On SPRT and RSPRT for the Unknown Mean in a Normal Distribution Whose Variance Is a Multiple of the Mean

Nitis Mukhopadhyay[1] and Debanjan Bhattacharjee[1]

[1] Department of Statistics, University of Connecticut,
215 Glenbrook Road, U-4120
Storrs, CT 06269-4120, USA
nitis.mukhopadhyay@uconn.edu
debanjan.bhattacharjee@gmail.com

Abstract. First, we consider a sequence of independent observations from a $N(\theta, c\theta)$ distribution with $-\infty < x < \infty, 0 < \theta, c < \infty$. We assume that θ remains unknown but c is known and the problem is to test H_0: $\theta = \theta_0$ vs. H_1: $\theta = \theta_1$ where $\theta_0, \theta_1 (\theta_0 \neq \theta_1)$ are specified positive numbers with target type-I and type-II error probabilities $0 < \alpha < 1$ and $0 < \beta < 1$ respectively, $\alpha + \beta < 1$. It may be tempting to adapt Wald's *sequential probability ratio test* (SPRT) available in the case of a $N(\theta, \sigma^2)$ population with σ^2 known, but there is a fundamental difference between that scenario and the present scenario. Given a fixed number of independent observations $X_1, ..., X_n$ from $N(\theta, \sigma^2)$, $\sum_{i=1}^n X_i$ is complete and sufficient for θ. But, given $X_1, ..., X_n$ from $N(\theta, c\theta)$, $\sum_{i=1}^n X_i^2$ is complete and sufficient for θ and $\sum_{i=1}^n X_i^2$ has a non-central χ_n^2 distribution with its non-centrality parameter involving θ. We begin with the SPRT methodology for the $N(\theta, c\theta)$ distribution and describe some crucial characteristics.

In some situations, however, observations may become available in random group sizes arriving sequentially. Hence, Mukhopadhyay and de Silva (2008) developed the notion of a *random* SPRT (RSPRT) along with its general theory. In the context of a $N(\theta, c\theta)$ distribution, we explore the role of RSPRT. During the presentation, we will emphasize statistical data analysis for both the SPRT and RSPRT.

Keywords. ASN, non-central chi-square, OC, random group-size sampling, truncated SPRT, UMP test.

1 Introduction

Let us consider a sequence of independent observations $X_1, X_2, ...$ from a $N(\theta, c\theta)$ distribution, that is, having a common *probability density function* (p.d.f.) given by

$$f(x; \theta) = (2\pi c\theta)^{-1/2} \exp\left\{-\frac{1}{2c\theta}(x-\theta)^2\right\}, \tag{1}$$

with $-\infty < x < \infty, 0 < \theta, c < \infty$. Here, we assume that θ remains unknown but c is known. We can rewrite (1) as follows:

$$f(x; \theta) = (2\pi c\theta)^{-1/2} \exp\left\{-\frac{1}{2c\theta}x^2 + \frac{1}{c}x - \frac{1}{2c}\theta\right\}, \tag{2}$$

which shows that the p.d.f. $f(x; \theta)$ belongs to a one-parameter exponential family.

Since $c(> 0)$ is assumed known, we may equivalently work with the Y-data, namely, $Y_1, Y_2, ...$ from a $N(\theta, \theta)$ distribution instead of the original X-data $X_1, X_2, ...$ from a $N(\theta, c\theta)$ distribution by letting $Y = c^{-1}X$. Hence, without any loss of generality, we assume having independently distributed observations $X_1, X_2, ...$ with a common p.d.f. $f(x; \theta)$ from (1) where $c = 1$.

The problem is one of testing H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ where $\theta_0, \theta_1 (\theta_0 \neq \theta_1)$ are specified positive numbers with target type-I and type-II error probabilities respectively given by two prespecifed numbers $0 < \alpha < 1$ and $0 < \beta < 1$ respectively, $\alpha + \beta < 1$. One may feel tempted to adapt Wald's (1947) *sequential probability ratio test* (SPRT) for H_0: $\theta = \theta_0$ vs. H_1: $\theta = \theta_1$ available in the case of a $N(\theta, \sigma^2)$ population with σ^2 known. One may refer to Mukhopadhyay and de Silva (2009, Chapter 3) and other sources for reviewing SPRT methodologies.

However, there is a fundamental difference between that $N(\theta, \sigma^2)$ scenario and the present scenario. If we have a fixed number of independent observations $X_1, \ldots, X_n$ from a $N(\theta, \sigma^2)$ population, then $\sum_{i=1}^n X_i$ is a complete and sufficient statistic for θ. Instead, if we have independent observations $X_1, \ldots, X_n$ from a $N(\theta, \theta)$ population, then $\sum_{i=1}^n X_i^2$ turns out to be a complete and sufficient statistic for θ.

Observe that $\sum_{i=1}^n X_i^2$ has a non-central χ_n^2 distribution for each fixed n with its non-centrality parameter involving the true value of θ. This creates some unexpected complications in implementing the Wald's SPRT found in the case of $N(\theta, \sigma^2)$ with $0 < \sigma^2 < \infty$ known. In Section 2, we describe the SPRT in the present situation and some of its associated characteristics.

In some situations, however, Wald's SPRT may not be implementable because observations become available in random group sizes arriving sequentially. In such circumstances, Mukhopadhyay and de Silva (2008) developed the notion and practicality of a *random SPRT* (RSPRT) along with its general theory. In a $N(\theta, \theta)$ distribution, we explore the role of RSPRT. Suppose that we observe data from $\{M_i, X_{i1}, \ldots, X_{iM_i}, i = 1, 2, \ldots\}$ sequentially. Here, we assume that the M_i's are random group sizes, and the identical distribution of $M_1, M_2, \ldots$ remains independent of the X's. One may view this situation as a hierarchical one: At each step i, we must record M_i observations, $i = 1, 2, \ldots$, but the M-process itself is generated by the nature independently of the X's. Section 3 briefly addresses the newly developed RSPRT in the context of a $N(\theta, \theta)$ distribution.

2 Wald's SPRT

Let us denote $\mathbf{X}_j = (X_1, \ldots, X_j), j = 1, 2, \ldots$. Given θ, the likelihood function for $X_1, \ldots, X_j$ is given by

$$L(\theta; j, \mathbf{X}_j) = \prod_{i=1}^{j} f(X_i; \theta) = (2\pi\theta)^{-j/2} \exp\left\{-\frac{1}{2\theta} \sum_{i=1}^{j} (X_i - \theta)^2\right\}, \tag{3}$$

$j = 1, 2, \ldots$. We wish to test H_0: $\theta = \theta_0$ against an alternative hypothesis H_1: $\theta = \theta_1$ where $\theta_0, \theta_1 (\theta_0 \neq \theta_1)$ are specified positive numbers with our target type-I and type-II error probabilities α, β respectively with $\alpha + \beta < 1$. Let us denote $L_0(j) = L(\theta_0; j, X_j)$ and $L_1(j) = L(\theta_1; j, X_j)$ respectively under H_0 and H_1. Clearly, we have:

$$Z_i \equiv \ln \frac{f(X_i; \theta_1)}{f(X_i; \theta_0)} = \frac{1}{2} \ln\left(\frac{\theta_0}{\theta_1}\right) + \frac{(\theta_1 - \theta_0)}{2\theta_0\theta_1} X_i^2 - \frac{(\theta_1 - \theta_0)}{2}, \tag{4}$$

which are independent and identically distributed observations with

$$E_\theta(Z_1) = \tfrac{1}{2} \ln\left(\frac{\theta_0}{\theta_1}\right) + \frac{(\theta_1 - \theta_0)}{2\theta_0\theta_1} (\theta + \theta^2) - \frac{(\theta_1 - \theta_0)}{2}, \tag{5}$$

$i = 1, 2, \ldots$.

In order to attain the desired levels for both error probabilities, we implement Wald's (1947) SPRT and successively consider the likelihood ratios given by

$$\Lambda_j \equiv \Lambda_j(\mathbf{X}_j) = \frac{L_1(j)}{L_0(j)}, j = 1, 2, \ldots . \tag{6}$$

We fix:

$$A = (1 - \beta)/\alpha,\ B = \beta/(1 - \alpha),\ a = \log A,\ b = \log B, \tag{7}$$

along the lines of Wald (1947). One may refer to Mukhopadhyay and de Silva (2009, Chapter 3) and other sources for reviewing SPRT methodologies.

The customary SPRT is then implemented as follows. We successively observe $\mathbf{X}_j$ thereby obtaining $\Lambda_j(\mathbf{X}_j)$ from (6) with $j = 1, 2, ...$ and let the stopping time be defined as follows:

$$\begin{aligned} N \equiv & \text{ first integer } (j \geq 1) \text{ for which } \Lambda_j(\mathbf{X}_j) \notin (B, A) \text{ that is when} \\ & \textstyle\sum_{i=1}^{j} Z_i \leq b \text{ or } \sum_{i=1}^{j} Z_i \geq a \text{ occurs for the first time. And we} \\ & \text{decide in favour of } H_0(\text{or } H_1) \text{ if } \textstyle\sum_{i=1}^{N} Z_i \leq b \text{ (or } \sum_{i=1}^{j} Z_i \geq a \text{),} \end{aligned} \tag{8}$$

with a, b determined by (7) where

$$\textstyle\sum_{i=1}^{N} Z_i = \frac{N}{2} \ln\left(\frac{\theta_0}{\theta_1}\right) + \frac{(\theta_1 - \theta_0)}{2\theta_0\theta_1} \sum_{i=1}^{N} X_i^2 - \frac{N(\theta_1 - \theta_0)}{2}.$$

The SPRT defined in (8) would clearly terminate w.p.1, that is $P_\theta\{N < \infty\} = 1$, whatever be the true θ value. Indeed, the *moment generating function* (m.g.f.) of N is finite whatever be the true θ value. This implies that both mean and the variance of the stopping variable N from (8) are finite whatever be the true θ value. These follow from Stein (1946). Also refer to Mukhopadhyay and de Silva (2009, Chapter 3).

2.1 Approximations for the OC and ASN Functions

To come up with the approximations for the *operating characteristic* (OC) function and the *average sample number* (ASN) function, first we need the m.g.f. of Z_1. We express $E_\theta[\exp(tZ_1)]$ as follows:

$$\exp\left(\tfrac{1}{2}t \ln\left(\frac{\theta_0}{\theta_1}\right) - \tfrac{1}{2}t(\theta_1 - \theta_0)\right)\left(1 - \frac{t(\theta_1 - \theta_0)\theta}{\theta_0\theta_1}\right)^{-1/2} \exp\left(\frac{t(\theta_1 - \theta_0)\theta^2}{2[\theta_0\theta_1 - t(\theta_1 - \theta_0)\theta]}\right), \tag{9}$$

since

$$E_\theta\left[\exp\left(tX_1^2\right)\right] = \frac{1}{\sqrt{2\pi\theta}} \textstyle\int_{-\infty}^{\infty} \exp\left(tx^2 - \frac{1}{2\theta}(x - \theta)^2\right) = (1 - 2\theta t)^{-1/2} \exp\left(\frac{\theta^2 t}{(1 - 2\theta t)^{1/2}}\right),$$

for $t < (2\theta)^{-1}$. Now, we want to locate $t = t_0 \equiv t_0(\theta) \neq 0$ such that $E_\theta\left[\exp(t_0(\theta)Z_1)\right] = 1$.

An equivalent expression for $t_0 \equiv t_0(\theta)$ is the solution $t \equiv t_0(\theta)$ satisfying the following implicit equation involving t and $\theta, \theta_0, \theta_1$:

$$t \ln\frac{\theta_0}{\theta_1} - t(\theta_1 - \theta_0) + \frac{\theta^2 t(\theta_1 - \theta_0)}{\theta_0\theta_1 - t(\theta_1 - \theta_0)\theta} = \ln\left(1 - \frac{t(\theta_1 - \theta_0)\theta}{\theta_0\theta_1}\right). \tag{10}$$

Clearly, there is no closed form expression for $t_0(\theta)$ in terms of θ. Hence, finding such $t_0(\theta)$ for some given value of θ is a difficult computational problem. We will present some numerical values of $t_0(\theta)$ obtained by using an appropriate computing search engine for selected choices of $\theta, \theta_0, \theta_1$ at the conference.

The OC function is defined as $L(\theta) = P_\theta\{$ Accept the null hypothesis $H_0\}$. Its approximation is given by

$$L(\theta) \approx \frac{1 - \exp(t_0 a)}{\exp(t_0 b) - \exp(t_0 a)} \text{ where } t_0 \equiv t_0(\theta) \neq 0 \text{ such that} E_\theta\left\{\exp\left(t_0 Z_1\right)\right\} = 1. \tag{11}$$

This expression is obtained from Wald's fundamental identity which can be found in Wald (1947), Ghosh et al. (1997, p. 35), Mukhopadhyay and de Silva (2009, p.51) and other sources. The recent discussion paper of Lai (2004) is filled with an incredible breadth of information on Wald-type identities and their numerous applications. At the conference, we will present some approximate numerical values of $L(\theta)$ or the approximate OC curve itself obtained by using the expression shown in (11) for selected choices of $\theta, \theta_0, \theta_1$ and α, β.

The ASN function is $E_\theta\{N\}$ associated with the stopping rule (8) and its approximation works out as follows:

$$E_\theta\{N\} \approx \begin{cases} \dfrac{bL(\theta) + a(1 - L(\theta))}{E_\theta(Z_1)} & \text{if } E_\theta(Z_1) \neq 0, \text{ that is } \theta \neq \theta^* \\ -\dfrac{ab}{\theta} & \text{if } E_\theta(Z_1) = 0, \text{ that is } \theta = \theta^* \end{cases} \tag{12}$$

where

$$\theta^* = -\frac{1}{2} + \sqrt{\frac{1}{4} + \theta_0\theta_1 + \theta_0\theta_1\frac{\ln(\theta_1) - \ln(\theta_0)}{\theta_1 - \theta_0}}. \quad (13)$$

Again, at the conference, we will present some approximate numerical values of $E_\theta\{N\}$ or the approximate ASN curve itself obtained by using the expression shown in (12) for selected choices of $\theta, \theta_0, \theta_1$ and α, β.

2.2 Truncated SPRT

Clearly, the SPRT (8) is open-ended. That is, even though the stopping rule (8) will terminate w.p.1, there is a distinct possibility that the stopping time would require more observations than one may be able to afford. In these situations, we propose to implement a truncated version of the SPRT (8). Suppose that one is not willing to go beyond n_0 observations where n_0 is determined by an experimenter after taking into account available resources including the cost, the waiting time to gather data, and relevant logistics.

A truncated SPRT would then work as follows: Implement the SPRT (8) and proceed by recording an additional observation at each step as needed. Now, if sampling stops naturally for the first time at step 1 or at step 2, or ... or at step n_0 followed by either acceptance of H_0 or H_1, then that stopped value of N and the associated decision are both noted. On the other hand, sampling may continue from step 1 to step 2, ... reaching all the way up to and including step n_0, but one may still need to continue with more observation at the step n_0 according to the stopping rule (8). Now, that may not be allowed considering non-affordability. Once the stopping time reaches the step n_0 but does not stop here, we must intervene to stop sampling right away.

Along the line of Wald (1947), we define the stopping time associated with a *truncated* SPRT as follows:

Let $T = \min\{N, n_0\}$. If $T = N$, we decide in favor of H_0 (or H_1) if $\sum_{i=1}^N Z_i \le b$ (or $\sum_{i=1}^N Z_i \ge a$). If $T = n_0$, we decide in favor of H_0 (or H_1) if $b < \sum_{i=1}^{n_0} Z_i \le 0$ (or $0 < \sum_{i=1}^{n_0} Z_i < a$). (14)

2.2.1 Three Ideasof Where to Truncate the SPRT

In view of the optimal character of Wald's SPRT, its expected sample size will beat the sample size required by the *uniformly most powerful* (UMP) test for H_0 vs. H_1 with comparable α, β under both H_0, H_1. See, Wald and Wolfowitz (1948). So, first, let us look at the UMP test for H_0 vs. H_1 with error probabilities α, β.

For a fixed sample size n, observe that the statistic $T_n \equiv \Sigma_{i=1}^n X_i^2$ has a non-central chi-square distribution with the degree of freedom n and its non-centrality parameter, $\lambda = \frac{1}{2}n\theta^2$ when θ is the true value. We write $\Sigma_{i=1}^n X_i^2 \sim \chi_n^2[\lambda]$ when θ is the true parameter value.

Clearly, under H_i, the statistic $\Sigma_{i=1}^n X_i^2 \sim \chi_n^2[\lambda_i]$ where $\lambda_i = \frac{1}{2}n\theta_i^2$, $i = 0, 1$. Without any loss of generality, let us assume that $\theta_1 > \theta_0$. Then, for a fixed sample size n, the level α UMP test will reject H_0, that is it will decide in favor of H_1, if and only if

$$\Sigma_{i=1}^n X_i^2 > \chi_{n,\alpha}^2[\lambda_0] \text{ where } \chi_{n,\alpha}^2[\lambda_0] \text{ is the upper } 100\alpha\% \text{ point of } \chi_{n,\alpha}^2[\lambda_0] \text{ distribution.} \quad (15)$$

Now, the probability of type-II error probability associated with the level α UMP test (15) will not exceed β provided that

$$P_{\theta_1}\left\{\Sigma_{i=1}^n X_i^2 < \chi_{n,\alpha}^2[\lambda_0]\right\} \le \beta \Leftrightarrow n = n_0 \text{ is the smallest positive integer such that } \chi_{n,\alpha}^2[\lambda_0] \le \chi_{n,1-\beta}^2[\lambda_1]. \quad (16)$$

Idea #1: Such n_0, if readily available for the fixed values of θ_0, θ_1 and α, β under consideration, then we will implement the truncated SPRT from (14) with this n_0. But, since tables for $\chi_{n,\gamma}^2[\lambda]$ corresponding to a non-central chi-square distribution may not be readily available for a wide range of fixed values of λ, n, γ corresponding to θ_0, θ_1 and α, β under consideration, we also present two large-sample methods to determine other reasonable choices for the truncation point n_0.

Idea #2: For a fixed but large sample size n, by the central limit theorem, one has: With $\mu \equiv \mu(\theta) = \theta + \theta^2$ and $\sigma^2 \equiv \sigma^2(\theta) = 2\theta^2(1+2\theta)$,

$$\frac{\sqrt{n}(T_n - \mu(\theta))}{\sigma(\theta)} \xrightarrow{\pounds} N(0,1) \text{ as } n \to \infty \text{ for true } \theta. \tag{17}$$

Hence, combining (15) with (17), we may express a large-sample approximate level α UMP test H_0 vs. H_1 as follows: We will reject H_0, that is we will decide in favor of H_1, if and only if

$$\frac{\sqrt{n}(T_n - \mu(\theta_0))}{\sigma(\theta_0)} > z_\alpha \text{ where } z_\alpha \text{ is the upper } 100\alpha\% \text{ point of a } N(0,1) \text{ distribution.} \tag{18}$$

Then, the probability of type-II error associated with the approximate level α UMP test (18) will not exceed β provided that we have:

$$n = n_0 \text{ is the smallest positive integer} \geq \frac{2\theta_1^2(1+2\theta_1)\left[z_\alpha \frac{\theta_0}{\theta_1}\sqrt{\frac{1+2\theta_0}{1+2\theta_1}} + z_\beta\right]^2}{(\theta_1-\theta_0)^2(1+\theta_0+\theta_1)^2}. \tag{19}$$

Then, we may implement the truncated SPRT from (14) with n_0 defined by (19).

Idea #3: From (17), since the variance of the asymptotic distribution of $\sqrt{n}(T_n - \mu(\theta))$ involves θ as seen from the expression of $\sigma^2(\theta)$, it may be tempting to first come up with the associated variance stabilizing transformation. Now, for a non-zero differentiable real valued function $g(.)$, the variance of the asymptotic distribution of $\sqrt{n}\,(g(T_n) - g(\mu(\theta)))$ would be given by the expression of $g^{'2}(\theta)\sigma^2(\theta)$. We want to determine the function $g(.)$ such that $g^{'2}(\theta)\sigma^2(\theta)$ becomes free from θ. That is, we must have:

$$g(\theta) = \int g^{'}(\theta)d\theta = \int \left\{2\theta^2(1+2\theta)\right\}^{-1/2} d\theta. \tag{20}$$

In (19), we make the following substitution: $\theta = \frac{1}{2}\tan^2(\gamma)$ so that we obtain:

$$g(\theta) = \sqrt{2}\int \text{cosec}(\gamma)d\gamma = \sqrt{2}\log\left(\left[\sin(\tan^{-1}(\sqrt{2\theta})\right]^{-1} - [2\theta]^{-1/2}\right). \tag{21}$$

Now, in view of (21), we can claim that $\sqrt{n}\,(g(T_n) - g(\mu(\theta))) \xrightarrow{\pounds} N(0,1)$ as $n \to \infty$.

At this point, the large-sample approximate level α UMP test (18) for H_0 vs. H_1 may be restated as follows: We will reject H_0, that is we will decide in favor of H_1, if and only if

$$\sqrt{n}\,(g(T_n) - g(\mu(\theta_0))) > z_\alpha \text{ where } z_\alpha \text{ is the upper } 100\alpha\% \text{ point of a } N(0,1) \text{ distribution.} \tag{22}$$

The probability of type-II error associated with the approximate level α UMP test (22) will not exceed β provided that

$$\begin{aligned} & n = n_0 \text{ is the smallest positive integer} \geq \left[\frac{z_\alpha + z_\beta}{\mu_1 - \mu_0}\right]^2 \text{ where } \mu_i \\ &= g(\mu(\theta_i)) = \sqrt{2}\log\left(\left[\sin(\tan^{-1}(\sqrt{2(\theta_i+\theta_i^2)})\right] - \left(2(\theta_i+\theta_i^2)\right)^{-1/2}\right), i = 0,1. \end{aligned} \tag{23}$$

Then, we may implement the truncated SPRT from (14) with n_0 defined by (23).

2.3 Performances and Comparisons

During presentation, we will highlight characteristics of the SPRT (8) and the truncated SPRT (14). We also hope to examine how the truncated SPRT (14) compares with each other when n_0 is determined from (16), (19), and (23) respectively.

3 Mukhopadhyay and de Silva's RSPRT

We consider the following independent observations

$$\{M_i, X_{i1}, X_{i2}, \ldots, X_{iM_i}; i = 1, 2, \ldots, k, \ldots\}$$

where we assume that (A1) the M_i's constitute a sequence of independent observations with its *probability mass function* (p.m.f.)

$$g_i(m_i) = \mathcal{P}(M_i = m_i), m_i = 0, 1, 2, \ldots, i = 1, 2, \ldots,$$

and (A2) the X's constitute a sequence of conditionally i.i.d. observations, given the M's, with the common p.d.f. $f(x; \theta)$ from (1). One immediately notes that we do not rule out the possibility that in some time-intervals, there may be no observations available,

In general, let us denote $\mathbf{M}^{(k)} = (M_1, \ldots, M_k)$ for the associated data on the number of arriving data in k consecutive time-intervals where k may have been fixed in advance. Then, the likelihood function of the observations $\{M_i, X_{i1}, X_{i2}, \ldots, X_{iM_i};\ i = 1, 2, \ldots, k\}$ is expressed as

$$Q_{k,\mathbf{M}^{(k)}}(\theta) = \Pi_{i=1}^{k} g_i(M_i) \Pi_{j=1}^{M_i} f(X_{ij}; \theta).$$

For the same testing problem of deciding in favor of H_0 or H_1 under consideration, Mukhopadhyay and de Silva (2008) proposed the following RSPRT: A likelihood ratio that falls under $B\,(= \beta/(1-\alpha))$ or above $A\,(= (1-\beta)/\alpha)$ would be deemed "too small" or "too large" respectively. We start with the first group of observations $\{X_{11}, \ldots, X_{1M_1}\}$ and continue taking one additional group of observations at-a-time as long as

$$B < \Lambda_{l,\mathbf{M}^{(l)}} \equiv \left[\Pi_{i=1}^{l} \Pi_{j=1}^{M_i} f(X_{ij}; \theta_1)\right] \left[\Pi_{i=1}^{l} \Pi_{j=1}^{M_i} f(X_{ij}; \theta_0)\right]^{-1} < A, l = 1, 2, \ldots. \tag{24}$$

We stop observing further additional groups of observations as soon as the likelihood ratio at some stage becomes too small ($\leq B$) or too large ($\geq A$). In other words, we continually watch for the likelihood ratio to go out of the interval (B, A) for the first time. At termination, if the likelihood ratio appears too small (too large), then we decide in favor of the hypothesis $H_0 (H_1)$. The stopping time is defined as follows: Stop sampling with K groups of observations where K is the first integer $k(\geq 1)$ such that $\Lambda_{k,\mathbf{M}^{(k)}} \notin (B, A)$, or equivalently, let

$$\begin{array}{c} K = \inf\left\{k \geq 1 : \Sigma_{i=1}^{k} S_{i,M_i} \notin (b, a)\right\} \text{ whereas we accept } H_0 (\text{or } H_1) \text{ if} \\ \text{and only if } \ \Sigma_{i=1}^{K} S_{i,M_i} \leq b \left(\text{or } \Sigma_{i=1}^{K} S_{i,M_i} \geq a\right) \ \text{ when sampling terminates.} \end{array} \tag{25}$$

where

$$Z_{ij} = \ln\left[f(X_{ij}; \theta_1)\right] - \ln\left[f(X_{ij}; \theta_0)\right], a = \ln(A), b = \ln(B), \text{ and } S_{i,M_i} = \Sigma_{j=1}^{M_i} Z_{ij}, i = 1, 2, \ldots\ .$$

Mukhopadhyay and de Silva (2008) developed the theory of such sequential tests in general. They proved, among other properties, that the RSPRT (25) terminates w. p.1. During presentation, we will highlight characteristics of the RSPRT (25) in the context of a $N(\theta, \theta)$ population and compare them with those of the SPRT described in (8).

References

Ghosh, M., Mukhopadhyay, N., and Sen, P. K. (1997). *Sequential Estimation*. Wiley, New York.

Lai, T. L. (2004). Likelihood Ratio Identities and Their Applications to Sequential Analysis (with discussions). *Sequential Analysis*, **23**, 467-497.

Mukhopadhyay, N. and de Silva, B. M. (2008)). Theory and Applications of a New Methodology for the Random Sequential Probability Ratio Test. *Statistical Methodology,* **5**, 424-453.

Mukhopadhyay, N. and de Silva, B. M. (2009). *Sequential Methodologies and Their Applications*. CRC/Chapman & Hall, Boca Raton.

Stein, C. (1946). A Note on Cumulative Sums. *Annals of Mathematical Statistics*, **17**, 498-499.

Wald, A. (1947). *Sequential Analysis*. Wiley, New York.

Wald, A. and Wolfowitz, J. (1948). Optimum Character of the Sequential Probability Ratio Test. *Annals of Mathematical Statistics*, **19**, 326-339.