failures encountered prior to the occurrence of the r_1^{th} success.

Similarly, for the second kind of bacteria, let

Y = number of failures encountered prior to the occurrence of the r_2^{th} success.

Then

$$X \sim NB(r_1,\ 1-e^{-\lambda_1 v}),$$

$$Y \sim NB(r_2,\ 1-e^{-\lambda_2 v}).$$

Therefore, by taking $k_1 = k_2 = 1$, $W_1 = X$, $V_1 = X + Y$, $p_1 = e^{-\lambda_1 v}$ and $p_1' = e^{-\lambda_2 v}$ in (2.6), we have for $x = 0,1,2,\ldots, k$, $k = 0,1,2,\ldots,$

$$P\{X = x \mid X + Y = k\} = \frac{\binom{r_1+x-1}{x}\binom{r_2+k-x-1}{k-x}e^{(\lambda_1-\lambda_2)vx}}{\sum_{j=0}^{k}\binom{r_1+j-1}{j}\binom{r_2+k-j-1}{k-j}e^{(\lambda_1-\lambda_2)vj}}$$

which depends on the desired difference of two bacterial densities $\lambda_1 - \lambda_2$ only. Thus a confidence interval for $(\lambda_1 - \lambda_2)$ can be found by the usual method.

The success of the above method hinges on the fact that the probability of success p can be expressed in terms of the exponential

function $1 - e^{-\lambda v}$. In general, if one is interested in constructing confidence intervals for the difference of two Bernoulli parameters $p_1 - p_2$ or equivalently $q_2 - q_1$, and if both p_1 and p_2 are small (or both q_1 and q_2 are small) so that it is reasonable to assume that $q_i = 1 - p_i = e^{-p_i}$, $p_i = 1 - e^{-p_i}$, $i = 1,2$, then using the IBS plan to take observations from the two Bernoulli populations, one can use the above described method to find an approximate confidence interval for e^{-p_2}/e^{-p_1} (or e^{-q_2}/e^{-q_1}), and hence for $p_1 - p_2$ (or $q_2 - q_1$), as desired. If p_1, p_2 are neither close to 0 nor close to unity, then the usual normal approximation is applicable.

6.5 Comparing Two Yule's Birth Processes.

Suppose we have a population of objects which can generate (or give "birth" to) new objects, and that objects do not disappear (or "die") from the population. Let $P_X(x; m, t)$ be the probability distribution of the r.v. X which represents the number of objects in the population at time t. Suppose $x = m$ when $t = 0$. If $\lambda > 0$ is the growth rate, then

$$P_X(x; m, t) = \binom{x-1}{m-1} e^{-m\lambda t}(1-e^{-\lambda t})^{x-m}, \quad x = m, m+1, \ldots.$$

This is called Yule's birth process.

If there are two populations whose growth patterns can be described by Yule's process, one interesting problem is to compare the growth rates λ_1, λ_2 of the two populations. For this purpose, we change the form of $P_X(x; m, t)$ by a simple translation $Y = X - m$. Then

$$P_Y(y; m, t) = \binom{y+m-1}{y} (e^{-\lambda t})^m (1-e^{-\lambda t})^y, \quad y = 0,1,2,\ldots.$$

Thus $Y \sim NB(m, e^{-\lambda t})$. For population 1, let

$$Y_1 \sim NB(m_1, e^{-\lambda_2 t}).$$

For population 2, let

$$Y_2 \sim NB(m_2, e^{-\lambda_2 t}).$$

Then by taking $k_1 = k_2 = 1$, $W_1 = Y_1$, $V_1 = Y_2 + Y_1$, $p_1 = 1-e^{-\lambda_1 t}$, and $p_1' = 1-e^{-\lambda_2 t}$ in (2.6), we have for $y_1 = 0,1,2,\ldots, k$; $k = 0,1,2,\ldots,$

$$(6.1) \qquad P\{Y_1 = y_1 | Y_1 + Y_2 = k\} = c(y_1, k)/\sum_{j=0}^{k} c(j, k)$$

where

$$c(j, k) = \binom{m_1+j-1}{j}\binom{m_2+k-j-1}{k-j}\left(\frac{p_2}{p_1}\right)^j$$

and where $p_i = 1-e^{-\lambda_i t}$, $i = 1,2$. The conditional distribution (6.1) can be used to test hypotheses about λ_1 and λ_2. For example, take

$$H_0: \lambda_1 \leq \lambda_2 \text{ versus } H_1: \lambda_1 > \lambda_2.$$

We note that $\lambda_1 \leq \lambda_2$ iff

$$(1-e^{-\lambda_2 t})/(1-e^{-\lambda_1 t}) < 1.$$

Thus the problem is reduced to testing:

$$H_0: p_2/p_1 \leq 1 \text{ versus } H_1: p_2/p_1 > 1.$$

Clearly the distribution (6.1) can be used. Under the hypothesis $p_2/p_1 = 1$, we use the identity(Feller (1968), p. 65).

$$\sum_{y_1=0}^{k} \binom{m_1+y_1-1}{y_1}\binom{m_2+k-y_1-1}{k-y_1} = \binom{m_1+m_2+k-1}{k}$$

and the distribution (6.1) takes a much simpler form:

$$(6.2) \qquad P\{Y_1 = y_1 | Y_1 + Y_2 = k\} = \frac{\binom{m_1+y_1-1}{y_1}\binom{m_2+k-y_1-1}{k-y_1}}{\binom{m_1+m_2+k-1}{k}}, \quad y_1 = 0,1,\ldots, k; \quad k = 0,1,2,\ldots.$$

We recognize this distribution as the so-called "negative hypergeometric distribution" with parameters $m_1 + m_2$, m_1 and k. It is closely related to the "hypergeometric waiting time distribution."

6.5 Comparison of Two Proportions--$\theta = p_1/p_2$. Percentage Change and Effective Indices.

In various applied statistical fields, particularly in biomedical investigations, one of the most often encountered problems is to compare two proportions p_1 and p_2 of Bernoulli populations π_1, π_2. For example, one may wish to compare the success rate such as cure rate of an experimental method or drug with that of a standard method or drug. Two measures of comparison appear to be intuitively reasonable. One may measure the difference of the two proportions $\Delta = p_1 - p_2$, or equivalently $\Delta = q_2 - q_1$. Alternatively, one may measure the ratio of the two proportions $\theta \equiv p_1/p_2$. Whether Δ or θ should be used is not always clear. (A similar problem occurs in comparing two Poisson parameters.) We note θ can be employed to define a useful quantity φ:

$$\varphi = 1 - \theta = \frac{p_2 - p_1}{p_2}.$$

φ is interpreted as the percentage increase (or decrease) between the two populations. A second useful quantity related to θ is the so-called effectiveness index ζ. Let p_2 and p_1 be the success rates of an experimental method and a standard method respectively. Let $q_i = 1 - p_i$, $i = 1,2$, be the corresponding failure rates. Then

$$\zeta = \frac{p_2 - p_1}{1 - p_1} \tag{6.3}$$

$$= \frac{q_1 - q_2}{q_1} . \tag{6.4}$$

The two forms are easily seen to be equivalent. In both forms, the denominator $1 - p_1 = q_1$ is the failure rate of the standard method and is also the maximum difference possible between the experimental and the standard. Thus in form (6.4), ζ measures the reduction of failure rate made by the experimental over the standard $(q_1 - q_2)$ relative to the maximum possible improvement $(1-p_1)$. While in form (6.3), ζ measures the increase of success rate in terms of the maximum possible improvement. Greenwood and Yule (1915) called ζ "efficiency." The name "effectiveness index" is proposed by Hovland, Lumsdaine and Sheffield (1949). The readers are referred to these two papers for detailed discussion of ζ. We want to establish confidence limits for θ, φ and ζ. Since $\zeta = 1 - q_2/q_1$, the problem reduces to that of constructing confidence limits for the ratio of two Bernoulli parameters. We shall illustrate the case of θ.

Bross (1954) found an approximate confidence interval for θ using a Poisson approximation; Noether (1957) employed a normal approximation. An easy application of the theory developed in

Chapter II provides an exact solution. We employ IBS plan on populations π_i, $i = 1,2$, until r_i^{th} failure occurs. If X_i denotes the number of successes prior to the r_i^{th} failure, then

$$X_i \sim NB(r_i, q_i),\ i = 1,2.$$

By taking $k_1 = k_2 = 1$, $W_1 = X_1$, $V_1 = X_1 + X_2$ in (2.6), we obtain for $x_2 = 0,1,2,\ldots, k$; $k = 0,1,2,\ldots$

$$(6.5) \qquad P\{X_1 = x_1 | X_1 + X_2 = k\} = h(x_1; k)\theta^{x_1} / \sum_{t=0}^{k} h(t; k)\theta^t,$$

where

$$h(t; k) = \binom{r_1+t-1}{t}\binom{r_2+k-t-1}{k-t}$$

$$\theta = p_1/p_2.$$

Since (6.5) depends on θ only, it can be used to construct confidence limits for θ. By relabeling X_i's to represent number of failures prior to the r_i^{th} success, the same procedure can be used to find confidence limits for the ratio q_1/q_2. A binomial approximation will be briefly mentioned.

Let $r_1 \to \infty$, $r_2 \to \infty$, $r_1/(r_1 + r_2) \to p$, where $0 < p < 1$ is a constant. Then, by Stirling's formulas, (6.5) has the limiting distribution $B(k, \frac{\theta p}{\theta p + q})$ (see Appendix B). It is interesting to note that the same limiting distribution can also be obtained under different setof assumptions. $r_i \to \infty$, $p_i \to 0$, such that $r_i p_i \to \lambda_i$ where λ_i is a positive constant, $i = 1,2$. We simply note these assumptions make it possible to approximate X_i by V_i, where

$$V_i \sim P_0(\lambda_i),\ i = 1,2.$$

Now it is well known that

$$P\{V_1 = x_1 | V_1 + V_2 = k\}$$

is $B(k, \lambda)$, where $\lambda = \frac{\lambda_1}{\lambda_1 + \lambda_2}$. We see that λ is the same as $\theta p/(\theta p+q)$.

Another approximation is to use the fact mentioned at the end of Section 3.4 that when p_i is small, $X_i \sim NB(r_i, p_i)$, $i = 1,2$, can be approximated by Chi-square or gamma distribution. The distribution of the quotient of two independent gamma r.v.'s is known. Using this, it can be shown $W = X_1/X_2$ has approximately the density

$$f_W(w) = \frac{1}{B(r_1, r_2)} \frac{(\theta w)^{r_1}}{w(1+\theta w)^{r_1+r_2}}, \quad w > 0$$

where $\theta = p_1/p_2$.

Sometimes it is desired to compare two effectiveness indices ζ_1 and ζ_2. For example, a new medical treatment is administered to two groups of subjects (examples: two ethnic groups or two groups of people belonging to the same race but living in different localities such as Japanese living in Japan and American Japanese). Each type of subject has a control group and an experimental group. Thus there are four groups with success rates p_1, p_2 and p_1', p_2', say. Then ζ for type 1 subjects is:

$$\zeta_1 = \frac{p_2 - p_1}{1 - p_1},$$

and for type 2 subjects:

$$\zeta_2 = \frac{p_2' - p_1'}{1 - p_1'}.$$

Let $q_i = 1 - p_i$, $q_i' = 1 - p_i'$, $i = 1,2$. To compare ζ_1 and ζ_2, we note that $\zeta_1 > \zeta_2$ iff $1 - \zeta_1 < 1 - \zeta_2$ iff $q_2/q_1 < q_2'/q_1'$, which is equivalent to

$$(6.6) \qquad \gamma \equiv (q_1 q_2')/(q_1' q_2) > 1.$$

The quantity γ is the ratio of the products of Bernoulli parameters; therefore it is a special case of $\theta = p_1 p_2 \cdots p_{k_1} / p_1' p_2' \cdots p_{k_2}'$ in Chapter 2. For each of the four groups of subjects, we observe the number of failures prior to the r_i^{th} success, $i = 1,2,3,4$. We shall obtain four negative binomial r.v.'s:

$$X_i \sim NB(r_i, p_i), \ i = 1,2$$
$$X_3 \sim NB(r_3, p_1')$$
$$X_4 \sim NB(r_4, p_2').$$

Let $u = (w_4, v_2, v_3)$. Then from equation (2.6) we have:

$$(6.7) \qquad P\{X_1 = x_1 | X_4 - X_1 = w_4, X_2 + X_1 = v_2, X_3 + X_1 = v_3\}$$
$$= b(x_1, u)\gamma^{x_1} / \sum_t b(t, u)\gamma^t,$$

where

$$b(t, u) = \prod_{j=2}^{3} \binom{r_1+x_1-1}{x_1} \binom{r_4+w_4+x_1-1}{w_4+x_1} \binom{r_j+v_j-x_1-1}{v_j-x_1}$$

and the ranges of values of the variables are:

$w_4 = -x_1, -x_1+1, -x_1+2, \ldots$

$v_j = x_1, x_1+1, x_1+2, \ldots, \ j = 2,3$

x_1 is a non-negative integer satisfying $\max(0, -w_4) \leq x_1 \leq \text{Min}(v_2, v_3)$

t has the same range of values as x_1.

Clearly we can use (6.7) to test the hypothesis

$$H_0: \zeta_1 = \zeta_2$$

because H_0 is equivalent to $\gamma = 1$.

APPENDIX A

PROOF OF THEOREM 3.1

Using (3.7) and (3.8), (3.31) can be written as:

$$(A.1) \qquad P\{X_1 = x_1 \mid X_2 - X_1 = w\} = A(r_1, r_2; \mu, x_1 - \alpha) / \sum_{j=0}^{\infty} A(r_1, r_2; \mu, j)$$

where

$$A(r_1, r_2; \mu, t) = \frac{\Gamma(r_1+\alpha+t)}{\Gamma(r_1+\alpha)} \frac{\Gamma(r_2+\beta+t)}{\Gamma(r_2+\beta)} \frac{\mu^t}{\Gamma(1+|w|+t)t!} .$$

Suppose for $n = 1,2,\ldots$, X_{1n}, X_{2n} are negative binomial r.v.'s, $X_{in} \sim NB(r_{in}, p_{in})$, $i = 1,2$, where $r_{in} \to \infty$, $q_{in} = 1 - p_{in} \to 0$ such that $(r_{1n}+\alpha)(r_{2n}+\beta)q_1q_2 = \lambda$ as $n \to \infty$, and X_{1n} is independent of X_{2n} for each n. Then from (1):

$$(A.2) \qquad P\{X_{1n} = x_1 \mid X_{2n} - X_{1n} = w\} = A(r_{1n}, r_{2n}; \mu, x_1-\alpha) / \sum_{j=0}^{\infty} A(r_{1n}, r_{2n}; \mu, j$$

Now

$$A(r_{1n}, r_{2n}; \mu, t) = \frac{\Gamma(r_{1n}+\alpha+t)}{\Gamma(r_{1n}+\alpha)} / (r_{1n}+\alpha)^t \cdot \frac{\Gamma(r_{2n}+\beta+t)}{\Gamma(r_{2n}+\beta)} / (r_{2n}+\beta)^t \frac{\lambda^t}{\Gamma(1+|w|+t)t!} .$$

By Lemma 1 of Appendix B, for any $a > 0$, as $x \to \infty$

$$(A.3) \qquad \frac{\Gamma(x+a)}{x^a\Gamma(x)} \to 1.$$

Therefore

$$(A.4) \qquad \lim_{n \to \infty} A(r_{1n}, r_{2n}; \mu, t) = \frac{\lambda^t}{\Gamma(1+|w|+t)t!} .$$

If we can justify that

$$\text{(A.5)} \qquad \lim_{n\to\infty}\{\lim_{m\to\infty} g_m(n)\} = \lim_{m\to\infty}\{\lim_{n\to\infty} g_m(n)\}$$

where $g_m(n) = \sum_{y=0}^{m} A(r_{1n}, r_{2n}; \mu, y)$, then using (A.4), we shall have:

$$\text{(A.6)} \qquad \lim_{n\to\infty} P\{X_{1n} = x_1 \mid X_{2n} - X_{1n} = w\} = \frac{\lambda^{x_1-\alpha}}{\Gamma(1+|w|+x_1-\alpha)(x_1-\alpha)!} \div \sum_{y=0}^{\infty} \frac{\lambda^{y}}{\Gamma(1+|w|+y)y!}$$

$$= \frac{\lambda^{x_1+(w/2)}}{x_1!(x_1+w)!\,I_{|w|}(2\sqrt{\lambda})}$$

as claimed by the theorem.

Now we show that the interchange of the two limit operations in (A.5) is permissible. For this purpose we appeal to the "iterated limit theorem" (see, for example, Apostle (1957), p. 374, Theorem 13.2) by showing that on the left side of (A.5):

A. The iterated limit $\lim_{n\to\infty}\{\lim_{m\to\infty} g_m(n)\}$ exists.

B. $\lim_{m\to\infty} g_m(n)$ converges uniformly for all n.

For A: We note for any fixed n, the limit

$$\text{(A.7)} \qquad g(n) = \lim_{m\to\infty} g_m(n)$$

exists, $n = 1,2,\ldots$. This is so because the RHS of (A.7) is $\frac{1}{\Gamma(1+|w|)} F(r_1+\alpha, r_2+\beta, 1+|w|; \mu)$, which is finite. Next observe that

$$g(n) \geq g(n+1) \geq 0 \quad \text{for all } n.$$

(Apply the observation that in (A.3), $\frac{\Gamma(x+a)}{\Gamma(x)x^a}$ monotonically decreases to unity as x increases.) Therefore $\lim_{n\to\infty} g(n) = a$ say, exists, and assertion A. is proved. For B., we see that for all $n = 1,2,\ldots,$ and all $m = 1,2,\ldots$

$$(A.8) \qquad |g_m(n) - g(n)| = \Big| \sum_{y=m+1}^{\infty} A(r_{1n}, r_{2n}; \mu, y) \Big|$$

$$< \Big| \sum_{y=m+1}^{\infty} A(r_{11}, r_{21}; \mu, y) \Big|$$

$$= |g_m(1) - g(1)|.$$

But from (A.7), we have $\lim_{m\to\infty} g_m(1) = g(1)$. Therefore, given $\epsilon > 0$ there exists $M(\epsilon)$ such that

$$(A.9) \qquad |g_m(1) - g(1)| < \epsilon, \text{ for all } m \geq M(\epsilon).$$

Applying this result to (A.8), we have that for any $\epsilon > 0$, there exists $M(\epsilon)$ such that $|g_m(n) - g(n)| < \epsilon$, for all $m \geq M(\epsilon)$ and all $n = 1,2,\ldots.$ Thus $g_m(n)$ converges to $g(n)$ uniformly for all n. This proves B., and hence the theorem.

APPENDIX B

BINOMIAL APPROXIMATION TO THE DISTRIBUTION (6.5)

The following lemma is a consequence of Stirling's formula.

Lemma 1.

Let $A(a, x) = \Gamma(x+a)/x^a\Gamma(x)$, where $a > 0$, $x > 0$. Then $\lim_{x \to \infty} A(a, x) = 1$ for any $a > 0$.

To show that (6.5) can be approximated by a binomial distribution, let

$$f(r_1, r_2, t, k) = \binom{r_1+t-1}{t}\binom{r_2+k-t-1}{k-t}$$

$$f(r, k) = r^k/k!$$

Then

$$f(r_1, r_2, t, k)/f(r_1+ r_2, k) = \frac{k!}{t!(k-t)!}\left(\frac{r_1}{r_1+ r_2}\right)^t\left(\frac{r_2}{r_1+ r_2}\right)^{k-t} \cdot A(t, r_1)A(k-t, r_2)A(k, r_1+ r_2).$$

By letting $r_1 \to \infty$, $r_2 \to \infty$, $r_1/(r_1+ r_2) \to p$, $0 < p < 1$. We have by Lemma 1

$$\lim f(r_1, r_2, t, k)/f(r_1+ r_2, k) = \binom{k}{t}p^t(1-p)^{k-t}.$$

Now, if we divide numerator and denominators of the right hand side of (6.5) by $f(r_1 + r_2, k)$ we obtain

$$P\{X_1= x_1 | X_1+ X_2= k\} \to \frac{\binom{k}{x_1}p^{x_1}(1-p)^{k-x_1}\theta^{x_1}}{\sum_{t=0}^{k}\binom{k}{t}p^t(1-p)^{k-t}\theta^t}$$

$$= \binom{k}{x_1}\lambda^{x_1}(1-\lambda)^{k-x_1},$$

which is a $B(k, \lambda)$ distribution with $\lambda = (\frac{p\theta}{p\theta+q})$, $q = 1 - p$.

REFERENCES

1. Apostol, T. M. (1957). Mathematical Analysis. Addison-Wesley Publishing Co., Reading, Mass.
2. Barlow, R. E. and Proschan, F. (1965). Mathematical Theory of Reliability. John Wiley and Sons, New York.
3. Birnbaum, A. (1954). Statistical methods for Poisson processes and exponential populations. J. Amer. Statist. Assoc. 49 254-266.
4. Birnbaum, Z. W., Esary, J. D., and Saunders, S. C. (1961). Multi-component systems and structures and their reliability. Technometrics 3 55-71.
5. Bross, Irwin (1954). A confidence interval for a percentage increase. Biometrics 10 245-250.
6. Buehler, R. J. (1957). Confidence limits for the product of two binomial parameters. J. Amer. Statist. Assoc. 52 482-493.
7. Chapman, D. G. (1952). Inverse multiple and sequential sample censuses. Biometrics 8 286-306.
8. Cochran, William G. (1950). Estimation of bacterial densities by means of the most probable number. Biometrics 6 105-116.
9. De Cicco, Henry (1960). The Error in Linearized Estimates of the Variance of Products. Technical Note No. 2, Office for Reliability, Research and Effects (Reliability Branch - ORDSW-DR).
10. El Mawaziny, A. H. (1965). Chi-Square Distribution Theory with Applications to Reliability Problems. PhD Thesis, Iowa State University, Ames.

11. Feller, W. (1968). An Introduction to Probability Theory and Its Applications, Vol. 1, 3rd Ed. John Wiley and Sons, New York.

12. Ferguson, T. S. (1965). A characterization of the geometric distribution. Amer. Math. Monthly 72 256-260.

13. General Electric Co. (1962). Tables of the Individual and Cumulative Terms of the Poisson Distribution. Van Nostrand, New York.

14. Greenwood, M. and Yule, G. W. (1915). The statistics of anti-typhoid and anti-cholera innoculations and the interpretation of such statistics in general. Proc. Royal Society of Medicine 8 Part 2, 113-194.

15. Haldane, J. B. S. (1945). On a method of estimating frequencies. Biometrika 33 222-225.

16. Harris, Bernard (1968). Hypothesis Testing and Confidence Intervals for Products and Quotients of Poisson Parameters with Applications to Reliability. MRC Technical Summary Report No. 923, U.S. Army Mathematics Research Center. The University of Wisconsin, Madison.

17. Hovland, Carl I., Lumsdaine, A. A., and Sheffield, F. D. (1949). Experiments in Mass Communication. Volume III of Studies in Social Psychology in World War II. Princeton University Press, Princeton, New Jersey.

18. Lehmann, E. L. (1959). Testing Statistical Hypotheses. John Wiley and Sons, New York.

19. Lentner, M. M. and Buehler, R. J. (1963). Some inferences about gamma parameters with an application to a reliability problem. J. Amer. Statist. Assoc. 58 670-677.

20. Lieberman, G. J. and Ross, S. M. (1970). Confidence intervals for independent exponential series systems. Department of Operations Research and Department of Statistics, Stanford University. Technical Report No. 130.

21. Madansky, Albert (1965). Approximate confidence limits for the reliability of series and parallel systems. Technometrics 7 495-503.

22. Malik, H. B. (1968). Exact distribution of the product of independent generalized gamma variables with the same shape parameter. Ann. Math. Stat. 39 1751-1752.

23. Mann, N. R. (1970). Computer aided selection of prior distributions for generating Monte Carlo confidence bounds on system reliability. Naval Research Logistics Quarterly 17 41-54.

24. Mine, H. (1959). Reliability of physical system. Transactions of the 1959 International Symposium on Circuit and Information Theory.

25. Myrhe, Janet and Saunders, Sam C. (1965). On confidence limits for the reliability of systems. Ann. Math. Stat. 39 1463-1472.

26. Nadler, J. (1960). Inverse binomial sampling plans when an exponential distribution is sampled with censoring. Ann. Math. Stat. 31 1201-1204.

27. Noether, G. E. (1957). Two confidence intervals for the ratio of two probabilities and some measures of effectiveness. J. Amer. Statist. Assoc. 52 36-45.

28. Parzen, E. (1962). Stochastic Processes. Holden-Day, Inc., San Francisco.

29. Paulson, Edward (1967). Sequential procedures for selection of the best one of several binomial populations. Ann. Math. Stat. 38 117-123.

30. Pearson, Karl (1932). Tables of the Incomplete Beta Functions. Cambridge Univ. Press, Cambridge, England.

31. Rand Corporation (1955). A Million Random Digits with 100,000 Normal Deviates. Free Press, Glencoe, Illinois.

32. Rosenblatt, Joan Raup (1963). Confidence limits for the reliability of complex systems. Statistical Theory of Reliability, 115-137. Ed. Marvin Zelen. The University of Wisconsin Press, Madison.

33. Sandelius, D. M. (1950). An inverse sampling procedure for bacterial place counts. Biometrics 6 291-292.

34. Shooman, M. L. (1968). Probabilistic Reliability: An Engineering Approach. McGraw-Hill Book Co., New York.

35. Thomas, Ralph E. (1960). An Improved Formula for the Standard Deviation of the Reliability Product Rule. Technical Report No. 2, Battelle Memorial Institute.

36. Tucker, H. G. (1967). A Graduate Course in Probability. Academic Press, New York.

37. Whittaker, E. T. and Watson, G. N. (1915). A Course of Modern Analysis, 2nd. ed. Cambridge Univ. Press, Cambridge, England.