ON THE EFFICIENCY OF CERTAIN TESTS FOR 2 x 2 TABLES

by

John H. MacKay
University of North Carolina

This research was supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command under contract No. AF 18(600) - 458.

Institute of Statistics
Mimeograph Series No. 146
April, 1956

A C K N O W L E D G E M E N T S

Few individuals can carry out protracted programs of graduate study without the aid and encouragement of a great many persons. I am happy to acknowledge my indebtedness to all those who have assisted me.

I want particularly to thank Professor Wassily Hoeffding, who proposed the problem, read rough drafts promptly and carefully, and gave many helpful suggestions.

The financial assistance of the Department of Statistics and the United States Air Force is gratefully acknowledged.

Finally, I wish to thank my wife, Patricia, whose faith and understanding have been constant sources of inspiration to me.

John H. MacKay

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

The theorems and lemmas in this paper were developed in an attempt to find a "best" way of analyzing the results of an experiment of the type which Barnard [1] calls a "2 x 2 comparative trial". For instance, suppose there are two manufacturing processes such that for each process a particular unit of product will be defective with some fixed but unknown probability. In order to decide in which process this probability is the least (or possibly to decide that the probabilities are equal) we will take samples of sizes m and n from the two, and count the numbers u and v (say) of defectives. The results of the experiment may then be put in tabular form.

	Defective	Not Defective	Totals
Process I	u	$m - u$	m
Process II	v	$n - v$	n
Totals	S	$N - S$	N

Thus, we are led to a 2 x 2 contingency table with one set of marginal totals fixed. The analysis of such tables has occupied the attention of statisticians for more than half a century [12].

In the terminology of Barnard [1], such a table contains data resulting from one of three kinds of experiments:

(i) Independence Trial (Both sets of marginal totals variable),

(ii) 2 x 2 Comparative Trial (One set of marginal totals variable),

1 Numbers in square brackets refer to bibliography

(iii) Double Dichotomy (Marginal totals fixed).

The first kind (i) is exemplified by the "inoculated or not-inoculated, infected or not-infected" problem, where the sample is chosen at random from the population. Fisher's [6] lady tasting tea is an example of the third kind (iii).

One of the puzzling questions for a long time was whether or not the same test should be used in all three situations. The approximate χ^2 test, which Cochran [3] calls the X^2 test, is equally valid in all three cases for large samples, as Cochran points out. This is not so in small samples because the distribution of X^2 is then different in the three cases. The Cochran article cited, incidentally, has an excellent bibliography of the whole history of the 2 x 2 contingency table problem. Fisher [5] recommended that the same "exact" test be used in all three situations. Finally, Tocher [14] showed that a modified form of Fisher's test (Fisher-Tocher test) is a most powerful one-tailed test (in the Neyman-Pearson sense) with any of the three types of data. One result which we here obtain is that the two-sided Fisher-Tocher test and the X^2 test are both asymptotically efficient in a certain sense.

Let the probabilities of defectives in the two populations be ξ and η respectively. We define three types of problems for $0 < \delta < 1$:

<u>Type I</u> $\omega_1(\delta) = \{\xi-\eta \leq -\delta\}$ vs. $\omega_2(\delta) = \{\xi-\eta \geq \delta\}$,

<u>Type II</u> $\omega_1 = \{\xi-\eta = 0\}$ vs. $\omega_2(\delta) = \{|\xi-\eta| \geq \delta\}$,

<u>Type III</u> $\omega_1 = \{\xi-\eta \leq 0\}$ vs. $\omega_2(\delta) = \{\xi-\eta \geq \delta\}$.

We obtain various tests, all depending only on linear functions of the sample percentages, which are asymptotically "efficient" in certain senses (as subsequently defined), for the following situations in which we have preassigned upper bounds α and β to the error probabilities.

(a) For Equal Sample Sizes

<u>Type I</u>. Exact efficient test when $\alpha = \beta$ (Chapter II).

<u>Type III</u>. Asymptotically efficient test where $\frac{\alpha}{\beta}$ remains fixed and α tends to zero (Chapter V).

(b) For Unequal Sample Sizes m and n, $\frac{m}{n}$ Fixed

<u>Types I, II, and III</u>. Asymptotically efficient tests where α, β remain fixed and δ tends to zero. Asymptotic expressions for the most "economical" ratios $\frac{m}{n}$, which minimize the total cost of sampling when such tests are used (Chapter III).

<u>Type I</u>. Asymptotically efficient test where $\frac{\alpha}{\beta}$ remains fixed and α tends to zero. Asymptotic expression for the minimum sample size (Chapter V).

We define a class of bayes tests which have certain desirable properties, and in some special cases where the sample sizes are equal and small, compare these bayes tests with the corresponding Fisher-Tocher tests previously mentioned (Chapter II).

Finally, we show (Chapter IV) that the Fisher-Tocher test is asymptotically efficient for the type III problem as $\delta \to 0$, and that the two sided Fisher-Tocher test and the X^2 test are both asymptotically efficient for the type II problem as $\delta \to 0$.

We will begin by formulating a slightly more general problem than the one posed above.

Let $x_k = (y_1, y_2, \ldots, y_k)$ represent k independent observations of a random variable Y. Suppose the random variable X_k has the generalized density $f_\theta(x, k)$ with respect to a fixed measure μ_k for x in $\mathcal{X}_k$ and all θ in some closed Euclidean space Θ.

Suppose there are two alternative courses of action associated with two decisions D_1, D_2, and subsets ω_1, ω_2 of Θ, such that D_i is preferred for θ in $\omega_i (i = 1, 2)$, and it does not particularly matter which decision is taken for θ in $\Theta - \omega_1 - \omega_2$. The decision is to be taken on the basis of a sample of size k and a critical function ϕ_k' : when the sample size is k and x is observed, take D_2 with probability $\phi_k'(x)$ and D_1 with probability $1 - \phi_k'(x)$. For this sample size the error probabilities are $E_\theta \phi_k'$ when θ is in ω_1 and $E_\theta \{ 1-\phi_k' \}$ when θ is in ω_2. In general, we will require that these probabilities be bounded above by preassigned numbers α and β. We will choose k large enough so that

$$(1.1) \qquad \sup_{\omega_1} E_\theta \phi_k' \leq \alpha , \quad \sup_{\omega_2} E_\theta \{ 1-\phi_k' \} \leq \beta \quad ,$$

where $\alpha > 0$, $\beta > 0$, and $\alpha + \beta < 1$.

In order to simplify the notation we will define the following risk functions for any test based on k observations, and related useful concepts.

$$(1.2) \qquad r_k^*(\theta,\phi') = \begin{cases} E_\theta \phi' & \text{for } \theta \text{ in } \omega_1 \\ E_\theta \left\{ 1-\phi' \right\} & \text{for } \theta \text{ in } \omega_2 \\ 0 & \text{otherwise} \end{cases} ,$$

$$r_{ik}^*(\phi') = \sup_{\omega_i} r_k^*(\theta,\phi') \quad \text{for } i = 1,\ 2 ,$$

$$a = \beta/(\alpha + \beta), \quad b = \alpha/\alpha + \beta) ,$$

$$r_k(\theta,\phi') = \begin{cases} a\, E_\theta \phi' & \text{for } \theta \text{ in } \omega_1 \\ b E_\theta \left\{ 1-\phi' \right\} & \text{for } \theta \text{ in } \omega_2 \\ 0 & \text{otherwise} \end{cases} ,$$

$$r_{ik}(\phi') = \sup_{\omega_i} r_k(\theta,\phi') \quad \text{for } i = 1,\ 2 ,$$

$$r_k(\phi') = \sup_{\Omega} r_k(\theta,\phi') = \max \left\{ r_{1k}(\phi'), r_{2k}(\phi') \right\} ,$$

$$r_k = \inf_{\phi'} r_k(\phi')$$

We will show that conditions (1.1) are equivalent to the following condition:

$$(1.2.1) \qquad r_k(\phi_k') \leq a\alpha .$$

To prove the statement, suppose $r_k(\phi_k') = \max \left\{ r_{1k}(\phi_k'), r_{2k}(\phi_k') \right\} \leq a\alpha = b\beta$. Then $r_{1k}^*(\phi_k') = (\frac{1}{a}) r_{1k}(\phi_k') \leq \alpha$ and $r_{2k}^*(\phi_k') = (\frac{1}{b}) r_{2k}(\phi_k') \leq \beta$, which imply (1.1). On the other hand, $r_{1k}^*(\phi_k') \leq \alpha$ implies $r_{1k}(\phi_k') \leq a\alpha$, and $r_{2k}^*(\phi_k') \leq \beta$ implies $r_{2k}(\phi_k') \leq b\beta = a\alpha$. Hence

$r_k(\phi'_k) \leq a\alpha$.

We can therefore restrict ourselves to finding tests satisfying (1.2.1), and in what follows we will assume there is at least one finite k and a test ϕ'_k satisfying the condition (1.2.1). There will, in general, be many tests and sample sizes for which this condition is satisfied. One optimal way of choosing among these is to pick a test, not necessarily unique, satisfying (1.2.1), for which the sample size k is a minimum.

(1.3) Definitions.

N^* = least sample size k for which condition (1.2.1) can be satisfied.

Let $\{\phi'_\lambda\}$ be a sequence of tests, one for each finite sample size $k = \lambda$.

Let T be the set of sequences $\{\phi'_\lambda\}$, such that $\{\phi'_\lambda\}$ is in T if and only if there exists a finite k such that $\phi'_k \varepsilon \{\phi'_\lambda\}$ satisfies condition (1.2.1).

Let $N^*\{\phi'_\lambda\}$ = least integer k such that $\phi'_k \varepsilon \{\phi'_\lambda\}$ satisfies condition (1.2.1).

It then follows that

$$N^* = \min_{\{\phi'_\lambda\} \varepsilon T} N^*\{\phi'_\lambda\} \quad . \tag{1.3.1}$$

Let N = least integer k such that $r_k \leq a\alpha$.

Our first problem is to find a test ϕ_{N^*} which is based on a size N^* and at the same time satisfies condition (1.2.1). We will now define an efficient test.

(1.3.2) A test will be called efficient if it is based on a sample size N^* and satisfies condition (1.2.1).

We now present some lemmas.

Lemma (1.4)

For fixed k, if there is a test ϕ_k such that $r_k(\phi_k) = r_k$, then,

$$r_{1k}(\phi_k) = r_{2k}(\phi_k) \quad .$$

Proof of Lemma (1.4)

Suppose the lemma is false. Then there exists $\varepsilon > 0$, and a test ϕ_k, such that $r_{ik}(\phi_k) + \varepsilon = r_{jk}(\phi_k) = r_k$ for $i \neq j$ (= 1,2). We may suppose without loss of generality that $i = 1$, $j = 2$.

Let $\phi_{yk} = y + (1-y)\phi_k$. This is a test for $0 \leq y \leq 1$. Then $1 - \phi_{yk} = (1-y)(1-\phi_k)$. Determine y so that $r_{1k}(\phi_{yk}) = r_{2k}(\phi_{yk})$. Then, $ay + (1-y)r_{1k}(\phi_k) = (1-y)(r_{1k}(\phi_k) + \varepsilon)$, or $y = \frac{\varepsilon}{a+\varepsilon}$, and $0 < y < 1$. For such y, $r_k(\phi_k)-r_k(\phi_{yk}) = r_{2k}(\phi_k)-r_{2k}(\phi_{yk}) = yr_{2k}(\phi_k) \geq y\,\varepsilon > 0$. Hence, $r_k(\phi_{yk}) < r_k(\phi_k) = r_k$, which is impossible.

Lemma (1.5)

If, for some fixed k, there is a test ϕ_{1k} which satisfies conditions (1.1), and for the same k there is a test ϕ_k such that $r_k(\phi_k) = r_k$, then ϕ_k also satisfies the conditions (1.1).

Proof of Lemma (1.5)

$r_k(\phi_k) \leq (\phi_{1k}) \leq a\alpha = b\beta$ because ϕ_{1k} also satisfies condition

(1.2.1). Therefore, $\emptyset_k$ satisfies condition (1.2.1), which is equivalent to conditions (1.1). The proof is complete.

It is clear that there is a sequence $\{\emptyset_\lambda^*\}$ in T such that

(1.6) $$N^* = N^*\{\emptyset_\lambda^*\} .$$

Lemma (1.7)

If, for every positive integer k, there is a test $\emptyset_k$ such that $r_k(\emptyset_k) = r_k$, then $N = N^*$.

Proof of Lemma (1.7)

By definition of N^*, there is a test $\emptyset_{N^*}^*$ such that $r_{N^*} \leq r_{N^*}\left(\emptyset_{N^*}^*\right) \leq a\alpha$. Hence, $N \leq N^*$. By definition of N and assumption of the lemma, there is a test $\emptyset_N$ such that $r_N\left(\emptyset_N\right) \leq a\alpha$. Hence, $N^* \leq N^*\left\{\emptyset_\lambda\right\} \leq N$. This completes the proof.

An Extended Problem

Suppose we may take any number of observations from each of γ populations. The cost of a single observation from the i^{th} population is $c_i > 0$, independent of the number of observations, for $i = 1, 2, \ldots, \gamma$. The problem is to minimize the total sampling cost subject to conditions (1.1) or analogous conditions. It will be convenient in what follows to assume fixed ratios among the sample sizes. For instance, let $n_i/n_1 = \nu_i$ for $i = 1, 2, \ldots, \gamma$, where n_i is the number of observations from the i^{th} population. Then we may be able, for each fixed $\nu = (\nu_1, \nu_2, \ldots, \nu_\gamma)$, to determine our best test and minimum sample size. Subsequently we might be able to choose ν in such a way that the total cost of sampling is minimized. For such a setup the total cost of sampling (if $n_1 = k$) is

$$C(k,\nu) = c_1n_1+c_2n_2+\ldots+c_\gamma n_\gamma=n_1(c_1\nu_1+\ldots+c_\gamma\nu_\gamma)=k(c_1\nu_1+\ldots+c_\gamma\nu_\gamma).$$

We may assume $\nu_1, \ldots, \nu_\gamma$ are integers, since, for $k = k'n_1(k', n_1$ positive integers), we have $C(k,\nu) = k'(c_1n_1 + c_2n_2 + \ldots + c_\gamma n_\gamma)$.

Let V be the γ-dimensional space of vectors ν with non-negative integral components and $\nu_1 = 1$, and such that for each ν in V, condition (1.2.1) can be satisfied for some finite k, where $n_i = k\nu_i$ for $i = 1, \ldots, \gamma$. For each ν in V we define quantities like those in sections (1.2) and (1.3). For instance, we replace N^* by N^*_ν , T by T_ν, ϕ_λ by $\phi^{(\nu)}_\lambda$, r_k by $r_k^{(\nu)}$, etc.

We assume conditions (1.1) can be satisfied with finite cost and define C to be the minimum cost for which condition (1.2.1) can be satisfied. Then,

$$C = \inf_V C(N^*_\nu,\nu) = \inf_V \ (N^*_\nu)(c_1\nu_1+\ldots+c_\gamma\nu_\gamma) \quad . \tag{1.8}$$

Lemma (1.9)

If, for all positive integral k and all ν in V, there are tests $\phi_k^{(\nu)}$, such that $r_k^{(\nu)}\left(\phi_k^{(\nu)}\right) = r_k^{(\nu)}$, then,

(a) $C = \inf_V \ (N_\nu)(c_1\nu_1+\ldots+c_\gamma\nu_\gamma)$,

(b) There is a test which satisfies conditions (1.1 or 1.2.1) with cost C.

Proof of Lemma (1.9)

The truth of statement (a) follows directly from lemma (1.7).

To prove part (b), let $V_1 = \{\nu : c_i\nu_i \leq C+1\}$. Since $\alpha + \beta < 1$, we have $N_\nu \geq 1$. Hence, ν in $V - V_1$ implies

$$N_\nu(c_1\nu_1+\ldots+c_\gamma\nu_\gamma) > \gamma(C+1) > C$$

Hence, $C = \min_{V_1} C(N_\nu,\nu) = C(N_{\nu_0},\nu_0)$ for some ν_0 in V_1(a finite set).

By assumption, $r_{N_{\nu_0}}\left\{\emptyset^{(\nu_0)}_{N_{\nu_0}}\right\} \leq a\alpha$, and $\emptyset^{(\nu_0)}_{N_{\nu_0}}$ satisfies condition (1.2.1) with cost C. This completes the proof.

The question naturally arises, "When can conditions (1.1) be satisfied?" The following simple lemma gives some information on this point.

Lemma (1.10)

If, for any test $\emptyset'_k$, $E_\theta\emptyset'_k$ is a continuous function of $\theta \in \underline{\Omega}$, $\alpha + \beta < 1$, and ω_1 and ω_2 have a limit point in common, then conditions (1.1) cannot be satisfied for any finite k.

Proof of Lemma (1.10)

If the lemma were false there would exist a test $\emptyset$ satisfying (1.1) for some finite k. There exist sequences $\{\theta_{ij}\} \in \omega_i$ for $i = 1, 2$ and all $j = 1, 2, \ldots$, such that $\theta_{ij} \longrightarrow \theta$ as $j \longrightarrow \infty$ for $i = 1, 2$. Then,

$$E_{\theta_{1j}}\emptyset \leqq \sup_{\omega_1} E_\theta\emptyset \leqq \alpha \quad \text{for } j = 1, 2, \ldots ,$$

and

$$E_{\theta_{2j}}(1-\emptyset) \leqq \sup_{\omega_2} E_\theta(1-\emptyset) \leqq \beta \text{ for } j = 1, 2, \ldots .$$

Taking the limits on the left in the two preceding equations as $j \longrightarrow \infty$, and making use of the continuity property of $E_\theta\emptyset$, we have $E_\theta\emptyset \leq \alpha$ and $E_\theta(1-\emptyset) \leq \beta$. Hence $\alpha + \beta \geq 1$, a contradiction.

CHAPTER II

CHOOSING BETWEEN TWO BINOMIAL POPULATIONS WHEN THE SAMPLE SIZES ARE EQUAL AND THE LOSS FUNCTION IS SYMMETRICAL

In this chapter we consider tests based on two samples of equal size from two binomial populations, and we assume that the loss function satisfies certain symmetry conditions. We show that, when looking for minimax tests, we may restrict ourselves to a class T_0 of tests having certain symmetry properties, and we show (Theorem 2.3) that there is a unique test in T_0 which uniformly minimizes the risk among all tests in T_0.

When the loss function is of a certain simple type we show that the above mentioned test is a bayes test and can differ from any other minimax test only on the set where the two sample percentages are equal. We obtain (Theorem 2.6.7) simple upper and lower bounds to the maximum risk of this bayes test.

We conclude by comparing the risk of a certain bayes test with that of a randomized version of Fisher's "exact" test as defined by Tocher [14], obtaining some general results and considering some special cases in which the sample sizes are small.

2.1 The Problem.

Let $Y_1, Y_2, \ldots$ ad. inf. be a sequence of independently and identically distributed random vectors $Y_i = (X_{1i}, X_{2i})$, where X_{1i} and

X_{2i} are independent for all i, and

(2.1.1)
$$X_{1i} = \begin{Bmatrix} 1 & \text{with probability } \xi \\ 0 & \text{with probability } 1-\xi \end{Bmatrix},$$

$$X_{2i} = \begin{Bmatrix} 1 & \text{with probability } \eta \\ 0 & \text{with probability } 1-\eta \end{Bmatrix}.$$

Let $\theta = (\xi,\eta)$ and $\Omega = \{ \theta: 0 \leq \xi, \eta \leq 1 \}$.

On the basis of an observation, $(y_1, y_2, \ldots, y_k)$, we want to make one of the two decisions (D_1, D_2). We consider test functions ϕ' based on sample size k: when $(y_1, \ldots, y_k)$ is observed, take D_2 with probability $\phi'(y_1, \ldots, y_k)$, and take D_1 with probability $1 - \phi'(y_1, \ldots, y_k)$. Let $w_i(\theta)$ be the "loss" associated with taking decision D_i $(i = 1, 2)$ when θ is the true parameter.

(2.1.2) We will consider only loss functions which have the following properties:

(2.1.2.1) $w_1(\xi,\eta) = w_2(1-\xi,1-\eta)$,

(2.1.2.2) $w_1(\xi,\eta) = w_2(\eta,\xi)$,

(2.1.2.3 $(\xi-\eta)[w_1(\xi,\eta)-w_2(\xi,\eta)] \geq 0$ for all $(\xi,\eta) \in \Omega$,

and inequality is strict for some $\theta = (\xi,\eta)$, where θ is in the interior of Ω.

In particular, the following loss functions satisfy conditions (2.1.2).

(2.1.3)
$$w_i(\theta) = \begin{Bmatrix} 1 & \text{if } \theta \in \omega_j \\ 0 & \text{otherwise} \end{Bmatrix} \text{ for } i \neq j = 1, 2,$$

where $0 < \delta < 1$, $\omega_1 = \{ \theta: \xi-\eta \leq -\delta \}$, and $\omega_2 = \{ \theta: \xi-\eta \geq \delta \}$.

The risk functions when the sample size is k and the test ϕ' is used are given by

$$(2.1.4) \qquad r_k(\theta,\phi') = w_1(\theta)E_\theta(1-\phi') + w_2(\theta)E_\theta\phi' \quad ,$$

$$r_k(\phi') = \sup_{\Omega} r_k(\theta,\phi') \ .$$

Let $U_k = \sum_{i=1}^{k} X_{1i}$ and $V_k = \sum_{i=1}^{k} X_{2i}$. Then U_k has the frequency function $f_\xi(u,k) = \binom{k}{u}\xi^u(1-\xi)^{k-u}$, and V_k has the frequency function $f_\eta(v,k) = \binom{k}{v}\eta^v(1-\eta)^{k-v}$. Let $X_k = (U_k, V_k)$. Then X_k has the frequency function $f_\theta(x,k) = f_\xi(u,k)\, f_\eta(v,k)$. Since X_k is a sufficient statistic for $\theta \in \Omega$, we may and shall confine ourselves to tests $\phi' = \phi'(X_k)$, depending only on X_k.

Our aim is to find a test ϕ, based on sample size k, satisfying

$$(2.1.5) \qquad r_k(\phi) = \inf_{\phi'} r_k(\theta,\phi') \quad \text{for every } k \ .$$

We will show that the test

$$\phi = \begin{Bmatrix} 1 & \text{if} & u > v \\ 1/2 & \text{if} & u = v \\ 0 & \text{if} & u < v \end{Bmatrix} ,$$

satisfies the condition (2.1.5) for every k whenever the loss functions satisfy conditions (2.1.2).

We first introduce a lemma.

Let X be a random variable with the generalized density $f_\theta(x)$

with respect to a fixed measure μ for $x \varepsilon \mathcal{X}$ and θ in some parameter space Θ. Let g and g^* be transformations of Θ onto Θ and $\mathcal{X}$ onto $\mathcal{X}$ respectively, with the properties:

(a) $g(g\theta) = \theta$, $g^*(g^*x) = x$,

(b) $f_{g\theta}(g^*x) = f_\theta(x)$ almost everywhere (μ) on $\mathcal{X}$ for all $\theta \varepsilon \Theta$,

(c) the loss functions have the property, $w_1(\theta) = w_2(g\theta)$ for all $\theta \varepsilon \Theta$,

(d) the measure μ is invariant under g^*.

Lemma (2.2)

If conditions (a) through (d) hold and ϕ' is any test, then there is a test ϕ_1 belonging to the class $T = \{\phi' : \phi'(x) = 1 - \phi'(g^*x)\}$, with the properties:

(e) $r(\phi_1) \leq r(\phi')$,

(f) $r(\theta,\phi_1) = r(g\theta,\phi_1)$ for all $\theta \varepsilon \Theta$, where $r(\theta,\phi')$ and $r(\phi')$ are defined by (2.1.4).

Proof of Lemma (2.2)

Let ϕ' be any test and define $\phi_1(x) = \dfrac{\phi'(x)+1-\phi'(g^*x)}{2}$. Then

$1 - \phi_1(g^*x) = 1 - \left[\dfrac{\phi'(g^*x)+1-\phi'(g^*g^*x)}{2}\right]$, and by assumption (a),

$1 - \phi_1(g^*x) = \dfrac{\phi'(x)+1-\phi(g^*x)}{2} = \phi_1(x)$. Hence $\phi_1 \varepsilon T$. Now

$$2r(\theta,\phi_1) = w_1(\theta)E_\theta\left[\phi'(g^*X)+1-\phi'(X)\right]+w_2(\theta)E_\theta\left[\phi'(X)+1-\phi'(g^*X)\right],$$

$$= r(\theta,\phi')+w_1(\theta)\int\phi'(g^*x)f_\theta(x)d\mu+w_2(\theta)\int\left[1-\phi'(g^*x)\right]f_\theta(x)d\mu .$$

By assumption (b),

$$2r(\theta,\phi_1)=r(\theta,\phi')+w_1(\theta)\int\phi'(g^*x)f_{g\theta}(g^*x)d\mu+w_2(\theta)\int[1-\phi'(g^*x)]f_{g\theta}(g^*x)d\mu .$$

By assumption (d),

$$2r(\theta,\phi_1)=r(\theta,\phi')+w_1(\theta)E_{g\theta}\phi'+w_2(\theta)E_{g\theta}(1-\phi') \quad .$$

By assumption (c),

$$2r(\theta,\phi_1)=r(\theta,\phi')+w_2(g\theta)E_{g\theta}\phi'+w_1(g\theta)E_{g\theta}(1-\phi')$$

$$=r(\theta,\phi')+r(g\theta,\phi') \ .$$

Now, $2r(g\theta,\phi_1)=r(g\theta,\phi')+r(gg\theta,\phi')=r(g\theta,\phi')+r(\theta,\phi')$, by assumption (a). Hence, $r(g\theta,\phi_1)=r(\theta,\phi_1)$, which proves assertion (f). Finally, $2\sup_{\Omega} r(\theta,\phi_1)=\sup_{\Omega}[r(\theta,\phi')+r(g\theta,\phi')] \leq 2\sup_{\Omega} r(\theta,\phi')$, which completes the proof.

In order to apply the lemma to our problem, let $g_1(\xi,\eta) = (1-\xi,1-\eta)$, $g_1^*(u,v) = (k-u,k-v)$, $g_2(\xi,\eta)=(\eta,\xi)$, and $g_2^*(u,v) = (v,u)$. Let $T_1 = \{\phi' : \phi'(u,v) = 1 - \phi'(k-u,k-v)\}$, $T_2 = \{\phi' : \phi'(u,v) = 1 - \phi'(v,u)\}$, and $T_0 = T_1 \cap T_2$.

<u>Theorem (2.3)</u>

For the binomial problem (2.1) with loss functions satisfying conditions (2.1.2), the class T_0 has the properties of the class T of Lemma (2.2). The test

$$\text{(a)} \qquad \phi = \left\{ \begin{array}{l} 1 \text{ if } u > v \\ 1/2 \text{ if } u = v \\ 0 \text{ if } u < v \end{array} \right\} ,$$

is the unique member of T_0 which minimizes the risk uniformly in

$\theta \in \Omega$. Moreover, $r_k(\phi) = \inf_{\phi'} r_k(\phi')$. These results hold for every k.

Proof of Theorem (2.3)

Since conditions (2.1.2) hold, and the conditions of lemma (2.2) are satisfied for both pairs of transformations (g_1, g_1^*) and (g_2, g_2^*), the classes T_1 and T_2 both possess the properties of the class T of the lemma. The same is therefore true of T_0.

Any member of T_2 has the property that $\phi'(u,u) = 1 - \phi'(u,u) = 1/2$ for $u = 0, 1, \ldots, k$. Hence, $\phi \in T_0$ implies

(2.3.1) $$\phi(u,u) = 1/2 \quad \text{for} \quad u = 0, 1, \ldots, k .$$

Clearly, $r_k[(\xi,\eta),\phi] = r_k[(1-\xi,1-\eta),\phi]$ for $\phi \in T_0$, and in minimizing the risk we may suppose that $\theta = (\xi, \xi+\varepsilon)$ for $0 \leq \xi \leq 1 - \varepsilon$, $0 \leq \varepsilon \leq 1$.

We want to choose $\phi \in T_0$ so as to minimize

$$r(\theta,\phi) = w_1(\theta)E_\theta(1-\phi) + w_2(\theta)E_\theta\phi$$

$$= \frac{w_1(\theta)+w_2(\theta)}{2} P_\theta\left\{U = V\right\} + \sum_{u \neq v} w_1(\theta)[1-\phi(u,v)]f_\theta(u,v)$$

$$+ \sum_{u \neq v} w_2(\theta)\phi(u,v)f_\theta(u,v)$$

$$= [(\frac{w_1(\theta)+w_2(\theta)}{2}) P_\theta(U = V) + w_1(\theta)P_\theta(U > V) + w_2(\theta)P_\theta(U < V)]$$

$$+ [\sum_{u>v} \phi(u,v) \left\{w_2(\theta)-w_1(\theta)\right\} \left\{ f_\theta(u,v)-f_\theta(k-u,k-v)\right\}].$$

The first term in square brackets is independent of ϕ, and we want to minimize the second expression in square brackets. Now $\theta = (\xi, \xi+\varepsilon)$ implies $\xi - \eta \leq 0$ and by (2.1.2.3) we have $w_2(\theta) - w_1(\theta) \geq 0$. Hence we can solve our problem by determining ϕ so as to minimize $S(\theta)$ uniformly for $\theta = (\xi, \xi+\varepsilon)$, where $S(\theta) = S_1(\theta) + S_2(\theta)$, and

$$S_1(\theta) = \sum_{\substack{u>v \\ u=k-v}} \phi(u,v) \left[f_\theta(u,v) - f_\theta(k-u,k-v) \right] ,$$

$$S_2(\theta) = \sum_{\substack{u>v \\ u\neq k-v}} \phi(u,v) \left[f_\theta(u,v) - f_\theta(k-u,k-v) \right] .$$

In $S_1(\theta)$ we have $u = v + c$, where $c > 0$, and the coefficient of $\phi(u,v)$ has the same sign as

$$\xi^v(1-\xi)^v(\xi+\varepsilon)^v(1-\xi-\varepsilon)^v \left[\xi^c(1-\xi-\varepsilon)^c - (1-\xi)^c(\xi+\varepsilon)^c \right] ,$$

which non-positive for all ξ and ε, and is strictly negative for $0 < \xi < 1-\varepsilon$, $0 < \varepsilon < 1$. For some ξ and ε satisfying the latter conditions we have $w_2(\theta) - w_1(\theta) > 0$. To minimize $S_1(\theta)$ and $\left[w_2(\theta) - w_1(\theta) \right] S_1(\theta)$ uniformly in $\theta = (\xi, \xi + \varepsilon)$, we must take

(2.3.2) $\qquad \phi(u,v) = 1$ whenever $u > v$ and $u = k - v$.

In order to minimize $S_2(\theta)$ we first write it in the form:

$$S_2(\theta) = \sum_{\substack{u>v \\ u<k-v}} \phi(u,v) \left[f_\theta(u,v) - f_\theta(k-u,\ k-v) \right] +$$
$$+ \sum_{\substack{u>v \\ u>k-v}} \phi(u,v) \left[f_\theta(u,v) - f_\theta(k-u,\ k-v) \right] .$$

The second sum in the last expression is equal to

$$\sum_{\substack{u>v\\u<k-v}} \phi(u,v) \left[f_\theta(k-v,\ k-u) - f_\theta(v,u) \right] ,$$

and

$$S_2(\theta) = \sum_{\substack{u>v\\u<k-v}} \phi(u,v) \left[f_\theta(u,v)-f_\theta(k-u,k-v)+f_\theta(k-v,k-u)-f_\theta(v,u)\right].$$

In $S_2(\theta)$ we have $u = v + c$ and $k - u = v + d$, where both c and d are positive. The coefficient of $\phi(u,v)$ in $S_2(\theta)$ has the same sign as

$$\left[\xi(1-\xi)(\xi+\varepsilon)(1-\xi-\varepsilon)\right]^v \left[(1-\xi)^d(1-\xi-\varepsilon)^d+\xi^d(\xi+\varepsilon)^d\right]\left[\xi^c(1-\xi-\varepsilon)^c-(1-\xi)^c(\xi+\varepsilon)^c\right].$$

This last expression is non-positive for all ξ and ε, and is strictly negative for $0 < \xi < 1 - \varepsilon$, $0 < \varepsilon < 1$. For some ξ and ε satisfying these last conditions we have $w_2(\theta) - w_1(\theta) > 0$. To minimize $S_2(\theta)$ and $\left[w_2(\theta) - w_1(\theta)\right] S_2(\theta)$ uniformly in $\theta = (\xi,\xi+\varepsilon)$, we must take

(2.3.3) $\qquad \phi(u,v) = 1$ whenever $u > v$ and $u < k - v$.

The test satisfying (2.3.1 - 2.3.3) and belonging to T_0 is clearly the test defined by (a) of the theorem, and is unique. That $r_k(\phi) = \inf_{\phi'} r_k(\phi')$ is obvious from the above arguments. The results obtained do not depend on k or δ. The proof is complete.

2.4 Points of Maximum Risk for the Best Test when the $w_i(\theta)$ are given by (2.1.3).

Here we may restrict ourselves to ω_1 where the risk is now simply $E_\theta\phi(X_k)$. Any θ in ω_1 is of the form $(\xi,\xi+\varepsilon)$ for $0 \le \xi \le 1-\varepsilon$ and $\delta \le \varepsilon \le 1$.

$$(2.4.1)\quad \frac{\partial}{\partial\varepsilon} E_{\xi,\xi+\varepsilon}\phi(X_k) = k\sum_{u,v}[\phi(u,v+1)-\phi(u,v)]f_\xi(u,k)f_{\xi+\varepsilon}(v,k-1) < 0\ ,$$

and the maximum risk occurs on the boundary of ω_1, that is, for $\theta = (\xi,\xi+\delta)$ where $0 \le \xi \le 1-\delta$.

$$\frac{\partial}{\partial\xi} E_{\xi,\xi+\delta}\phi(X_k)$$

$$= k\left[\sum_{u,v} D_1(u,v)f_\xi(u,k-1)f_{\xi+\delta}(v,k)\right] + k\left[\sum_{u,v} D_2(u,v)f_\xi(u,k)f_{\xi+\delta}(v,k-1)\right],$$

where

$$D_1(u,v) = \phi(u+1,v) - \phi(u,v) = \begin{cases} 1/2 \text{ for } u = v \text{ or } u = v-1 \\ 0 \quad \text{otherwise} \end{cases},$$

$$D_2(u,v) = \phi(u,v+1) - \phi(u,v) = \begin{cases} -1/2 \text{ for } u = v \text{ or } u = v+1 \\ 0 \quad \text{otherwise} \end{cases}.$$

We can factor out $k/2$ from every term, and the derivative is proportional to

$$(2.4.2)\quad P_\theta(U_{k-1} = V_k) + P_\theta(U_{k-1} = V_k - 1) - P_\theta(U_k = V_{k-1}) - P_\theta(U_k - 1 = V_{k-1})\ .$$

Let $U = U_{k-1}$, $V = V_{k-1}$, $W = X_{1k}$, and $Z = X_{2k}$. All these are independent, and omitting the subscript θ temporarily, we may write (2.4.2) in the form:

$$(2.4.3)\quad \left\{\begin{array}{l} P(U=V)P(Z=0) + P(U=V+1)P(Z=1) + P(U+1=V)P(Z=0) + P(U=V)P(Z=1) \\ - P(U=V)P(W=0) - P(U+1=V)P(W=1) - P(U=V)P(W=1) - P(U=V+1)P(W=0) \end{array}\right\}$$

$$= P(U=V+1)P(Z=1) + P(U+1=V)P(Z=0) - P(U+1=V)P(W=1) - P(U=V+1)P(W=0)\ .$$

Hence,

(2.4.4) $\frac{\partial}{\partial \xi} E_{\xi,\xi+\delta}\phi = (k/2)(1-2\xi-\delta)[P_\theta(U_{k-1}+1=V_{k-1})-P_\theta(U_{k-1}=V_{k-1}+1)]$,

which is negative for $\xi = 1 - \delta$ and $k > 1$. For $k = 1$, the derivative (2.4.4) vanishes identically. For $k > 1$, let $k' = k - 1$. Then the factor in square brackets in (2.4.4) is, for $0 \leq \xi < 1 - \delta$,

$$\left(\frac{\xi+\delta}{1-\xi-\delta} - \frac{\xi}{1-\xi}\right) \sum_j \binom{k'}{j}^2 \xi^j(1-\xi)^{k'-j}(\xi+\delta)^j(1-\xi-\delta)^{k'-j} > 0 .$$

Hence, the derivative (2.4.4) has the same sign as $(1-2\xi-\delta)$. By Theorem (2.3), $r_k[(\xi,\eta),\phi] = r_k[(\eta,\xi),\phi]$. We have proved the following theorem.

Theorem (2.4.5)

For $k > 1$ and $0 < \delta < 1$, $r_k(\theta,\phi)$ attains its maximum only at the points $\theta_1 = (\frac{1-\delta}{2}, \frac{1+\delta}{2}) \varepsilon\ \omega_1$ and $\theta_2 = (\frac{1+\delta}{2}, \frac{1-\delta}{2}) \varepsilon\ \omega_2$.

2.5 Admissibility and Almost-Uniqueness of the Best Test.

Let $P(\theta)$ assign probability $1/2$ to each of the points θ_1 and θ_2 as defined in (2.4.5). It is easy to show that the test ϕ given in (2.3) is bayes with respect to P, and can differ from any other bayes solution only on the set $\{(u,v): u = v\}$. Any minimax test must also be bayes with respect to P, and can differ from ϕ only on this same set.

To show that ϕ is admissible, we assume that ϕ is not admissible. Then there is a test ϕ_1 with $r_k(\theta,\phi) - r_k(\theta,\phi_1) \geq 0$ for all θ, and inequality is strict for some θ which, we may assume without loss of generality, lies in ω_1. The point of strict inequality is then $\theta = (\xi,\xi+\varepsilon)$, where $0 \leq \xi \leq 1 - \varepsilon$ and $\delta \leq \varepsilon \leq 1$. ϕ_1 differs from ϕ only on the set $\{(u,v): u=v\}$, and we have

$$E_\theta\phi - E_\theta\phi_1 = \sum_j \binom{k}{j}^2 a_j \xi^j(1-\xi)^{k-j}(\xi+\varepsilon)^j(1-\xi-\varepsilon)^{k-j} > 0 ,$$

where $a_j = \phi(j,j) - \phi_1(j,j)$, for $j = 0, 1, \ldots, k$. Let $\theta' = (\xi+\varepsilon,\xi) \in \omega_2$. Then,

$$r(\theta',\phi) - r(\theta',\phi_1) = E_{\theta'}(\phi_1-\phi)$$

$$= \sum_j \binom{k}{j}^2 a_j(\xi+\varepsilon)^j(1-\xi-\varepsilon)^{k-j}\xi^j(1-\xi)^{k-j} < 0,$$

contradicting our assumption. We have proved the following theorem, when the loss functions are given by (2.1.3).

Theorem (2.5.1)

The test ϕ is admissible for all k. For any k, the test ϕ can differ from any other minimax test only on the set $\{(u,v):u = v\}$.

2.6 Bounds to the Maximum Risk and Minimum Sample Size for the Best Test.

We again assume that the loss functions are given by (2.1.3). Let $r_k(\delta) = E_{\theta_1}\phi$, where θ_1 is given by (2.4.5), and the sample size is k. Using the notation of (2.4.1) and (2.4.2) we have,

$$(2.6.1)\quad r_{k+1}(\delta) = P_{\theta_1}(W=Z)E_{\theta_1}\phi(U_k,V_k)+P_{\theta_1}(W > Z)E_{\theta_1}\phi(U_k+1,V_k) + P_{\theta_1}(W < Z)E_{\theta_1}\phi(U_k,V_k+1) .$$

$$r_{k+1}(\delta) = r_k(\delta)+P_{\theta_1}(W=1,Z=0)E_{\theta_1}D_1(U_k,V_k) + P_{\theta_1}(W=0,Z=1)E_{\theta_1}D_2(U_k,V_k) .$$

$$r_{k+1}(\delta) = r_k(\delta)+\tfrac{1}{2}P_{\theta_1}(U_k=V_k)\left[P_{\theta_1}(Y=1,Z=0) - P_{\theta_1}(Y=0,Z=1)\right]$$

$$= r_k(\delta) + \frac{1}{2}\left[\left(\frac{1-\delta}{2}\right)^2 - \left(\frac{1+\delta}{2}\right)^2\right]P_{\theta_1}(U_k = V_k)$$

$$= r_k(\delta) - \frac{\delta}{2}\sum_j \binom{k}{j}^2 \left(\frac{1-\delta}{2}\right)^k \left(\frac{1+\delta}{2}\right)^k \quad .$$

$$\text{(2.6.2)} \qquad r_{k+1}(\delta) = r_k(\delta) - \frac{\delta}{2}\binom{2k}{k}\left(\frac{1-\delta^2}{2}\right)^k < r_k(\delta) \quad .$$

It is easy to show that $r_1(\delta) = (1-\delta)/2$, and we have,

$$\text{(2.6.3)} \qquad r_k(\delta) = (1-\delta)/2 - \frac{\delta}{2}\sum_{j=1}^{k-1}\binom{2j}{j}\left(\frac{1-\delta^2}{2}\right)^j = \frac{\delta}{2}\sum_{j=k}^{\infty}\binom{2j}{j}\left(\frac{1-\delta^2}{2}\right)^j \quad .$$

Define $\rho(x)$ by the relation: $x! = \sqrt{2\pi}\, x^{x+1/2} e^{-x+\rho(x)}$.

Then $\binom{2j}{j} = (\pi j)^{-1/2} 4^j e^{D(j)}$, where $D(j) = \rho(2j) - 2\rho(j)$. Uspensky [15], page 353, gives results which imply that $-1/8j < D(j) < 0$, for $j = 1, 2, \ldots,$. Let $a = -\log(1-\delta^2)$ and $f(x) = (\pi x)^{-1/2}e^{-ax}$. Then

$$\text{(2.6.4)} \qquad \frac{\delta}{2}(1-1/8k)\sum_{j=k}^{\infty} f(j) < r_k(\delta) < \frac{\delta}{2}\sum_{j=k}^{\infty} f(j) \quad .$$

Now $f'(x) < 0$ and $f''(x) > 0$ for $x > 0$. A straight line tangent to f at $x' > 0$ lies below f for all other positive x. Let $h(t) = f(x) + (t-x)f'(x) \le f(t)$. Inequality is strict for all $t \ne x$ where t and x are assumed positive.

$$\text{(2.6.5)} \qquad \int_x^{x+1} f(t)dt < f(x) = \int_{x-1/2}^{x+1/2} h(t)dt < \int_{x-1/2}^{x+1/2} f(t)dt \quad ,$$

for $x > 1/2$.

$$\frac{\delta}{2}(1-1/8k)\int_{k}^{\infty} f(t)dt < r_k(\delta) < \frac{\delta}{2}\int_{k-1/2}^{\infty} f(t)dt \quad . \tag{2.6.6}$$

Now $\int_{c}^{\infty} f(t)dt = \frac{2}{\sqrt{a}} \left[1 - \Phi(\sqrt{2ac}) \right]$ for $c > 0$. Let

$$L(\delta,k) = (1 - 1/8k)\frac{\delta}{\sqrt{a}}\left\{1 - \Phi(\sqrt{2ak})\right\} ,$$

$$U(\delta,k) = \frac{\delta}{\sqrt{a}}\left\{1 - \Phi(\sqrt{2ak-a})\right. \quad .$$

We have proved the following theorem.

Theorem (2.6.7)

For all $k = 1, 2, \ldots$, and all $0 < \delta < 1$,

$$L(\delta,k) < r_k(\delta) < U(\delta,k) \; .$$

Simpler but cruder bounds may be obtained by using the fact that $\delta\sqrt{1-\delta^2/2} < \sqrt{a} < \delta$.

We define $N(\delta)$ as the least integer k such that $r_k(\delta) \leq \alpha$, where $0 < \alpha < 1/2$. Let $N'(\delta) = N(\delta) - 1$ and

$k_1(\delta)$ = largest integer k such that $L(\delta,k) \geq \alpha$,

$k_2(\delta)$ = least integer k such that $U(\delta,k) \leq \alpha$.

The relation (2.6.2) shows that $r_k(\delta)$ is strictly monotone decreasing in k. Hence, $N'(\delta)$ is the largest integer k such that $r_k(\delta) > \alpha$. Now $\alpha \leq L(\delta,k_1) < r_{k_1}(\delta)$ implies $k_1 \leq N'$, and $\alpha \geq U(\delta,k_2) > r_{k_2}(\delta)$ implies $k_2 \geq N$. We have proved the following theorem.

Theorem (2.6.8)

For every $0 < \delta < 1$ we have

$$k_1(\delta) + 1 \leq N(\delta) \leq k_2(\delta) \quad .$$

2.7 Comparison of a Bayes Test with the One-Sided Fisher-Tocher Test when the Sample Sizes are Equal and $\alpha < 1/2$.

The bayes test discussed in the preceding sections is a member of the class of tests defined by

$$(2.7.1) \qquad \phi = \left\{ \begin{array}{ll} 1 & \text{if } u - v > d \\ b & \text{if } u - v = d \\ 0 & \text{otherwise} \end{array} \right\} \quad .$$

Our previous results suggest that such tests might have some optimum properties. In particular, suppose that the sample size is k, and that ϕ is of level α for testing $\theta_o = (1/2, 1/2)$. Then d and b are defined by the relations:

$$(2.7.2) \qquad g(j) = \binom{2k}{k+j} 4^{-k} \quad ,$$

$$\sum_{j>d} g(j) \leq \alpha < \sum_{j \geq d} g(j) \quad ,$$

$$\sum_{j>d} g(j) + bg(d) = \alpha \ .$$

We will show that if $\alpha < (1/2)(1 - g(0))$, then ϕ is of level α for testing $\omega_1 = \{\theta: \xi - \eta \leq 0\}$. Then it seems reasonable to compare ϕ with the Fisher-Tocher test, ψ, which is most powerful unbiased [14] for testing ω_1 against $\omega_2 = \{\theta: \xi - \eta > 0\}$. ψ is defined by the relations:

$$(2.7.3) \qquad \psi = \left\{ \begin{array}{ll} 1 & \text{if } u - v > c(s) \\ a(s) & \text{if } u - v = c(s) \\ 0 & \text{otherwise} \end{array} \right\}, \text{ where } s = u + v \ ,$$

$$h(j,s) = \binom{k}{j}\binom{k}{s-j}/\binom{2k}{s} \quad ,$$

$$\sum_{j>c(s)} h(j,s) \leqq \alpha < \sum_{j \geqq c(s)} h(j,s) \quad ,$$

$$\sum_{j>c(s)} h(j,s) + a(s)h(c(s),s) = \alpha$$

One way to compare the two tests is to compare their power functions. As we will show, ϕ suffers by such a comparison because $E_\theta \phi < \alpha$ in ω_1, except at a single point where equality holds, while $E_\theta \psi = \alpha$ everywhere in $L = \{\theta : \xi = \eta\}$. In three of the special cases treated below we show that $E_\theta(\psi - \phi) > 0$ everywhere in ω_1, except at the points $(1/2, 1/2)$ and $(0,1)$, where the expected value vanishes.

If we say that a test ϕ_1 is "better" at θ than another test ϕ_2, we will mean that $E_\theta(\phi_1 - \phi_2) > 0$ if $\theta \varepsilon \omega_2$, and $E_\theta(\phi_1 - \phi_2) < 0$ if $\theta \varepsilon \omega_1$. Roughly speaking, we will show that ϕ is "better" than ψ if θ is "close" to $L^* = \{\theta : \xi = 1 - \eta\}$ or θ is "close" to L and in ω_1, and ψ is "better" than ϕ if θ is "close" to L and in ω_2. If θ is "close" to L and also "close" to L^*, then θ is close to $(1/2, 1/2)$ where $E_\theta(\phi - \psi)$ is close to zero, and one test is about as good as the other.

We will define a class of contours such that any point inside $\bar{\Gamma}$ is on one of the contours of the class, and show that $E_\theta(\phi - \psi)$ can vanish at most once on any such contour. By studying special cases we will obtain some idea of how large a difference in power may be expected in certain subsets of Ω.

We begin with two lemmas.

Lemma (2.7.4)

Let W be a discrete random variable assuming with positive probabilities only the values $0, \pm 1, \pm 2, \ldots, \pm k$, and such that

(a) $P(W = j) = P(W = -j) > 0$ for $j = 0, 1, \ldots, k$,

(b) $0 < j \leq k$ implies $P(W = j - 1) > P(W = j)$.

Let Z be another discrete random variable independent of W which assumes with positive probabilities only the values $0, \pm 1$, and such that

(c) $P(Z = 0) \geq 1/2$

(d) $P(Z = -1) = p(Z = 1) > 0$.

Then $T = (W + Z)$ assumes with positive probabilities only the values $0, \pm 1, \ldots, \pm (k + 1)$ and

(e) $P(T = i) = P(T = -i) > 0$ for $i = 0, 1, \ldots, (k + 1)$,

(f) $0 < i \leq (k + 1)$ implies $P(T = i - 1) > P(T = i)$.

Proof of Lemma (2.7.4)

To prove statement (e) we note that W has the same distribution as $-W$ and Z has the same distribution as $-Z$. Since W and Z are independent, $(W + Z)$ has the same distribution as $(-W - Z)$.

To prove statement (f), we first assume that $j > 0$ and compare $P(T = j)$ with $P(T = j+ 1)$.

$$P(T=j) = P(W=j+1)P(Z=-1) + P(W+j)P(Z=0) + P(W=j-1)P(Z=1) \quad .$$

$$P(T=j+1) = P(W=j+2)P(Z=-1) + P(W=j+1)P(Z=0) + P(W=j)P(Z=1) \quad .$$

Applying (b) to each pair of terms in the above expressions we conclude that

$$1 < i \leq k + 1 \quad \text{implies} \quad P(T=i-1) > P(T=i) > 0 \ .$$

Now we want to show that (f) holds when $i = 1$, that is, to show that $g = P(T=0) - P(T=1) > 0$.

$$g = hP(Z=1) + P(Z=0)[P(W=0) - P(W=1)] \ ,$$

where $h = 2P(W=1) - P(W=2) - P(W=0) > P(W=1) - P(W=0)$. Hence,

$$g \geq [P(Z=0) - P(Z=1)][P(W=0) - P(W=1)]$$

$$= \left[\frac{3P(Z=0)-1}{2}\right][P(W=0) - P(W=1)] > 0 \ ,$$

which completes the proof.

<u>Lemma (2.7.5)</u>

Let c be any non-negative integer and

$$\phi_c = \phi_c(u,v) = \begin{Bmatrix} 1 & \text{if} \quad u - v > c \\ 0 & \text{otherwise} \end{Bmatrix} \ .$$

Let $H(\xi) = E_{\xi\xi}\phi_c(U_k,V_k)$ where k is fixed. Then $H(\xi)$ has a unique maximum for $\xi \in (0, 1)$ at the point $\xi = 1/2$, or else $H(\xi) = 0$ for all $\xi \in (0, 1)$.

<u>Proof of Lemma (2.7.5)</u>

Let $U = U_{k-1}$, $V = V_{k-1}$, $X = X_{1k}$ and $Y = X_{2k}$. The four variables are mutually independent, and it is easy to show that

$$H'(\xi) = k[P\{U - (V + Y) = c\} - P\{U + X - V = c + 1\} \ ,$$

$$H'(\xi) = k(1 - 2\xi)[P(U - V = c) - P(U - V = c + 1)] \ ,$$

(2.7.5.1) $$H'(\xi) = k(1 - 2\xi)[P(T = c) - P(T = c + 1)] \ ,$$

where $T = U - V = \sum_{i=1}^{k-1} Z_i$, and

$$Z_i = X_{1i} - X_{2i} = \left\{ \begin{array}{l} -1 \text{ with probability } \xi(1-\xi) \\ \;\;0 \text{ with probability } \xi^2+(1-\xi)^2 \\ \;\;1 \text{ with probability } \xi(1-\xi) \end{array} \right\} .$$

The Z's are independent, and finite induction using lemma (2.7.4) discloses that the factor $[P(T = c) - P(T = c + 1)]$ in (2.7.5.1) is strictly positive for $c \leq k - 1$, and vanishes identically for $c \geq k$. In the latter case it is clear from the definition that $H(\xi)$ vanishes identically. For $c \leq k - 1$ the derivative has the same sign as $(1 - 2\xi)$ and vanishes uniquely at $\xi = 1/2$, which is a true maximum point. This completes the proof.

(2.7.6) ϕ and $E_\theta(\psi - \phi)$ in Relation to a Class of Contours.

(2.7.6.1) Definitions:

$$\omega(\xi_0) = \left\{ \theta:\ \left(\frac{1-\xi}{\xi}\right)\left(\frac{1-\eta}{\eta}\right) = \left(\frac{1-\xi_0}{\xi_0}\right)^2 ;\ 0 < \xi_0, \xi, \eta < 1 \right\} ,$$

$$\omega(0) = \left\{ \theta:\ 0 < \xi < 1,\ \eta = 0 \right\} ,$$

$$\omega(1) = \left\{ \theta:\ \xi = 1,\ 0 < \eta < 1 \right\} .$$

For any two distinct points, $\theta_1 = (\xi_1, \eta_1)$ and $\theta_2 = (\xi_2, \eta_2)$, both in $\omega(\xi_0)$, we say $\theta_1 < \theta_2$ if and only if $\xi_1 < \xi_2$.

The test ϕ defined in (2.7.1) may be written in the form

$$\phi = b\phi_{d-1} + (1-b)\,\phi_d ,$$

where the two tests on the right are defined by Lemma (2.7.5). We will assume in what follows that $\alpha < (1/2)\,[1 - \binom{2k}{k}4^{-k}]$, which tends to

1/2 as k becomes infinite. This implies that $d \geq 1$. In particular $d \geq 1$ if $\alpha < 1/4$ or if $k \geq \dfrac{1}{\pi(1-2\alpha)^2}$. Then we have from lemma (2.7.5) ,

(2.7.6.2) $$E_{\xi,\xi}\phi \text{ has a unique maximum at } \xi = 1/2 \ .$$

It follows directly from (2.4.1) that ϕ is of level α for testing $\omega_1 = \left\{\theta\colon \xi - \eta \leq 0\right\}$.

For any two points θ_1 and θ_2 of $\omega(\xi_0)$, where $\theta_1 < \theta_2$ and $0 \leq \xi_0 \leq 1$, it is easy to show that ϕ is a most powerful test of level $E_{\theta_1}\phi$ for testing θ_1 against θ_2. It is clear that ϕ differs from ψ at both of the values $(u,v) = (0, 0)$, and $(u,v) = (k, k)$. Using the Neyman-Pearson lemma as formulated by Lehmann [10], we obtain $E_{\theta_1} < E_{\theta_2}\phi$ and

(2.7.6.3) $$E_{\theta_1}\psi \leq E_{\theta_1}\phi \text{ implies } E_{\theta_2}\psi < E_{\theta_2}\phi \ .$$

Therefore $E_\theta(\psi - \phi)$ can vanish at most once in $\omega(\xi_0)$. Since $\psi(0,0) = \alpha$ and $\phi(0, 0) = 0$ we also have, because of symmetry, $E_\theta(\psi - \phi) > 0$ for $\theta \in \left\{\theta\colon \xi = 0;\ 0 \leq \eta < 1\right\}$ or $\theta \in \left\{\theta\colon 0 < \xi \leq 1;\ \eta = 1\right\}$.

From the above arguments we conclude that there is a region containing the set $\left\{\theta\colon \xi = 1 - \eta,\ 1/2 < \xi < 1\right\}$, throughout which $E_\theta\phi > E_\theta\psi$, and ϕ has higher power than ψ . There is a region containing the set $\left\{\theta\colon \xi = 1 - \eta,\ 0 < \xi < 1/2\right\}$, within which ϕ is "better" than

Ψ, because it has smaller first kind of error. Finally, there is a region, R, containing the set $\{\theta: \xi = \eta; \xi \neq 1/2\}$ such that $E_\theta \Psi > E_\theta \phi$ in this region, and therefore Ψ is more powerful than ϕ if $\theta \varepsilon R$ and $\xi > \eta$, but ϕ is "better" than Ψ if $\theta \varepsilon R$ and $\xi \leq \eta$, because then ϕ has smaller first kind of error. In the special cases studied below (2.7.8) we will graph the contours $E_\theta(\Psi - \phi) = 0$, and obtain some idea of the shapes of these various regions.

2.7.7 <u>Some Local Properties of</u> $E_\theta \phi$

For ϕ_c defined by lemma (2.7.5) and $c \geq 0$, let $G(\eta) = E_\theta \phi_c(X_k)$, where $\theta = (\eta + \delta, \eta)$, $0 \leq \eta \leq 1 - \delta$ and $0 < \delta < 1$. A little manipulation similar to that carried out in section (2.4) leads to

$$G'(\eta) = k(1 - 2\eta - \delta)\left[P_\theta(U_{k-1} - V_{k-1} = c) - P_\theta(U_{k-1} - V_{k-1} = c + 1)\right] .$$

If we evaluate the expression in square brackets at $\eta_\delta = \frac{1-\delta}{2}$, we obtain

$$\binom{2k}{k+c}\left(\frac{1+\delta}{2}\right)^{k+c}\left(\frac{1-\delta}{2}\right)^{k-c} - \binom{2k}{k+c+1}\left(\frac{1+\delta}{2}\right)^{k+c+1}\left(\frac{1-\delta}{2}\right)^{k-c-1} ,$$

which is negative, zero, or positive according as $\left(\frac{2c+1}{2k+1} - \delta\right)$ is negative, zero or positive. Since $G'(\eta_\delta) = 0$, it follows that $G(\eta)$ has a local maximum at η_δ for small δ and a local minimum at η_δ for large δ.

Now $\phi = b\phi_{d-1} + (1-b)\phi_d$, where $d-1 \geq 0$, and we conclude that for fixed δ and $\theta = (\eta + \delta, \eta)$, $E_\theta \phi$ has a local maximum at $\eta = \eta_\delta$ for small values of δ, and a local minimum at $\eta = \eta_\delta$ for large values of δ.

2.8 $E_\theta(\phi - \psi)$ for Selected Small Sample Sizes and $\alpha = .05$.

Let $\theta_\Delta = (\frac{1}{2} + \Delta, \frac{1}{2} - \Delta)$, where $0 \leq \Delta \leq \frac{1}{2}$. In the special cases ($k = 1, 3, 6$ and 10) below, we will find the approximate value of $S = \sup_{0 \leq \Delta \leq 1/2} E_{\theta_\Delta}(\phi - \psi)$, and approximate the unique value Δ_0 at which it occurs. In each case we will also determine a few points on the contour $E_\theta(\phi - \psi) = 0$, and present a graph.

We define:

$$\phi_j = \sum_{u-v=j} \binom{k}{u}\binom{k}{v}\phi(u,v) \quad ,$$

$$\psi_j = \sum_{u-v=j} \binom{k}{u}\binom{k}{v}\psi(u,v) \quad ,$$

$$D_j = \phi_j - \psi_j \quad ,$$

$$E_j = \phi(j,0) - \psi(j,0),$$

$$y = \frac{1-2\Delta}{1+2\Delta} \quad .$$

Then,

$$E_{\theta_\Delta}(\phi - \psi) = (\frac{1}{4} - \Delta^2)^k \left\{ \sum_{j=0}^{k} D_j (\frac{1+2\Delta}{1-2\Delta})^j \right\}, \text{ for } \Delta \neq 1/2 \quad .$$

For $0 \leq \Delta < 1/2$, we may write $E_{\theta_\Delta}(\phi - \psi)$ in the form,

$$g(y) = E_{\theta_\Delta}(\phi - \psi) = \left\{ y(1+y)^{-2} \right\}^k \left\{ \sum_{j=0}^{m} D_j y^{-j} \right\} ,$$

where m is the largest integer j such that $D_j \neq 0$. If $E_{\theta_\Delta}(\phi - \psi)$ has its maximum for Δ in $(0, 1/2)$, then $g'(y)$ vanishes at

$y_0 \in (0, 1)$, and the required maximum is $g(y_0)$, provided y_0 is unique. $g'(y)$ vanishes on $(0,1)$ if and only if

$$f(y) = \sum_{j=0}^{m+1} [(k-m+j) D_{m-j} - (k-m+1-j) D_{m+1-j}] y^j = 0 \quad ,$$

where we define $D_{-1} = D_{m+1} = 0$. We may also write,

$$(2.8.1) \qquad f(y) = A_0 + A_1 y + \ldots + A_{m+1} y^{m+1} = 0$$

In the three examples ($k = 3$, 6 and 10) for which the maximum of $E_{\theta_\Delta}(\phi - \psi)$ occurs for Δ in $(0, \frac{1}{2})$, we will show that the coefficients in the polynomial (2.8.1) satisfy the conditions:

$$(2.8.2) \qquad \sum_i A_i < 0, \quad A_i > 0 \text{ if } i \neq 1 \quad .$$

These conditions imply $f(0) > 0$, $f(1) < 0$, and $f''(y) > 0$, for all y on $[0, 1]$. Hence, there is a unique y_0 on $[0, 1]$ such that $f(y_0) = 0$. Hence $\Delta_0 = \frac{1-y_0}{2+2y_0}$.

Let $\eta(\xi)$ be the value of $\eta > 1 - \xi$ for which $E_{\xi,\eta}(\phi - \psi) = 0$. In the cases $k = 1$, 3, and 6, we have: $\eta(\xi)$ exists and is unique for $1/2 \leq \xi \leq 1$. The results of our investigations of the special cases are tabulated below. A sketch of the computations will be found in the appendix.

(2.8.3)

TABLE OF RESULTS IN SPECIAL CASES

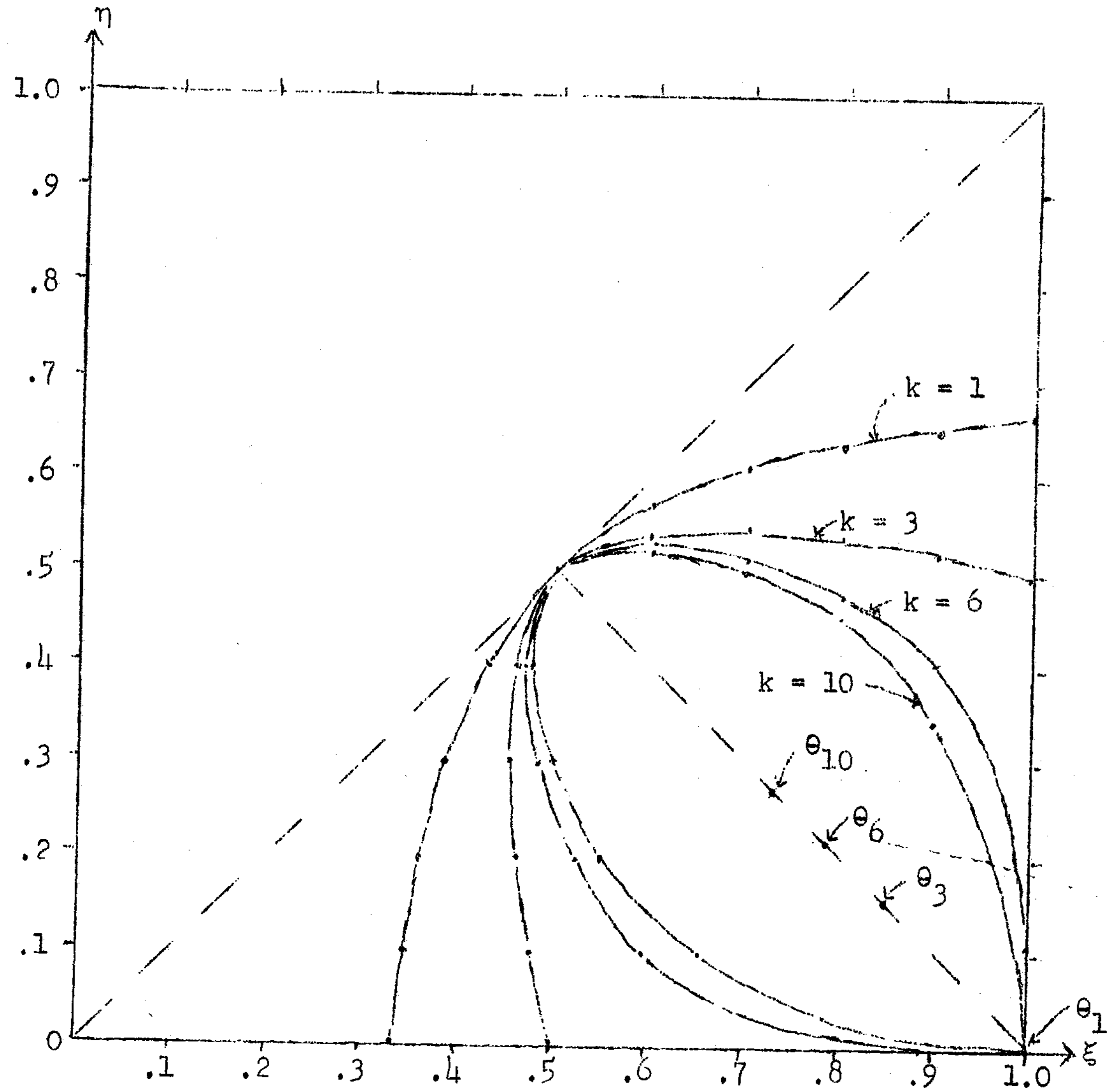
k	1	3	6	10
Δ_0	.500	.345	.284	.230
S	.100	.040	.030	.028
$\eta(.5)$	.500	.500	.500	.500*
$\eta(.6)$	.571	.540	.530	.525*
$\eta(.7)$	.611	.543	.513	.504*
$\eta(.8)$	.636	.534	.475	.454*
$\eta(.9)$	.654	.520	.405	.347*
$\eta(.10)$	.667	.500	.112	.000*

* These values have not been proved unique.

On the following page we present the above material in the form of a graph.

Graph of $E_\theta(\phi - \psi) = 0$

$\theta_k = (1/2+\Delta_0, 1/2-\Delta_0)$ for Sample Size k



CHAPTER III

ASYMPTOTIC SOLUTIONS AS THE DISTANCE BETWEEN THE ALTERNATIVES TENDS TO ZERO

Now we want to investigate the case where the sets ω_1 and ω_2 depend on some positive quantity δ. Suppose we have a continuous function $h(\theta)$ defined for all θ in the closed Euclidean parameter space Ω. Our theorems will apply to three types of problems.

Type I. $\omega_1(\delta) = \{\theta: h(\theta) \leqq -\delta\}$; $\omega_2(\delta) = \{\theta: h(\theta) \geqq \delta\}$.

Type II. $\omega_1(\delta) = \omega_1 = \{\theta: h(\theta) \leqq 0\}$; $\omega_2(\delta) = \{\theta: h(\theta) \geqq \delta\}$.

Type III. $\omega_1(\delta) = \omega_1 = \{\theta: h(\theta) = 0\}$; $\omega_2(\delta) = \{\theta: h(\theta) \geq \delta\}$.

Throughout we will assume that δ is confined to some set $(0 < \delta \leqq \delta_o$, for instance) so that neither of the sets ω_i is empty. We will judge a test ϕ', based on a sample of size k, on the basis of its maximum risk, $r_k(\phi',\delta)$, as defined in (1.2) of Chapter I, where the additional argument δ is included to reflect the fact that the sets $\omega_i = \omega_i(\delta)$ may depend on δ $(i = 1, 2)$.

We will derive some theorems which will rest on the following assumptions expressed in the terminology of Chapter I.

Assumption (a): For every k and every μ_k-measurable set A of $\mathcal{X}_k$ we have

$$P_\theta(A) = \int_A f_\theta(x,k)d\mu_k \text{ is continuous in } \theta \text{ for all } \theta \text{ in } \Omega .$$

Assumption (b). $0 < \alpha$; $0 < \beta$; $\alpha + \beta < 1$ (end of assumption (b)).

Let $Z_k = z_k(X_k)$ be a random variable defined for every k and $x_k \in \mathcal{X}_k$.

Let $M_{1k}(x,\delta) = \inf_{\theta \in \omega_1(\delta)} P_\theta(Z_k \leq x)$; $M_{2k}(x,\delta) = \sup_{\theta \in \omega_2(\delta)} P_\theta(Z_k \leq x)$.

Assumption (c). There exists a constant $r'(0 < r' < \infty)$ and functions $H_1(x,d)$, $H_2(x,d)$, such that for every $d > 0$,

$M_{1k}(x,dk^{-r'}) \longrightarrow H_1(x,d)$ as $k \longrightarrow \infty$ at every continuity point of $H_1(x,d)$,

$M_{2k}(x,dk^{-r'}) \longrightarrow H_2(x,d)$ as $k \longrightarrow \infty$ at every continuity point of $H_2(x,d)$.

Note that M_{1k} and H_1 are non-decreasing in d, M_{2k} and H_2 are non-increasing in d for problems of types I, II and III. Also M_{1k}, M_{2k} are non-decreasing in x.

Assumption (d). The simultaneous equations, $H_1(x,d) = 1-\alpha, H_2(x,d) = \beta$, have the unique solution $(x,d) = (c,D)$ (both finite). The functions H_1 and H_2 are continuous in (x,d) in some neighborhood of (c,D), and in particular continuous in d at $x = c$ in some neighborhood $\mathcal{D}$ of D. (End of assumption (d)).

Let $P_{k,\delta}(\theta)$ be an a priori distribution which assigns probability one to a finite set in Θ. More specifically let $P_{k,\delta}$ assign probability $p_{ij}(k,\delta)$ to $\theta_{ij} = \theta_{ij}(k,\delta) \in \omega_i(\delta)$; $i = 1, 2$; $j = 1, 2, \ldots, R_i$.

Let $\quad p(k,\delta) = \sum_{j=1}^{R_1} p_{1j}(k,\delta) > 0$; $\quad 1 - p(k,\delta) = \sum_{j=1}^{R_2} p_{2j}(k,\delta) > 0$.

Let $\phi = \phi(Z_k)$ be defined by

$$\phi(z) = \begin{cases} 1 \text{ if } z > c \\ 0 \text{ if } z \leq c \end{cases}, \quad \text{where } c \text{ is defined by assumption (d).}$$

<u>Assumption (e)</u>. For every (k,δ) the test ϕ is bayes with respect to $P_{k,\delta}(\theta)$, where for any test ϕ' based on a sample of size k the risk function is

$$r_k(\theta,\phi',\delta) = \begin{cases} aE_\theta\phi' & \text{for } \theta\varepsilon\, \omega_1(\delta); \quad a = \dfrac{\beta}{\alpha+\beta} \\ bE_\theta(1-\phi') & \text{for } \theta\varepsilon\, \omega_2(\delta); \quad b = \dfrac{\alpha}{\alpha+\beta} \\ 0 \quad \text{otherwise} \end{cases}.$$

<u>Assumption (f)</u>. For every k, $\theta_{1i}(k,\delta) \longrightarrow \theta_k$ and $\theta_{2j}(k,\delta) \longrightarrow \theta_k \varepsilon\, \Omega$ as $\delta \longrightarrow 0$ for $i = 1, \ldots, R_1$; $j = 1, \ldots, R_2$.

For every k, $p(k,\delta) \longrightarrow b = \alpha/(\alpha+\beta)$ as $\delta \longrightarrow 0$.

<u>Assumption (g')</u>. For every pair of positive sequences $\{\delta_\lambda\}$ and $\{k_\lambda\}$ such that $k_\lambda \longrightarrow \infty$ and $\delta_\lambda k_\lambda^{r'} \longrightarrow D' < D$ as $\lambda \longrightarrow \infty$ we have,

Every limit point of $r_{k_\lambda}(P_{k_\lambda,\delta_\lambda}, \phi, \delta_\lambda)$ is greater than $a\alpha$.

<u>Definitions.</u>

$$r_k(\phi',\delta) = \sup_{\Omega} r_k(\theta, \phi', \delta),$$

$$r_k(\delta) = \inf_{\phi'} \sup_{\Omega} r_k(\theta, \phi', \delta),$$

$N(\delta)$ = least integer k such that $r_k(\delta) \leq a\alpha = b\beta = \alpha\beta/(\alpha+\beta)$,

$$r_k(P,\emptyset',\delta) = \int_{\Omega} r_k(\theta,\emptyset',\delta)dP(\theta) \quad \text{for any a priori distribution } P(\theta);$$

$N_p(\delta)$ = least integer k such that $r_k(P_{k,\delta}, \emptyset, \delta) \leqq a\alpha$,

$N_c(\delta)$ = least integer k such that $r_k(\emptyset, \delta) \leq a\alpha$.

Lemma (3.1)

If assumptions (a), (b), and (f) hold, then $N_p(\delta) \longrightarrow \infty$ as $\delta \longrightarrow 0$.

Proof of Lemma (3.1)

Suppose the lemma is false. Then there exists a sequence $\{\delta_\lambda\}$ such that $\delta_\lambda \longrightarrow 0$ as $\lambda \longrightarrow \infty$, and for $k_\lambda = N_p(\delta_\lambda)$, we have $k_\lambda \longrightarrow k$ (finite) as $\lambda \longrightarrow \infty$. Since k_λ takes on only integral values we have

(3.1.1) $k_\lambda = k$ for $\lambda > \lambda_0$, i. e. for λ sufficiently large.

$$r_{k_\lambda}(P_{k_\lambda,\delta_\lambda},\emptyset,\delta_\lambda) = r_k(P_{k,\delta_\lambda},\emptyset,\delta_\lambda)$$

$$= \sum_i p_{1i}(k,\delta_\lambda)aE_{\theta_{1i}(k,\delta_\lambda)}\emptyset + \sum_j p_{2j}(k,\delta_\lambda)bE_{\theta_{2j}(k,\delta_\lambda)}(1-\emptyset)$$

$$\leqq a\alpha \quad \text{for } \lambda > \lambda_0$$

Now by assumption (f), $\theta_{1i}(k,\delta_\lambda)$ and $\theta_{2j}(k,\delta_\lambda)$ tend to θ_k for $i = 1, \ldots, R_1$; $j = 1, \ldots, R_2$, as $\lambda \longrightarrow \infty$. Taking the limit as $\lambda \longrightarrow \infty$ in the above inequality, and appealing to the continuity assumption (a) we obtain

(3.1.2) $$(aE_{\theta_k}\emptyset)(\lim_{\lambda \to \infty} \sum_i p_{1i}(k,\delta_\lambda)) + (bE_{\theta_k}(1-\emptyset))(\lim_{\lambda \to \infty} \sum_j p_{2j}(k,\delta_\lambda)) \leqq a\alpha.$$

The two sums are $p_1(k,\delta_\lambda)$ and $1 - p_1(k,\delta_\lambda)$, respectively, and their limits are b and $1-b = a$, by assumption (f). We obtain the relation $abE_{\theta_k}\phi + abE_{\theta_k}(1-\phi) = ab \leq a\alpha$, contradicting assumption (b).

Lemma (3.2)

If assumptions (c) and (d) hold, then $N_c(\delta)$ is finite for every $\delta > 0$.

Proof of Lemma (3.2)

If the lemma were false there would exist $\delta > 0$ such that:

(3.2.1) $$r_k(\phi,\delta) > a\alpha \quad \text{for all } k .$$

Here, we know that $M_{1k}(x,\delta)$ is non-decreasing in δ; $M_{2k}(x,\delta)$ is non-increasing in δ. Hence, $H_1(c,d)$ is increasing in d at $d = D$, by the uniqueness assumption (d). There exists $d \in \Omega$, $d > D$, such that

(3.2.2) $$H_1(c,d) > H_1(c,D) = 1-\alpha \text{ . Similarly, } H_2(c,d) < \beta .$$

For k sufficiently large we have $dk^{-r'} < \delta$, and because $\omega_i(\delta)$ are non-expanding in δ $(i = 1,2)$,

(3.2.3) $$r_k(\phi,\delta) \leq r_k(\phi,dk^{-r'}) = \max\left\{ a[1-M_{1k}(c,dk^{-r'})], bM_{2k}(c,dk^{-r'}) \right\},$$

for k sufficiently large.

By assumption (c) the last expression on the right converges to $\max\left\{ a[1-H_1(c,d)], bH_2(c,d) \right\}$, which by (3.2.2) is less than $\max(a\alpha, b\beta)$. Hence, $\lim_{k \to \infty} r_k(\phi,dk^{-r'}) < \max(a\alpha,b\beta)$. Taking limits on both sides of (3.2.3) and applying (3.2.1), we obtain the

result, $\max(a\alpha, b\beta) > \lim_{k \to \infty} r_k(\phi, \delta) \geq a\alpha(=b\beta)$, which is impossible.

Lemma (3.3)

If assumption (e) holds we have: $N_p(\delta) \leq N(\delta) \leq N_c(\delta)$ for all $\delta > 0$.

Proof of Lemma (3.3)

$r_k(P_{k,\delta}, \phi, \delta) \leq r_k(P_{k,\delta}, \phi', \delta) \leq r_k(\phi', \delta)$ for all ϕ' by assumption (e) and a property of a priori distributions. Hence

$$r_k(P_{k,\delta}\phi, \delta) \leq \inf_{\phi'} r_k(\phi', \delta) = r_k(\delta) \leq r_k(\phi, \delta),$$

$$N_p(\delta) \leq N(\delta) \leq N_c(\delta) \text{ for all } \delta > 0 .$$

Theorem (3.4A)

If assumptions (a) through (f) and assumption (g') hold, then asymptotically as $\delta \to 0$,

(1) $N(\delta) \sim (D/\delta)^{1/r'}$, where $D = D(\alpha,\beta)$ defined by assumption (d),

(2) The test ϕ is asymptotically efficient, that is, $N_c(\delta) \sim N(\delta)$.

Proof of Theorem (3.4A)

We have to show that $\delta N^{r'} \to D$ as $\delta \to 0$, and $N_c/N \to 1$ as $\delta \to 0$, where $N = N(\delta)$; $N_c = N_c(\delta)$; $N_p = N_p(\delta)$.

Assumption (e) implies that Lemma (3.3) holds, and we can prove both assertions of the theorem by showing two things, as $\delta \to 0$,

(3.4.1) Every limit point of $\delta N_c^{r'}$ is less than or equal to D,

(3.4.2) Every limit point of $\delta N_p^{r'}$ is greater than or equal to D.

We first prove (3.4.1). For every positive sequence $\{\delta_\lambda\}$ converging to zero, $n_\lambda+1 = N_c(\delta_\lambda)$ is finite by lemma (3.2) for all λ. Now suppose (3.4.1) is false. Then for some positive sequence $\{\delta_\lambda\}$ converging to zero as $\lambda \longrightarrow \infty$, we have $\delta_\lambda(n_\lambda+1)^{r'} \longrightarrow D'$ and therefore

$$(3.4.3) \qquad \delta_\lambda n_\lambda^{r'} \longrightarrow D' > D \text{ as } \lambda \longrightarrow \infty .$$

Choose D'' such that $D < D'' < D'$; $D'' \varepsilon \mathcal{D}$ (which can be done by assumption (d)). For λ sufficiently large, $\delta_\lambda n_\lambda^{r'} > D''$. The monotonicity of $\omega_i(\delta)$ in $\delta (i = 1, 2)$, and the definition of n_λ yield

$$(3.4.4) \qquad a\alpha < r_{n_\lambda}(\emptyset,\delta_\lambda) \leq r_{n_\lambda}(\emptyset,D''n_\lambda^{-r'}) \quad \text{for } \lambda \text{ sufficiently large.}$$

By lemmas (3.1) and (3.3), n_λ becomes infinite with λ. By assumption (c) and equation (3.2.3), the right-hand side of (3.4.4) converges to max $\left\{a(1-H_1(c,D'')),bH_2(c,D'')\right\}$, which, by the monotonicity of H_1 and H_2 in $\mathcal{D}$, is less than $\text{Max}(a\alpha, b\beta) = a\alpha$. Hence $a\alpha < a\alpha$, and (3.4.1) is verified.

To prove (3.4.2), we note that for any sequence $\{\delta_\lambda\}$ converging to zero, lemmas (3.3) and (3.2) assure finite existence of $N_p(\delta_\lambda) = m_\lambda$ for every λ. Now suppose (3.4.2) were false. Then there would exist some sequences, $\{\delta_\lambda, m_\lambda\}$; $\delta_\lambda \longrightarrow 0$ as $\lambda \longrightarrow \infty$, such that $\delta_\lambda m_\lambda^{r'} \longrightarrow D' < D$ as $\lambda \longrightarrow \infty$. We also have

$$[illegible] \qquad a\alpha \geq r_{m_\lambda}(P_{m_\lambda}, \delta_\lambda, \emptyset,\delta_\lambda) \text{ for all } \lambda .$$

But by assumption (g') every limit point of the expression on the right is greater than $a\alpha$. We again obtain $a\alpha > a\alpha$, which completes the proof.

We now state two more assumptions.

Assumption (g). For every k, $E_{\theta_{1i}} \phi$ and $E_{\theta_{2j}} \{1-\phi\}$ are non-increasing in δ for $i = 1, 2, \ldots, R_1$ and $j = 1, 2, \ldots, R_2$. (θ_{1i} and θ_{2j} are functions of δ.)

Assumption (h). For every d in $\mathcal{D}$ we have,

$$E_{\theta'_{1i}} \phi \longrightarrow 1 - H_1(c,d) \text{ as } k \longrightarrow \infty, \text{ where } \theta'_{1i}=\theta_{1i}(k,dk^{-r'}); i=1,2,\ldots,R_1,$$

and

$$E_{\theta'_{2j}} \{1-\phi\} \longrightarrow H_2(c,d) \text{ as } k \longrightarrow \infty, \text{ where } \theta'_{2j}=\theta_{2j}(k,dk^{-r'}); j=1,2,\ldots,R_2.$$

Theorem (3.4)

If assumptions (a) through (h) are satisfied, then asymptotically as $\delta \longrightarrow 0$,

(1) $N(\delta) \sim (D/\delta)^{1/r'}$, where $D = D(\alpha,\beta)$ as defined by assumption (d),

(2) The test ϕ is asymptotically efficient, that is, $N_c(\delta) \sim N(\delta)$.

Proof of Theorem (3.4)

To prove the theorem we need only show that conditions (g) and (h) imply condition (g'). Suppose assumptions (g) and (h) hold but that (g') is false. Then there exist positive sequences $\{\delta_\lambda\}$ and $\{k_\lambda\}$ such that $\delta_\lambda k_\lambda^{r'} \longrightarrow D' < D$ as $\lambda \longrightarrow \infty$, $k_\lambda \longrightarrow \infty$ with λ, and at the same time,

(3.4.5) $$r_{k_\lambda}(P_{k_\lambda,\delta_\lambda},\phi,\delta_\lambda) \longrightarrow r_0 \leq a\alpha \text{ as } \lambda \longrightarrow \infty \quad .$$

By an argument similar to that following (3.4.3) we have D'' in $\mathcal{D}$ such that $D' < D'' < D$ and for λ large enough,

(3.4.6) $$\delta_\lambda < D'' k_\lambda^{-r'} \quad .$$

Now
$$r_{k_\lambda}(P_{k_\lambda}, \delta_\lambda, \emptyset, \delta_\lambda) = \sum_i p_{1i} a E_{\theta_{1i}} \emptyset + \sum_j p_{2j} b E_{\theta_{2j}} \left\{ 1 - \emptyset \right\},$$

and by assumption (g) the expectations in this last expression will not be increased if we replace δ_λ by $D'' k_\lambda^{-r'}$, provided λ is large enough. If we consider any arbitrary subsequence $\{\lambda'\}$ of $\{\lambda\}$ such that the sums $\sum_i p_{1i}(m_{\lambda'}, \delta_{\lambda'})$ and $\sum_j p_{2j}(m_{\lambda'}, \delta_{\lambda'})$ converge to limits p' and $(1-p')$, we have by assumption (h),

(3.4.7) $$a\alpha \geq ap' \left\{ 1 - H_1(c, D'') \right\} + b(1 - p') H_2(c, D'') \quad .$$

By the uniqueness assumption (d) and the monotonicity properties of H_1 and H_2, the right-hand side is greater than $a\alpha$. The proof is complete.

In the applications of the preceding theorems to particular cases, the following lemmas and theorems will be useful.

Let U_k and V_k be independent random variables with the binomial frequency functions $\binom{m}{u}\xi^u(1-\xi)^{m-u}$ and $\binom{n}{v}\eta^v(1-\eta)^{n-v}$ for $\eta = \xi + \dfrac{\varepsilon}{\sqrt{k}}$, where $m = km'$, $n = kn'$; m', n' and k are positive integers.

Let
$$W_k = \sqrt{k}\,(U_k/m - V_k/n) \qquad ,$$

and
$$\phi_k(t;\xi,\varepsilon) = E_{\xi,\varepsilon}\, e^{it[W_k - E_{\xi,\varepsilon} W_k]} \quad .$$

Lemma (3.5.1)

If $B = \left\{(\xi,\varepsilon):\ a \leq \xi \leq 1 - a;\ \varepsilon_0 \leq \varepsilon \leq \varepsilon_1\right\}$ where $0 < a < 1/2$; $-\infty < \varepsilon_0 \leq \varepsilon_1 < \infty$, then for every fixed real t, $\phi_k(t;\xi,\varepsilon) \longrightarrow e^{-(t\sigma)^2/2}$ uniformly in B as $k \longrightarrow \infty$, where $\sigma^2 = \xi(1-\xi)\left\{\frac{1}{m'} + \frac{1}{n'}\right\}$.

Proof of Lemma (3.5.1)

Cramer [4], page 199, gives results which yield

$$\phi_k(t;\xi,\varepsilon) = \left\{1 + \frac{y_{1k}(\xi)}{km'}\right\}^{km'}\left\{1 + \frac{y_{2k}(\xi,\varepsilon)}{kn'}\right\}^{kn'},$$

where

$$y_{1k}(\xi) = -t^2\xi(1-\xi)/2m' + \vartheta_1 k^{-1/2}; \quad \mathrm{Mod}(\vartheta_1) < 1,$$

$$y_{2k}(\xi,\varepsilon) = -t^2\eta(1-\eta)/2n' + \vartheta_2 k^{-1/2} \quad \mathrm{Mod}(\vartheta_2) < 1,$$

$$y_{1k}(\xi) \longrightarrow y_1(\xi) = -t^2\xi(1-\xi)/2m' \text{ uniformly in B.}$$

For $y_2(\xi) = -t^2\xi(1-\xi)/2n'$,

$$|y_{2k}(\xi,\varepsilon)-y_2(\xi)| = \left|\vartheta_2 k^{-1/2} + \frac{t^2k^{-1/2}}{2n'}(\xi\varepsilon - \varepsilon(1-\xi) + \varepsilon^2 k^{-1/2})\right|$$

$$\leq k^{-1/2}\left[1 + \frac{t^2}{2n'}\,\mathrm{Max}\left\{(2|\varepsilon_0| + \varepsilon_0^2),(2|\varepsilon_1| + \varepsilon_1^2)\right\}\right],$$

and $y_{2k}(\xi,\varepsilon) \longrightarrow y_2(\xi)$ uniformly in B as $k \longrightarrow \infty$. Hence,

$$\phi_k(t;\xi,\varepsilon) \longrightarrow e^{\{y_1(\xi)+y_2(\xi)\}} = e^{-(t\sigma)^2/2}$$

uniformly in B as $k \longrightarrow \infty$.

Lemma (3.5.2)

If W_k is a random variable and $a_k = EW_k \longrightarrow +\infty$ as $k \longrightarrow \infty$, $\sigma_k^2 = VW_k \leq M < \infty$ for all k, and X_k is a random variable converging in probability to a value $x < +\infty$, then ,

$$P\left\{W_k \leq X_k\right\} \longrightarrow 0 \text{ as } k \longrightarrow \infty \quad .$$

Proof of Lemma (3.5.2)

Given $\varepsilon > 0$ and $\delta > 0$, there exists $K_1 = K_1(\varepsilon,\delta)$ such that $k > K_1$ implies $P\left\{X_k \leq x + \delta\right\} > 1 - \varepsilon/2$,

$$P\left\{W_k \leq X_k\right\} = P\left\{W_k \leq X_k, X_k \leq x + \delta\right\} + P\left\{W_k \leq X_k, X_k > x+\delta\right\}$$

$$< P\left\{W_k \leq x+\delta\right\} + \varepsilon/2 = P\left\{W_k - a_k \leq x + \delta - a_k\right\} + \varepsilon/2.$$

For k sufficiently large we have $x + \delta - a_k < 0$. For such k,

$$P\left\{W_k \leq X_k\right\} < P\left\{ |W_k - a_k| \geq a_k - x - \delta\right\} + \varepsilon/2 \leq \frac{M}{(a_k-x-\delta)^2} + \varepsilon/2 .$$

For k sufficiently large we have $\dfrac{M}{(a_k-x-\delta)^2} < \varepsilon/2$. Then,

$P\left\{W_k \leq X_k\right\} < \varepsilon$, which completes the proof.

Lemma (3.5.3)

Let $g(y)$, $g_1(y)$, $g_2(y)$, ... ad. inf., be functions defined for $y\varepsilon A$, and let $B \subset A_k \subset A$ for all k. If

(1) $\lim\limits_{k \longrightarrow \infty} g_k(y) = g(y)$ uniformly for y in B,

(2) $\underline{\lim\limits_{k \longrightarrow \infty}} \inf\limits_{A_k - B} g_k(y) > \inf\limits_{B} g(y)$,

then,

$$(3)\qquad \lim_{k \to \infty} \inf_{A_k} g_k(y) = \inf_B g(y) \qquad .$$

Proof of Lemma (3.5.3)

For any sufficiently small positive ε we can choose K so that $k > K$ implies

$$\left| g_k(y) - g(y) \right| < \varepsilon \qquad \text{for all } y \text{ in } B ,$$

$$\inf_{A_k - B} g_k(y) > \inf_B g(y) + \varepsilon \qquad .$$

Also,

$$\inf_{A_k} g_k(y) = \min \left\{ \inf_B g_k(y), \inf_{A_k - B} g_k(y) \right\} \quad .$$

For $k > K$, $\inf_B g_k(y) = \inf_B \left\{ g(y) + [\, g_k(y) - g(y) \,] \right\} = \inf_B g(y) + \Theta_k \varepsilon$; $|\Theta_k| \leq 1$.

Hence, $\inf_{A_k} g_k(y) = \inf_B g_k(y)$ and we have, $k > K$ implies

$$\left| \inf_{A_k} g_k(y) - \inf_B g(y) \right| < \varepsilon, \qquad \text{which implies (3) .}$$

Lemma (3.5.3A)

Let $h(y)$; $h_1(y)$, $h_2(y)$, ... ad. inf. be functions defined for $y \varepsilon A$, and let $B \subset A_k \subset A$ for all k. If

$$(1)\qquad \lim_{k \to \infty} h_k(y) = h(y) \quad \text{uniformly for } y \text{ in } B, \quad .$$

$$(2)\qquad \overline{\lim_{k \to \infty}} \sup_{A_k - B} h_k(y) < \sup_B h(y) \qquad ,$$

then,

$$(3)\qquad \lim_{k \to \infty} \sup_{A_k} h_k(y) = \sup_B h(y) \qquad .$$

<u>Proof of Lemma (3.5.3A)</u>

Let $g(y) = -h(y)$; $g_i(y) = -h_i(y)$ (for $i = 1, 2, \ldots$). Then conditions (1) and (2) are equivalent to the corresponding conditions of Lemma (3.5.3). Condition (3) is equivalent to the corresponding conclusion of Lemma (3.5.3).

We state without proof a theorem due to Parzen [11]:

<u>Theorem (3.5.4)</u> (Theorem 7c of Parzen)

Let $F_o(x,\theta)$, $F_1(x,\theta)$, $F_2(x,\theta)$, ... ad. inf., be distribution functions defined for all x in $[-\infty, \infty]$ and all θ in $(\ )$ (Euclidean). Let the $F_o(x,\theta)$ form an equicontinuous θ-family for all x in $[-\infty, \infty]$. Let $\phi_k(t,\theta) = E_{F_k} e^{itx}$ for $i = 0, 1, 2, \ldots$.

Then the necessary and sufficient condition that, at each x in $[-\infty, \infty]$, $F_k(x,\theta) \longrightarrow F_o(x,\theta)$ uniformly in $\theta \varepsilon (\ )$ as $k \longrightarrow \infty$, is:

(1) for every real t, $\phi_k(t,\theta) \longrightarrow \phi_o(t,\theta)$ uniformly in $\theta \varepsilon (\ )$.

It then follows that convergence is uniform in (x, θ) .

Note: $F(x,\theta)$ are equicontinuous at $x = \pm \infty$ if $\lim_{x \longrightarrow \pm \infty} F(x,\theta) = F(\pm \infty, \theta)$ uniformly in θ.

<u>Lemma (3.5.5)</u>

If $F_k(x,\theta)$ are distribution functions for all θ in some set B and $k = 0, 1, \ldots$ ad inf. such that $P_\theta \left\{ W_k \leq x \right\} = F_k(x,\theta)$ for $k = 1, 2,$ ad inf., and

(1) $F_k(x,\theta) \longrightarrow F_o(x,\theta)$ uniformly in θ for each fixed x as $k \longrightarrow \infty$,

(2) The functions $F_o(x,\theta)$ for θ in B are an equicontinuous

family,

(3) There is a function $\sigma(\theta)$ defined for all $\theta \varepsilon B$ such that, given any $\varepsilon > 0$ and $\delta > 0$, there exists $K = K(\varepsilon, \delta)$ independent of θ, with the property, $k > K$ implies

$$P \left\{ \left| D_k - \sigma(\theta) \right| > \delta \right\} < \varepsilon \text{ for all } \theta \varepsilon B ,$$

then it follows that at each fixed x,

(4) $F_k(xD_k, \theta) \longrightarrow F_o[x\sigma(\theta), \theta]$ uniformly in B.

Proof of Lemma (3.5.5)

For $x = 0$ or $\pm \infty$ the lemma is trivially true. Otherwise we may, without loss of generality, assume that $0 < x < \infty$. For any $\delta > 0$,

(3.5.5.1) $P_\theta\left\{W_k \leq xD_k\right\}$

$$= P_\theta\left\{W_k \leq xD_k, \left|D_k - \sigma(\theta)\right| \leq \delta\right\} + P_\theta\left\{W_k \leq xD_k, \left|D_k - \sigma(\theta)\right| > \delta\right\} .$$

Now given $\varepsilon > 0$, we can pick $K_1 = K_1(\varepsilon, \delta)$ such that $k > K_1$ implies

$$P_\theta\left\{ \left|D_k - \sigma(\theta)\right| > \delta \right\} < \varepsilon/3 \quad \text{for all } \theta \text{ in } B .$$

(3.5.5.2) $P_\theta\left\{W_k \leq xD_k, \left|D_k - \sigma(\theta)\right| \leq \delta\right\}$

$$\leq P_\theta\left\{W_k \leq xD_k\right\} \leq P_\theta\left\{W_k \leq xD_k, \left|D_k - \sigma(\theta)\right| \leq \delta\right\} + \varepsilon/3 .$$

(3.5.5.3) $P_\theta\left\{W_k \leq x[\sigma(\theta) - \delta], \left|D_k - \sigma(\theta)\right| \leq \delta\right\}$

$$\leq P_\theta\left\{W_k \leq xD_k\right\} \leq P_\theta\left\{W_k \leq x[\sigma(\theta) + \delta]\right\} + \varepsilon/3 .$$

$$P_\theta\left\{W_k \leq x[\sigma(\theta) - \delta]\right\} - \varepsilon/3 \leq P_\theta\left\{W_k \leq xD_k\right\} \leq P_\theta\left\{W_k \leq x[\sigma(\theta) + \delta]\right\} + \varepsilon/3,$$

uniformly in θ for $k > K_1$.

By the remark of theorem (3.5.4) (which is easily proved) convergence of $F_k(x,\theta)$ is uniform in the vector (x,θ), and we may choose $K_2(\varepsilon) = K_2$ such that $F_o\{x[\sigma(\theta)-\delta],\theta\} -\varepsilon/3 < P_\theta\{W_k \leq x[\sigma(\theta)-\delta]\}$ and

$$P_\theta\{W_k \leq x[\sigma(\theta)+\delta]\} < F_o\{x[\sigma(\theta)+\delta],\theta\} + \varepsilon/3$$

for all θ when $k > K_2$.

For all θ, when $k > K_1, K_2$, we have

$$(3.5.5.4) \qquad F_o\{x[\sigma(\theta)-\delta],\theta\} - \frac{2\varepsilon}{3} < P_\theta\{W_k \leq xD_k\}$$

$$< F_o\{x[\sigma(\theta)+\delta],\theta\} + \frac{2\varepsilon}{3} \quad .$$

Because of the equicontinuity of the family we may pick $\Delta(\varepsilon)$ such that $|y_1 - y_2| < \Delta(\varepsilon)$ implies $|F_o(y_1,\theta) - F_o(y_2,\theta)| < \varepsilon/3$ for all θ in B.

Up to this point the only requirement on δ has been that it be positive, and we now specify that $\delta = \frac{\Delta(\varepsilon)}{2x}$ which is positive because x is finite and positive. Now take $y_1 = x\sigma(\theta)$ and $y_2 = x\sigma(\theta) \pm x\delta$. Then $|y_1 - y_2| = x\delta = \frac{\Delta(\varepsilon)}{2}$, and we have

$$(3.5.5.5) \quad F_o\{x\sigma(\theta),\theta\} -\varepsilon < F_k(xD_k,\theta) < F_o\{x\sigma(\theta),\theta\} +\varepsilon,$$

$$\text{for } k > K_1,K_2 \text{ and all } \theta\varepsilon B \quad .$$

This completes the proof.

In preparation for the next lemma consider the equation

$$\frac{u+v}{u-v} = e^{2uv} \qquad \text{for fixed } v; \; 0 < v < \infty \quad .$$

The left-hand side is monotone decreasing in $u > v$ and passes from $+\infty$

to 1, while the right-hand side is monotone increasing and passes from e^{2v^2} to $+\infty$, as u goes from $v+0$ to $+\infty$. Hence the equation

$$u = v\,\frac{e^{2uv}+1}{e^{2uv}-1} \quad ,$$

has a unique root $u(v) > v$ for every fixed v such that $0 < v < \infty$.

Lemma (3.5.6)

For $0 < v < \infty$, let $u(v)$ be the unique root $u > v$ of the equation, $u = v\,\frac{e^{2uv}+1}{e^{2uv}-1}$. [$u(v) > v$ is uniquely defined.] Let $\alpha(v) = 2\,\Phi(-v)$; $\beta(v) = \Phi(u+v) - \Phi(u-v) > 0$ for finite v; $R(v) = \alpha(v) + \beta(v)$. Then $R(v) > 1/2$ for $0 < v < \infty$.

Proof of Lemma (3.5.6)

We may assume $v \geqq .67$ because $\alpha(v) > 1/2$ for $0 \leqq v \leqq .67$.

$u - v = \frac{2v}{e^{2uv}-1} < 2v/2uv = 1/u < 1/v \longrightarrow 0$ as $v \longrightarrow \infty$. Hence $R(\infty) = 1/2$. If $R(v) \leqq 1/2$ for finite v, then there is some finite v such that $R'(v) = 0$. The supposed value of v must be greater than .67.

Now, $\frac{du}{dv} = \frac{u}{v}\,\frac{(1 - u^2 + v^2)}{(1 + u^2 - v^2)}$ and if $R'(v) = 0$, we have

$$(3.5.6.1) \qquad \frac{2u(u+v)}{1 + (u^2 - v^2)} = e^{uv + u^2/2} \quad .$$

Now $u < v + 1/(u(1+v^2))$ and $u \geqq 2v$ implies $2v^2 \leqq 1/(1+v^2)$ which implies $v < .61$. Hence we may assume that $u \leqq 2v$. We have from

(3.5.6.1),

$$(3.5.6.2) \qquad e^{uv+u^2/2} < 2u^2 + 2uv \leq 3(uv + u^2/2),$$

which implies [13],

$$3v^2/2 < (uv+u^2/2) < 1.5122 \quad \text{or} \quad v < 1.009 < 1.01 .$$

$z = u/v$ is monotone decreasing in v and $v < 1.01$ implies

$$u/v > u(1.01)/(1.01) > 1.2/(1.01) \quad \text{or} \quad u > 1.188v,$$

from the same table [13]. Combining with $(u > v)$, we obtain: $1 + (u^2-v^2) > 1 + .41344v^2 > 1 + .43144(.67)^2$ and $(1 + (u^2-v^2))^{-1} < .85$. From (3.5.6 .1) we obtain

$$(3.5.6.3) \qquad e^{uv+u^2/2} < (.85)(2u^2+2uv) < 2.55(uv+u^2/2),$$

which is impossible.

3.6 Application of Theorem (3.4) to a Particular Problem.

(3.6.1) The Problem

Let $X_k = (U_m, V_n)$ where $(m,n) = (km', kn')$ and k, m', and n' are positive integers. U_m has the frequency function $f_\xi(u,m) = \binom{m}{u}\xi^u(1-\xi)^{m-u}$ and V_n has the frequency function $f_\eta(v,n) = \binom{n}{v}\eta^v(1-\eta)^{n-v}$. Let $\theta = (\xi,\eta)$ in $\Omega = \{\theta: 0 \leq \xi,\eta \leq 1\}$. We assume that U_m and V_n are independent. Then X_k takes on values $x = (u, v)$ in $\mathcal{X}_k$, where $\mathcal{X}_k = \{(u,v): u=0, 1, \ldots, m; v = 0, 1, \ldots, n\}$. X_k has the frequency function $f_\theta(x,k) = f_\xi(u,m)f_\eta(v,n)$.

Note that X_k is not of the form suggested in Chapter I but is a sufficient statistic based on an observation which is of that form. In this and the succeeding examples we may and will confine ourselves to tests based on such sufficient statistics.

For $0 < \delta < 1$ and $h(\theta) = (\xi-\eta)$, let $\omega_1(\delta) = \{\theta : h(\theta) \leq -\delta\}$ and $\omega_2(\delta) = \{\theta : h(\theta) \geq \delta\}$. Suppose our problem is to decide that θ is in one of these two sets where it does not matter which decision is made if θ is in $\Omega - \omega_1 - \omega_2$. We want to find a test which is asymptotically efficient in the sense of Theorem (3.4), where $0 < \alpha < 1/2$ and $0 < \beta < 1/2$. We will do this by showing that the conditions of the theorem can be satisfied.

Assumptions (a) and (b) are satisfied. This is obvious.

(3.6.2) <u>Assumption (c) is satisfied</u>.

Let $Z_k = z_k(X_k) = \sqrt{k}\,(U_m/m - V_n/n)$ and $r' = 1/2$. First we study

(3.6.2.1) $$M_{1k}(x,\delta) = \inf_{\omega_1(\delta)} P_\theta(Z_k \leq x) \quad .$$

Now $\theta \varepsilon\, \omega_1$ is of the form $\theta = (\xi, \xi + \varepsilon)$ where $\delta \leq \varepsilon \leq 1$ and $0 \leq \xi \leq 1 - \varepsilon$. We will first show that the minimum in (3.6.2.1) occurs for $\varepsilon = \delta$. Let (x,k) be fixed and define

(3.6.2.2) $$R(u,v) = \begin{Bmatrix} 1 \text{ if } z_k(u,v) \leq x \\ 0 \text{ otherwise} \end{Bmatrix} = \begin{Bmatrix} 1 \text{ if } \frac{u}{m} - \frac{v}{n} \leq \frac{x}{\sqrt{k}} \\ 0 \text{ otherwise} \end{Bmatrix} \quad .$$

Now $R(u,v)$ is non-increasing in u and non-decreasing in v. We have

(3.6.2.3) $$P_\theta(Z_k \leq x) = \sum_{u,v} R(u,v)\, f_\xi(u,m)\, f_{\xi+\varepsilon}(v,n) \quad ,$$

and the derivative of the last expression with respect to ε is

$$n\left\{ \sum_{u,v} \left[R(u,v+1)-R(u,v) \right] f_{\xi}(u,m)\, f_{\xi+\varepsilon}(v,n-1)\right\} \geq 0 \quad .$$

Hence

$$(3.6.2.4) \qquad M_{1k}(x,\delta) = \inf_{0 \leq \xi \leq 1-\delta} P_{\xi,\xi+\delta}(Z_k \leq x) \quad .$$

We are interested in the limiting properties of $M_{1k}(x,dk^{-r'}) = M_{1k}(x,dk^{-1/2})$.

For the case $x + d < 0$, let $\theta_{ok} = (0, dk^{-1/2})$. $E_{\theta_{ok}} Z_k = -d$, and $V_{\theta_{ok}} Z_k = dk^{-1/2}(1-dk^{-1/2})/n' \longrightarrow 0$ as $k \longrightarrow \infty$. Therefore $Z_k \longrightarrow -d$ in probability as $k \longrightarrow \infty$, when θ_{ok} is the parameter. Hence, for $x + d < 0$,

$$(3.6.2.5) \qquad M_{1k}(x,dk^{-1/2}) \leq P_{\theta_{ok}} \left\{ Z_k \leq x \right\} \longrightarrow 0 \text{ as } k \longrightarrow \infty \quad .$$

For the case $x + d > 0$, let $\sigma^2(\xi) = \xi(1-\xi)(1/m' + 1/n')$ and $\sigma = \sigma(1/2)$. Then there exists some constant a, $0 < a < 1/2$, such that $\sigma^2(a) < \frac{(x+d)^2}{4} \left\{ 1 - \Phi(\frac{x+d}{\sigma}) \right\}$. Let $B = \left\{ \xi : a \leq \xi \leq 1-a \right\}$ and $H_1(x,\xi,d) = \Phi(\frac{x+d}{\sigma(\xi)})$. It is easy to verify that the $H_1(x,\xi,d)$ form an equicontinuous family for ξ in B. $Z_k - E_{\theta_k} Z_k$ has, for $\theta_k = (\xi,\xi+dk^{-1/2})$, the characteristic function of lemma (3.5.1) which converges uniformly for ξ in B to $e^{-(t\sigma[\xi])^2/2}$. Hence by Parzen's Theorem (3.5.4), $P_{\theta_k}\left\{ Z_k \leq x \right\} \longrightarrow H_1(x,\xi,d)$ uniformly in B as

$k \to \infty$. Let $A = \{\xi : 0 \leq \xi \leq 1\}$ and $t = (x+d)/2 > 0$. Now $x > t-d$ and $P_{\theta_k}(Z_k \leq x) \geq P_{\theta_k}(-t-d \leq Z_k \leq t-d) \geq 1 - (V_{\theta_k} Z_k)/t^2$. Hence,

$$\varliminf_{k \to \infty} \inf_{A-B} P_\theta(Z_k \leq x) \geq 1 - \frac{4\sigma^2(a)}{(x+d)^2} > \Phi\left(\frac{x+d}{\sigma}\right) = \inf_B H_1(x,\xi,d) = H_1(x,d).$$

The conditions of Lemma (3.5.3) are satisfied, and for $x + d > 0$, $M_{1k}(x, dk^{-1/2}) \to H_1(x,d)$ as $k \to \infty$. Finally, M_{1k} is monotone non-decreasing in x and $M_{1k}(x,dk^{-1/2}) \to H_1(x,d)$ as $k \to \infty$, where $(x+d) \neq 0$, $d > 0$, and

$$(3.6.2.6)\quad H_1(x,d) = \left\{\begin{array}{ll} \text{Zero} & \text{if } x + d < 0 \\ \Phi\left(\frac{x+d}{\sigma}\right) & \text{if } x+d \geq 0, \text{ where } \sigma^2 = (1/4)(1/m'+1/n') \end{array}\right\}.$$

We can use arguments analogous to those just completed to show that if $M_{2k}(x,\delta) = \sup_{\omega_2(\delta)} P_\theta(Z_k \leq x)$, then $M_{2k}(x,dk^{-1/2}) \to H_2(x,d)$ as $k \to \infty$, at every continuity point of $H_2(x,d)$, where $d > 0$, and

$$(3.6.2.7)\quad H_2(x,d) = \left\{\begin{array}{ll} \Phi\left(\frac{x-d}{\sigma}\right) & \text{if } x \leq d \\ 1 & \text{if } x > d \end{array}\right\}.$$

(3.6.3) <u>Assumption (d) is satisfied.</u>

We want to demonstrate the uniqueness of the solution to the simultaneous equations, $H_1(x,d) = 1-\alpha$ and $H_2(x,d) = \beta$. We have assumed that $\alpha < 1/2$ and $\beta < 1/2$. Hence, $\Phi\left(\frac{x+d}{\sigma}\right) = 1-\alpha$ and $\Phi\left(\frac{x-d}{\sigma}\right) = \beta$ is an equivalent set of equations. The solution $(x,d) = (c,D)$ is given by

$$c = \frac{\sigma}{2}\left[\Phi^{-1}(1-\alpha) + \Phi^{-1}(\beta)\right], \quad D = \frac{\sigma}{2}\left[\Phi^{-1}(1-\alpha) - \Phi^{-1}(\beta)\right] > 0.$$

It is clear that the solution is unique and that $-D < c < D$. The neighborhood $\mathcal{J}^{\theta}$ is determined in an obvious way.

(3.6.4) Assumptions (e) and (f) are satisfied.

In order to construct the a priori distribution we first define $\xi = \xi(\delta)$ for $0 < \delta < 1/2$ by the equation:

(3.6.4.1) $$\left(\frac{\xi}{1-\xi}\right)^{m'} = \left(\frac{1-\xi-\delta}{\xi+\delta}\right)^{n'} .$$

It follows that $1/2 - \delta < \xi < 1/2$. Let

(3.6.4.2) $$\theta_1(\delta) = (\xi, \xi+\delta),\ \theta_2(\delta) = (1-\xi, 1-\xi-\delta),$$

(3.6.4.3) $$p(k,\delta) = \left\{ \frac{b}{a\left(\frac{1-\xi-\delta}{\xi+\delta}\right)^{2c\sqrt{kn'}} + b} \right\} ,$$

where c is defined by (3.6.3), $\xi = \xi(\delta)$, $a = \frac{\beta}{\alpha+\beta}$, and $b = \frac{\alpha}{\alpha+\beta}$.

a Priori Distribution

Let $P_{k,\delta}(\theta)$ assign probability $p(k,\delta)$ to $\theta_1(\delta)$ and probability $1 - p(k,\delta)$ to $\theta_2(\delta)$.

Bayes Solution

The average risk with respect to $P_{k,\delta}(\theta)$ and any test ϕ' is $paE_{\theta_1}\phi' + b(1-p)E_{\theta_2}(1-\phi')$, and it is easy to show that

$$\frac{f_{\theta_2}}{f_{\theta_1}} = \left(\frac{\xi+\delta}{1-\xi-\delta}\right)^{2n(u/m-v/n)} .$$

Since $\xi + \delta > 1/2$, it follows almost immediately that the test ϕ is bayes with respect to $P_{k,\delta}(\theta)$, where $x = (u, v)$, and

$$(3.6.4.4) \qquad \phi = \phi(x) = \left\{ \begin{array}{ll} 1 & \text{if } z_k(x) > c \\ 0 & \text{otherwise} \end{array} \right\} .$$

Therefore assumption (e) is satisfied.

From the remark following (3.6.4.1) we conclude that $\xi(\delta) \longrightarrow 1/2$ as $\delta \longrightarrow 0$. Hence, $\theta_1(\delta)$ and $\theta_2(\delta)$ both converge to $\theta_k = (1/2, 1/2)$ as δ tends to zero, no matter what the positive values m' and n' may be. Hence $\left(\frac{1-\xi-\delta}{\xi+\delta}\right)^{2c\sqrt{k}\, n'} \longrightarrow 1$ as $\delta \longrightarrow 0$ for each fixed $n = n'k$. Therefore, formula (3.6.4.3) implies that $p(k,\delta) \longrightarrow \frac{b}{a+b} = b$ as $\delta \longrightarrow 0$, for each fixed k, and assumption (f) is satisfied.

<u>(3.6.5) Assumption (g) is satisfied.</u>

Since ϕ is non-decreasing in u and non-increasing in v,

$$(3.6.5.1) \qquad \frac{\partial}{\partial \xi} E_{\xi,\eta}\phi = m \sum_{u,v} \left[\phi(u+1,v) - \phi(u,v) \right] f_\xi(u,m-1) f_\eta(v,n) \geq 0 ,$$

$$\frac{\partial}{\partial \eta} E_{\xi,\eta}\phi = n \sum_{u,v} \left[\phi(u,v+1) - \phi(u,v) \right] f_\xi(u,m) f_\eta(v,n-1) \leq 0 .$$

Now if $\eta(\delta) = \xi(\delta) + \delta$, formula (3.6.4.1) yields the identity

$$\xi^{m'} \eta^{n'} = (1-\xi)^{m'} (1-\eta)^{n'} .$$

Differentiating both sides of the last expression with respect to ξ and doing a little algebra reveals that

$$\frac{d\eta}{d\xi} \left[\frac{n'}{\eta} + \frac{n'}{1-\eta} \right] = - m' \left(\frac{1}{\xi} + \frac{1}{1-\xi} \right) < 0 ,$$

which implies $\frac{d\eta}{d\xi} < 0$. By differentiating the identity (3.6.4.1) with respect to δ, factoring, and rearranging terms we obtain

$$\frac{d\xi}{d\delta}\left[\frac{m'}{\xi}+\frac{n'}{\xi+\delta}+\frac{m'}{1-\xi}+\frac{n'}{1-\xi-\delta}\right] = -n'\left[\frac{1}{\xi+\delta}+\frac{1}{1-\xi-\delta}\right] < 0 .$$

Hence $\frac{d\xi}{d\delta} < 0$. Thus $\xi(\delta)$ is a decreasing function of δ, and $\eta(\delta)$ an increasing function of δ. Let $\delta > 0$, $\Delta > 0$, $\delta + \Delta < 1$, and define $\xi = \xi(\delta+\Delta)$, $\xi' = \xi(\delta)$, $\eta = \eta(\delta+\Delta)$, $\eta' = \eta(\delta)$. Then $E_{\xi,\eta}\phi \leq E_{\xi',\eta}\phi \leq E_{\xi',\eta'}\phi$, and $E_{\theta_1}\phi$ is a non-increasing function of δ, where $\theta_1 = \theta_1(\delta)$. Similarly $E_{\theta_2}(1-\phi)$ is a non-increasing function of δ, where $\theta_2 = \theta_2(\delta)$. Assumption (g) is therefore satisfied.

(3.6.6) Assumption (h) is satisfied.

For $\theta_{1k} = (\xi(\delta), \xi(\delta)+\delta)$ and $\delta = dk^{-1/2}$, we have $E_{\theta_{1k}} Z_k = -d$ and $\xi(\delta) \longrightarrow 1/2$ as $k \longrightarrow \infty$. From our study of the characteristic function in (3.6.2), we conclude that for $c > -d$, $E_{\theta_{1k}}\phi = P_{\theta_{1k}}(Z_k > c)$ converges to $1 - \Phi(\frac{c+d}{\sigma}) = 1-H_1(c,d)$ as $k \longrightarrow \infty$. Similarly, for $c < d$, $E_{\theta_{2k}}(1-\phi) \longrightarrow \Phi(\frac{c-d}{\sigma}) = H_2(c,d)$ as $k \longrightarrow \infty$. The neighborhood $\mathcal{S}$ has just the properties, $-d < c < d$, and assumption (h) is satisfied.

We have proved

Theorem (3.6.7).

The test ϕ defined in (3.6.4.4) is asymptotically efficient in the sense of Theorem (3.4). An asymptotic expression for $N(\delta)$ as $\delta \longrightarrow 0$ is given by

$$N(\delta) \sim (D/\delta)^2 \quad ,$$

where D is defined in section (3.6.3).

3.7 Application of Theorem (3.4A) to a Particular Case.

(3.7.1) The Problem.

Let X_k be defined as in (3.6.1) with the same frequency function, Ω, etc. Let $h(\theta) = (\xi-\eta)$, $\omega_1(\delta) = \omega_1 = \{\theta{:}h(\theta) = 0\}$, and $\omega_2(\delta) = \{\theta{:}\ |h(\theta)| \geq \delta\}$, where $0 < \delta < 1$. Again we want to decide that θ is in one of these two sets where it does not matter which decision is made when θ is in $\Omega - \omega_1 - \omega_2$. We will find a test which is asymptotically efficient in the sense of Theorem (3.4A), where $0 < \alpha$, $0 < \beta$, and $\alpha + \beta \leq 1/2$. We will show that the conditions of the theorem can be satisfied.

Assumptions (a) and (b) are satisfied. This is clear.

(3.7.2) Assumption (c) is satisfied.

Let $W_k = \sqrt{k}\,(U_m/m - V_n/n)$ and $Z_k = |W_k|$.

Now

(3.7.2.1) $$M_{1k}(x,\delta) = \inf_{0 \leq \xi \leq 1} P_{\xi,\xi}(Z_k \leq x)$$

is independent of δ, and may be written $M_{1k}(x)$.

For $x < 0$, $M_{1k}(x) = 0$ for all k and $M_{1k}(x) \longrightarrow 0$ as $k \longrightarrow \infty$.

For $x > 0$, let $\theta = (\xi,\xi)$, $\sigma^2(\xi) = \xi(1-\xi)(1/m'+1/n')$, $\sigma^2 = \sigma^2(1/2)$, and $h(x) = \Phi(x/\sigma) - \Phi(-x/\sigma)$. Then there exists some constant a, $0 < a < 1/2$, so that

(3.7.2.2) $$0 < \sigma^2(a) < x^2 \left\{ 1 - h(x) \right\} \quad .$$

Let $B = \{\xi: a \leq \xi \leq 1 - a\}$, and $A = \{\xi: 0 \leq \xi \leq 1\}$. We note that the functions $\Phi(y/\sigma\{\xi\})$ form an equicontinuous family for ξ in B. Application of lemma (3.5.1) with $A_k = A$, and Parzen's Theorem (3.5.4), shows almost immediately that, as $k \to \infty$, $P_\theta\{W_k \leq y\} \to \Phi(y/\sigma\{\xi\})$ uniformly for ξ in B. Hence for $x > 0$, $P_\theta\{|W_k| \leq x\} \to H_1(x,\xi) = \Phi(x/\sigma\{\xi\}) - \Phi(-x/\sigma\{\xi\})$ uniformly for ξ in B, as $k \to \infty$. It is now evident that

$$H_1(x) = h(x) = \inf_B H_1(x,\xi) \quad . \tag{3.7.2.3}$$

For any positive x, $\inf_{A-B} P_\theta\{|W_k| \leq x\} \geq \inf_{A-B}\left[1 - \frac{\sigma^2(\xi)}{x^2}\right] > h(x)$ for all k, and the conditions of Lemma (3.5.3) are satisfied. Therefore $M_{1k}(x) \to H_1(x) = h(x)$ as $k \to \infty$, for all $x > 0$. $M_{1k}(x)$ is monotone non-decreasing in x, and

$$M_{1k}(x, dk^{-1/2}) = M_{1k}(x) \to H_1(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ \Phi(\frac{x}{\sigma}) - \Phi(\frac{-x}{\sigma}) & \text{for } x > 0 \end{cases} \quad \text{as} \tag{3.7.2.4}$$

$k \to \infty$, where $\sigma^2 = (1/4)(1/m' + 1/n')$.

We now consider $M_{2k}(x, dk^{-1/2})$, where d is positive, and

$$M_{2k}(x,\delta) = \sup_{\omega_2(\delta)} P_\theta\{Z_k \leq x\} \quad . \tag{3.7.2.5}$$

The set $\omega_2(\delta)$ is closed, and we may write

$$M_{2k}(x, dk^{-1/2}) = P_{\theta_k}(Z_k \leq x) \quad ,$$

where $\theta_k = (\xi_k, \eta_k)$ is some point in $\omega_2(dk^{-1/2})$. In studying the limiting behavior of this latter probability, we may assume that θ_k is in the set $\left\{\theta: \eta-\xi \geq dk^{-1/2}\right\} \subset \omega_2(dk^{-1/2})$. To prove this, we assume that $\theta_k = (1-\xi, 1-\xi-\varepsilon)$, where $0 \leq \xi \leq 1-\varepsilon$, $dk^{-1/2} \leq \varepsilon \leq 1$. Then θ_k is not in the above mentioned set, but does belong to $\omega_2(dk^{-1/2})$.

$$P_{\theta_k}(Z_k \leq x) = \sum_{\left|\frac{u}{m} - \frac{v}{n}\right| \leq \frac{x}{\sqrt{k}}} f_{1-\xi}(u,m) f_{1-\xi-\varepsilon}(v,n) \quad ,$$

$$= \sum_{\left|\frac{m-u'}{m} - \frac{n-v'}{n}\right| \leq \frac{x}{\sqrt{k}}} f_{1-\xi}(m-u',m)\, f_{1-\xi-\varepsilon}(n-v',n) \quad ,$$

$$= \sum_{\left|\frac{u}{m} - \frac{v}{n}\right| \leq \frac{x}{\sqrt{k}}} f_{\xi}(u,m)\, f_{\xi+\varepsilon}(v,n) \quad ,$$

$$= P_{\theta_k'}(Z_k \leq x) \quad ,$$

where $\theta_k' = (\xi, \xi+\varepsilon)$ does belong to the set in question.

For $x < 0$, $P_\theta\left\{\left|W_k\right| \leq x\right\} = 0$ for all θ and all k. Hence $M_{2k}(x, dk^{-1/2}) \longrightarrow 0$ as $k \longrightarrow \infty$, for all $x < 0$.

For $x > d > 0$, let $\theta_{ok} = (0, dk^{-1/2})$. $E_{\theta_{ok}} W_k = -d$ for all k, $V_{\theta_{ok}} W_k = dk^{-1/2}(1-dk^{-1/2})/n' \longrightarrow 0$ as $k \longrightarrow \infty$. Therefore $W_k \longrightarrow -d$ in probability as $k \longrightarrow \infty$, when θ_{ok} is the parameter. Since M_{2k} is a supremum, $M_{2k}(x, dk^{-1/2}) \geq P_{\theta_{ok}}\left\{\left|W_k\right| \leq x\right\} \longrightarrow 1$ as $k \longrightarrow \infty$, for all $x > d > 0$.

For $0 < x < d$, let $\theta_{1k} = (1/2, 1/2 + dk^{-1/2})$. We have, from (3.7.2.5), and the definition of θ_k,

(3.7.2.6) $M_{2k}(x, dk^{-1/2}) = P_{\theta_k}\left\{-x \leq W_k \leq x\right\} \geq P_{\theta_{1k}}\left\{-x \leq W_k \leq x\right\}$ for all k, where $\theta_k \in \{\theta: \xi - \eta \leq -dk^{-1/2}\}$ and $E_{\theta_{1k}} W_k = -d$.

Let:

$$\mu_k = E_{\theta_k} W_k \leq -d \text{ for all } k,$$

$$\sigma_k^2 = V_{\theta_k} W_k \ ,$$

$$\sigma^2(\xi) = \xi(1-\xi)(1/m' + 1/n') \quad .$$

Then, $V_{\theta_{1k}} W_k \longrightarrow \sigma^2$ as $k \longrightarrow \infty$, where $\sigma^2 = \sigma^2(1/2)$. We now assert that the limit points of μ_k are bounded. If this were not true, we could pick a subsequence $\{k'\}$ of $\{k\}$ such that $\mu_{k'} \longrightarrow -\infty$ and $\sigma_{k'} \longrightarrow \sigma_1$ (finite) as $k' \longrightarrow \infty$. The inequality (3.7.2.6) holds for the sequence $\{k'\}$, and the left-hand side converges to zero as $k' \longrightarrow \infty$, by lemma (3.5.2). On the other hand, an application of lemma (3.5.1) shows that the right-hand side of (3.7.2.6) converges to $\Phi(\frac{d+x}{\sigma}) - \Phi(\frac{d-x}{\sigma}) > 0$ as $k' \longrightarrow \infty$, and we get a contradiction. Hence, there exists a finite d', $0 < d < d'$, and a set $A = \{(\xi,\varepsilon): 0 \leq \xi \leq 1;\ d \leq \varepsilon \leq d'\}$, such that for $0 < x < d$ and $\eta = \xi + \varepsilon/\sqrt{k}$,

(3.7.2.7) $$\lim_{k \longrightarrow \infty} H_{2k}(x, dk^{-1/2}) = \lim_{k \longrightarrow \infty} \sup_{A} P_{\xi,\varepsilon}\{|W_k| \leq x\} \ .$$

Let $B = \{(\xi,\varepsilon): a \leq \xi \leq 1-a;\ 0 < a < 1/2\} \cap A$, and $g(\varepsilon,\sigma) = \Phi(\frac{\varepsilon+x}{\sigma(\xi)}) - \Phi(\frac{\varepsilon-x}{\sigma(\xi)})$. We can apply Lemma (3.5.1) and Theorem (3.5.4) to obtain

$$P_{\xi,\varepsilon}(W_k \leq y) \longrightarrow \Phi(\frac{\varepsilon+y}{\sigma(\xi)}) \text{ uniformly in } B \text{ as } k \longrightarrow \infty .$$

The condition of equicontinuity of the family $\Phi(\frac{\varepsilon+y}{\sigma(\xi)})$ is easily demonstrated. Since the convergence is also uniform in y we have

$$P_{\xi,\varepsilon}\left\{|W_k| \leq x\right\} \longrightarrow g(\varepsilon,\sigma) \text{ as } k \longrightarrow \infty \text{ uniformly in } B\ ,$$

where $\sigma = \sigma(\xi)$. The function g is decreasing in ε and we have $\sup_B g(\varepsilon,\sigma(\xi)) = \sup_{\xi\,\varepsilon\,B} g(d,\sigma(\xi))$. For fixed d, $\frac{dg}{d\sigma} = C[(d-x)e^{dx/\sigma^2} - (d+x)e^{-dx/\sigma^2}]$, where $C > 0$. Thus if $0 < x < d$,

$$\frac{dg}{d\sigma}\begin{Bmatrix} > \\ = \\ < \end{Bmatrix} 0 \text{ if } \sigma^2 \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} \left(\frac{2dx}{\log(\frac{d+x}{d-x})}\right) .$$

This implies almost immediately that

$$(3.7.2.8)\quad \sup_B g(\varepsilon,\sigma(\xi)) = g(d,\sigma^*),\text{ where } \sigma^* = \min\left[\sqrt{\frac{2dx}{\log(\frac{d+x}{d-x})}}\ ,\ \sigma(\tfrac{1}{2})\right] .$$

The precise value of the number a plays no role in the foregoing argument and we may, in particular, choose a so that $\sigma^2(a) < \frac{(d-x)^2}{4} g(d,\sigma^*)$. Let $t = \frac{d-x}{2} > 0$ and $\nu_k = E_\theta W_k \leq -d$, where $\theta = (\xi, \xi + \frac{\varepsilon}{\sqrt{k}})$ and $(\xi,\varepsilon)\,\varepsilon\,\Lambda$.

For all k and all θ we have

$$P_\theta\left\{|W_k| \leq x\right\} \leq P_\theta\left\{|W_k - \nu_k| > t\right\} \leq \frac{V_\theta W_k}{t^2} = \frac{4V_\theta W_k}{(d-x)^2} .$$

But $\frac{4V_\theta W_k}{(d-x)^2} \longrightarrow \frac{4\sigma^2(\xi)}{(d-x)^2}$ as $k \longrightarrow \infty$. This implies

$$(3.7.2.9)\quad \overline{\lim}_{k\longrightarrow\infty}\ \sup_{\theta\,\varepsilon\,\Lambda-B} P_\theta\left\{|W_k| \leq x\right\} \leq \frac{4\sigma^2(a)}{(d-x)^2} < g(d,\sigma^*) .$$

We may now apply lemma (3.5.3a) with $A_k = A$ to obtain, for $0 < x < d$,

$$(3.7.2.10) \qquad M_{2k}(x, dk^{-1/2}) \longrightarrow g(d,\sigma^*) \text{ as } k \longrightarrow \infty .$$

M_{2k} is monotone non-decreasing in x and we conclude that for every fixed d, $0 < d < \infty$, $M_{2k}(x,dk^{-1/2}) \longrightarrow H_2(x,d)$ as $k \longrightarrow \infty$, at every continuity point of $H_2(x,d)$. $H_2(x,d)$ is given by

$$(3.7.2.11) \qquad H_2(x,d) = \left\{ \begin{array}{ll} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > d \\ 1/2 & \text{if } x = d \\ \Phi(\frac{d+x}{\sigma*}) - \Phi(\frac{d-x}{\sigma*}) & \text{for } 0 < x < d \end{array} \right\} ,$$

$$\text{where } \sigma* = \sigma(x,d) = \min \left\{ \left(\frac{2dx}{\log(\frac{d+x}{d-x})} \right)^{1/2} , (1/2)\left(\frac{1}{m'} + \frac{1}{n'}\right)^{1/2} \right\} .$$

Assumption (c) is therefore satisfied.

<u>Assumption (d) is satisfied</u>.

We have assumed that $0 < \alpha < 1/2$. The equation $H_1(c,d) = 1-\alpha$ is, according to (3.7.2.4), the same as $\Phi(\frac{c}{\sigma}) - \Phi(\frac{-c}{\sigma}) = 1-\alpha$. A unique positive solution clearly exists. Now we want to show that, given c, we can find a unique D such that $H_2(c,D) = \beta$. We have $0 < \beta < 1/2$, and it is clear from (3.7.3.6) that our D, if it exists, has the property $D > c > 0$. We will, in what follows, consider only values $d > c$. In the first place,

$$\frac{\partial}{\partial d} \left[\Phi(\frac{d+c}{\sigma}) - \Phi(\frac{d-c}{\sigma}) \right] = \frac{1}{\sigma} \left[d\Phi(\frac{d+c}{\sigma}) - d\Phi(\frac{d-c}{\sigma}) \right] < 0 .$$

It is clear from the definition that $\sigma^* = \sigma(c,d)$ is continuous and monotone non-decreasing in d. Hence, $d' > d$ implies

$$H_2(c,d) \geq \Phi\left(\frac{d+c}{\sigma(c,d')}\right) - \Phi\left(\frac{d-c}{\sigma(c,d')}\right) > H_2(c,d') \quad ,$$

and $H_2(c,d)$ is continuous and monotone decreasing in d.

For d large enough, $\sigma(c,d)$ is independent of d. Therefore, $\frac{d+c}{\sigma(c,d)}$ and $\frac{d-c}{\sigma(c,d)}$ both become infinite with d, and $H_2(c,d) \longrightarrow 0$ as $d \longrightarrow \infty$. As d approaches c from above, $\sigma(c,d) \longrightarrow 0$ and $\frac{d-c}{\sigma(c,d)} \longrightarrow 0$. Hence, $H_2(c,d) \longrightarrow \Phi(\infty) - \Phi(\frac{1}{2}) = \frac{1}{2}$ as $d \longrightarrow c$ from above.

There is a unique value D of d such that $H_2(c,D) = \beta$. Assumption (d) is satisfied.

<u>(3.7.3) Assumption (e) is satisfied.</u>

For $0 < \delta < 1$, let $\xi = \xi(\delta)$ be given by the equation:

$$\left(\frac{\xi}{1-\xi}\right)^{m'} = \left(\frac{1-\xi-\delta}{\xi+\delta}\right)^{n'} \quad .$$

Then $\frac{1}{2} - \delta < \xi(\delta) < \frac{1}{2}$. Define:

$$\theta_{11} = \theta_1 = \left(\frac{1}{2}, \frac{1}{2}\right) ,$$

$$\theta_{21} = (\xi, \xi+\delta) \text{ and } \theta_{22} = (1-\xi, 1-\xi-\delta) ,$$

$$a = \beta/(\alpha+\beta) \text{ and } b = \alpha/(\alpha+\beta) ,$$

$$Q(w) = 2^{m+n-1}\left(\frac{b}{a}\right)\xi^m(\xi+\delta)^n \cdot \left\{ \left(\frac{\xi}{1-\xi}\right)^{mw/\sqrt{k}} + \left(\frac{1-\xi}{\xi}\right)^{mw/\sqrt{k}} \right\} \quad ,$$

$$p = p(k,\delta) = \frac{Q(c)}{1+Q(c)} \quad .$$

The a priori distribution is given by

$$dP_{k,\delta}(\theta) = \left\{ \begin{array}{ll} p & \text{if } \theta = \theta_1 \\ \frac{1-p}{2} & \text{if } \theta = \theta_{21} \text{ or } \theta = \theta_{22} \\ 0 & \text{otherwise} \end{array} \right\} .$$

For any test ϕ' based on k observations the bayes risk is

$$r_k(P_{k,\delta},\phi',\delta) = apE_{\theta_1}\phi' + \frac{b(1-p)}{2}\left\{ E_{\theta_{21}}(1-\phi')+E_{\theta_{22}}(1-\phi') \right\} .$$

A little algebra, using the definitions of $Q(c)$ and $\xi(\delta)$, shows that the test ϕ minimizes the bayes risk, where ϕ is given by

$$\phi = \left\{ \begin{array}{l} 1 \text{ if } Q(c) < (\frac{b}{a})2^{m+n-1}\xi^m(\xi+\delta)^n\left[(\frac{\xi}{1-\xi})^{u-\frac{mv}{n}} + (\frac{1-\xi}{\xi})^{u-\frac{mv}{n}}\right] \\ 0 \quad \text{otherwise} \end{array} \right\} .$$

Let $w = \sqrt{k}(\frac{u}{m} - \frac{v}{n})$. Then $\frac{mw}{\sqrt{k}} = u - mv/n$, and we have

$$\phi = \left\{ \begin{array}{l} 1 \text{ if } Q(c) < Q(w) \\ 0 \text{ otherwise} \end{array} \right\} .$$

Now $Q(w)$ is of the form $Q(w) = \text{constant}\cdot(e^{\gamma w} + e^{-\gamma w})$, where $\gamma > 0$. $Q'(w)$ is proportional to $(e^{\gamma w} - e^{-\gamma w})$, which has the same sign as w. Therefore, $Q(w)$ is a monotone increasing function of $z = |w|$. Our bayes solution is

$$\phi = \phi(z) = \left\{ \begin{array}{ll} 1 & \text{if } z > c \\ 0 & \text{if } z \leq c \end{array} \right\} , \tag{3.7.3.1}$$

where c is given by $H_1(c) = 1 - \alpha$ and $z = \left| \sqrt{k}\ (u/m - v/n) \right|$. Assumption (e) is therefore satisfied.

(3.7.4) Assumption (f) is satisfied.

First we want to show that θ_1, θ_{21}, and θ_{22} converge to a common limit as $\delta \to 0$. $\theta_1 = (1/2,\ 1/2)$ is independent of δ. $\theta_{21} = [\xi(\delta),\ \xi(\delta) + \delta]$ and $\theta_{22} = [1 - \xi(\delta),\ 1 - \xi(\delta) - \delta]$, where $1/2 - \delta < \xi(\delta) < 1/2$. The result follows immediately.

To show that $p(k,\delta) \to b$ as $\delta \to 0$, we note that $Q(c) \to \frac{b}{a}$ as $\xi = \xi(\delta) \to 1/2$. Hence $p(k,\delta) = \frac{Q(c)}{1+Q(c)} \to \frac{b}{a+b} = b$ as $\delta \to 0$, and assumption (f) is satisfied.

Assumption (g') is satisfied.

It follows from the symmetry argument following (3.7.2.5) that $E_{\theta_{21}}(1-\phi) = E_{\theta_{22}}(1-\phi)$. Hence

$$r_k(P_{k,\delta},\phi,\delta) = paE_{\theta_1}\phi + b(1-p)E_{\theta_{21}}(1-\phi) \quad .$$

Now $E_{\theta_1}W_k = 0$, and $V_{\theta_1}(W_k) = (1/4)(\frac{1}{m'} + \frac{1}{n'}) = \sigma^2$ for all k. Also, $E_{\theta_{21}}W_k = \sqrt{k}\ \delta$ and $V_{\theta_{21}}(W_k) \to \sigma^2$ as $k \to \infty$.

Now suppose that there is a pair of positive sequences $\{\delta_\lambda\}$ and $\{k_\lambda\}$, such that $k_\lambda \to \infty$ and $\delta_\lambda k_\lambda^{1/2} \to D' < D$ as $\lambda \to \infty$. Then $\delta_\lambda \to 0$ as $k \to \infty$. Let $\theta_{1\lambda} = \theta_1(\delta_\lambda)$ and $\theta_{2\lambda} = \theta_{21}(\delta_\lambda)$. By applying Lemma (3.5.1) we can readily obtain: $E_{\theta_{1\lambda}}\phi = P_{\theta_{1\lambda}}(W_{k_\lambda} < -c)$ $+ P_{\theta_{1\lambda}}(W_{k_\lambda} > c) \to \Phi(\frac{-c}{\sigma}) + 1 - \Phi(\frac{c}{\sigma}) = \alpha$ as $k \to \infty$, and

$$(3.7.4.1)\quad E_{\theta_{2\lambda}}(1-\phi) = P_{\theta_{2\lambda}}(-c \leq W_{k_\lambda} \leq c) \longrightarrow \Phi\left(\frac{D'+c}{\sigma}\right) - \Phi\left(\frac{D'-c}{\sigma}\right) \text{as } k \longrightarrow \infty.$$

In order to complete the proof in this section we must prove the two statements:

(1) Every limit point of $p(k_\lambda, \delta_\lambda)$ is less than unity,

(2) $\Phi\left(\frac{D'+c}{\sigma}\right) - \Phi\left(\frac{D'-c}{\sigma}\right) > \beta$.

Since $p(k,\delta) = \frac{Q(c)}{1+Q(c)}$, we can prove statement (1) by showing that $Q(c)$ is bounded. $Q(c)$ is given by

$$Q(c) = 2^{m+n-1}\left(\frac{b}{a}\right)\xi^m(\xi+\delta)^n \left\{ \left(\frac{\xi}{1-\xi}\right)^{mc/\sqrt{k}} + \left(\frac{1-\xi}{\xi}\right)^{mc/\sqrt{k}} \right\} .$$

For all δ in $(0, 1)$ and $\xi = \xi(\delta)$ we have, $\xi(1-\xi) < 1/4$ and $(\xi+\delta)(1-\xi-\delta) < 1/4$. Also, by definition of ξ ,

$$\xi^{m/2}(\xi+\delta)^{n/2} = (1-\xi)^{m/2}(1-\xi-\delta)^{n/2} .$$

Hence,

$$\xi^m(\xi+\delta)^n = \left\{\xi(1-\xi)\right\}^{m/2} \left\{(\xi+\delta)(1-\xi-\delta)\right\}^{n/2} < 2^{-m-n} ,$$

and

$$Q(c) < 2^{-1}\left(\frac{b}{a}\right)\left[\left(\frac{\xi}{1-\xi}\right)^{mc/\sqrt{k}} + \left(\frac{1-\xi}{\xi}\right)^{mc/\sqrt{k}}\right]$$

$$\leq 2^{-1}\left(\frac{b}{a}\right)\left[1 + \left(\frac{1-\xi}{\xi}\right)^{mc/\sqrt{k}}\right] .$$

The relation, $1/2 - \delta < \xi < 1/2$, implies $\left(\frac{1-\xi}{\xi}\right) < \frac{1+2\delta}{1-2\delta}$. If $D'' > D'$, then $mc/\sqrt{k_\lambda} = m'c\sqrt{k_\lambda} < m'cD''/\delta_\lambda$, for λ sufficiently large. For such values of k_λ,

$$\left(\frac{1-\xi}{\xi}\right)^{mc/\sqrt{k_\lambda}} < \left(\frac{1+2\delta_\lambda}{1-2\delta_\lambda}\right)^{m'cD''/\delta_\lambda} .$$

The right-hand side of the above inequality is bounded. Hence $Q(c)$ is bounded and statement (1) is verified.

Now, we wish to prove statement (2). $\frac{\partial}{\partial D'}\left[\Phi\left(\frac{D'+c}{\sigma}\right)-\Phi\left(\frac{D'-c}{\sigma}\right)\right] \leq 0$ because $D' \geq 0$, and inequality is strict if $D' > 0$. Hence $D' < D$ implies

$$\Phi\left(\frac{D'+c}{\sigma}\right) - \Phi\left(\frac{D'-c}{\sigma}\right) > \Phi\left(\frac{D+c}{\sigma}\right) - \Phi\left(\frac{D-c}{\sigma}\right) \quad .$$

Our proof will be completed by showing that the right-hand side of the last inequality is greater than or equal to β. Suppose $\Phi\left(\frac{D+c}{\sigma}\right)-\Phi\left(\frac{D-c}{\sigma}\right) < \beta$. By the definition of $H_2(c,D)$(3.7.2.11) this implies $\sigma(c,D) < \sigma$. By the definition of $\sigma(c,D)$ this in turn implies that $D < d_0(c)$ where d_0 is defined by

$$2cd_0 = \sigma^2 \log\left(\frac{d_0+c}{d_0-c}\right) \text{ , } d_0 > c \quad .$$

Our supposition leads to the inequalities,

$$(3.7.4.2) \qquad \beta > \Phi\left(\frac{D+c}{\sigma}\right) - \Phi\left(\frac{D-c}{\sigma}\right) \geq \Phi\left(\frac{d_0+c}{\sigma}\right) - \Phi\left(\frac{d_0-c}{\sigma}\right) = \beta' \quad .$$

Let $u = d_0/\sigma$ and $v = c/\sigma$. $u = u(v)$ and v satisfy the conditions of Lemma (3.5.6), and an application of the Lemma leads to the inequality, $\alpha + \beta' > 1/2$. Therefore, $\alpha + \beta > 1/2$, contrary to the assumption of the problem. Hence,

$$\Phi\left(\frac{D'+c}{\sigma}\right)- \Phi\left(\frac{D'-c}{\sigma}\right) > \Phi\left(\frac{D+c}{\sigma}\right)- \Phi\left(\frac{D-c}{\sigma}\right) \geq \beta \quad ,$$

and statement (2) is verified. Therefore, assumption (g') is satisfied.

The assumptions of Theorem (3.4A) are satisfied in this case, and we are led to the following theorem.

Theorem (3.7.5)

The test ϕ defined by (3.7.3.1) is asymptotically efficient in the sense of Theorem (3.4A) for the problem of this section. An asymptotic expression for $N(\delta)$ in this case is

$$N(\delta) \sim (D/\delta)^2 \quad ,$$

where D is defined by $H_2(c,D) = \beta$, and $H_2(c,d)$ is given by (3.7.2.11).

3.8 Application of Theorem (3.4) to a Second Problem.

Let $X_k = (U_m, V_n)$ be defined as in section (3.6.1). Again, we let $h(\theta) = (\xi-\eta)$, $0 < \delta < 1$, and suppose we want to decide that θ is in $\omega_1(\delta) = \omega_1 = \left\{ \theta : h(\theta) \leq 0 \right\}$ or that θ is in $\omega_2(\delta) = \left\{ \theta : h(\theta) \geq \delta \right\}$. We suppose that it does not particularly matter which decision is made when $\theta \in \Omega - \omega_1 - \omega_2$.

We will find a test which is asymptotically efficient in the sense of Theorem (3.4) by showing that the conditions of the theorem can be satisfied. We again assume that $0 < \alpha, \beta < 1/2$.

Assumptions (a) and (b) are satisfied.

This is obvious.

(3.8.1) Assumption (c) is satisfied.

We will define our test in terms of the random variable $Z_k = \sqrt{k}\,(U_m/m - V_n/n)$. The argument of section (3.6.2) shows immediately that

$$M_{1k}(x, dk^{-1/2}) = \inf_{\omega_1(dk^{-1/2})} P_\theta(Z_k \leq x) = \inf_{0 \leq \xi \leq 1} P_{\xi,\xi}(Z_k \leq x) \quad ,$$

and the argument of the section still holds good for $d = 0$. Hence $M_{1k}(x, dk^{-1/2}) \longrightarrow H_1(x,d) = H_1(x)$ as $k \longrightarrow \infty$, at every continuity point of $H_1(x)$, where $H_1(x)$ is given by

$$(3.8.1.1) \qquad H_1(x) = \left\{ \begin{array}{ll} \Phi(\frac{x}{\sigma}) & \text{for } x \geq 0 \\ 0 & \text{otherwise} \end{array} \right\} ,$$

where $\sigma^2 = (1/4)(\frac{1}{m'} + \frac{1}{n'})$.

Similarly, for

$$M_{2k}(x, dk^{-1/2}) = \sup_{\omega_2(dk^{-1/2})} \left\{ P_\theta(Z_k \leq x) \right\} ,$$

we have $M_{2k}(x, dk^{-1/2}) \longrightarrow H_2(x,d)$ as $k \longrightarrow \infty$, at every continuity point of $H_2(x,d)$, where $H_2(x,d)$ is given by

$$(3.8.1.2) \qquad H_2(x,d) = \left\{ \begin{array}{ll} \Phi(\frac{x-d}{\sigma}) & \text{for } x \leq d \\ 1 & \text{otherwise} \end{array} \right\} .$$

Therefore, assumption (c) is satisfied.

<u>(3.8.2) Assumption (d) is satisfied.</u>

The solution of the simultaneous system: $H_1(c) = 1 - \alpha$ and $H_2(c,D) = \beta$, is given by

$$c = \sigma \Phi^{-1}(1-\alpha) \quad ,$$
$$D = \sigma \left[\Phi^{-1}(1-\alpha) - \Phi^{-1}(\beta) \right] > c \quad .$$

<u>(3.8.3) Assumptions (e) and (f) are satisfied.</u>

Let $\xi = \xi(\delta)$ be defined by the equation

$$\left(\frac{\xi}{1-\xi}\right)^{m'} = \left(\frac{1-\xi-\delta}{\xi+\delta}\right)^{n'} \quad ,$$

where $0 < \delta < 1/2$. Then $1/2 - \delta < \xi < \delta$. Let $\theta_{11}(k,\delta) = \theta_1 = (1/2, 1/2)$ and $\theta_{21}(k,\delta) = \theta_2(\delta) = (1-\xi, 1-\xi-\delta)$. Let $a = \beta/(\alpha+\beta)$, $b = \alpha/(\alpha+\beta)$ and

$$Q(z) = \left(\frac{b}{a}\right) 2^{m+n} \xi^m (\xi+\delta)^n \left(\frac{1-\xi}{\xi}\right)^{mz/\sqrt{k}} \quad ,$$

$$p = p(k,\delta) = \frac{Q(c)}{1+Q(c)} \quad .$$

A Priori Distribution

Let

$$P_{k,\delta}(\theta) = \left\{ \begin{array}{ll} p(k,\delta) & \text{for } \theta = \theta_1 \\ 1-p(k,\delta) & \text{for } \theta = \theta_2(\delta) \\ 0 & \text{otherwise} \end{array} \right\} .$$

Bayes Solution

The bayes risk for any test ϕ' based on k observations is $apE_{\theta_1}\phi' + b(1-p)E_{\theta_2}(1-\phi')$. To minimize the bayes risk we can use the test

$$\phi = \left\{ \begin{array}{l} 1 \text{ if } \frac{Q(c)}{1+Q(c)} 2^{-m-n} < \left(\frac{b}{a}\right)\left(\frac{1}{1+Q(c)}\right)(1-\xi)^u \xi^{m-u} (1-\xi-\delta)^v (\xi+\delta)^{n-v} \\ 0 \text{ otherwise} \end{array} \right\} .$$

Substituting $u = mz/\sqrt{k} + mv/n$ from the definition of $z = \sqrt{k}(u/m - v/n)$, we obtain

$$\phi = \left\{ \begin{array}{l} 1 \text{ if } Q(c) < \left(\frac{b}{a}\right) 2^{m+n} \xi^m (\xi+\delta)^n \left(\frac{1-\xi}{\xi}\right)^{mz/\sqrt{k}+mv/n} \left(\frac{1-\xi-\delta}{\xi+\delta}\right)^v \\ 0 \text{ otherwise} \end{array} \right\} .$$

But from the definitions of ξ and $Q(z)$, this is equivalent to

$$\emptyset = \begin{cases} 1 & \text{if } Q(c) < \emptyset(z) \\ 0 & \text{otherwise} \end{cases} .$$

The function $Q(z)$ is monotone increasing in z because $(\frac{1-\xi}{\xi}) > 1$.

Hence

$$(3.8.3.1) \qquad \emptyset = \begin{cases} 1 & \text{if } z_k > c \\ 0 & \text{otherwise} \end{cases} , \text{ where } z_k = \gamma\sqrt{k}\left(\frac{u_m}{m} - \frac{v_n}{n}\right) .$$

$\emptyset$ is the required bayes solution, and assumption (e) is satisfied.

It is clear that $\theta_2(\delta) \longrightarrow \theta_1$ (independent of k) as $\delta \longrightarrow 0$. Now $Q(c) \longrightarrow \frac{b}{a}$ as $\xi \longrightarrow 1/2$, and therefore $p(k,\delta) \longrightarrow \frac{b}{a+b} = b$ as $\delta \longrightarrow 0$. Assumption (f) is therefore satisfied.

Assumption (g) is satisfied.

$E_{\theta_1}\emptyset$ is independent of δ. That $E_{\theta_2}(1-\emptyset)$ is non-increasing in δ follows immediately from the argument of section (3.6.5).

Assumption (h) is satisfied.

This follows directly from the comments of section (3.6.6).

We have shown that the assumptions of Theorem (3.4) are satisfied, and we conclude this section with a theorem.

Theorem (3.8.4).

The test $\emptyset$ defined in (3.8.3.1) is asymptotically efficient in the sense of Theorem (3.4). The constants c and D are given by (3.8.2), and an asymptotic expression for $N(\delta)$ as $\delta \longrightarrow 0$ is

$$N(\delta) \sim (D/\delta)^2 .$$

3.9 Asymptotically Economical Sub-Sample Ratios.

Suppose that we are studying a group of populations, γ in number, and that we may take as many independent observations as we like from each population. Let the number of observations from the i^{th} population be $n_i > 0$, for $i = 1, \ldots, \gamma$. Let k be the largest common divisor of the n_i, and define $\nu_i = n_i/k$ for $i = 1, \ldots, \gamma$.

Suppose that the cost of a single observation from the i^{th} population is $c_i > 0$, for $i = 1, 2, \ldots, \gamma$, independently of the number of such observations. The total sampling cost is then $k(c_1\nu_1 + \ldots + c_\gamma\nu_\gamma)$. We will study the asymptotic cost as the n_i become infinite and $\nu = (\nu_1, \nu_2, \ldots, \nu_\gamma)$ remains fixed. We will then try to choose ν so that the asymptotic expression for the cost is minimized. This may lead us to consider non-integral or even irrational values of $(\nu_1, \ldots, \nu_\gamma)$. Such a situation will cause no difficulty since our only requirement is that $(n_1, \ldots, n_\gamma)$ be proportional to $(\nu_1, \ldots, \nu_\gamma)$. We can achieve this proportionality to any desired degree of accuracy by taking the n_i large enough.

If, for the class of decision functions based on $(n_1, n_2, \ldots, n_\gamma) = k(\nu_1, \nu_2, \ldots, \nu_k)$ observations, there is a minimum attainable risk, $r_k(\nu,\delta)$, depending on some positive quantity δ, let us define $N(\nu,\delta)$ as the least integer k such that $r_k(\nu,\delta) \leq \alpha$. We may define $C(\nu,\delta)$, for each fixed ν, as the minimum cost for which $r_k(\nu,\delta) \leq \alpha$. Then $C(\nu,\delta) = N(\nu,\delta)(c_1\nu_1 + \ldots + c_\gamma\nu_\gamma)$. It is clear that for any positive integral λ we have, $\lambda N(\lambda\nu,\delta) = N(\nu,\delta)$, and therefore, $C(\lambda\nu,\delta) = C(\nu,\delta)$. Thus, the cost is determined by a set of ratios ν_1/ν_j, for $j = 2, \ldots, \gamma$.

Lemma (3.9.1)

If there is a constant $r' > 0$ and a function $D(\nu) = \left[\frac{1}{\nu_1} + \frac{1}{\nu_2} + \dots + \frac{1}{\nu_\gamma} \right]^{r'} g$, where g is independent of ν for every $\nu \in V_0 = \{ \nu : \nu$ has positive integral components $\}$, and such that

(a) $N(\nu,\delta) \sim \left[D(\nu)/\delta \right]^{1/r'}$ as $\delta \to 0$,

Then

(b) $$\inf_{V_0} C(\nu,\delta) \sim \left(\frac{g}{\delta}\right)^{1/r'} \left(\sum_{i=1}^{\gamma} \sqrt{c_i} \right)^2 = C(\delta) \text{ as } \delta \to 0 ,$$

and

(c) $C(\nu',\delta) \sim C(\delta)$ as $\delta \to 0$, if and only if $\nu_1'/\nu_j' = \sqrt{c_j/c_1}$ for $j = 2, 3, \dots, \gamma$.

Proof of Lemma (3.9.1)

The relation (a) implies

$$(3.9.1.1) \quad C(\nu,\delta) \sim \left(\frac{g}{\delta}\right)^{1/r'} \left[\frac{1}{\nu_1} + \dots + \frac{1}{\nu_\gamma} \right] \left[c_1\nu_1 + \dots + c_\gamma \nu_\gamma \right] .$$

By Cauchy's inequality,

$$(\Sigma\, c_i^{1/2})^2 = \left[\Sigma\, (c_i\nu_i)^{1/2} \nu_i^{-1/2} \right]^2 \leq (\Sigma\, c_i\nu_i)\, (\Sigma\, \nu_i^{-1}) ,$$

and equality holds if and only if $\nu_i = \text{const.}\ c_i^{-1/2}$ for $i = 1, \dots, \gamma$. This condition is the same as $\nu_1/\nu_j = \sqrt{c_j/c_1}$ for $j = 2,3,\dots,\gamma$, which establishes the truth of (c). The condition $\nu_1/\nu_j = \sqrt{c_j/c_1}$ can be approximated with any desired accuracy by some $\nu \in V_0$, which implies (b). The proof is complete.

(3.9.2) Applications of Lemma (3.9.1) to the Examples (3.6), (3.7), and (3.8).

In all three cases the sample sizes are km' and kn' where m', n' and k, are integers. The sampling cost is $c_1 m + c_2 n$. In all three cases we may divide the risk by the constant $a = \beta/(\alpha+\beta)$, and write $N(\delta)$ = least integer k such that $\inf_{\phi'} r_k(\phi',\delta) \leq \alpha$. This has no effect on the results.

The constant $r' = 1/2$ in all three cases and the function $D(\nu)$ is of the proper form. For the three cases, the best sub-sample ratios are

(3.9.2.1) $$m'/n' = \sqrt{c_2/c_1} \quad ,$$

and the asymptotic cost as $\delta \longrightarrow 0$ is

(3.9.2.2) $$C(\delta) \sim \left\{ \frac{g(\alpha,\beta)}{\delta} \right\}^2 (\sqrt{c_1} + \sqrt{c_2})^2, \text{ for suitable } g \; .$$

We give the definitions of $g(\alpha,\beta)$ for the three cases below.

For the example in (3.6) ,

$$g(\alpha,\beta) = (1/4)\left\{ \Phi^{-1}(1-\alpha) - \Phi^{-1}(\beta) \right\} \; .$$

For the example in (3.7) we show that $\sigma(c,D) = (1/2)(\frac{1}{m'} + \frac{1}{n'})^{1/2}$.

For $y_0 = \Phi^{-1}(1-\alpha/2)$, $g(\alpha,\beta)$ is defined by

$$\Phi(2g+y_0) - \Phi(2g-y_0) = \beta \; .$$

For the example in (3.8),

$$g(\alpha,\beta) = 1/2 \left\{ \Phi^{-1}(1-\alpha) - \Phi^{-1}(\beta) \right\} \; .$$

CHAPTER IV

ASYMPTOTIC EFFICIENCY OF FISHER-TOCHER TESTS AND THE APPROXIMATE CHI SQUARE TEST

In this chapter we will be dealing with samples from two binomial populations in which the probabilities of "rejection" in a single trial are ξ and η respectively. We assume that sampling is carried on in such a way that the ratio of the two sample sizes remains fixed.

We first consider the one-sided Fisher-Tocher test [14] as a method of deciding either that $\xi \leq \eta$ or that $\xi - \eta \geq \delta$, where $0 < \delta < 1$. We show that this test is asymptotically efficient as δ tends to zero (Theorem 4.2.4).

Next we treat a two-sided version of the Fisher-Tocher test as a method of deciding either that $\xi = \eta$ or that $\xi - \eta \geq \delta$ for $0 < \delta < 1$. We show that this test is asymptotically efficient as δ tends to zero (Theorem 4.3.2). It follows as an immediate consequence of this last result that the approximate chi square test (4.4) is also asymptotically efficient as $\delta \longrightarrow 0$.

The structure of the tests considered makes the proofs of the theorems somewhat complicated. In order to use some results of Hoeffding ([7] and [8]) we first show that the Fisher-Tocher tests are also tests based on permutations of observations. Using these results we obtain certain convergence properties which are essential to the proofs. These convergence properties do not necessarily hold when (ξ,η) is close to the boundary of $\Theta = \left\{ \theta: 0 \leq \xi , \eta \leq 1 \right\}$. When (ξ,η) is close to the boundary of $\bigcap$, certain alternative methods may be used. The proofs of Theorems (4.2.4) and (4.3.2) are therefore divided into two parts, treating the cases (ξ, η) away from the boundary and (ξ, η) close to the boundary separately.

(4.1) Introduction.

Let $X_{11}, X_{12}, \ldots,$ and $X_{21}, X_{22}, \ldots,$ be two sequences of mutually independent random variables where

$$X_{1i} = \begin{cases} 1 \text{ with probability } \xi \\ \\ 0 \text{ with probability } 1-\xi \end{cases}, \text{ for } i = 1,2,\ldots,$$

and

$$X_{2j} = \begin{cases} 1 \text{ with probability } \eta \\ \\ 0 \text{ with probability } 1-\eta \end{cases}, \text{ for } j = 1, 2, \ldots .$$

Let $\theta = (\xi,\eta) \varepsilon \bigcap = \{\theta : 0 \leq \xi \leq \eta\}$. If ω_1 and ω_2 are disjoint subsets of $\bigcap$ we may, in certain situations, compare a most powerful unbiased test of ω_1 against ω_2 with one of the asymptotically efficient bayes tests obtained in Chapter III. Let $U_m = \sum_{i=1}^{m} X_{1i}$ and $V_n = \sum_{j=1}^{n} X_{2j}$. Let the test ψ be defined by the relations:

$$(4.1.1) \quad \psi(u,v) = \begin{cases} 1 \quad \text{if } u - v > c(s) \\ a(s) \text{ if } u - v = c(s) \\ 0 \quad \text{otherwise} \end{cases}, \text{ where } s=u+v ,$$

$$h(j, s) = \binom{m}{j}\binom{n}{s-j}/\binom{m+n}{s} \qquad ,$$

$$\sum_{j > c(s)} h(j,s) \leq \alpha < \sum_{j \geq c(s)} h(j,s) \quad ,$$

$$\sum_{j > c(s)} h(j,s) + a(s)\, h\, [\, c(s), s\,] = \alpha \quad .$$

The Fisher-Tocher test, ψ, is most powerful unbiased of level α for testing $\{\theta: \xi - \eta \leq 0\}$ against $\{\theta: \xi - \eta > 0\}$ [14]. For δ close to zero we may regard ψ as a good test of $\{\theta: \xi - \eta \leq 0\}$ against $\{\theta: \xi - \eta \geq \delta\}$ and compare it with the asymptotically efficient bayes test for this problem found in section (3.8) of Chapter III.

A two-sided extension of the test ψ with $\alpha/2$ assigned to each "tail" may be regarded as a good test of $\{\theta: \xi - \eta = 0\}$ against $\{\theta: |\xi - \eta| \geq \delta\}$ for small values of δ, and compared with the corresponding asymptotically efficient bayes test obtained in section (3.7) of Chapter III. We may also compare the test called the X^2 test by Cochran [3] with this latter bayes test.

(4.2) Asymptotic Efficiency of the One-Sided Fisher-Tocher Test.

We begin with some definitions. We have a sample of $m = km'$ X_1's and $n = kn'$ X_2's where m', n' and k are positive integers. The number of observations, $N = k(m' + n')$, depends only on k where m' and n' are fixed. Let

$$X = X_k = (Y_1, Y_2, \ldots, Y_N) = (X_{11}, X_{12}, \ldots, X_{1m}, X_{21}, \ldots, X_{2n}) \quad , \tag{4.2.1}$$

$$W = W_k = \sqrt{k}\,(U_m/m - V_n/n) \quad ,$$

$$S = S_k = \sum_{i=1}^{N} Y_i \quad ,$$

$$D^2 = D_k^2 = (\frac{N}{N-1})(\frac{1}{m'} + \frac{1}{n'}) \frac{S}{N}(1 - \frac{S}{N}) \quad ,$$

$$t(X) = t_k(X_k) = W(X_k)/D(X_k) \quad .$$

Let $\mathcal{G} = \mathcal{G}_k$ be the permutation group of $(y_1, y_2, \ldots, y_N)$. $\mathcal{G}$ contains $M = N!$ elements. $t^{(1)}(x) \leq t^{(2)}(x) \leq \ldots \leq t^{(M)}(x)$ be the ordered values of $t(gx)$ for g in $\mathcal{G}$. Define

$[M\alpha]$ = largest integer less than or equal to $M\alpha$, $\gamma = M - [M\alpha]$,

$M^+(x)$ = number of values $t^{(j)}$ which are greater than $t^{(\gamma)}$,

$M^0(x)$ = number of values $t^{(j)}$ which are equal to $t^{(\gamma)}$,

$$a(x) = \frac{M\alpha - M^+(x)}{M^0(x)} \quad .$$

It is easy to show that the test ψ (4.1.1) has the alternative definition

$$(4.2.2) \qquad \psi = \begin{cases} 1 & \text{if } t(x) > t^{(\gamma)}(x) \\ a(x) & \text{if } t(x) = t^{(\gamma)}(x) \\ 0 & \text{otherwise} \end{cases} \quad .$$

(4.2.3) <u>Risk Function and Asymptotic Efficiency of the One-Sided Fisher-Tocher Test.</u>

We are given two positive quantities α and β with $\alpha + \beta < 1$. When the sample size is k, the risk function for ψ is

$$r_k(\theta, \psi, \delta) = \begin{cases} aE_\theta \psi & \text{for } \theta \text{ in } \omega_1 \\ bE_\theta(1-\psi) & \text{for } \theta \text{ in } \omega_2(\delta) \\ 0 & \text{otherwise} \end{cases} ,$$

where $\omega_1 = \omega_1(\delta) = \{ \theta: \xi - \eta \leq 0 \}$, $\omega_2(\delta) = \{ \theta: \xi - \eta \geq \delta \}$, $a = \beta/(\alpha+\beta)$, and $b = \alpha/(\alpha+\beta)$.

Define

$$B_k = \Omega \cap \{ \theta: 2k^{-3/4} \leq \xi, \eta \leq 1-2k^{-3/4} \} , \tag{4.2.3.1}$$

$$A_k = \Omega - B_k ,$$

$$r_{Ak}(\psi, \delta) = \sup_{A_k \cap \omega_2(\delta)} r_k(\theta, \psi, \delta) ,$$

$$r_{Bk}(\psi, \delta) = \sup_{B_k \cap \omega_2(\delta)} r_k(\theta, \psi, \delta) ,$$

$$r_k(\psi, \delta) = \operatorname*{Sup}_{\Omega} r_k(\theta, \psi, \delta) ,$$

$N_{A\psi}(\delta)$ = least integer k such that $r_{Ak}(\psi, \delta) \leq a\alpha$,

$N_{B\psi}(\delta)$ = least integer k such that $r_{Bk}(\psi, \delta) \leq a\alpha$,

$N_\psi(\delta)$ = least integer k such that $r_k(\psi, \delta) \leq a\alpha$,

$N(\delta)$ = least integer k such that $\inf_{\phi'} \operatorname*{Sup}_{\Omega} r_k(\theta, \phi', \delta) \leq a\alpha$.

Now ψ is unbiased, and it follows almost immediately that

$$N_\psi(\delta) = \operatorname{Max} \{ N_{A\psi}(\delta), N_{B\psi}(\delta) \} .$$

We obtained an asymptotic expression for $N(\delta)$ as δ tends to zero in Section (3.8) of Chapter III. We will prove the following theorem.

Theorem (4.2.4)

$\dfrac{N(\delta)}{N_{\psi}(\delta)} \longrightarrow 1$ as $\delta \longrightarrow 0$. In words, the one-sided Fisher-Tocher test is asymptotically efficient in the sense of Theorem (3.4).

The remainder of this section (4.2.4) is devoted to proving the theorem just stated.

For fixed k let $MF_k(y,s)$ be the number of elements of $\mathcal{G}$ Such that $t(gx) \leq y$ when $s = u + v$. But $t(x) \leq y$ if and only if $\sqrt{k}\, Nu/mn \leq yD + \sqrt{k}\,(s/m)$. For $h(j,s) = \dfrac{\binom{m}{j}\binom{n}{s-j}}{\binom{N}{s}}$, a little algebra shows that

$$F_k(y,s) = \sum_{j=0}^{[mnyD/N\sqrt{k}\, + ms/N]} h(j,s) \quad .$$

Let

$$a_N(i,s) = \left\{ \begin{array}{l} 1 \text{ for } i = 1, \ldots, s \\ \\ 0 \text{ for } i = s+1, \ldots, N \end{array} \right\} \text{ for } k=1,2,\ldots, \text{ and } S=0,\ldots,N,$$

$$b_N(i) = \left\{ \begin{array}{l} 1 \text{ for } i = \text{ for } i = 1, \ldots, m \\ \\ 0 \text{ for } i = \text{ for } i = m+1, \ldots, N \end{array} \right\} \quad .$$

Let $r_1, r_2, \ldots, r_M = (r_{11}, \ldots, r_{1N}), (r_{21}, \ldots, r_{2N}), \ldots, (r_{M1}, \ldots, r_{MN})$ be the $M = N!$ permutations of $(1, \ldots, N)$, and let Z_{Ni} be a random vector, independent of $(Y_1, \ldots, Y_N)$, which takes on the values $r_1, \ldots, r_M$ with equal probabilities $1/M$. Let

$$T_N(S) = \sum_{i=1}^{N} a_N(i,S)\, b_N(Z_{Ni}) \quad .$$

We now consider the conditional distribution of $T_N(S)$ for $S = s$ (fixed). $T_N(s) = j$ with probability $h(j,s)$. Let $\bar{a}_N(s) = \sum_{i=1}^{N} a_N(i,s)/N$ and $\bar{b}_N = \sum_{i=1}^{N} b_N(i)/N$. By Theorem (4) of Hoeffding [7] we have: $T_N(s)$ is asymptotically normally distributed if

(4.2.4.1)
$$\lim_{N \to \infty} \frac{N\left[\max_{1 \le i \le N}\left\{a_N(i,s)-\bar{a}_N(s)\right\}^2\right]\left[\max_{1 \le i \le N}\left\{b_N(i)-\bar{b}_N\right\}^2\right]}{\left[\sum_{i=1}^{N}\left\{a_N(i,s)-\bar{a}_N(s)\right\}^2\right]\left[\sum_{i=1}^{N}\left\{b_N(i)-\bar{b}_N\right\}^2\right]} = 0 .$$

Condition (4.2.4.1) is equivalent to the condition:

$$\lim_{N \to \infty} \frac{\operatorname{Max}\left\{s^2,(N-s)^2\right\}}{Ns(N-s)} = 0.$$ This last is true whenever

$\lim_{N \to \infty} \operatorname{Max}\left\{\frac{1}{N-s}, \frac{1}{s}\right\} = 0$. Now $ET_N(s) = \frac{ms}{N}$ and $\operatorname{Var} T_N(s) = \frac{mns(1-s/N)}{N(N-1)} = \sigma_{Ns}^2$. Let $T_N^*(s) = \left\{T_N(s) - ET_N(s)\right\}/\sigma_{Ns}$. It follows that

(4.2.4.2)
$$F_k(y,s) = P\left\{T_N^*(s) \le y\right\} .$$

We conclude that $F_k(y,s) - \Phi(y) \to 0$ as $N \to \infty$ (i.e. as $k \to \infty$). Let $g_k(s) = \operatorname{Max}\left\{\frac{1}{N-s}, \frac{1}{s}\right\}$. Now we will show that for every sequence of points $\theta_k \in B_k$ we have $g_k(S_k) \to 0$ in probability as $k \to \infty$. Now, $P\left[\operatorname{Max}\left\{S^{-1},(N-S)^{-1}\right\} > \varepsilon\right] \le P\left\{S < \frac{1}{\varepsilon}\right\} + P\left\{(N-S) < \frac{1}{\varepsilon}\right\}$, and because of symmetry it will be sufficient to prove that one of the terms on the right tends to zero. $S_k = U_k + V_k$, and $P(S_k < \frac{1}{\varepsilon}) \leq (P(U_k < \frac{1}{\varepsilon})$. For $\theta_k = (\xi_k, \eta_k)$ in B_k, $a_k = EU_k = km'\xi_k$, where $2k^{-3/4} \leq \xi_k \leq 1 - 2k^{-3/4}$, and $\sigma_k^2 = \mathrm{Var}(U_k) = km'\xi_k(1-\xi_k)$. Let $c_k = k^{1/8}\sqrt{m'}(1-k^{-3/4})^{-1/2} > 0$. Then

$$P(a_k - c_k\sigma_k \leq U_k) \geq 1 - c_k^{-2} \quad .$$

It is easy to show that

$$a_k - c_k\sigma_k > 2m'k^{1/4}\left[1 - \sqrt{\frac{k^{3/4}-2}{2k^{3/4}-2}}\right] \quad \text{for all } k \quad .$$

Therefore $a_k - c_k\sigma_k \longrightarrow +\infty$ as $k \longrightarrow \infty$. For k sufficiently large we have $a_k - c_k\sigma_k > \frac{1}{\varepsilon}$. For such k, $P(U_k \geq \frac{1}{\varepsilon}) \geq P(U_k \geq a_k - c_k\sigma_k) \geq 1 - c_k^{-2}$. But c_k tends to ∞ with k and $P\left\{U_k \geq \frac{1}{\varepsilon}\right\} \longrightarrow 0$ as $k \longrightarrow \infty$. Hence, $g_k(S_k) \longrightarrow 0$ in probability as $k \longrightarrow \infty$ for every sequence of points $\theta_k \in B_k$.

We now apply Theorem (9.1) of Hoeffding [8], which asserts, "If $g_k(S_k) \longrightarrow 0$ implies $f_k(S_k) \longrightarrow 0$, then $g_k(S_k) \longrightarrow 0$ in probability implies $f_k(S_k) \longrightarrow 0$ in probability." Let $f_k(S_k) = F_k(y, S_k) - \Phi(y)$. Then, for every y and every sequence of points $\theta_k \in B_k$, we have

(4.2.4.3) $$F_k(y, S_k) \longrightarrow \Phi(y) \text{ in probability as } k \longrightarrow \infty \quad .$$

Now we apply Theorem (3.1) of Hoeffding [8], which may be stated in the form, "If (4.2.4.3) holds and $\Phi(y) = 1 - \alpha$ has the unique solution $y = y_0$, then $t_k^{(\gamma)}(X_k) \longrightarrow y_0$ in probability as $k \longrightarrow \infty$, for every sequence of points $\theta_k \varepsilon B_k$."

Let $t_k'(X_k) = t_k(X_k) - t_k^{(\gamma)}(X_k) + y_0$. Our test in (4.2.2) may be written in the form

$$(4.2.4.4) \qquad \psi = \left\{ \begin{array}{ll} 1 & \text{if } t_k' > y_0 \\ a(x) & \text{if } t_k' = y_0 \\ 0 & \text{otherwise} \end{array} \right\} .$$

Let $B_{2k} = B_k \bigcap \omega_2(dk^{-1/2})$ where $d > 0$ is fixed. The set B_{2k} is closed and there exist θ_k and θ_k' such that

$$\sup_{B_{2k}} P_\theta(t_k \leq x) = P_{\theta_k}(t_k \leq x) \quad ,$$

and

$$\sup_{B_{2k}} P_\theta(t_k' \leq x) = P_{\theta_k'}(t_k' \leq x) \quad .$$

Let $M_{B2k}(x,d) = P_{\theta_k'}(t_k' \leq x)$. We want to show that if $P_{\theta_k}(t_k \leq x) \longrightarrow H_2(x,d)$ as $k \longrightarrow \infty$, and x is a continuity point of $H_2(x,d)$, then $M_{B2k}(x,d) \longrightarrow H_2(x,d)$ as $k \longrightarrow \infty$. By definition of θ_k and θ_k' we have

$$P_{\theta_k'}(t_k' \leq x) \geq P_{\theta_k}(t_k' \leq x) \text{ for all } k \quad .$$

We have shown above that $t_k - t_k' \longrightarrow 0$ in probability for every sequence of points in B_k, and in particular for the sequence $\{\theta_k\}$ as defined above. Hence, by Theorem (20.6) of Cramér [4], $P_{\theta_k}(t_k \leq x) \longrightarrow H_2(x,d)$ implies $P_{\theta_k}(t_k' \leq x) \longrightarrow H_2(x,d)$, as $k \longrightarrow \infty$. Therefore, all the limit points of $P_{\theta_k'}(t_k' \leq x)$ are greater than or equal to $H_2(x,d)$. Suppose one of the limit points is actually greater than $H_2(x,d)$. Then there is a subsequence $\{j\}$ of $\{k\}$ and $\delta > 0$ such that

$$P_{\theta_j'}(t_j' \leq x) \longrightarrow H_2(x,d) + 2\delta \text{ as } j \longrightarrow \infty \quad .$$

Given any $\varepsilon > 0$ it is easy to show that for j sufficiently large,

$$P_{\theta_j'}(t_j' \leq x) \leq P_{\theta_j'}(t_j \leq x + \varepsilon) + \delta \quad .$$

For ε sufficiently small we have

$$\lim_{j \longrightarrow \infty} P_{\theta_j'}(t_j' \leq x) \leq H_2(x + \varepsilon,\ d) + \delta \quad ,$$

where x is assumed to be a continuity point of H_2. But H_2 is monotone non-decreasing in x and we obtain

$$\delta \leq H_2(x + \varepsilon, d) - H_2(x,d) \text{ for all sufficiently small } \varepsilon \ .$$

Hence, x is not a continuity point of $H_2(x,d)$ and we get a contradiction.

We will now show that $P_{\theta_k}(t_k \leq x)$ converges. It is easy to demonstrate that $P_{\theta_k}(D_k = 0) \longrightarrow 0$ as $k \longrightarrow \infty$ for any sequence of points θ_k in B_{2k}. By definition of t_k, we may therefore restrict ourselves to the consideration of $\sup\limits_{B_{2k}} P_\theta(W_k \leq xD_k) = P_{\theta_k}(W_k \leq xD_k)$, where the new sequence $\theta_k \varepsilon B_{2k}$ may differ from the old.

Let $\theta_{1k} = (1/2, 1/2 - dk^{-1/2})$. $\theta_{1k} \varepsilon B_{2k}$ for large enough k .
We have

(4.2.4.5) $P_{\theta_k}(W_k \leq xD_k) \geq P_{\theta_{1k}}(W_k \leq xD_k)$ for all k. It is easy to show, using Lemmas (3.5.1) and (3.5.5), that the right-hand side of the last inequality converges to $\bar{\Phi}(x - \frac{d}{\sigma}) > 0$, as $k \longrightarrow \infty$, where $\sigma^2(\xi) = \xi(1-\xi)(\frac{1}{m'} + \frac{1}{n'})$ and $\sigma^2 = \sigma^2(1/2)$.

We now assert that the limit points of $E_{\theta_k}(W_k)$ are bounded. Suppose the statement is false. Then there exists a subsequence $\{j\}$ of $\{k\}$ such that $E_{\theta_j}(W_j) \longrightarrow + \infty$ as $j \longrightarrow \infty$, and D_j tends to some constant in probability as $j \longrightarrow \infty$. By Lemma (3.5.2) this implies that $P_{\theta_j}(W_j \leq xD_j)$ tends to zero as $j \longrightarrow \infty$. Hence the left-hand side of (4.2.4.4) tends to zero (with j replacing k). But the right-hand side of (4.2.4.5) has a positive limit, and we obtain a contradiction. Hence the limit points of $E_{\theta_k}(W_k)$ are bounded.

As a consequence of the above argument there exists d', $0 < d < d' < \infty$, and a sequence of sets $C_k = \{(\xi,\varepsilon): 2k^{-3/4} \leq \xi \leq 1 - 2k^{-3/4}; d \leq \varepsilon \leq d'\}$, such that for $\eta = \xi - \frac{\varepsilon}{\sqrt{k}}$ we have,

$$\lim_{k \longrightarrow \infty} P_{\theta_k}(W_k \leq xD_k) = \lim_{k \longrightarrow \infty} \sup_{C_k} P_{\xi,\varepsilon}(W_k \leq xD_k) .$$

We can define a positive constant a, $a < 1/2$, so that $\sigma^2(a) <$

$\min \left[\frac{d^2}{16x^2}, \frac{d^2}{4} \Phi\left(x - \frac{d}{\sigma}\right) \right]$, where x and d are assumed fixed. Let $B = \left\{ (\xi,\varepsilon): a \leq \xi \leq 1 - a \right\} \cap C_k$, which contains the point $\xi = a$ and is independent of k for k sufficiently large. The characteristic function of $(W_k - E_\theta W_k) = (W_k - \varepsilon)$ is of the form given in Lemma (3.5.1) for $\theta \in B$. Hence, $P_{\xi,\varepsilon}(W_k \leq y) \longrightarrow \Phi\left(\frac{y-\varepsilon}{\sigma(\xi)}\right)$ uniformly in B as $k \longrightarrow \infty$.

An application of Chebyshev's inequality shows that given any $\Delta > 0$ and $\delta > 0$, we can take k large enough so that

$$P_{\xi,\varepsilon} \left\{ \left| D_k - \sigma(\xi) \right| > \Delta \right\} < \delta \quad \text{for all } (\xi,\varepsilon) \text{ in } C \quad ,$$

where $C = \lim\limits_{k \longrightarrow \infty} C_k$. By Lemma (3.5.5) this implies that, as $k \longrightarrow \infty$,

$$P_{\xi,\varepsilon}(W_k \leq xD_k) \longrightarrow \Phi\left(x - \frac{\varepsilon}{\sigma(\xi)}\right) \text{ uniformly in } B \quad .$$

The function $\Phi(x - y)$ is decreasing in positive y, and we have $\sup\limits_{B} \Phi\left(x - \frac{\varepsilon}{\sigma(\xi)}\right) = \Phi\left(x - \frac{d}{\sigma}\right)$. Let $H_2(x,d) = \Phi\left(x - \frac{d}{\sigma}\right)$. Let $\Delta = \frac{d}{4}|x|$ and $\delta > 0$. For k sufficiently large we have, $P_{\xi,\varepsilon}\left\{|D_k-\sigma(\xi)| < \Delta\right\} > 1 - \delta$ for all (ξ,ε) in C.

$$\begin{aligned} P_{\xi,\varepsilon}(W_k \leq xD_k) &\leq P_{\xi,\varepsilon}(W_k \leq |x| D_k) \\ &\leq P_{\xi,\varepsilon}(W_k - EW_k \leq |x| D_k - d) \\ &\leq P_{\xi,\varepsilon}\left\{ W_k - EW_k \leq |x|D_k - d, |D_k - \sigma(\xi)| < \delta \right\} + \delta \end{aligned}$$

$$\leq P_{\xi,\varepsilon}\left\{ W_k - EW_k \leq |x| \left[\sigma(\xi) + \Delta \right] - d \right\} + \delta .$$

Now suppose $(\xi,\varepsilon) \varepsilon C - B$. Then $\sigma^2(\xi) \leq \sigma^2(a) < \frac{d^2}{16x^2}$. Hence, for large enough k,

$$P_{\xi,\varepsilon}\left\{ W_k \leq xD_k \right\} \leq P_{\xi,\varepsilon}\left\{ W_k - EW_k \leq -\frac{d}{2} \right\} + \delta .$$

Now $V_{\xi,\varepsilon} W_k \longrightarrow \sigma^2(\xi)$ as $k \longrightarrow \infty$. Hence, for k large enough and δ small enough,

$$\operatorname*{Sup}_{C_k - B} P_{\xi,\varepsilon}\left\{ W_k \leq xD_k \right\} \leq 2\delta + \frac{4\sigma^2(a)}{d^2} < H_2(x,d) \quad .$$

We now apply Lemma (3.5.3a) and obtain

$$P_{\theta_k}(t_k \leq x) = \operatorname*{Sup}_{B_{2k}} P_{\theta}(t_k \leq x) \longrightarrow H_2(x,d) \text{ as } k \longrightarrow \infty \quad .$$

But $H_2(x,d)$ is everywhere continuous in x, and consequently, for every $d > 0$,

(4.2.4.6) $$P_{\theta'_k}(t'_k \leq x) = \operatorname*{Sup}_{B_{2k}} P_{\theta}(t'_k \leq x) \longrightarrow H_2(x,d) \text{ as } k \longrightarrow \infty \quad ,$$

where $H_2(x,d)$ is given by

(4.2.4.7) $$H_2(x,d) = \Phi\left(x - \frac{d}{\sigma}\right) \quad .$$

In virtue of (4.2.4.4) we can use arguments similar to those of Lemmas (3.1) and (3.2) of Chapter III, to show that $N_{B\psi}(\delta)$ is finite for every positive δ and becomes infinite as $\delta \longrightarrow 0$. We may now apply the proof of Theorem (4.1) of Hoeffding-Rosenblatt [9] to show that, asymptotically as $\delta \longrightarrow 0$, (4.2.4.8) $$N_{B\psi}(\delta) \sim (D/\delta)^2 ,$$

where $\Phi(y_0 - D/\sigma) = \beta$, $\Phi(y_0) = 1 - \alpha$, and $\sigma^2 = (1/4)(1/m' + 1/n')$. This asymptotic value is the same as that of $N(\delta)$ given in (3.8.4) of Chapter III.

We now consider $N_{A\psi}(\delta)$. We first note that for any sequence of points θ_k in $\omega_2(dk^{-1/2})$, we have $P_{\theta_k}(D_k = 0) \longrightarrow 0$ as $k \longrightarrow \infty$.

Let $W_k^{(\gamma)} = D_k t_k^{(\gamma)}$. Since we are concerned only with limits as $k \longrightarrow \infty$, we may assume that ψ is given by

$$\psi = \left\{ \begin{array}{ll} 1 & \text{if} \quad W_k > W_k^{(\gamma)} \\ a(s_k) & \text{if} \quad W_k = W_k^{(\gamma)} \\ 0 & \text{otherwise} \end{array} \right\} .$$

Omitting subscripts k we observe that if G is a random variable taking on the values $g \in \mathcal{G}$ with equal probabilities, then, $Et(Gx) = 0$ and $Et^2(Gx) = 1$. By Theorem (2.1) of Hoeffding [8], we conclude that $t^{(\gamma)} \leq [(1 - \alpha)/\alpha]^{1/2}$. Let $C = [2(\frac{1}{m'} + \frac{1}{n'})(1 - \alpha)/2]^{1/2}$ and $R_k = [(S_k/N)(1 - S_k/N)]^{1/2}$. Then,

$$E_\theta(1 - \psi) \leq P_\theta(W_k \leq W_k^{(\gamma)}) \leq P_\theta(W_k \leq CR_k), \text{ for all } k.$$

Let $E_k(d) = A_k \cap \omega_2(dk^{-1/2})$, where A_k is defined by (4.2.3.1) and d is fixed. It is an easy matter to show that, given $\varepsilon > 0$ and $\delta > 0$, we can take K large enough so that $k > K$ implies

$$P_{\theta_k}(|R_k| > \varepsilon) < \delta \quad \text{for all } \theta_k \in E_k .$$

Hence, given $\delta > 0$ we have,

$$P_{\theta_k}(W_k \leq CR_k) \leq P_{\theta_k}(W_k - EW_k \leq CR_k - d) \quad ,$$

$$P_{\theta_k}(W_k \leq CR_k) < P_{\theta_k}(W_k - EW_k \leq -\frac{d}{2}) + \delta \; ,$$

$$P_{\theta_k}(W_k \leq CR_k) < \frac{4V_{\theta_k}(W_k)}{d^2} + \delta, \text{ for all } \theta_k \varepsilon E_k \quad .$$

The right-hand side of the last inequality tends to δ as $k \longrightarrow \infty$, uniformly for $\theta_k \varepsilon E_k$, and δ is arbitrary. Hence

(4.2.4.9) $$\sup_{E_k(d)} E_\theta(1 - \psi) \longrightarrow 0 \text{ as } k \longrightarrow \infty \text{ for all } d > 0 \quad .$$

We wish to show that $\varepsilon > 0$ and δ sufficiently small implies $\frac{N_{A\psi}(\delta)}{N_{B\psi}(\delta)} < 1+\varepsilon$. Suppose this last statement is false. Then there exists a positive sequence $\{\delta_k\}$ converging to zero as $k \longrightarrow \infty$, and a positive quantity ε, such that for k sufficiently large,

$$\frac{N_{A\psi}(\delta_k)}{N_{B\psi}(\delta_k)} > 1 + \varepsilon \quad .$$

This implies $\delta_k^2 \, N_{A\psi}(\delta_k) > D^2$ for k sufficiently large, where D is defined by (4.2.4.8). Let $n_k + 1 = N_{A\psi}(\delta_k)$. Clearly, n_k tends to infinity with k, and for k sufficiently large we have $\delta_k > Dn_k^{-1/2}$. The risk functions are monotone non-increasing in δ, and by definition of $N_{A\psi}(\delta)$ we have, for k sufficiently large,

$$r_{An_k}(\psi, Dn_k^{-1/2}) \geq r_{An_k}(\psi, \delta_k) > b\beta \quad .$$

Considering the definition of the risk function we obtain

$$\sup_{E_k(D)} E_\theta(1 - \psi) > \beta \quad \text{for all } k,$$

which implies, from (4.2.4.9), that $\beta = 0$, contrary to our assumption. Hence, $N_\psi(\delta) = \text{Max}\,[N_{A\psi}(\delta), N_{B\psi}(\delta)] \sim N_{B\psi}(\delta)$ as $\delta \longrightarrow 0$. We have shown in (4.2.4.8) that $N_{B\psi}(\delta) \sim N(\delta)$ as $\delta \longrightarrow 0$. This completes the proof of Theorem (4.2.4).

(4.3) <u>The Asymptotic Efficiency of the Two-Sided Fisher-Tocher Test.</u>

We assume that $0 < \alpha, \beta$ and $\alpha + \beta < 1$. Our random variable X_k is the one defined in section (4.2.1). Let $\omega_1(\delta) = \omega_1 = \{\theta : \xi = \eta\}$, $\omega_2^*(\delta) = \{\theta : |\xi - \eta| \geq \delta\}$, and $\omega_2(\delta) = \{\theta : \xi - \eta \geq \delta\}$.

Let

$$\psi_1 = \begin{Bmatrix} 1 \text{ if } u - v < c_1(s) \\ a_1(s) \text{ if } u - v = c_1(s) \\ 0 \text{ otherwise} \end{Bmatrix},$$

$$\psi_2 = \begin{Bmatrix} 1 \text{ if } u - v > c_2(s) \\ a_2(s) \text{ if } u - v = c_2(s) \\ 0 \text{ otherwise} \end{Bmatrix}.$$

We suppose the definitions of ψ_1 and ψ_2 to parallel the definition of ψ in (4.1.1) so that each has size $\alpha/2$ in ω_1. The test we consider in this section is

$$\psi = \psi_1 + \psi_2 \qquad (4.3.1)$$

The rest of this section will be devoted to proving the following Theorem.

Theorem (4.3.2)

$\frac{N(\delta)}{N_{\cup}(\delta)} \longrightarrow 1$ as $\delta \longrightarrow 0$. In words, the two-sided Fisher-Tocher test is asymptotically efficient in the sense of Theorem (3.4A).

In proving the theorem we will use many of the results and most of the notation of the previous section in which we treated the one-sided Fisher-Tocher test. For instance, we write ψ_1 and ψ_2 by analogy with (4.2.2), in terms of the group $\mathcal{G}$.

$$\psi_1 = \begin{Bmatrix} 1 \text{ if } t(x) < t^{(\gamma_1)}(x) \\ a_1(x) \text{ if } t(x) = t^{(\gamma_1)}(x) \\ 0 \text{ otherwise} \end{Bmatrix},$$

$$\psi_2 = \begin{Bmatrix} 1 \text{ if } t(x) > t^{(\gamma_2)}(x) \\ a_2(x) \text{ if } t(x) = t^{(\gamma_2)}(x) \\ 0 \text{ otherwise} \end{Bmatrix}.$$

All the arguments leading to (4.2.4.3) hold here, and $t_k^{(\gamma_2)} \longrightarrow \Phi^{-1}(1 - \alpha/2) = y_0$ in probability, $t_k^{(\gamma_1)} \longrightarrow -y_0$ in probability, as $k \longrightarrow \infty$, for any sequence of points $\theta_k \in B_k \cap \omega_2^*(dk^{-1/2})$.

It is clear that $t_k(u,v) = -t_k(m-u, n-v)$, and we readily deduce that $r_k[(\xi,\eta), \psi, \delta] = r_k[(1-\xi,1-\eta), \psi, \delta]$. In what follows we may and will replace $\omega_2^*(\delta)$ by $\omega_2(\delta)$ everywhere, and then use all the definitions of (4.2.3.1).

Let $t_{1k} = \min(-t_k^{(\gamma_1)}, t_k^{(\gamma_2)})$ and $t_{2k} = \max(-t_k^{(\gamma_1)}, t_k^{(\gamma_2)})$. Let

$t_{3k} = \max(0, t_{1k})$ and $t_{4k} = \max(0, t_{2k})$. It is easy to show that

$$P_\theta(|t_k| < t_{3k}) \leq E_\theta(1-\psi) \leq P_\theta(|t_k| \leq t_{4k}) \ .$$

Let $t' = |t_k| - t_{3k} + y_0$, $t'' = |t_k| - t_{4k} + y_0$, and

$$\psi' = \begin{Bmatrix} 1 & \text{if } t_k' \geq y_0 \\ 0 & \text{otherwise} \end{Bmatrix} ,$$

$$\psi'' = \begin{Bmatrix} 1 & \text{if } t_k'' > y_0 \\ 0 & \text{otherwise} \end{Bmatrix} .$$

It follows from the definition of the risk function and from (4.2.3.1) that

(4.3.2.1) $$N_{B\psi'}(\delta) \leq N_{B\psi}(\delta) \leq N_{B\psi''}(\delta) \qquad .$$

We will obtain an asymptotic expression for $N_{B\psi}(\delta)$ by finding one for $N_{B\psi''}(\delta)$ and showing that $N_{B\psi'}(\delta)$ has the same asymptotic value.

We begin by considering $\sup\limits_{B_{2k}} P_\theta(|t_k| \leq x)$, where $B_{2k} = B_k \cap \omega_2(dk^{-1/2})$. We will show that for every positive d this supremum tends to a function $H_2(x,d)$ which is everywhere continuous in x. Since $t_k'' - t_k \longrightarrow 0$ in probability for every sequence of points θ_k in B_{2k}, as $k \longrightarrow \infty$, an argument like that following (4.2.4.4) will show that $\sup\limits_{B_{2k}} P_\theta(t_k'' \leq x) \longrightarrow H_2(x,d)$ as $k \longrightarrow \infty$.

For every sequence of points θ_k in B_{2k}, we have $P_{\theta_k}(D_k = 0) \longrightarrow 0$ as $k \longrightarrow \infty$. Hence, in all the limiting arguments, we may replace $P_{\theta_k}(|t_k| \leq x)$ by $P_{\theta_k}(|W_k| \leq xD_k)$ for $\theta_k \varepsilon B_{2k}$. Now B_{2k} is closed and there exists $\theta_k \varepsilon B_{2k}$ such that

$$(4.3.2.2) \qquad P_{\theta_k}(|W_k| \leq xD_k) = \operatorname*{Sup}_{B_{2k}} P_\theta(|W_k| \leq xD_k),$$

where d is fixed. θ_k is not necessarily unique.

In order to show that (4.3.2.2) converges, we begin by defining $\theta_{1k} = (1/2,\ 1/2 - dk^{-1/2}) \ \varepsilon\ B_{2k}$ for all k sufficiently large, and assuming $x > 0$. From the arguments following (4.2.4.5) we may conclude immediately that $P_{\theta_{1k}}(|W_k| \leq xD_k)$ has a positive limit as $k \longrightarrow \infty$, and also that the limit points of $E_{\theta_k}(W_k)$ are bounded. Hence, there exists d', $0 < d < d' < \infty$ and a sequence of sets $C_k = \left\{ (\xi,\varepsilon):\ 2k^{-3/4} \leq \xi \leq 1 - 2k^{-3/4},\ d \leq \varepsilon \leq d' \right\}$, such that for $\eta = \xi - \frac{\varepsilon}{\sqrt{k}}$ we have

$$(4.3.2.3) \qquad \lim_{k \longrightarrow \infty} P_{\theta_k}(|W_k| \leq xD_k) = \lim_{k \longrightarrow \infty} \operatorname*{Sup}_{C_k} P_{\xi,\varepsilon}(|W_k| < xD_k).$$

Let $\sigma^2(\xi) = \xi(1-\xi)(\frac{1}{m'} + \frac{1}{n'})$, and $\sigma^2 = \sigma^2(1/2)$. Choose the positive number a so that $\sigma^2(a) < \min\left\{ \frac{d^2}{16x^2},\ \frac{d^2}{4} h(x,d) \right\}$ where $h(x,d) = \Phi(\frac{d+x}{\sigma}) - \Phi(\frac{d-x}{\sigma})$. Let $B = C_k \cap \left\{ \xi : a \leq \xi \leq 1 - a \right\}$. By lemma (3.5.1), $P_{\xi,\varepsilon}(|W_k| \leq y) \longrightarrow \Phi(\frac{\varepsilon+y}{\sigma(\xi)}) - \Phi(\frac{\varepsilon-y}{\sigma(\xi)})$ as $k \longrightarrow \infty$, uniformly in $B (y > 0)$. Given any $\Delta > 0$, $\delta > 0$, we can find a value of K independent of (ξ,ε) $\varepsilon\ C = \lim_{k \longrightarrow \infty} C_k$, such that $P_{\xi,\varepsilon}\left\{ |D_k - \sigma(\xi)| > \Delta \right\} < \delta$ for all $k > K$.

Hence, by lemma (3.5.5), as $k \longrightarrow \infty$,

$$P_{\xi,\varepsilon}(\,|W_k| \leq xD_k) \longrightarrow \Phi\left(x + \frac{\varepsilon}{\sigma(\xi)}\right) - \Phi\left(-x + \frac{\varepsilon}{\sigma(\xi)}\right) \text{ uniformly in } B .$$

The function on the right in the last expression is decreasing in $\frac{\varepsilon}{\sigma(\xi)}$, and its maximum in B is $h(x,d)$. Now take $\Delta = \frac{d}{4x}$.

$$P_{\xi,\varepsilon}(\,|W_k| \leq xD_k) \leq P_{\xi,\varepsilon}(W_k \leq xD_k) ,$$

$$\leq P_{\xi,\varepsilon}\left\{W_k - EW_k \leq x\,[\,\sigma(\xi) + \Delta\,] - d\right\} + \delta ,$$

for large values of k. For such values of k we have,

$$P_{\xi,\varepsilon}(\,|W_k| \leq xD_k) \leq \delta + \frac{4V_{\xi\varepsilon}(W_k)}{d^2} ,$$

$$\operatorname*{Sup}_{C_k - B} P_{\xi,\varepsilon}(\,|W_k| \leq xD_k) \leq \delta + \frac{4}{d^2} \operatorname*{Sup}_{C_k - B} V_{\xi,\varepsilon}(W_k) .$$

The right-hand side of the above inequality converges to $\delta + \frac{4\sigma^2(a)}{d^2} <$ $h(x,d)$ as $k \longrightarrow \infty$, provided δ is small enough. We now apply lemma (3.5.3a), and conclude that for $x > 0$,

$$\operatorname*{Sup}_{B_{2k}} P_\theta(\,|W_k| \leq xD_k) \longrightarrow h(x,d) \quad \text{as } k \longrightarrow \infty \quad . \tag{4.3.2.4}$$

For $x < 0$, the left-hand side of the last expression clearly tends to zero. The same expression is monotone non-decreasing in x, and $h(0,d) = 0$. Hence,

$$\operatorname*{Sup}_{B_{2k}} P_\theta(\,|t_k| \leq x) \longrightarrow H_2(x,d) \quad \text{as} \quad k \longrightarrow \infty ,$$

where

$$(4.3.2.5) \quad H_2(x,d) = \begin{Bmatrix} 0 \text{ if } x \leq 0 \\ \Phi\left(\frac{d+x}{\sigma}\right) - \Phi\left(\frac{d-x}{\sigma}\right) \text{ if } x > 0 \end{Bmatrix} .$$

But $H_2(x,d)$ is everywhere continuous in x. Therefore,

$\sup_{B_{2k}} P_\theta(|t_k''| \leq x) \longrightarrow H_2(x,d)$ as $k \longrightarrow \infty$.

Now we may use arguments like those of lemmas (3.1) and (3.2) of Chapter III to show that $N_{B\psi''}(\delta)$ is finite for positive δ and becomes infinite as $\delta \longrightarrow 0$. By using the proof of Theorem (4.1) of Hoeffding - Rosenblatt [9] almost verbatim, we can now show that $N_{B\psi''}(\delta) \sim (D/\delta)^2$ as $\delta \longrightarrow 0$, where D is defined by $\Phi\left(\frac{D+y_0}{\sigma}\right) - \Phi\left(\frac{D-y_0}{\sigma}\right) = \beta$, and $y_0 = \Phi^{-1}(1 - \alpha/2)$.

It is clear that we can apply all of the foregoing arguments to the test ψ' and obtain the same result. Hence $N_{B\psi}(\delta)$ has the same asymptotic expression, and from (3.8.2) of Chapter III we get an identical result for $N(\delta)$. We have therefore shown that

$$(4.3.2.6) \qquad \frac{N(\delta)}{N_{B\psi}(\delta)} \longrightarrow 1 \quad \text{as } \delta \longrightarrow 0 \quad .$$

In order to complete the proof of Theorem (4.3.2) we have to show that for every sequence of δ's converging to zero, the limit points of $\frac{N_{A\psi}(\delta)}{N_{B\psi}(\delta)}$ are all less than or equal to unity. We will prove this last statement by the method of contradiction, first making some definitions and establishing some preliminary results.

For $G \in \mathcal{G}$ we have $Et_k(Gx) = 0$, $Et_k^2(Gx) = 1$, and by Theorem

(2.1) of Hoeffding [8] this implies $-\sqrt{\frac{2-\alpha}{\alpha}} \leq t_k^{(\gamma_1)} \leq t_k^{(\gamma_2)} \leq \sqrt{\frac{2-\alpha}{\alpha}}$.

Hence,

$$E_\theta(1-\psi) \leq P_\theta(t_k^{(\gamma_1)} \leq t_k \leq t_k^{(\gamma_2)}) \leq P_\theta(W_k \leq D_k \sqrt{\frac{2-\alpha}{\alpha}}) .$$

Let $R_k = \sqrt{(S_k/N)(1-S_k/N)}$ and $C = \sqrt{2(\frac{1}{m'} + \frac{1}{n'})(2-\alpha)/\alpha}$.

Then,

$$E_\theta(1-\psi) \leq P_\theta(W_k \leq CR_k) .$$

The result (4.2.4.9) holds, with the same supporting argument, and we have

(4.3.2.7) $$\sup_{E_k(\delta)} E_\theta(1-\psi) \longrightarrow 0 \quad \text{as} \quad k \longrightarrow \infty ,$$

where

$$E_k(\delta) = A_k \cap \omega_2(dk^{-1/2}) .$$

Now suppose there is a positive sequence $\{\delta_k\}$ tending to zero as $k \longrightarrow \infty$ and some $\varepsilon > 0$ such that $\frac{N_{A\psi}(\delta_k)}{N_{B\psi}(\delta_k)} > 1 + \varepsilon$ for k sufficiently large. That portion of Section (4.2) which follows (4.2.4.9) may be taken verbatim to yield $\beta = 0$, which is a contradiction. This completes the proof of Theorem (4.3.2).

(4.4) Asymptotic Efficiency of the Approximate χ^2 Test.

For the problem (4.3) of testing $\omega_1 = \{\theta : \xi = \eta\}$ against $\omega_2^* = \{\theta : |\xi - \eta| \geq \delta\}$, the approximate χ^2 test which Cochran [3] calls the X^2 test, is sometimes used. Let $X_k^2 = \{\sqrt{k}(u/m - v/n)\}^2/C_k^2$, where $C_k^2 = (\frac{1}{m'} + \frac{1}{n'}) \frac{s_k}{N}(1 - \frac{s_k}{N}) = \frac{N-1}{N} D_k^2$. Then $|X_k| = \sqrt{\frac{N}{N-1}}\, |t_k(x)|$.

If $y_0 = \Phi^{-1}(1 - \alpha/2)$, the X^2 test is

$$T = \begin{cases} 1 & \text{if } X_k^2 > y_0^2 \\ 0 & \text{otherwise} \end{cases},$$

or

$$T = \begin{cases} 1 & \text{if } |t_k(x)| > \sqrt{\frac{N-1}{N}}\, y_0 \\ 0 & \text{otherwise} \end{cases}.$$

Theorem (4.4.1)

The X^2 test is asymptotically efficient in the sense of Theorem (3.4A).

Proof of Theorem (4.4.1)

The theorem is an immediate consequence of Theorem (4.3.2) because of the fact that $\sqrt{\frac{N-1}{N}}\, y_0 \longrightarrow y_0$ as $k \longrightarrow \infty$. This completes the proof.

CHAPTER V

ASYMPTOTIC SOLUTIONS AS THE ERROR PROBABILITIES TEND TO ZERO

We will treat the case where the risk functions are defined as in (1.2) of Chapter I. We define $a = \beta/(\alpha+\beta)$, and will let α and β tend to zero in such a way that α/β remains fixed. For an a priori distribution $P(\theta)$, and bayes solution ϕ with respect to $P(\theta)$ we define:

(5.1.1) $\quad N_p(\alpha) =$ least integer k such that $r_k(P,\phi) \leq a\alpha$,

$N_c(\alpha) =$ least integer k such that $r_k(\phi) \leq a\alpha$,

$N(\alpha) =$ least integer k such that $r_k \leq a\alpha$.

It then follows that $N_p(\alpha) \leq N(\alpha) \leq N_c(\alpha)$ for all α. Our object is to choose P and ϕ so that $\dfrac{N(\alpha)}{N_c(\alpha)} \longrightarrow 1$ as $\alpha \longrightarrow 0$. In such case we will say that ϕ is asymptotically efficient as $\alpha \longrightarrow 0$.

We begin with some lemmas.

<u>Lemma (5.2)</u>

Let $0 < t_0 < 1$, $\alpha > 0$, $n(\alpha) \geq m(\alpha) \geq 0$ for all α. If

(a) $$\lim_{\alpha \to 0} \left(\frac{t^{m(\alpha)}}{\alpha}\right) = 0 \quad \text{whenever} \quad 0 < t < t_0 \quad ,$$

and

(b) $$\varliminf_{\alpha \to 0} \left(\frac{t_0^{n(\alpha)}}{\alpha}\right) > 0 \quad ,$$

then

(c) $$m(\alpha) \sim n(\alpha) \quad \text{and} \quad m(\alpha) \sim \frac{\log \alpha}{\log t_0} \quad \text{as} \quad \alpha \longrightarrow 0 \quad .$$

<u>Proof of Lemma (5.2)</u>

Let $w(\alpha) = \dfrac{\log \alpha}{m(\alpha)\log t_0}$ and $z(\alpha) = \dfrac{\log \alpha}{n(\alpha)\log t_0}$.

Then $z(\alpha) \leq w(\alpha)$ for all α. We can prove the lemma by showing that

(d) The limit points of $w(\alpha)$ are less than or equal to unity as $\alpha \longrightarrow 0$,

and

(e) The limit points of $z(\alpha)$ are greater than or equal to unity as $\alpha \longrightarrow 0$.

Suppose statement (d) is false. Then there is a sequence $\{\alpha_j\}$ tending to zero as $j \longrightarrow \infty$, and $\varepsilon > 0$, such that if $m_j = m(\alpha_j)$ we have

$$(5.2.1) \qquad \lim_{j \longrightarrow \infty} \left(\frac{t^{m_j}}{\alpha_j}\right) = 0 \quad \text{for all} \quad t \in (0,t_0) \quad ,$$

and $\lim_{j \longrightarrow \infty} w(\alpha_j) = 1 + 2\varepsilon$. For j sufficiently large we have $\log \alpha_j < m_j(1+\varepsilon)\log t_0 = m_j \log t$, for some $t \in (0,t_0)$. Hence, $t^{m_j}/\alpha_j > 1$ for j sufficiently large. This contradicts (5.2.1).

Now suppose statement (e) is false. Then there is a sequence $\{\alpha_j\}$ tending to zero as $j \longrightarrow \infty$, and a positive number b such that for $n_j = n(\alpha_j)$ we have

$$\lim_{j \longrightarrow \infty} \left(\frac{t_0^{n_j}}{\alpha_j}\right) = b \quad ,$$

and

$$\lim_{j \longrightarrow \infty} z(\alpha_j) < 1 \quad .$$

Let

$$\lambda = \begin{cases} \dfrac{\log(2/b)}{\log t_0} & \text{if } b \leq 2 \\ 0 & \text{if } b > 2 \end{cases} ,$$

and $n_j' = n_j + \lambda$. It then follows that

(5.2.2)
$$\lim_{j \to \infty} \left(\frac{t_0^{n_j'}}{\alpha_j} \right) \geq 2 .$$

Now λ is finite and $\dfrac{\log \alpha_j}{n_j' \log t_0}$ has the same limit points as $z(\alpha_j)$ as $j \to \infty$. Hence for j large enough, $\dfrac{\log \alpha_j}{n_j' \log t_0} < 1$, and $\alpha_j > t_0^{n_j'}$.

This implies $\dfrac{t_0^{n_j'}}{\alpha_j} < 1$ for all j, a contradiction of (5.2.2).

Lemma (5.3)

Let $0 < \delta < 1/2$, $-\delta < s < \delta$, $t(s) = \delta - s$, $g(t) = \left(\frac{1+t}{1-t}\right)^{(1/t-t)}$,

and $y(s) = \dfrac{\log g\{t(s)\}}{\log g\{t(-s)\}}$. It follows that

$$1 - 2\delta < y(s) < (1-2\delta)^{-1} .$$

Proof of Lemma (5.3)

For $s \in (-\delta,\delta)$ we have $t(s) \in (0,1)$ and $(1/t-t) > 0$. Hence $g\{t(s)\} > 1$ and $\frac{dg}{dt} = g(t)\, f(t)$ has the same sign as $f(t)$, where

$$\begin{aligned} t^2 f(t) &= (-1-t^2)\left\{\log(1+t)-\log(1-t)\right\} + (1-t^2)\left\{\frac{t}{1+t} - \frac{t}{1-t}\right\}, \\ &= (1+t^2)\left\{\log(1-t) + \log(1+t)\right\} + 2t , \\ &= \left\{-2t^3/3-2t^5/5- \dots\right\} + \left\{-2t^3 \;-2t^5/3- \;\dots\right\} < 0 . \end{aligned}$$

It follows that $g(t)$ is monotone decreasing in $t \varepsilon (0,1)$ and $g\{t(s)\}$ monotone increasing in $s \varepsilon (-\delta,\delta)$. Hence $y(s)$ is monotone increasing in $s \varepsilon (-\delta,\delta)$. It is clear that $y(s) = 1/y(-s)$, and the lemma will be proved if we can show that $\lim_{s \to \delta} y(s) \leq \frac{1}{1-2\delta}$, that is,

$$\frac{2}{\log \left[\left(\frac{1+2\delta}{1-2\delta}\right)^{1/2\delta - 2\delta} \right]} \leq \frac{1}{1-2\delta} .$$

Let $x = 2\delta$. then $0 < x < 1$, and we want to show that $\log(\frac{1+x}{1-x}) \geq 2x/(1+x)$. Since $x/(1+x) < x$ on $(0, 1)$ it is sufficient to observe that $(\frac{1+x}{1-x}) - e^{2x} > 0$ for x on $(0, 1)$. This completes the proof.

The main result of this chapter is a theorem in which the following definitions and assumptions will be used.

Let $X_1, X_2, \ldots$ ad. inf. be a sequence of real valued, independent, and identically distributed random variables, each of which has the moment generating function $M(t,\theta) = E_\theta e^{tX_1}$ for all θ in some parameter space Ω. Let L_1 and L_2 be closed subsets of Ω, $L_1 \subset \omega_1$, $L_2 \subset \omega_2$, and define

$$\text{(5.4.1)} \qquad \begin{aligned} M(c,\theta,t) &= e^{-ct} M(t,\theta) , \\ m(c,\theta) &= \inf_t M(c,\theta,t) , \\ m_i(c) &= \sup_{L_i} m(c,\theta), \text{ for } i = 1, 2 . \end{aligned}$$

Let $P(\theta)$ be an a priori distribution assigning probability $p' = p'(\alpha)$ to $\theta_1 \varepsilon L_1$ and probability $1 - p'$ to $\theta_2 \varepsilon L_2$. Let $S_k = X_1 + X_2 + \ldots + X_k$, and

$$(5.4.2) \quad \phi = \phi(s_k) = \left\{ \begin{array}{ll} 1 & \text{if } s_k > kc \\ 0 & \text{otherwise} \end{array} \right\} .$$

Assumption (a)

The test ϕ is bayes with respect to $P(\theta)$, that is, it minimizes

$$r_k(P,\phi') = ap'E_\theta\phi' + b(1-p')E_\theta(1-\phi'), \text{ for all } k,$$

where ϕ' is any test based on k observations, $a = \beta/(\alpha+\beta)$, and $b = \alpha/(\alpha+\beta)$.

Assumption (b)

$m_1(c) = m(c,\theta_1)$, $m_2(c) = m(c,\theta_2)$, and $0 < m_1(c) = m_2(c) = m(c) < 1$.

Assumption (c)

$$\sup_{\omega_i} r_k(\theta,\phi) = \sup_{L_i} r_k(\theta,\phi) \text{ for } i = 1, 2, \text{ and all } k.$$

Assumption (d)

$E_{\theta_1}X_1$ is finite for all $\theta \varepsilon \underline{(\,)}$. $\theta \varepsilon L_1$ implies $E_\theta X_1 \leq c$ and $\theta \varepsilon L_2$ implies $E_\theta X_1 \geq c$.

Assumption (e)

As α tends to zero, every limit point of $p'(\alpha)$ is less than unity.

Assumption (f)

$$N_p(\alpha) \longrightarrow \infty \quad \text{as} \quad \alpha \longrightarrow 0 .$$

Theorem (5.5)

If assumptions (a) through (f) hold, then the test ϕ is asymptotically efficient as $\alpha \longrightarrow 0$. That is,

$$\frac{N(\alpha)}{N_c(\alpha)} \longrightarrow 1 \quad \text{as} \quad \alpha \longrightarrow 0 \quad ,$$

where $N(\alpha)$ and $N_c(\alpha)$ are defined by (5.1.1). An asymptotic expression for $N(\alpha)$ is

$$N(\alpha) \sim \frac{\log \alpha}{\log \{m(c)\}} \quad \text{as} \quad \alpha \longrightarrow 0 \quad .$$

Preof of Theorem (5.5)

From assumption (d) and Theorem (1) of Chernoff [2] we have:

$$\theta \in L_1 \text{ implies } \quad P_\theta(S_k \geq kc) \leq \{m(c,\theta)\}^k ,$$

$$\theta \in L_2 \text{ implies } \quad P_\theta(S_k \leq kc) \leq \{m(c,\theta)\}^k .$$

Applying assumptions (c) and (b),

$$r_{1k}(\emptyset) = \sup_{\omega_1} r_k(\theta,\emptyset) \leq a \sup_{L_1} P_\theta(S_k \geq kc) \leq a \{m(c)\}^k ,$$

$$r_{2k}(\emptyset) = \sup_{\omega_2} r_k(\theta,\emptyset) = b \sup_{L_2} P_\theta(S_k \leq kc) \leq b \{m(c)\}^k .$$

We may assume without loss of generality that $b \leq a$, hence

$$r_k(\emptyset) = \sup_{\Omega} r_k(\theta,\emptyset) \leq a \{m(c)\}^k, \text{ for all } k .$$

Let $\quad N_c' = N_c'(\alpha) = N_c(\alpha) - 1$. Then for every α we have

$$a\alpha < r_{N_c'}(\emptyset) \leq a \{m(c)\}^{N_c'} .$$

Hence, $\{m(c)\}^{N_c'+1} > \alpha m(c)$ for all α, and

$$(5.5.1) \qquad \underline{\lim}_{\alpha \to 0} \frac{\{m(c)\}^{N_c(\alpha)}}{\alpha} \geq m(c) > 0 \ .$$

From Theorem (1) of Chernoff [2] we also obtain for all $0 < \varepsilon < m(c,\theta)$,

$$\lim_{k \to \infty} \frac{\{ m(c,\theta)-\varepsilon \}^k}{P_\theta \{S_k \leq kc\}} = 0 \text{ for all } \theta \in L_2 \ .$$

Hence, for $\theta = \theta_2$ and $0 < \varepsilon < m(c)$, assumption (b) implies that

$$(5.5.2) \qquad \lim_{k \to \infty} \frac{\{m(c) - \varepsilon\}^k}{P_{\theta_2}\{S_k \leq kc\}} = 0 \ .$$

By definition of $N_p = N_p(\alpha)$, we have

$$a\alpha \geq r_{N_p}(P,\emptyset) = ap' E_{\theta_1}\emptyset + b(1-p')E_{\theta_2}(1-\emptyset) \ ,$$

and $b(1-p')P_{\theta_2}(S_{N_p} \leq cN_p) \leq a\alpha$. Hence,

$$\frac{1}{P_{\theta_2}(S_{N_p} \leq cN_p)} \geq \frac{b(1-p')}{a\alpha} \quad \text{for all } \alpha > 0 \ .$$

By assumption (f), $N_p \to \infty$ as $\alpha \to 0$, and combining the last inequality with (5.5.2) we obtain

$$\frac{b}{a} \lim_{\alpha \to 0} \frac{[1-p'(\alpha)][m(c)-\varepsilon]^{N_p}}{\alpha} = 0 \ .$$

By assumption (e), every limit point of $1 - p'(\alpha) > 0$ as $\alpha \to 0$. Hence, for every $0 < \varepsilon < m(c)$,

$$(5.5.3) \qquad \lim_{\alpha \to 0} \frac{\{m(c)-\varepsilon\}^{N_p(\alpha)}}{\alpha} = 0 \quad .$$

Let $t_0 = m(c)$, $t = m(c)-\varepsilon$, $N_p(\alpha) = m(\alpha)$, and $N_c(\alpha) = n(\alpha)$. By assumption (a), $m(\alpha) \le n(\alpha)$ for all α. Then from (5.5.1) and (5.5.3) we see that the conditions of Lemma (5.2) are satisfied. Hence $\frac{N_p(\alpha)}{N_c(\alpha)} \to 1$ as $\alpha \to 0$, which implies that $\frac{N(\alpha)}{N_c(\alpha)} \to 1$ as $\alpha \to 0$. An asymptotic expression for $N_c(\alpha)$ and $N(\alpha)$ as $\alpha \to 0$, is given by Lemma (5.2) and is, $N(\alpha) \sim \frac{\log \alpha}{\log \{m(c)\}}$. This completes the proof of Theorem (5.5).

(5.6) An Application of Theorem (5.5) to a Particular Problem

Let $X_{11}, X_{12}, \ldots$ ad. inf., and $X_{21}, X_{22}, \ldots$ ad. inf., be mutually independent random variables as defined by

$$X_{1i} = \begin{cases} 1 & \text{with probability } \xi \\ 0 & \text{with probability } 1-\xi \end{cases} \quad \text{for } i = 1, 2, \ldots,$$

and

$$X_{2j} = \begin{cases} 1 & \text{with probability } \eta \\ 0 & \text{with probability } 1-\eta \end{cases} \quad \text{for } j = 1, 2, \ldots, .$$

Let $\theta = (\xi,\eta) \in \Theta = \{\theta: 0 \le \xi,\eta \le 1\}$, $0 < \delta < 1$,
$\omega_1 = \{\theta: \xi - \eta \le -\delta\}$, and $\omega_2 = \{\theta: \xi - \eta \ge \delta\}$.

On the basis of a sample of m X_1's and n X_2's, we want to find a test which is asymptotically efficient in the sense of Theorem (5.5). We will keep $\frac{m}{n}$ fixed as $\alpha \to 0$, and show that a test which depends only on a linear function of the sample percentages is asymptotically efficient. This linear function depends on the ratio $\frac{m}{n}$ only.

Let $m = km'$ and $n = kn'$, where m', n', and k are positive integers. Define:

$$(5.6.1) \qquad U_{1r} = \sum_{i=1}^{m'} X_{1[(r-1)m'+i]}, \quad \text{for} \quad r = 1, 2, \ldots,$$

$$U_k = \sum_{r=1}^{k} U_{1r}, \quad \text{for } k = 1, 2, \ldots,$$

$$V_{2r} = \sum_{j=1}^{n'} X_{2[(r-1)n'+j]}, \quad \text{for} \quad r = 1, 2, \ldots,$$

$$V_k = \sum_{r=1}^{k} V_{2r} \quad \text{for} \quad k = 1, 2, \ldots, \quad .$$

The linear function we wish to use in defining our test is $yU_k/m - V_k/n - c$, where y and c are constants. We will first define some constants y and c, and then show that the resulting test is asymptotically efficient, by demonstrating that the assumptions of Theorem (5.5) are satisfied.

Let ξ_0 be the unique root on the interval $[0, 1 - \delta]$, of the equation

$$(5.6.2) \quad g(\xi, m', n') = m'(\xi+\delta)(1-\xi-\delta)(1-2\xi) - n'\xi(1-\xi)\left\{2(\xi+\delta)-1\right\} = 0 \quad .$$

We have, $g(-\delta) = -n'\delta(1+\delta) < 0$, $g(0) = m'\delta(1-\delta) > 0$, $g(1/2-\delta) = m'\,\delta/2 > 0$, $g(1/2) = -\,n'\,\delta/2 < 0$, $g(1-\delta) = -\,n'\,\delta(1-\delta) < 0$. Hence ξ_0 exists, is unique, and lies on $(1/2 - \delta, 1/2)$. It is easy to show that $g'(-\delta) > 0$, $g'(1/2 - \delta) < 0$, and $g'(1/2) < 0$. Hence $g(\xi)$ is monotone decreasing on

$(1/2 - \delta, 1/2)$. For $m'' > 0$, $g(\xi, m' + m'', n') > g(\xi, m', n')$ for all $\xi \varepsilon (1/2 - \delta, 1/2)$. Hence ξ_0 is a monotone increasing function of $\frac{m'}{n'}$. In what follows we will assume that $m' > n'$. Then it is easy to verify that

$$(5.6.3) \qquad \frac{1-\delta}{2} < \xi_0 < \frac{1}{2} - \frac{n'\delta}{m'+n'} \quad .$$

Define:

$$(5.6.4) \qquad Q_1 = (1-\xi_0)/\xi_0 > 1 \quad ,$$

$$Q_2 = \frac{1-\xi_0-\delta}{\xi_0+\delta} < 1 \quad ,$$

$$y = y(\xi_0) = \frac{-m' \log Q_1}{n' \log Q_2} \quad .$$

From (5.6.2) we have $\frac{m'}{n'} = \left. \frac{\xi(1-\xi)\{2(\xi+\delta)-1\}}{(\xi+\delta)(1-\xi-\delta)(1-2\xi)} \right|_{\xi = \xi_0}$.

Let $\xi_0 = (1-\delta+s)/2$. Then $s \varepsilon (-\delta,\delta)$, and for $0 < \delta < 1/2$, $y = y(s)$ satisfies the conditions of lemma (5.3) and we have

$$(5.6.5) \qquad 1 - 2\delta < y < \frac{1}{1-2\delta} \ , \quad \text{for} \quad 0 < \delta < 1/2 \quad .$$

The random variable $X_k = (U_k, V_k)$ is a sufficient statistic for $\theta \varepsilon \underline{(\overline{\ })}$ and we will confine ourselves to tests based on this statistic. X_k has the frequency function $f_\theta(u, v; k) = \binom{m}{u}\xi^u(1-\xi)^{m-u}\binom{n}{v}\eta^v(1-\eta)^{n-v}$.

<u>Assumptions (a) and (e) are satisfied.</u>

Let

$$dP(\theta) = \begin{cases} 1/2 & \text{if } \theta = \theta_1 = (\xi_0, \xi_0 + \delta) \\ 1/2 & \text{if } \theta = \theta_2 = (1-\xi_0, 1-\xi_0 - \delta) \\ 0 & \text{otherwise} \end{cases} .$$

For any test $\emptyset'$ based on k observations of the form (u_{1r}, v_{2r}), the bayes risk with respect to P is

$$r_k(P,\emptyset') = 1/2 + 1/2 \sum_{u,v} \emptyset' \left[f_{\theta_1}(u,v;k) - f_{\theta_2}(u,v;k) \right] .$$

A test $\emptyset$ minimizing the above risk is

$$\emptyset = \begin{cases} 1 & \text{if } \left(\frac{1-\xi_0}{\xi_0}\right)^{2u} \left(\frac{\xi_0+\delta}{1-\xi_0-\delta}\right)^{-2v} > \left(\frac{1-\xi_0}{\xi_0}\right)^{m} \left(\frac{1-\xi_0-\delta}{\xi_0+\delta}\right)^{n} \\ 0 & \text{otherwise} \end{cases} .$$

Let $p = u/m$, $q = v/n$, and $y = y(\xi_0)$. Then

$$\emptyset = \begin{cases} 1 & \text{if } \left(\frac{\xi_0+\delta}{1-\xi_0-\delta}\right)^{2nyp-2nq} > \left(\frac{\xi_0+\delta}{1-\xi_0-\delta}\right)^{n(y-1)} \\ 0 & \text{otherwise} \end{cases} .$$

But $\left(\frac{\xi_0+\delta}{1-\xi_0-\delta}\right) = Q_2^{-1} > 1$. Hence, for $c = \frac{y-1}{2}$,

$$\emptyset = \begin{cases} 1 & \text{if } yp - q > c \\ 0 & \text{otherwise} \end{cases} .$$

Let $Z_r = yU_{1r}/m' - V_{2r}/n'$ and $S_k = \sum_{r=1}^{k} Z_r$. Then for $S_k = s_k$

we have

$$(5.6.6)\qquad \emptyset = \begin{Bmatrix} 1 & \text{if} & s_k > c \\ 0 & \text{otherwise} & \end{Bmatrix}, \text{ where } c = \frac{y-1}{2} .$$

Hence $\emptyset$ is bayes with respect to P and is of the proper form. Assumption (e) is satisfied because $p' = 1/2$ independently of α.

Assumption (b) is satisfied.

The moment generating function of Z_r is

$$M(t,\theta) = E_\theta e^{tZ_r} = [(1-\xi)+\xi e^{ty/m'}]^{m'} [(1-\eta)+\eta e^{-t/n'}]^{n'} .$$

Let $L_1 = \{\theta: \eta=\xi+\delta\}$ and $L_2 = \{\theta: \xi=\eta+\delta\}$. In L_1, $M(c,\theta,t)$ is of the form

$$M_1(y,\xi,t) = e^{-t(y-1)/2}[(1-\xi)+\xi e^{ty/m'}]^{m'}[(1-\xi-\delta)+(\xi+\delta)e^{-t/n'}]^{n'} .$$

In L_2, θ is of the form $(1-\xi-\delta, 1-\xi)$, and we may write $M(c,\theta,t)$ in the form

$$M_2(y,\xi,t) = e^{-t(y-1)/2}[(\xi+\delta)+(1-\xi-\delta)e^{ty/m'}]^{m'}[\xi+(1-\xi)e^{-t/n'}]^{n'}.$$

Now $y = y(\xi_0) = -m' \log Q_1/n' \log Q_2$, and we have

$$M_1(y,\xi,2n't \log Q_2) = [(1-\xi)Q_1^t+\xi Q_1^{-t}]^{m'} [(1-\xi-\delta)Q_2^t +(\xi+\delta)Q_2^{-t}]^{n'} ,$$

$$M_2(y,\xi',2n't \log Q_2) = [(\xi'+\delta)Q_1^t+(1-\xi'-\delta)Q_1^{-t}]^{m'} [\xi' Q_2^t+(1-\xi')Q_2^{-t}]^{n'}.$$

Now it follows that $\inf_t M_1(y,\xi,t) = \inf_t M_2(y,1-\xi-\delta,t)$, and taking the suprema for $\xi \in (0, 1-\delta)$ we obtain

$$m_1(c) = \sup_{L_1} \inf_t M(c,\theta,t) = \sup_{L_2} \inf_t M(c,\theta,t) = m_2(c).$$

To complete the proof that assumption (b) is satisfied we need only show that $\sup_\xi \inf_t M_1(y,\xi,t) = \inf_t M_1(y,\xi_0,t)$. Since y is fixed, we can simplify the notation by writing $M(\xi,t) = M_1(y,\xi,2n't \log Q_2)$.

Let $f(\xi) = [2\sqrt{\xi(1-\xi)}]^{m'} [2\sqrt{(\xi+\delta)(1-\xi-\delta)}]^{n'}$. What we want to show is that $\sup_\xi \inf_t M(\xi,t) = \inf_t M(\xi_0,t) = f(\xi_0)$.

By Schwarz's inequality, $M(\xi,t) \geq f(\xi)$ for all ξ and t. Hence

$$\sup_\xi \inf_t M(\xi,t) \geq \sup_\xi f(\xi) .$$

Now $f(0) = f(1-\delta)$, and it is easy to show that $f'(\xi)$ is proportional to $g(\xi)$ for $\xi \varepsilon (0, 1-\delta)$, where $g(\xi)$ is defined by (5.6.2). Now f is positive on $(0, 1-\delta)$ and $f'(\xi) = 0$ if and only if $\xi = \xi_0$. Hence $\sup_\xi f(\xi) = f(\xi_0)$, and

$$\sup_\xi \inf_t M(\xi,t) \geq f(\xi_0) .$$

Now $\sup_\xi \inf_t M(\xi,t) \leq \sup_\xi M(\xi,-1/2)$, and $\frac{d}{d\xi} M(\xi,-1/2)$ has the same sign as $U(\xi)$, where

$$U(\xi)=m'[Q_1^{1/2}-Q_1^{-1/2}][(1-\xi-\delta)Q_2^{-1/2}+(\xi+\delta)Q_2^{1/2}] +$$
$$n'[Q_2^{1/2}-Q_2^{-1/2}][(1-\xi)Q_1^{-1/2}+ \xi Q_1^{1/2}] .$$

But $U'(\xi) < 0$. Therefore, the vanishing point of $\frac{d}{d\xi} M(\xi,-1/2)$, if any, is unique, and represents a true maximum of $M(\xi,-1/2)$. Now

$(1/2)\ \sqrt{\xi_0(1-\xi_0)(\xi_0+\delta)(1-\xi_0-\delta)}\ \ U(\xi_0) = g(\xi_0) = 0$, which implies $U(\xi_0) = 0$. Hence

$$\sup_{\xi}\ \inf_{t}\ M(\xi,t) \leq M(\xi_0, -1/2) \quad .$$

It is easy to show that $(\xi,t) = (\xi_0, -1/2)$ implies that equality holds in Schwarz's inequality, and $M(\xi_0,-1/2) = f(\xi_0)$. Let

$$(5.6.7) \qquad m(c) = \sup_{0 < \xi < 1-\delta} \left[2\sqrt{\xi(1-\xi)}\right]^{m'} \left[2\sqrt{(\xi+\delta)(1-\xi-\delta)}\right]^{n'} .$$

It follows that $0 < m(c) = \sup_{\xi}\ \inf_{t}\ M(\xi,t) = M(\xi_0,-1/2) = \inf_{t} M(\xi_0,t) < 1$, and assumption (b) is satisfied.

Assumption (c) is satisfied.

We have shown in section (3.6.2) of Chapter III, that tests which are monotone non-decreasing in u_k and monotone non-increasing in v_k attain their maximum risks $r_{1k}(\phi)$ and $r_{2k}(\phi)$ on the boundaries L_1 and L_2. Hence assumption (c) is satisfied.

Assumption (d) is satisfied.

$E_\theta(yU_{1r}/m' - V_{2r}/n') = \mu$, clearly finite for all $\theta \in \Theta$.

Suppose $\delta \geq 1/2$. Then in L_1, $\mu = y\xi - \xi - \delta \leq y\xi - \xi - 1/2 \leq y(1-\delta) - 1/2 \leq \frac{y-1}{2} = c$. In L_2, $\mu = (y-1)\xi + y\delta \geq y\delta - \xi \geq \delta(y+1) - 1 \geq \frac{y-1}{2} = c$.

For $0 < \delta < 1/2$, θ in L_1 implies $\mu = (y-1)\xi - \delta$ and $(y-1)\xi - \delta \leq c = \frac{y-1}{2}$, because, by (5.6.5), we have $(1-2\delta) \leq y \leq (1-2\delta)^{-1}$. A similar argument shows that $\mu \geq c$ for θ in L_2.

Assumption (f) is satisfied.

Suppose the assumption is not satisfied. Then there exists a sequence $\{\alpha_j\}$ tending to zero as $j \longrightarrow \infty$, and a finite N, such that $N_p(\alpha_j) \longrightarrow N$ as $j \longrightarrow \infty$. For j large enough we have $N_p(\alpha_j) = N$ and

$$r_N(P,\phi) = \frac{a}{2} E_{\theta_1}\phi + \frac{b}{2} E_{\theta_2}(1-\phi) \leq \alpha_j \quad .$$

Taking the limit as $j \longrightarrow \infty$, we obtain

$$P_{\theta_1}\{ S_N > Nc \} = P_{\theta_2}\{ S_N \leq Nc \} = 0 \quad ,$$

which is clearly false. All the assumptions of Theorem (5.5) are satisfied and we obtain the following theorem.

Theorem (5.6.8)

For the problem of this section, the test ϕ defined by (5.6.6) is asymptotically efficient as $\alpha \longrightarrow 0$ and $\frac{\alpha}{\beta}$ remains fixed. The constant y is given by (5.6.4). An asymptotic expression for $N(\alpha)$ as $\alpha \longrightarrow 0$ is $N(\alpha) \sim \dfrac{\log \alpha}{\log \{ m(c) \}}$, where $m(c)$ is given by (5.6.7).

(5.7) Application of Theorem (5.5) to a Second Problem.

Let $X_{11}, X_{12}, \ldots$ ad. inf. and $X_{21}, X_{22}, \ldots$ ad. inf. be mutually independent random variables where

$$X_{1i} = \left\{ \begin{array}{l} 1 \text{ with probability } \xi \\ 0 \text{ with probability } 1-\xi \end{array} \right\} \text{ for } i = 1,2,\ldots ,$$

$$X_{2i} = \begin{cases} 1 \text{ with probability } \eta \\ 0 \text{ with probability } 1-\eta \end{cases} \quad \text{for } i = 1,2,\ldots , \quad .$$

Let $\theta = (\xi,\eta) \in \Theta = \{\theta: 0 \leq \xi, \eta \leq 1\}$, $0 < \delta < 1$, $\omega_1 = \{\theta: \xi-\eta \leq 0\}$, and $\omega_2 = \{\theta: \xi-\eta \geq \delta\}$. We want to find a test based on k observations from each population which is asymptotically efficient in the sense of Theorem (5.5), where the risk function is $r_k(\theta,\phi)$ as defined in (1.2) of Chapter I. Let $U_k = \sum_{i=1}^{k} X_{1i}$, $V_k = \sum_{i=1}^{k} X_{2i}$, and $Z_i = X_{1i} - X_{2i}$, for $i, k = 1, 2, \ldots$ ad. inf. We will show that a test based on the linear function $(\frac{u_k}{k} - \frac{v_k}{k} - c)$,for suitably chosen constant c, is asymptotically efficient, by showing that the conditions of Theorem (5.5) are satisfied.

<u>Assumptions (a) and (e) are satisfied.</u>

Let $p = \frac{\alpha}{\alpha+\beta}$ and define $P(\theta)$ by

$$P(\theta) = \begin{cases} p & \text{if } \theta = \theta_1 = (1/2,\ 1/2) \in \omega_1 \\ 1-p & \text{if } \theta = \theta_2 = (\frac{1+\delta}{2}, \frac{1-\delta}{2}) \in \omega_2 \\ 0 & \text{otherwise} \end{cases} \quad .$$

(U_k, V_k) is a sufficient statistic with frequency function $f_\theta(u,v;k) = \binom{k}{u}\xi^u(1-\xi)^{k-u}\eta^v(1-\eta)^{k-v}$, and we may restrict ourselves to tests depending on (U_k, V_k). For any test ϕ' based on k observations the bayes risk is

$$r_k(P,\phi') = apE_{\theta_1}\phi' + b(1-p)E_{\theta_2}(1-\phi') \quad ,$$

$$= \alpha\beta(\alpha+\beta)^{-2} \left[E_{\theta_1}\phi' + E_{\theta_2}(1-\phi') \right] .$$

A bayes solution is given by

$$\phi = \begin{cases} 1 \text{ if } f_{\theta_2}/f_{\theta_1} > 1 \\ 0 \text{ otherwise} \end{cases} ,$$

$$\phi = \begin{cases} 1 \text{ if } \left[\frac{1+\delta}{1-\delta} \right]^{u-v} > (1-\delta^2)^k \\ 0 \text{ otherwise} \end{cases} .$$

Define c by the equation

$$(5.7.1) \qquad \left[\frac{1+\delta}{1-\delta} \right]^{c} = (1-\delta^2)^{-1} .$$

It is easy to show that c is unique and $0 < c < \delta$. Let $S_k = \sum_{i=1}^{k} Z_i$.

Then for $S_k = s_k$,

$$(5.7.2) \qquad \phi = \begin{cases} 1 \text{ if } s_k > kc \\ 0 \text{ otherwise} \end{cases} ,$$

and assumption (a) is satisfied. Assumption (e) is satisfied because

$p = \frac{1}{1+\beta/\alpha}$ is fixed for $\frac{\alpha}{\beta}$ fixed.

<u>Assumption (b) is satisfied.</u>

Let $L_1 = \{\theta: \xi=\eta\} \subset \omega_1$ and $L_2 = \{\theta: \xi=\eta+\delta\} \subset \omega_2$. The moment generating function of Z_i is

$$M(t,\theta) = E_\theta e^{tZ_i} = (1-\xi+\xi e^t)(1-\eta+\eta e^{-t}) .$$

In L_1, $M(c,\theta,t)$ is of the form

$$M_1(c,\xi,t) = e^{-ct}(1-\xi+\xi e^{t})(1-\xi+\xi e^{-t}).$$

In L_2, $M(c,\theta,t)$ is of the form

$$M_2(c,\eta,t) = e^{-ct}[(1-\eta-\delta)+(\eta+\delta)e^{t}][1-\eta+\eta e^{-t}] \; .$$

We first consider $m_1(c) = \sup_{\xi} \inf_{t} M_1(c,\xi,t)$. $m_1(c) \geq \inf_{t} M_1(c,1/2,t)$. $\frac{d}{dt} M_1(c,1/2,t)$ has the same sign as $[-c+e^{t}/(1+e^{t}) - e^{-t}/(1+e^{-t})]$, and can vanish only if $e^{t} = \frac{(1+c)}{(1-c)}$. This is possible only if $-1 < c < 1$, which is true. The vanishing point is clearly a true minimum and we have

$$m_1(c) \geq (\tfrac{1}{4})\left[\frac{1-c}{1+c}\right]^{c}(1+\frac{1+c}{1-c})(1+\frac{1-c}{1+c}) = \left[\frac{1-c}{1+c}\right]^{c}(1-c^2)^{-1} \; .$$

$m_1(c) \leq \sup_{\xi} M_1(c,\xi,t_0)$, where $e^{t_0} = \frac{1+c}{1-c}$. It is easy to show that $\frac{d}{d\xi} M_1(c,\xi,t_0) = 0$ if and only if $\xi = 1/2$.

$$\frac{d^2}{d\xi^2} M_1(c,\xi,t_0) = 2e^{-ct_0}(\frac{1+c}{1-c} - 1)(\frac{1-c}{1+c} - 1) \quad ,$$

which has the same sign as $-c^2/(1-c^2)<0$ because $-1 < c < 1$. Hence $m_1(c) \leq M_1(c,1/2,\, t_0) = \left[\frac{1-c}{1+c}\right]^{c}(1-c^2)^{-1}$, and

$$m_1(c) = \left[\frac{1-c}{1+c}\right]^{c}(1-c^2)^{-1} \quad .$$

Now we consider $m_2(c) = \sup_{\xi} \inf_{t} M_2(c,\xi,t)$. $m_2(c) \leq \sup_{\xi} M_2(c,\xi,t')$, where $e^{t'} = (\frac{1+c}{1-c})(\frac{1-\delta}{1+\delta}) > 0$. It is not difficult to verify that $\frac{d}{d\xi} M_2(c,\xi,t') = 0$ if and only if $\xi = \frac{1-\delta}{2}$.

$$\frac{d^2}{d\xi^2} M_2(c,\xi,t') = 2e^{-ct'}[(e^{t'}-1)(e^{-t'}-1)] ,$$

which has the same sign as $(1-\cosh t') < 0$ because $t' \neq 0$. Hence

$$m_2(c) \leq M_2(c, \frac{1-\delta}{2}, t') .$$

$m_2(c) \geq \inf_t M_2(c, \frac{1-\delta}{2}, t)$. $\frac{d}{dt} M_2(c, \frac{1-\delta}{2}, t) = M_2(c, \frac{1-\delta}{2}, t)[-c + Qe^t/(1+Qe^t) - 1/(1+Qe^t)]$, where $Q = \frac{1+\delta}{1-\delta}$. This derivative vanishes only if $Qe^t = (\frac{1+c}{1-c})$, that is, if $t = t'$. This t' is clearly a true minimum point and we have $m_2(c) \geq M_2(c, \frac{1-\delta}{2}, t')$. Hence

$$m_2(c) = M_2(c, \frac{1-\delta}{2}, t') .$$

A little algebra using the definitions of t' and c reveals that $M_2(c, \frac{1-\delta}{2}, t') = (\frac{1-c}{1+c})^c (1-c^2)^{-1}$. Let

$$m(c) = (\frac{1-c}{1+c})^c (1-c^2)^{-1} . \tag{5.7.3}$$

We have shown that $0 < m(c) = \sup_{L_1} \inf_t M(c,\theta,t) = \sup_{L_2} \inf_t M(c,\theta,t)$, and $m(c) = \inf_t M(c,\theta_1,t) = \inf_t M(c,\theta_2, t)$. To complete the proof that assumption (b) is satisfied we need only show that $m(c) < 1$. Let $a_j = \frac{1}{j(2j-1)} > 0$ for $j = 1,2,\ldots,$. $\operatorname{Log} m(c) = c[\log(1-c) - \log(1+c)] -$

$\log(1-c^2) = -\sum_{j=1}^{\infty} c^{2j} a_j < 0$, and assumption (b) is satisfied.

Assumption (c) is satisfied.

The test ϕ is monotone non-decreasing in u_k and monotone non-increasing in v_k. We may again use the results in (3.6.2) of Chapter III to conclude that the maximum risks $r_{1k}(\phi)$ and $r_{2k}(\phi)$ occur on the boundaries L_1 and L_2.

Assumption (d) is satisfied.

$E_\theta Z_1 = 0$ for θ in L_1 and $E_\theta Z_1 = \delta$ for θ in L_2. To show that assumption (d) is satisfied we merely point out that $0 < c < \delta$.

Assumption (f) is satisfied.

The corresponding argument of Section (5.6) applies directly to the present case.

All the assumptions of Theorem (5.5) are satisfied in this case and we present the results in the form of a theorem.

Theorem (5.7.4)

For the problem of this section, the test ϕ defined by (5.7.1) and (5.7.2) is asymptotically efficient as $\alpha \to 0$ and α/β remains fixed. An asymptotic expression for $N(\alpha)$ as $\alpha \to 0$ is

$$N(\alpha) \sim \frac{\log \alpha}{\log\{m(c)\}}$$, where $m(c)$ is given by (5.7.3).

APPENDIX

1. General

The expression $E_{\xi,\eta}(\phi - \psi)$, for fixed ξ, is a polynomial in η which we can write as $P_\xi(\eta) = p_k(\xi)(1-\eta)^k + p_{k-1}(\xi)\eta(1-\eta)^{k-1} + \ldots + P_0(\xi)$. If $P_1(\eta) = 0$ for some $0 < \eta \leq 1$, then such η is unique and we call it $\eta(1)$. Otherwise, let $\eta(1) = 0$. We define $\eta(\xi)$ by the relations: $1-\xi \leq \eta \leq 1$ and $P_\xi(\eta) = 0$. The considerations of (2.7.6) tell us that $\eta(\xi)$ exists for $1/2 \leq \xi \leq 1$, and that $\eta(1) \neq 0$ implies $\eta(\xi)$ is unique for $1/2 \leq \xi \leq 1$. $P_1(\eta)$ is given by

$$\sum_{j=0}^{k} \left[E_j (1-\eta)^j \eta^{k-j} \right] ,$$

where E_j is defined in section (2.8). If $\eta(1) \neq 0$, there are no points θ in $\left\{ \theta: \xi < \eta, (\xi,\eta) \neq (0,1) \right\}$ for which $E_\theta(\phi - \psi)$ vanishes.

In determining $\eta(\xi)$ for values $k = 6$ and $k = 10$, we determine two values, $\eta' = c/100$ and $\eta'' = (c+1)/100$, such that $P_\xi(\eta') < 0 < P_\xi(\eta'')$, where $0 < c < 100$ is an integer. Our approximate value of $\eta(\xi)$ is then $\eta(\xi) = (10c + c^*)/1000$, where c^* is obtained by linear interpolation. The polynomials $P_\xi(\eta)$ are not presented here.

In determining Δ_0 (see section 2.8) for $k > 1$, we solve equation (2.8.1) for y_0, by the same method we used in finding $\eta(\xi)$. Then we round y_0 off to two significant figures, and set $\Delta_0 = (1-y_0)/(2+2y_0)$. We obtain $S = \left[y_0/(1+y_0)^2 \right]^k \sum_k D_j y_0^{-j}$ by substitution.

We present below the constants used in determining $\eta(1)$ and y_0. For $k > 1$, the columns headed Δ_j show that conditions (2.8.2) are satisfied, and y_0 is unique.

2. <u>(k=1)</u>

Let $T(u,v) = \phi(u,v) - \psi(u,v)$. Then $0 < \alpha \leq 1/4$ implies $T(0,0) = T(1,1) = -\alpha$, $T(0,1) = 0$, and $T(1,0) = 2\alpha$. $1/4 \leq \alpha < 1/2$ implies $T(0,0) = T(1,1) = \alpha - 1/2$, $T(0,1) = 0$, and $T(1,0) = 1 - 2\alpha$. Hence the function $\eta(\xi)$ is independent of α, and is given by

$$\eta(\xi) = (3\xi - 1)/(4\xi - 1)$$

For $0 < \alpha < 1/4$, $E_{\theta_\Delta} T$ is increasing in Δ, and $\Delta_0 = 1/2$. In such case, $S = E_{1,0}(\phi - \psi) = 2\alpha$. $\alpha = .05$ implies $S = .10$.

3. <u>(k=3)</u>

j	ψ_j	ϕ_j	D_j	A_j	E_j
0	.10		-.1	.7	-.05
1	.60		-.6	-4.7	-.30
2	1.50	2.2	.7	2.1	.35
3	1.00	1.0		.3	
Totals	3.2	3.2	0	-1.6	0

$\eta(1) \neq 0$ is given by $h(1-\eta) = -.05\eta^3 - .3\eta^2(1-\eta) + .35\eta(1-\eta)^2 = 0$, and $\eta(1) = 1/2$. The function $\eta(\xi)$ is determined by semi-graphical means from the relation

$$\frac{h(\xi)}{\xi^3} = \frac{-h(1-\eta)}{(1-\eta)^3} \quad .$$

$f(y_0) = A_0 + \ldots + A_3 y_0^3$, $y_0 \approx .16$, $y_0^{-1} \approx 6.25$, $y_0/(1+y_0)^2 \approx .119$, and

$$\Delta_0 \approx .345,$$
$$S \approx (.119)^3 \sum_j D_j (6.25)^j \approx .0396 .$$

4. <u>(k=6)</u>

j	ψ_j	ϕ_j	D_j	A_j	E_j
0	.1		-.1	109.8	-.05
1	1.2		-1.2	-470.6	-.60
2	35.3		-35.3	276.4	-3.30
3	89.2	125.8	36.6	7.8	(24/55)
4	66.0	66.0		.6	
5	12.0	12.0			
6	1.0	1.0			
Totals	204.8	204.8	0	-76.0	XX

(4.1) $\eta(1) \approx .112$

Here, $y_0 \approx .28$, $1/y_0 \approx 3.57$, $y/(1+y)^2 \approx .17$, and

$$\Delta_0 \approx .284,$$
$$S \approx (.17)^6 \sum_j D_j (3.57)^j \approx .0296.$$

5. (k=10)

j	D_j	A_j	E_j
0	-.1	57,804.0	-.05
1	-2.0	-201,966.8	-1.00
2	-83.5	123,929.2	-9.50
3	-9,548.4	984.0	-57.00
4	9,634.0	21.0	$210\left\{\frac{30,728.8}{38,760}-1\right\}$
5	-	1.0	-
Totals	0	-19,227.6	

Since all the E_j are negative, we have

(5.1) $$\eta(1) = 0$$

Here, $y_0 \approx .37$, $1/y_0 \approx 2.7$, $y/(1+y)^2 \approx .197$, and

$$\Delta_0 \approx .230,$$

$$S \approx (.197)^{10} \sum_j D_j (2.7)^j \approx .0285 .$$

BIBLIOGRAPHY

[1] Barnard, G. R., "Significance Tests for 2 x 2 Tables", Biometrika, 34(1947), 123-138.

[2] Chernoff, H., "A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations", Annals of Mathematical Statistics, 23(1952), 493-507.

[3] Cochran, W.G., "The χ^2 Test of Goodness of Fit", Annals of Mathematical Statistics, 23(1952), 315-345.

[4] Cramer, H., Mathematical Methods of Statistics, Princeton, Princeton University Press, 1946.

[5] Fisher, R.A., Statistical Methods for Research Workers, 5th and subsequent editions, Edinburgh, Oliver and Boyd Ltd., 1934.

[6] Fisher, R.A., The Design of Experiments, Edinburgh, Oliver and Boyd Ltd., 1935.

[7] Hoeffding, W., "A Combinatorial Central Limit Theorem", Annals of Mathematical Statistics, 22(1951), 558-566.

[8] Hoeffding, W., "The Large-Sample Power of Tests Based on Permutations of Observations", Annals of Mathematical Statistics, 23(1952), 169-192.

[9] Hoeffding, W., and Rosenblatt, J.R., "The Efficiency of Tests", Annals of Mathematical Statistics, 26(1955), 52-63.

[10] Lehmann, E. L., "Some Principles of the Theory of Testing Hypotheses", Annals of Mathematical Statistics, 21(1950), 1-26.

[11] Parzen, E., "On Uniform Convergence of Sequences of Random Variables", University of California Publications in Statistics, 2(1954), 23-52.

[12] Pearson, K., "On the Criterion that a Given System of Deviation from the Probable in the Case of a Correlated System of Variables Is Such That It Can Be Reasomably Supposed to Have Arisen from Random Sampling", Philosophy Magazine, Series 5, 50(1900), 157-172.

[13] Tables of the Exponential Function, National Bureau of Standards Applied Mathematics Series, 14 (1951).

[14] Tocher, K.D., "Extension of the Neyman-Pearson Theory of Tests to Discontinuous Variates", Biometrika, 37(1950), 130-144.

[15] Uspensky, J. V., Introduction to Mathematical Probability, New York and London, McGraw-Hill Book Company, Inc., 1937.