DETECTING A CHANGE IN THE MEAN
OF A NORMAL DISTRIBUTION AT
AN UNKNOWN TIME POINT

By

HENRY FELICIANO MAGALIT

Bachelor of Science
University of the Philippines
Quezon City, Philippines
1954

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1966

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, 1970

DETECTING A CHANGE IN THE MEAN
OF A NORMAL DISTRIBUTION AT
AN UNKNOWN TIME POINT

Thesis Approved:

L. D. Broemeling
Thesis Adviser

P. Larry Claypool

Robert D Morrison

Joe W White Emma

D. Durham
Dean of the Graduate College

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to Dr. Lyle D. Broemeling for his guidance and invaluable assistance in the preparation of this thesis. I express my gratitude to Dr. Robert D. Morrison for his careful reading of and suggestions in this thesis, and for serving as chairman of my advisory committee.

I express my appreciation to Dr. Joe V. Whiteman and Dr. P. Larry Claypool for serving on my committee. I wish also to thank Dr. Dioscoro L. Umali, Dean, College of Agriculture, University of the Philippines for making it possible for me to study abroad. And most of all, I wish to acknowledge my wife, Connie, for encouraging me to work for a doctoral degree.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Detecting a shift in distribution of a sequence of observations is important in such applications as medical diagnosis, quality control, and "tracking" problems. In analyzing a sequence of observations, we often encounter a shift or change in the parameters of the distribution. If the point or time at which the shift takes place is known, the magnitude of the shift is easily estimated by classical statistical methods. However, if the point at which the shift occurs is not known, two problems arise: (1) that of detecting the shift and estimating the point at which the shift occurs and (2) that of estimating the magnitude of the shift.

Two general approaches, namely a Bayes procedure and a non-parametric technique have been applied to these problems. Page [1], working in the area of quality control, developed a non-parametric procedure to detect a change in the location parameter of a sequence of observations having arbitrary distribution functions. Chernoff and Zachs [2] used a Bayes procedure to detect a change in the mean of a normal distribution. The latter authors constructed an "ad hoc" procedure to estimate the time of shift. They also compared the power function of Page's test procedure with the power function of a Bayes test procedure for different alternatives involving a binomial population. In a later work, Kander and Zachs [3] generalized the distribution to a one-parameter exponential family by deriving the Bayes procedure for detecting a shift at an unknown time point. Mustafi [4] extended the work of Kander and Zachs under the assumption that the probabilities of change remain constant over time. He derived a sequence of estimators of the current mean.

Page's procedure does not give any estimate of the magnitude of the shift nor of the time at which the shift takes place. The form of Bayes estimator is very complex, and only an "ad hoc" procedure was developed to estimate the point of shift. Thus there is a need to develop simple estimation procedures for estimating the magnitude of the shift and the point of shift to be of practical use, which are among the objectives of this research.

Statement of Problem

Let $x_1, x_2, \ldots, x_n$ be a sequence of n independent random variables, where we are interested in detecting a change or shift in the mean of these observations. Suppose we know one of these two situations is true, either there is no change or there is only one change in this distribution. Then we must distinguish between the following hypotheses:

$$H_0: x_i \sim \text{NID}(\theta_0, \sigma^2) \quad i = 1, 2, \ldots, n,$$

and

$$H_1: x_i \sim \text{NID}(\theta_0, \sigma^2) \quad i = 1, 2, \ldots, m \quad (1 \leq m \leq n-1)$$

$$x_i \sim \text{NID}(\theta_1, \sigma^2) \quad i = m+1, \ldots, n$$

where m and θ_1 are assumed unknown.

If H_0 holds, there has been no change, while if H_1 is true, there has been exactly one change in the mean, and it is located between the m^{th} and $(m+1)^{th}$ observations. We refer to m as the shift point or time of shift.

In this research a likelihood-ratio test (LRT) is developed to detect a shift, and the maximum likelihood principle is used to estimate the time where the shift occurs and the magnitude of the shift. This study tests H_0 against H_1 for the following four cases:

(1) Initial mean and variance known.

(a) One-tailed test.

(b) Two-tailed test.

(2) Initial mean unknown but variance known.

(a) One-tailed test.

(b) Two-tailed test.

(3) Initial mean known but variance unknown. Two-tailed test.

(4) Initial mean and variance unknown. Two-tailed test.

Cases (3) and (4) involve the nuisance parameter problem, which was not considered by previous authors using the parametric approach.

For each of the four cases, the likelihood-ratio test statistic is derived, the critical region of the test is defined for a given type I error, and the power function is derived for various alternative hypotheses. The one-tailed tests for cases (1) and (2) are compared to the corresponding Bayes procedure derived by Chernoff and Zacks [2]. Modified likelihood-ratio tests (MLRT) which are independent of the distribution of the estimated shift point are also constructed for cases (1) and (2). A multivariate extension is also discussed in this dissertation.

Bayes Test Procedure

Chernoff and Zacks [2] used the same testing problem as is given in the previous section. Assigning equal probabilities to the shift points m and using the Bayesian approach, they derived the test statistic and power for the following cases where the variance is one:

When the initial mean $\theta_0 = 0$, reject H_0 iff (if and only if)

$$T(x_1, x_2, \ldots, x_n) = \sum_1^n (i-1)\, x_i \geq k(\alpha) \qquad (1.1)$$

where $k(\alpha)$ is a constant depending on α, the type I error.

Under H_0, T is distributed $N[0,\ n(n-1)(2n-1)/6]$, and the power function is

$$\beta_m(\theta_1) = \Pr[\, z \geq z_\alpha - \theta_1 \{3n(n-1)/4n-2\}^{\frac{1}{2}} \{1 - m(m-1)/n(n-1)\}\,] \qquad (1.2)$$

for $0 \leq \theta_1 < \infty$, where z is distributed $N[0, 1]$ and z_α is its upper $100\alpha\%$ point.

When the initial mean θ_0 is unknown, reject H_0 iff

$$T^*(x_1, x_2, \ldots, x_n) = \sum_1^n (i-1)(x_i - \bar{x}_n) \geq k'(\alpha) \qquad (1.3)$$

where

$$\bar{x}_n = \sum_1^n x_i / n$$

and $k'(\alpha)$ is a constant depending on α. Under H_0, T^* is distributed $N[0,\ n(n-1)(n+1)/12]$, and the power function is

$$\beta^*_m(\theta_0, \theta_1) = \Pr[\, z \geq z_\alpha - (\theta_1 - \theta_0)\, m\,(n-m)/\{n(n^2-1)/3\}^{\frac{1}{2}}\,] \qquad (1.4)$$

for $0 \leq \theta_1 - \theta_0 < \infty$, where z and z_α are defined above.

CHAPTER II

SHIFT PROBLEMS WHEN THE VARIANCE IS KNOWN

In this chapter we will develop, as was done by Chernoff and Zacks using a Bayesian approach, tests to detect a change or shift in the mean of a sequence of observations from a normal distribution with a known variance. We will use a likelihood-ratio statistic in order to detect at most one change in the mean of the distribution; while the principle of maximum likelihood is used to estimate the shift point and magnitude of the shift. This estimator because of its simplicity, is more desirable than the "ad hoc" estimator proposed by Chernoff and Zacks. As we will see, using the likelihood-ratio test (LRT) involves a complicated distribution for the power function. This was avoided by using the Bayes procedure.

Case 1: Initial Mean and Variance Known

Consider

$$H_0 : x_i \sim N(\theta_0, \sigma^2) \quad i = 1, 2, \ldots, n \qquad (2.1)$$

against the

$$H_1 : x_i \sim N(\theta_0, \sigma^2) \quad i = 1, 2, \ldots, m$$
$$x_i \sim N(\theta_1, \sigma^2) \quad i = m+1, \ldots, n$$

where m and θ_1 are unknown, $1 \leq m \leq n-1$.

Under H_1, the likelihood function for θ_1 $(\sigma^2 = 1)$ is

$$f_1(\theta_1, m) = (2\pi)^{-n/2} \exp - 1/2 [\sum_{i=1}^{m} (x_i - \theta_0)^2 + \sum_{i=m+1}^{n} (x_i - \theta_1)^2] \quad (2.2)$$

Suppose m is some particular value, say s, the conditional likelihood function is

$$f_{1s}(\theta_1, s) = (2\pi)^{-n/2} \exp - 1/2 [\sum_{i=1}^{s} (x_i - \theta_0)^2 + \sum_{i=s+1}^{n} (x_i - \theta_1)^2]. \quad (2.3)$$

Differentiating $\ln f_{1s}$ with respect to θ_1 gives the conditional maximum likelihood estimate (MLE) of θ_1 as

$$\sum_{s+1}^{n} x_i/(n-s),$$

denoted as $\bar{x}_{s+1,n}$.

Now consider the maxima of the estimated likelihood functions

$$\max_m \hat{f}_{1m} = \max_m \{(2\pi)^{-n/2} \exp - 1/2 [\sum_{i=1}^{m} (x_i - \theta_0)^2 + \sum_{i=m+1}^{n} (x_i - \bar{x}_{m+1,n})^2]\}$$

$$= (2\pi)^{-n/2} \exp - 1/2 [\sum_{i=1}^{s} (x_i - \theta_0)^2 + \sum_{i=s+1}^{n} (x_i - \bar{x}_{s+1,n})^2]$$

$$= f_{1s}(\hat{\theta}_1, s). \quad (2.4)$$

It can be shown that one can compute the MLE of $m, \hat{m}$, by finding

$$\max_m [(n-m)(\bar{x}_{m+1,n} - \theta_0)^2].$$

Thus,

$$\Pr[\hat{m} = s] = \Pr[\max_m (n-m)(\bar{x}_{m+1,n} - \theta_0)^2 = (n-s)(\bar{x}_{s+1,n} - \theta_0)^2]. \quad (2.5)$$

If for a given sequence of observations the estimate, $\hat{m}$, of the shift point is s, then the estimate of the change in the mean is $\overline{x}_{s+1,n} - \theta_0$.

Under H_0, the likelihood function ($\sigma^2 = 1$) is

$$f_0 = (2\pi)^{-n/2} \exp - 1/2 [\sum_{i=1}^{n} (x_i - \theta_0)^2].$$

Now consider the likelihood ratio

$$L = \frac{f_0}{f_{1\hat{m}}(\theta_1, \hat{m})}$$

$$= \exp - 1/2 [\sum_{i=1}^{n} (x_i - \theta_0)^2 - \sum_{i=1}^{\hat{m}} (x_i - \theta_0)^2 - \sum_{i=\hat{m}+1}^{n} (x_i - \overline{x}_{\hat{m}+1,n})^2]$$

$$= \exp - 1/2 [(n - \hat{m}) (\overline{x}_{\hat{m}+1,n} - \theta_0)^2].$$

Thus,

$$-2\ln L = (n - \hat{m}) (\overline{x}_{\hat{m}+1,n} - \theta_0)^2. \qquad (2.6)$$

The test statistic is given by Equation (2.6).

Therefore, reject H_0 if the test statistic, denoted as T_n,

$$T_n = (n - \hat{m}) (\overline{x}_{\hat{m}+1,n} - \theta_0)^2 \geq k(\alpha), \qquad (2.7)$$

where $k(\alpha)$ is a constant depending on α, the type I error.

Two-Tailed Test of Case 1 for $H_1 : \theta_1 \neq \theta_0$

To find the test statistic for this case we need to find the conditional distribution of T_n; i.e., the distribution of

$$[T_n | \hat{m}=s] = [(n-s)(\overline{x}_{s+1,n} - \theta_0)^2],$$

denoted as T_s.

In matrix notation T_s may be written as

$$(n-s)(\bar{x}_{s+1,n} - \theta_0)^2 = X_1' A X_1 \qquad (2.8)$$

where

$$\underset{n \times n}{A} = \frac{1}{n-s} \left(\begin{array}{c|c} \phi & \phi \\ \hline \phi & J_{n-s}^{n-s} \end{array} \right)$$

and

$$\underset{1 \times n}{X_1'} = [(x_1 - \theta_0), \ldots, (x_n - \theta_0)]$$

and J_n^t is an $n \times t$ matrix of ones.

In order to get the distribution of T_s we shall use the following corollary by Graybill [5]:

Corollary I: If X_1 is distributed $N(\mu, \sigma^2 I)$, then $X_1' A X_1$ is distributed as noncentral chi-square, $\chi'^2(K, \lambda)$, where $\lambda = \mu' A \mu / 2\sigma^2$, if and only if A is idempotent of rank K.

Let μ' be a $1 \times n$ vector where the first m elements are zeros and the last (n-m) elements are $\theta_1 - \theta_0$.

For each s, where $1 \leq s \leq m$, it can be shown that $\mu' A \mu / 2\sigma^2$ is

$$\lambda = (n-m)^2 (\theta_1 - \theta_0)^2 / 2(n-s)\sigma^2. \qquad (2.9)$$

For each s, where $m + 1 \leq s \leq n - 1$, it can also be shown that $\mu' A \mu / 2\sigma^2$ is

$$\lambda^* = (n-s)(\theta_1 - \theta_0)^2 / 2\sigma^2. \qquad (2.10)$$

Under H_0, $\theta_1 = \theta_0$, thus $\lambda = \lambda^* = 0$. Since $A^2 = A$ and trace $A = 1$, A is idempotent of rank 1. It follows from Corollary I, that

$X_1'AX_1$ is distributed as a central chi-square variate, $\chi^2(1)$, with one degree of freedom.

The null hypothesis will be rejected if and only if

$$T_n = (n - \hat{m})(\bar{x}_{\hat{m}+1,n} - \theta_0)^2 \geq k(\alpha), \quad (2.11)$$

where $k(\alpha)$ is the upper $100\alpha\%$ point, $\chi_\alpha^2(1)$, of the chi-square distribution with one degree of freedom.

Under H_1, $X_1'AX_1$ is distributed $\chi'^2(1, \lambda)$ for $1 \leq s \leq m$ and $\chi'^2(1, \lambda^*)$ for $m+1 \leq s \leq n-1$, where λ and λ^* are given by (2.9) and (2.10), respectively.

Using the above result, the power of the test $\beta_{\hat{m}}(\theta_1)$ for given θ_1 and m is

$$\beta_m(\theta_1) = \Pr[(n-\hat{m})(\bar{x}_{\hat{m}+1,n} - \theta_0)^2 \geq \chi_\alpha^2(1)]$$

$$= \sum_{s=1}^{m} \Pr[(n-s)(\bar{x}_{s+1,n} - \theta_0)^2 \geq \chi_\alpha^2(1)] \cdot \Pr[\hat{m}=s]$$

$$+ \sum_{s=m+1}^{n-1} \Pr[(n-s)(\bar{x}_{s+1,n} - \theta_0)^2 \geq \chi_\alpha^2(1)] \cdot \Pr[\hat{m}=s] \quad (2.12)$$

where $\Pr[\hat{m}=s]$ is given by (2.5). To obtain the power, one must know the distribution of $\hat{m}$. This will be discussed later in the chapter.

One-Tailed Test of Case 1 for $H_1: \theta_1 > \theta_0$

From Equation (2.11) we know that T_n is distributed $\chi^2(1)$ under H_0.

Thus, $(T_n)^{1/2}$ is distributed N(0, 1), and we have

$$\alpha = \Pr[(n-\hat{m})^{1/2}(\bar{x}_{\hat{m}+1,n} - \theta_0) \geq k'(\alpha)] = \Pr[z \geq z_\alpha] \quad (2.13)$$

where z is distributed N(0, 1) and $k'(\alpha)$ is the upper $100\alpha\%$ point, z_α, of the standard normal distribution.

The power of the test for given m and θ_1 is

$$\beta_m(\theta_1) = \Pr[(n-\hat{m})^{1/2}(\bar{x}_{\hat{m}+1,n} - \theta_0) \geq z_\alpha]$$

$$= \sum_{s=1}^{m} \Pr[(n-s)^{1/2}(\bar{x}_{s+1,n} - \theta_0) \geq z_\alpha] \cdot \Pr[\hat{m}=s]$$

$$+ \sum_{s=m+1}^{n-1} \Pr[(n-s)^{1/2}(\bar{x}_{s+1,n} - \theta_0) \geq z_\alpha] \cdot \Pr[\hat{m}=s] \tag{2.14}$$

where $\Pr[\hat{m}=s]$ is given by Equation (2.5). Let us proceed to simplify Equation (2.14).

Consider $1 \leq s \leq m$, $x_i \sim N(\theta_0, 1)$, $i = 1, 2, \ldots, m$, and $x_i \sim N(\theta_1, 1)$, $i = m+1, \ldots, n$. It can be shown that

$$(\bar{x}_{s+1,n} - \theta_0) \sim N[(n-m)(\theta_1 - \theta_0)/(n-s),\ 1/(n-s)] \tag{2.15}$$

and

$$\Pr[(n-s)^{1/2}(\bar{x}_{s+1,n} - \theta_0) \geq z_\alpha] = \Pr[z \geq z_\alpha - (n-m)(\theta_1 - \theta_0)/(n-s)^{1/2}]. \tag{2.16}$$

When $m+1 \leq s \leq n-1$, it can be shown that

$$(\bar{x}_{s+1,n} - \theta_0) \sim N[(\theta_1 - \theta_0),\ 1/(n-s)] \tag{2.17}$$

and

$$\Pr[(n-s)^{1/2}(\bar{x}_{s+1,n} - \theta_0) \geq z_\alpha] = \Pr[z \geq z_\alpha - (n-s)^{1/2}(\theta_1 - \theta_0)]. \tag{2.18}$$

Using results from Equations (2.16) and (2.18) in Equation (2.14), we have

$$\beta_m(\theta_1) = \sum_{s=1}^{m} \Pr[z \geq z_\alpha - (n-m)(\theta_1 - \theta_0)/(n-s)^{1/2}] \cdot \Pr[\hat{m}=s]$$

$$+ \sum_{s=m+1}^{n-1} \Pr[z \geq z_\alpha - (n-s)^{1/2}(\theta_1 - \theta_0)] \cdot \Pr[\hat{m}=s]. \qquad (2.19)$$

The distribution of $\hat{m}$ and the evaluation of power will be discussed later.

Case 2: Initial Mean Unknown but Variance Known

If we consider the null hypothesis against the alternative hypothesis as given in (2.1), the likelihood function for θ_0, assuming $\sigma^2=1$, is

$$f_0 = (2\pi)^{-n/2} \exp - 1/2 \sum_{1}^{n} (x_i - \theta_0)^2,$$

and the MLE of θ_0 is $\bar{x}$, the sample mean.

Under H_1 the conditional likelihood function for θ_0 and θ_1, given m=s, is

$$f_{1s}(\theta_0, \theta_1, s) = (2\pi)^{-n/2} \exp - 1/2 \left[\sum_{1}^{s} (x_i - \theta_0)^2 + \sum_{s+1}^{n} (x_i - \theta_1)^2 \right].$$

The conditional MLE of θ_0 is $\bar{x}_{1,s}$ and the condition MLE of θ_1 is $\bar{x}_{s+1,n}$. The MLE of m, is given by $\hat{m}$=s if and only if

$$f_{1s}(\bar{x}_{1,s}, \bar{x}_{s+1,n}, s) = \max_m \hat{f}_{1m}$$

$$= \max_m \left[(2\pi)^{-n/2} \exp - 1/2 \left\{ \sum_{1}^{m} (x_i - \bar{x}_{1,m})^2 \right.\right.$$

$$\left.\left. + \sum_{m+1}^{n} (x_i - \bar{x}_{m+1,n})^2 \right\} \right].$$

One can obtain the maximum of the above by finding

$$\max_m \,[m\,\bar{x}^2_{1,m} + (n-m)\bar{x}^2_{m+1,n}].$$

Thus,

$$\Pr[\hat{m}=s] = \Pr[\max_m\{m\bar{x}^2_{1,m} + (n-m)\,\bar{x}^2_{m+1,n}\} = s\bar{x}^2_{1,s} + (n-s)\bar{x}^2_{s+1,n}] \tag{2.20}$$

The likelihood ratio is

$$L = \exp - 1/2[\sum_1^n (x_i - \bar{x})^2 - \sum_1^{\hat{m}} (x_i - \bar{x}_{1,\hat{m}})^2 - \sum_{\hat{m}+1}^n (x_i - \bar{x}_{\hat{m}+1,n})^2],$$

where

$$\bar{x} = [\hat{m}\bar{x}_{1,\hat{m}} + (n-\hat{m})\,\bar{x}_{\hat{m}+1,n}] \div n.$$

If one takes the log of the likelihood ratio and substitutes the value for $\bar{x}$ in ln L, he can obtain by algebraic simplification

$$-2 \ln L = \hat{m}\,\frac{(n-\hat{m})}{n}\,(\bar{x}_{\hat{m}+1,n} - \bar{x}_{\hat{m}})^2 \tag{2.21}$$

The right hand side of Equation (2.21) gives the test statistic T_n.

Thus reject H_0 if

$$T_n \geq k(\alpha). \tag{2.22}$$

Two-Tailed Test of Case 2 for $H_1: \theta_1 \neq \theta_0$

To find the distribution of T_n we need to find the conditional distribution of T_n, given m=s. This can be written in matrix notation as

$$T_s = \frac{s(n-s)}{n}\,(\bar{x}_{s+1,n} - \bar{x}_{1,s})^2 = X'AX, \tag{2.23}$$

where

$$\underset{n \times n}{A} = \left(\begin{array}{c|c} \frac{n-s}{sn} J_s^s & -\frac{1}{n} J_s^{n-s} \\ \hline -\frac{1}{n} J_{n-s}^s & \frac{s}{(n-s)n} J_{n-s}^{n-s} \end{array} \right)$$

and

$$\underset{1 \times n}{X'} = (x_1, x_2, \ldots, x_n),$$

$A^2 = A$ and trace $A = 1$. By letting μ' be a $1 \times n$ vector where the first m elements are θ_0, and the last (n-m) elements are θ_1, and using corollary I we have

$$X'AX \sim \chi'^2(1, \lambda_1), \tag{2.24}$$

where $\lambda_1 = \mu' A \mu / 2\sigma^2$.

For each s, when $1 \leq s \leq m$, it can be shown that

$$\lambda_1 = s(n-m)^2(\theta_1 - \theta_0)^2 / 2(n-s)n\sigma^2, \tag{2.25}$$

and when $m+1 \leq s \leq n-1$

$$\lambda_1^* = (n-s)m^2(\theta_1 - \theta_0)^2 / 2sn\sigma^2. \tag{2.26}$$

Under H_0, $\theta_1 = \theta_0$, and $\lambda_1 = \lambda_1^* = 0$, $X'AX$ is distributed $\chi^2(1)$ by Corollary 2.1. $K(\alpha) = \chi_\alpha^2(1)$ is the upper $100\alpha\%$ point of a chi-square with one degree of freedom.

Given a set (θ_0, θ_1, m), the power of the test is

$$\beta_m(\theta_0, \theta_1) = \sum_{s=1}^{m} \Pr\left[\frac{s(n-s)}{n}(\bar{x}_{s+1,n} - \bar{x}_{1,s})^2 \geq \chi_\alpha^2(1)\right] \cdot \Pr[\hat{m}=s]$$

$$+ \sum_{s=m+1}^{n-1} \Pr\left[\frac{s(n-s)}{n}(\bar{x}_{s+1,n} - \bar{x}_{1,s})^2 \geq \chi_\alpha^2(1)\right] \cdot \Pr[\hat{m}=s]. \tag{2.27}$$

Using (2.24), (2.25), (2.26) in (2.27) we have

$$\beta_m(\theta_0, \theta_1) = \sum_{s=1}^{m} \Pr[\chi'^2(1,\lambda_1) \geq \chi^2_\alpha(1)] \cdot \Pr[\hat{m}=s]$$

$$+ \sum_{s=m+1}^{n-1} \Pr[\chi'^2(1, \lambda_1^*) \geq \chi^2_\alpha(1)] \cdot \Pr[\hat{m}=s], \tag{2.28}$$

where $\Pr[\hat{m}=s]$ is given by Equation (2.20). The distribution of $\hat{m}$ and the evaluation of the above power function, as in the previous cases, will be discussed later.

One-Tailed Test of Case 2 for $H_1: \theta_1 > \theta_0$

From Equations (2.24), (2.25), and (2.26) we know that T_n is distributed $\chi^2(1)$ under H_0. Thus $(T_n)^{1/2}$ is distributed N(0, 1), and we have

$$\alpha = \Pr\{[\hat{m}(n-\hat{m})/n]^{1/2} [\bar{x}_{\hat{m}+1,n} - \bar{x}_{1,\hat{m}}] \geq k'(\alpha)\}$$

$$= \Pr[z \geq z_\alpha] \tag{2.29}$$

where z is distributed N(0, 1) and $k'(\alpha) = z_\alpha$ is its upper $100\alpha\%$ point.

The power of the test $\beta_m(\theta_0, \theta_1)$ depends on θ_0, θ_1 and m and is given by

$$\beta_m(\theta_0, \theta_1) = \Pr[\{\hat{m}(n-\hat{m})/n\}^{1/2} (\bar{x}_{\hat{m}+1,n} - \bar{x}_{1,\hat{m}}) \geq z_\alpha]$$

$$= \sum_{s=1}^{m} \Pr[\{s(n-s)/s\}^{1/2} (\bar{x}_{s+1,n} - \bar{x}_{1,s}) \geq z_\alpha] \cdot \Pr[\hat{m}=s]$$

$$+ \sum_{s=m+1}^{n-1} \Pr[\{s(n-s)/s\}^{1/2} (\bar{x}_{s+1,n} - \bar{x}_{1,s}) \geq z_\alpha] \cdot \Pr[\hat{m}=s] \tag{2.30}$$

Let us proceed to simplify Equation (2. 30). Consider the case when $1 \le s \le m$. It can be shown that

$$(\bar{x}_{s+1,n} - \bar{x}_{1,s}) \sim N[(n-m)(\theta_1 - \theta_0)/(n-s),\ n/s(n-s)], \qquad (2.31)$$

For the case $m+1 \le s \le n-1$, it can be verified that

$$(\bar{x}_{s+1,n} - \bar{x}_{1,s}) \sim N[m(\theta_1 - \theta_0)/s,\ n/s(n-s)]\,. \qquad (2.32)$$

Using the information in (2.31) and (2.32), Equation (2.30) can be simplified to

$$\beta_m(\theta_0, \theta_1) = \sum_{s=1}^{m} \Pr[z > z_\alpha - (n-m)\{s/n(n-s)\}^{1/2} (\theta_1 - \theta_0)] \cdot \Pr[\hat{m}=s]$$

$$+ \sum_{s=m+1}^{n-1} \Pr[z \geq z_\alpha - m\ \{(n-s)/ns\}^{1/2} (\theta_1 - \theta_0)] \cdot \Pr[\hat{m}=s] \qquad (2.33)$$

where $\Pr[\hat{m}=s]$ is given by Equation (2. 20).

Simulating the Distribution of $\hat{m}$ and Evaluation of the Power Function

From the results of Equations (2. 5) and (2. 20) one can estimate the distribution of $\hat{m}$ by simulation on a computer.

For a given pair (m, θ_1) generate m NID(0, 1) deviates $x_1, x_2, \ldots, x_m$ and n-m NID$(\theta_1, 1)$ deviates $x_{m+1}, x_{m+2}, \ldots, x_n$. Find

(a) $\max_m (n-m)\, \bar{x}^2_{m+1,n}$

and

(b) $\max_m [(n-m)\, \bar{x}^2_{m+1,n} + m\, \bar{x}^2_{1,m}]$

and record the values as $\hat{m}$ and $\hat{m}^*$, respectively.

Repeat the above procedure 250 times. We get a frequency count of the possible values of $\hat{m}$ and $\hat{m}*$.

Repeat the above procedures for different values of (m, θ_1).

This procedure gives an estimate of the distribution of $\hat{m}$ in (2.5), for which the initial mean θ_0 is known. At the same time we obtain an estimate of the distribution of $\hat{m}$ in Equation (2.20) in which the initial mean θ_0 is unknown.

Using the above simulation for the distribution of $\hat{m}$, the power function of each test in this chapter is evaluated for the following set of parameter values: $n = 12$; $m = 1, 3, 5, 7, 9, 11$; $\theta_0 = .3, .6, .9, 1.2$; $\sigma^2 = 1$; and type I error $\alpha = .05$.

Table I gives the powers for the one-tailed LRT, and the Bayes test for the case where the initial mean and variance are known. Table II gives these powers when the initial mean is unknown and variance is known to be 1. Table III shows the power of the two-tailed LRTs when the initial mean and variance are known and when the initial mean is unknown but variance is known. The power function for the two-tailed Bayes test procedure is not available.

The exact distribution of $\hat{m}$ for a sample of size three is derived in the Appendix.

TABLE I

THE POWERS OF THE ONE-TAILED ($\theta_1 > \theta_0$) LRT AND THE BAYES TEST WHEN THE INITIAL MEAN ($\theta_0 = 0$) AND VARIANCE ($\sigma^2 = 1$) ARE KNOWN FOR n = 12 AND α = .05

θ_1	m	LRT*	BAYES**
.3	1	.1766	.2222
	3	.1602	.2105
	5	.1476	.1846
	7	.1265	.1480
	9	.1017	.1066
	11	.0705	.0670
.6	1	.5082	.5459
	3	.4327	.5141
	5	.3646	.4399
	7	.2814	.3283
	9	.1932	.1991
	11	.1009	.0882
.9	1	.8303	.8403
	3	.7635	.8094
	5	.6563	.7243
	7	.5156	.5618
	9	.3410	.3283
	11	.1439	.1141
1.2	1	.9674	.9697
	3	.9442	.9569
	5	.8779	.9103
	7	.7510	.7750
	9	.5269	.4822
	11	.2072	.1450

* Obtained from Equation (2.19)
** Obtained from Equation (1.2)

TABLE II

THE POWERS OF THE ONE-TAILED ($\theta_1 > \theta_0$) LRT AND THE BAYES TEST WHEN THE INITIAL MEAN IS UNKNOWN BUT VARIANCE ($\sigma^2 = 1$) IS KNOWN FOR n = 12 AND α = .05

$\theta_1 - \theta_0$	m	LRT*	BAYES**
.3	1	.0752	.0659
	3	.0896	.0957
	5	.0946	.1139
	7	.0925	.1139
	9	.0798	.0957
	11	.0645	.0659
.6	1	.1105	.0855
	3	.1614	.1666
	5	.1761	.2216
	7	.1774	.2216
	9	.1431	.1666
	11	.1841	.0855
.9	1	.1610	.1092
	3	.2709	.2647
	5	.3190	.3715
	7	.3119	.3715
	9	.2440	.2647
	11	.1203	.1092
1.2	1	.2257	.1372
	3	.4250	.3858
	5	.5088	.5442
	7	.4878	.5442
	9	.3869	.3858
	11	.1666	.1372

* Obtained from Equation (2.33)
** Obtained from Equation (1.4)

TABLE III

THE POWERS OF THE TWO-TAILED LRT WHEN INITIAL MEAN ($\theta_0 = 0$) AND VARIANCE ($\sigma^2 = 1$) ARE KNOWN; AND WHEN INITIAL MEAN IS UNKNOWN AND VARIANCE ($\sigma^2 = 1$) IS KNOWN FOR n = 12 AND $\alpha = .05$

$\theta_1 - \theta_0$	m	Initial Mean θ_0 Known*	Initial Mean θ_0 Unknown**
.3	1	.1136	.0525
	3	.1021	.0574
	5	.0932	.0610
	7	.0801	.0616
	9	.0665	.0572
	11	.0594	.0526
.6	1	.3940	.0611
	3	.3211	.0885
	5	.2590	.1044
	7	.1894	.1088
	9	.1233	.0910
	11	.0673	.0607
.9	1	.7515	.0802
	3	.6661	.1564
	5	.5427	.2021
	7	.3971	.2071
	9	.2386	.1604
	11	.0928	.0812
1.2	1	.9455	.1093
	3	.9050	.2700
	5	.8094	.3684
	7	.6487	.3586
	9	.4094	.2795
	11	.1368	.1119

* Obtained from Equation (2.12)
** Obtained from Equation (2.28)

CHAPTER III

SHIFT PROBLEMS WHEN THE VARIANCE IS UNKNOWN

The sequence of observations in which we are interested in detecting a change in the mean is drawn very often from a normal population with an unknown variance. In this chapter we will develop tests to detect this change. We will encounter here a more complicated distribution which is the ratio of two independent noncentral chi-squares. This distribution will reduce to a mixture of Beta distributions in which the mixing distributions are the products of two Poissons having means equal to the noncentralities of the two chi-squares. The distribution of the two-tailed LRT is again a mixture of the distribution just described with another mixing distribution for $\hat{m}$. Two cases will be considered. Case 3 considers the problem when the initial mean is known but variance is unknown, and Case 4 considers the situation when the initial mean and variance are unknown.

Case 3: Initial Mean Known but Variance Unknown

Under H_1 the likelihood function for θ_1, σ^2 and m is similar to that in (2.2) except σ^2 appears in the function as seen in the following equation.

$$f_1(\theta_1, \sigma^2, m) = (2\pi\sigma^2)^{-n/2} \exp - 1/2\sigma^2 \left[\sum_{i=1}^{m} (x_i - \theta_0)^2 + \sum_{i=m+1}^{n} (x_i - \theta_1)^2 \right]. \tag{3.1}$$

In order to find the MLE of θ_1, σ^2, and m, we find the conditional maximum likelihood estimators of θ_1, and σ^2 given m=s, s=1,2,...,n-1. For m=s, let

$$\max_{\theta_1, \sigma^2} \{f_1(\theta_1, \sigma^2, s)\} = f_1(\hat{\theta}_{1s}, \hat{\sigma}^2_{1s}, s)$$

where $\hat{\theta}_{1s}$ and $\hat{\sigma}^2_{1s}$ are the conditional MLE of θ_1 and σ^2 respectively. Thus to find the MLE of m, we find

$$f_1(\hat{\theta}_1, \hat{\sigma}^2_1, \hat{m}) = \max_m [f_1(\hat{\theta}_{1m}, \hat{\sigma}^2_{1m}, m)]$$

$$= \max_m [(2\pi\hat{\sigma}^2_{1m})^{-n/2} e^{-n/2}] \qquad (3.2)$$

The index m, which maximizes the conditional maxima, $f_1(\hat{\theta}_{1m}, \hat{\sigma}^2_{1m}, m)$, is the MLE of m, denoted $\hat{m}$, and the corresponding estimates $\hat{\theta}_{1\hat{m}}$ and $\hat{\sigma}^2_{1\hat{m}}$ are the MLE of θ_1 and σ^2, denoted as $\hat{\theta}_1$ and $\hat{\sigma}^2_1$, respectively. It can be shown that

$$\hat{\theta}_1 = \bar{x}_{\hat{m}+1, n} \qquad (3.3)$$

and

$$\hat{\sigma}^2_1 = 1/n \left[\sum_1^{\hat{m}} (x_i - \theta_0)^2 + \sum_{\hat{m}+1}^{n} (x_i - \bar{x}_{\hat{m}+1, n})^2\right] \qquad (3.4)$$

It can be shown that the $\hat{m}$ found in (3.2) can also be obtained by finding

$$\max_m [(n-m)(\bar{x}_{m+1, n} - \theta_0)^2].$$

Thus,

$$\Pr[\hat{m}=s] = \Pr[\max (n-m)(\bar{x}_{m+1, n} - \theta_0)^2 = (n-s)(\bar{x}_{s+1, n} - \theta_0)^2] \qquad (3.5)$$

The same result was obtained in Equation (2.5).

Under H_0 the likelihood function is

$$f_0(\sigma^2) = (2\pi\sigma^2)^{-n/2} \exp - 1/2\sigma^2 \left[\sum_1^n (x_i - \theta_0)^2 \right].$$

When

$$\hat{\sigma}_2^2 = 1/n \left\{ \sum_1^n (x_i - \theta_0)^2 \right\}$$

is substituted in the above expression, the maximum of the likelihood function becomes

$$f_0(\hat{\sigma}_2^2) = (2\pi\hat{\sigma}_2^2)^{-n/2} e^{-n/2}.$$

The likelihood ratio L is

$$\frac{f_0(\hat{\sigma}_2^2)}{f_1(\hat{\theta}_1, \hat{\sigma}_1^2, \hat{m})} = \left(\frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2} \right)^{-n/2}.$$

Thus,

$$L^{-2/n} = \frac{\sum_1^n (x_i - \theta_0)^2}{\sum_1^{\hat{m}} (x_i - \theta_0)^2 + \sum_{\hat{m}+1}^n (x_i - \bar{x}_{\hat{m}+1,n})^2}$$

$$= 1 + \frac{(n-\hat{m})(\bar{x}_{\hat{m}+1,n} - \theta_0)^2}{\sum_1^{\hat{m}} (x_i - \theta_0)^2 + \sum_{\hat{m}+1}^n (x_i - \bar{x}_{\hat{m}+1,n})^2}.$$

The test statistic is

$$T_n = \frac{(n-\hat{m})(\bar{x}_{\hat{m}+1,n} - \theta_0)^2}{\sum_1^{\hat{m}} (x_i - \theta_0)^2 + \sum_{\hat{m}+1}^n (x_i - \bar{x}_{\hat{m}+1,n})^2}. \qquad (3.6)$$

Now let us find the distribution of $T_s = [T_n | \hat{m}=s]$. First consider the numerator of T_s. From Equation (2.8) it is seen that $X_1'AX_1$ is distributed $\chi'^2(1, \lambda)$, where

$$\underset{n \times n}{A} = \frac{1}{(n-s)} \left(\begin{array}{c|c} \phi & \phi \\ \hline \phi & J_{n-s}^{n-s} \end{array} \right),$$

and λ is the noncentrality as shown in (2.9) and (2.10).

Now consider the denominator of T_s, whose terms can be written in matrix notation as

$$\sum_{1}^{s} (x_i - \theta_0)^2 = X_1'BX_1$$

where

$$\underset{1 \times n}{X_1'} = [(x_1 - \theta_0), (x_2 - \theta_0), \ldots, (x_n - \theta_0)],$$

$$\underset{n \times n}{B} = \left(\begin{array}{c|c} I_s & \phi \\ \hline \phi & \phi \end{array} \right),$$

and

$$\sum_{s+1}^{n} (x_i - \bar{x}_{s+1,n})^2 = X'CX,$$

where

$$\underset{1 \times n}{X'} = (x_1, x_2, \ldots, x_n),$$

$$\underset{n \times n}{C} = \left(\begin{array}{c|c} \phi & \phi \\ \hline \phi & I_{n-s} - \frac{1}{(n-s)} J_{n-s}^{n-s} \end{array} \right).$$

B and C are both idempotent with trace of B equal to s and trace of C equal to n-s-1. By Corollary I $X_1'BX_1$ is distributed $\chi'^2(s, \lambda_2)$, and $X'CX$ is distributed $\chi'^2(n-s-1, \lambda_3)$, where λ_2 and λ_3 are given below.

For $1 \leq s \leq m$, it can be shown that

$$\lambda_2 = 0$$

and (3.7)

$$\lambda_3 = (n-m)(m-s)(\theta_1 - \theta_0)^2/2(n-s)\sigma^2.$$

For $m+1 \leq s \leq n-1$, it is seen that

$$\lambda_2^* = (s-m)(\theta_1 - \theta_0)^2/2(n-s)\sigma^2$$

and (3.8)

$$\lambda_3^* = 0.$$

To show that $(X_1'BX_1 + X'CX)$ is distributed $\chi'^2(n-1, \lambda_2 + \lambda_3)$, we only have to show that $X_1'BX_1$ and $X'CX$ are independent. Since $X_1'BX_1$ and $X'CX$ involve different components of the vector X, they are independent.

We have to show that the numerator and denominator of T_s are independent. The numerator $X_1'AX_1$ is a function of $\bar{x}_{s+1,n}$. The components of the denominator, $X_1'BX_1$ are functions of $x_1 - \theta_0$, $x_2 - \theta_0, \ldots, x_s - \theta_0$, and $X'CX$ is a function of $x_i - \bar{x}_{s+1,n}$, $s+1 \leq i \leq n$.

Let us prove that $\bar{x}_{s+1,n}$ is independent of

$$\begin{pmatrix} x_1 - \theta_0 \\ x_2 - \theta_0 \\ \vdots \\ x_s - \theta_0 \\ \hline x_{s+1} - \bar{x}_{s+1,n} \\ \vdots \\ x_n - \bar{x}_{s+1,n} \end{pmatrix} = \begin{pmatrix} \underset{s \times 1}{X^{(1)}} \\ \hline \underset{(n-s)\times 1}{X^{(2)}} \end{pmatrix}$$

Since $X^{(1)}$ and $\bar{x}_{s+1,n}$ involve different components of the vector X, they are independent. To prove that $\bar{x}_{s+1,n}$ and $x_i - \bar{x}_{s+1,n}$, for $s+1 \leq i \leq n$, are independent, we only have to prove that $\text{cov}[\bar{x}_{s+1,n}, (x_i - \bar{x}_{s+1,n})]$ equals zero.

$$\text{cov}[\bar{x}_{s+1,n}, (x_i - \bar{x}_{s+1,n})] = \frac{\sigma^2}{(n-s)} - \frac{\sigma^2}{(n-s)} = 0$$

for $s + 1 \leq i \leq n$. This completes the proof that the numerator of T_s is independent of its denominator.

The next problem is to find the distribution of T_s, the ratio of two independent noncentral chi-square distributions. The solution to this problem is given by Lukacs and Laha [6]. They showed that if x_1 and x_2 are two independent random variables, where x_i is distributed $\chi'^2(n_i, \lambda_i)$, the quotient of $w = x_1/x_2$ has the density function

$$f_w(x) = \sum_{r=0}^{\infty} \sum_{t=0}^{\infty} \left\{ \left(\frac{e^{-(\lambda_1+\lambda_2)/2} (\lambda_1/2)^r (\lambda_2/2)^t}{r!\, t!} \right) \left(\frac{1}{B[(n_1+2r)/2, (n_2+2t)/2]} \right) \cdot \left(\frac{x^{(n_1+2r-2)/2}}{(1+x)^{(n_1+n_2+2r+2t)/2}} \right) \right\} \quad \text{for } x > 0$$

$$= 0 \text{ otherwise,} \tag{3.9}$$

where

$$B[(n_1+2r)/2, (n_2+2t)/2] = \frac{\Gamma[(n_1+2r)/2]\, \Gamma[(n_2+2t)/2]}{\Gamma[(n_1+n_2+2r+2t)/2]} .$$

$f_w(x)$ is a mixture of random variables having Beta distributions of the second kind. Letting $x = y/(1-y)$, in (3.9) we get

$$f(y) = \sum_{r=0}^{\infty} \sum_{t=0}^{\infty} \frac{e^{-(\lambda_1+\lambda_2)/2}(\lambda_1/2)^r(\lambda_2/2)^t}{r!\ t!} \cdot \frac{y^{(n_1+2r-2)/2}(1-y)^{(n_2+2t-2)/2}}{B[(n_1+2r)/2,\ (n_2+2t)/2]} \tag{3.10}$$

for $\quad 0 < y < 1$

$= 0$ otherwise.

Since f(y) is continuous and uniformly convergent, we can interchange integral and summation signs. For any real number u, $0 < u < 1$, we have

$$\Pr(y \geq u) = \sum_{r=0}^{\infty} \sum_{t=0}^{\infty} \frac{e^{-(\lambda_1+\lambda_2)/2}(\lambda_1/2)^r(\lambda_2/2)^t}{r!\ t!} \cdot \left[\int_u^1 \frac{y^{(n_1+2r-2)/2}(1-y)^{(n_2+2t-2)/2}}{B\{(n_1+2r)/2,\ (n_2+2t)/2\}} \, dy \right]. \tag{3.11}$$

The quantity inside the bracket can be evaluated by using the incomplete Beta function.

Under H_0, T_n can be written

$$\frac{\chi_1^2(1)}{\chi_2^2(n-1)} = \frac{F(1, n-1)}{n-1} \tag{3.12}$$

since the noncentralities are zero. Thus $k(\alpha) = [F_\alpha(1, n-1)/(n-1)]$, where $F_\alpha(1, n-1)$ is the upper $100\alpha\%$ point of the central F distribution with 1 and n-1 degrees of freedom in the numerator and denominator, respectively. The critical value in the transformed variable, y, is $k'(\alpha) = k(\alpha)/[1+k(\alpha)] = [F_\alpha(1, n-1)/\{(n-1) + F_\alpha(1, n-1)\}]$.

The power function for m, θ_1, σ^2 is

$$\beta_m(\theta_1, \sigma^2) = \sum_{s=1}^{m} \{\Pr[F''(1, n-1; \lambda, \lambda_2+\lambda_3) \geq F_\alpha(1, n-1)] \cdot \Pr[\hat{m}=s]\}$$

$$+ \sum_{s=m+1}^{n-1} \{\Pr[F''(1, n-1; \lambda^*, \lambda_2^* + \lambda_3^*) \geq F_\alpha(1, n-1)] \cdot \Pr[\hat{m}=s]\}$$

$$= \sum_{s=1}^{m} \{\Pr[y \geq k'(\alpha)] \cdot \Pr[\hat{m}=s]\} + \sum_{s=m+1}^{n-1} \{\Pr[y \geq k'(\alpha)] \cdot \Pr[\hat{m}=s]\} \tag{3.13}$$

where F'' is a noncentral F with two noncentralities defined in Equations (2.9), (2.10), (3.7), and (3.8). $\Pr(y \geq k'(\alpha)$ is defined in Equation (3.11). $\Pr(\hat{m}=s)$ is defined in Equation (3.5) and was estimated by simulation. The power is tabulated in Table IV at the end of the chapter.

Case 4: Initial Mean and Variance Unknown

Under H_0, the likelihood function is

$$f_0(\theta_0, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp - 1/2\sigma^2 \sum_{i=1}^{n} (x_i - \theta_0)^2.$$

The MLE of θ_0 and σ^2, respectively, are

$$\hat{\theta}_0 = \overline{x}$$

$$\hat{\sigma}_1^2 = 1/n \sum_{i=1}^{n} (x_i - \overline{x})^2.$$

Under H_1, assuming the shift occurs at $m=s$, the conditional likelihood function is

$$f_1(\theta_{0s}, \theta_1, \sigma^2, s) = (2\pi\sigma^2)^{-n/2} \exp - 1/2\sigma^2 \sum_{i=1}^{s} (x_i - \theta_0)^2 + \sum_{i=s+1}^{n} (x_i - \theta_1)^2 .$$

The conditional MLE of θ_0, θ_1, and σ^2 respectively, are

$$\hat{\theta}_{0s} = \bar{x}_{1,s},$$

$$\hat{\theta}_{1s} = \bar{x}_{s+1,n},$$

and

$$\hat{\sigma}^2_{2s} = 1/n\,[\sum_{i=1}^{s}(x_i - \bar{x}_{1,s})^2 + \sum_{i=s+1}^{n}(x_i - \bar{x}_{s+1,n})^2].$$

To find the MLE of m we find

$$f_1(\hat{\theta}_{0\hat{m}}, \hat{\theta}_1, \hat{\sigma}^2_2, \hat{m}) = \max_m [f_1(\hat{\theta}_{0m}, \hat{\theta}_{1m}, \hat{\sigma}^2_{2m}, m)]$$

$$= \max_m [(2\pi\hat{\sigma}^2_{2m})^{-n/2}\, e^{-n/2}] \tag{3.14}$$

It can be shown that the $\hat{m}$ found in (3.14) can also be obtained by finding

$$\max_m [m\,\bar{x}^2_{1,m} + (n-m)\,\bar{x}^2_{m+1,n}].$$

Thus,

$$\Pr[\hat{m}=s] = \Pr[\max_m \{m\bar{x}^2_{1,m} + (n-m)\bar{x}^2_{m+1,n}\} = s\bar{x}^2_{1,s} + (n-s)\bar{x}^2_{s+1,n}] \tag{3.15}$$

The likelihood ratio is

$$L = \frac{f_0(\hat{\theta}_0, \hat{\sigma}^2_1)}{f_1(\hat{\theta}_{0\hat{m}}, \hat{\theta}_1, \hat{\sigma}^2_2, \hat{m})} = \left(\frac{\hat{\sigma}^2_1}{\hat{\sigma}^2_2}\right)^{-n/2},$$

and

$$L^{-2/n} = \frac{\hat{\sigma}^2_1}{\hat{\sigma}^2_2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{\sum_{i=1}^{\hat{m}}(x_i - \bar{x}_{1,\hat{m}})^2 + \sum_{i=\hat{m}+1}^{n}(x_i - \bar{x}_{\hat{m}+1,n})^2} \tag{3.16}$$

From the derivation of Equation (2.21) it can be deduced that the numerator of (3.16) is

$$\sum_{i=1}^{n}(x_i-\bar{x})^2 = \sum_{i=1}^{n}(x_i-\bar{x}_{1,\hat{m}})^2+\sum_{i=\hat{m}+1}^{n}(x_i-\bar{x}_{\hat{m}+1,n})^2+\frac{\hat{m}(n-\hat{m})}{n}(\bar{x}_{\hat{m}+1,n}-\bar{x}_{1,\hat{m}})^2.$$

After making the above substitution and simplifying,(3.16) may be written as

$$L^{-2/n} = 1+\frac{\frac{\hat{m}(n-\hat{m})}{n}(\bar{x}_{\hat{m}+1,n}-\bar{x}_{1,\hat{m}})^2}{\sum_{i=1}^{\hat{m}}(x_i-\bar{x}_{1,\hat{m}})^2+\sum_{i=\hat{m}+1}^{n}(x_i-\bar{x}_{\hat{m}+1,n})^2}.$$

The test statistic is

$$T_n = \frac{\frac{\hat{m}(n-\hat{m})}{n}(\bar{x}_{\hat{m}+1,n}-\bar{x}_{1,\hat{m}})^2}{\sum_{i=1}^{\hat{m}}(x_i-\bar{x}_{1,\hat{m}})^2+\sum_{i=\hat{m}+1}^{n}(x_i-\bar{x}_{\hat{m}+1,n})^2}. \tag{3.17}$$

Let us proceed to find the distribution of $T_s = [T_n | \hat{m} = s]$ in order to determine the critical region and power function of the test.

Consider in matrix notation the denominator of T_s.

$$\sum_{i=1}^{s}(x_i-\bar{x}_{1,s})^2+\sum_{i=s+1}^{n}(x_i-\bar{x}_{s+1,n})^2 = X'A_1X, \tag{3.18}$$

where

$$\underset{n\times n}{A_1} = \left(\begin{array}{c|c} I_s-\frac{1}{s}J_s^s & \phi \\ \hline \phi & I_{n-s}-\frac{1}{n-s}J_{n-s}^{n-s} \end{array}\right)$$

and

$$\underset{1\times n}{X'} = (x_1, x_2, \ldots, x_n).$$

It can be shown that $A_1A_1 = A_1$ and trace $A_1 = n-2$, thus A_1 is idempotent of rank n-2. By Collorary I, $X'A_1X$ is distributed $\chi'^2(n-2, \lambda_4)$, where λ_4 is the noncentrality.

Now look at the noncentrality, $\lambda_4 = \mu'A_1\mu/2\sigma^2$, of $X'A_1X$. Let μ' be a $1 \times n$ vector whose first m elements are θ_0 and the last n-m are θ_1. It can be shown that

$$\lambda_4 = \frac{(m-s)(n-m)(\theta_1 - \theta_0)^2}{2(n-s)\sigma^2} \quad \text{for } 1 < s \leq m \tag{3.19}$$

and

$$\lambda_4^* = \frac{m(s-m)(\theta_1 - \theta_0)^2}{2s\,\sigma^2} \quad \text{for } m+1 \leq s \leq n-1. \tag{3.20}$$

From (2.24) the numerator of T_s is $X'AX$ and

$$X'AX \sim \chi'^2(1, \lambda_1)$$

where λ_1 is defined in Equations (2.25) and (2.26).

$X'AX$ and $X'A_1X$ are independent because A_1 and A are idempotent and $A_1A = \phi$.

Therefore, the distribution of $T_s = (X'AX/X'A_1X)$ is similar to Equation (3.9) except for the noncentralities which are defined by Equations (2.25), (2.26), (3.19) and (3.20).

Under H_0, from (3.17) we have

$$T_n = \frac{\chi_1^2(1)}{\chi_2^2(n-2)} = \frac{F(1, n-2)}{n-2}$$

since the noncentralities are all zero. Thus

$$k(\alpha) = [F_\alpha(1, n-2)/(n-2)] .$$

The critical point in the transformed variable y is

$$k'(\alpha) = [F_\alpha(1, n-2)/\{(n-2) + F_\alpha(1, n-2)\}] .$$

The power of the test for m, θ_0, θ_1, and σ^2 is

$$\beta_m(\theta_0, \theta_1, \sigma^2) = \sum_{s=1}^{m} \{\Pr[F''(1, n-2; \lambda_1, \lambda_4) \geq F_\alpha(1, n-2)] \cdot \Pr[\hat{m}=s]\}$$

$$+ \sum_{s=m+1}^{n-1} \{\Pr[F''(1, n-2; \lambda_1^*, \lambda_4^*) \geq F_\alpha(1, n-2)] \cdot \Pr[\hat{m}=s]\}$$

$$= \sum_{s=1}^{m} \{\Pr[y \geq k'(\alpha)] \cdot \Pr[\hat{m}=s]\}$$

$$+ \sum_{s=m+1}^{n-1} \{\Pr[y \geq k'(\alpha)] \cdot \Pr[\hat{m}=s]\} \qquad (3.21)$$

where $\Pr[\hat{m}=s]$ is given by (3.15). λ_1, λ_4, λ_1^* and λ_4^* are defined in Equations (2.25), (2.26), (3.19) and (3.20). $\Pr[y \geq k'(\alpha)]$ is given by Equation (3.11) using the appropriate noncentralities.

The powers for Cases 3 and 4 are given in Table IV.

TABLE IV

THE POWERS OF TWO-TAILED LRT WHEN INITIAL ($\theta_0 = 0$) IS KNOWN AND VARIANCE IS UNKNOWN; AND WHEN THE INITIAL MEAN AND VARIANCE ARE UNKNOWN FOR n = 12 AND α = .05

$\theta_1 - \theta_0$	m	Initial Mean Known*	Initial Mean Unknown**
	1	.0978	.0515
	3	.0893	.0547
.3	5	.0831	.0574
	7	.0731	.0579
	9	.0626	.0546
	11	.0528	.0515
	1	.3203	.0569
	3	.2582	.0761
.6	5	.2098	.0874
	7	.1555	.0913
	9	.1051	.0784
	11	.0625	.0565
	1	.6251	.0700
	3	.5465	.1246
.9	5	.4369	.1576
	7	.3159	.1622
	9	.1917	.1277
	11	.0812	.0709
	1	.7460	.0903
	3	.6999	.2075
1.2	5	.6541	.2842
	7	.5219	.2730
	9	.3244	.2169
	11	.1142	.0923

* Obtained from Equation (3.13)

** Obtained from Equation (3.21)

CHAPTER IV

OTHER TEST PROCEDURES AND EXTENSIONS

Up to this point, as in Chapters II and III, the distributions of the LRT is a function of $\hat{m}$ and hence the power functions were evaluated by estimating the distribution of $\hat{m}$. We will consider in this chapter test statistics with distributions that do not depend on $\hat{m}$. For sample size of n we will encounter n-1 correlated noncentral chi-squares distribution. This is the distribution of the two-tailed modified likelihood-ratio test (MLRT) for Case 1 and 2. A multivariate extension of Case 1 is discussed.

A MLRT When the Initial Mean and Variance are Known for $H_1 : \theta_1 \neq \theta_0$

If the shift point is known, say at m = s, a significant difference in the change of the mean may be detected by a likelihood ratio statistic,

$$T_s = \frac{(n-s)(\bar{x}_{s+1,n} - \theta_0)^2}{\sigma^2}, \quad s = 1, 2, \ldots, n-1 .$$

A procedure to detect a shift at an unknown time point can be devised by averaging the test statistic, T_s, over all possible shifts. If the arithmetic average is used, this implies that the shift can occur with equal probability at any one of the n-1 possible time points. To this end, for $\sigma^2 = 1$, we define

$$T = \sum_{s=1}^{n-1} (n-s)(\bar{x}_{s+1,n} - \theta_0)^2 \tag{4. 1}$$

as a modified likelihood-ratio statistic, which is the sum of n-1 chi-squares, each with one degree of freedom.

The distribution of T is approximated by equating the first two moments of T to that of a scaled chi-square distribution having the form $a\chi^2(b)$ and then solving the two equations for a and b. The first and second moments are

$$E\left(\sum_{s=1}^{n-1} X'A_sX\right) = E[a\chi^2(b)] = ab \tag{4. 2}$$

and

$$\text{Var}\left[\sum_{s=1}^{n-1} X'A_sX\right] = \text{Var}[a\chi^2(b)] = 2a^2b. \tag{4. 3}$$

Writing Equation (4. 1) in matrix notation we get

$$\sum_{s=1}^{n-1} (n-s)(\bar{x}_{s+1,n} - \theta_0)^2 = \sum_{s=1}^{n-1} X'A_sX \tag{4. 4}$$

where

$$\underset{n \times n}{A_s} = \frac{1}{(n-s)} \left(\begin{array}{c|c} \phi & \phi \\ \hline \phi & J_{n-s}^{n-s} \end{array}\right)$$

and

$$\underset{1 \times n}{X'} = [(x_1 - \theta_0), \ldots, (x_n - \theta_0)] .$$

Koch [7] has shown that if X is distributed $N(\mu, I)$ then

$$E(X'AX) = \text{tr}A + \mu'A\mu \tag{4. 5}$$

and

$$\mathrm{Cov}(X'A\ X, X'B\ X) = 2\mathrm{tr}AB + 4\mu'AB\mu, \tag{4.6}$$

where A and B are real symmetric matrices.

From (4.5) we have

$$ab = \sum_{s=1}^{n-1} E(X'A_sX) = \sum_{s=1}^{n-1} \mathrm{tr}A_s + \sum_{s=1}^{n-1} \mu'A_s\mu \tag{4.7}$$

$$ab = n - 1 + \sum_{s=1}^{n-1} \mu'A_s\mu \tag{4.8}$$

since trace $A_s = 1$ for all s. Let μ' be a $1 \times n$ vector whose first m elements are zeros and last (n-m) are $\theta_1 - \theta_0$. It can be shown that

$$\mu'A_s\mu = \frac{(n-m)^2}{n-s} (\theta_1 - \theta_0)^2 \quad \text{if } 1 \leq s \leq m$$

and

$$= (n-s)(\theta_1 - \theta_0)^2 \quad \text{if } m+1 \leq s \leq n-1.$$

Thus

$$\sum_{s=1}^{n-1} \mu'A_s\mu = \sum_{s=1}^{m} \frac{(n-m)^2}{n-s} (\theta_1 - \theta_0)^2 + \sum_{s=m+1}^{n-1} (n-s)(\theta_1 - \theta_0)^2. \tag{4.9}$$

Using (4.8) and (4.9) we have

$$ab = (n-1) + \sum_{s=1}^{m} \frac{(n-m)^2}{n-s} (\theta_1 - \theta_0)^2 + \sum_{s=m+1}^{n-1} (n-s)(\theta_1 - \theta_0)^2. \tag{4.10}$$

In order to evaluate (4.3) we need the identity

$$\mathrm{Var}\left(\sum_{s=1}^{n-1} X'A_sX\right) = \sum_{s=1}^{n-1} \mathrm{Var}(X'A_sX) + 2 \sum_{s=1}^{n-2} \sum_{t=s+1}^{n-1} \mathrm{Cov}(X'A_sX, X'A_tX) \tag{4.11}$$

and

$$\underset{s<t}{A_sA_t} = \frac{1}{(n-s)}\left(\begin{array}{c|c} \phi & \phi \\ \hline \phi & J^{n-t}_{n-s} \end{array}\right).$$

Let us proceed by evaluating the second term of the right-hand side of (4.11). From (4.6) we get

$$2\sum_{s=1}^{n-2}\sum_{t=s+1}^{n-1} \mathrm{Cov}(X'A_sX, X'A_tX) = 4\sum_{s=1}^{n-2}\sum_{t=s+1}^{n-1} \mathrm{tr}A_sA_t + 8\sum_{s=1}^{n-2}\sum_{t=s+1}^{n-1} \mu'A_sA_t\mu. \tag{4.12}$$

In order to simplify the second term of (4.12) we consider

$$\mu'A_sA_t\mu = \frac{(n-m)^2}{(n-s)}(\theta_1-\theta_0)^2 \quad \text{for} \quad 1 \le s \le m$$

and

$$= (n-t)(\theta_1-\theta_0)^2 \quad \text{for } m+1 \le s \le n-1.$$

Equation (4.12) then simplifies to

$$2\sum_{s=1}^{n-2}\sum_{t=s+1}^{n-1} \mathrm{Cov}(X'A_sX, X'A_tX) = 4\sum_{s=1}^{n-2}\sum_{t=s+1}^{n-1}\left(\frac{n-t}{n-s}\right) + 8\sum_{s=1}^{m}\sum_{t=s+1}^{n-1}\frac{(n-m)^2}{(n-s)}(\theta_1-\theta_0)^2$$

$$+8\sum_{s=m+1}^{n-2}\sum_{t=s+1}^{n-1}(n-t)(\theta_1-\theta_0)^2 \tag{4.13}$$

Now consider the first term of (4.11). If we let s = t, then the first term becomes a special case of the second term. Thus we have

$$\sum_{s=1}^{n-1} \mathrm{Var}(X'A_sX) = 2\sum_{s=1}^{n-1} \mathrm{tr}\,A_s^2 + 4\sum_{s=1}^{n-1}\mu'A_s\mu \tag{4.14}$$

$$= 2\sum_{s=1}^{n-1}\left(\frac{n-s}{n-s}\right) + 4\sum_{s=1}^{m}\frac{(n-m)^2}{(n-s)}(\theta_1-\theta_0)^2$$

$$+ 4\sum_{s=m+1}^{n-1}(n-s)(\theta_1-\theta_0)^2. \tag{4.15}$$

By substituting the information given in (4. 13) and (4. 15), (4. 11) becomes

$$\mathrm{Var}\left(\sum_{s=1}^{n-1} X'A_sX\right) = 2(n-1)+4\sum_{s=1}^{m}\frac{(n-m)^2}{(n-s)}(\theta_1-\theta_0)^2+4\sum_{s=m+1}^{n-1}(n-s)(\theta_1-\theta_0)^2$$

$$+4\sum_{s=1}^{n-2}\sum_{t=s+1}^{n-1}\left(\frac{n-t}{n-s}\right)+8\sum_{s=1}^{m}\sum_{t=s+1}^{n-1}\frac{(n-m)^2}{(n-s)}(\theta_1-\theta_0)^2$$

$$+8\sum_{s=m+1}^{n-2}\sum_{t=s+1}^{n-1}(n-t)(\theta_1-\theta_0)^2$$

$$= 2a^2b \qquad (4.16)$$

Therefore using (4. 10) and (4. 16) we can solve for a and b. Under H_0 the solution is

$$a_0 = n/2$$

and

$$b_0 = \frac{2(n-1)}{n} .$$

Thus reject H_0 iff

$$T = \sum_{s=1}^{n-1}[(n-s)(\bar{x}_{s+1,n}-\theta_0)^2] \geq a_0\chi^2(b_0), \qquad (4.17)$$

where $\chi^2_\alpha(b_0)$ is the upper $100\alpha\%$ point of the chi-square distribution with b_0 degrees of freedom.

The approximate power function for θ_1 and m is

$$\beta_m(\theta_1) = \Pr\left[\sum_{s=1}^{n-1}(n-s)(\bar{x}_{s+1,n}-\theta_0)^2 \geq a_0\chi^2(b_0)\right]$$

$$= \Pr[a\chi^2(b) \geq a_0\chi^2(b_0)] \qquad (4.18)$$

where a and b are functions of θ_1, m and n. The approximate power is tabulated in Table V at the end of the chapter.

A MLRT When Initial Mean and Variance

Are Known for $H_1 : \theta_1 > \theta_0$

If the shift occurs at m=s (s=1, 2, . . . , n-1), the test statistic,

$$T_s = \sqrt{n-s} \; (\bar{x}_{s+1,n} - \theta_0) \quad 1 \le s \le n-1$$

may be used to detect a difference in the means, then averaging over the n-1 possible values of s,

$$T = \sum_{s=1}^{n-1} [\sqrt{n-s} \; (\bar{x}_{s+1,n} - \theta_0)]$$

$$= \sum_{s=1}^{n-1} \left[\frac{1}{\sqrt{n-s}} \sum_{i=s+1}^{n} (x_i - \theta_0) \right]$$

$$= \sum_{i=2}^{n} \left[\left(\sum_{j=1}^{i-1} \frac{1}{\sqrt{n-j}} \right) (x_i - \theta_0) \right] \tag{4.19}$$

may be used to detect a shift at an unknown time point.

Under H_0 we have

$$E(T) = 0$$

since $E(x_i - \theta_0) = 0$ for all i, and

$$\mathrm{Var}(T) = \mathrm{Var} \left[\sum_{i=2}^{n} \left(\sum_{j=1}^{i=1} \frac{1}{\sqrt{n-j}} \right) (x_i - \theta_0) \right]$$

$$= \sum_{i=2}^{n} \left(\sum_{j=1}^{i-1} \frac{1}{\sqrt{n-j}} \right)^2$$

since variance x_i equals one for all i.

If we let

$$T' = \frac{T}{\sqrt{\sum_{i=2}^{n}\left(\sum_{j=1}^{i-1}\frac{1}{\sqrt{n-j}}\right)^2}}, \tag{4.20}$$

H_0 is rejected iff $T' \geq z_\alpha$, where z_α is the upper $100\alpha\%$ point of $N(0, 1)$.

To investigate the power of the test statistic, T is written as

$$T = \sum_{i=2}^{m}\left[\left(\sum_{j=1}^{i-1}\frac{1}{\sqrt{n-j}}\right)(x_i - \theta_0)\right] + \sum_{i=m+1}^{n}\left[\left(\sum_{j=1}^{i-1}\frac{1}{\sqrt{n-j}}\right)(x_i - \theta_0)\right], m \geq 2.$$

Consider the mean and variance of T. We have

$$E(T) = \sum_{i=m+1}^{n}\left[\left(\sum_{j=1}^{i-1}\frac{1}{\sqrt{n-j}}\right)(\theta_1 - \theta_0)\right], \quad m \geq 1$$

denoted by μ.

$$\text{Var } T = \sum_{i=2}^{n}\left(\sum_{j=1}^{i-1}\frac{1}{\sqrt{n-j}}\right)^2$$

denoted as v.

Therefore,

$$z = \frac{T-\mu}{\sqrt{v}} \sim N(0, 1) \tag{4.21}$$

and the power for m and θ_1 is

$$\beta_m(\theta_1) = \Pr[z \geq z_\alpha - \frac{\mu}{\sqrt{v}}] \tag{4.22}$$

where z_α is the upper $100\alpha\%$ point of $N(0, 1)$.

Refer to Table V for power calculations.

A MLRT When the Initial Mean is Unknown but Variance is Known for $H_1 : \theta_1 \neq \theta_0$

In a similar manner to the case when the initial mean is known,

$$T_s = \frac{s(n-s)}{n} (\bar{x}_{s+1,n} - \bar{x}_{1,s})^2$$

is the LRT for detecting a change in the mean of the shift point $m=s(s=1,2,\ldots,n-1)$.

Letting $T_s = X'A_sX$, where X is the $1 \times n$ vector with components x_i, and A_s is defined in Equation (2.23),

$$T = \sum_{s=1}^{n-1} X'A_sX \tag{4.23}$$

is used to detect a shift in the mean at an unknown time point.

To approximate the power of the test, we equate the first two moments of T to those of $a\chi^2(b)$, where a and b are constants to be determined. Equating the moments gives (4.7) and (4.11). Let us proceed to simplify (4.7) and (4.11).

First consider the second term of (4.11) which is (4.12). To simplify (4.12) we need the identity

$$\underset{s<t}{A_sA_t} = \frac{1}{n^2} \left(\begin{array}{c|c} \frac{n(n-t)}{t} J_s^t & -n \; J_s^{n-t} \\ \hline \frac{-sn(n-t)}{t(n-s)} J_{n-s}^t & \frac{ns}{n-s} J_{n-s}^{n-t} \end{array} \right) \tag{4.24}$$

and

$$\text{tr } A_sA_t = \frac{(n-t)s}{t(n-s)} . \tag{4.25}$$

Let μ' be a $1 \times n$ vector whose first m elements are θ_0 and the last n-m elements are θ_1. It can be shown that the second term of

(4. 12) simplifies to

$$8 \sum_{s=1}^{n-2} \sum_{t=s+1}^{n-1} \mu' A_s A_t \mu = 8 \sum_{s=1}^{m} \sum_{t=s+1}^{n-1} \frac{s(n-m)^2}{n(n-s)} (\theta_1 - \theta_0)^2$$

$$+ 8 \sum_{s=m+1}^{n-2} \sum_{t=s+1}^{n-1} \frac{m^2(n-t)}{nt} (\theta_1 - \theta_0)^2. \tag{4.26}$$

Similarly one can show that the first term of (4. 12) is

$$4 \sum_{s=1}^{n-2} \sum_{t=s+1}^{n-1} \operatorname{tr} A_s A_t = 4 \sum_{s=1}^{n-2} \sum_{t=s+1}^{n-1} \frac{(n-t)s}{t(n-s)} . \tag{4.27}$$

Before combining the terms let us first consider the first term of (4. 11) which is (4. 14). The second term of (4. 14) simplifies to

$$4 \sum_{s=1}^{n-1} \mu' A_s \mu = 4 \left[\sum_{s=1}^{m} \frac{s(n-m)^2}{n(n-s)} + \sum_{s=m+1}^{n-1} \frac{(n-s)m}{ns} \right] (\theta_1 - \theta_0)^2, \tag{4.28}$$

and its second term

$$\sum_{s=1}^{n-1} \operatorname{tr} A_s^2 = 2(n-1). \tag{4.29}$$

Combining the results of (4. 26), (4. 27), (4. 28), and (4. 29), (4. 11) can be written as

$$2a^2 b = 2(n-1) + 4 \left[\sum_{s=1}^{m} \frac{s(n-m)^2}{n(n-s)} + \sum_{s=m+1}^{n-1} \frac{(n-s)m}{ns} \right] (\theta_1 - \theta_0)^2$$

$$+ 4 \sum_{s=1}^{n-2} \sum_{t=s+1}^{n-1} \frac{(n-t)s}{t(n-s)} + 8 \sum_{s=1}^{m} \sum_{t=s+1}^{n-1} \frac{s(n-m)^2}{n(n-s)} (\theta_1 - \theta_0)^2$$

$$+ 8 \sum_{s=m+1}^{n-2} \sum_{t=s+1}^{n-1} \frac{m^2(n-t)}{nt} (\theta_1 - \theta_0)^2. \tag{4.30}$$

Now we proceed to simplify (4.7). The first term can be shown to be

$$\sum_{s=1}^{n-1} \operatorname{tr} A_s = n-1 \tag{4.31}$$

and its second term

$$\sum_{s=1}^{n-1} \mu' A_s \mu = \left[\sum_{s=1}^{m} \frac{(n-m)^2 s}{n(n-s)} + \sum_{s=m+1}^{n-1} \frac{(n-s)m^2}{ns} \right] (\theta_1 - \theta_0)^2 \tag{4.32}$$

Using the results of (4.31) and (4.32), Equation (4.7) simplifies to

$$ab = (n-1) + \left[\sum_{s=1}^{m} \frac{(n-m)^2 s}{n(n-s)} + \sum_{s=m+1}^{n-1} \frac{(n-s)m^2}{ns} \right] (\theta_1 - \theta_0)^2. \tag{4.33}$$

Let a_0 and b_0 be the solution to (4.30) and (4.33) under H_0. Therefore reject H_0 iff

$$T \geq a_0 \chi_\alpha^2(b_0)$$

where $\chi_\alpha^2(b_0)$ is the upper $100\alpha\%$ point of the $\chi^2(b_0)$ distribution.

The approximate power function for θ_0, θ_1, and m is

$$\beta_m(\theta_0, \theta_1, m) = \Pr[a\chi^2(b) \geq a_0 \chi_\alpha^2(b_0)] \tag{4.34}$$

where a and b are functions of θ_0, θ_1, m and n and are solved from (4.30) and (4.33). It should be noted that the accuracy of this moment approximation to the true power function was not investigated. The approximate power was tabulated for n = 12 for various values of θ_0, θ_1 and m as can be seen in Table V at the end of the chapter.

A MLRT When Initial Mean is Unknown but Variance is Known for $H_1 : \theta_1 > \theta_0$

From (2.29) under H_0 we have

$$T_s = \left[\sqrt{\frac{s(n-s)}{n}} \; (\bar{x}_{s+1,n} - \bar{x}_{1,s})\right] \text{ is distributed } N(0,1) \text{ for } s=1,2,\ldots,n-1. \tag{4.35}$$

Now consider the sum of T_s over all values of s, namely

$$\begin{aligned}
T_s &= \sum_{s=1}^{n-1}\left[\sqrt{\frac{s(n-s)}{n}} \; (\bar{x}_{s+1,n} - \bar{x}_{1,s})\right] \\
&= \sum_{s=1}^{n-1}\left[\left\{\sqrt{\frac{s}{n(n-s)}}\right\} \sum_{i=s+1}^{n} x_i - \left\{\sqrt{\frac{n-s}{ns}}\right\} \sum_{i=1}^{s} x_i\right] \\
&= \sum_{i=2}^{n}\left[\sum_{j=1}^{i-1} \sqrt{\frac{j}{n(n-j)}}\right] x_i - \sum_{i=1}^{n-1}\left[\sum_{j=1}^{n-1} \sqrt{\frac{n-j}{nj}}\right] x_i \\
&= \sum_{i=2}^{n-1}\left[\sum_{j=1}^{i-1} \sqrt{\frac{j}{n(n-j)}} - \sum_{j=1}^{n-1} \sqrt{\frac{n-j}{nj}}\right] x_i \\
&\quad + \left[\sum_{j=1}^{n-1} \sqrt{\frac{j}{n(n-j)}}\right] x_n - \left[\sum_{j=1}^{n-1} \sqrt{\frac{n-j}{nj}}\right] x_1 .
\end{aligned} \tag{4.36}$$

Without loss of generality let $\theta_0 = 0$, then

$$E(T) = \sum_{i=m+1}^{n-1}\left[\sum_{j=1}^{i-1} \sqrt{\frac{j}{n(n-j)}} - \sum_{j=1}^{n-1} \sqrt{\frac{n-j}{nj}}\right] \theta_1 + \left[\sum_{j=1}^{n-1} \sqrt{\frac{j}{n(n-j)}}\right] \theta_1, \tag{4.37}$$

denoted as μ and

$$\text{Var}\,(T) = \sum_{i=2}^{n-1}\left[\sum_{j=1}^{i-1} \sqrt{\frac{j}{n(n-j)}} - \sum_{j=1}^{n-1} \sqrt{\frac{n-j}{nj}}\right]^2 + \left[\sum_{j=1}^{n-1} \sqrt{\frac{j}{n(n-j)}}\right]^2 + \left[\sum_{j=1}^{n-1} \sqrt{\frac{n-j}{nj}}\right]^2, \tag{4.38}$$

denoted as v.

Since the x_is are independent and each have variance 1, T is distributed $N(\mu, v)$.

Therefore the power function of the test for θ_0, θ_1 and m is

$$\beta_m(\theta_0, \theta_1, m) = Pr[z \geq z_\alpha - \frac{\mu}{\sqrt{v}}] \tag{4.39}$$

where μ and v are defined in Equations (4.37) and (4.38), and z_α is the upper $100\alpha\%$ of $N(0, 1)$.

Refer to Table VII for some power calculations.

TABLE V

THE POWERS OF THE TWO-TAILED MODIFIED LRT WHEN THE INITIAL MEAN ($\theta_0 = 0$) AND VARIANCE ($\sigma^2 = 1$) ARE KNOWN; AND WHEN THE INITIAL MEAN IS UNKNOWN BUT VARIANCE ($\sigma^2 = 1$) IS KNOWN FOR n = 12 AND α = .05

$\theta_1 - \theta_0$	m	MLRT When Initial Mean is Known*	MLRT When Initial Mean Is Unknown**
	1	.1453	.0572
	3	.1420	.0765
.3	5	.1274	.0866
	7	.1045	.0838
	9	.0769	.0698
	11	.0545	.0534
	1	.3897	.0764
	3	.3556	.1354
.6	5	.3029	.1651
	7	.2317	.1617
	9	.1458	.1226
	11	.0658	.0666
	1	.7135	.1033
	3	.6471	.2074
.9	5	.5545	.2652
	7	.4219	.2661
	9	.2518	.2001
	11	.0848	.0865
	1	.9299	.1338
	3	.8810	.2929
1.2	5	.7980	.3886
	7	.6445	.3969
	9	.3940	.3004
	11	.1119	.1137

* Obtained from Equation (4.18)
**Obtained from Equation (4.34)

TABLE VI

THE POWERS OF THE ONE-TAILED ($\theta_1 > \theta_0$) MLRT AND THE BAYES TEST WHEN THE INITIAL MEAN ($\theta_0 = 0$) AND VARIANCE ($\sigma^2 = 1$) ARE KNOWN FOR n = 12 AND α = .05

θ_1	m	Modified LRT*	Bayes**
.3	1	.2087	.2222
	3	.2002	.2105
	5	.1804	.1846
	7	.1502	.1480
	9	.1120	.1066
	11	.0704	.0670
.6	1	.5091	.5459
	3	.4854	.5141
	5	.4276	.4399
	7	.3348	.3283
	9	.2156	.1991
	11	.0967	.0882
.9	1	.0842	.8403
	3	.7786	.8094
	5	.7084	.7243
	7	.5725	.5618
	9	.3600	.3283
	11	.1295	.1141
1.2	1	.9546	.9697
	3	.9420	.9569
	5	.8997	.9103
	7	.7858	.7750
	9	.5281	.4822
	11	.1694	.1450

* Obtained from Equation (4.22)
**Obtained from Equation (1.2)

TABLE VII

THE POWERS OF THE ONE-TAILED ($\theta_1 > \theta_0$) MLRT AND THE BAYES TEST WHEN THE INITIAL MEAN IS UNKNOWN AND THE VARIANCE ($\sigma^2 = 1$) IS KNOWN FOR n = 12 AND α = .05

$\theta_1 - \theta_0$	m	Modified LRT*	Bayes**
.3	1	.0730	.0659
	3	.1011	.0957
	5	.1106	.1139
	7	.1049	.1139
	9	.0878	.0957
	11	.0636	.0659
.6	1	.1035	.0855
	3	.1826	.1666
	5	.2114	.2216
	7	.1940	.2216
	9	.1436	.1666
	11	.0798	.0855
.9	1	.1422	.1092
	3	.2961	.2647
	5	.3520	.3715
	7	.3182	.3715
	9	.2195	.2647
	11	.0991	.1092
1.2	1	.1898	.1372
	3	.4341	.3858
	5	.5167	.5442
	7	.4674	.5442
	9	.3143	.3858
	11	.1216	.1372

*Obtained from Equation (4.39).

**Obtained from Equation (1.4).

Some Extensions

Estimating the Shift Point

As seen from previous considerations, the LRT first estimates the shift point then uses the relevant observations as a test statistic. Thus the properties of the estimator of the shift point should be investigated.

One-Tailed Test When the Variance is Unknown

The one-tailed test when the variance is unknown can be derived to be a doubly noncentral t-distribution and the result of Krishnan [8] can be used to evaluate the power of this test.

A Multivariate Extension

Consider a set of N independent random vectors $X_1, X_2, \ldots, X_N$. Suppose we are interested in detecting a change in the mean vector of these observations, assuming a common known covariance matrix V.

Thus we consider the following hypotheses:

$$H_0 : X_i \sim \text{NID}\,(\mu_0, V) \quad i=1,2,\ldots,N$$

against

$$\begin{aligned} H_1 : X_i &\sim \text{NID}(\mu_0, V) \quad i=1,2,\ldots,M \\ X_i &\sim \text{NID}(\mu_1, V) \quad i=M+1,\ldots,N \end{aligned} \qquad (1 \leq M \leq N-1)$$

where M and μ are unknown, the latter is a P-component vector.

Consider the multivariate extension of Case 1. We can prove that H_0 is rejected if

$$(N-\hat{M})(\overline{X}_{\hat{M}} - \mu_0)' V^{-1}(\overline{X}_{\hat{M}} - \mu_0) \geq \chi_{\alpha}^2(P) \tag{4.32}$$

where

$$\overline{X}_{\hat{M}} = \sum_{i=\hat{M}+1}^{N} X_i/N-\hat{M})$$

is the sample mean vector based on the last N-$\hat{M}$ vector observations, $\hat{M}$ is the MLE of M, and $\chi_{\alpha}^2(P)$ is the upper $100\alpha\%$ point of the chi-square distribution with P degrees of freedom.

The distribution of $\hat{M}$ is given as, $\hat{M}$=s if and only if

$$\max_M[(N-M)(\overline{x}_M - \mu_0)'V^{-1}(\overline{x}_M - \mu_0)] = (N-s)(\overline{x}_s - \mu_0)'V^{-1}(\overline{x}_s - \mu_0). \tag{4.33}$$

The power function for shift point M and mean vector μ_1 is

$$\beta_M(\mu_1) = \sum_{s=1}^{M} \Pr[\chi'^2(P, \lambda_1) \geq \chi_{\alpha}^2(P)] \cdot \Pr[\hat{M}=s]$$

$$+ \sum_{s=M+1}^{N-1} \Pr[\chi'^2(P, \lambda_1^*) \geq \chi_{\alpha}^2(P)] \cdot \Pr[\hat{M}=s]$$

where

$$\lambda_1 = (N-M)^2(\mu_1 - \mu_0)'V^{-1}(\mu_1 - \mu_0)/2(N-s), \tag{4.34}$$

$$\lambda_1^* = (N-s)(\mu_1 - \mu_0)'V^{-1}(\mu_1 - \mu_0)/2$$

As seen from Case 1, extension to the multivariate case is straight forward. The power function in the multivariate situation is a mixture of non-central chi-square distributions with P degrees of freedom and the above formulas reduce to the univariate case by letting P=1. The other multivariate extensions of Cases 2, 3, and 4 will not be discussed in this investigation.

CHAPTER V

SUMMARY AND CONCLUSIONS

This research developed test procedures for detecting a shift in the mean of a normal distribution when the time the shift occurs is unknown. It also developed a simple estimation procedure for estimating the location or time of the shift and the magnitude of the shift. Since the likelihood-ratio test first estimates the time where the shift takes place, it uses only the relevant observations. Thus the estimate of the magnitude of shift is $\bar{x}_{\hat{m}+1, n} - \theta_0$ if θ_0 is known, where $\hat{m}$ is the estimate of the location or time of shift. This is the index m which is found by evaluating

$$\max_m \left[(n-m)(\bar{x}_{m+1, n} - \theta_0)^2\right].$$

When the initial mean, θ_0, of the population is not known $\bar{x}_{1, \hat{m}}$ becomes its estimate, where $\hat{m}$ is given by maximizing the likelihood function which can also be found by evaluating

$$\max_m \left[m\bar{x}^2_{1, m} + (n-m)\bar{x}^2_{m+1, n}\right].$$

When the variance of a normal distribution is known, in Chapter II, the one-tailed test has a test statistic which is normally distributed either when the initial mean is known or when it is unknown. However, the two-tailed test statistic is distributed as a noncentral chi-square either when the initial mean is known or when it is not known. The Bayes test procedure has better power compared to the

LRT for Case 1 if the shift does not occur late in the sequence as shown in Table I. From Table II, Case 2, the LRT has better power than the Bayes test procedure when the magnitude of shift is quite large and when the shift occurs late or early. The Bayes test procedure has more power than the LRT when the magnitude of shift is small and when the shift does not occur early.

In Cases 3 and 4 where the variance is unknown, it is seen from Table IV that for each value of θ_1, as m increases the power decreases strictly when the initial mean is known. But when the initial mean is unknown, the power increases then decreases symmetrically about m equal to n/2. The power function of the test of Case 3 is a finite mixture of ratios of noncentral chi-square distributions, where the mixing distribution is that of $\hat{m}$, the MLE of m. The power function of Case 4 has a similar distribution.

In Chapter IV, a modified LRT was constructed which does not depend on the distribution of $\hat{m}$. However, this procedure as with the Bayes and non-parametric does not give an estimate of the shift point. The test statistic in this case was approximated by a scaled chi-square distribution. The one-tailed test of the modified LRT has a test statistic which is normally distributed. The one-tailed test compared very well with the Bayes test since it is more powerful than the Bayes test at half of the shift points. When the initial mean is known, the Bayes test is better if the shift does not occur later than n/2. The situation is reversed for Case 2; the MLRT is better if the shift occurs earlier than n/2.

A SELECTED BIBLIOGRAPHY

(1) Page, E. S. "A test for a change in a parameter occuring at an unknown point." Biometrika, Vol. 42, 1955, pp. 523-526.

(2) Chernoff, H. and Zacks, S. "Estimating the current mean of a normal distribution which is subjected to changes in time." Annals of Mathematical Statistics, Vol. 35, 1964, pp. 999-1018.

(3) Kander, Z. and Zacks, S. "Test procedures for possible changes in parameters of statistical distributions occuring at unknown time points." Annals of Mathematical Statistics, Vol. 37, 1966, pp. 1196-1210.

(4) Mustafi, C. K. "Inference problems about parameters which are subjected to changes over time." Annals of Mathematical Statistics, Vol. 39, 1968, pp. 840-854.

(5) Graybill, F. A. An Introduction to Linear Statistical Models, Vol. I, McGraw-Hill Book Co., Inc., New York, 1961.

(6) Lukacs, E. and Laha, R. G. "Applications of characteristic functions." No. 14 Griffin's Statistical Monographs and Courses, Hafner Publishing Co., New York 1964.

(7) Koch, G. G. "A general approach to the estimation of variance components." Technometric, Vol. 9, 1967, pp. 93-118.

(8) Krishnan, M. "Series representation of the doubly non-central t-distribution." Journal of the American Statistical Association, Vol. 63, 1968, pp. 1004-1011.

APPENDIX

THE EXACT DISTRIBUTION OF $\hat{m}$ FOR SAMPLE SIZE n EQUALS THREE

We consider Case 1, when the initial mean and variance are known. Let $\theta_0 = 0$ and $\sigma^2 = 1$. The MLE of m is $\hat{m} = s$ iff

$$\max_m [(n-m)\bar{x}^2_{m+1,n}] = (n-s)\,\bar{x}^2_{s+1,n}.$$

Under H_0 we have if $\hat{m} = 1$ then

$$\max_m [(n-m)\,\bar{x}^2_{m+1,n}] = 2\bar{x}^2_{2,3}$$

and $2\,\bar{x}^2_{2,3} \geq x_3^2$. This inequality is equivalent to the four sets of inequalities as follow:

$$(1) \quad \sqrt{2}\ \frac{(x_2+x_3)}{2} \geq x_3,\ x_2+x_3 \geq 0,\ x_3 \geq 0,$$

or

$$(2) \quad \sqrt{2}\ \frac{(x_2+x_3)}{2} + x_3 \geq 0,\ x_2+x_3 \geq 0,\ x_3 < 0$$

or

$$(3) \quad -\sqrt{2}\ \frac{(x_2+x_3)}{2} + x_3 \geq 0,\ x_2+x_3 < 0,\ x_3 \geq 0$$

or

$$(4) \quad -\sqrt{2}\ \frac{(x_2+x_3)}{2} - x_3 \geq 0,\ x_2+x_3 < 0,\ x_3 \geq 0$$

From set 1 we get

$$\Pr[x_2 + (1-\sqrt{2})\,x_3 \geq 0,\ x_2+x_3 \geq 0,\ x_3 \geq 0] = \Pr(A_1)$$

where A_1 in Figure 1 is the area between $x_2 > 0$ and the line $x_2 + (1-\sqrt{2})x_3 = 0$.

From set 2 we get

$$\Pr[x_2 + (1+\sqrt{2})\,x_3 \geq 0,\ x_2 + x_3 \geq 0,\ x_3 < 0] = \Pr(A_2)$$

where A_2 is the area between $x_2 < 0$ and the line $x_2 + (1-\sqrt{2})\,x_3 = 0$.

From set 3 we get

$$\Pr[x_2 + (\ 1 - \sqrt{2})\,x_3 \leq 0,\ x_2 + x_3 \leq 0,\ x_3 < 0] = \Pr(A_3)$$

where A_3 is the area between $x_2 > 0$ and the line $x_2 + (1-\sqrt{2})\,x_3 = 0$.

From set 4 we get

$$\Pr[x_2 + (1+\sqrt{2})\,x_3 \leq 0,\ x_2 + x_4 < 0,\ x_3 > 0] = \Pr(A_4).$$

Since tan ϕ_1 equals tan ϕ_3 and tan ϕ_2 equals tan ϕ_4, as shown in Figure 1, then $\Pr(A_1)$ equals $\Pr(A_3)$ and $\Pr(A_2)$ equals $\Pr(A_4)$. We also know that tan ϕ_1 equals -cot ϕ_2, and tan ϕ_3 equals -cot ϕ_4. Thus $(\phi_1 + \phi_2)$ equals 90^o and $(\phi_3 + \phi_4)$ equals 90^o.

Therefore,

$$\Pr(A_1) + \Pr(A_4) = 90/360 = 1/4$$

and

$$\Pr(A_3) + \Pr(A_2) = 1/4\ .$$

Since the four sets are mutually exclusive therefore

$$\sum_{i=1}^{4} \Pr(A_i) = 1/2,$$

$$\Pr(\hat{m}=1) = 1/2,$$

and

$$\Pr(\hat{m}=2) = 1/2.$$

This result shows that $\hat{m}$ under H_0 has a uniform distribution for n = 3. The extension of the proof for $n \geq 4$ is very tedious and will not be attempted in this research.

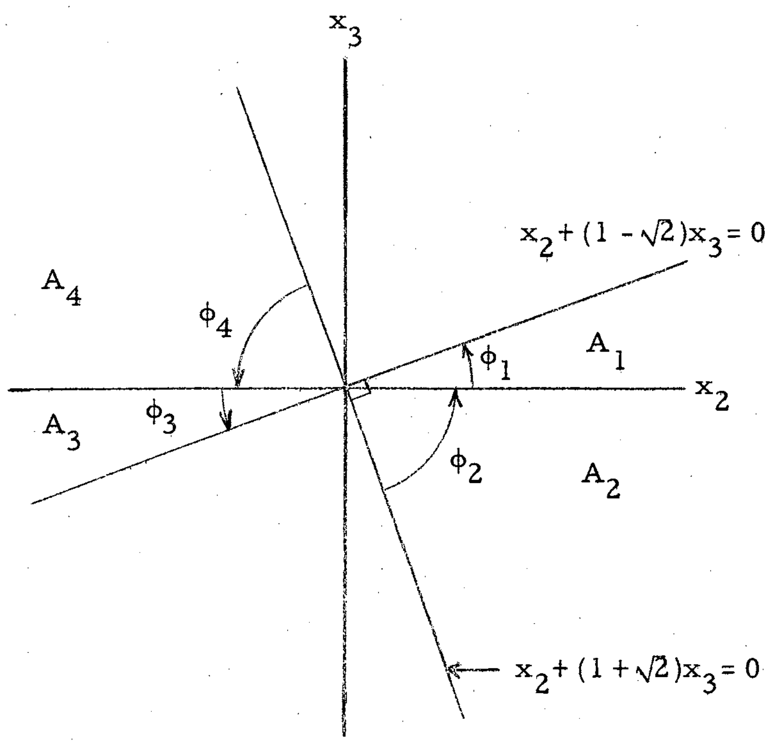

Figure 1.

VITA 3

Henry Feliciano Magalit

Candidate for the Degree of

Doctor of Philosophy

Thesis: DETECTING A CHANGE IN THE MEAN OF A NORMAL DISTRIBUTION AT AN UNKNOWN TIME POINT

Major Field: Statistics

Biographical:

Personal Data: Born in Numancia, Aklan, Philippines, February 27, 1933, the son of Iluminada Feliciano and Valeriano R. Magalit.

Education: Finished high school in 1950 at Central Philippine University, Iloilo City, Philippines; received the Bachelor of Science in Mechanical Engineering from the University of the Philippines in October, 1954; took graduate courses at the Statistical Center, University of the Philippines from June 1960 to October 1962. Received the Master of Science degree from Oklahoma State University, Stillwater, Oklahoma, in July 1966, with a major in mathematics and statistics; completed the requirements for the Doctor of Philosophy degree in May, 1970.

Professional Experience: Student assistant, College of Engineering, University of the Philippines, March to October 1954; passed the Philippine Civil Service examination for Junior Mechanical Engineers in February 1955; appointed assistant instructor in mathematics June 1955 and promoted to instructor in mathematics October 1957 at the College of Agriculture, University of the Philippines; on "special detail" at Oklahoma State University from September 1964 to January 1970; employed as a test engineer by Stewart Warner Corporation, Chicago, Illinois from June to September 1965; worked as graduate assistant at Oklahoma State University from September 1965 to January 1970.