ON A THEORY OF SIGNIFICANCE TESTING

By

BILLY J. MOORE

Bachelor of Science
Oklahoma State University
Stillwater, Oklahoma
1965

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1967

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
May, 1969

ON A THEORY OF SIGNIFICANCE

TESTING

Thesis Approved:

J. Leroy Folks
Thesis Adviser

Carl E. Marshall

Robert D. Morrison

David E. Bee

James E. Shamblin

D. D. Durham
Dean of the Graduate College

ACKNOWLEDGMENT

I wish to express my sincere gratitude to Dr. J. Leroy Folks for serving as Chairman of my advisory committee and directing the preparation of this thesis.

I extend special appreciation to Dr. Carl E. Marshall, Director of the Statistical Laboratory, for serving on my advisory committee and for counseling me during most of my college career.

I further express appreciation to Dr. Robert D. Morrison, Dr. David E. Bee, and Dr. James Shamblin for serving on my advisory committee.

Acknowledgment is due also to the National Aeronautics and Space Administration for providing financial assistance under grant number NASA NSG(T) - 67 - S2 9821 and the Department of Mathematics and Statistics for providing a graduate assistantship.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

A mathematical theory of significance testing in which test statistics are derived and compared by methods paralleling those of Neyman and Pearson hypothesis testing has been investigated by Finley (5). His formulation of the theory is based on an approach to significance testing suggested by Dempster and Schatzoff (4). The present study is intended to further the development of this theory.

Significance testing is a relatively old concept in statistics, the history dating back as far as 1735 when Bernoulli used the procedure in studying hypotheses in astronomy. Fisher (6), however, is given credit for introducing in 1925 the term "test of significance" and promoting its general use in scientific research. He described a procedure for assessing the significance of an apparent discrepancy between the observations and the "null" hypothesis H_0 , and this was done without mention of alternative hypotheses to H_0 . It appears the basic ingredients to the test of significance, according to Fisher and his followers, are the sample space, the observations, and the hypothesis under question. It was argued, and still is (Anscombe (1), Dempster and Schatzoff), that no formal decision rules be incorporated into significance testing.

A general theory of hypothesis testing was formally presented by Neyman and Pearson (9) in 1933. Since that time, this theory has become predominant in the teaching of statistics. There are several major differences between the hypothesis testing theory of Neyman and Pearson and the significance testing of Fisher. Hypothesis testing requires adherence to strict formal decision rules, and it insists upon the formal statement of an hypothesis alternative to H_0. In other words, the researcher is compelled either to accept the null hypothesis (a term evidently borrowed from Fisher) or to reject it in favor of the alternative, according to some mathematical function of the observed data.

Finley wrote that although Neyman-Pearson theory is almost universally accepted over significance testing in statistical textbooks, many applied statisticians and researchers actually evaluate their hypothesis problems by computing significance levels, i.e., performing significance tests. More than likely, this is done by using test functions derived by Neyman-Pearson theory, which in this case is not truly applicable. Therefore, it seems reasonable that if significance testing is going to be done in practice, then the mathematical development of a theory of significance testing is in order.

The formulation of the present theory is based on the paper of Dempster and Schatzoff. A test statistic is defined so that small values of the statistic are inconsistent with the null hypothesis, and the distribution of the statistic is known exactly under the null hypothesis. For a particular hypothesis problem a class A (not necessarily unique) of achievable significance levels is procured in some manner. The search is then undertaken for a test statistic which has the largest power, in the sense of Neyman and Pearson, for each significance level in A. For the class A this statistic is termed "best" for the hypothesis problem.

It is recalled that in classical hypothesis testing all α - levels in the interval (0, 1) are achievable, even when the probability density is discrete. This is due to the admission of an extraneous random experiment which permits tests to be of exact size α . Since this independent experiment contains absolutely no information about the null hypothesis and contributes nothing to the computation of a significance level, this randomization procedure appears to have no place in significance testing. Therefore, all α are not in general achievable, and the set A defined above need not contain all values in the unit interval.

One of the first steps in selecting a significance test for a hypothesis problem is the selection of a statistic on which to base the test statistic. It is argued by Lehmann (8') that if randomization is permitted, as it is in Neyman-Pearson theory, there is no loss in generality in restricting consideration to a sufficient statistic. Lehmann states, "Given any procedure based on x , it is possible to construct an equivalent one ... which can be viewed as a randomized procedure based solely on the sufficient statistic." Thus, if randomization is permitted in the theory, and if a sufficient statistic exists, then test statistic candidates can be restricted to those based on the sufficient statistic. However, since the independent randomized experiment is not used in significance testing, then no such justification for basing tests on sufficient statistics is afforded the theory.

Finley recognized the need for more investigation into the role of the sufficient statistics in significance testing. Working with the one-parameter exponential family of densities, $p_\theta(x) = C(\theta)\exp[\theta T(x)] \cdot h(x)$, he found that optimum test statistics $T^*(x)$ were indeed based on the sufficient statistic $T(x)$ for one-sided alternatives. He remarked that this result did not seem to depend on the properties of a sufficient statistic. However, the optimum test statistic was the likelihood ratio statistic, and the likelihood ratio depends on the observations only through the sufficient statistic. Assume $T(x)$ is sufficient for θ, and $p_\theta(x)$ is the density of x. By the factorization criterion there exists a factorization such that $p_\theta(x) = g_\theta[T(x)] \cdot h(x)$ where the first factor may depend on θ but depends on x only through $T(x)$, and $h(x)$ is independent of θ. The ratio $p_{\theta_1}(x)/p_{\theta_2}(x)$ then depends on the observations only through $T(x)$.

Even with the above correspondence between the likelihood ratio statistic and the sufficient statistic for one sided alternatives, there still remain unsolved similar problems for the two-sided alternative hypothesis $(\theta \neq \theta_0)$. In studying the two-sided case with the one-parameter exponential family, Finley derived two examples of unbiased test statistics, both unimodal functions of the sufficient statistic. He suspected both test statistics to enjoy optimum properties, but he did not mathematically justify his suspicions. He suggested that further research be done on the necessity and sufficiency of optimum unbiased test statistics to be unimodal functions of the sufficient statistics.

It has also been suggested that more work is needed on the concept of unbiasedness in general for significance testing. In particular, the usefulness of unbiased test statistics in significance testing is not clear when the probability density is discrete. It has been shown that its application is quite restricted in this case. Another question is the necessity and sufficiency of two-tail unbiased test statistics for two-sided alternatives.

A Brief Review of Significance Testing

Let X denote a random variable, either vector or scalar, and assume X has a probability density $f_\theta(x)$, or cumulative distribution function (c.d.f.) given by $F_\theta(x)$. The parameter θ may be a vector or scalar belonging to some parameter space Θ. The null and alternative hypotheses are given by

$$H_0:\ \theta \in \Theta_0 \quad \text{where } \Theta_0 \subset \Theta \text{, and}$$

$$H_A:\ \theta \in \Theta_A \quad \text{where } \Theta_A \subset \Theta \text{ .}$$

It is required that $\Theta_0 \cap \Theta_A = \phi$ and $\Theta_0 \cup \Theta_A \subset \Theta$.

Let $T^*(x)$ denote a test statistic calculated from X with c.d.f. $G_\theta(t^*)$. It is required that $G_\theta(t^*)$ be completely specified for $\theta \in \Theta_0$; in other words, the null hypothesis must be simple as far as $T^*(x)$ is concerned. The test statistics are chosen so that small values of the statistic are inconsistent with the null hypothesis. Then the significance level associated with $T^*(x)$, denoted by $SL(T^*)$, is given by

$$\begin{aligned} SL(T^*) &= G_0(t^*) \\ &= P_0[T^* \le \text{observed}] \\ &= \alpha \text{, say .} \end{aligned}$$

Since the significance level is a random variable, it has a distribution function, which is denoted by

$$H^*_\theta(\alpha) = P_\theta[SL(T^*) \le \alpha]$$
$$= P_\theta[T^*(x) \le t^*_\alpha]$$

where $t^*_\alpha = G_0^{-1}(\alpha)$ for all achievable α. If $\theta \varepsilon \Theta_0$ then $H^*_\theta(\alpha)$ will be written as $H^*_0(\alpha)$. It will be understood that only non-trivial achievable significance levels $(\alpha \neq 0, 1)$ will be considered unless otherwise indicated.

The significance test corresponding to $T^*(x)$ is said to be unbiased if

$$H^*_\theta(\alpha) \ge \alpha , \quad \theta \varepsilon \Theta_A ,$$

and $T^*(x)$ is said to be an unbiased test statistic. In this study unbiased significance tests will be considered only for hypothesis problems with the two-sided alternatives, $H_A: \theta \neq \theta_0$.

Statement of Problem

The purpose of this investigation is to continue the development of a theory of significance testing. The emphasis is on properties and characteristics of unbiased significance tests for the one-parameter case, and on significance testing in general for the two-parameter normal distribution.

In Chapter II, examples of unbiased test statistics in both the discrete and continuous case are given. In Chapters III and IV an investigation is made on necessary and sufficient conditions for unbiased test statistics. Chapter V considers null hypotheses for one of the two parameters of the normal distribution when the other parameter is assumed unknown.

CHAPTER II

EXAMPLES OF UNBIASED TEST STATISTICS

The basic purpose of this chapter is to give examples of unbiased test statistics to which reference may be made in later chapters. In his study, Finley gave several examples concerning unbiased test statistics. He showed that, without the aid of the randomized test of Neyman-Pearson theory, there did not exist, in general, two-tail test statistics for the parameters of the binomial and Poisson densities. It is clear that the development of this new theory of significance testing will be somewhat restricted in the discrete cases; it is not clear, though, if the theory will be restricted in the continuous cases. It is hoped the examples presented here will help the theory.

Continuous Case

Example 2.1: Consider a random sample of size n from $N(0, \sigma^2)$ and the hypothesis problem $H_0: \sigma = \sigma_0$ versus $H_A: \sigma \neq \sigma_0$. Finley shows an unbiased test statistic may be based on the sufficient statistic $T(x) = \Sigma X_i^2$, where $T(x)/\sigma_0^2$ has the chi-square distribution with n degrees of freedom. The test statistics he proposed is

$$T^*(t) = t^{n/2} \, e^{-t/2\sigma_0^2}$$

We note that $T^*(t)$ is unimodal in t for all n even though for $n \leq 2$ the density of t is not unimodal in t.

Example 2.2: Consider a random sample of size n from $N(\mu, 1)$ and the hypothesis problem $H_0: \mu = \mu_0$ versus $H_A: \mu \neq \mu_0$. The assertion is that an unbiased test statistic is given by

$$T^*(\overline{x}) = e^{-\frac{n(\overline{x} - \mu_0)^2}{2}} .$$

To show this is true we must show $H_\mu^*(\alpha)$ is minimized by setting $\mu = \mu_0$.

Since $T^*(\overline{x})$ is a unimodal function of $\overline{x}$ about $\overline{x} = \mu_0$, we have

$$\begin{aligned} H_\mu^*(\alpha) &= P_\mu[T^*(\overline{x}) \leq t_\alpha^*] \\ &= P_\mu[\overline{X} \leq c_1] + P_\mu[\overline{X} \geq c_2] \\ &= 1 - \int_{c_1}^{c_2} f_\mu(\overline{x})\, d\overline{x} . \end{aligned}$$

Taking the derivative of $H_\mu^*(\alpha)$ with respect to μ, we obtain

$$\begin{aligned} \frac{dH_\mu^*(\alpha)}{d\mu} &= -\int_{c_1}^{c_2} n(\overline{x} - \mu)\sqrt{\frac{n}{2\pi}}\, e^{-\frac{n}{2}(\overline{x} - \mu)^2}\, d\overline{x} \\ &= \left\{\sqrt{\frac{n}{2\pi}}\, e^{-\frac{n}{2}(\overline{x} - \mu)^2}\right\}_{c_1}^{c_2} \\ &= \sqrt{\frac{n}{2\pi}} \left\{ e^{-\frac{n}{2}(c_2 - \mu)^2} - e^{-\frac{n}{2}(c_1 - \mu)^2} \right\} \end{aligned}$$

which equals zero when $\mu = \mu_0$. Hence we have

$$\left.\frac{dH^*_\mu(\alpha)}{d\mu}\right|_{\mu=\mu_0} = 0 \ .$$

To ascertain a minimum does occur at $\mu = \mu_0$, we obtain the second derivative,

$$\frac{d^2H^*_\mu(\alpha)}{d\mu^2} = \frac{\sqrt{n}}{2\pi}\left\{n[c_2-\mu]e^{-\frac{n}{2}(c_2-\mu)^2} - n[c_1-\mu]e^{-\frac{n}{2}(c_1-\mu)^2}\right\}$$

Since $c_2 > c_1$ and $e^{-\frac{n}{2}(c_1-\mu_0)^2} = e^{-\frac{n}{2}(c_2-\mu_0)^2}$, it follows that

$$\left.\frac{d^2H^*_\mu(\alpha)}{d\mu^2}\right|_{\mu=\mu_0} > 0 \ ;$$

hence $H^*_\mu(\alpha)$ is minimized at $\mu = \mu_0$, and $T^*(\bar{x})$ is an unbiased test statistic.

<u>Example 2.3</u>: Consider the density

$$\begin{aligned} f_\theta(x) &= 1 + \theta(x - 1/2) \ , & 0 \le x \le 1 \ , \\ &= 0 & \text{elsewhere} \ , \\ & & -2 \le \theta \le 2 \ . \end{aligned}$$

The test statistic $T^*(x) = -|x - 1/2|$ is unbiased for the hypothesis problem $H_0\colon \theta = \theta_0$ versus $H_A\colon \theta \ne \theta_0$. To verify this, write

$$H^*_\theta(\alpha) = 1 - \int_{c_1}^{c_2} [1 + \theta(x-1/2)]dx$$

where $c_1 + c_2 = 1$.

$$H^*_\theta(\alpha) = 2(1-c_2)$$
$$= 2c_1$$

which does not depend on any choice of $\theta = \theta_0$ nor any $\theta \neq \theta_0$. Since $H^*_\theta(\alpha)$ is constant for all θ then $T^*(x)$ is an unbiased two-tail test statistic.

Furthermore $T^*(x) = |x-1/2|$ is an unbiased test statistic whose significance regions do not include either tail region of the x axis. It is a no-tail test statistic.

Example 2.4 (Wilks (10)): Consider a random sample of size n from a $N(\mu, \sigma^2)$ population. The hypothesis problem is a composite null hypothesis versus a composite alternative hypothesis,

$$H_0: \mu = \mu_0 \quad , \quad \sigma^2 > 0 \quad ,$$
$$H_A: \mu \neq \mu_0 \quad , \quad \sigma^2 > 0 \quad .$$

Define a test statistic as

$$T^*(t) = \left[1 + \frac{t^2}{n-1}\right]^{-\frac{n}{2}}$$

where t has Student's distribution with $n - 1$ degrees of freedom. The density of t is given by

$$f_{n-1}(t) = \frac{1}{\sqrt{\pi(n-1)}} \cdot \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})} \cdot \left[1 + \frac{t^2}{n-1}\right]^{-\frac{n}{2}}$$
$$-\infty < t < \infty .$$

For any point $\theta_1 = (\mu_1, \sigma_1^2) \in \Theta_A$ we have

$$P_{\theta_1}[SL(T^*) \le \alpha] = P_{\theta_1}[\left|\frac{\sqrt{n}\ (\bar{X}-\mu_0)}{S}\right| > t_\alpha]$$

$$= P_{\theta_1}[\left|\frac{\sqrt{n}\ (\bar{X}-\mu_1) + \sqrt{n}\ (\mu_1-\mu_0)}{S}\right| > t_\alpha]$$

$$= P_{\theta_1}[\left|T + \frac{\delta}{\sqrt{U}}\right| > t_\alpha]$$

where

$$\delta = \sqrt{n(n-1)}\ (\mu_1-\mu_0)/\sigma_1 \quad ,$$

and t and u have the following joint density:

$$f(t,\ u) = \frac{(\frac{u}{2})^{\frac{n-2}{2}} e^{-\frac{u}{2}[1 + \frac{t^2}{n-1}]}}{2\sqrt{\pi(n-1)}\ \Gamma(\frac{n-1}{2})} \quad , \quad \begin{array}{c} -\infty < t < \infty \\ 0 < u < \infty \end{array} .$$

We note the symmetry of the density function about $t = 0$ for any value of u. It follows then that

$$\int_{-t_\alpha - \frac{\delta}{\sqrt{u}}}^{t_\alpha - \frac{\delta}{\sqrt{u}}} f(t,\ u)dt = \int_{-t_\alpha - \frac{\delta}{\sqrt{u}}}^{-t_\alpha} f(t,\ u)dt + \int_{-t_\alpha}^{t_\alpha - \frac{\delta}{\sqrt{u}}} f(t,\ u)dt$$

$$= \int_{t_\alpha}^{t_\alpha + \frac{\delta}{\sqrt{u}}} f(t,\ u)dt + \int_{-t_\alpha}^{t_\alpha - \frac{\delta}{\sqrt{u}}} f(t,\ u)dt$$

$$< \int_{t_\alpha - \frac{\delta}{\sqrt{u}}}^{t_\alpha} f(t,\ u)dt + \int_{-t_\alpha}^{t_\alpha - \frac{\delta}{\sqrt{u}}} f(t,\ u)dt$$

$$= \int_{-t_\alpha}^{t_\alpha} f(t,\ u)dt$$

from which we obtain

$$P_{\theta_1}[SL(T^*) \le \alpha] = 1 - P_{\theta_1}[|T + \frac{\delta}{\sqrt{U}}| \le t_\alpha]$$

$$> 1 - P_{\theta_1}[|t| \le t_\alpha]$$

$$= \alpha \quad .$$

Therefore $T^*(t)$ is unbiased for the hypothesis problem.

Discrete Case

Example 2.5: Consider n independent Bernoulli trials and the hypothesis $H_0: \theta = 1/2$ versus $H_A: \theta \ne 1/2$. The statistic $T = \sum_1^n X_i$ is sufficient, and it has probability function,

$$p_\theta(t) = \binom{n}{t} \theta^t (1-\theta)^{n-t} \quad , \quad t = 0, 1, \ldots, n \quad .$$

Finley shows that

$$T^*(t) = -|t - \frac{n}{2}|$$

is an unbiased test statistic for the problem. The probability function under θ_0 is symmetric in t, as is the test statistic.

Example 2.6: Let the random variable X have probability function

$$p_\theta(x) = \frac{2(x + \theta)}{(n+1)(n+2\theta)} \quad , \quad x = 0, 1, \ldots, n$$
$$0 \le \theta < \infty \quad .$$

We wish to show $T^*(x) = -|x - n/2|$ gives an unbiased significance test for $H_0: \theta = \theta_0$ versus $H_A: \theta \ne \theta_0$.

Let α_j be the significance level attained by observing either $x = j$ or $x = n-j$. Then

$$H_\theta^*(\alpha_j) = \frac{2}{(n+1)(n+2\theta)} \quad \sum_{x=0}^{j} (x+\theta) + \sum_{n-j}^{n} (x+\theta)$$

$$= \frac{2}{(n+1)(n+2\theta)} \quad 2(j+1)\theta + \frac{j(j+1)}{2} + n(j+1) - \frac{j(j+1)}{2}$$

$$= \frac{2}{(n+1)(n+2\theta)} [(j+1)(2\theta + n)]$$

$$= \frac{2(j+1)}{(n+1)} \quad , \quad j = 0, 1, \ldots, [n/2] \quad .$$

Since $H_\theta^*(\alpha_j)$ is constant for each achievable α for all θ, then it trivially follows that $T^*(x)$ is unbiased. It should be noted that $p_\theta(x)$ is not symmetric in x nor is it unimodal.

Furthermore, by letting $T^*(x) = |x - \frac{n}{2}|$ we obtain a no-tail unbiased test statistic.

<u>Example 2.7</u>: Let x have probability function

$$p_\theta(x) = \frac{e^{-\sin\theta}(\sin\theta)^x}{x!} \quad , \quad x = 0, 1, 2, \ldots$$

$$0 < \theta < \pi \quad .$$

For the hypothesis problem H_0: $\theta = \pi/2$ versus H_A: $\theta \neq \pi/2$ the test statistic $T^*(x) = x$ is unbiased.

$$H_\theta^*(\alpha) = \sum_{x=0}^{j} \frac{e^{-\sin\theta}(\sin\theta)^x}{x!}$$

$$= 1 - \int_0^{\chi^2} \frac{t^{\frac{k}{2}-1} e^{-\frac{t}{2}} dt}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})}$$

where $\sin\theta = \frac{1}{2}\chi^2$, and $j + 1 = \frac{k}{2}$. We can then write

$$H_\theta^*(\alpha_j) = 1 - F_k\ (2\sin\theta)$$

where $F_k(2\sin\theta)$ is chi-square distribution function for k degrees of freedom evaluated at $\chi^2 = 2\sin\theta$.

By taking the first derivative with respect to θ , we obtain

$$\frac{d\ H_\theta^*(\alpha_j)}{d\theta} = -F_k'(2\sin\theta)\cdot 2\cos\theta\ , \quad \text{or}$$

$$= 0\ , \text{ when } \theta = \frac{\pi}{2}\ ;$$

and by taking the second derivative with respect to θ ,

$$\frac{d^2 H_\theta^*(\alpha_j)}{d\theta^2} = -\left[F_k''(2\sin\theta)[2\cos\theta]^2 - F_k'(2\sin\theta)2\sin\theta\right], \quad \text{or}$$

$$= 2F_k'(2) > 0 \quad \text{when} \quad \theta = \frac{\pi}{2}\ .$$

Therefore $H_\theta^*(\alpha_j)$ has its minimum at $\theta = \frac{\pi}{2}$, and $T^*(x) = x$ is an unbiased test statistic. We note here that an infinite number of achievable α levels is possible; we also note $T^*(x)$ is a one-tail test statistic for a 2-sided alternative hypothesis.

<u>Example 2.8:</u> Let x be a random variable with probability function

$$p_\theta(x) = \begin{cases} \frac{1}{2} \dfrac{e^{-\sin\theta}(\sin\theta)^x}{x!}, & x = 1,2,3,\ldots \\ \frac{1}{2} \dfrac{e^{-\frac{1}{\sin\theta}}(\sin\theta)^x}{|x|!}, & x = -1,-2,\ldots \\ \frac{1}{2}\left(e^{-\sin\theta} + e^{-\frac{1}{\sin\theta}}\right), & x = 0 \\ & 0 < \theta < \pi \end{cases}$$

The test statistic

$$T^*(x) = \begin{cases} 0 & x = -1,0,1,2 \\ -(x - 1/2)^2 + \frac{9}{4} & \text{all other integers,} \end{cases}$$

is unbiased for the hypothesis problem $H_0: \theta = \frac{\pi}{2}$ versus $H_A: \theta \neq \frac{\pi}{2}$.

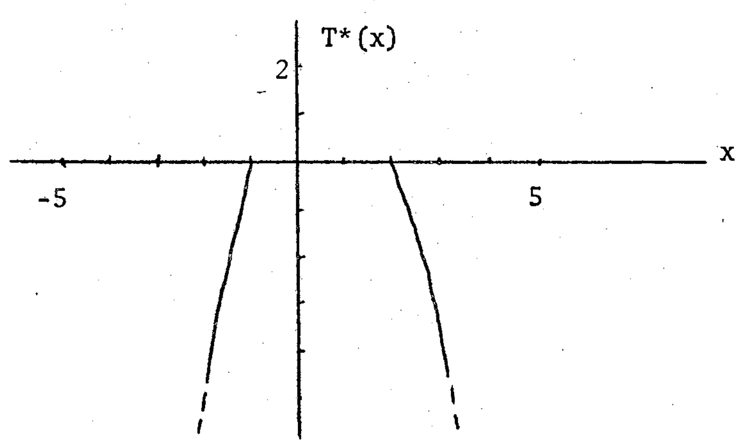

Fig. 2.1: An Unbiased Test Statistic for Ex. 2.8.

Consider positive integers a and b such that for $\alpha < 1$,

$$H^*_\theta(\alpha) = \sum_{-\infty}^{-a} \frac{e^{-\frac{1}{\sin\theta}} (\sin\theta)^x}{2(|x|!)} + \sum_{b}^{\infty} \frac{e^{-\sin\theta}(\sin\theta)^x}{2(x!)}$$

Since $\sum_{0}^{x'} (e^{-m} m^x)/x! = 1 - F_n(\chi^2)$ with $m = \frac{1}{2}\chi^2$, $x' + 1 = \frac{n}{2}$, and $F_n(\chi^2)$ the chi-square distribution, then

$$2H^*_\theta(\alpha) = F_{2a}(\chi^2_1) + F_{2b}(\chi^2_2)$$

where $1/\sin\theta = \frac{1}{2}\chi^2_1$, $\sin\theta = \frac{1}{2}\chi^2_2$, and 2a and 2b are the respective degrees of freedom.

As before we proceed to take first and second derivatives to determine if a minimum occurs at $\theta = \pi/2$.

$$\frac{1}{2}\,\frac{dH^*_\theta(\alpha)}{d\theta} = \left[\frac{-2\cos\theta}{\sin^2\theta}\right] F'_{2a}\left(\frac{2}{\sin\theta}\right) + [2\cos\theta]\, F'_{2b}(2\sin\theta)$$

which equals zero when $\theta = \pi/2$.

$$\frac{1}{2}\,\frac{d^2H^*_\theta(\alpha)}{d\theta^2} = \left[\frac{-2\cos\theta}{\sin^2\theta}\right]^2 F''_{2a}\left(\frac{2}{\sin\theta}\right) + \left[\frac{2\sin\theta}{\sin^2\theta} + \frac{4\cos^2\theta}{\sin^3\theta}\right] F'_{2a}\left(\frac{2}{\sin\theta}\right)$$

$$+ [2\cos\theta]^2 F''_{2b}\left(\frac{2}{\sin\theta}\right) + [-2\sin\theta]\, F'_{2b}(2\sin\theta)$$

and at $\theta = \pi/2$, we have

$$\frac{1}{2}\,\frac{d^2H^*_\theta(\alpha)}{d\theta^2} = 2[F'_{2a}(2) - F'_{2b}(2)] \quad .$$

Since $F'_{2a}(2)$ is the chi-square density function for $2a$ degrees of freedom evaluated at $\chi^2 = 2$, similarly for $F'_{2b}(2)$, then $F'_{2a}(2) - F'_{2b}(2) > 0$ for all a and b such that $2 < 2a < 2b$. The test statistic $T^*(x)$ is defined such that for achievable $\alpha < 1$ the integers a and b satisfy this inequality.

We note that this example gives a two-tail unbiased test statistic with an infinite number of achievable α's. Furthermore, the probability function is not symmetric in x.

A brief summary of the important characteristics of the preceding examples is in order. It is hoped some questions will be answered regarding the nature of density functions $p_\theta(x)$ which admit an unbiased test statistic. This summary is offered in the form of a short list.

1. It is not necessary in either the continuous or discrete case that $p_\theta(x)$ be unimodal in x or symmetric in x (examples 2.3 and 2.6).

2. It is possible to have a two-tail test statistic based on x when $p_\theta(x)$ is not symmetric in x in the discrete case (examples 2.6 and 2.8).

3. It is possible to have a one-tailed test statistic when the alternative hypothesis is two-sided $(\theta \neq \theta_0)$ in the discrete case (example 2.7).

4. It is possible to have a no-tail unbiased test statistic in both the continuous and discrete cases (examples 2.3 and 2.6).

5. It is possible to have an infinite number of achievable α's in the discrete case (examples 2.7 and 2.8).

6. It is possible to have an unbiased test statistic when the null hypothesis is composite (example 2.4).

The work in the following chapters will investigate necessary and sufficient conditions for unbiased test statistics to exist in the context of significance testing.

CHAPTER III

NECESSARY CONDITIONS FOR UNBIASEDNESS FOR FAMILIES WITH STRICT MONOTONE LIKELIHOOD RATIO

When a uniformly most sensitive (UMS) test statistic cannot be found for a hypothesis problem, one may wish to search for a "best" test statistic in a smaller class. One such class is that of the unbiased test statistics. This chapter investigates necessary conditions for test statistics to be unbiased when the density function has monotone likelihood ratio (MLR) or strict monotone likelihood ratio (SMLR) in some real-valued statistic. An important family of distributions which have the latter property is a one-parameter exponential family whose densities are defined by

$$p_\theta(x) = C(\theta)\, e^{Q(\theta)T(x)} h(x) \tag{3.1}$$

where Q is strictly monotone. The primary goals of this and the next chapter are to achieve results applicable to this family.

It is recalled that we shall consider unbiased test statistics only in the cases of two-sided alternative hypothesis $(\theta \neq \theta_0)$. One of the major questions, then, is whether an unbiased test statistic must be two-tail. It was emphasized in Chapter II the answer to this question in general is no. However, we need to explore the question further for the more interesting density functions.

It is seen from example 2.6 that a one-tail unbiased test statistic is possible in the one-parameter exponential case. However, the function $Q(\theta) = \sin\theta$ was unimodal in θ for $0 < \theta < \pi$; it was not strictly monotone. The first theorem shows that a one-tail unbiased test statistic does not exist when $Q(\theta)$ is strictly monotone in θ.

We first need a lemma close to that given by Lehmann:

Lemma 3.1: Let $p_\theta(x)$ belong to a family of densities on the real line with MLR in x.

i) If ψ is a monotone (nondecreasing or nonincreasing) function of x, then $E_\theta\psi$ is monotone (nondecreasing or nonincreasing, respectively) in θ.

ii) For any $\theta < \theta'$, the distribution function of x satisfies

$$F_{\theta'}(x) \leq F_\theta(x) \qquad , \qquad \text{for all } x .$$

iii) If $p_\theta(x)$ is SMLR in x, then for $\theta < \theta'$,

$$F_{\theta'}(x) < F_\theta(x) \qquad , \qquad \text{for all } x .$$

Proof: Lehmann gives the proof for i) and ii) for the case when $\psi(x)$ is nondecreasing. We follow his method of proof for the case when $\psi(x)$ is nonincreasing.

Let $\theta < \theta'$ and let the sets A and B be defined as

$$A = \{x: \; p_{\theta'}(x)/p_\theta(x) < 1\} \quad , \quad \text{and}$$

$$B = \{x: \; p_{\theta'}(x)/p_\theta(x) > 1\} \quad .$$

Let $a = \inf_A \psi(x)$, $b = \sup_B \psi(x)$, then since ψ is nonincreasing in x , $a - b \geq 0$ or $b - a \leq 0$.

$$E_{\theta'} \psi - E_\theta \psi = \int_A \psi(p_{\theta'} - p_\theta)d\mu + \int_B \psi(p_{\theta'} - p_\theta)d\mu$$

$$\leq a\int_A (p_{\theta'} - p_\theta)d\mu + b\int_B (p_{\theta'} - p_\theta)d\mu$$

By adding and subtracting the integral $a\int_B (p_{\theta'} - p_\theta)d\mu$ on the right side of the inequality, and combining terms we obtain

$$E_{\theta'}\psi - E_\theta \psi \leq (b-a)\int_B (p_{\theta'} - p_\theta)d\mu$$

$$\leq 0 \quad .$$

This proves i).

Now let $\psi(x) = 0$ for $x > x_0$ and $\psi(x) = 1$ otherwise. $E_\theta\psi(x) = F_\theta(x_0)$ which implies $F_{\theta'}(x) \leq F_\theta(x)$ for $\theta < \theta'$. This proves part ii).

For part iii) we assume that $p_\theta(x)$ is SMLR in x and $\psi(x)$ is nondecreasing and non-constant. It is shown in the appendix that $E_{\theta'}\psi(x) > E_\theta\psi(x)$ for $\theta' > \theta$.

$$\text{Let } \psi(x) = 0 \quad , \quad x \leq x_0$$
$$= 1 \quad , \quad x > x_0 \quad ,$$

then $P_{\theta'}[X > x_0] > P_\theta[X > x_0]$ which implies $F_{\theta'}(x_0) < F_\theta(x_0)$.

If we now assume a sample of size one is taken, then we can prove the following theorem.

<u>Theorem 3.1</u>: Let $p_\theta(x)$ be a family of densities on the real line

with SMLR in x . There does not exist an unbiased test statistic which is a monotone function of x for the hypothesis problem $H_0: \theta = \theta_0$ versus $H_A: \theta \neq \theta_0$.

Proof: The proof is by contradiction. Assume an unbiased test statistic $T^*(x)$ is a nondecreasing function of x . Then for any θ ,

$$
\begin{aligned}
H_\theta^*(\alpha) &= P_\theta[SL(T^*) \leq \alpha] \\
&= P_\theta[T^*(X) \leq t_\alpha^*] \\
&= P_\theta[\; X \leq x_0 \;]
\end{aligned}
$$

where we define $x_0 = \max\{x | T^*(x) \leq t_\alpha^*\}$; hence $H_\theta^*(\alpha) = F_\theta(x_0)$, and for $\theta > \theta_0$ it follows from lemma 3.1 (iii) that

$$
H_\theta^*(\alpha) < H_0^*(\alpha) = \alpha \quad .
$$

This contradicts the assumption that $T^*(x)$ is unbiased.

Assume $T^*(x)$ is a nonincreasing function of x . Then for any θ ,

$$
\begin{aligned}
H_\theta^*(\alpha) &= P_\theta[T^*(X) \leq t_\alpha^*] \\
&= P_\theta[\; X \geq x_0 \;] \\
&= 1 - P_\theta[X < x_0]
\end{aligned}
$$

where $x_0 = \min\{x | T^*(x) < t_\alpha^*\}$. We next consider the difference

$$
H_\theta^*(\alpha) - H_0^*(\alpha) = P_0[X < x_0] - P_\theta(X < x_0) \quad ,
$$

when $\theta < \theta_0$. If $p_\theta(x)$ were continuous in x then

$$
\begin{aligned}
H_\theta^*(\alpha) - H_0^*(\alpha) &= F_0(x_0) - F_\theta(x_0) \\
&< 0 \quad ,
\end{aligned}
$$

which contradicts the unbiasedness of $T^*(x)$. If $p_\theta(x)$ is discrete in x , then

$$H^*_\theta(\alpha) - H^*_0(\alpha) = P_0[X \leq x_1] - P_\theta[X \leq x_1] \quad ,$$

where $x_1 < x_0$ and there exists no value of x with positive probability such that $x_1 < x < x_0$. Hence

$$H^*_\theta(\alpha) - H^*_0(\alpha) = F_0(x_1) - F_\theta(x_1)$$
$$< 0 \quad , \quad \theta < \theta_0 \quad .$$

again there is a contradiction and the theorem is proved.

The following corollary may be stated directly from this theorem:

Corollary 3.1: Let $p_\theta(x)$ be a family of densities on the real line with SMLR in x . For the two-sided hypothesis problem there does not exist an unbiased one-tail test statistic.

Thus, it has been shown that for densities which are SMLR in a real statistic x there does not exist a one-tail unbiased significance test based on x . This still does not imply we must have two-tail statistics in these cases, even though this is suspected to be true. It may be possible that there exist multi-modal test statisitics which are unbiased, but this has neither been proved or disproved.

Let $T^*(x)$ be a test statistic for the hypothesis problem H_0: $\theta = \theta_0$ versus H_A: $\theta \neq \theta_0$ where x has a density $p_\theta(x)$. If $R = \{x | T^*(x) \leq t^*_\alpha\}$, then for $\theta \neq \theta_0$,

$$H^*_\theta(\alpha) = P_\theta[SL(T^*) \leq \alpha]$$

$$= P_\theta[T^*(X) \le t^*_\alpha]$$

$$= \int_R p_\theta(x)\ d\mu(x) = \int_R \frac{p_\theta(x)}{p_0(x)} \cdot p_0(x) d\mu(x) \quad .$$

Now let $r_\theta(x) = p_\theta(x)/p_0(x)$, and write $H^*_\theta(\alpha) = E_0[r_\theta(x) I_R]$. If $T^*(x)$ is unbiased, then

$$E_0[r_\theta(x) I_R] \ge \alpha \quad , \quad \text{or equivalently}$$

$$E_0[r_\theta(x) I_R] \ge \int_R p_0(x) d\mu(x) \quad , \text{which upon}$$

dividing both sides of the inequality by the right-hand side gives

$$E_0[r_\theta(x) | T(\overset{*}{x}) \le t^*_\alpha] \ge 1 \quad . \tag{3.2}$$

Furthermore, if (3.2) holds then $T^*(x)$ is unbiased. This implies that (3.2) is a necessary and sufficient condition for $T^*(x)$ to be unbiased. Using the notation $\overline{R}$ for the complement of R , we state and prove the following theorem:

<u>Theorem 3.2</u>: Let $r_\theta(x) = p_\theta(x)/p_0(x)$. Then $E_0[r_\theta(X) | T^*(X) \le t^*_\alpha] \ge 1$ if and only if $E_0[r_\theta(X) | T^*(X) > t^*_\alpha] \le 1$, for every θ .

<u>Proof:</u>

$$E_0[r_\theta(X) | T^* \le t^*_\alpha] = \frac{\int_R r_\theta(x)\ p_0(x) d\mu}{\int_R p_0(x) d\mu}$$

$$= \frac{\int_R r_\theta \cdot p_0 \, d\mu + \int_{\bar{R}} r_\theta \cdot p_0 d\mu - \int_{\bar{R}} r_\theta \cdot p_0 \, d\mu}{1 - \int_{\bar{R}} p_0 \, d\mu}$$

$$= \frac{E_0[r_\theta(x)] - \int_{\bar{R}} r_\theta \cdot p_0 \, d\mu}{1 - \int_{\bar{R}} p_0 \, d\mu}$$

$$= \frac{1 - \int_{\bar{R}} r_\theta \, p_0 \, d\mu}{1 - \int_{\bar{R}} p_0 \, d\mu}$$

Therefore, $E_0[r_\theta(X)|T^*(X) \leq t_\alpha^*] \geq 1$ if and only if

$$\int_{\bar{R}} r_\theta \cdot p_0 \, d\mu \leq \int_{\bar{R}} p_0 \, d\mu \quad , \quad \text{or}$$

$$\frac{\int_{\bar{R}} r_\theta \cdot p_0 \, d\mu}{\int_{\bar{R}} p_0 \cdot d\mu} \leq 1 \quad ,$$

which can be written as

$$E_0[r_\theta(X)|T^*(x) > t_\alpha^*] \leq 1 \quad ,$$

and the theorem is proved.

It may be of interest to study $p_\theta(x)$ as a function of θ for fixed values of x. Perhaps in doing so some clue can be found as to the existence and/or nature of an unbiased test statistic.

Let $p_\theta(x)$ be SMLR in x, and assume that for fixed x_0, $p_\theta(x_0)$ is a unimodal continuous function of θ. The restriction of continuity is not necessary, but it simplifies the discussion without a great loss in application. We shall also restrict Θ to be an open set on the real line. Then for fixed θ_0, there exists $\theta_1 < \theta_0$ and $\theta_2 > \theta_0$ such that

$$p_{\theta_1}(x_0) = p_{\theta_2}(x_0) < p_{\theta_0}(x_0) \quad ,$$

from which we obtain

$$r_{\theta_1}(x_0) = r_{\theta_2}(x_0) < 1 \quad .$$

Since $p_\theta(x)$ is SMLR in x, then $r_{\theta_1}(x)$ is strictly decreasing in x, and $r_{\theta_2}(x)$ is strictly increasing in x.

We now prove the following theorem:

Theorem 3.3: Let $p_\theta(x)$ be SMLR in x, and let R be any subset of the domain of $p_\theta(x)$ such that $b = \sup R$ is real. If $p_\theta(x)$ is unimodal and continuous in θ at $x = b$, then there exists a θ such that $E_0[r_\theta(X)|x\varepsilon R] < 1$. If the same conditions hold for $a = \inf R$ the result is identical.

Proof: By the previous discussion there exists $\theta_2 > \theta_0$ such that

$$r_{\theta_2}(b) = c < 1 \quad ,$$

where $r_{\theta_2}(x)$ is strictly increasing in x. Therefore, for all x in R, $r_{\theta_2}(x) \leq r_{\theta_2}(b) = c < 1$. This implies

$$E_0[r_{\theta_2}(x)|x\varepsilon R] = \frac{\int_R r_{\theta_2}(x)\, p_0(x)\, d\mu(x)}{\int_R p_0(x)\, d\mu(x)}$$

$$\leq c$$

$$< 1 \ .$$

It follows that if $p_\theta(x)$ is unimodal and continuous in θ at $a = \inf R$, then for $\theta_1 < \theta_0$ we would have $r_{\theta_1}(a) < 1$ where $r_{\theta_1}(x)$ is strictly decreasing in x . Then we obtain the desired result by proceding as above.

This theorem can be used in the following manner. Suppose x has a density $p_\theta(x)$ such as the binomial, poisson, or normal, and we wish to find an unbiased significance test for $H_0: \theta = \theta_0$ versus $H_A: \theta \neq \theta_0$. Theorem 3.3 simply tells us we need not consider test statistics based on x which allow no-tail or one-tail significance tests.

If $p_\theta(x)$ is unimodal in θ at $x = a$, denote by θ_a the value of θ at which $p_\theta(a)$ is maximized. Let $R = \{x | T^*(x) \leq t^*_\alpha\}$ and let $\bar{R}$ be the complement of R . We now state and prove the following lemma:

Lemma 3.2: Let $p_\theta(x)$ be SMLR in x , and let $T^*(x)$ be an unbiased test statistic based on x , where $R = \{x | T^*(x) \leq t^*_\alpha\}$. If $p_\theta(x)$ is unimodal and continuous in θ at

i) $a = \inf \bar{R}$, then $\theta_a \leq \theta_0$,

ii) $b = \sup \bar{R}$, then $\theta_b \geq \theta_0$.

Proof: The proof is by contrapositive.

i) Assume $\theta_a > \theta_0$, then $p_{\theta_a}(a) > p_0(a)$, or $r_{\theta_a}(a) = c > 1$ where $r_{\theta_a}(x)$ is strictly increasing in x . Therefore, for all x in $\bar{R}$ we have $r_{\theta_a}(x) \geq c > 1$, which implies $E_0[r_{\theta_a}(X)|x\varepsilon\bar{R}] > 1$. This is a contradiction, since by unbiasedness of $T^*(x)$ we must have $E_0[r_{\theta_a}(X)|x\varepsilon\bar{R}] \leq 1$. Therefore, $\theta_a \leq \theta_0$.

ii) Assume $\theta_b < \theta_0$, then $r_{\theta_b}(b) > 1$ where $r_{\theta_b}(x)$ is strictly decreasing in x. Again it follows that $E_0[r_{\theta_b}(X)|x\varepsilon\bar{R}] > 1$, which contradicts the hypothesis of unbiasedness.

Theorem 3.4: Let the conditions of lemma 3.2 hold. If x_0 is a point at which $\theta = \theta_0$ maximizes $p_\theta(x_0)$ as a function of θ, then $a \leq x_0 \leq b$.

Proof: The proof is again by contrapositive. Assume that $x_0 > b$, and let θ_b be the maximizing θ at $x = b$. Then $p_{\theta_b}(b) \geq p_0(b)$, which implies $r_{\theta_b}(x_0) \leq 1$. It follows that

$$r_{\theta_b}(b) \geq 1 \geq r_{\theta_b}(x_0) \quad .$$

Since $p_\theta(x)$ is SMLR in x and $x_0 > b$, then

$$r_{\theta_b}(b) > r_{\theta_b}(x_0)$$

which implies $r_{\theta_b}(x)$ is strictly decreasing and, thus, θ_b must be less than θ_0. By lemma 3.2 (ii) this is a contradiction. Therefore $x_0 \leq b$.

Similarly it can be shown that $x_0 \geq a$.

In example 2.1 it was shown the statistic $T^*(t) = t^{n/2}\exp[-t/2\sigma_0^2]$ is unbiased, where $T = \sum_1^n X_i^2$. It is well known that when $\mu = 0$, T/n is the maximum likelihood estimate for σ^2. Therefore $t_0 = n\sigma_0^2$ is the point at which $\sigma = \sigma_0$ maximizes $p_\theta(t_0)$ as a function of

$\theta = \sigma^2$. It is seen that for every level of $\alpha(<1)$ the closed interval $[a, b]$ contains t_0 .

In example 2.5 the statistic $T^*(x) = -|x - n/2|$ is unbiased. The maximum likelihood estimate of θ in the binomial is x/n . Therefore, if $\theta_0 = 1/2$ and n is even, then $x_0 = n\theta_0 = n/2$ is the point at which $\theta = \theta_0$ maximizes $p_\theta(x)$ as a function of θ . For $n = 6$ and $\theta_0 = 1/2$, Theorem 3.4 states that $x = 3$ will be a point in $\overline{R}$ for all $\alpha < 1$. If n be odd, the application of the theorem in this discrete case appears to fail; for in this case, $x_0 = n/2$ is not an integer and $p_\theta(x_0) = 0$.

It was shown by Finley that for the family (3.1) a necessary condition for unbiasedness is

$$E_0[T(X)|T^*[T(X)] \leq t_\alpha^*] = E_0[T(X)] \quad . \tag{3.2}$$

This condition may be written in a form closely resembling that given by Lehmann ,

$$E_0[T(X)I_R] = \alpha\ E_0[T(X)] \tag{3.3}$$

where $R = \{x|T^*[T(X)] \leq t_\alpha^*\}$. We exhibit an example to show this condition is not sufficient.

Consider a sample of size n to be taken from a $N(\mu, 1)$ population. Let the hypothesis problem be $H_0: \mu = 0$ versus $H_A: \mu \neq 0$. By letting $T(x) = \overline{X}$, it is seen for example 2.2 that the unbiased test statistic $T^*(\overline{x}) = \exp[-n\overline{x}^2/2]$ satisfies (3.3). We shall show that the bimodal statistic

$$T'(\overline{x}) = \overline{x}^2\ e^{-\overline{x}^2/2}$$

also satisfies (3.3), but it is not unbiased.

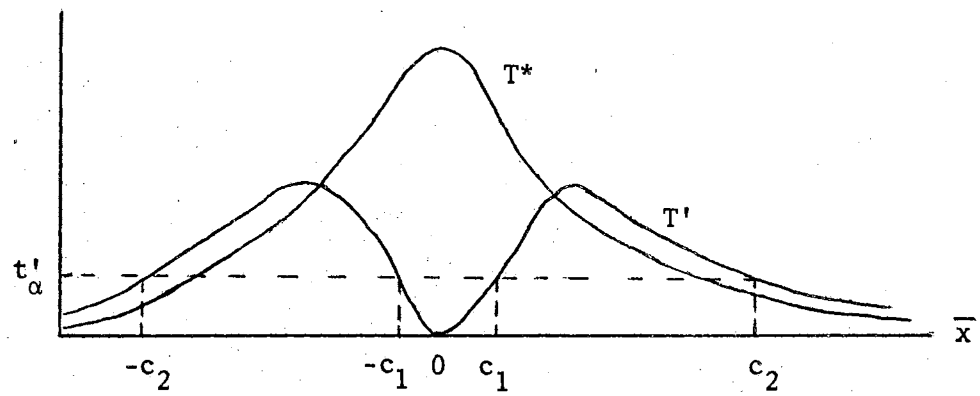

Fig. 3.1

By the symmetry of $T'(\bar{x})$ it is clear that

$$E_0[\bar{X}|T'(\bar{x}) \leq t'_\alpha] = E_0[\bar{X}\ I_{R'}]$$

$$= 0 \qquad ,$$

thus $T'(\bar{x})$ satisfies the necessary condition. To investigate its unbiasedness we examine $H'_\mu(\alpha)$ and its derivative with respect to μ. With c_1 and c_2 defined in figure 3.1 ,

$$\frac{H'_\mu(\alpha)}{2} = \int_0^{c_1} f_\theta(\bar{x})\ d\bar{x} + \int_{c_2}^{\infty} f_\theta(\bar{x})\ d\bar{x}$$

$$= 1 - \int_{c_1}^{c_2} \frac{n^{1/2}}{(2\pi)^{1/2}} \exp -[n(\bar{x}-\mu)^2/2]d\bar{x}$$

It is then straightforward to show

$$\left.\frac{dH'_\mu(\alpha)/2}{d\mu}\right|_{\mu=0} = \frac{n^{1/2}}{(2\pi)^{1/2}} [e^{-nc_2^2/2} - e^{-nc_1^2/2}] \quad .$$

The derivative does not equal zero at $\mu = 0$ since

$$e^{-nc_2^2/2} \neq e^{-nc_1^2/2} \quad ;$$

therefore $T'(\bar{x})$ cannot be unbiased.

One can see at first glance that $T'(\bar{x})$ is not a "good" test statistic; for one could observe a value of $\bar{x} = 0$ and obtain a significance level of $\alpha = 0$. A statistic such as this would be useless.

CHAPTER IV

SUFFICIENT CONDITIONS FOR UNBIASED AND UMSU TEST STATISTICS IN THE ONE-PARAMETER EXPONENTIAL FAMILY

The preceding chapter gave several necessary conditions for the existence of an unbiased test statistic involving densities with strict monotone likelihood ratio. One of these necessary conditions will now be used in producing a sufficient condition for a test statistic to be unbiased. Attention will be mainly restricted to the one-parameter exponential family of densities.

We first state and prove, in the context of significance testing, a portion of Lehmann's generalization of the fundamental lemma.

<u>Theorem 4.1</u>: Let $f_1, \ldots, f_{m+1}$ be real-valued functions on a Euclidean space $\mathfrak{X}$. For a hypothesis problem let $T^*(x)$ be a test statistic, and let $R = \{x | T^*(x) \leq t^*_\alpha\}$. Assume there exists a $T^*(x)$ such that for constants $c_1, c_2, \ldots, c_m$,

$$\int_R f_i \, d\mu = c_i \quad , \quad i = 1, \ldots, m \quad . \tag{4.1}$$

Finally, let $\mathcal{C}$ be the set of all such test statistics which satisfy (4.1). A sufficient condition for a number of $\mathcal{C}$ to maximize $\int_R f_{m+1} d\mu$ is the existence of constants $k_1, \ldots, k_m$ such that

$$f_{m+1}(x) \geq \sum_1^m k_i f_i(x) \quad , \quad x\varepsilon R$$

$$< \sum_1^m k_i f_i(x) \quad , \quad \text{otherwise} \qquad (4.2)$$

Proof: Assume there exists $T^*(x) \varepsilon \mathcal{C}$ such that (4.2) holds. Let $T'(x)$ be any other test statistic belonging to $\mathcal{C}$, where $R' = \{x|T'(x) \leq t'_\alpha\}$. Furthermore, define

$$S^+ = \{x|I_R - I_{R'} > 0\} \quad , \quad \text{and}$$

$$S^- = \{x|I_R - I_{R'} < 0\} \quad .$$

Then

$$\int (I_R - I_{R'})(f_{m+1} - \sum_1^m k_i f_i)d\mu$$

$$= \int_{S^+} (I_R - I_{R'})(f_{m+1} - \sum_1^m k_i f_i)d\mu + \int_{S^-} (I_R - I_{R'})(f_{m+1} - \sum_1^m k_i f_i)d\mu$$

When $x\varepsilon S^+$, $f_{m+1}(x) \geq \sum_1^m k_i f_i(x)$; similarly, when $x\varepsilon S^-$, $f_{m+1}(x) < \sum_1^m k_i f_i(x)$. Therefore, we have that

$$\int (I_R - I_{R'})(f_{m+1} - \sum_1^m k_i f_i)d\mu \geq 0 \quad , \quad \text{or}$$

$$\int (I_R - I_{R'})f_{m+1}\, d\mu \geq \sum_1^m k_i \int (I_R - I_{R'})f_i\, d\mu \quad .$$

Since both $T^*(x)$ and $T'(x)$ belong to $\mathcal{C}$,

$$\int (I_R - I_{R'})f_i \; d\mu = \int I_R \; f_i \; d\mu - \int I_{R'} \; f_i \; d\mu$$

$$= c_i - c_i$$

$$= 0 \quad , \quad i = 1, \ldots, m \quad .$$

Therefore

$$\int_R f_{m+1} \; d\mu \geq \int_{R'} f_{m+1} \; d\mu$$

as was to be proved.

Assume that $f_{m+1}(x) = p_\theta(x)$ for $\theta \varepsilon \Theta_A$; then

$$\int_R f_{m+1} \; d\mu = \int_R p_\theta \; d\mu$$

$$= P_\theta[T^*(x) \leq t^*_\alpha]$$

$$= H^*_\theta(\alpha) \quad , \quad \theta \varepsilon \Theta_A \quad . \tag{4.3}$$

Therefore, if $T^*(x)$ belongs to $\mathscr{C}$ and maximizes (4.3) for all achievable α , $T^*(x)$ is most sensitive among those statistics in $\mathscr{C}$. If $T^*(x)$ is most sensitive for all θ in Θ_A then $T^*(x)$ would be termed UMS in $\mathscr{C}$.

Consider the hypothesis problem H_0: $\theta = \theta_0$ versus H_A: $\theta \neq \theta_0$ when x has the one-parameter exponential density (3.1). Let $\mathscr{C}$ be the class of all test statistics $T^*(x)$ which satisfy

i) $P_0[T^*(x) \leq t^*_\alpha] = \int_R p_0(x) d\mu(x) = \alpha$, and

$$\text{ii)} \quad E_0[T(x)I_R(x)] = \int_R t(x)p_0(x)d\mu$$

$$= \alpha \cdot E_0[T(x)] \quad .$$

Condition i) requires that all test statistics in $\mathcal{E}$ must be comparable. Condition ii) states that all test statistics in $\mathcal{E}$ satisfy a necessary condition for $T^*(x)$ to be unbiased for this family of densities.

For $m=2$, let $f_1 = p_0(x)$, $f_2 = t(x)p_0(x)$, $f_3 = p_\theta(x)$, $c_1 = \alpha$, and $c_2 = \alpha \cdot E_0[T(x)]$, where $t(x)$ is the mathematical notation for the statistic $T(x)$. It is seen then that a sufficient condition for a test statistic $T^*(x)$ in $\mathcal{E}$ to be uniformly most sensitive is that for every θ in Θ_A there exists constants $k_1(\theta, \alpha)$ and $k_2(\theta,\alpha)$ such that

$$p_\theta(x) \geq k_1 \cdot p_0(x) + k_2 \cdot t(x)p_0(x) \quad , \quad x\varepsilon R \qquad (4.4)$$

where the values of k_1 and k_2 may depend on the values of θ and α , and $R = \{x | T^*(x) \leq t^*_\alpha\}$.

Define $\mathcal{U}$ to be a set of comparable unbiased test statistics for this hypothesis problem. Clearly, $\mathcal{U}$ is a subset of $\mathcal{E}$. Assume $\mathcal{U}$ is not empty, that is, there exists at least one unbiased test statistic for the problem. If $T^*(x)$ is UMS among those in $\mathcal{E}$, then it necessarily follows that $T^*(x)$ is unbiased, for if $T'(x) \varepsilon\, \mathcal{U} \subset \mathcal{E}$ then

$$H^*_\theta(\alpha) \geq H'_\theta(\alpha) \geq \alpha \quad , \quad \theta\varepsilon\Theta_A \quad .$$

We can summarize the preceding paragraph as a sufficient condition for unbiasedness:

> Let $\mathcal{U}$ and $\mathcal{E}$ be defined as above. If $\mathcal{U}$ is not empty, and $T^*(x)$ is UMS in $\mathcal{E}$, then $T^*(x)$ is unbiased, and, hence UMSU.

This sufficient condition appears to be more academic than useful in significance testing. After all, one of our most difficult problems is determining whether or not an unbiased test statistic exists. It should be noted that in the Neyman-Pearson approach to hypothesis testing the conditional clause "If $\mathcal{U}$ is not empty" would have no meaning, for there would always exist the trivial unbiased test statistic $T^*(x) = \alpha$.

The next theorem is intended to be more useful in application. We first must define what is meantby a function being concave upwards.

<u>Definition 4.1</u>: Let x_1 and x_2 be points at which $f(x)$ is defined, and consider $\gamma \varepsilon (0, 1)$ such that $f(x)$ is defined at $\gamma x_1 + (1-\gamma)x_2$. Then $f(x)$ is concave upwards if

$$\gamma f(x_1) + (1-\gamma)f(x_2) \geq f[\gamma x_1 + (1-\gamma)x_2] \quad .$$

Note that $r_\theta(x) = p_\theta(x)/p_0(x)$ for the density (3.1) satisfies this definition.

<u>Theorem 4.2</u>: For the hypothesis problem $H_0: \theta = \theta_0$ versus $H_A: \theta \neq \theta_0$, let $p_\theta(x)$ be a density such that $r_\theta(x)$ be concave upwards in $T(x)$ for $\theta \neq \theta_0$. Assume a test statistic $T^*(x)$ is such that

> i) $T^*(x) = F[T(x)]$, where F is unimodal in $T(x)$, and

$$\text{ii)}\quad E_0[T(x)|T^*(x) \le t^*_\alpha] = E_0[T(x)] \quad .$$

Then $T^*(x) = F[T(x)]$ is a UMSU test statistic for the hypothesis problem.

<u>Proof</u>: Let $\mathcal{C}$ be the set of test statistics which satisfy i) and ii). We first show that $T^*(x)$ is UMS among those in $\mathcal{C}$.

By hypothesis we may write

$$\begin{aligned} R &= \{x|T^*(x) \le t^*_\alpha\} \\ &= \{x|F[T(x)] \le t^*_\alpha\} \\ &= \{x|T(x) \le t_1\} \cup \{x|T(x) \ge t_2\} \end{aligned}$$

where $F(t_1) = F(t_2) = t^*_\alpha$. Also by hypothesis we have

$$\int_R p_0(x)d\mu(x) = \alpha \quad , \quad \text{and}$$

$$\int_R T(x)\; p_0(x)d\mu(x) = \alpha \cdot E_0[t(x)] \quad .$$

Therefore we may apply Theorem 4.1 which implies $T^*(x)$ is most sensitive if there exists constants $k_1(\theta, \alpha)$ and $k_2(\theta, \alpha)$ such that

$$p_\theta(x) \ge k_1 \cdot p_0(x) + k_2 \cdot t(x)\; p_0(x) \quad , \quad x\varepsilon R \quad , \quad \text{or}$$

$$r_\theta(x) \ge k_1 + k_2\; t(x) \quad , \quad x\varepsilon R \quad .$$

Consider the curve $r_\theta(x)$ graphed as a function of the real-valued $t(x)$; it is concave upwards by hypothesis. A straight line can be passed through the coordinates $(t_1, r_\theta(t_1))$ and $(t_2, r_\theta(t_2))$ where $r_\theta(t_i)$ is understood to be the value of r_θ evaluated at

$T(x) = t_i$, $i = 1, 2$, (See Fig. 4.1). Let the equation of this straight line be $y = a + bt(x)$. By the concavity of $r_\theta(x)$, we have

$$r_\theta(x) \geq a + bt(x) \quad , \quad x\varepsilon R$$
$$< a + bt(x) \quad , \quad \text{otherwise} .$$

Let $k_1(\theta, \alpha) = a$ and $k_2(\theta, \alpha) = b$. Since this can be done for each θ in Θ_A and all achievable α , then $T^*(x)$ is UMS in $\mathcal{C}$.

We now show $T^*(x)$ is unbiased. In the preceding argument we found constants $k_1(\theta, \alpha)$ and $k_2(\theta, \alpha)$ such that

$$r_\theta(x) \geq k_1 + k_2 t(x) \quad , \quad x\varepsilon R$$
$$< k_1 + k_2 t(x) \quad , \quad \text{otherwise} .$$

Define $\Theta^* = \{ \theta | k_1 (\theta, \alpha) + k_2(\theta, \alpha) E_0[T(x)|T^* \leq t^*_\alpha] \geq 1 \}$, and $\Theta\text{-}\Theta^* = \{ \theta | k_1 (\theta, \alpha) + k_2 (\theta, \alpha) E_0[T(x)|T^* \leq t^*_\alpha] < 1 \}$. Then for any θ in Θ^* ,

$$E_0[r_\theta(x)|T^* \leq t^*_\alpha] = \frac{\int_R r_\theta(x) p_0(x) d\mu(x)}{\int_R p_0(x) d\mu(x)}$$
$$\geq \frac{\int_R [k_1 + k_2 t(x)] p_0(x) d\mu(x)}{\alpha}$$
$$= k_1 + k_2 E_0[T(x)|T^* \leq t^*_\alpha]$$
$$\geq 1 .$$

Therefore, for any $\theta\varepsilon\Theta^*$, a sufficient condition for $T^*(x)$ to be unbiased is satisfied.

Consider θ in $\Theta\text{-}\Theta^*$; then

$$E_0[r_\theta(x)|T^* > t^*_\alpha] \leq \frac{\int_R [k_1+k_2\ t(x)]p_0(x)d\mu}{\int_R p_0(x)d\mu}$$

$$= k_1 + k_2\ E_0[T(x)|T^* > t^*_\alpha]$$

$$= k_1 + k_2\ E_0[T(x)|T^* \leq t^*_\alpha]$$

$$< 1 \ .$$

Therefore, by Theorem 3.2, a sufficient condition for $T^*(x)$ to be unbiased is satisfied for all $\theta\varepsilon\Theta-\Theta^*$. The statistic is then unbiased and, in conjunction with the first part of the proof, it is UMSU.

It is clear that a density $p_\theta(x)$ from the one-parameter exponential family satisfies the initial conditions of Theorem 4.2, where $T(x)$ is sufficient for θ . This theorem imp;ies that, in our search for UMSU test statistics for this family, it is sufficient to consider only unimodal functions of the sufficient statistic. The use of the sufficient statistic has not been proved necessary, nor has the use of a unimodal function of any statistic been proved necessary. However, it is still suspected by this writer that both conditions are necessary, at leastin "regular" cases.

It has been mentioned previously that the real-valued statistic $T(x)$ in the exponential density has a density which also belongs to the one-parameter exponential family. There may exist other statistics $Y(x)$ which are not sufficient for θ , but which also have densities in this family. For example in the normal case (example 2.1) when $\mu = 0$ and $\theta_0 = \sigma_0^2$, the density for $T = \sum_1^n X_i^2$ is

$$p_0(t) = \frac{t^{(n/2)-1} e^{-(t/2\sigma_0^2)}}{(2\sigma_0^2)^{n/2} \Gamma(\frac{n}{2})} , \tag{4.6}$$

and the density for $y_1 = n\bar{x}^2$ is

$$p_0(y_1) = \frac{y_1^{(1/2)-1} e^{-(y_1/2\sigma_0^2)}}{(2\sigma_0^2)^{1/2} \Gamma(\frac{1}{2})} , \tag{4.7}$$

and the density for $y_2 = \sum_1^n (x_i - \bar{x})^2$ is

$$p_0(y_2) = \frac{y_2^{((n-1)/2) - 1} e^{-(y_2/2\sigma_0^2)}}{(2\sigma_0^2)^{((n-1)/2)} \Gamma(\frac{n-1}{2})} . \tag{4.8}$$

All three densities are of the exponential family involving $Q(\sigma^2) = -1/(2\sigma^2)$, but only T is sufficient for σ_0^2 .

We now state and prove the following:

<u>Theorem 4.3</u>: For the hypothesis problem $H_0: \theta = \theta_0$ versus $H_A: \theta \neq \theta_0$, let T be a real-valued statistic such that a) $p_\theta(t)$ belongs to the one-parameter exponential family b) $T^*(t)$ is unimodal in t , c) $E_0[T | T^*(t) \leq t_\alpha^*] = E_0[T]$ for every achievable α , then

i) $T^*(t)$ is a UMSU test statistic among all test statistics based on t , and

ii) if T is sufficient for θ , then $T^*(t)$ is UMSU among all test statistics.

<u>Proof</u>: From a) a necessary condition for unbiasedness of a test

statistic T^* based on T is given by c). By using the density $p_\theta(t)$ in place of $p_\theta(x)$ in theorems 4.1 and 4.2 and letting $R = \{x | T^*(t) \leq t^*_\alpha\}$, one can argue i) in almost exactly the same manner as that used in the two theorems. Therefore, for any real-valued statistic T satisfying a), it follows that any $T^*(t)$ satisfying b) and c) is UMSU among all test statistics based on T. Part i) is proved.

If T is sufficient for θ, then there exists a function $g_\theta[t(x)]$ which depends on x only through the statistic T and a function $h(x)$ such that

$$p_\theta(x) = g_\theta[T(x)] \cdot h(x) \quad .$$

In our case, however,

$$p_\theta(x) = C(\theta)\, e^{T(x)Q(\theta)}\, h(x) \quad .$$

Therefore, if T is sufficient then it must occur as the statistic $T(x)$ in the one-parameter exponential joint density. Theorem 4.2 gives $T^*[T(x)] = T^*(t)$ as UMSU, and the theorem is proved.

By applying this theorem to the statistics $Y_1 = n\overline{X}^2$, $Y_2 = \sum_1^n (X_i - \overline{X})^2$, and $T = \Sigma X_i^2$ of the densities 4.7, 4.8, and 4.6 respectively, we have

$$T^*(y_1) = y_1^{(3/2) - 1}\, e^{-(y_1/2\sigma_0^2)}$$

is UMSU among all test statistics based on y_1;

$$T^*(y_2) = y_2^{[(n+1)/2] - 1}\, e^{-(y_2/2\sigma_0^2)}$$

is UMSU among all test statistics based on y_2; and

$$T^*(t) = t^{n/2}\, e^{-(t/2\sigma_0^2)}$$

is UMSU among all test statistics since $T = \sum_1^n X_i^2$ is sufficient for σ_0^2. The test statistics $T^*(y_1)$ and $T^*(y_2)$ were obtained in the same manner as Finley obtained $T^*(t)$.

It may be of interest to the reader to compare the statistics T and Y_1. Suppose a random sample of size 6 is taken from a normal population with mean 0 and σ^2 unknown. Let $\sigma_0^2 = 2$ under H_0. From the sample, $T = \Sigma X_i^2$ is computed to be, say, 15.6. By using the chi-square density with 6 degrees of freedom, one can compute $SL = .65$. At $\sigma^2 = 1$, $H_\theta^{y_1}(\alpha) = .69$ and $H_\theta^t(\alpha) = .84$. Therefore we see, as expected, the probability of obtaining a SL less than or equal to .65 when $\sigma^2 = 1$ is greater when using the sufficient statistic. Note that it would be possible for a sample with $\overline{x} = 0$ and $s^2 = 2$ to give a significance level of $\alpha = 0$ when using the statistic Y_1.

According to the results given in this chapter it may be concluded that the test statistics proposed by Finley in examples 2.1 and 2.4 are indeed UMSU. The test statistic in example 2.2 is also UMSU. Therefore we now have UMSU test statistics for the binomial when $p_0 = 1/2$, the normal $\mu = \mu_0 (\sigma^2$ known), and the normal $\sigma^2 = \sigma_0^2$ (μ known). The next step may naturally seem to be the normal case when both parameters are unknown. This is the purpose of the next chapter.

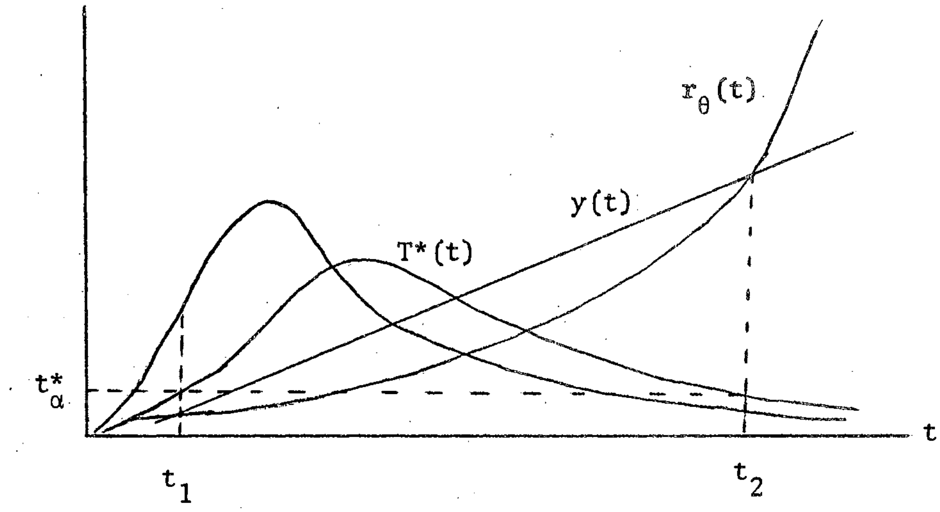

Fig. 4.1 Normal case: H_0: $\sigma = \sigma_0$ vs. H_A: $\sigma \neq \sigma_0$ ($\mu = 0$); T*(t) is chi-square density with n + 2 df.

CHAPTER V

TEST STATISTICS FOR PARAMETERS OF NORMAL DISTRIBUTION IN COMPOSITE NULL HYPOTHESES

Discussion of significance testing previously has been restricted to simple hypotheses concerning only one parameter. This, of course, is due to the requirement that the distribution of a SL be known exactly under H_0. There are still several important null hypotheses in statistics which are composite. One of these is Student's problem. In this chapter we will consider Student's problem in the context of significance testing. More generally, we will examine hypotheses involving the two parameters of the normal distribution.

Hypothesis problems considered in this chapter are

$$\text{a)} \quad H_0: \mu = 0 \ , \ \sigma^2 > 0 \qquad (5.2)$$

$$H_A: \mu > 0 \ , \ \sigma^2 > 0$$

$$\text{b)} \quad H_0: \mu = 0 \ , \ \sigma^2 > 0$$

$$H_A: \mu < 0 \ , \ \sigma^2 > 0$$

$$\text{c)} \quad H_0: \mu = 0 \ , \ \sigma^2 > 0$$

$$H_A: \mu \neq 0 \ , \ \sigma^2 > 0$$

$$\text{d)} \quad H_0: \sigma = \sigma_0 \ , \ -\infty < \mu < \infty$$

$$H_A: \sigma > \sigma_0 \ , \ -\infty < \mu < \infty$$

e) H_0: $\sigma = \sigma_0$, $-\infty < \mu < \infty$

H_A: $\sigma < \sigma_0$, $-\infty < \mu < \infty$

f) H_0: $\sigma = \sigma_0$, $-\infty < \mu < \infty$

H_A: $\sigma \neq \sigma_0$, $-\infty < \mu < \infty$

The general approach to the problems in Chapters III and IV has been parallel to Neyman-Pearson theory as treated by Lehmann. The same is true in this chapter, except that no attempt is made to give necessary and sufficient conditions for significance testing in the multiparameter case. The main objective here is to find test statistics for the hypothesis problems in (5.2) with desirable properties.

Suppose $T^*(x)$ and $T'(x)$ are two test statistics for one of the above hypothesis problems. Let $R = \{x | T^*(x) \leq t^*_\alpha\}$ and $S = \{x | T'(x) \leq t'_\alpha\}$. Note for any statistic t and any $\theta = (\mu, \sigma^2)$,

$$\begin{aligned} H^*_\theta(\alpha) &= P_\theta[SL(T^*) \leq \alpha] \\ &= E_\theta[I_R(x)] \\ &= E_\theta[E_\theta[I_R(x) | t]] \\ &= E_\theta[P_\theta[SL(T^*) \leq \alpha | t]] \quad . \end{aligned}$$

If $P_\theta[SL(T^*) \leq \alpha | t] \geq P_\theta[SL(T') \leq \alpha | t]$ for every value of t , for every achievable α, and all $\theta \varepsilon \Theta_A$, then $T^*(x)$ would be UMS.

How can such a test statistic $T^*(x)$ be found? For definiteness and clarity, let Θ_0 be the set of θ's corresponding to H_0 of 5.2a). Suppose the statistic t is sufficient for σ^2 under H_0 , and consider the conditional sample space for a fixed value of t .

It may be possible to find a conditional test statistic on this space which is UMS for the hypothesis problem. In other words, we now examine the hypothesis problem on the conditional space rather than on the unconditional sample space. The distribution of the conditional test statistic may or may not depend on t for $\theta\varepsilon\Theta_0$. If it should happen that the distribution under H_0 does <u>not</u> depend on t, then the conditional statistic has a distribution which does not depend on σ^2 nor t for $\theta\varepsilon\Theta_0$. Therefore, the distribution of this statistic is exactly known under H_0, so it satisfies the definition of a test statistic. Since it is UMS for each value of t it is UMS for the hypothesis problem.

To briefly summarize, the search is for a test statistic $T^*(x)$ such that for any $T'(x)$

$$P_\theta[SL(T^*) \le \alpha|t] \ge P_\theta[SL(T') \le \alpha|t] \quad , \quad \theta\varepsilon\Theta_A \qquad (5.3)$$

for each value of the sufficient statistic t, and

$$P_0[SL(T^*) \le \alpha|t] = P_0[SL(T') \le \alpha|t]$$

$$= \alpha \qquad , \qquad \text{a.e.t.} \qquad (5.4)$$

The search will be initiated in the conditional sample space for a fixed but arbitrary t.

UMS Test Statistics for $H_A: \mu > 0$
and $H_A: \mu < 0$, σ^2 unknown

Consider the hypothesis problem 5.2.a). For mathematical convenience, we make an orthogonal transformation from the sample space $\mathcal{X}$ to a new space $\mathcal{Y}$. In matrix notation, the transformation is defined to be

$$Y = AX \quad ,$$

where X is the $n\times 1$ column vector of the x_i, Y is the $n\times 1$ column vector of the y_i, $i = 1, \ldots, n$, and the $n\times n$ matrix A is

$$A = \begin{bmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 & \cdots & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & 0 & \cdots & 0 & 0 \\ \vdots & & & & & \vdots & \vdots \\ \frac{1}{\sqrt{n(n-1)}} & \cdots & & \frac{1}{\sqrt{n(n-1)}} & \cdots & \frac{1}{\sqrt{n(n-1)}} & \frac{-(n-1)}{\sqrt{n(n-1)}} \end{bmatrix}$$

The joint density of the y_i's is then

$$f_\theta(y_1, \ldots y_n) = (2\pi\sigma^2)^{-n/2} \exp - \frac{1}{2\sigma^2} \left\{ (y_1 - \sqrt{n}\mu)^2 + \sum_2^n y_i^2 \right\} \quad ,$$

and some interesting relationships between the x's and the y's are

$$\text{a)} \quad y_1 = \sqrt{n}\,\bar{x} \quad , \quad y_i = \frac{\sum_{j=1}^{i-1} X_j - (i-1)X_i}{\sqrt{i(i-1)}} \quad , \; i = 2,\ldots,n \; . \tag{5.5}$$

b) $E(y_1) = n\,\mu$, $E(y_i) = 0$; $i = 2, \ldots, n$,

c) $\sum_1^n y_i^2 = \sum_1^n x_i^2$, $\sum_2^n y_i^2 = \sum_1^n (x_i - \bar{x})^2$,

d) the y_i's are statistically independent with variance σ^2 , and

e) the statistic $t = \sum_1^n y_i^2$ is a sufficient and complete statistic when $\mu = 0$.

The transformation just defined is a one-to-one mapping from the original sample space $\mathcal{X}$ to the $\mathcal{Y}$ space. A hypothesis problem about the parameters for the $\mathcal{X}$ space is equivalent to the problem in the new space since the parameters are μ and σ^2 in both cases.

Since $t = \sum_1^n y_i^2$ is sufficient for θ under H_0 , we are now interested in finding the conditional distribution of the y_i's given this choice of t . To do this we make the transformation $y = y_1$, $y_2 = y_2$, ..., $y_{n-1} = y_{n-1}$, and $t = \sum_1^n y_i^2$. The absolute value of the Jacobian is

$$|J| = (\frac{1}{2})(t - \sum_1^{n-1} y_i^2)^{-1/2} ,$$

and the conditional density is given by

$$g_\theta(y_1, \ldots, y_n | t) = \frac{f_\theta(y_1, \ldots, y_{n-1}, t)}{\chi'_\theta(t)}$$

$$= \frac{\exp - \frac{1}{2\sigma^2}[(y_1 - \sqrt{n}\,\mu)^2 + t - y_1^2]}{(2\pi\sigma^2)^{n/2}[t - \sum_1^{n-1} y_i^2]^{1/2} \chi'_\theta(t)}$$

where $\chi'_\theta(t)$ is the density for the non-central chi-square distribution with n degrees of freedom and noncentrality parameter $\lambda = n\mu^2/(2\sigma^2)$.

On the conditional space consider the composite null hypothesis versus a simple alternative,

$$H_0\colon\ \mu = 0 \quad , \quad \sigma^2 > 0$$

$$H_A\colon\ \mu = \mu_1 (>0) \quad , \quad \sigma^2 = \sigma_1^2 \quad .$$

By the sufficiency of t for θ in Θ_0, the conditional density of the y_i's does not depend on σ^2. Therefore, on this conditional space H_0 is in reality a simple hypothesis. For this case of a simple hypothesis versus a simple alternative, the Neyman-Pearson lemma for significance testing gives the most sensitive test statistic to be the likelihood ratio statistic. This would be

$$LR(\cdot|t) = \frac{g_0(y_1, \ldots, y_{n-1}|t)}{g_{\theta_1}(y_1, \ldots, y_{n-1}|t)} \quad .$$

Let ℓ_α be the value of the conditional likelihood ratio statistic such that

$$P_0[LR(\cdot|t) \le \ell_\alpha|t] = \alpha \quad .$$

Since $LR(\cdot|t)$ is the most sensitive test statistic on this conditional space among all comparable test statistics, then

$$P_{\theta_1}[LR(\cdot|t) \le \ell_\alpha|t] \ge P_{\theta_1}[S^* \le s_\alpha^*|t]$$

for any other test statistic S^* on the conditional space.

For this simple alternative we check to see if the distribution of $LR(\cdot|t)$ depends on t under H_0. If it does not, then $LR(\cdot|t)$ will give a most sensitive test statistic for the hypothesis problem. If, in addition, the distribution under H_0 does not depend on the choice of $\theta_1 \varepsilon \Theta_A$, then the statistic will be UMS.

It is easy to show

$$LR(\cdot|t) = e^{-[(\sqrt{n}\mu_1/\sigma_1^2)\cdot y_1]} e^{n\mu_1^2/2\sigma_1^2} \cdot \frac{\chi'_{\theta_1}(t)}{\chi'_0(t)} ;$$

and it follows that for any θ ,

$$\begin{aligned} P_\theta[LR(\cdot|t) \le \ell_\alpha|t] &= P_\theta[e^{-\sqrt{n}\mu_1 y_1/\sigma_1^2} \le e^{-\sqrt{n}\mu_1 c_\alpha/\sigma_1^2} \mid t] \\ &= P_\theta[(-\sqrt{n}\mu_1/\sigma_1^2)y_1 \le (-\sqrt{n}\mu_1/\sigma_1^2)c_\alpha \mid t] \\ &= P_\theta[\, y_1 \ge c_\alpha \mid t] \end{aligned} \qquad (5.6)$$

when $\mu_1 > 0$. When $\theta\varepsilon\Theta_0$ this probability is equal to α .

It is seen from (5.6) that $LR(\cdot|t)$ has a conditional distribution which does not depend on the choice of θ_1 in Θ_A , therefore the conditional statistic $LR(\cdot|t)$ is UMS on the conditional space. The equation also implies the statistic $S = -y_1$ is UMS on the conditional space, but it will not do for a test statistic because its distribution is not exactly known under H_0 .

Consider the conditional distribution of y_1 given t . We recall that this is the distribution of the sample mean (apart from a constant) given the uncorrected sum of squares. It can be shown that this density is given by

$$h_\theta(y_1|t) = \frac{(2\pi\sigma^2)^{-1/2} e^{-\frac{(y_1-\sqrt{n}\mu)^2}{2\sigma^2}} \cdot t(-y_1^2)^{-3/2} e^{-\left[\frac{t-y_1^2}{2\sigma^2}\right]}}{(2\sigma^2)^{(n-1)/2}\, \Gamma(\frac{n-1}{2}) \cdot \chi'_\theta(t)} ,$$

$$-t^{1/2} < y_1 < t^{1/2} \qquad (5.7)$$

By making the change of variable $z = y_1/t^{1/2}$, (5.7) can be written as a density $h_\theta(z|t)$ with the range on z being the open interval (-1, 1). Therefore we have

$$P_\theta[LR(\cdot|t) \leq \ell_\alpha|t] = P_\theta[Z \geq z_\alpha|t] \quad . \qquad (5.8)$$

When $\theta\varepsilon\Theta_0$ it follows from (5.7)

$$h_0(z|t) = h_0(z)$$

$$= \frac{\Gamma(\frac{n}{2})(1-z^2)^{(n-3)/2}}{\sqrt{\pi}\;\;\Gamma(\frac{n-1}{2})} \quad , -1 < z < 1$$

which does not depend on t . Therefore, for $\mu = 0$ the density of $z = y_1/t^{1/2}$ does not depend on σ^2 nor t . This suggests letting our test statistic be

$$S^*(y) = \frac{-y_1}{t^{1/2}} \quad , \quad \text{where } t = \sum_1^n y_i^2 \quad .$$

It satisfies the condition that its distribution be known exactly under H_0 , and by 5.8) it is UMS for the hypothesis problem 5.2.a).

Furthermore, since

$$\frac{y_1}{t^{1/2}} = \frac{y_1}{[\sum_2^n y_i^2 + y_1^2]^{1/2}}$$

$$= \frac{y_1/(\sum_2^n y_i^2)^{1/2}}{[1 + y_1^2/\sum_2^n y_i^2]^{1/2}}$$

$$= \frac{u}{[1 + u^2]^{1/2}} \quad , \quad \text{say} \; ,$$

then

$$P_0[SL(S^*) \le \alpha] = P_0[y_1/t^{1/2} \ge c_\alpha]$$

$$= P_0\left[\frac{U}{(1+u^2)^{1/2}} \ge c_\alpha\right]$$

where u is define above. Since $u/(1+u^2)^{1/2}$ is a monotone function of u, we may write

$$P_0[SL(S^*) \le \alpha] = P_0[U \ge u_\alpha] \quad .$$

The statistic u has Student's distribution with $n-1$ degrees of freedom. We may now say a UMS test statistic for 5.2a) is given by

$$T^*(y) = \frac{-y_1}{(\sum_2^n y_i^2)^{1/2}} \quad ;$$

or in the $\mathcal{X}$ space, a UMS test statistic is

$$T^*(x) = \frac{-\sqrt{n}\,\bar{x}}{[\sum_1^n (x_i-\bar{x})^2]^{1/2}} \quad .$$

It is clear that if $\mu_1 < 0$ then the inequality in 5.6) would be reversed. It follows that a UMS test statistic for 5.2.b) would be given by

$$T^*(x) = \frac{\sqrt{n}\,\bar{x}}{[\sum_1^n (x_i-\bar{x})^2]^{1/2}} \quad .$$

For the hypothesis problem 5.2.c) we refer to the argument presented by Finley for the two-sided alternative hypotheses $H_A: \theta \neq \theta_0$ in one parameter. He shows that a UMS test statistic for that problem does not exist. His argument also applies to 5.2.c), so it can be stated that no UMS test statistic exists for $H_0: \mu = 0$, versus $H_A: \mu \neq 0$, σ^2 unknown. The search for a UMS test statistic in the smaller class of unbiased test statistics is carried out in the next section.

UMSU Test Statistic for $H_A: \mu \neq 0$, σ^2 Unknown

It was pointed out in the preceding section that a UMS test statistic for the hypothesis problem

$$H_0: \mu = 0 \quad , \quad \sigma^2 > 0$$

$$H_A: \mu \neq 0 \quad , \quad \sigma^2 > 0$$

does not exist. In this section the restriction of unbiasedness is added to the test statistics, and the existence or nonexistence of a UMSU test statistic is investigated.

It is again mathematically easier to work in the $\mathcal{Y}$ space defined in the preceding section. The plan of attack in obtaining a test statistic is essentially the same as before, except for the restriction of unbiasedness. We first go to the conditional space given the sufficient statistic to look for a test statistic which is UMSU for each value of t. In symbols, we seek a statistic $T^*(y)$ such that for each value of t and any unbiased test statistic $T'(y)$,

$$P_\theta[T^*(y) \le t^*_\alpha | t] \ge P_\theta[T'(y) \le t'_\alpha | t] \ge \alpha \quad , \quad \theta \varepsilon \Theta_A \quad ,$$

when

$$P_0[T^*(y) \le t^*_\alpha | t] = P_0[T'(y) \le t'_\alpha | t]$$

$$= \alpha \quad , \quad \text{a.e.t.} \quad .$$

It then follows that

$$P_\theta[T^*(y) \leq t^*_\alpha] \geq P_\theta[T'(y) \leq t'_\alpha] \geq \alpha \quad , \qquad \theta \varepsilon \Theta_A$$

which implies $T^*(y)$ is a UMSU test statistic.

Let $S^*(\cdot|t)$ be a test statistic derived on the conditional space for a fixed t, and let $R_t = \{y | S^*(\cdot|t) \leq s^*_\alpha , \ t = \sum_1^n y_i^2\}$. Then we have

$$P_0[S^*(\cdot|t) \leq s^*_\alpha \mid t] = \alpha \quad , \quad \text{and}$$

$$H^*_\theta(\alpha|t) = \int_{R_t} f_\theta(y_1, \ldots, y_{n-1}|t) dy_1 \ldots dy_{n-1} \quad ,$$

$$= \int_{R_t} \frac{\exp - [1/2\sigma^2][t - 2\sqrt{n}\, \mu y_1 + n\mu^2]}{(2\pi\sigma^2)^{n/2} \ (t - \sum_1^{n-1} y_i^2)^{1/2} \cdot \chi'_\theta(t)} \, dy_1, \ldots \, dy_{n-1}$$

Since it is desired to impose the restriction that our test statistic be unbiased on this conditional space, then for each σ^2 we must have $H^*_\theta(\alpha|t)$ to be a minimum at $\mu = 0$, or

$$\left. \frac{\partial H^*_\theta(\alpha|t)}{\partial \mu} \right|_{\mu = 0} = 0 \quad .$$

It can be shown that

$$\frac{\partial H^*_\theta(\alpha|t)}{\partial \mu} = \int_{R_t} (\sqrt{n}\, y_1/\sigma^2) f_\theta(y_1, \ldots, y_{n-1}|t) \, dy_1 \ldots dy_{n-1}$$

$$- \int_{R_t} (n\mu/\sigma^2) K(\theta, t) \cdot f_\theta(y_1, \ldots, y_{n-1}|t) \, dy_1 \ldots, dy_{n-1} \quad ,$$

where $K(\theta, t)$ is defined at $\mu = 0$, then this implies we must have for unbiasedness of $S^*(\cdot|t)$,

$$E_0[Y_1 I_{R_t}(y)|t] = 0 \quad .$$

On the conditional space consider two conditions on the test statistics,

$$\text{i)} \int_{R_t} f_0(y_1, \ldots, y_{n-1}|t)dy_1, \ldots, dy_{n-1} = \alpha \quad , \quad \text{and}$$

$$\text{ii)} \int_{R_t} y_1 f_0(y_1, \ldots, y_{n-1}|t)dy_1, \ldots, dy_{n-1} = 0 \quad .$$

Restriction i) just states that only comparable test statistics are considered; ii) gives a necessary condition for $S^*(\cdot|t)$ to be unbiased in the conditional space.

Let $\mathcal{C}_t$ be the set of all test statistics which satisfy the necessary conditions i) and ii). For fixed t, we attempt to apply the theorems of Chapter IV to the conditional space. A sufficient condition for $S^*(\cdot|t)$ to be UMS among those in $\mathcal{C}_t$ is that there exist constants $k_1(\theta,\alpha,t)$ and $k_2(\theta,\alpha,t)$ such that

$$\frac{f_\theta(y_1, \ldots, y_{n-1}|t)}{f_0(y_1, \ldots, y_{n-1}|t)} \geq k_1 + k_2 y_1 \quad , \quad y \in R_t \quad ,$$

or

$$c(\theta,t) \exp\left[\left(\frac{\sqrt{n}\mu}{\sigma^2}\right) \cdot y_1\right] \geq k_1 + k_2 y_1 \quad , \quad y \in R_t \quad , \qquad (5.9)$$

where $\theta \varepsilon \Theta_A$.

Let $S^*(\cdot|t) = F[u(y)]$ where $u(y) = y_1$ for fixed t. Then since $R_t = \{y|F(y_1) \leq f_\alpha \; , \; t = \sum_1^n y_i^2\}$, and letting $R_t^* = \{y_1|y\varepsilon R_t\}$, it follows that

$$E_0[Y_1 I_{R_t}|t] = \int_{R_t} y_1 \left\{\frac{f_0(y_1,\ldots,y_{n-1}, t)}{\chi'_\theta (t)}\right\} dy_2,\ldots,dy_{n-1}\, dy_1$$

$$= \int_{R'_t} y_1 \left[\frac{f_0(y_1, t)}{\chi'_\theta (t)}\right] dy_1$$

$$= E_0[Y_1 I_{R'_t}|t] \quad .$$

Condition ii) now can be viewed as requiring the conditional expectation of y_1 over the significance region to be equal to zero when the statistic is based on y_1 .

Observe now the symmetry of the conditional density of y_1 about $y_1 = 0$ when $\theta\varepsilon\Theta_0$ (or equivalently, when $\mu = 0$) ,

$$f_0(y_1|t) = \frac{\Gamma(\frac{n}{2}) \left[1 - \frac{y_1^2}{t}\right]^{(n-3)/2}}{\sqrt{\pi}\ \Gamma(\frac{n-1}{2})} \quad , \quad -t^{1/2} < y_1 < t^{1/2} \quad .$$

If one choses $S^*(\cdot|t) = F[u(y)] = (1 - y_1^2/t)^{(n-3)/2}$, then $R'_t = \{y_1: c_\alpha \le |y_1| \le t^{1/2}\}$ is a two-tail region symmetric about $y_1 = 0$. It is clear that for this choice of $S^*(\cdot|t)$ both conditions i) and ii) are satisfied. Therefore $S^*(\cdot|t)$ is a member of $\mathcal{C}_t$, and it is a unimodal function of y_1 .

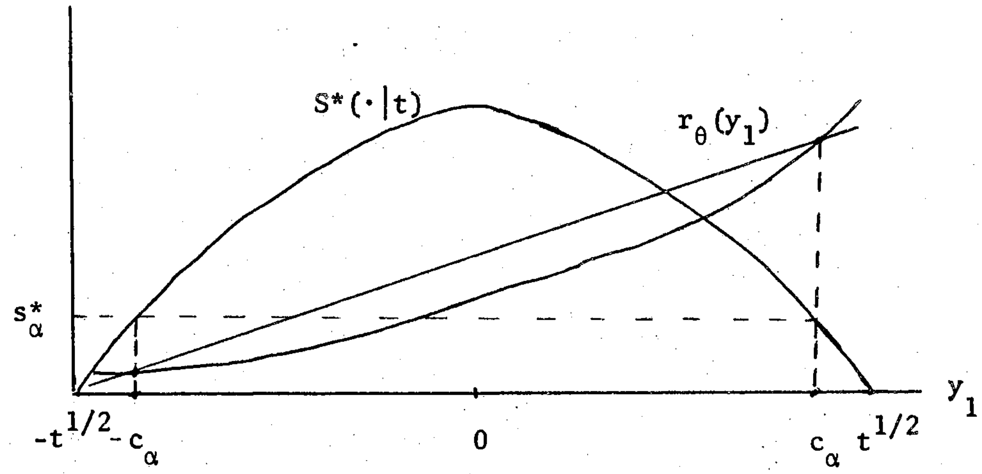

Fig 5.1 UMSU Conditional Test Statistic $S^*(\cdot|t)$

From inequality (5.9) it is seen that $f_\theta(\cdot|t)/f_0(\cdot|t)$ is concave upwards in y_1 for fixed t. Since this is true, and since $S^*(\cdot|t)$ is unimodal in y_1, then application of Theorem 4.2 on the conditional space gives $S^*(\cdot|t) = (1 - y_1^2/t)^{(n-3)/2}$ to be a UMSU conditional test statistic.

The distribution of this test statistic for arbitrary θ is given by

$$P_\theta[S^*(\cdot|t) \leq s^*_\alpha|t] = P_\theta[|Y_1| \geq c_\alpha|t]$$

$$= P_\theta\left[\frac{|Y_1|}{t^{1/2}} \geq \frac{c_\alpha}{t^{1/2}}\middle|t\right]$$

$$= P_\theta[|Z| \geq z_\alpha|t]$$

where $z = y_1/t^{1/2}$ is discussed in the preceding section. Therefore

$$P_\theta[S^*(\cdot|t) \leq s^*_\alpha|t] = P_\theta[|U| \geq u_\alpha|t]$$

where u has Student's distribution with $n-1$ degrees of freedom.

Since the distribution of u is independent of both t and σ^2 for $\theta \varepsilon \Theta_0$, then this implies a statistic based on u which allows all α will satisfy the requirements of a test statistic for the hypothesis problem. Furthermore, the above results state that if the statistic is unimodal and symmetric about $u = 0$, then the statistic will be UMSU for the problem.

The density for Student's distribution for $n-1$ degrees of freedom is

$$f_{n-1}(u) = \frac{1}{\sqrt{\pi(n-1)}} \cdot \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})} \cdot [1 + u^2/(n-1)]^{-n/2} .$$

If we let

$$T^*(y) = [1 + u^2/(n-1)]^{-n/2} , \quad u = \frac{y_1}{(\sum_{2}^{n} y_i^2)^{1/2}} ,$$

then $T^*(y)$ is a UMSU test statistic for the hypothesis problem 5.2.c).

UMS Test Statistic For H_A: $\sigma > \sigma_0$, H_A: $\sigma < \sigma_0$ When μ Is Unknown

The purpose of this section is to find a UMS test statistic for hypothesis problems 5.2.d) and 5.2.e). There will be no need to make a transformation on the $\mathcal{X}$ space in this and the following section. The procedure is again the same as in the two preceding sections. The statistic $\bar{X}$ is sufficient for θ when $\theta \varepsilon \Theta_0$; hence the search for a test statistic on the conditional space can be undertaken since the conditional density under H_0 is independent of μ. The goal is to find a test statistic $T^*(x)$ such that for any test statistic T'

$$P_\theta[SL(T^*) \le \alpha | t] \ge P_\theta[SL(T') \le \alpha | t] \quad , \quad \theta \varepsilon \Theta_A$$

for each value of t , when

$$P_0[SL(T^*) \le \alpha | t] = P_0[SL(T') \le \alpha | t]$$

$$= \alpha \qquad\qquad a.e.t.$$

For fixed $t = \overline{x}$, consider the following composite null hypothesis and simple alternative hypothesis in the reduced space:

$$H_0: \sigma = \sigma_0 \quad , \quad \mu \varepsilon (-\infty, \infty)$$

$$H_A: \sigma = \sigma_1 (>\sigma_0) \quad , \quad \mu = \mu_1 \quad .$$

The most sensitive test statistic for this problem in the conditional space is the likelihood ratio statistic,

$$LR(\cdot | \overline{x}) = \frac{f_0(x_1, \ldots, x_{n-1} | \overline{x})}{f_1(x_1, \ldots, x_{n-1} | \overline{x})} \quad .$$

For arbitrary θ ,

$$f_\theta(x_1, \ldots, x_{n-1} | \overline{x}) = \frac{\exp \; (-1/2\sigma^2) \sum_1^n (x_i - \overline{x})^2}{n^{1/2} \; (2\pi\sigma^2)^{(n-1)/2}} \tag{5.10}$$

and

$$P_\theta[LR(\cdot | \overline{x}) \le \ell_\alpha \mid \overline{x}]$$

$$P_\theta \left[(\sigma_1/\sigma_0)^{n-1} \exp\{(-1/2) \sum_1^n (X_i - \overline{x})^2 \, [1/\sigma_0^2 - 1/\sigma_1^2]\} \right.$$
$$\left. \le (\sigma_1/\sigma_0)^{n-1} \exp\{(-c_\alpha/2)[1/\sigma_0^2 - 1/\sigma_1^2]\} \mid \overline{x} \right]$$

Therefore, the most sensitive conditional statistic $LR(\cdot | \overline{x})$ is UMS on the conditional space, and its distribution is the same as that of a statistic based on the corrected sum of squares. Our next step would be to find the conditional density function of the corrected sum of squares given the sample mean, but they are statistically independent in this case. Therefore we may write for arbitrary θ

$$P_\theta[LR(\cdot|\overline{x}) \le \ell_\alpha|\overline{x}] = P_\theta[\sum_1^n (X_i - \overline{X})^2 \ge c_\alpha]$$

$$= P_\theta\left[\sum_1^n \frac{(X_i - \overline{X})^2}{\sigma_0^2} \ge c_\alpha/\sigma_0^2\right]$$

$$= P_\theta[V \ge v_\alpha]$$

where v has the chi-square distribution with $n-1$ degrees of freedom.

It follows that a UMS test statistic for hypothesis problem 5.2.d) is given by

$$T^*(x) = \frac{-\sum_1^n (x_i - \overline{x})^2}{\sigma_0^2} \quad ,$$

and a UMS test statistic for 5.2.e) is given by

$$T^*(x) = \frac{\sum_1^n (x_i - \overline{x})^2}{\sigma_0^2} \quad .$$

UMSU Test Statistic For H_A: $\sigma \neq \sigma_0$, μ Unknown

It has been discussed that a UMS significance test is not possible for the hypothesis problem

$$H_0: \sigma = \sigma_0 \quad , \quad \mu \in (-\infty,\infty)$$

$$H_A: \sigma \neq \sigma_0 \quad , \quad \mu \in (-\infty,\infty)$$

As in the third section of this chapter, we shall add the restriction of unbiasedness and look for a UMS test statistic in that class of statistics. This investigation will be initiated in the conditional space of $\mathcal{X}$ for fixed $\bar{x}$, where it is hoped to find a test statistic that is UMSU for each value of $\bar{x}$.

Let $\bar{x} = t$, and let $S^*(\cdot|t)$ be a conditional test statistic. Define $R_t = \{x | S^*(\cdot|t) \leq s^*_\alpha \, , \, t = \sum_1^n x_i/n\}$, and consider

$$P_0[S^*(\cdot|t) \leq s^*_\alpha|t] = \alpha \quad , \quad \text{and}$$

$$H^*_\theta(\alpha|t) = \int_{R_t} f_\theta(x_1, \ldots, x_{n-1}|t) dx_1, \ldots, dx_{n-1} \quad .$$

In order for $S^*(\cdot|t)$ to be unbiased on the conditional space it must be true that $H^*_\theta(\alpha|t|)$ is a minimum at $\sigma = \sigma_0$. After substituting the conditional density found in (5.10), one can easily show

$$\frac{\partial H^*_\theta(\alpha|t)}{\partial \sigma^2} = \int_{R_t} \frac{\left\{ -\frac{(n-1)}{2\sigma^2} + \frac{\sum_1^n (x_i-\bar{x})^2}{2(\sigma^2)^2} \right\} e^{\frac{-\sum_1^n (x_i-\bar{x})^2}{2\sigma^2}} dx_1, \ldots, dx_{n-1}}{(2\pi\sigma^2)^{(n-1)/2} n^{1/2}}$$

$$= E_\theta \left\{ \left[\frac{\sum_1^n (x_i-\bar{x})^2}{2(\sigma^2)^2} - \frac{(n-1)}{2\sigma^2} \right] I_{R_t} \;\Big|\; t \right\}$$

Setting this equal to zero at $\sigma = \sigma_0$ implies a necessary condition for $S^*(\cdot|t)$ to be unbiased on the conditional space is

$$E_0 \left[\sum_1^n (X_i - \bar{x})^2/\sigma_0^2 \; I_{R_t} \;\Big|\; t \right] = \alpha(n-1)$$

But since $\Sigma(x_i-\overline{x})^2$ is statistically independent of $t = \overline{x}$, this condition is equivalent to

$$E_0[W/\sigma_0^2 \; I_S] = \alpha(n-1) \tag{5.11}$$

where $w(x) = \sum_1^n (x_i - \overline{x})^2$, and $S = w(R_t)$.

Since w/σ_0^2 has the chi-square distribution with $n-1$ degrees of freedom, then it follows from Chapter IV that the statistic

$$T^*(x) = w^{\frac{n+1}{2} - 1} e^{-w/2\sigma_0^2} \quad , \quad w = \sum_1^n (x_i - \overline{x})^2 \quad ,$$

satisfies (5.11). Furthermore, since its distribution does not depend on $\overline{x}$ nor μ, $T^*(x)$ is a test statistic for the hypothesis problem, and by Theorem 4.3, $T^*(x)$ is a UMSU test statistic. This is almost the same result as that found for example 2.1, except that here the significance levels are computed from the chi-square density with $n-1$ degrees of freedom instead of n. This is in accordance with the findings of Lehmann.

CHAPTER VI

EXTENSIONS

Many of the more important hypothesis problems have been attempted. It is admitted that not all questions originally asked have been answered adequately. A prime example is the role of the sufficient statistic in significance testing. As Finley pointed out, there does not seems to be an argument in significance testing analogous to the one in Neyman-Pearson theory for basing a test on the sufficient statistic. But this writer feels there still may be possibilities for favorable results on this subject. One approach is to write the joint density in the factored form $g_\theta[T(x)]\cdot h(x)$ as mentioned in the introduction and to study $H_\theta^*(\alpha)$ under the assumption a most sensitive test statistic exists.

There is still need for work on unbiasedness in significance testing. The question of the necessity of two-tail test statistics has not been adequately answered, either in the continuous or discrete case. Arguments for necessary and sufficient conditions for the existence of unbiased test statistics are very limited.

There are interesting problems for this theory which have not been considered. One is comparing the means of two independent normal populations with equal but unknown variances. Next comes the question of an analogy in significance testing to the F-test in the analysis of variance. That is, are we justified in using the F-test in our context of significance testing? One could always tackle the Behrens-Fisher problem, where the variances are assumed unequal. A significance test for independence in a bivariate normal population would also be of interest.

The above suggestions are for additional counterparts of Neyman-Pearson theory in the context of significance testing. There are methods of comparing test statistics other than by the basic criterion "most sensitive" used here. Another way is to study and compare the asymptotic properties of various test statistics. An approach of this kind has been advanced by Bahadur [2].

Bahadur Efficiency

Bahadur has developed a theory for comparing test statistics asymptotically when they hold certain properties. Consider the test statistic $T^{(1)}(x) = T_n^{(1)}(x)$ as a term of a sequence $\{T_n^{(1)}(x)\}$. If two test sequences satisfy the proper conditions for increasing n, then a comparison of statistic 1 versus statistic 2 can be made by the asymptotic properties of the sequences and can be stated in terms of asymptotic efficiency.

Consider the following definition of a <u>standard sequence</u> by Bahadur. The sequence $\{T_n\}$ is a standard sequence for testing $\theta = \theta_0$ is (I) $\lim_{n\to\infty} F_n(x) = F(x)$ for every x, where F is a continuous c.d.f. (II) There exists a constant k, $0 < k < \infty$, such that $\log[1 - F(x)] \to -kx^2/2$ as $x \to \infty$. (III) There exists a function b on Θ such that $b(\theta_0) = 0$ and $0 < b(\theta) < \infty$ for $\theta \neq \theta_0$, and such that $\{T_n/\sqrt{n}\}$ is a consistent estimate of b.

Suppose $\{T_n\}$ is a standard sequence. Bahadur argues T_n has the asymptotic distribution F when $\theta = \theta_0$, but otherwise $T_n \to \infty$ in probability. Consequently, large values of T_n are significant when T_n is regarded as a test statistic for H_0. Note that this is in contrast to the definition of a test statistic in this paper. Bahadur then defines $SL_n = 1 - F[T_n]$ to be the level attained by T_n for a given sample; in other words, the significance level is computed from the asymptotic distribution of T_n, not from the actual distribution. Even though SL_n is only an approximate level Bahadur argues the study of such levels appear legitimate and useful.

Bahadur observes that in typical cases SL_n is asymptotically uniformly distributed over (0, 1) when $\theta \varepsilon \Theta_0$. When $\theta \varepsilon \Theta_A$, he maintains that it is also typical that

$$SL_n \to 0 \quad \text{as} \quad n \to \infty \tag{6.1}$$

with probability one. It is asserted that the "rate" at which (6.1) occurs when $\theta \neq \theta_0$ is an indication of asymptotic efficiency of T_n against that θ.

It is pointed out that in typical cases (6.1) occurs exponentially fast. This is, in essence, what condition II is stating:

$$SL_n = e^{-\frac{aT_n^2}{2}[1 + 0(1)]} \qquad \text{as} \quad T_n \to \infty \ ,$$

where it is recalled that for $\theta \neq \theta_0$, $T_n \to \infty$ in probability. Suppose that there exists a parametric function $c(\theta)$ defined over Θ_A such that $0 < c(\theta) < \infty$ and

$$\frac{\log SL_n}{n} \to -\frac{1}{2}c(\theta) \quad \text{as} \quad n \to \infty \tag{6.2}$$

with probability one when $\theta \neq \theta_0$. The function $c(\theta)$ is called the "slope" of the sequence $\{T_n\}$.

An intuitive reason for using slopes in comparing test statistics can now be given. Let $T_n^{(1)}$ and $T_n^{(2)}$ be two test sequences with slopes $c_1(\theta)$ and $c_2(\theta)$. Then suppose $c_1(\theta) > c_2(\theta)$ where

$$\frac{-2 \log SL_n^{(1)}}{n} \to c_1(\theta) \quad ,$$

$$\frac{-2 \log SL_n^{(2)}}{n} \to c_2(\theta) \quad , \quad \text{as} \quad n \to \infty \quad .$$

It is reasoned that $-\log SL_n^{(1)} \to \infty$ at a "faster" rate than $-\log SL_n^{(2)}$ or $SL_n^{(1)} \to 0$ faster than $SL_n^{(2)}$. Hence, it would be judged that $T_n^{(1)}$ is better, asymptotically, than $T_n^{(2)}$.

Bahadur argues that the ratio $c_1(\theta)/c_2(\theta)$ serves as the asymptotic efficiency of $T_n^{(1)}$ relative to $T_n^{(2)}$ when $\theta \neq \theta_0$. If $c_1/c_2 > 1$, then $T_n^{(1)}$ is judged more efficient than $T_n^{(2)}$ i.e., it will take a larger sample size for $T_n^{(2)}$ to attain a given significance level than $T_n^{(1)}$.

It would be highly desirable to compare some of the test statistics offered in this paper by the Bahadur method, for instance, the unbiased test statistics $T_1^*(t) = t^{n/2} e^{-t/2\sigma_0^2}$, and $T_2^*(y) = y^{-1/2} e^{-y/2\sigma_0^2}$ mentioned in Chapter IV. Unfortunately, this writer has not been able to show these statistics satisfy the necessary conditions of the standard sequence. It would seem that the relatively simple unbiased test statistic $T^*(x) = -|x - n/2|/(n/4)^{1/2}$ for the binomial problem might be a likely candidate for a standard sequence, but this has not been established by this writer.

Thus, it seems a major difficulty in incorporating Bahadur efficiency into this theory of significance testing may be in showing the test statistics satisfy the requirements for Bahadur's test sequences. It is felt that more investigation in this may bring observations which are more definite.

If significance testing does not lend easily to standard sequences then it may be possible to get a measure of asymptotic efficiency in some other manner. Bahadur notes that there is a formal connection between the asymptotic slope of a standard sequence and the asymptotic power of the corresponding sequence of tests. This presents the question of a corresponding relationship in an asymptotic theory of significance testing involving asymptotic sensitivity of test statistics.

It has been suggested that other asymptotic comparisons of test statistics may be possible. Consider two test statistics $T_n^{(1)}$ and $T_n^{(2)}$, where for a given sample

$$SL_n^{(1)} = P[T_n^{(1)} \leq \text{observed}]$$

$$SL_n^{(2)} = P[T_n^{(2)} \leq \text{observed}] \quad .$$

One may then define the statistic $D_n^{(1,2)} = SL_n^{(1)} - SL_n^{(2)}$ and study its asymptotic properties when $\theta \neq \theta_0$. For instance, if for $\theta \neq \theta_0$

$$D_n^{(1,2)} \to d \quad , \quad \text{as } n \to \infty$$

with probability one, then we could rate $T_n^{(1)}$ inferior, equal, or superior to $T_n^{(2)}$ when d is $>$, $=$, or <0 , respectively. Similarly, it may be possible to study ratios instead of differences. Another interesting question might be whether or not two test statistics $T_n^{(1)}$ and $T_n^{(2)}$ being equally efficient according to Bahadur implies that $D_n^{(1,2)} \to 0$ in probability.

CHAPTER VII

SUMMARY

The purpose of this investigation was to advance a theory of significance testing. Primary emphasis was placed on examining the concept of unbiasedness for the two-sided alternative hypothesis. Work was done on significance tests for composite null hypotheses in the normal case with no parameters assumed known. No attempt was made to discuss significance testing and the philosophies of statistical inference.

Examples of unbiased test statistics for both continuous and discrete densities were given in Chapter II. Attention was called to certain properties and characteristics of the examples. In particular, it was shown there exists an unbiased test statistic for the discrete case which allowed an infinite number of achievable significance levels. There also exist one-tail and no-tail unbiased statistics for the two-sided alternative.

Necessary conditions for unbiased test statistics were examined in Chapter III. Investigation was concentrated on probability densities with strict monotone likelihood ratio. It was shown for densities with this property a one-tail unbiased test statistic does not exist. Unbiasedness was studied when the probability density was a continuous unimodal function of θ for fixed value of x.

Chapter IV dealt with sufficient conditions for a test statistics to be uniformly most sensitive unbiased. It was shown for the one-parameter exponential density that a unimodal function of the sufficient statistic satisfying certain conditions was UMSU. In particular, the test statistics in examples 2.1, 2.2, and 2.5 are UMSU.

All work on this theory prior to Chapter V had been restricted to hypotheses concerning one parameter. In that chapter hypotheses are considered concerning one of the parameters in the normal distribution with the other assumed unknown, thereby creating a composite null hypotheses. It was shown that for one-sided alternative hypotheses on the mean μ, the familiar one-sided t-test is UMS in significance testing. The equal-tails t-test is UMSU for the two-sided alternative $H_A: \mu \neq 0$, $\sigma^2 > 0$. For testing σ^2 the familiar one-sided chi-square tests were shown to be optimum. For the hypothesis problem $H_A: \sigma \neq \sigma_0$, $\mu \varepsilon (-\infty, \infty)$, a two-tail test statistic based on the chi-square density with $n+1$ degrees of freedom was shown to be UMSU.

A SELECTED BIBLIOGRAPHY

Anscombe, F. J. "Tests of Goodness of Fit." Journal of the Royal Statistical Society. Series B. Vol. 25 (1963), 81-94.

Bahadur, R. R. "Stochastic Comparison of Tests." Annals of Mathematical Statistics. Vol. 31 (1960), 276-295.

Bahadur, R. R. "Rates of Convergence of Estimates and Test Statistics." Annals of Mathematical Statistics. Vol. 38 (1967), 303-324.

Dempster, A. P. and Schatzoff, M. "Expected Significance Level as a Sensitivity Index for Test Statistics." Journal of the American Statistical Association. Vol. 60 (1965), 420-436.

Finley, R. D. "A Theory of Significance Testing." Unpublished Ph.D. Thesis (1968). Oklahoma State University.

Fisher, R. A. Statistical Methods and Scientific Inference. Hafner Publishing Co. New York, N. Y. (1956).

Fraser, D. A. S. Nonparametric Methods in Statistics. John Wiley and Sons Inc., New York, N. Y. (1957).

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

Neyman, J. and Pearson, E. S. "On the Problem of the Most Efficient Tests of Statistical Hypotheses." Transactions of the Royal Society of London. Series A. Vol. 231 (1933), 289-337.

Wilks, S. S. Mathematical Statistics. John Wiley and Sons, Inc., New York, N. Y. (1963).

APPENDIX

Lemma 1: Let $f(x)$ be strictly increasing and let $g(x)$ be nondecreasing and non-constant in a real random variable x with probability density $p_\theta(x)$. If $E_\theta[f(x)] = \mu_f$ and $E_\theta\, g(x) = \mu_g$ are finite, then $\text{cov}[f(x), g(x)] > 0$.

Proof: Let x_0 be the point such that

$$f(x) - \mu_f < 0 \quad , \quad x \le x_0$$

$$f(x) - \mu_f \ge 0 \quad , \quad x > x_0 \quad .$$

Suppose there exists an $x_1 < x_0$ such that $g(x_1) < g(x_0)$, then $g(x) - g(x_0) < 0$ for every $x \le x_1$. Since $\text{cov}(f, g) = E_\theta[g(f-\mu_f)]$ can be expressed as

$$\text{cov}[f(x), g(x)] = E_\theta[g(x) - g(x_0)][f(x) - \mu_f] \quad ,$$

then by denoting $g_0 = g(x_0)$ it follows that

$$\text{cov}(f,g) = \int_{-\infty}^{x_1} (g-g_0)(f-\mu_f)\, p_\theta \, d\mu + \int_{x_1}^{x_0} (g-g_0)(f-\mu_f)\, p_\theta \, d\mu$$

$$+ \int_{x_0}^{\infty} (g-g_0)(f-\mu_f)\, p_\theta \, d\mu \, .$$

For the interval $(-\infty, x_1]$ the integrand is positive, and for the two intervals $(x_1, x_0]$, (x_0, ∞) the integrand is nonnegative, which implies

$$\operatorname{cov}(f,g) \geq \int_{-\infty}^{x_1} (g-g_0)(f-\mu_f)\, p_\theta \, d\mu$$

$$> 0 \quad .$$

Suppose there exists an $x_2 > x_0$ such that $g(x_2) > g(x_0)$, then $g(x)-g(x_0) > 0$ for all $x \geq x_2$. With modification of the above argument the result again follows.

Lemma 2: If $p_\theta(x)$ is SMLR in x and $\psi(x)$ is nondecreasing and non-constant with finite expected value, then $E_{\theta'}\psi > E_\theta\psi$ for $\theta' > \theta$.

Proof: Since

$$E_{\theta'}\psi = \int \psi p_{\theta'} \, d\mu$$

$$= \int \psi r_{\theta'} \, p_\theta \, d\mu$$

$$= E_\theta[\psi \, r_{\theta'}] \quad ,$$

then $E_{\theta'}\psi > E_\theta\psi$ if and only if $E_\theta[\psi \, r_{\theta'}] > E_\theta\psi$, and since $E_\theta[r_{\theta'}]=1$ this is equivalent to $E_{\theta'}\psi > E_\theta\psi$ if and only if $\operatorname{cov}(r_{\theta'}, \psi) > 0$. Direct application of the preceding lemma completes the proof.

VITA

Billy John Moore

Candidate for the Degree of

Doctor of Philosophy

Thesis: ON A THEORY OF SIGNIFICANCE TESTING

Major Field: Statistics

Biographical:

Personal Data: Born in San Pedro, California, June 26, 1943, the son of Clyde Joseph and Mildred Moore.

Education: Attended elementary and junior high school in Hot Springs, Arkansas; graduated from Arlington High School in Arlington, Texas, 1961; received the Bachelor of Science degree from Oklahoma State University, with a major in mathematics and statistics in May, 1965; received the Master of Science degree from Oklahoma State University, with a major in mathematics and statistics in May, 1967; completed requirements for the Doctor of Philosophy degree in May, 1969, at Oklahoma State University.

Professional Experience: Graduate assistant in the Department of Mathematics and Statistics, Oklahoma State University, 1968-1969.

Professional Organizations: American Statistical Association.