A STUDY ON ERROR SIZES AND REQUIRED SAMPLE SIZE IN SEQUENTIAL PROBABILITY RATIO TEST

A MONTE CARLO APPROACH

By

Gemechis Dilba

A THESIS SUBMITTED IN PARTIAL FULFILMENT
FOR THE DEGREE OF MASTER OF SCIENCE IN STATISTICS
OF THE ADDIS ABABA UNIVERSITY

December, 1993
Addis Ababa

ACKNOWLEDGMENTS

I wish to express my deep gratitude to Dr. Abebe Tessera, my advisor, without whom this research would have been not possible. He has given me invaluable comments and suggestions at each stage of the research.

I am very grateful to my family and the family of Ato Hambissa Deressa for their constant encouragement throughout my study.

Finally, I am greatly indebted to the German Academic Exchange Service (DAAD) for granting me an in-country scholarship. My thanks is also due to Prof. Leo Knüsel for sponsoring me to conduct part of the research at a German University (Munich). He provided me with every necessary materials.

ABSTRACT

In this paper computer simulation is employed to investigate the nature of the error sizes and the distribution of sample size required in a sequential probability ratio test (SPRT). In the test we can distinguish between two kinds of error probabilities. The first kind are the specified error probabilities which are usually denoted by α and β, and the second kind are the true error probabilities which may be denoted by α_t and β_t. Wald has shown that the relation $\alpha_t + \beta_t \leq \alpha + \beta$ holds true. The objective of this project is to investigate the relation between α_t and α, and between β_t and β. As sample size required in SPRT is a random variable, its distribution is also studied. Two probability distributions: Bernoulli and normal (known variance) are selected for the study.

The study shows that in the case of normal distribution, when the parameters under H0 and H1 are slightly far apart and $\alpha = \beta$, the estimates of true error probabilities are less than their corresponding specified error probabilities and that they are close to each other. The estimates of α_t and β_t also decrease as $d = \theta_1 - \theta_0$ increases. This shows that the actual risks are by far less than the specified value for large d. When α and β are not equal, there are times when the estimate of α_t or β_t exceeds its corresponding specified error size as observed. But, still if the parameters under H0 and H1 are far apart,the estimates indicate that $\alpha_t \leq \alpha$ and $\beta_t \leq \beta$.

For Bernoulli distribution, the results are not very far from those of normal except that in some cases the estimates of α_t or β_t are found to be greater than α or β which led to disobeying the inequality $\alpha_t + \beta_t \leq \alpha + \beta$. This may be attributed to sample fluctuation.

And finally, sample size distribution is observed to depend mainly on $d = \theta_1 - \theta_0$ $(\theta_0 < \theta_1)$. The mean and variance of the required sample size increase rapidly as d decreases. Further, the distributions are all positively skewed. It is also observed that in rare occasions there is a chance that the sample size in SPRT exceeds the size one needs in non-sequential test.

CONTENTS

CHAPTER I

INTRODUCTION

The methods of sequential analysis were developed simultaneously in the United States and Great Britain in response to demands for more efficient sampling inspection procedures during World War II. Darling(1976) and Wallis(1980) have studied the development of the methods and they note the great contribution of Abraham Wald to what we call today " Sequential probability ratio test (SPRT)". Actually, the double sampling inspection which was proposed by Dodge and Romig(1929) is the most elementary form of sequential analysis.

Wald(1947) describes the method as statistical inference whose characteristic feature is that the number of observation required by the procedure is not determined in advance of the experiment. Decision at any time depends on the observations drawn up to that time, and if there is no sufficient information to reach a decision additional observation(s) are taken. Wetherill(1975) explains the usefulness of such a method by saying that "... the whole of life is sequential, for our future actions are conditioned to some extent by our past experience." Of course, the method may not yield best results always. Dixon and Massey(1983) enumerate circumstances under which it is appropriate (p 423-24) to apply the method. It is proper to apply the method when each individual may be tested separately, each test is expensive (increasing the importance of minimum sample size), the response time is short, test conditions do not permit testing of more than one individual at a time, or test cases are available only at wide intervals in time (as for example in the cases of rare diseases). And sequential procedure may not be appropriate when test materials are inexpensive, long response times extend the total experimental time greatly, a decision must be reached in a specified short time, or when a loss in uniformity of tests occurs if the tests are not all done simultaneously,as, for example from seasonal variations in response or decay in drug strength. Very often the method is applied to sampling inspection and sequential medical trials.

In the last five decades sequential procedure has gone through many theoretical developments. Suggestions and modifications have been given after studying some of its properties. Most studies are made on the basis of simulation or lengthy numerical computations.

Baker(1950) conducted an experiment on testing the hypothesis H0: μ=0 versus H1: μ=1 for normal distribution with known variance, σ^2=1. The total number of tests performed was 2003. He considered four different combinations of α and β (α=β), and observed that the estimate of α_t, the estimate of the true probability of Type I error is about 0.7 times the specified value. Corneliussen and Ladd(1970) conducted a similar experiment. They concluded that the average sample number is not a good measure for assessing the effort required to conduct a sequential test. Schneiderman and Armitage(1962) also carried out some empirical sampling trials to compare three sequential plans: open(Wald, or Sobel-Wald) plan, the restricted plan (Armitage), and the wedge plan (Schneiderman-Armitage) with respect to their average sample number (ASN) and operating characteristics. They considered the hypothesis μ = 0 against μ = 1 of a normal distribution with σ = 1 (α = β = 0.05). The open plan shows an advantage in ASN at H0, and a disadvantage at certain values of μ between H0 and H1. However, the risk on H1 is closer to the specified value (β) for the restricted and wedge plans than for the open plan. It was also observed that for all three schemes the true errors of the first and second kinds tended to be smaller than specified.

Other researchers studied the properties of sequential method analytically. Anderson(1960) modified the method after observing the relatively large expected sample size for values of the parameter between the two specified values in testing a simple against simple alternative hypothesis. The other notable result is the asymptotic expansion derived by Lotov(1988) for the error probabilities α(A,B) and β(A,B) corresponding to the demarkation points A and B.

In this paper, computer simulation will be used to assess the pattern of error probabilities and sample size distribution associated with the open (Wald) test or SPRT. It is essentially what Baker had started earlier, but on broader basis both in scope and the number of tests performed. Here two probability distributions, five different cases of α's and β's, and a number of different parameter values will be tried.

The paper is organized into four parts. Chapter 2 deals with describing the method (SPRT) and deriving some of the most important results. Chapter 3 deals with the objectives of the project, the methods and materials used. Chapter 4 is devoted to analysis of the data generated by the simulation program and in chapter 5 the results obtained are discussed and conclusions are given.

CHAPTER II
LITERATURE REVIEW

Before introducing the general method of sequential probability ratio test, let us highlight its point of departure from the most powerful test for testing a simple hypothesis against a simple alternative. In the latter method we have two regions, the rejection and acceptance regions to make a decision. That is, after computing the value of the test statistic from a sample of fixed size, decision is automatic depending on where the value falls. However, in SPRT, in addition to the rejection and acceptance regions, there is a third region called continue sampling region. It is a gap between the rejection and acceptance regions. Whenever the value of the test statistic falls in this region, it implies that the sample is not sufficiently suggestive to either accept or reject the hypothesis. This requires one to draw additional observation(s) to make the evidence stronger. And the process continues until a decision is reached. As a result, in SPRT sample size can not be predetermined.

2.1 The Sequential Probability Ratio Test.

Suppose we want to test the simple versus simple alternative about parameter θ of a certain probability distribution f(θ,x),

$$H0 : \theta = \theta_0$$

$$H1 : \theta = \theta_1$$

Wald(1947) has developed a method, sequential probability ratio test (SPRT), by which we can discriminate between the two hypotheses.

Let $X_1, X_2, \ldots$ be independent observations from f(θ,x). The likelihood ratio $\lambda_n(X)$ of the first n observations is defined as

$$\lambda_n(X) = \frac{\prod_{i=1}^{n} f(\Theta_0, x_i)}{\prod_{i=1}^{n} f(\Theta_1, x_i)} \qquad n = 1,2,\ldots$$

The procedure is as follows

i) Fix the probabilities of type-I and type-II errors at α and β respectively

ii) Obtain demarkation points A and B (A < B) depending on α and β of step (i)

The decision rule is

reject H0 if $\lambda_n(x) \le A$

accept H0 if $\lambda_n(x) \ge B$

continue sampling otherwise

The sample size required to reach at a decision is a random variable and can be written as

$$N = \min(n:\ \lambda_n(x) \notin (A,B))$$

2.1.1 Demarkation points and the True Error Probabilities

It is almost impossible to solve exactly for A and B for given α and β. Wald(1947) obtained a simple approximation bounds for A and B in terms of the error probabilities.

$$\frac{\alpha}{(1-\beta)} \le A < B \le \frac{(1-\alpha)}{\beta} \qquad \ldots\ldots\ 2.1$$

The derivation is as follows :

Let E_k denote the set of X_i's i.e $(X_1,\ldots,X_k)$ that calls for termination of the procedure at stage k with a decision to reject H0.

$$\alpha = P\{\text{ rejecting H0 / H0 true }\} = \sum_k P\{ E_k/H0 \}$$

$$1-\beta = P\{\text{ rejecting H0 / H1 true }\} = \sum_k P\{ E_k/H1 \}$$

H0 will be rejected if

$$\lambda_n(x) = \frac{f_X(\Theta_0;x)}{f_X(\Theta_1;x)} \le A$$

$$\Rightarrow \quad f_X(\theta_0;X) \le A\, f_X(\theta_1;X) \qquad \text{.......} \; 2.2$$

which is true for all E_k. $f_X(\theta_i;X)$ is the joint distribution of

$X_1, X_2, \ldots, X_n$ under H_i. i = 0,1

Now, using (2.2) we can write

$$P(E_k/H_0) = \int_{E_k} fX(\theta_0;X)\, d_x \le A \int_{E_k} fX(\theta_1;X)\, d_x = A\, P(E_k/H_1)$$

i.e $\quad P\{ E_k / H0 \} \le A\ P\{ E_k / H1 \}$ 2.3

Taking summation over all possible k's we get

$$\alpha = \sum_k P(E_k/H0) \le A \sum_k P(E_k/H1) = A(1-\beta)$$

$$\Rightarrow \frac{\alpha}{(1-\beta)} \le A \qquad \text{.......} \; 2.4$$

Likewise, let F_k denote the set of X_i's i.e $(X_1, \ldots, X_k)$ that calls for termination of the procedure at stage k with a decision to accept H0.

$$1-\alpha = P\{ \text{accepting H0 / H0 true} \} = \sum_k P\{ F_k / H0 \}$$

$$\beta = P\{ \text{accepting H0 / H1 true} \} = \sum_k P\{ F_k / H1 \}$$

We accept H0 if

$$\lambda_n(X) \ge B \Rightarrow \text{if } f_X(\theta_0;X) \ge B f_X(\theta_1;X)$$

This is true for all F_k. Therefore, we can write

$$P(F_k/H0) = \int_{F_k} f_X(\theta_0;X)\, d_x \ge B \int_{F_k} f_X(\theta_1;X)\, d_x = B\, P(F_k/H1)$$

i.e $\quad P\{ F_k / H0 \} \ge B\ P\{ F_k / H1 \}$ 2.5

Taking summation over all k's,

$$1-\alpha = \sum_k P(F_k/H0) \le B\sum_k P(F_k/H1) = B\beta$$

$$\Rightarrow \frac{(1-\alpha)}{\beta} \ge B \qquad \ldots\ldots\ 2.6$$

From Equations (2.4) and (2.6) the desired inequality follows.

Wald mentions that the approximation of A by $\alpha/(1-\beta)$ and B by $(1-\alpha)/\beta$ suffices for most practical purposes. The approximation widens the continue sampling region which in turn has an effect of slightly increasing the number of observations necessary to reach at a decision.

After the above approximations, the error probabilities corresponding to the test will no more be α and β. Let us denote the true probabilities of Type I and Type II errors by α_t and β_t respectively. It is true that

$$\frac{\alpha_t}{1-\beta_t} \le A = \frac{\alpha}{1-\beta} \qquad \ldots\ldots\ 2.7$$

$$\frac{1-\alpha_t}{\beta_t} \ge B = \frac{1-\alpha}{\beta} \qquad \ldots\ldots\ 2.8$$

Rearranging Equations (2.7) and (2.8) we have

$$\alpha_t + \beta_t \le \alpha + \beta \qquad \ldots\ldots\ 2.9$$

This tells us that the sum of the probabilities of Type I and Type II errors for the choice of $A = \alpha/(1-\beta)$ and $B = (1-\alpha)/\beta$ can not exceed the sum of the specified probabilities of Type I and Type II errors.

Regarding the bounds on α_t and β_t, it is clear that

$$\alpha_t \le \frac{\alpha_t}{1-\beta_t} \qquad \ldots\ldots\ 2.10$$

$$\beta_t \le \frac{\beta_t}{1-\alpha_t} \Rightarrow \frac{1}{\beta_t} \ge \frac{1-\alpha_t}{\beta_t} \qquad \ldots\ldots\ 2.11$$

Combining (2.7) with (2.10) and (2.8) with (2.11) it follows

$$\alpha_t \le \frac{\alpha}{1-\beta} = A \qquad \text{....... 2.12}$$

$$\beta_t \le \frac{\beta}{1-\alpha} = \frac{1}{B} \qquad \text{....... 2.13}$$

The inequalities (2.12) and (2.13) can be written as

$$\alpha_t \le \frac{\alpha}{1-\beta} \doteq \alpha(1+\beta) \quad \text{and} \quad \beta_t \le \frac{\beta}{1-\alpha} \doteq \beta(1+\alpha)$$

for small α and β.

2.1.2 The Average Sample Number (ASN).

From the description of SPRT, we see that sample size is a random variable. Therefore, we can talk about the expected number of observations required to reach at a decision. Wald has derived an approximate expression:

$$E(N) \approx \frac{\pi(\theta)\ln(A) + (1-\pi(\theta))\ln(B)}{E(Z)}$$

Where

$$Z = \ln\left(\frac{f(\theta_0, x)}{f(\theta_1, x)}\right)$$

and $\pi(\theta)$ is the power function.

Specifically, the approximate expected sample size under H0 and H1 are respectively

$$E(N/H0) \approx \frac{\alpha\ln(A) + (1-\alpha)\ln(B)}{E(Z/H0)} \qquad \text{....... 2.14}$$

$$E(N/H1) \approx \frac{(1-\beta)\ln(A) + \beta\ln(B)}{E(Z/H1)} \qquad \text{....... 2.15}$$

2.2 Relative Efficiency of SPRT.

We can compare SPRT with the most powerful test of Neyman-Pearson. Wald used the expected sample size as a basis to see the relative performance of SPRT. In the non-sequential case, sample size is fixed and it can be computed for given values α and β, say $n(\alpha,\beta)$. The relative efficiency of SPRT is defined as

$$R_{\theta}(\theta) = \frac{n(\alpha,\beta)}{E_{\theta}(N)} \qquad \text{....... 2.16}$$

Where $E_{\theta}(N)$ is the expected sample size in SPRT when the true state of nature is θ.

Here we compare the average sample size it requires one to reach a decision in SPRT with the size one needs in non-sequential test using the same specified error probabilities as applied in SPRT.

In next section, the above results will be applied to specific probability distributions which are selected for the study.

CHAPTER III

OBJECTIVE OF THE RESEARCH

3.1 Introduction.

In the preceding chapter some of the basic theories of sequential analysis developed by Wald were briefly presented. In particular, it was shown that by using the approximate demarkation points, the sum of the true error probabilities are less than or equal to the sum of the specified error sizes. But this does not tell us exactly the relation of α_t to α and β_t to β. The first objective of this research is to study these relationships.

The other most important result proved by Wald is that SPRT requires on the average fewer observations than a fixed-sample size test having the same error probabilities. This is actually the strong optimal property of SPRT. After computing efficiency for normal distribution, he also reached at the conclusion that the average decrease in the number of observations by using SPRT amounts to 50%. This is sometimes taken as that there would always be a saving in the number of observations and hence cost. However, even though the process terminates with probability 1, under some condition ($p(Z=0) < 1$), there is a chance that SPRT may require many more observations than a comparable non-sequential test. The optimal property of SPRT takes into account an average property. Here a study on the nature of the sample size distribution, which is the second main objective of the research, could be useful.

Thus, among others, the specific objectives of this research project are:

i) To make a comparative study of the true and specified error sizes

ii) To study the sample size distribution

3.2 Materials and Methods

Monte Carlo method is extensively used. The program tests the simple versus simple alternative hypotheses about the parameter of a specified distribution. For the study, the Bernoulli and normal distributions are selected. In the former distribution inference would be on p, probability of success.

While for normal distribution the interest is in test about the mean under the assumption of known variance.

The simulation program consists of repeatedly selecting observations from the distributions to perform tests sequentially. And generation of independent observations from the distributions is by the use of a random number generator.

3.2.1 Pseudo-Random Number.

There are two basic techniques by which one can generate independent observations from the interval (0,1). They are known as the congruential and feedback shift register method. Both methods consist of first generating integers over some known range and then transforming to the interval (0,1). Let us consider the former method which can be described as follows.

$$x_i = (a x_{i-1} + b)\ (Mod\ M) \quad i=1,2,...$$

where $0 \leq x_i \leq M - 1$. M, a and b are parameters describing the generator and x_0 is the seed. Now by the transformation $u_i = x_i/M$ random numbers in (0,1) will be generated. Undesirable statistical properties of this generator are in order.

i) It picks only the numbers 0,1/M,...,(M-1)/M from [0,1]. Further more, the sample distribution of the observations generated by this method was found to be not appreciably close to the theoretical distribution, uniform(0,1).

ii) Since the equation given to generate observations is a deterministic one, the u_i's are not truly random or independent. Subsequent observations are predictable given the seed and the parameters.

The Pascal language in which the simulation program is written has a built-in random number generator called RANDOM. It suffers from the above problems. Gebre-Egziabher(1992) wrote a program (Function) which alleviates the problems some how. The program makes use of the built-in random number generator. In the simulation program this function is used to generate observations from uniform(0,1). In the next two sections we shall see how this is in turn used to generate observations from the Bernoulli and normal distributions.

3.2.1.1 Independent Observations from Bernoulli Distribution

Let $U_1, U_2, \ldots$ be independent observations from Uniform(0,1).

Independent observations can be generated from the Bernoulli distribution with parameter p as follows.

If $U_i > p$ then $X_i = 0$

otherwise $X_i = 1$ $\quad\quad i = 1,2,\ldots$

In this way $X_1, X_2, \ldots$ constitute independent observations from

$$f(p;x) = p^x(1-p)^{1-x}, \quad x = 0,\ 1.$$

3.2.1.2 Independent Observations from Normal Distribution.

The Box-Müller method consists of generating independent observations X_1 and X_2 from Normal(μ,σ^2) as

$$X_1 = \mu + \sigma\,[-2\ln(U_1)]^{1/2}\cos(2\pi U_2)$$

$$X_2 = \mu + \sigma\,[-2\ln(U_1)]^{1/2}\sin(2\pi U_2)$$

where U_1 and U_2 are independent observations from Uniform(0,1).

The Polar Marsaglia method is a modification of this method, and it avoids the use of time-consuming sine and cosine functions. The procedure for Polar Marsaglia is as follows:

Compute

$$V_1 = 2U_1 - 1$$

$$V_2 = 2U_2 - 1$$

$$W = V_1^2 + V_2^2$$

If $W < 1$ then

$$X_1 = \mu + \sigma v_1\sqrt{\frac{-2\ln(W)}{W}} \quad \text{and} \quad X_2 = \mu + \sigma v_2\sqrt{\frac{-2\ln(W)}{W}}$$

constitute independent observations from $N(\mu,\sigma^2)$. If $w \geq 1$, we select another U_i's and repeat the procedure. This method is used in the program.

3.2.2 SPRT for Bernoulli Distribution.

Consider test of hypothesis about the parameter p, H0: $p = p_0$ against H1: $p = p_1$ ($p_0 < p_1$) of a bernoulli distribution.

$$f(p;x) = p^x(1-p)^{1-x} \quad x = 0,1$$

If $x_1, x_2, \ldots, x_n$ are independent observations from this distribution, the likelihood ratio is given by

$$\lambda_n(x) = \frac{p_0^{\sum_{i=1}^{n} x_i}(1-p_0)^{n-\sum_{i=1}^{n} x_i}}{p_1^{\sum_{i=1}^{n} x_i}(1-p_1)^{n-\sum_{i=1}^{n} x_i}}$$

Substituting this in the decision rule:

reject H0 if $\lambda_n(x) \leq A$ 3.1

accept H0 if $\lambda_n(x) \geq B$ 3.2

continue sampling otherwise ,

taking the natural logarithm of both sides and after some algebraic manipulations, we get

$$\textit{reject } H_0 \textit{ if } \sum_{i=1}^{n} x_i \geq a_2 + bn$$

$$accept\ H_0\ if\ \sum_{i=1}^{n} x_i \le a_1 + bn$$

Continue sampling otherwise n = 1,2,...

where

$$a_1 = \frac{lnB}{d}\ ,\ a_2 = \frac{lnA}{d},\ A = \frac{\alpha}{1-\beta},\ B = \frac{1-\alpha}{\beta}$$

$$d = \ln(\frac{p_0(1-p_1)}{p_1(1-p_0)})\ ,\ b = -\frac{\ln((1-p_0)/(1-p_1))}{d}$$

$\sum x_i = a_1 + bn$ and $\sum x_i = a_2 + bn$ are parallel lines in the ($n,\Sigma x_i$) plane. We reach a decision when $\sum x_i$ touches or crosses one of the two lines. The decision rule for the case $p_0 > p_1$ can also be similarly obtained.

Regarding the theoretical average sample size required, $Z = \ln(\frac{f(p_0;x)}{f(p_1;x)}) = x\ln(\frac{p_0(1-p_1)}{p_1(1-p_0)}) + \ln(\frac{1-p_0}{1-p_1})$. And the expectation of this under H0 and H1 are respectively

$$E(Z/H0) = p_0\ln(\frac{p_0(1-p_1)}{p_1(1-p_0)}) + \ln(\frac{1-p_0}{1-p_1}) \qquad \text{....... 3.3}$$

$$E(Z/H1) = p_1\ln(\frac{p_0(1-p_1)}{p_1(1-p_0)}) + \ln(\frac{1-p_0}{1-p_1}) \qquad \text{....... 3.4}$$

Replacing Equations (3.3) and (3.4) in (2.14) and (2.15) respectively, we get

$$E(N/H_0) \approx \frac{\alpha \ln(A) + (1-\alpha)\ln(B)}{p_0 \ln(\frac{p_0(1-p_1}{p_1(1-p_0)}) + \ln(\frac{1-p_0}{1-p_1})}$$

$$E(N/H_1) \approx \frac{(1-\beta)\ln(A) + \beta \ln(B)}{p_1 \ln(\frac{p_0(1-p_1}{p_1(1-p_0)}) + \ln(\frac{1-p_0}{1-p_1})}$$

The relative efficiency of SPRT for this distribution when p0 is close to p1 may be computed using central limit theorem. That is, we compute n(α,ß), the sample size required in non-sequential procedure corresponding to a given α and ß using a normal approximation.

$$n(\alpha, \beta) = [\frac{Z_\alpha \sqrt{p_0(1-p_0)} + Z_\beta \sqrt{p_1(1-p_1)}}{p_1 - p_0}]^2$$

3.2.3 SPRT for Normal Distribution

The probability density function of a normal distribution with mean μ and variance σ^2 is given by

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}, \quad -\infty < x < \infty$$

Suppose we want to test H0:$\mu=\mu_0$ versus H1:$\mu=\mu_1$ ($\mu_0 < \mu_1$) for known variance σ^2. The likelihood function is given by

$$\lambda_n(X) = \frac{\prod_{i=1}^{n} f(x_i;\mu_0,\sigma^2)}{\prod_{i=1}^{n} f(x_i;\mu_1,\sigma^2)} = e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}[\mu_0^2-\mu_1^2+2(\mu_1-\mu_0)x_i]}$$

Substituting this in Equations (3.1) and (3.2), and after some simplifications, the decision rule reduces to reject H0 if $\sum_{i=1}^{n} x_i \geq -\frac{\sigma^2 lnA}{\mu_1-\mu_0} + \frac{1}{2}(\mu_0 + \mu_1)n$

accept H1 if $\sum_{i=1}^{n} x_i \leq -\frac{\sigma^2 lnB}{\mu_1-\mu_0} + \frac{1}{2}(\mu_0+\mu_1)n$

Continue sampling otherwise

An expression for the theoretical average sample size required is as follows :

$$z = \ln\left(\frac{f(x;\mu_0,\sigma^2)}{f(x;\mu_1,\sigma^2)}\right) = -\frac{1}{2\sigma^2}(2(\mu_1-\mu_0)x + \mu_0^2 - \mu_1^2)$$

Which implies that

$$E(Z/H_0) = \frac{(\mu_1-\mu_0)^2}{2\sigma^2} \quad \text{and} \quad E(Z/H_1) = -\frac{(\mu_1-\mu_0)^2}{2\sigma^2}$$

Substituting these in Equations (2.14) and (2.15) it follows

$$E(N/H_0) \approx \frac{2\sigma^2}{(\mu_1-\mu_0)^2}[\alpha\ln(A) + (1-\alpha)\ln(B)]$$

$$E(N/H_1) \approx -\frac{2\sigma^2}{(\mu_1-\mu_0)^2}[(1-\beta)\ln(A) + \beta\ln(B)]$$

For the fixed-sample size approach, the sample size corresponding to specified error sizes α and β is

$$n(\alpha,\beta) = \left[\frac{\sigma(Z_\alpha + Z_\beta)}{\mu_1 - \mu_0}\right]^2$$

Therefore, the relative efficiency of SPRT for normal distribution is

$$R_e(H_0) = \frac{1}{2}(Z_\alpha + Z_\beta)^2[\alpha\ln(\frac{\alpha}{1-\beta}) + (1-\alpha)\ln(\frac{1-\alpha}{\beta})]^{-1} \quad \ldots\ldots\ 3.5$$

$$R_e(H_1) = -\frac{1}{2}(Z_\alpha + Z_\beta)^2[(1-\beta)\ln(\frac{\alpha}{1-\beta}) + \beta\ln(\frac{1-\alpha}{\beta})]^{-1} \quad \ldots\ldots\ 3.6$$

When $\alpha = \beta$, (3.5) and (3.6) reduce to

$$R_e(H_0) = R_e(H_1) = 2\,Z_{1-\alpha}^2[(1-2\alpha)\ln(\frac{1-\alpha}{\alpha})]^{-1} \quad \ldots\ldots\ 3.7$$

From Equations 3.5 to 3.7, we see that relative efficiency does not depend on d = $\mu_1 - \mu_0$. The effect of d on relative efficiency will also be studied.

3.2.4 The Simulation Program.

The tasks of the computer program given in Appendix A can be summarized as follows. It has two main procedures. Namely, SIMSPRTB and SIMSPRTN which perform tests for Bernoulli and normal distributions respectively. To test a single hypothesis, say $H0{:}\theta=\theta_0$ against $H1{:}\theta=\theta_1$, for either distributions, test will be done 32000 times. Each time the type of decision made and the sample size required will be registered. The program is also capable of running under H0 and H1 alternatively. That is assuming the true parameter is θ_0 or θ_1. The other subprograms are UNIFORMR, RCRDPEDD and RCRDNEDD. UNIFORMR is a function which draws observation randomly from Uniform(0,1) on request. The procedures RCRDPEDD and RCRDNEDD are responsible for recording results at some iteration steps.

The inputs of the program are :

i) The type of distribution (Bernoulli or normal)

ii) Error sizes (α and β). Five cases are chosen
(α,β) = (0.05,0.05) , (0.05,0.01) , (0.01,0.05) , (0.01,0.01) and (0.05,0.10)

iii) Parameter under H0 and H1. For Bernoulli distribution, there are 26 cases while for normal distribution the variance is set to 1 and 11 different combinations of means are considered.

And the outputs are:

i) Record of the sample size required to reach at a decision

ii) The proportion of times H0 is erroneously rejected or accepted out of 32000 SPRTs. In other words, estimates of the true error probabilities

iii) The empirical average sample size required

iV) Record of proportion of times H0 is erroneously rejected or accepted at iteration steps 1000,1500,2000,...,32000

V) Record of number of times H0 is erroneously rejected or accepted out of each subsequent 100 SPRTs. In this way we have a set of 320 observations.

CHAPTER IV

ANALYSIS

In this section the data generated by the program described in the previous chapter will be analyzed. The outputs for the different cases are given in summary form in Tables B1 to B10 (Appendix B). Here some particular results are picked from the Tables to study the effect of the different parameters under H0 and H1 on the estimates of α_t and β_t , distribution of sample size required and the average sample size and hence the efficiency of SPRT. Specially, emphasis will be given to the effect of $d = \theta_1 - \theta_0$ on the different estimates.

4.1 Estimates of the True Error Probabilities.

Computation of the estimates of α_t or β_t requires counting the number of times a true hypothesis is erroneously rejected or accepted out of the total number of tests performed. For instance, consider computing an estimate of α_t. In the program we assume that the parameter under H0 is true, and hence observations will be generated from the distribution under consideration using this parameter to perform the tests. The desired estimate is then the ratio of the number of times H0 is rejected to the total number of tests performed (32000). Similarly, to compute an estimate of β_t, we assume H1 is true and count the number of times H0 is accepted.

$$\hat{\alpha}_t = \frac{\textit{Number of times H0 is rejected when it is true}}{32000}$$

$$\hat{\beta}_t = \frac{\textit{Number of times H0 is accepted when it is false}}{32000}$$

These are obtained for the different combinations of parameters and the specified error sizes. Columns 3 and 4 in Tables B1 to B10 contain the estimates. The plot of the estimates as a function of the number of tests performed is given below for some cases. From Figure 4.1 to 4.4 we see that the proportion of erroneous decisions stabilizes as the number of tests increases.

Bernoulli, H0:p=0.01 versus H1:p=0.05
(alpha = beta = 0.05)

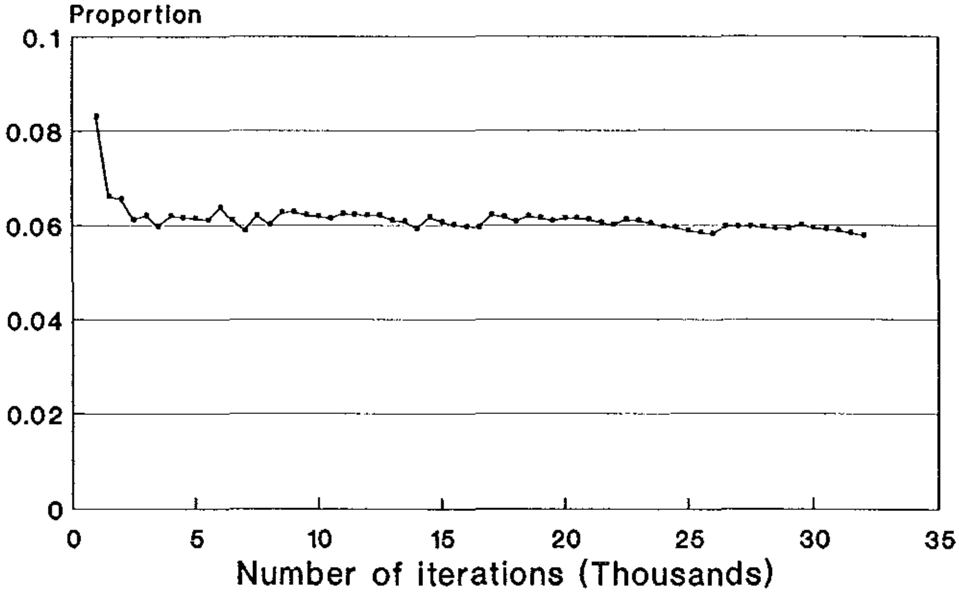

Fig.4.1: Estimates of αt Versus Number of Iterations.

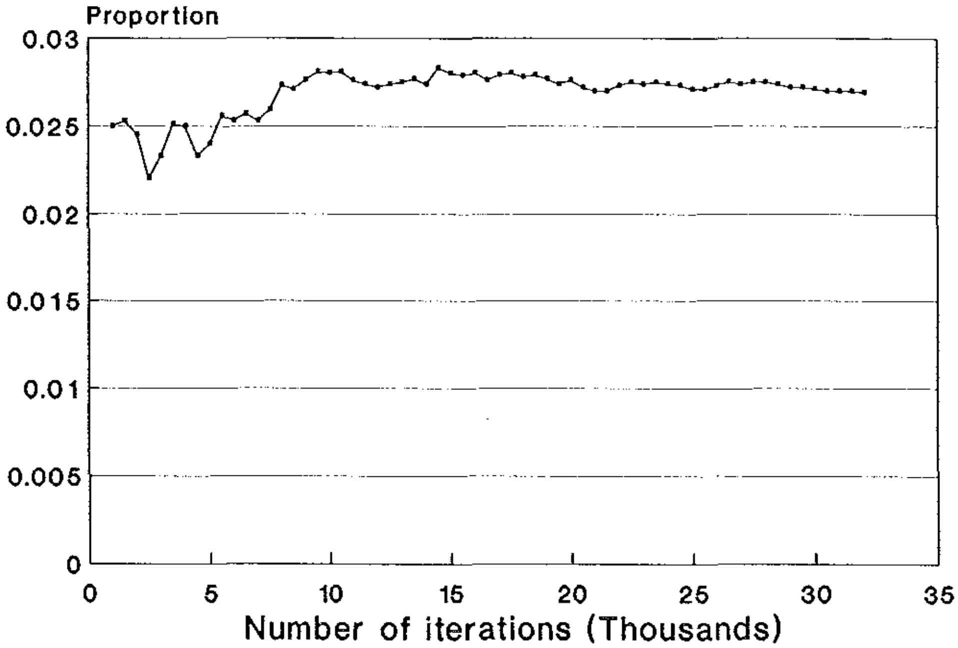

Fig.4.2: Estimates of Bt Versus Number of Iterations.

Normal, H0:μ=0 versus H1:μ=1
(alpha = beta = 0.05)

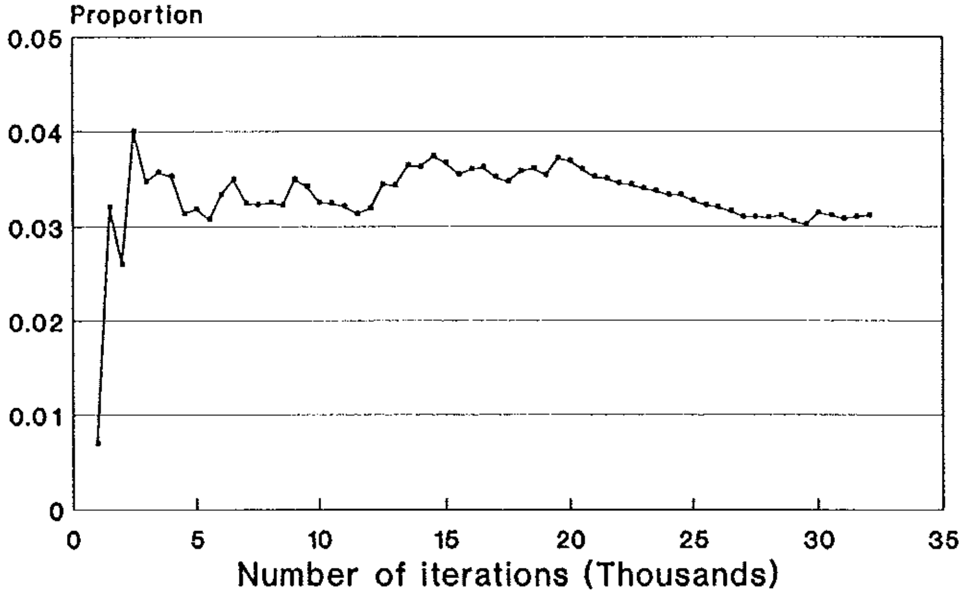

Fig.4.3: Estimates of αt Versus Number of Iterations.

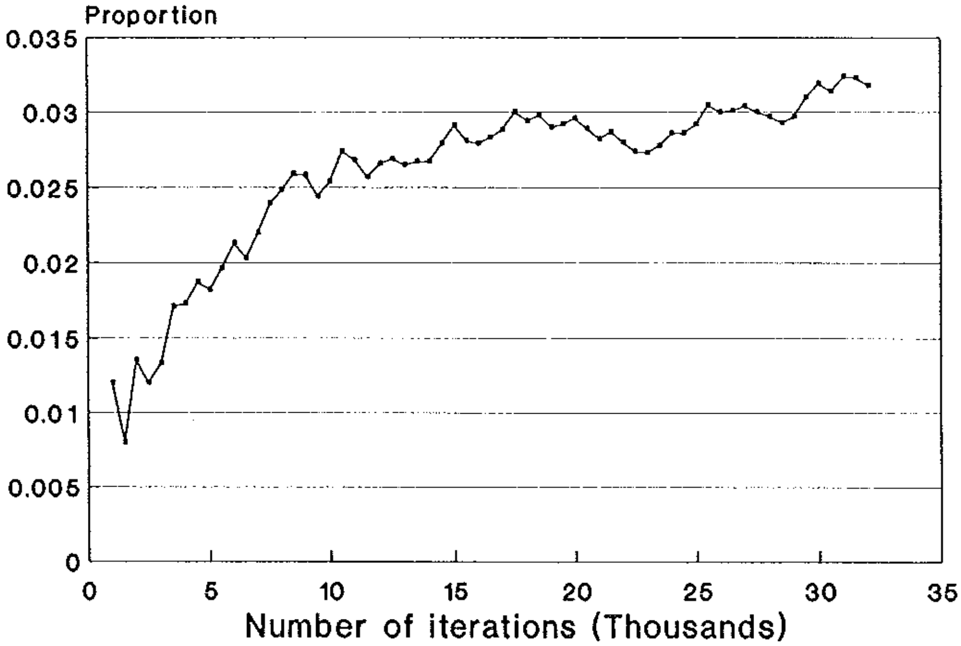

Fig.4.4: Estimates of βt Versus Number of Iterations.

4.1.1 The Estimates Versus $d = \theta_1 - \theta_0$

The magnitude of d has an impact on the estimates of α_t and β_t. As can be discerned from Figure 4.5 to 4.8, the estimates tend to decrease as d increases. In some cases they are much less than the corresponding specified error sizes when d is large. For example in testing $H0:\mu=0$ against $H1:\mu=3$ for normal distribution with $\sigma^2=1$ and $\alpha=\beta=0.05$, the estimates of α_t and β_t are observed to be 0.0068 and 0.0061 respectively. We have almost a similar result for Bernoulli distribution. Consider when $\alpha = \beta = 0.05$ (Table B1). If $d = p_1-0.1$ $(0.1 = p_0)$, the estimates have a tendency to increase as d decreases. Also fixing p_0 at 0.4 or 0.7 and varying p_1 gives the same result.

The other thing observed from the estimates is that for normal distribution, when $\alpha = \beta$, the estimates of α_t and β_t are very close to one another whatever d may take. This indicates α_t is close to β_t when $\alpha = \beta$. For Bernoulli distribution this property does not hold true in general. For example, in testing $H0 : p = 0.01$ versus $H1 : p = 0.05$ or $H0 : p = 0.01$ versus $H1 : p = 0.10$ with $\alpha = \beta = 0.05$, the result is quite different. The estimate of α_t is near α , but that of β_t, in both tests is less than β by a large magnitude. This indicates that α_t and β_t may not be equal despite $\alpha = \beta$.

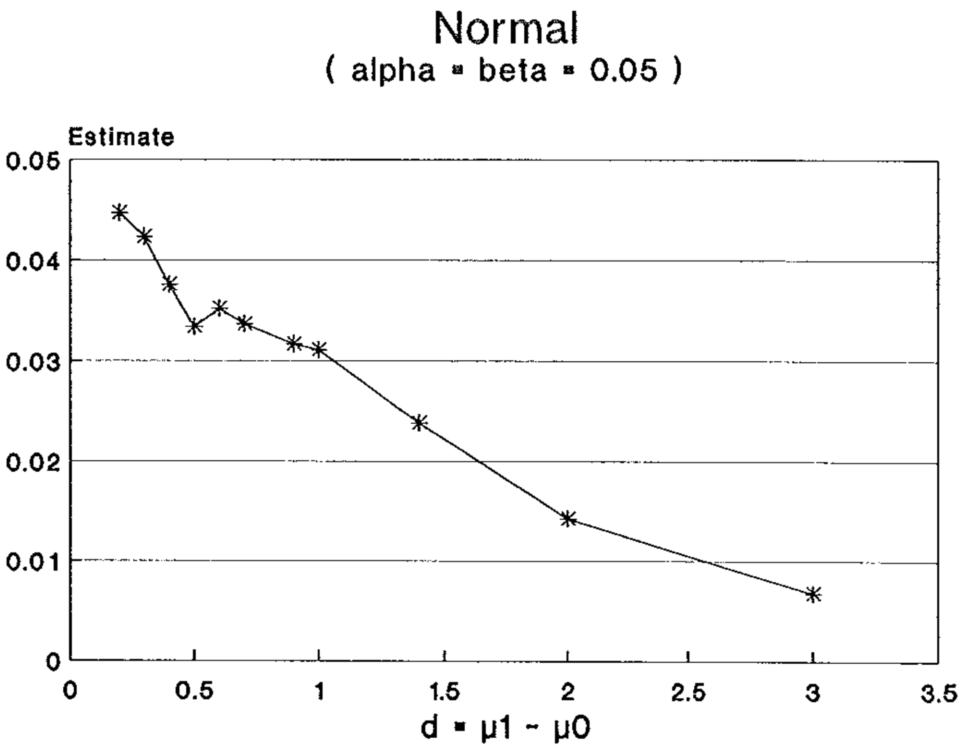

Fig.4.5: Estimate of αt versus d.

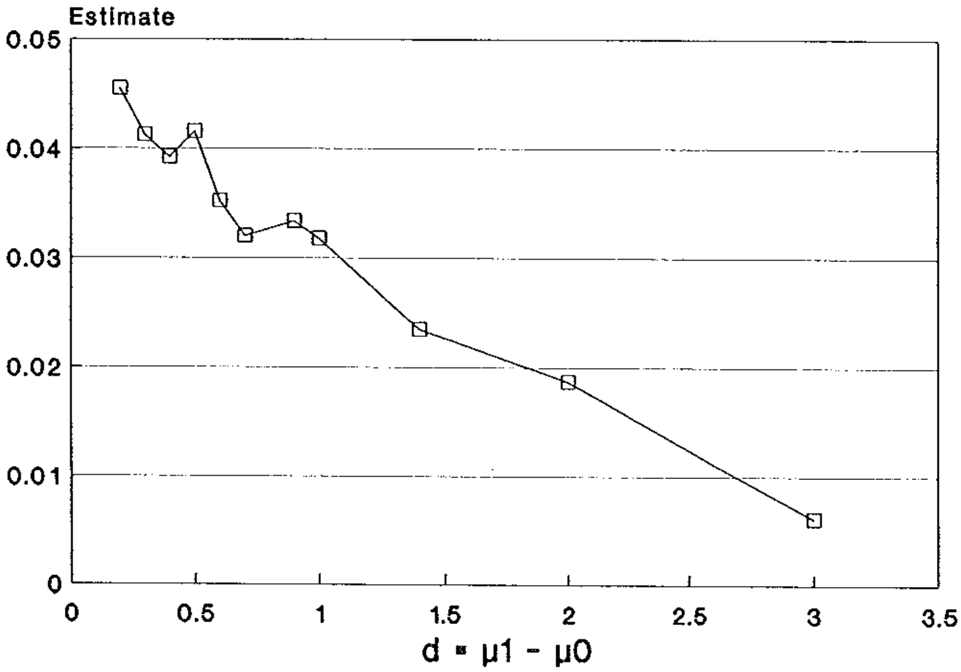

Fig.4.6: Estimate of βt versus d.

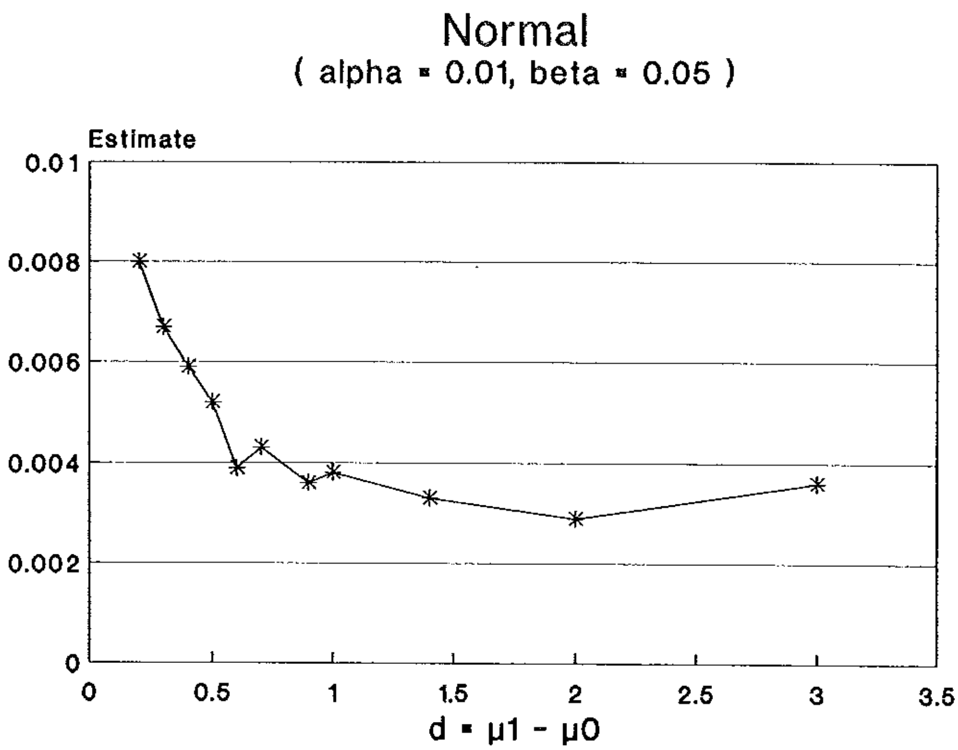

Fig.4.7: Estimate of α_t versus d.

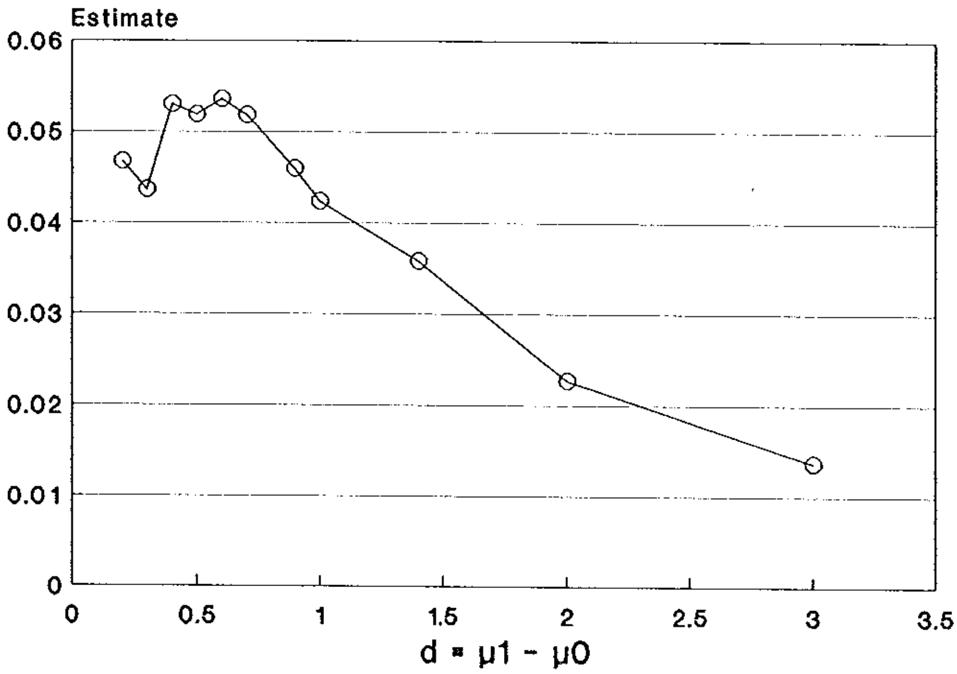

Fig.4.8: Estimate of β_t versus d.

4.1.2 Distribution of Number of Wrong Decisions

The other way of comparing the specified and the true error probabilities is by observing the distribution of number of erroneous decisions. Let us define the random variable Y as taking the value 1 if wrong decision is made and 0 otherwise. The probability distribution of Y is given by

$$P(Y=y) = \alpha_t^y (1-\alpha_t)^{1-y} \ , \ y = 0, 1 \quad if\ H_0\ is\ true$$

$$P(Y=y) = \beta_t^y (1-\beta_t)^{1-y} \ , \ y = 0, 1 \quad if\ H_1\ is\ true$$

That is E(Y/H0 true) = α_t and E(Y/H1 true) = β_t . In fact the estimators given in section 4.1 are estimators of the parameters of the above Bernoulli distributions. The estimators are simply the arithmetic mean of the observations $Y_1, Y_2, \ldots, Y_{32000}$. Where the Y_i's are random observations from one of the above distributions depending on whether H0 or H1 is true.

Now let us define a new random variable, say W, which counts the number of erroneous decisions in each subsequent 100 tests.

$$W_i = \sum_{j=100(i-1)+1}^{100i} Y_j \ , \quad i = 1,2,\ldots,320$$

It has a binomial distribution with number of trials 100 and the probability of success equal to α_t or β_t depending on whether H0 or H1 is true. In this way $W_1, W_2, \ldots, W_{320}$ constitute a random sample of size 320 from binomial(100,α_t) if H0 is true ,or from binomial(100,β_t) if H1 is true. Data on W_i's is available from the program and the frequency distribution constructed for 8 cases is given below (Table 4.1). In the Table w denotes the number of erroneous decisions out of 100 SPRTs and f_i is the corresponding frequency in case i.

Where

i=1	is normal with	$\alpha=\beta=0.05$,	$\mu_0=0$,	$\mu_1=1$	and H0	true
i=2	>>	>>	>>	>>	H1	>>
i=3	>>	$\alpha=0.01, \beta=0.05$	>>	>>	H0	>>
i=4	>>	>>	>>	>>	H1	>>
i=5	Bernoulli with	$\alpha=\beta=0.05$,	$p_0=0.01$,	$p_1=0.05$	H0	>>
i=6	>>	>>	>>	>>	H1	>>
i=7	>>	>>	$p_0=0.01$,	$p_1=0.10$	H0	>>
i=8	>>	>>	>>	>>	H1	>>

The idea is to compare the f_i's with expected frequencies of binomial(100,α) or binomial(100,β) for a sample of size 320. Table 4.2 and 4.3 contain the expected frequencies of binomial(100,0.05) and binomial(100,0.01) respectively.

Now if we compare $f_1, f_2, f_4, f_5, f_6, f_7$ and f_8 with the expected frequencies in Table 4.2, we see that there is a large discrepancy between the two. In particular the frequency corresponding to w = 0 in f_1, f_2, f_4,f7 and f_8 is very large as compared with 1.89. This indicates that a right decision is made in many experiments and by more than 95 per cent of the time. As a result α_t and β_t are less than α and β respectively in most cases. Comparing f_3 with Table 4.3 also yields the same result.

In summary, the high positive skewness of most of the frequency distributions in Table 4.1 with mode at 0 tells us that α_t and β_t are less than their corresponding specified error sizes.

Table 4.1 : Frequency distribution of number of erroneous decisions in 100 SPRTs.

w	f1	f2	f3	f4	f5	f6	f7	f8
0	212	213	299	172	14	34	118	144
1	1	3	0	4	50	64	8	13
2	5	0	0	9	36	72	12	33
3	13	10	5	12	50	54	27	33
4	14	13	4	12	30	42	13	43
5	14	7	4	13	29	23	26	10
6	11	15	2	15	19	18	23	10
7	6	4	0	17	18	8	18	15
8	1	8	3	7	15	1	7	11
9	5	5	1	3	5	3	13	2
≥10	38	42	2	56	54	1	55	6

Table 4.2: Expected frequency of bin(100,0.05) for a sample of size 320.

x	E.F
0	1.89
1	9.97
2	25.98
3	44.66
4	57.01
5	57.61
6	48.00
7	33.93
8	20.76
9	11.17
10	9.02
	320.00

Table 4.3: Expected frequency of bin(100,0.01) for a sample of size 320.

x	E.F
0	117.13
1	118.31
2	59.16
3	19.52
4	4.78
5	1.10
	320.00

4.2 Distribution of Sample Size Required (N)

From the program, a set of 32000 observations (sample sizes) are available for a single test. Frequency distributions are constructed and the pictorial representations are given below for some cases. In general, the distributions are positively skewed.

Figure 4.9 to 4.13 are for Bernoulli distribution. It has a strange pattern. The relative frequencies rise and fall within certain intervals of sample size and diminish slowly as sample size increases. When p_0 and p_1 are close to one another and are in the proximity of 0.5, the ups and downs will get closer and closer as in Figure 4.13. And in other tests, as can be seen from Figures 4.9 and 4.11, some integers have negligible chance of occurrence. For example, if we consider Figure 4.9, the integers 72,112,152,192,232,272,312,... have relatively high chances and their corresponding chances decrease as the size increases. It was also observed that N will not take some integer values in some tests. A case in point is when testing H0: p=0.1 against H1: p=0.9 with $\alpha = \beta = 0.05$. The sample size cannot be an odd integer.

In the case of the normal distribution (Fig. 4.14 to 4.18), the relative frequencies fall slowly. Moreover, when $\alpha = \beta$ the distributions under H0 and H1 appear to be the same as can be seen from the graphs and the various statistics (mean, variance and skewness). And when $\alpha < \beta$, the skewness for the distribution under H0 is less than the skewness of the distribution under H1.

Bernoulli, H0:p=0.01 versus H1:p=0.05
(alpha = beta = 0.05)

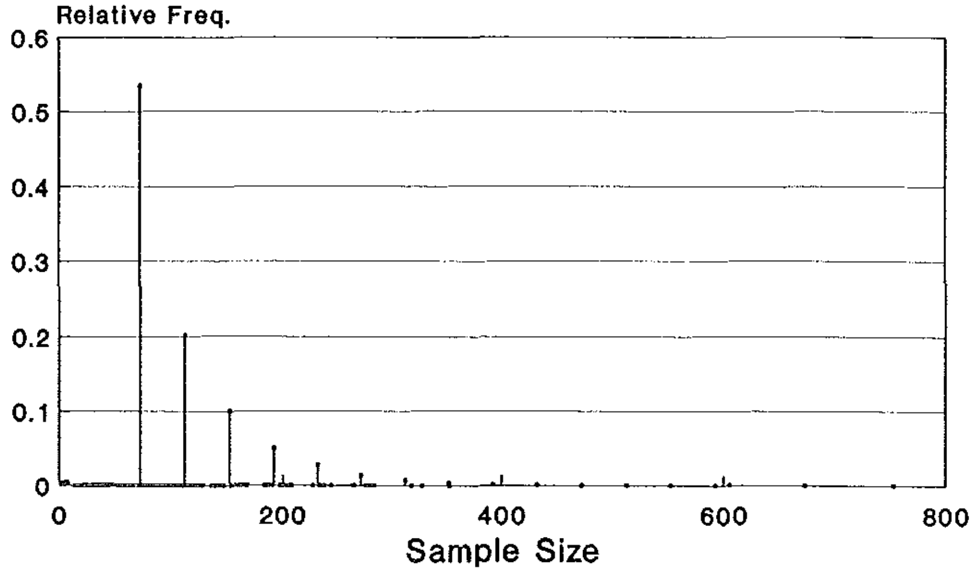

Fig.4.9 : Empirical sample size distribution when H0 is true.

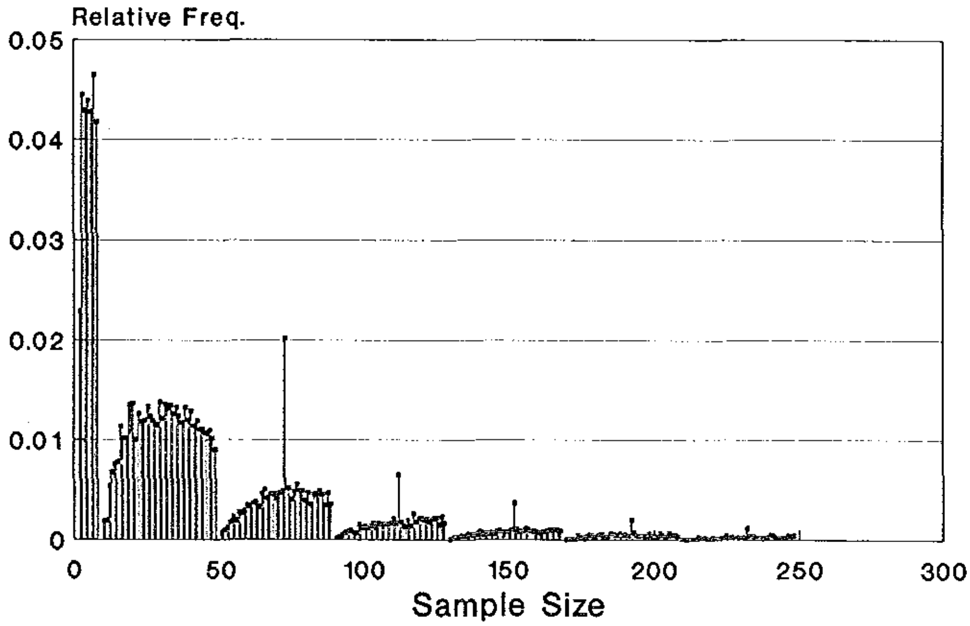

Fig.4.10 : Empirical sample size distribution when H1 is true.

Bernoulli, H0:p=0.01 versus H1:p=0.10
(alpha = beta = 0.05)

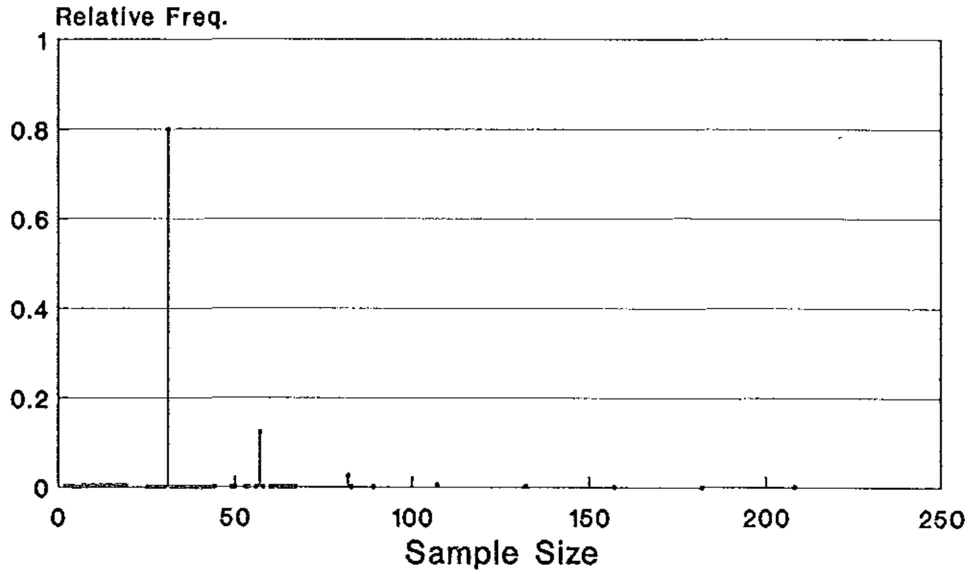

Fig.4.11: Empirical sample size distribution when H0 is true.

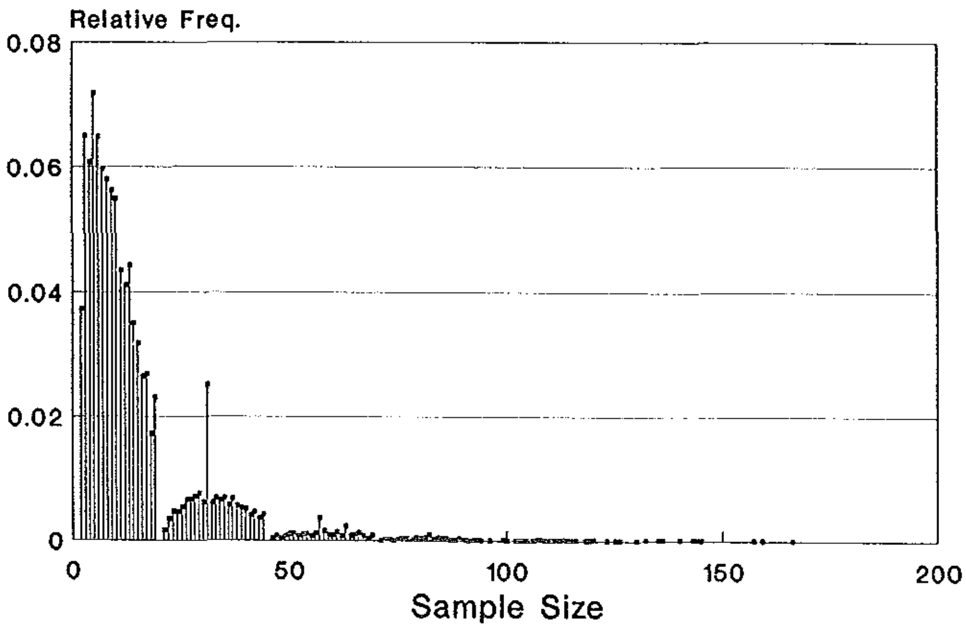

Fig.4.12: Empirical sample size distribution when H1 is true.

Bernoulli, H0:p=0.4 versus H1:p=0.5
(alpha = beta = 0.05)

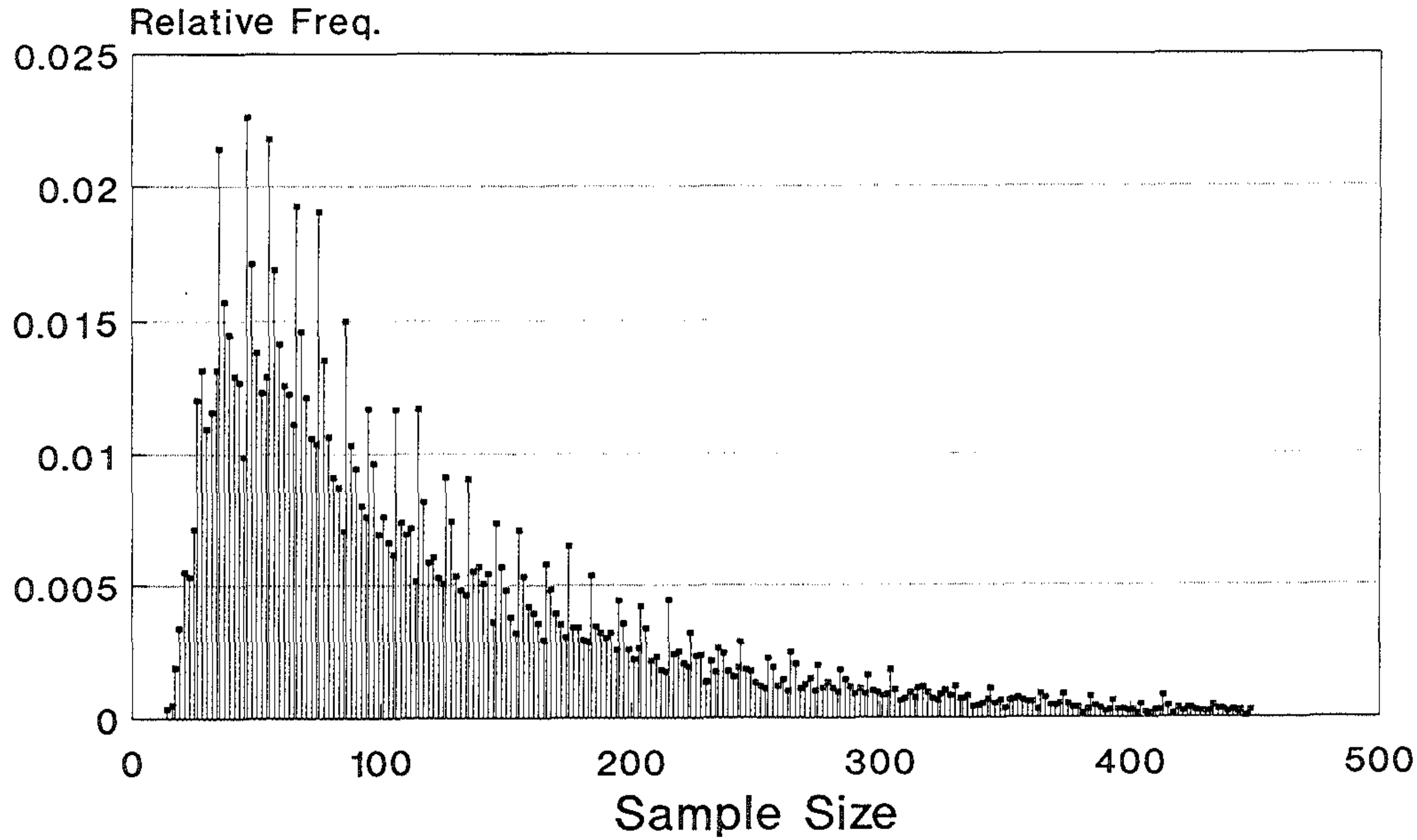

Fig.4.13: Empirical sample size
distribution when H1 is true.

Normal, H0:μ=0 versus H1:μ=1, sigma = 1
(alpha = beta = 0.05)

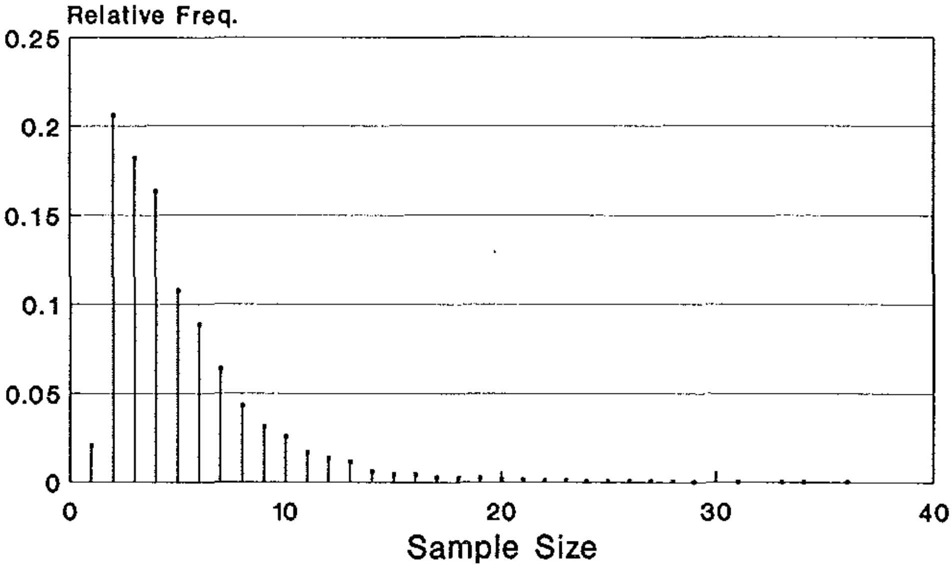

Fig.4.14: Empirical sample size
distribution when H0 is true.

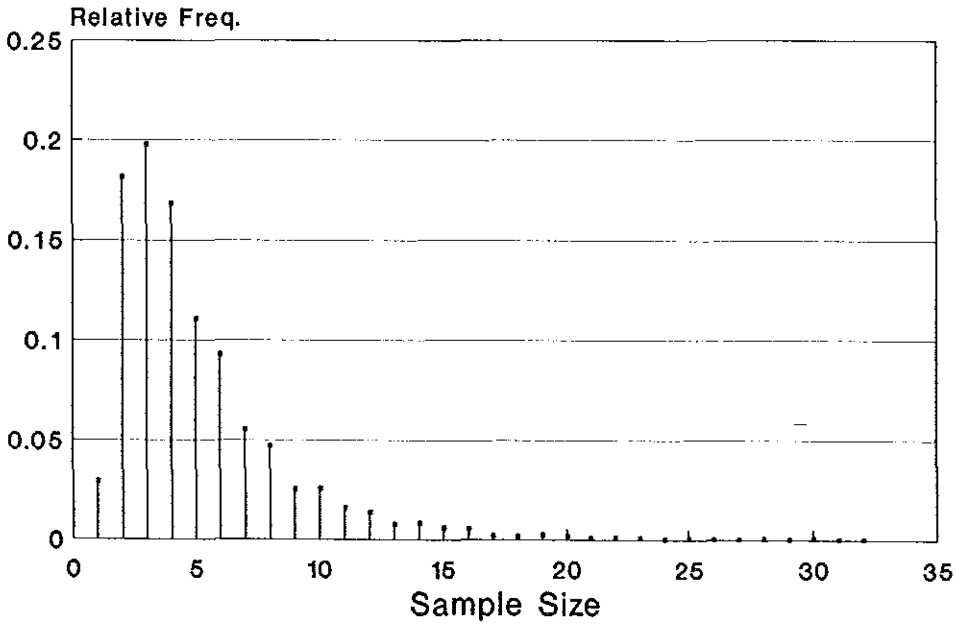

Fig.4.15: Empirical sample size
distribution when H1 is true.

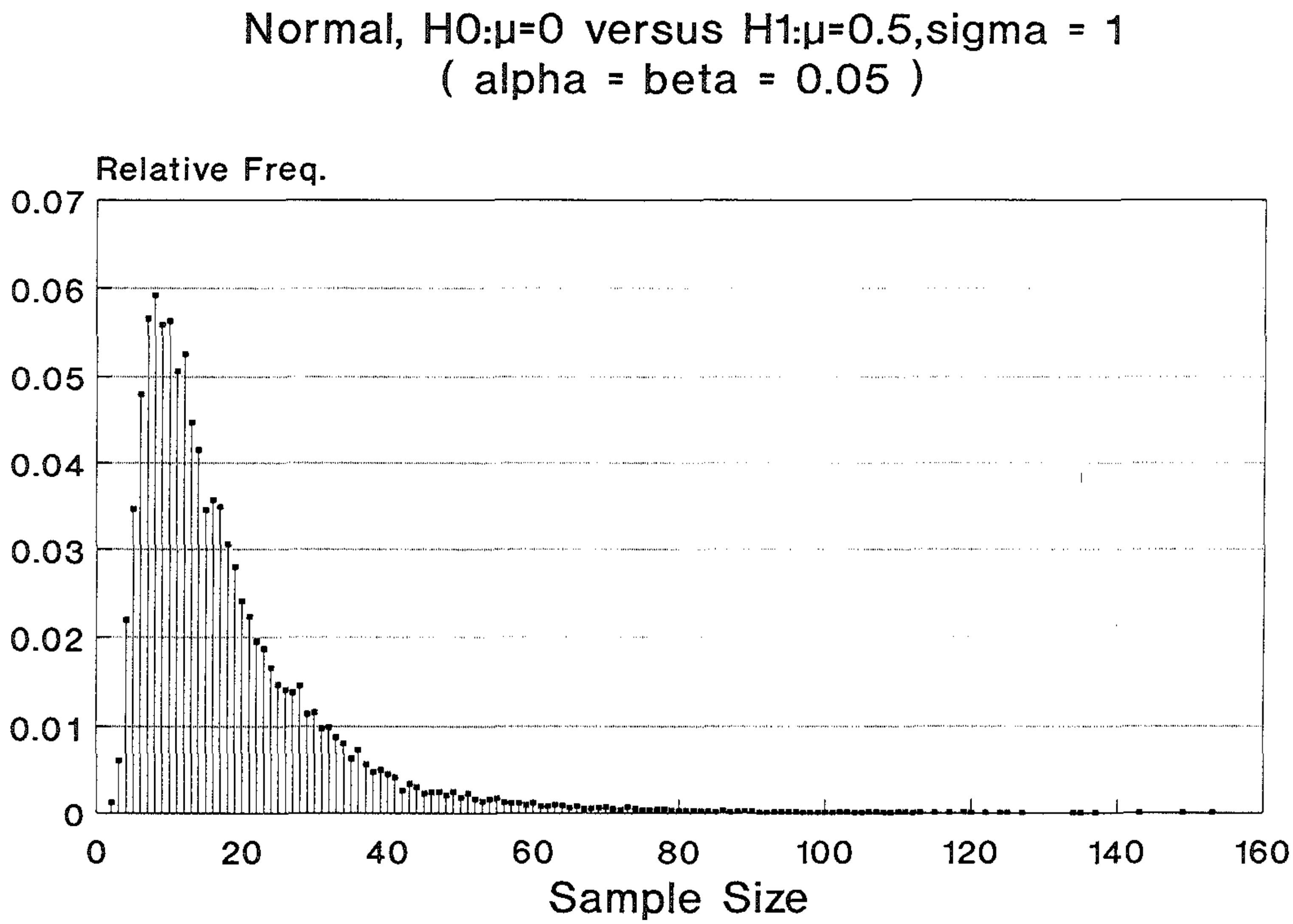

Fig.4.16: Empirical sample size distribution when H0 is true.

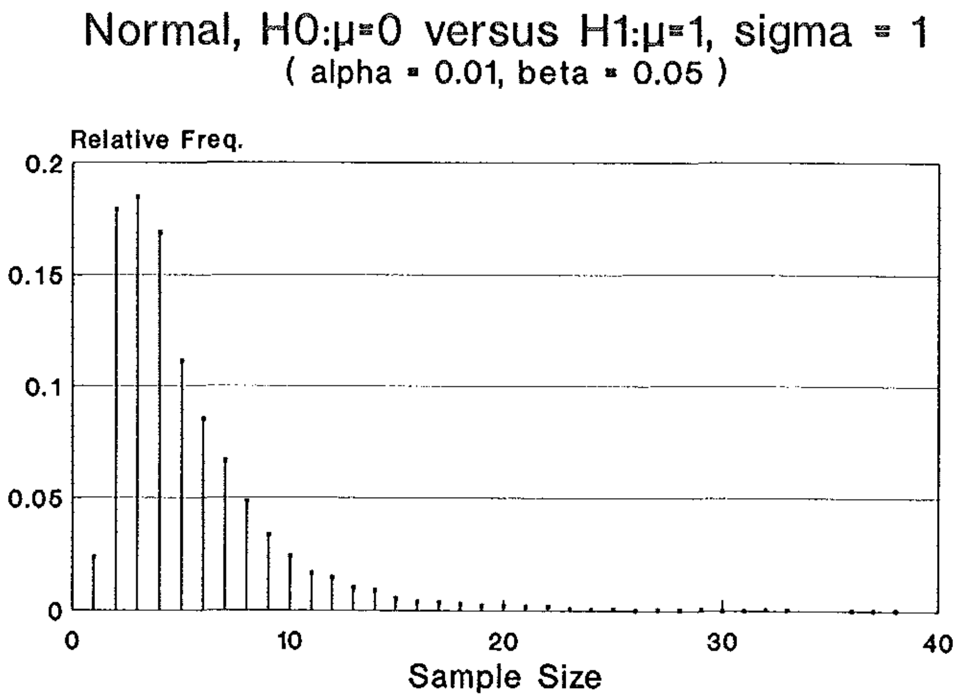

Fig.4.17: Empirical sample size distribution when H0 is true.

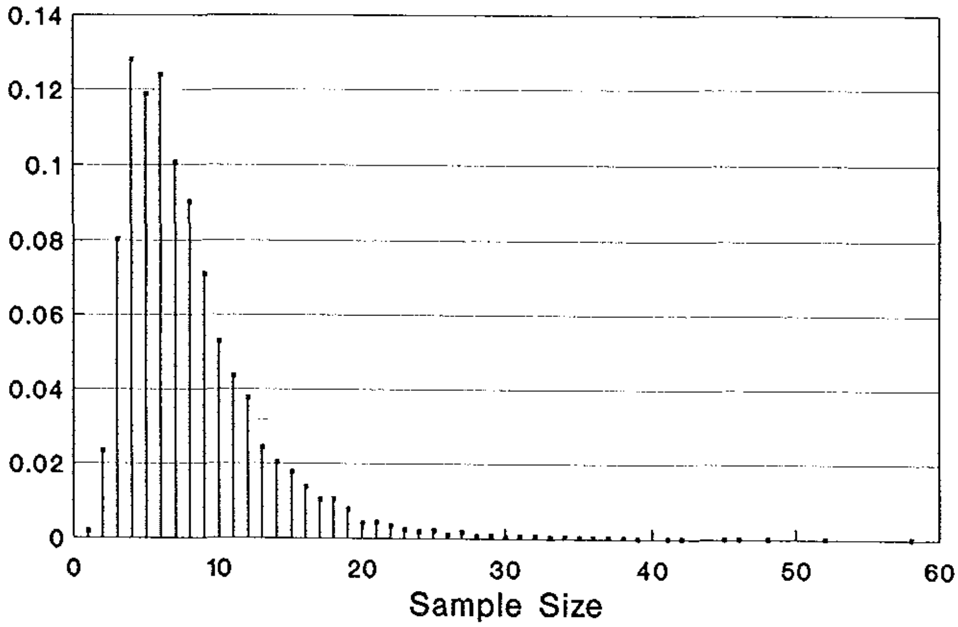

Fig.4.18: Empirical sample size distribution when H1 is true.

4.2.1 Empirical Mean and Variance of N

The mean, ASN by large depends on $d = \theta_1-\theta_0$. For the two distributions it increases as d decreases. The formula derived by Wald for ASN sometimes substantially underestimates or overestimates the actual value. To assess the discrepancy, an empirical ASN is computed for each test. In the Tables in Appendix B both the empirical and theoretical ASNs are displayed. In the vast majority of cases, the formula leads to overestimations. Exceptions are for normal distribution ($\sigma^2 = 1$) when the parameters under H0 and H1 are very close ($d \leq 0.2$) or very far apart ($d \geq 1.4$) and Bernoulli distribution when the parameters are far apart ($d \geq 0.7$). In general, for normal distribution, the ASN decreases rapidly as d increases (Figure 4.19 to 4.22). When $\alpha = \beta$, the theoretical ASN under H0 and H1 are equal and the empirical ASNs are very close to one another. When $\alpha < \beta$, the ASN under H0 is less than the ASN under H1, and when $\alpha > \beta$ that under H0 is greater.

And for the Bernoulli distribution, it is true that for fixed p_0 the ASN decreases as d increases. However, this does not mean that the ASN in testing $H0{:}p=p_0$ against $H1{:}p=p_1$ is less than the ASN in testing $H0{:}p=p_2$ against $H1{:}p=p_3$ for all p_0,p_1,p_2 and p_3 satisfying $p_1-p_0 > p_3-p_2$. For example, if we consider testing H0:p=0.4 against H1:p=0.5 and H0:p=0.01 against H1:p=0.10, even though d in the former test is greater, the latter test has a smaller ASN. Thus, for Bernoulli distribution, in addition to the magnitude of d, the relative positions of p_0 and p_1 (near 0 or 0.5) matters.

As far as the variance is concerned, it has almost the same property as the mean for both distributions. It decreases rapidly as d increases. This implies that in tests with small d, the chance is high that one draws a large number of observations before reaching at a decision. The next section also deals with a matter related to this.

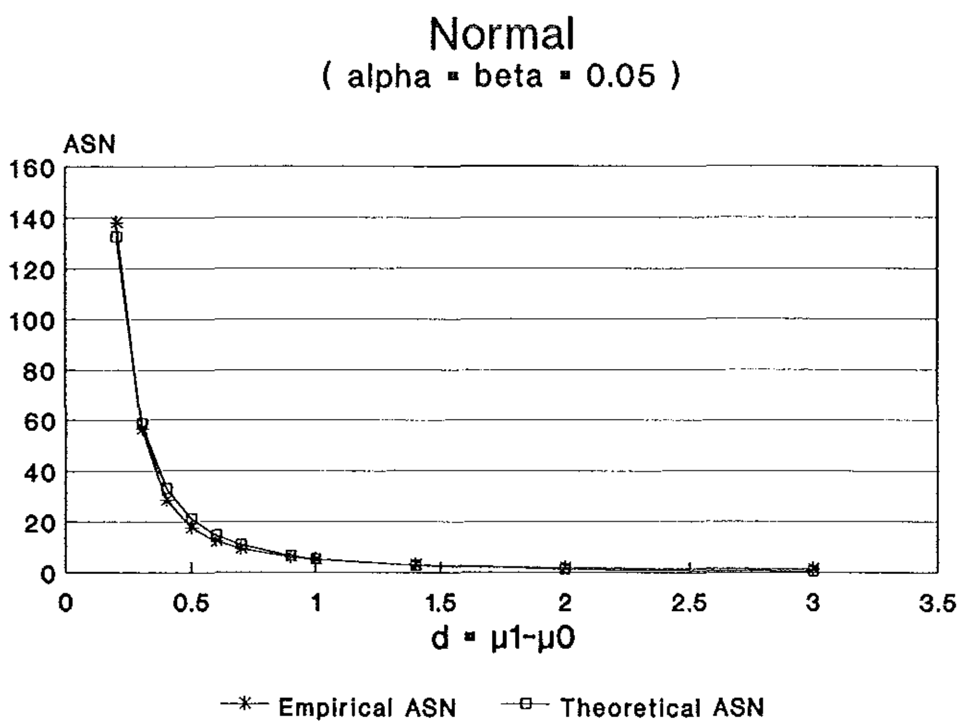

Fig 4.19: ASN under H0 versus d.

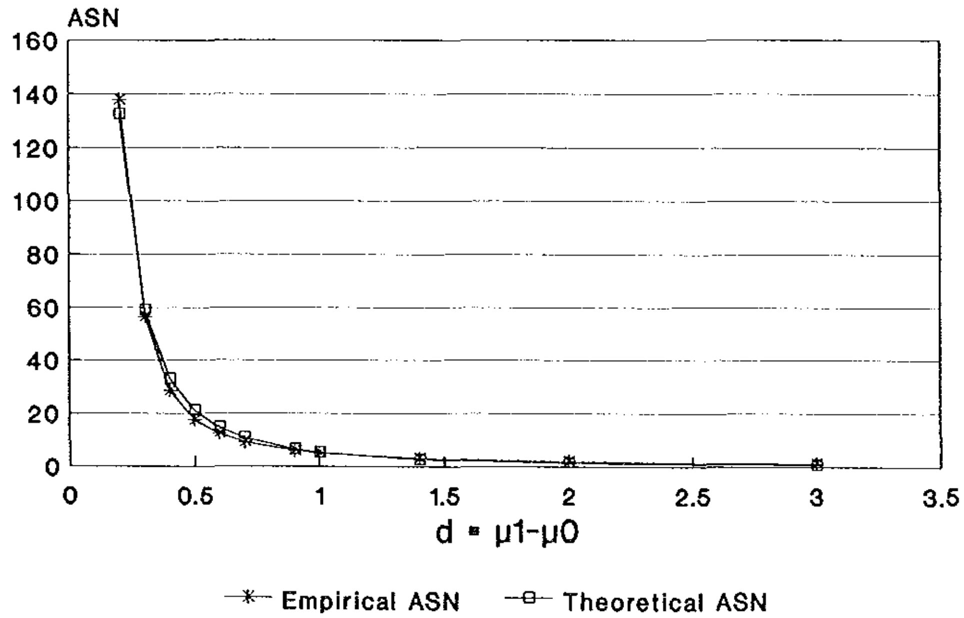

Fig. 4.20: ASN under H1 versus d.

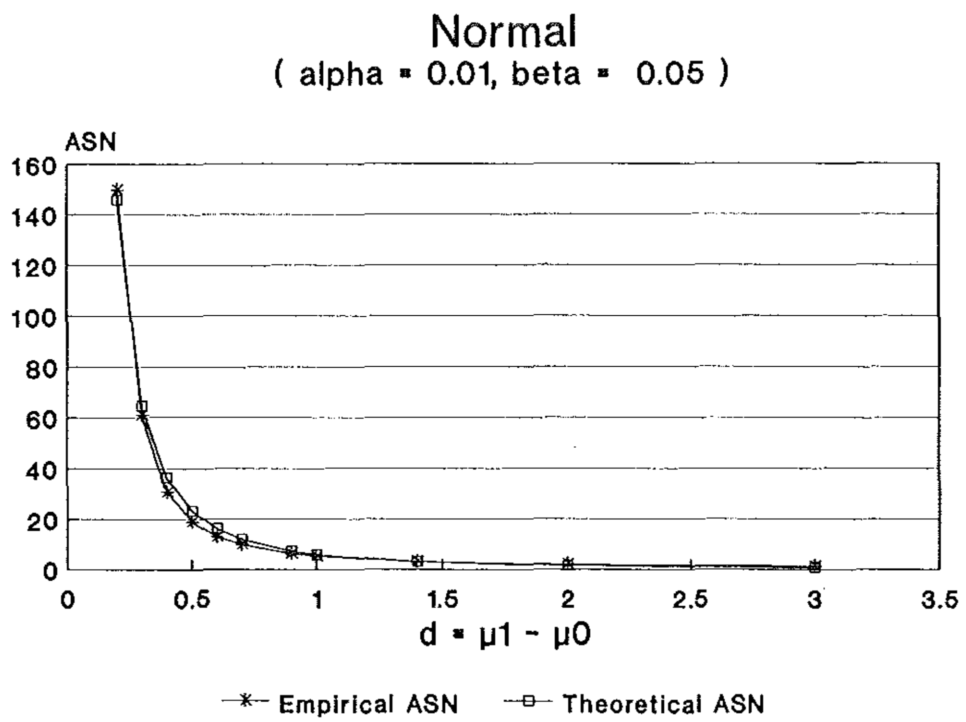

Fig.4.21: ASN under H0 versus d.

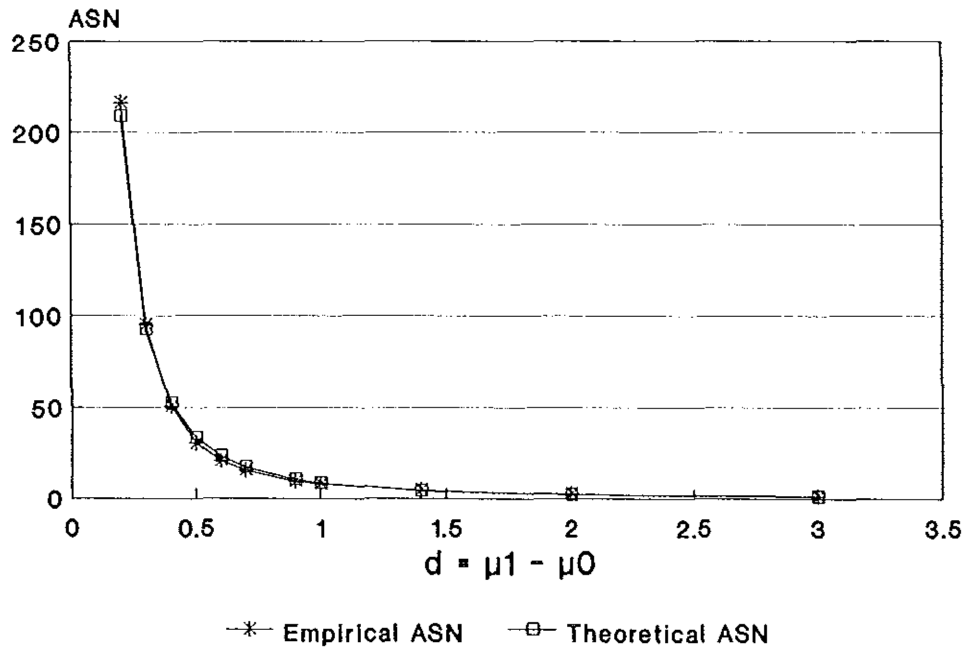

Fig.4.22: ASN under H1 versus d.

4.2.2 Distribution of N as Compared With the Sample Size Required in Non-sequential Test.

The optimal property of SPRT that comes first to mind is that it has small ASN as compared to sample size, $n(\alpha,\beta)$, of a comparable fixed-sample size test having the same error probabilities. But in reality, it is likely that one comes across sequences of observations which lead to drawing many more observations than $n(\alpha,\beta)$. The ASN sometimes gives us a false impression about the termination of the process. That is, even though ASN is less than or equal to $n(\alpha,\beta)$, there are chances that the sample size required in SPRT is by far greater than $n(\alpha,\beta)$. To see this effect, the empirical frequency distribution of N is compared below (Tables 4.4 and 4.5) with some multiples of $n(\alpha,\beta)$. For each test the proportion of times N exceeded some multiple of $n(\alpha,\beta)$ is computed. In both Tables distribution under H0 is considered.

Table 4.4

Percent of N greater than some multiple of $n(\alpha,\beta)$ for bernoulli distribution. ($\alpha = \beta = 0.05$)

			Percent of N greater than			
p_0	p_1	$n=n(\alpha,\beta)$	n	1.25 n	1.5 n	2 n
0.4	0.5	265	8.2	4.2	2.1	0.5
0.4	0.6	65	4.9	2.1	0.9	0.2
0.01	0.05	170	10.8	5.7	2.9	0.8
0.01	0.10	53	15.5	3.2	3.2	0.7

In the last two tests of the above Table we see that there is a chance greater than 1 in 10 of dealing with a sample of size greater than $n(\alpha,\beta)$.

Table 4.5

Percent of N greater than some multiple of n(α,β) for normal distribution. (α = β = 0.05)

			Percent of N greater than			
μ_0	μ_1	n=n(α,β)	n	1.25 n	1.5 n	2 n
0	0.2	271	27	5.0	2.4	0.6
0	0.3	120	8.7	4.4	2.3	0.6
0	0.4	68	6.3	3.2	1.6	0.4
0	0.5	43	4.6	2.4	1.2	0.3
0	0.6	30	4.7	2.0	0.9	0.2
0	0.7	22	4.9	2.4	1.1	0.2
0	0.9	13	5.6	3.0	1.7	0.4
0	1	11	5.3	2.9	1.4	0.3
0	1.4	6	4.9	2.9	1.1	0.3
0	2	3	7.2	7.2	2.6	0.3
0	3	1	24.6	24.6	24.6	2.7

And from Table 4.5 we see that except when $\mu 1-\mu 0$ is small or large, the chance is nearly 1 in 20 that one deals with a sample of size greater than the sample size necessary in non-sequential test.

4.3 The Empirical Relative Efficiency of SPRT.

The expression given before for the computation of relative efficiency makes use of the ratio of n(α,β) to E(N/θ). However, the sample size one needs in fixed-sample approach to maintain equal error probabilities as that of SPRT for given α and β is not really n(α,β) but $n(\alpha_t,\beta_t)$. Moreover, the theoretical ASN (E(N/θ)) was observed to underestimate or overestimate the actual parameter. In what follows we estimate efficiency using the empirical results and compare it with that of Wald.

The computation is based on

$$\hat{R}_e(\theta) = \frac{n(\hat{\alpha}_t,\hat{\beta}_t)}{[Empirical\ ASN/\theta]} \qquad \ldots\ldots\ 4.1$$

For normal distribution the results are given below in Table 4.6. Note that the efficiency is 2.055 using the formula defined by Wald (Equation 3.7) and it remains the same for all $d = \mu_0 - \mu_1$.

Table 4.6

Empirical estimate of the efficiency of SPRT for normal distribution. ($\alpha = \beta = 0.05$)

			Empirical ASN		Estimate of Effi.	
μ_0	μ_1	$n(\hat{\alpha}_t, \hat{\beta}_t)$	Under H0	Under H1	Under H0	Under H1
0	0.2	287.1	138.1	137.9	2.0788	2.0818
0	0.3	133.1	56.7	56.5	2.3481	2.3563
0	0.4	78.3	28.5	28.7	2.7470	2.7278
0	0.5	50.9	17.6	17.5	2.8909	2.9075
0	0.6	36.4	12.5	12.5	2.9099	2.9099
0	0.7	27.7	9.4	9.2	2.9421	3.0060
0	0.9	16.8	6.0	6.0	2.8007	2.8007
0	1	13.8	5.0	5.0	2.7674	2.7674
0	1.4	8.0	3.0	3.0	2.6766	2.6766
0	2	4.6	1.8	1.8	2.5360	2.5360
0	3	2.7	1.3	1.3	2.1146	2.1146

From the above Table we see that the relative efficiency rises until reaching some number and then falls. For $|\mu 1 - 0.7| < 0.3 = d$, it is nearly 3 times as efficient as the fixed-sample approach.

CHAPTER V

DISCUSSION AND CONCLUSION

In this research two probability distributions and five common cases of the specified error probabilities were considered for simulation purposes. The parameters of the distributions were selected as much as possible in a way that enables one to study the desired effects. The hope is then that something can be said about SPRT as applied to Bernoulli and normal distributions on the basis of the analysis made.

The parameters considered in the case of normal distribution may look far beyond what one may face in real problems. The test was about the equality of two normal distributions $N(\mu 0,1)$ and $N(\mu 1,1)$. Throughout the analysis, variance was assumed to be 1. The assumption was based on the fact that if all properties about the test of hypothesis for $N(\mu,1)$ is studied, it can be extended to test of hypothesis about the mean of $N(\mu,\sigma^2)$ for known σ^2 by a simple transformation. This is because the continue sampling region in testing the equality of $N(\mu 0,1)$ and $N(\mu 1,1)$ is the same as the continue sampling region in testing the equality of $N(\sigma\mu 0,\sigma^2)$ and $N(\sigma\mu 1,\sigma^2)$. Thus, it suffices to take the variance as 1. Further, the property of testing the equality of $N(\mu 0,1)$ and $N(\mu 1,1)$ is the same as the property of testing the equality of $N(\mu 0+c,1)$ and $N(\mu 1+c,1)$ for some constant c. Due to this property ,only a limited number of parameters were considered. Therefore, given a test $H0{:}\mu=\mu 0$ against $H1{:}\mu=\mu 1$ for known variance σ^2, one can get information on the estimates of ASN, α_t and β_t by referring to the test $H0{:}\mu=\mu 0/\sigma$ against $H1{:}\mu=\mu 1/\sigma$ in tables B6 - B10 depending on the chosen α and β. When we come to Bernoulli distribution, its asymmetry led to considering many combination of parameters.

The results of the analysis can now be summarized as follows:

i) For the normal distribution when $\alpha = \beta$, estimates of α_t and β_t decrease as d increases. And they are close to the common value $\alpha = \beta$ when d is small. When α is different from β, in most cases the estimates indicate that α_t and β_t are less than their corresponding specified error probabilities for large d. And in all cases the sum of the estimates of α_t and β_t is less than or equal to $\alpha + \beta$ as expected. In the case of the Bernoulli distribution, no such general statement can be made as the results extremely depend on the particular parameter values selected and not only on d = p1-p0. In other words, even if $\alpha = \beta$ it is not so easy to tell from the estimates the rate at which α_t and β_t fall or rise as d decreases. However, it is worth mentioning that the estimates computed for the different cases can help one to get at least a feeling about the magnitudes of α_t and β_t. In the case of Bernoulli distribution, in some tests, the estimates are observed to be much larger than their corresponding specified error sizes. This led to their sum being greater than $\alpha + \beta$, and this contradicts the theoretical result. The reason for this may be due to sampling fluctuation and the discrete steps (the 1 and 0's) involved in the test procedure.

ii) The approximate expression for the ASN in most cases overestimates the actual parameter.

iii) In the derivation of the relative efficiency of SPRT for normal distribution, it was observed that $d = \mu_1 - \mu_0$ has no effect on relative efficiency. But from the empirical results we see that d has some effect on it.

iV) Finally, regarding the sample size distribution, the empirical results revealed that in most cases the sample size (N) required in SPRT is less than the comparable sample size ,$n(\alpha,\beta)$, one needs in non-sequential test. For normal distribution when $\alpha = \beta = 0.05$, $\mu_0=0$, $0.4 \le \mu_1 \le 2$ and $\sigma^2=1$, about 95% of the time the sample size one needs in SPRT is less than the corresponding size necessary in non-sequential test. And for Bernoulli distribution, in testing H0: p=0.01 versus H1: p=0.05 and H0: p=0.4 versus H1: p=0.5 with $\alpha = \beta = 0.05$, about 90% of the time N is less than $n(\alpha,\beta)$. For the same α and β, when $p_0=0.01$ and $p_1=0.10$, the chance is nearly 85%. But, it should be noted that there is a chance that N can be greater than some factors (1, 1.25, 1.5 and 2) of $n(\alpha,\beta)$ as was observed from the empirical frequency distributions. And the chance decreases as the factor increases. N was realized to be even five times greater than $n(\alpha,\beta)$. That is, even though SPRT requires on average smaller sample size, there are still rare occasions when the situation may go worse.

APPENDIX A: THE SIMULATION PROGRAM

```
Program SimuSPRT (input,output);
{ This program
1.   Generates observations  (under H0 and H1) from bernoulli
     or normal distribution (depending on the choice of a
     user) to  test the simple versus  simple  alternative
     hypothesis about parameter of interest using SPRT ( 32000
     times )
2.   Records the sample size required  to terminate the process
3.   Screens estimates of ASN and the true error probabilities
4.   Records proportion H0 is erroneously rejected or accepted
     at some iteration steps
5.   Records number of error decisions in each 100 tests.}
{ }
uses Crt;
const
numgen = 32000;
type
  NumErrDec = array [1..numgen] of integer;{number of error
                                             decisions}
var { global variables }
    choice,i,j,n,countRej :integer;
    alpha,beta :real;{probabilities of Type-I and Type-II errors}
        A, B   :real;  { demarkation points }
Intr1,Intr2,slope,lowerltn,upperltn,n0,n1,AverageN,EstErr:real;
H0True :boolean;
  filevar1,filevar2,filevar3 :text;
  NumRej :NumErrDec; { Number of times H0 is rejected so far}
function uniformR:real;
  { Generates independent observations from Uniform(0,1). The
    method has been adopted from Algorithm AS 183, Wichmann and
    Will(1982,1984), McLeod(1985) and Zeisel (1986). Random is
    a built-in random number generator. }
var
  ix,iy,iz :integer;
  t :real;
  begin
    repeat
      ix := random(30000);
      iy := random(30000);
      iz := random(30000);
      while ix < 1 do ix := random(30000);
      while iy < 1 do iy := random(30000);
      while iz < 1 do iz := random(30000);
      ix := 171*(ix mod 177) - 2*(ix div 177);
      iy := 172*(iy mod 176) - 35*(iy div 176);
      iz := 170*(iz mod 178) - 63*(iz div 178);
      if (ix < 0) then ix := ix + 30269;
      if (iy < 0) then iy := iy + 30307;
      if (iz < 0) then iz := iz + 30323;
      t := frac(ix/30269.0 + iy/30307.0 + iz/30323.0);
    until t > 0.0;
    uniformR := t;
  end;  { of function uniformR }
```

```
procedure rcrdpedd;  {  This procedure records proportion H0 is
   erroneously  rejected  or  accepted  at the  iteration  steps:
   1000,1500,...,32000 }
var cnt :integer;
  ratio :real;
begin
  rewrite(filevar2);
  cnt := 500;
  for i := 1 to 63 do
    begin
      cnt := cnt + 500;
      ratio := NumRej[cnt]/cnt;
      if H0True then writeln(filevar2,cnt:7,ratio:9:4)
      else writeln(filevar2,cnt:7,(1 - ratio):9:4);
    end;
  close(filevar2);
end; { of procedure rcrdpedd }
{  -------------------------------------------------------- }
procedure rcrdnedd;   { This procedure records the number of
  erroneous decisions out of 100 SPRTs. The series consists of
  320 observations.}
var diff :integer;
begin
  rewrite(filevar3);
  if H0True then writeln(filevar3,NumRej[100]:4)
  else writeln(filevar3,(100 - NumRej[100]):4);
  for i := 1 to 319 do
    begin
      diff := NumRej[(i+1)*100] - NumRej[i*100];
      if H0True then writeln(filevar3,diff:4)
      else writeln(filevar3,(100 - diff):4);
    end;
  close(filevar3);
end; { of procedure rcrdnedd }
{  --------------------------------------------------------  }
procedure simsprtb;  { for bernoulli distribution }
var { local variables }
  obstn,sumobstn :integer;
  p0,p1,d,p,d0,d1 :real;
{ given : 1. demarkation points A and B
          2. parameters under H0 and H1, p0 and p1
   task : generates observations and tests hypothesis using sprt}
begin
  writeln('       Enter parameter of bernoulli distribution ');
  write('  Parameter under H0 = ');
  readln(p0);
  write('  Parameter under H1 = ');
  readln(p1);writeln;
  d := ln((p0*(1-p1))/(p1*(1-p0)));
  Intr1 := (ln(B))/d;
  Intr2 := (ln(A))/d;
  slope := -(ln((1-p0)/(1-p1)))/d;
  d0 := p0*d + ln((1-p0)/(1-p1));
  d1 := p1*d + ln((1-p0)/(1-p1));
  writeln('       E(N/H0) =',(n0/d0):7:3);
  writeln('       E(N/H1) =',(n1/d1):7:3);writeln;
```

```
writeln('*********** Iteration is going on ***********':54);
H0True := true;
repeat
  countRej := 0;
  AverageN := 0;
  if H0True
  then begin
         p := p0;
         assign (filevar1,'c:\B_SSDAT.H0');
         assign (filevar2,'c:\BER_PEDD.H0');
         assign (filevar3,'c:\BER_NEDD.H0');
      end
  else begin
         p := p1;
         assign (filevar1,'c:\B_SSDAT.H1');
         assign (filevar2,'c:\BER_PEDD.H1');
         assign (filevar3,'c:\BER_NEDD.H1');
      end;
  rewrite(filevar1);
  writeln;
  for i:=1 to numgen do
    begin { test }
      randomize;
      sumobstn := 0;
      n := 0;
      if p0 < p1
      then begin
             repeat
               begin  { selects observation from bernoulli
                        distribution   }
                 if uniformR > p then obstn := 0
                 else obstn := 1;
               end;
               n := n + 1;
               sumobstn := sumobstn + obstn;
               lowerltn := Intr1 + slope * n;
               upperltn := Intr2 + slope * n;
               if sumobstn <= lowerltn      {accept condition}
               then begin
                      writeln(filevar1,n);
                      AverageN := AverageN + (n/numgen);
                      NumRej[i] := countRej;
                   end
               else if sumobstn >= upperltn {reject condition}
               then begin
                      countRej := countRej + 1;
                      writeln(filevar1,n);
                      AverageN := AverageN + (n/numgen);
                      NumRej[i] := countRej;
                   end
          until (sumobstn <= lowerltn) or (sumobstn >= upperltn);
              end
      else begin  { when p0 > p1 }
             repeat
               begin { selects observation from bernoulli
                        distribution }
```

```
                    if uniformR > p then obstn := 0
                    else obstn := 1;
                  end;
                  n := n + 1;
                  sumobstn := sumobstn + obstn;
                  lowerltn := Intr2 + slope * n;
                  upperltn := Intr1 + slope * n;
                  if sumobstn <= lowerltn    { reject condition}
                  then begin
                         countRej := countRej + 1;
                         writeln(filevar1,n);
                         AverageN := AverageN + (n/numgen);
                         NumRej[i] := countRej;
                       end
                  else if sumobstn >= upperltn {accept condition}
                  then begin
                         writeln(filevar1,n);
                         AverageN := AverageN + (n/numgen);
                         NumRej[i] := countRej;
                       end
          until (sumobstn <= lowerltn) or (sumobstn >= upperltn);
          end
       end; { test }
       close(filevar1);
       rcrdpedd;
       rcrdnedd;
       EstErr := NumRej[numgen]/numgen;
       if H0True
       then begin
         writeln('Empirical A.S.N/H0 =':27,AverageN:7:3);
         writeln('  Estimate of true alpha = ',EstErr:5:4);
         end
       else begin
         writeln('Empirical A.S.N/H1 =':27,AverageN:7:3);
         writeln('  Estimate of true beta = ',(1-EstErr):5:4);
       end;
       sound(100);delay(100);nosound;
       H0True := not H0True;
   until H0True = true;
end; { of  procedure simsprtb}
{  -------------------------------------------------------  }
procedure simsprtn;  { for normal distribution }
label  37;
var { local }
  μ0,μ1,μ,d0,d1,obstn,obstn2,sumobstn,v1,v2,w :real;
  used,condition :boolean;
begin
  writeln('       Enter parameter of normal (μ,1) ');
  writeln('                  ( μ0 < μ1 )');
  write('  parameter under H0  = ');
  readln(μ0);
  write('  parameter under H1  = ');
  readln(μ1);
  Intr1 := -(ln(B))/(μ1-μ0);
  Intr2 := -(ln(A))/(μ1-μ0);
  slope := (μ0+μ1)/2;
```

```
   d0 := sqr(μ1-μ0)/2;
   d1 := -d0;
   writeln('        E(N/H0) =',(n0/d0):7:3);
   writeln('        E(N/H1) =',(n1/d1):7:3);writeln;
   writeln('*********** Iteration is going on ***********':54);
   H0True := true;
   repeat
     countRej := 0;
     AverageN := 0;
     if H0True
     then begin
            μ := μ0;
            assign (filevar1,'c:\N_SSDAT.H0');
            assign (filevar2,'c:\NOR_PEDD.H0');
            assign (filevar3,'c:\NOR_NEDD.H0');
          end
     else begin
            μ := μ1;
            assign (filevar1,'c:\N_SSDAT.H1');
            assign (filevar2,'c:\NOR_PEDD.H1');
            assign (filevar3,'c:\NOR_NEDD.H1');
          end;
     rewrite(filevar1);
     for i:=1 to numgen do
       begin { test }
         randomize;
         sumobstn := 0;
                n := 0;
         repeat
           used := false;
           repeat {    selects observation from a normal
               distribution, N(μ,1) using Polar Marsaglia method}
             v1 := 2*uniformR - 1;
             v2 := 2*uniformR - 1;
             w  := sqr(v1) + sqr(v2);
           until w < 1;
           obstn  := μ + v1*sqrt((-2*ln(w))/w);
           obstn2 := μ + v2*sqrt((-2*ln(w))/w);
37:        n := n + 1;
           sumobstn := sumobstn + obstn;
           lowerltn := Intr1 + slope * n;
           upperltn := Intr2 + slope * n;
           if sumobstn <= lowerltn
           then begin
                  writeln(filevar1,n);
                  AverageN := AverageN + (n/numgen);
                  NumRej[i] := countRej;
                end
           else if sumobstn >= upperltn
           then begin
                  countRej := countRej + 1;
                  NumRej[i] := countRej;
                  writeln(filevar1,n);
                  AverageN := AverageN + (n/numgen);
                end;
```

```
   condition := (sumobstn <= lowerltn) or (sumobstn >= upperltn);
        { condition of reaching at a decision }
          if not (condition or used)
          then begin
                 obstn := obstn2;
                 used  := true;
                 goto 37;
              end
        until condition;
      end; { test }
      close(filevar1);
      rcrdpedd;
      rcrdnedd;
      EstErr := NumRej[numgen]/numgen;
      if H0True
      then begin
             writeln('Empirical A.S.N/H0 =':27,AverageN:7:3);
             writeln('Estimate of true alpha = ',EstErr:5:4);
          end
      else begin
             writeln('Empirical A.S.N/H1 =':27,AverageN:7:3);
             writeln('Estimate of true beta = ',(1-EstErr):5:4);
          end;
      sound(100);delay(50);nosound;
      H0True := not H0True;
  until H0True = true;
end;  { of procedure simsprtn }
{     ==============================================     }
begin  { main }
  clrscr;
  writeln;writeln;
  writeln('For which distribution do you want to simulate ?':55);
  writeln;
  writeln('1   Bernoulli':30);
  writeln('2   Normal   ':30);
  write('Enter number corresponding to your choice  ':50);
  readln(choice);writeln;
  writeln(' Change  file names in which results will be stored');
  writeln('       when you change parameters !!!!  ');writeln;
  writeln('       Enter error probabilities ');
  write('  probability of type-I error  = ');
  readln(alpha);
  write('  probability of type-II error = ');
  readln(beta);
  A := alpha/(1-beta);
  B := (1-alpha)/beta;
  n0 := alpha*ln(A) + (1-alpha)*ln(B);
  n1 := (1-beta)*ln(A) + beta*ln(B);
  case choice of
    1: simsprtb;
    2: simsprtn;
  end;
  sound(75);delay(150);nosound;
end. { of main program }
```

APPENDIX B: RESULTS OF SIMULATION

TABLE B1: Result for Bernoulli Distribution
($\alpha = \beta = 0.05$)

Parameter		Estimate of		Average Sample Number Required: Empirical		Average Sample Number Required: Theoretical	
p_0	p_1	α_t	β_t	Under H0	Under H1	Under H0	Under H1
0.1	0.9	0.0150	0.0150	2.2	2.3	1.5	1.5
0.1	0.8	0.0160	0.0390	2.3	2.5	2.3	1.9
0.1	0.7	0.0250	0.0340	3.5	2.9	3.3	2.6
0.1	0.6	0.0290	0.0340	4.8	3.5	4.8	3.5
0.1	0.5	0.0430	0.0260	7.5	4.4	7.2	5.2
0.1	0.4	0.0430	0.0280	10.9	7.7	11.7	8.5
0.1	0.3	0.0440	0.0310	19.4	12.8	22.8	17.2
0.1	0.2	0.0560	0.0370	61.2	45.3	72.2	59.7
0.4	0.9	0.0370	0.0340	3.5	4.8	3.5	4.8
0.4	0.8	0.0270	0.0420	5.8	7.4	6.9	7.9
0.4	0.7	0.0380	0.0390	11.4	12.0	13.8	14.4
0.4	0.6	0.0400	0.0370	26.4	26.9	32.7	32.7
0.4	0.5	0.0460	0.0460	122.5	119.5	131.6	129.8
0.4	0.3	0.0439	0.0458	102.9	112.1	117.3	122.7
0.4	0.2	0.0307	0.0495	19.2	23.6	25.3	29.0
0.4	0.1	0.0278	0.0406	7.7	11.1	8.5	11.7
0.7	0.9	0.0320	0.0560	12.7	19.5	17.2	22.8
0.7	0.8	0.0423	0.0470	78.8	92.2	94.1	103.0
0.7	0.6	0.0481	0.0438	111.3	105.2	122.7	117.3
0.7	0.5	0.0393	0.0400	25.2	23.3	32.2	30.4
0.7	0.4	0.0454	0.0320	12.2	11.6	14.4	13.8
0.7	0.3	0.0330	0.0289	7.2	7.1	7.8	7.8
0.7	0.2	0.0244	0.0208	4.7	5.5	4.5	5.0
0.7	0.1	0.0391	0.0169	3.0	3.5	2.6	3.3
0.01	0.05	0.0578	0.0269	103.6	46.0	107.1	64.2
0.01	0.10	0.0474	0.0229	35.3	14.4	37.2	18.3

TABLE B2

Result for Bernoulli Distribution
(α = 0.05, β = 0.01)

Parameter		Estimate of		Average Sample Number Required			
				Empirical		Theoretical	
p_0	p_1	α_t	β_t	Under H0	Under H1	Under H0	Under H1
0.1	0.9	0.0120	0.0021	3.4	2.3	2.4	1.7
0.1	0.8	0.0278	0.0022	4.6	2.6	3.6	2.1
0.1	0.7	0.0293	0.0017	5.8	3.1	5.3	2.8
0.1	0.6	0.0371	0.0046	7.4	3.8	7.6	3.9
0.1	0.5	0.0682	0.0045	10.4	4.6	11.3	5.7
0.1	0.4	0.0554	0.0029	17.1	7.5	18.5	9.4
0.1	0.3	0.0763	0.0036	31.2	13.3	35.9	18.9
0.1	0.2	0.0564	0.0056	108.3	48.9	113.8	65.5
0.4	0.9	0.0422	0.0073	5.0	5.0	5.6	5.3
0.4	0.8	0.0518	0.0037	9.8	7.8	10.9	8.7
0.4	0.7	0.0531	0.0057	18.4	12.8	21.8	15.8
0.4	0.6	0.0612	0.0061	45.5	27.8	51.5	35.9
0.4	0.5	0.0517	0.0077	208.5	130.8	207.4	142.6
0.4	0.3	0.0536	0.0079	183.6	124.2	185.0	134.7
0.4	0.2	0.0481	0.0047	32.7	25.0	39.9	31.8
0.4	0.1	0.0455	0.0034	11.3	11.3	13.4	12.9
0.7	0.9	0.0511	0.0060	21.6	20.4	27.2	25.0
0.7	0.8	0.0542	0.0083	144.3	100.5	148.3	113.1
0.7	0.6	0.0515	0.0068	192.4	114.7	193.4	128.9
0.7	0.5	0.0667	0.0056	43.3	24.6	50.8	33.4
0.7	0.4	0.0618	0.0061	19.9	11.8	22.7	15.2
0.7	0.3	0.0537	0.0060	11.3	7.2	12.3	8.6
0.7	0.2	0.0318	0.0055	6.9	5.5	7.2	5.4
0.7	0.1	0.0414	0.0052	4.2	3.5	4.0	3.7
0.01	0.05	0.0573	0.0040	170.9	50.9	168.9	70.5
0.01	0.10	0.0665	0.0027	55.4	14.5	58.6	20.1

TABLE B3

Result for Bernoulli Distribution
($\alpha = 0.01$, $\beta = 0.05$)

Parameter		Estimate of		Average Sample Number Required			
				Empirical		Theoretical	
p_0	p_1	α_t	β_t	Under H0	Under H1	Under H0	Under H1
0.1	0.9	0.0001	0.0223	2.2	3.4	1.7	2.4
0.1	0.8	0.0032	0.0724	2.3	3.5	2.5	3.1
0.1	0.7	0.0032	0.0460	3.6	4.3	3.7	4.0
0.1	0.6	0.0057	0.0447	5.0	5.1	5.3	5.6
0.1	0.5	0.0062	0.0330	7.8	7.5	7.9	8.2
0.1	0.4	0.0060	0.0453	11.9	12.3	12.9	13.4
0.1	0.3	0.0068	0.0445	23.7	27.2	25.0	27.2
0.1	0.2	0.0076	0.0522	66.5	83.4	79.3	94.1
0.4	0.9	0.0033	0.0408	3.7	7.6	3.9	7.6
0.4	0.8	0.0041	0.0569	6.2	11.9	7.6	12.5
0.4	0.7	0.0058	0.0495	13.4	22.6	15.2	22.7
0.4	0.6	0.0064	0.0429	36.9	55.5	35.9	51.5
0.4	0.5	0.0079	0.0519	134.3	206.1	144.5	204.6
0.4	0.3	0.0076	0.0541	114.2	192.5	128.9	193.4
0.4	0.2	0.0038	0.0688	20.3	38.9	27.8	45.6
0.4	0.1	0.0041	0.0452	8.3	18.5	9.4	18.5
0.7	0.9	0.0041	0.0797	12.8	31.5	18.9	35.9
0.7	0.8	0.0067	0.0592	85.4	159.6	103.3	162.3
0.7	0.6	0.0084	0.0518	122.1	184.1	134.7	185.0
0.7	0.5	0.0071	0.0461	35.1	50.0	35.4	47.9
0.7	0.4	0.0063	0.0439	14.2	21.2	15.8	21.8
0.7	0.3	0.0038	0.0448	7.4	11.9	8.6	12.3
0.7	0.2	0.0056	0.0363	4.9	7.4	5.0	7.8
0.7	0.1	0.0030	0.0315	3.0	5.8	2.8	5.3
0.01	0.05	0.0066	0.0404	113.0	92.3	117.7	101.2
0.01	0.10	0.0063	0.0296	37.8	23.2	40.8	28.9

TABLE B4

Result for Bernoulli Distribution
($\alpha = \beta = 0.01$)

Parameter		Estimate of		Average Sample Number Required: Empirical		Average Sample Number Required: Theoretical	
p_0	p_1	α_t	β_t	Under H0	Under H1	Under H0	Under H1
0.1	0.9	0.0010	0.0018	3.4	3.5	2.6	2.6
0.1	0.8	0.0052	0.0040	4.7	3.7	3.9	3.3
0.1	0.7	0.0078	0.0037	5.9	4.4	5.7	4.4
0.1	0.6	0.0054	0.0065	7.8	5.8	8.2	6.0
0.1	0.5	0.0081	0.0071	11.4	7.7	12.2	8.8
0.1	0.4	0.0088	0.0054	19.4	12.9	19.9	14.5
0.1	0.3	0.0075	0.0068	34.1	23.1	38.7	29.3
0.1	0.2	0.0086	0.0078	115.8	90.6	122.7	101.4
0.4	0.9	0.0065	0.0068	5.7	7.9	6.0	8.2
0.4	0.8	0.0075	0.0072	10.8	12.5	11.8	13.5
0.4	0.7	0.0074	0.0076	23.0	23.8	23.4	24.5
0.4	0.6	0.0083	0.0085	46.8	46.6	55.5	55.5
0.4	0.5	0.0091	0.0099	224.3	221.1	223.6	220.6
0.4	0.3	0.0087	0.0091	198.3	208.4	199.4	208.5
0.4	0.2	0.0072	0.0091	33.9	41.4	43.0	49.2
0.4	0.1	0.0060	0.0078	13.0	19.4	14.5	19.9
0.7	0.9	0.0074	0.0073	29.2	39.9	29.3	38.7
0.7	0.8	0.0096	0.0088	154.3	172.3	159.9	175.0
0.7	0.6	0.0098	0.0102	206.9	198.2	208.5	199.4
0.7	0.5	0.0088	0.0076	45.7	40.7	54.7	51.7
0.7	0.4	0.0068	0.0081	24.1	22.8	24.5	23.4
0.7	0.3	0.0052	0.0068	12.6	12.5	13.3	13.3
0.7	0.2	0.0074	0.0058	7.3	7.7	7.7	8.4
0.7	0.1	0.0042	0.0052	4.4	6.0	4.4	5.7
0.01	0.05	0.0075	0.0070	184.4	100.4	182.1	109.1
0.01	0.10	0.0089	0.0044	59.8	23.8	63.1	31.2

TABLE B5

Result for Bernoulli Distribution
(α = 0.05, β = 0.10)

Parameter		Estimate of		Average Sample Number Required			
				Empirical		Theoretical	
p_0	p_1	α_t	β_t	Under H0	Under H1	Under H0	Under H1
0.1	0.9	0.0102	0.0159	2.2	2.3	1.1	1.1
0.1	0.8	0.0115	0.0568	2.2	2.5	1.7	1.7
0.1	0.7	0.0224	0.0413	3.4	2.9	2.5	2.3
0.1	0.6	0.0228	0.0861	3.6	3.4	3.6	3.2
0.1	0.5	0.0346	0.0784	5.1	4.3	5.4	4.7
0.1	0.4	0.0285	0.0833	8.3	7.3	8.8	7.6
0.1	0.3	0.0380	0.0873	15.6	14.1	17.1	15.5
0.1	0.2	0.0386	0.0958	56.6	57.4	54.4	53.5
0.4	0.9	0.0260	0.0603	3.1	4.7	2.7	4.3
0.4	0.8	0.0271	0.0780	5.3	7.3	5.2	7.1
0.4	0.7	0.0309	0.0897	9.4	12.3	10.4	12.9
0.4	0.6	0.0306	0.0979	23.3	31.9	24.6	29.3
0.4	0.5	0.0410	0.1016	81.6	105.7	99.0	116.4
0.4	0.3	0.0368	0.1054	70.4	100.0	88.3	110.0
0.4	0.2	0.0322	0.0939	17.1	26.4	19.1	26.0
0.4	0.1	0.0190	0.1258	5.0	9.9	6.4	10.5
0.7	0.9	0.0260	0.1127	10.0	18.0	13.0	20.4
0.7	0.8	0.0340	0.1185	53.5	80.6	70.8	92.3
0.7	0.6	0.0408	0.1058	75.9	92.8	92.3	105.2
0.7	0.5	0.0394	0.0927	22.8	26.8	24.2	27.3
0.7	0.4	0.0382	0.0755	9.6	10.5	10.9	12.4
0.7	0.3	0.0253	0.0908	5.2	6.9	5.9	7.0
0.7	0.2	0.0374	0.0852	3.1	4.2	3.4	4.4
0.7	0.1	0.0310	0.0350	2.7	3.5	1.9	3.0
0.01	0.05	0.0571	0.0620	74.1	41.6	80.6	57.5
0.01	0.10	0.0394	0.0535	26.7	13.6	28.0	16.4

TABLE B6

Result for Normal Distribution
($\alpha = \beta = 0.05$)

Parameter		Estimate of		Average Sample Number Required			
				Empirical		Theoretical	
μ_0	μ_1	α_t	β_t	Under H0	Under H1	Under H0	Under H1
0	0.2	0.0447	0.0455	138.1	137.9	132.5	132.5
0	0.3	0.0423	0.0412	56.7	56.5	58.9	58.9
0	0.4	0.0376	0.0392	28.5	28.7	33.1	33.1
0	0.5	0.0334	0.0415	17.6	17.5	21.2	21.2
0	0.6	0.0352	0.0352	12.5	12.5	14.7	14.7
0	0.7	0.0337	0.0320	9.4	9.2	10.8	10.8
0	0.9	0.0317	0.0334	6.0	6.0	6.5	6.5
0	1	0.0311	0.0318	5.0	5.0	5.3	5.3
0	1.4	0.0238	0.0235	3.0	3.0	2.7	2.7
0	2	0.0142	0.0187	1.8	1.8	1.3	1.3
0	3	0.0068	0.0061	1.3	1.3	0.6	0.6

TABLE B7

Result for Normal Distribution
($\alpha = 0.05$, $\beta = 0.01$)

Parameter		Estimate of		Average Sample Number Required: Empirical		Average Sample Number Required: Theoretical	
μ_0	μ_1	α_t	β_t	Under H0	Under H1	Under H0	Under H1
0	0.2	0.0465	0.0085	216.3	149.7	208.8	145.5
0	0.3	0.0452	0.0071	95.4	61.5	92.8	64.7
0	0.4	0.0512	0.0053	50.8	30.8	52.2	36.4
0	0.5	0.0517	0.0050	30.5	18.6	33.4	23.3
0	0.6	0.0542	0.0046	20.2	13.1	23.2	16.2
0	0.7	0.0472	0.0052	15.0	9.7	17.0	11.9
0	0.9	0.0444	0.0033	9.5	6.3	10.3	7.2
0	1	0.0383	0.0042	7.9	5.2	8.4	5.8
0	1.4	0.0334	0.0033	4.4	3.1	4.3	3.0
0	2	0.0241	0.0021	2.6	1.9	2.1	1.5
0	3	0.0076	0.0007	1.5	1.3	0.9	0.6

TABLE B8

Result for Normal Distribution
($\alpha = 0.01$, $\beta = 0.05$)

Parameter		Estimate of		Average Sample Number Required			
				Empirical		Theoretical	
μ_0	μ_1	α_t	β_t	Under H0	Under H1	Under H0	Under H1
0	0.2	0.0080	0.0467	149.5	216.2	145.5	208.8
0	0.3	0.0067	0.0436	61.2	95.4	64.7	92.8
0	0.4	0.0059	0.0530	30.8	50.7	36.4	52.2
0	0.5	0.0052	0.0519	18.8	30.4	23.3	33.4
0	0.6	0.0039	0.0536	12.9	20.7	16.2	23.2
0	0.7	0.0043	0.0518	9.7	15.3	11.9	17.0
0	0.9	0.0036	0.0459	6.2	9.5	7.2	10.3
0	1	0.0038	0.0423	5.2	7.9	5.8	8.4
0	1.4	0.0033	0.0358	3.1	4.3	3.0	4.3
0	2	0.0029	0.0227	1.9	2.5	1.5	2.1
0	3	0.0036	0.0137	1.3	1.5	0.6	0.9

TABLE B9

Result for Normal Distribution
($\alpha = \beta = 0.01$)

Parameter		Estimate of		Average Sample Number Required			
				Empirical		Theoretical	
μ_0	μ_1	α_t	β_t	Under H0	Under H1	Under H0	Under H1
0	0.2	0.0085	0.0090	231.4	231.0	225.2	225.2
0	0.3	0.0085	0.0081	102.0	102.1	100.1	100.1
0	0.4	0.0078	0.0082	54.3	54.3	56.3	56.3
0	0.5	0.0073	0.0067	32.2	32.1	36.0	36.0
0	0.6	0.0069	0.0063	21.8	21.7	25.0	25.0
0	0.7	0.0063	0.0079	15.8	16.1	18.4	18.4
0	0.9	0.0072	0.0052	9.8	9.9	11.1	11.1
0	1	0.0057	0.0073	8.2	8.2	9.0	9.0
0	1.4	0.0047	0.0026	4.5	4.5	4.6	4.6
0	2	0.0044	0.0035	2.6	2.6	2.3	2.3
0	3	0.0012	0.0011	1.5	1.5	1.0	1.0

TABLE B10

Result for Normal Distribution
($\alpha = 0.05$, $\beta = 0.10$)

Parameter		Estimate of		Average Sample Number Required			
				Empirical		Theoretical	
μ_0	μ_1	α_t	β_t	Under H0	Under H1	Under H0	Under H1
0	0.2	0.0421	0.0909	101.6	123.6	99.7	118.8
0	0.3	0.0367	0.0964	38.6	50.7	44.3	52.8
0	0.4	0.0350	0.0991	19.9	25.9	24.9	29.7
0	0.5	0.0319	0.0949	12.8	16.3	15.9	19.0
0	0.6	0.0303	0.0912	9.1	11.4	11.1	13.2
0	0.7	0.0267	0.0775	7.0	8.7	8.1	9.7
0	0.9	0.0270	0.0638	5.9	7.3	4.9	5.9
0	1	0.0276	0.0636	4.8	5.9	4.0	4.8
0	1.4	0.0224	0.0520	2.7	3.2	2.0	2.4
0	2	0.0181	0.0347	1.7	1.9	1.0	1.2
0	3	0.0084	0.0163	1.2	1.3	0.4	0.5

REFERENCES

Anderson,T.W (1960). Modification of the Sequential Probability Ratio Test to Reduce the Sample Size. Annals of Mathematical Statistics 31, 165-197.

Baker, A.G. (1950). Properties of Some Tests in Sequential Analysis. Biometrika 37, 334-346.

Besag, J. & Clifford, P. (1991). Sequential Monte Carlo P-Values. Biometrika 78, 301-304.

Burr,I.W. (1979). Elementary Statistical Quality Control. New York and Basel: Marcel Dekker, Inc.

Darling, D.A. (1976). The Birth, Growth and Blossoming of Sequential Analysis. In On the History of Statistics and Probability .ed. D.B. Owen. New York: Marcel Dekker.

Dixon, W.J. and Massey, Jr.F.J. (1983). Introduction to Statistical Analysis. International Student Edition: McGraw-Hill.

Edwards, D. (1985). Sequential Tests, Exact Simulation-Based Inference: A Survey, with Additions. Journal of Statistical Computation and Simulation 22, 307-326.

Edwards, D. and Hamson, M. (1989). A Guide to Mathematical Modelling. Macmillan.

Fisz, M. (1963). Probability Theory and Mathematical Statistics. John Wiley & Sons, Inc.

Gebre-Egziabher(1992). Computer Methods for Sampling from Univariate Continuous Distributions. An Ethiopian Journal of Science (SINET) 15,ISS 2, 99 - 115.

Govindarajulu, Z. (1985). Recent Developments in Sequential Analysis: Testing Hypothesis. Mathematical Scientist 10, 51-64.

Grant,E.L and Leavenworth,R.S (1988). Statistical Quality Control. McGraw-Hill Book Company.

Kennedy,Jr.W.J. and Gentle,J.E. (1980). Statistical Computing. New York and Basel: Marcel Dekker, Inc.

Lindgren, B.W. (1968), Statistical Theory. New York: Macmillan Publishing Co.,Inc.

Lotov, V.I. (1988). On the Accuracy of Approximation in the Sequential Wald Test. Translated by Ellis, M. Theory of Probability and its applications 33, ISS 2, 276-285.

Matloff,N.S (1988). Probability Modelling and Computer Simulation. Boston: PWS-KENT Publishing company.

Morgan, B.J.T. (1984). Elements of Simulation. New York: Chapman and Hall.

Schneiderman, M.A. and Armitage, P. (1962). A Family of Closed Sequential procedures. Biometrika 49, 41-56.

Siegmund, D. (1974). Error Probabilities and Average Sample Number of the Sequential Probability Ratio Test. Journal of the Royal Statistical Society (Series B) 37, 394-401.

Siegmund, D. (1985). Sequential Analysis. New York: Springer-Verlag.

Tremblay, J.P & Dedourek, J.M. (1989). Programming in Pascal. Toronto: McGraw-Hill.

Wald, A. (1947). Sequential Analysis. New York: Dover Publications Inc.

Wallis, A.W. (1980). The Statistical Research Group. Journal of American Statistical Association. 75, 320-333.

Wetherill, G.B. (1966). Sequential Methods in Statistics. London: Chapman and Hall.

Zaks, R. (1988). Introduction to Pascal. London: Sybex, Inc.