Sequential Analysis - Practical Problems
Wald's Sequential Probability Ratio Test

Somjit Roy*

Course Instructor: Prof. Rahul Bhattacharya

MSc. First Year
Department of Statistics
University of Calcutta

April 10, 2022

Contents

*somjit.roy2001@gmail.com

(All the computations have been done in `R`, refer to https://github.com/Roy-SR-007/Sequential-Analysis-STAT-203 for the detailed `R-Script: Sequential Practical Problem.R`)

1 Problem 1

Let θ be the probability of an item being *defective* and observations are drawn from a process sequentially.

1. Determine the *continuation region* of *SPRT* at the nth stage for testing $\mathcal{H}_0 : \theta = 0.5$ against $\mathcal{H}_1 : \theta = 0.8$ taking $\alpha = 0.2 = \beta$.
2. If the sequence of items are observed as DNNDDDDDD (N: non-defective & D: defective), what will be your decision on the ninth observation?
3. Calculate *Wald's Approximate OC & ASN* functions for different values of θ.

1.1 Solution to Problem 1

$\theta \Rightarrow$ *probability of an item being defective*. To determine the *Wald's SPRT* and its *continuation region* at the nth stage by taking $\alpha = 0.2 = \beta$, i.e., *strength* $= (\alpha, \beta) = (0.2, 0.2)$.

Define $X_n :=$ *denoting whether the* n*th item is defective or not*, $n \geq 1$. Clearly, $X_n \overset{iid}{\sim} \text{Ber}(\theta)$.

1.1.1 Testing Problem:

We are to test,

$$\mathcal{H}_0 : \theta = \theta_0 = 0.5 \text{ against } \mathcal{H}_1 : \theta = \theta_1 = 0.8 \tag{1.1.1.1}$$

1.1.2 To obtain the WALD'S SPRT for 1.1.1.1

The PMF of X_i's is given by,

$$p_\theta(x_i) = \theta^{x_i}(1-\theta)^{1-x_i}, x_i = 0, 1 \;\forall\; i, \theta \in (0,1)$$

We have,

$$Z_1 = \log\frac{p_{\theta_1}(x_1)}{p_{\theta_0}(x_1)} = x_1\left[\log\left(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)}\right)\right] + \log(\frac{1-\theta_1}{1-\theta_0}) \tag{1.1.2.1}$$

Using 1.1.2.1, we get,

$$\sum_{i=1}^{n} Z_i = n\log(\frac{1-\theta_1}{1-\theta_0}) + \left[\log\left(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)}\right)\right]\sum_{i=1}^{n} X_i \tag{1.1.2.2}$$

Stopping variable ($\mathfrak{N}$):

The *stopping variable* is given by,

$$\mathfrak{N} = \min\{n \geq 1 : \sum_{i=1}^{n} \notin (b, a)\} \tag{1.1.2.3}$$

where, $a = \log A, b = \log B$ and $A \approx \frac{1-\beta}{\alpha}, B \approx \frac{\beta}{1-\alpha}$ (for practical purposes).

Note 1. The statistic, i.e., the nth partial sum of X_i's, $\sum_{i=1}^{n} X_i$ is *sufficient* for θ under the family of distribution $\mathcal{P} := \{\text{Ber}(\theta), \theta \in (0,1)\}$. Therefore, we express $\mathfrak{N}$ in terms of $\sum_{i=1}^{n} X_i$.

$$\Leftrightarrow \sum_{i=1}^{n} X_i \leq d = \frac{b - n\log(\frac{1-\theta_1}{1-\theta_0})}{\log(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)})} \cap \sum_{i=1}^{n} X_i \geq \frac{a - n\log(\frac{1-\theta_1}{1-\theta_0})}{\log(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)})} = c \tag{1.1.2.4}$$

Therefore, from 1.1.2.4, we have,

$$\mathfrak{N} = \min\{n \geq 1 : \sum_{n=1}^{n} X_i \notin (d, c)\} \tag{1.1.2.5}$$

where, d and c are determined as above.

Decision Rule for the WALD'S SPRT:

The WALD'S SPRT for 1.1.1.1 is given by,

$$\phi(\tilde{x}) = \begin{cases} 1, & \sum_{i=1}^{n} X_i \geq d = \frac{b - n\log(\frac{1-\theta_1}{1-\theta_0})}{\log(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)})} \\ 0, & \sum_{i=1}^{n} X_i \leq c = \frac{a - n\log(\frac{1-\theta_1}{1-\theta_0})}{\log(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)})} \end{cases} \tag{1.1.2.6}$$

at strength $(\alpha, \beta) = (0.2, 0.2)$.

Computations for the WALD'S SPRT:

We get, $A \approx 4 \Rightarrow a = 1.386$ and $B \approx 0.25 \Rightarrow b = -1.386$. Therefore, for $\theta_0 = 0.5$ and $\theta_1 = 0.8$, the *test function* for the WALD'S SPRT as derived in 1.1.2.6 comes out as,

$$\phi(\tilde{x}) = \begin{cases} 1, & \sum_{i=1}^{n} X_i \geq d = \frac{-1.386 - n\log 0.4}{\log 4} \\ 0, & \sum_{i=1}^{n} X_i \leq c = \frac{1.386 - n\log 0.4}{\log 4} \end{cases} \tag{1.1.2.7}$$

at strength $(\alpha, \beta) = (0.2, 0.2)$.

1.1.3 (1) Continuation Region at the nth Stage:

We denote the *continuation region* at the nth stage by, $\mathcal{S}_n^c$, i.e.,

$$\Rightarrow \mathcal{S}_n^c = \{\tilde{x}_n : d = \frac{b - n\log(\frac{1-\theta_1}{1-\theta_0})}{\log(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)})} \leq \sum_{i=1}^{n} X_i \leq c = \frac{a - n\log(\frac{1-\theta_1}{1-\theta_0})}{\log(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)})}\} \tag{1.1.3.1}$$

Computations for Continuation Region at the nth Stage:

We get, $A \approx 4 \Rightarrow a = 1.386$ and $B \approx 0.25 \Rightarrow b = -1.386$. Therefore, for $\theta_0 = 0.5$ and $\theta_1 = 0.8$, the continuation region for the WALD's SPRT as derived in 1.1.3.1 comes out as a function of the *sample size*, n, i.e.,

$$\mathcal{S}_n^c = \{\tilde{x}_n : d = \frac{-1.386 - n\log 0.4}{\log 4} \leq \sum_{i=1}^{n} X_i \leq c = \frac{1.386 - n\log 0.4}{\log 4}\} \tag{1.1.3.2}$$

1.1.4 (2) Decision on the 9th Observation in the Sequence:

Given the trail/sequence of observations (*defective* or *non-defective*), we are to decide on the 9*th observation*. Note that,

$$\tilde{x}_9 = \{\text{DNNDDDDDD}\}$$

where D denotes *defective* and ND denotes *non-defective* item.

Here, 1.1.2.7 and 1.1.3.2 are respectively used to determine the *acceptance/rejection/continuation* status on the 9th observation for the given sequence of defective and non-defective items, i.e., $\tilde{x}_9$. Consider the following,

n	X_n	d	$S_n = \sum_{i=1}^{n} X_i$	c	Decision
1	1	-0.3391	1	1.661	CONTINUE
2	0	0.322	1	2.322	CONTINUE
3	0	0.983	1	2.983	CONTINUE
4	1	1.644	2	3.644	CONTINUE
5	1	2.305	3	4.305	CONTINUE
6	1	2.966	4	4.966	CONTINUE
7	1	3.627	5	5.627	CONTINUE
8	1	4.288	6	6.287	CONTINUE
9	1	4.949	7	6.948	REJECT

Table 1: *WALD'S SPRT DECISION RESULTS, 0:Non-Defective & 1:Defective*

Observation:

In reference to Table 2, it is to be observed that, as we move *downwards*, i.e., the *sample size increases* ↑, the *continuation region* of the WALD'S SPRT derived, becomes *weaker* and finally on the 9th observation of the trial/sequence of *defective* and *non-defective* items, the *null hypothesis* $\mathcal{H}_0 : \theta = \theta_0 = 0.5$ gets *rejected*. Consider the following figure 1, where the *continuation region* has been highlighted in *red*.

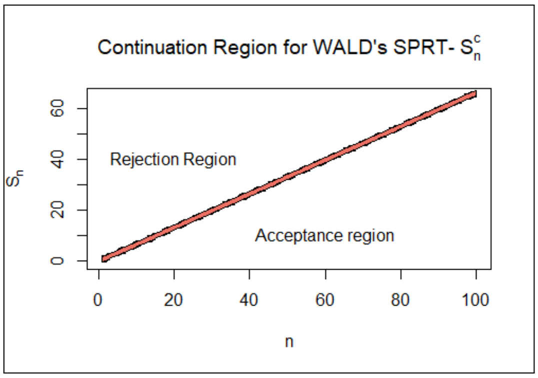

Figure 1: *Partitioned Regions (Acceptance/Rejection/Continuation) of the Sample Space $\mathcal{X}_n$ based on $\tilde{x}_n$, for $n = 100$*

1.1.5 (3.1) WALD'S Approximate OC Function:

The *OC Function* is given by,

$$\mathfrak{L}(\theta) \approx \begin{cases} \frac{A^{t_0}-1}{A^{t_0}-B^{t_0}}, & t_0 \neq 0 \\ \frac{\log A}{\log A - \log B}, & t_0 = 0 \end{cases} \tag{1.1.5.1}$$

where A and B are given by 1.1.2.3, also, t_0 is the *unique solution* of,

$$\mathcal{M}_{Z_1}(t_0) = \mathbb{E}_\theta(e^{t_0 Z_1}) = 1$$

$$\Rightarrow \mathbb{E}_\theta\left[e^{t_0\left(x_1[\log(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)})]+\log(\frac{1-\theta_1}{1-\theta_0})\right)}\right] = 1 \text{ (from 1.1.2.1)}$$

$$\Leftrightarrow (\frac{1-\theta_1}{1-\theta_0})^{t_0(\theta)}\left[1-\theta+\theta(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)})^{t_0(\theta)}\right] = 1 \tag{1.1.5.2}$$

$$\Leftrightarrow \theta = \frac{1-(\frac{1-\theta_0}{1-\theta_1})^{t_0(\theta)}}{1-(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)})^{t_0(\theta)}} \tag{1.1.5.3}$$

Note 2. For the time being, we consider the expression for $t_0(\theta)$ (refer to 1.1.5.2) as well as θ (refer to 1.1.5.3) for $t_0(\theta) \neq 0$, whereas, the limit could be easily evaluated when $t_0(\theta) \to 0$.

1.1.6 (3.2) WALD'S Approximate ASN Function:

The *ASN Function* is given by,

$$\mathbb{E}_\theta(\mathfrak{N}) = \begin{cases} \frac{b\cdot\mathfrak{L}(\theta)+a\cdot(1-\mathfrak{L}(\theta))}{\mathbb{E}_\theta(Z_1)}, & \mathbb{E}_\theta(Z_1) \neq 0 \\ \frac{b^2\cdot\mathfrak{L}(\theta)+a^2\cdot(1-\mathfrak{L}(\theta))}{V_\theta(Z_1)}, & \mathbb{E}_\theta(Z_1) = 0 \end{cases} \tag{1.1.6.1}$$

where, $\mathbb{E}_\theta(Z_1) = \theta\left[\log\left(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)}\right)\right] + \log(\frac{1-\theta_1}{1-\theta_0})$, $V_\theta(Z_1) = \theta(1-\theta)\left[\log\left(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)}\right)\right]^2$, a and b as defined in 1.1.2.3

1.1.7 Results - OC & ASN Curves:

This section gives an overview of the following results on the *Wald's Approximate OC & ASN Functions* and *corresponding parameters*, as given by 1.1.5.2, 1.1.5.1 and 1.1.6.1 respectively.

$t_0(\theta)$	θ	$\mathfrak{L}(\theta)$	$\mathbb{E}_\theta(Z_1)$	$V_\theta(Z_1)$	$\mathbb{E}_\theta(\mathfrak{N})$
-0.500	0.735	0.333	0.103	0.374	4.497
-0.300	0.706	0.398	0.063	0.399	4.514
-0.200	0.692	0.431	0.042	0.409	4.507
0.100	0.645	0.500	-0.022	0.439	4.424
0.200	0.629	0.569	-0.044	0.448	4.375
0.500	0.581	0.666	-0.111	0.468	4.176
0.700	0.549	0.725	-0.156	0.476	4.008

Table 2: *Results for Wald's Approximate OC & ASN Fuctions*

The OC and ASN curves are illustrated in the following diagram, i.e., Figure 2,

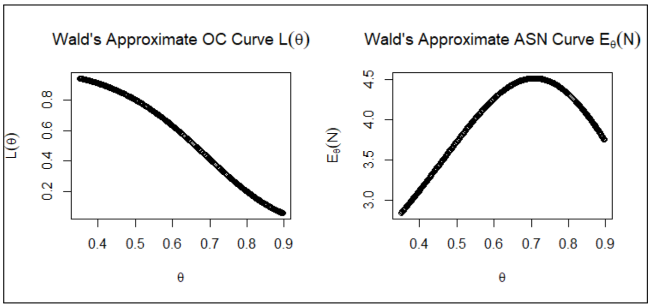

Figure 2: *Wald's Approximate OC and ASN Curves against θ*

2 Problem 2

If θ is the *expected number of defects*, determine the *continuation region* of *SPRT* at the nth stage for testing $\mathcal{H}_0 : \theta = 7$ against $\mathcal{H}_1 : \theta = 10$, taking $\alpha = 0.2$ and $\beta = 0.2$.

1. If the number of defects in a sequence of inspections are $\{7, 15, 11, 3, 4, 5, 7, 9, 12, 14, 0, 11\}$ respectively, what will your decision on the 12th observation.
2. Calculate *Wald's Aprroximate OC* and *ASN* functions for different values of θ.

2.1 Solution to Problem 2

$\theta \Rightarrow$ *expected # defects.* To determine the *Wald's SPRT* and its *continuation region* at the nth stage by taking $\alpha = 0.2 = \beta$, i.e., *strength* $= (\alpha, \beta) = (0.2, 0.2)$.

Define $X_n :=$ *# defects for the nth item*, $n \geq 1$. Clearly, $X_n \overset{iid}{\sim} \text{Poi}(\theta)$.

2.1.1 Testing Problem:

We are to test,

$$\mathcal{H}_0 : \theta = \theta_0 = 7 \text{ against } \mathcal{H}_1 : \theta = \theta_1 = 10 \tag{2.1.1.1}$$

2.1.2 To obtain the WALD'S SPRT for 2.1.1.1:

The PMF of X_i's is given by,

$$p_\theta(x_i) = \frac{e^{-\theta}\theta^{x_i}}{x_i!}, x_i \in \mathbb{N} \cup \{0\}, \theta \in \mathbb{R}^+, \forall\ i$$

We have,

$$\lambda_1 = \frac{p_{\theta_1}(x_1)}{p_{\theta_0}(x_1)} = e^{-(\theta_1 - \theta_0)}(\frac{\theta_1}{\theta_0})^{x_1} \tag{2.1.2.1}$$

Now,

$$Z_1 = \log\lambda_1 = -(\theta_1 - \theta_0) + x_1 \log(\frac{\theta_1}{\theta_0}) \tag{2.1.2.2}$$

Using 2.1.2.2, we get,

$$\sum_{i=1}^{n} Z_i = n(\theta_0 - \theta_1) + \log(\frac{\theta_1}{\theta_0}) \sum_{i=1}^{n} x_i \tag{2.1.2.3}$$

Stopping Variable ($\mathfrak{N}$):

The *stopping variable* ($\mathfrak{N}$) is given by,

$$\mathfrak{N} = \min\{n \geq 1 : \sum_{i=1}^{n} Z_i \notin (b, a)\} \tag{2.1.2.4}$$

where, $a = \log A$, $b = \log B$ and $A \approx \frac{1-\beta}{\alpha}$, $B \approx \frac{\beta}{1-\alpha}$ (for practical purposes).

Note 3. The statistic, i.e., the nth partial sum of X_i's, $\sum_{i=1}^{n} X_i$ is *sufficient* for θ under the family of distribution $\mathcal{P} := \{\text{Poi}(\theta), \theta \in \mathbb{R}^+\}$. Hence, we express $\mathfrak{N}$ in terms of $\sum_{i=1}^{n} X_i$.

We have,

$$\sum_{i=1}^{n} Z_i \notin (b,a) \Leftrightarrow \sum_{i=1}^{n} Z_i \leq b \ \cap \ \sum_{i=1}^{n} Z_i \geq a$$

$$\Leftrightarrow n(\theta_0-\theta_1)+\log{(\frac{\theta_1}{\theta_0})}\sum_{i=1}^{n} x_i \leq b \ \cap \ n(\theta_0-\theta_1)+\log{(\frac{\theta_1}{\theta_0})}\sum_{i=1}^{n} x_i \geq a$$

$$\Leftrightarrow \sum_{i=1}^{n} x_i \leq d = \frac{b-n(\theta_0-\theta_1)}{\log{(\frac{\theta_1}{\theta_0})}} \ \cap \ \sum_{i=1}^{n} x_i \geq \frac{a-n(\theta_0-\theta_1)}{\log{(\frac{\theta_1}{\theta_0})}} = c \tag{2.1.2.5}$$

Therefore, $\sum_{i=1}^{n} X_i \notin (d,c)$. Hence, the stopping variable, $\mathfrak{N}$ in terms of $\sum_{i=1}^{n} X_i$ is given as follows,

$$\mathfrak{N} = \min\{n \geq 1 : \sum_{i=1}^{n} X_i \notin (d,c)\} \tag{2.1.2.6}$$

where, $d = \frac{b-n(\theta_0-\theta_1)}{\log{(\frac{\theta_1}{\theta_0})}}$ and $c = \frac{a-n(\theta_0-\theta_1)}{\log{(\frac{\theta_1}{\theta_0})}}$.

Decision Rule for the WALD'S SPRT:

The WALD'S SPRT for 2.1.1.1 is given by,

$$\phi(\tilde{x}) = \begin{cases} 1, & \sum_{i=1}^{n} X_i \geq d = \frac{b-n(\theta_0-\theta_1)}{\log{(\frac{\theta_1}{\theta_0})}} \\ 0, & \sum_{i=1}^{n} X_i \leq c = \frac{a-n(\theta_0-\theta_1)}{\log{(\frac{\theta_1}{\theta_0})}} \end{cases} \tag{2.1.2.7}$$

at strength $(\alpha,\beta) = (0.2, 0.2)$.

Computations for the WALD'S SPRT:

We get, $A \approx 4 \Rightarrow a = 1.386$ and $B \approx 0.25 \Rightarrow b = -1.386$. Therefore, for $\theta_0 = 7$ and $\theta_1 = 10$, the *test function* for the WALD's SPRT as derived in 2.1.2.7 comes out as,

$$\phi(\tilde{x}) = \begin{cases} 1, & \sum_{i=1}^{n} X_i \geq d = \frac{-1.386+3n}{0.357} \\ 0, & \sum_{i=1}^{n} X_i \leq c = \frac{1.386+3n}{0.357} \end{cases} \tag{2.1.2.8}$$

at strength $(\alpha,\beta) = (0.2, 0.2)$.

2.1.3 Continuation Region at the nth Stage:

We denote the *continuation region* at the nth stage by, $\mathcal{S}_n^c$, i.e.,

$$\Rightarrow \mathcal{S}_n^c = \{\tilde{x}_n : b < \sum_{i=1}^{n} Z_i < a\}$$

$$\Leftrightarrow \mathcal{S}_n^c = \left\{\tilde{x}_n : d = \frac{b-n(\theta_0-\theta_1)}{\log{(\frac{\theta_1}{\theta_0})}} < \sum_{i=1}^{n} X_i < c = \frac{a-n(\theta_0-\theta_1)}{\log{(\frac{\theta_1}{\theta_0})}}\right\} \tag{2.1.3.1}$$

Computations for Continuation Region at the nth Stage:

We get, $A \approx 4 \Rightarrow a = 1.386$ and $B \approx 0.25 \Rightarrow b = -1.386$. Therefore, for $\theta_0 = 7$ and $\theta_1 = 10$, the continuation region for the WALD's SPRT as derived in 2.1.3.1 comes out as a function of the *sample size*, n, i.e.,

$$\Rightarrow \mathcal{S}_n^c = \left\{\tilde{x}_n : \frac{-1.386+3n}{0.357} < \sum_{i=1}^{n} X_i < \frac{1.386+3n}{0.357}\right\} \tag{2.1.3.2}$$

2.1.4 (1) Decision on the 12th Observation in the Sequence of Inspections:

Given, the trail/sequence of observations (inspections), we are to decide on the *12th observation*. Note that,

$$\tilde{x}_{12} = \{7, 15, 11, 3, 4, 5, 7, 9, 12, 14, 10, 11\}$$

We use, 2.1.2.7 and 2.1.3.1 to determine the *acceptance/rejection/continuation* status on the 12th observation in the given sequence of inspection readings, i.e., $\tilde{x}_{12}$. Consider the following,

n	X_n	d	$S_n = \sum_{i=1}^{n} X_i$	c	Decision
1	7	4.524	7	12.298	CONTINUE
2	15	12.935	22	20.709	REJECT
3	11	21.346	33	29.119	REJECT
4	3	29.757	36	37.531	CONTINUE
5	4	38.168	40	45.942	CONTINUE
6	5	46.579	45	54.353	REJECT
7	7	54.990	52	62.764	REJECT
8	9	63.401	61	71.175	REJECT
9	12	71.812	73	79.586	CONTINUE
10	14	80.223	87	87.997	CONTINUE-WEAKLY
11	0	88.635	87	96.408	ACCEPT
12	11	97.046	98	104.819	CONTINUE

Table 3: *WALD'S SPRT DECISION RESULTS*

Observation:

The results for each of the observations in the given sequence has been presented in Table 3. Note that, for the *second sample/sequential observation*, the *rejection region* is encountered, i.e., the WALD'S SPRT terminates. But, for the 12th observation in the sequence, the WALD'S SPRT indicates an *continuation region* in the light of the given observations and data. Now, this is only possible, if we work with each and every observation in the sequence of inspections provided, regardless of the termination of the SPRT, as in our case for the second observation. *Hence, we present our results here, assuming that, the SPRT is continued till the 11th observation.* The following Figure 3 illustrates the *acceptance*, *rejection* and *continuation* (highlighted portion) regions for the resultant WALD'S SPRT as given by 2.1.2.8.

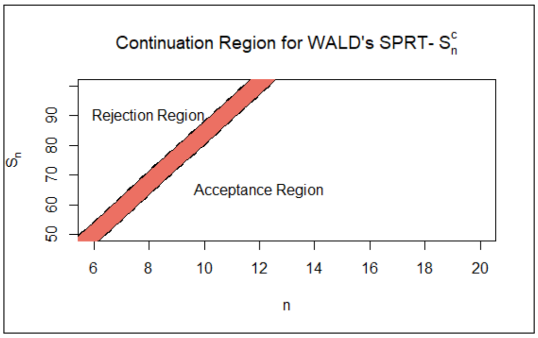

Figure 3: *Partitioned Regions (Acceptance/Rejection/Continuation) of the Sample Space $\mathcal{X}_n$ based on $\tilde{x}_n$, for* $n = 100$

2.1.5 (2.1) WALD'S Approximate OC Function:

The *OC Function* is given by,

$$\mathfrak{L}(\theta) \approx \begin{cases} \frac{A^{t_0}-1}{A^{t_0}-B^{t_0}}, & t_0 \neq 0 \\ \frac{\log A}{\log A - \log B}, & t_0 = 0 \end{cases} \tag{2.1.5.1}$$

where A and B are given by 2.1.2.4, also, t_0 is the *unique solution* of,

$$\mathcal{M}_{Z_1}(t_0) = \mathbb{E}_\theta(e^{t_0 Z_1}) = 1$$

$$\Rightarrow \mathbb{E}_\theta\left[e^{t_0(\theta_0-\theta_1)+t_0 x_1 \log\left(\frac{\theta_1}{\theta_0}\right)}\right] = 1 \text{ (from 2.1.2.2)}$$

$$\Rightarrow e^{t_0(\theta_0-\theta_1)}\mathbb{E}_\theta(e^{t_0 * X_1}) = 1$$

where, $t_0* = t_0 \log\left(\frac{\theta_1}{\theta_0}\right)$. Now, if $X \sim \text{Poi}(\theta)$, then, $\mathcal{M}_X(t) = \mathbb{E}(e^{tX}) = e^{-\theta}e^{\theta^t}, t \in \mathbb{R}$. Hence,

$$\Rightarrow e^{t_0(\theta_0-\theta_1)-\theta+\theta\left(\frac{\theta_1}{\theta_0}\right)^{t_0}} = 1$$

$$\Rightarrow t_0(\theta_0-\theta_1) - \theta + \theta\left(\frac{\theta_1}{\theta_0}\right)^{t_0} = 0$$

$$\Leftrightarrow t_0(\theta) = \frac{\theta\left(\left(\frac{\theta_1}{\theta_0}\right)^{t_0} - 1\right)}{\theta_1 - \theta_0} \tag{2.1.5.2}$$

Note 4. *Observe that, for $t_0(\theta) = 0$, we have,*

$$\Leftrightarrow \theta = \frac{\theta_1 - \theta_0}{\log\left(\frac{\theta_1}{\theta_0}\right)} \tag{2.1.5.3}$$

Note 5. In order to avoid computational complexity by using *Numerical Methods* for determining $t_0(\theta)$ as given by 2.1.5.2, we evaluate $\mathfrak{L}(\theta)$ as in 2.1.5.1 by finding θ after fixing t_0, as illustrated in Section 2.1.7.

2.1.6 (2.2) WALD'S Approximate ASN Function:

From 2.1.2.2, we have, $Z_1 = -(\theta_1 - \theta_0) + X_1 \log\left(\frac{\theta_1}{\theta_0}\right)$. Therefore, we have,

$$\Rightarrow \mathbb{E}_\theta(Z_1) = -(\theta_1 - \theta_0) + \theta \log\left(\frac{\theta_1}{\theta_0}\right) \tag{2.1.6.1}$$

$$\Rightarrow V_\theta(Z_1) = \theta\left(\log\left(\frac{\theta_1}{\theta_0}\right)\right)^2 \tag{2.1.6.2}$$

Therefore, using 2.1.6.1 and 2.1.6.2, the *ASN Function* is given by,

$$\mathbb{E}_\theta(\mathfrak{N}) = \begin{cases} \frac{b\cdot\mathfrak{L}(\theta)+a\cdot(1-\mathfrak{L}(\theta))}{\mathbb{E}_\theta(Z_1)}, & \mathbb{E}_\theta(Z_1) \neq 0 \\ \frac{b^2\cdot\mathfrak{L}(\theta)+a^2\cdot(1-\mathfrak{L}(\theta))}{V_\theta(Z_1)}, & \mathbb{E}_\theta(Z_1) = 0 \end{cases} \tag{2.1.6.3}$$

where a and b are defined as in 2.1.2.4.

2.1.7 Results - OC & ASN Curves:

This section gives an overview of the following results on the *Wald's Approximate OC & ASN Functions* and *corresponding parameters*, as given by 2.1.5.2, 2.1.5.3, 2.1.5.1, 2.1.6.1, 2.1.6.2 and 2.1.6.3 respectively.

$t_0(\theta)$	θ	$\mathfrak{L}(\theta)$	$\mathbb{E}_\theta(Z_1)$	$V_\theta(Z_1)$	$\mathbb{E}_\theta(\mathfrak{N})$
-0.500	9.183	0.333	0.275	1.168	1.677
-0.300	8.869	0.398	0.163	1.128	1.739
-0.200	8.715	0.431	0.108	1.109	1.764
0.000	8.411	0.500	0.000	1.070	1.796
0.200	8.115	0.569	-0.106	1.032	1.806
0.500	7.683	0.666	-0.259	0.977	1.780
0.700	7.405	0.725	-0.359	0.942	1.739

Table 4: *Results for Wald's Approximate OC & ASN Fuctions*

The OC and ASN curves are illustrated in the following diagram, i.e., Figure 4,

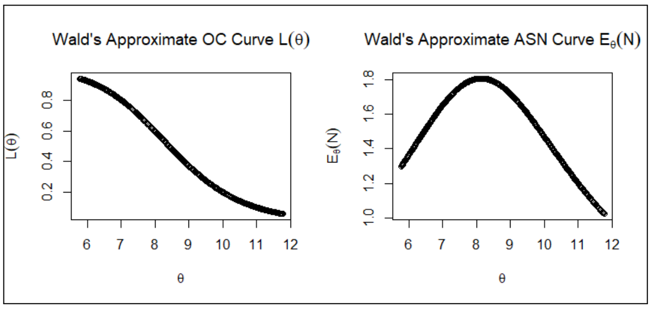

Figure 4: *Wald's Approximate OC and ASN Curves against θ*

2.1.8 Comparing WALD'S SPRT to the Fixed Sample Size Test:

We have already constructed a *Wald's SPRT* (refer to 2.1.2.8 and 2.1.2.7) for the testing problem as stated in 2.1.1.1. Now, we consider a *Fixed Sample Size Test* for the same testing problem, i.e., for 2.1.1.1.

Now, for the Fixed Sample Size Test, we are to find the *sample size* n, for which the following twp conditions holds, i.e.,

$$\mathbb{P}_{\mathcal{H}_0}(T_n \geq k) \leq \alpha \tag{2.1.8.1}$$

$$\mathbb{P}_{\mathcal{H}_1}(T_n \geq k) > 1 - \beta \tag{2.1.8.2}$$

where, $T_n = \sum_{i=1}^n X_i$. The conditions 2.1.8.1 and 2.1.8.2 are the *size/level* and *power* conditions respectively, which are to be satisfied. We already have the *strength* as $(\alpha, \beta) = (0.2, 0.2)$.

Also note that,

$$T_n = \sum_{i=1}^n X_i \sim \begin{cases} \text{Poi}(7 \cdot n), & \text{under } \mathcal{H}_0 \\ \text{Poi}(10 \cdot n), & \text{under } \mathcal{H}_1 \end{cases} \tag{2.1.8.3}$$

After, necessary computations, we get,

- $n = 4$
- $k = 33$
- $\mathbb{P}_{\mathcal{H}_0}(T_n \geq 33) = 0.195$ - size of the test
- $\mathbb{P}_{\mathcal{H}_1}(T_n \geq 33) = 0.885$ - power of the test.

Observations:

In case of the *Wald's SPRT*, under $\mathcal{H}_0 : \theta = \theta_0 = 7$, the *approximate Wald's ASN* comes out as $\mathbb{E}_{\theta_0=7}(\mathfrak{N}) \approx 1.653$ and under $\mathcal{H}_1 : \theta = \theta_1 = 10$, the it is $\mathbb{E}_{\theta_1=10}(\mathfrak{N}) \approx 1.468$ (refer to 2.1.6.3 and Figure 4). Therefore, it is to be observed that, the sample size required in case of the *Wald's Sequential Test* is much *lesser* than that of the *Fixed Sample Size Test*, where $n = 4$, as a result of which the *Wald's SPRT* developed as in 2.1.2.8 turns out to be *more economical in terms of cost and time consumption* than the *Fixed Sample Size Test*.

3 Problem 3

Suppose independent observations are drawn from a $N(\theta, 1)$ distribution.

1. Determine the *continuation region of SPRT* at the nth stage for testing $\mathcal{H}_0 : \theta = 0$ against $\mathcal{H}_1 : \theta = 1$ taking $\alpha = 0.01 = \beta$.
2. Calculate *Wald's Approximate OC* and *ASN* functions for different values of θ.

3.1 Solution to Problem 3

Consider $\mathbf{X} := (X_1, X_2, \ldots, X_n, \ldots) \overset{iid}{\sim} N(\theta, 1)$, $n \geq 1$. To determine the *Wald's SPRT* and its *continuation region* at the nth stage by taking $\alpha = 0.01 = \beta$, i.e., *strength* $= (\alpha, \beta) = (0.01, 0.01)$.

3.2 Testing Problem:

We are to test,

$$\mathcal{H}_0 : \theta = \theta_0 = 0 \text{ against } \mathcal{H}_1 : \theta = \theta_1 = 1 \tag{3.2.0.1}$$

3.2.1 (1) To obtain the WALD'S SPRT for 3.2.0.1:

The density or PDF of X_i's is given by,

$$p_\theta(x_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i - \theta)^2}, x_i \in \mathbb{R} \ \forall\ i, \theta \in \mathbb{R}$$

We have,

$$\lambda_n = \prod_{i=1}^{n} \frac{p_{\theta_1}(x_i)}{p_{\theta_0}(x_i)} = e^{n(\theta_1 - \theta_0)(\overline{x} - \frac{\theta_1 + \theta_0}{2})} \tag{3.2.1.1}$$

Therefore,

$$\sum_{i=1}^{n} Z_i = \log \lambda_n = \sum_{i=1}^{n} \log(\frac{p_{\theta_1}(x_i)}{p_{\theta_0}(x_i)}) = n(\theta_1 - \theta_0)(\overline{x} - \frac{\theta_1 + \theta_0}{2}) \tag{3.2.1.2}$$

Stopping Variable ($\mathfrak{N}$):

The *stopping variable* ($\mathfrak{N}$) is given by,

$$\mathfrak{N} = \min\{n \geq 1 : \sum_{i=1}^{n} Z_i \notin (b, a)\} \tag{3.2.1.3}$$

where, $a = \log A, b = \log B$ and $A \approx \frac{1-\beta}{\alpha}, B \approx \frac{\beta}{1-\alpha}$ (for practical purposes).

Note 6. *The statistic, i.e., the sample mean based on n observations, $\overline{X} = n^{-1} \sum_{i=1}^{n} X_i$ is* sufficient *for θ under the family of distribution $\mathcal{P} := \{N(\theta, 1), \theta \in \mathbb{R}\}$. Hence, we express $\mathfrak{N}$ in terms of $\overline{X}$.*

We have,

$$\sum_{i=1}^{n} Z_i \notin (b, a) \Leftrightarrow \sum_{i=1}^{n} Z_i \leq b \ \cap \ \sum_{i=1}^{n} Z_i \geq a$$

$$\Leftrightarrow \overline{X} \leq d = \frac{b}{n(\theta_1 - \theta_0)} + \frac{\theta_0 + \theta_1}{2} \ \cap \ \overline{X} \geq \frac{a}{n(\theta_1 - \theta_0)} + \frac{\theta_0 + \theta_1}{2} = c$$

Therefore, $\overline{X} \notin (d, c)$. Hence, the stopping variable, $\mathfrak{N}$ in terms of $\overline{X}$ is given as follows,

$$\mathfrak{N} = \min\{n \geq 1 : \overline{X} \notin (d, c)\} \tag{3.2.1.4}$$

where, $d = \frac{b}{n(\theta_1 - \theta_0)} + \frac{\theta_0+\theta_1}{2}$ and $c = \frac{a}{n(\theta_1-\theta_0)} + \frac{\theta_0+\theta_1}{2}$.

Decision Rule for the WALD'S SPRT:

The WALD'S SPRT for 3.2.0.1 is given by,

$$\phi(\tilde{x}) = \begin{cases} 1, & \overline{X} \geq d = \frac{b}{n(\theta_1-\theta_0)} + \frac{\theta_0+\theta_1}{2} \\ 0, & \overline{X} \leq c = \frac{a}{n(\theta_1-\theta_0)} + \frac{\theta_0+\theta_1}{2} \end{cases} \tag{3.2.1.5}$$

at strength $(\alpha, \beta) = (0.01, 0.01)$.

Computations for the WALD'S SPRT:

We get, $A \approx 99 \Rightarrow a = 4.595$ and $B \approx 0.0101 \Rightarrow b = -4.595$. Therefore, for $\theta_0 = 0$ and $\theta_1 = 1$, the *test function* for the WALD'S SPRT as derived in 3.2.1.5 comes out as,

$$\phi(\tilde{x}) = \begin{cases} 1, & \overline{X} \geq d = 0.50 - \frac{4.595}{n} \\ 0, & \overline{X} \leq c = 0.50 + \frac{4.595}{n} \end{cases} \tag{3.2.1.6}$$

at strength $(\alpha, \beta) = (0.01, 0.01)$.

3.2.2 Continuation Region at the nth Stage:

We denote the *continuation region* at the nth stage by, $\mathcal{S}_n^c$, i.e.,

$$\Rightarrow \mathcal{S}_n^c = \{\tilde{x}_n : b < \sum_{i=1}^{n} Z_i < a\}$$

$$\Leftrightarrow \mathcal{S}_n^c = \left\{\tilde{x}_n : d = \frac{b}{n(\theta_1 - \theta_0)} + \frac{\theta_0 + \theta_1}{2} < \overline{X} < c = \frac{a}{n(\theta_1 - \theta_0)} + \frac{\theta_0 + \theta_1}{2}\right\} \tag{3.2.2.1}$$

Therefore, putting the values of a, b, θ_0 and θ_1 in 3.2.2.1, we get the region as,

$$\Rightarrow \mathcal{S}_n^c = \left\{\tilde{x}_n : d = 0.50 - \frac{4.595}{n} < \overline{X} < 0.50 + \frac{4.595}{n}\right\} \tag{3.2.2.2}$$

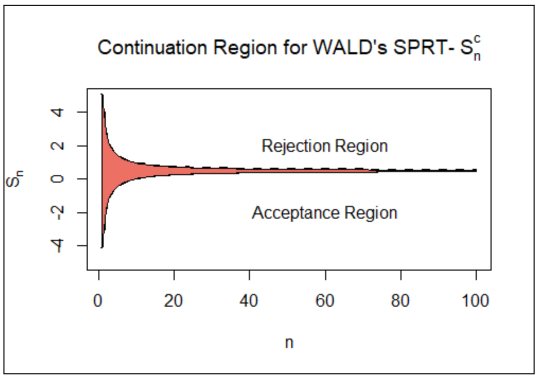

Figure 5: *Partitioned Regions (Acceptance/Rejection/Continuation) of the Sample Space $\mathcal{X}_n$ based on $\tilde{x}_n$, for $n = 100$. The red shaded area is the continuation region.*

3.2.3 (2.1) WALD'S Approximate OC Function:

The *OC Function* is given by,

$$\mathfrak{L}(\theta) \approx \begin{cases} \frac{A^{t_0}-1}{A^{t_0}-B^{t_0}}, & t_0 \neq 0 \\ \frac{\log A}{\log A - \log B}, & t_0 = 0 \end{cases} \tag{3.2.3.1}$$

where A and B are given by 3.2.1.3, also, t_0 is the *unique solution* of,

$$\mathcal{M}_{Z_1}(t_0) = \mathbb{E}_\theta(e^{t_0 Z_1}) = 1$$

We have, $Z_1 = (\theta_1 - \theta_0)(X - \frac{\theta_0+\theta_1}{2}) \sim N((\theta_1-\theta_0)(\theta - \frac{\theta_0+\theta_1}{2}), (\theta_1-\theta_0)^2)$. Therefore, $\mathcal{M}_{Z_1}(t_0)$ is given by,

$$\mathcal{M}_{Z_1}(t_0) = e^{t_0(\theta_1-\theta_0)(\theta-\frac{\theta_0+\theta_1}{2})+\frac{1}{2}t_0{}^2(\theta_1-\theta_0)^2}$$

Now, $\mathcal{M}_{Z_1}(t_0) = 1$, hence,

$$\Rightarrow t_0(\theta_1-\theta_0)(\theta - \frac{\theta_0+\theta_1}{2}) + \frac{1}{2}t_0^2(\theta_1-\theta_0)^2 = 0$$

$$\Leftrightarrow t_0(\theta) = \frac{\theta_0+\theta_1-2\theta}{\theta_1-\theta_0} \tag{3.2.3.2}$$

$$\Leftrightarrow \theta = \frac{(\theta_0+\theta_1)-(\theta_1-\theta_0)t_0(\theta)}{2} \tag{3.2.3.3}$$

Note 7. As in *Problem 2*, to avoid computational complexity by using *Numerical Methods* for determining $t_0(\theta)$ as given by 3.2.3.2, we calculate $\mathfrak{L}(\theta)$ as in 3.2.3.1 by finding θ after fixing t_0, as illustrated in Section 3.2.5.

3.2.4 (2.2) WALD'S Approximate ASN Function:

The *ASN Function* is given by,

$$\mathbb{E}_\theta(\mathfrak{N}) \approx \begin{cases} \frac{b\cdot\mathfrak{L}(\theta)+a\cdot(1-\mathfrak{L}(\theta))}{\mathbb{E}_\theta(Z_1)}, & \mathbb{E}_\theta(Z_1) \neq 0 \\ \frac{b^2\cdot\mathfrak{L}(\theta)+a^2(1-\mathfrak{L}(\theta))}{V(Z_1)}, & \mathbb{E}_\theta(Z_1) = 0 \end{cases} \tag{3.2.4.1}$$

where, $Z_1 = (\theta_1 - \theta_0)(X - \frac{\theta_0+\theta_1}{2}) \sim N((\theta_1-\theta_0)(\theta - \frac{\theta_0+\theta_1}{2}), (\theta_1-\theta_0)^2)$.

3.2.5 Results - OC & ASN Curves:

Here, the results on the *Wald's Approximate OC & ASN Functions* and *corresponding parameters*, as given by 3.2.3.2, 3.2.3.3, 3.2.3.1 and 3.2.4.1 respectively.

$t_0(\theta)$	θ	$\mathfrak{L}(\theta)$	$\mathbb{E}_\theta(Z_1)$	$V_\theta(Z_1)$	$\mathbb{E}_\theta(\mathfrak{N})$
-0.500	0.750	0.091	0.250	1.000	15.023
-0.300	0.650	0.201	0.150	1.000	18.304
-0.200	0.600	0.285	0.100	1.000	19.745
0.000	0.500	0.500	0.000	1.000	21.115
0.200	0.400	0.715	-0.100	1.000	19.745
0.500	0.250	0.909	-0.250	1.000	15.023
0.700	0.150	0.961	-0.350	1.000	12.117

Table 5: *Results for Wald's OC & ASN Functions*

The OC and ASN curves are illustrated in the following diagram, i.e., Figure 6.

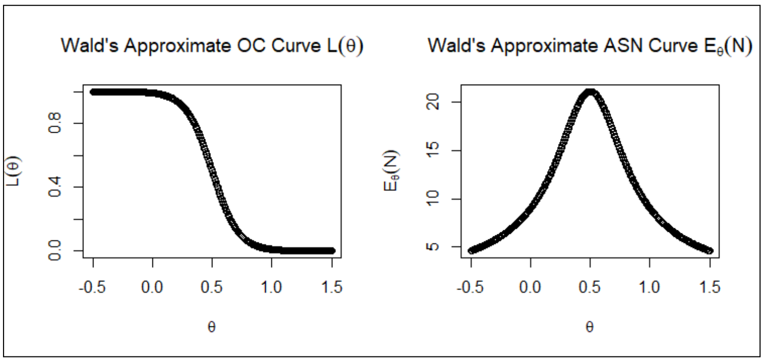

Figure 6: *Wald's Approximate OC and ASN Curves against* θ

4 Problem 4

If the observations are available from an *exponential distribution* with *mean* θ, what will be the modifications in *Problem 3* above?

4.1 Solution to Problem 4

In this problem, instead of presenting a detailed solution, the *modifications* will be thoroughly discussed, i.e., the alterations in the *set up*, *testing problem*, Z_1, *stopping variable* $\mathfrak{N}$, *test function*, *continuation region*, *boundary points*, *Wald's Approximate OC & ASN Functions.*

Consider $\mathbf{X} := (X_1, X_2, \ldots, X_n) \overset{iid}{\sim} \text{Exp}(\text{mean} = \theta)$, $n \geq 1$ and $\theta \in \Theta \equiv \mathbb{R}^+$. To determine the *Wald's SPRT* and its respective *continuation region* at the nth stage by taking $\alpha = 0.01 = \beta$, i.e., *strength* $= (\alpha, \beta) = (0.01, 0.01)$.

4.2 Testing Problem:

Unlike the testing problem as stated in *Problem 3*, here we consider the following, i.e., to test,

$$\mathcal{H}_0 : \theta = \theta_0 = 4 \text{ against } \mathcal{H}_1 : \theta = \theta_1 = 5 \tag{4.2.0.1}$$

4.3 To obtain the WALD'S SPRT for 4.2.0.1

The density or PDF of X_i's are given by,

$$p_\theta(x_i) = \frac{1}{\theta} e^{-\frac{x_i}{\theta}}, x_i \in \mathbb{R}^+ \ \forall\ i, \theta \in \mathbb{R}^+$$

We have,

$$Z_1 = \log \frac{p_{\theta_1}(x_1)}{p_{\theta_0}(x_1)} = \log(\frac{\theta_0}{\theta_1}) + x_1(\frac{1}{\theta_0} - \frac{1}{\theta_1}) \tag{4.3.0.1}$$

Therefore,

$$\sum_{i=1}^{n} Z_i = \log \lambda_n = \sum_{i=1}^{n} \log \frac{p_{\theta_1}(x_i)}{p_{\theta_0}(x_i)} = n \log(\frac{\theta_0}{\theta_1}) + (\frac{1}{\theta_0} - \frac{1}{\theta_1}) \sum_{i=1}^{n} X_i \tag{4.3.0.2}$$

Stopping Variable ($\mathfrak{N}$):

The *stopping variable* ($\mathfrak{N}$) is given by,

$$\mathfrak{N} = \min\{n \geq 1 : \sum_{i=1}^{n} Z_i \notin (b, a)\} \tag{4.3.0.3}$$

where, $a = \log A, b = \log B$ and $A \approx \frac{1-\beta}{\alpha}, B \approx \frac{\beta}{1-\alpha}$ (for practical purposes).

The stopping variable in terms of the *sufficient statistic* is given by,

$$\mathfrak{N} = \min\{n \geq 1 : \overline{X} \notin (d, c)\} \tag{4.3.0.4}$$

where, $d = \frac{b/n - \log(\frac{\theta_0}{\theta_1})}{\frac{1}{\theta_0} - \frac{1}{\theta_1}}$ and $c = \frac{a/n - \log(\frac{\theta_0}{\theta_1})}{\frac{1}{\theta_0} - \frac{1}{\theta_1}}$.

Decision Rule for the WALD'S SPRT:

The WALD'S SPRT for 4.2.0.1 is given by,

$$\phi(\tilde{x}) = \begin{cases} 1, & \overline{X} \geq d = \frac{b/n - \log(\frac{\theta_0}{\theta_1})}{\frac{1}{\theta_0} - \frac{1}{\theta_1}} \\ 0, & \overline{X} \leq c = \frac{a/n - \log(\frac{\theta_0}{\theta_1})}{\frac{1}{\theta_0} - \frac{1}{\theta_1}} \end{cases} \tag{4.3.0.5}$$

at strength $(\alpha, \beta) = (0.01, 0.01)$.

Computations for the WALD'S SPRT:

We get, $A \approx 99 \Rightarrow a = 4.595$ and $B \approx 0.0101 \Rightarrow b = -4.595$. Therefore, for $\theta_0 = 4$ and $\theta_1 = 5$, the *test function* for the WALD'S SPRT as derived in 4.3.0.5 comes out as,

$$\phi(\tilde{x}) = \begin{cases} 1, & \overline{X} \geq d = \frac{-\frac{4.595}{n} - \log 0.8}{1/4 - 1/5} \\ 0, & \overline{X} \leq c = \frac{\frac{4.595}{n} - \log 0.8}{1/4 - 1/5} \end{cases} \tag{4.3.0.6}$$

at strength $(\alpha, \beta) = (0.01, 0.01)$.

4.4 Continuation Region at the nth Stage:

We denote the *continuation region* at the nth stage by, $\mathcal{S}_n^c$, i.e.,

$$\Rightarrow \mathcal{S}_n^c = \left\{ \tilde{x}_n : d = \frac{b/n - \log(\frac{\theta_0}{\theta_1})}{\frac{1}{\theta_0} - \frac{1}{\theta_1}} < \overline{X} < c = \frac{a/n - \log(\frac{\theta_0}{\theta_1})}{\frac{1}{\theta_0} - \frac{1}{\theta_1}} \right\} \tag{4.4.0.1}$$

Therefore, putting the values of a, b, θ_0 and θ_1 in 4.4.0.1, we get the region as,

$$\Rightarrow \mathcal{S}_n^c = \left\{ \tilde{x}_n : d = \frac{-\frac{4.595}{n} - \log 0.8}{1/4 - 1/5} < \overline{X} < c = \frac{\frac{4.595}{n} - \log 0.8}{1/4 - 1/5} \right\} \tag{4.4.0.2}$$

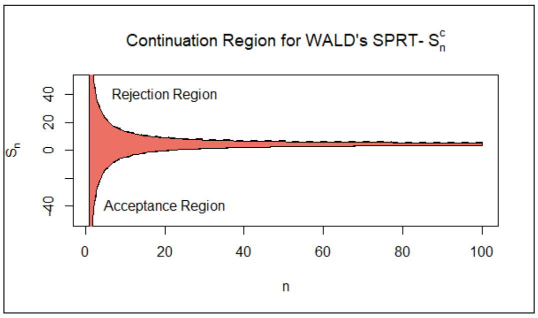

Figure 7: *Partitioned Regions (Acceptance/Rejection/Continuation) of the Sample Space $\mathcal{X}_n$ based on $\tilde{x}_n$, for $n = 100$. The red shaded area is the continuation region.*

4.5 WALD'S Approximate OC Function:

The *OC Function* is given by,

$$\mathfrak{L}(\theta) \approx \begin{cases} \frac{A^{t_0} - 1}{A^{t_0} - B^{t_0}}, & t_0 \neq 0 \\ \frac{\log A}{\log A - \log B}, & t_0 = 0 \end{cases} \tag{4.5.0.1}$$

where A and B are given by 4.3.0.3, also, t_0 is the *unique solution* of,

$$\mathcal{M}_{Z_1}(t_0) = \mathbb{E}_\theta(e^{t_0 Z_1}) = 1$$

$$\Leftrightarrow (\frac{\theta_0}{\theta_1})^{t_0(\theta)} = 1 - t_0(\theta)\theta(\frac{1}{\theta_0} - \frac{1}{\theta_1}) \tag{4.5.0.2}$$

$$\Leftrightarrow \theta = 1 \Big/ \left(\frac{t_0(\theta)(1/\theta_0 - 1/\theta_1)}{1 - (\frac{\theta_0}{\theta_1})^{t_0(\theta)}} \right) \tag{4.5.0.3}$$

4.6 WALD'S Approximate ASN Function:

The *ASN Function* is given by,

$$\mathbb{E}_\theta(\mathfrak{N}) \approx \begin{cases} \frac{b\cdot\mathfrak{L}(\theta)+a\cdot(1-\mathfrak{L}(\theta))}{\mathbb{E}_\theta(Z_1)}, & \mathbb{E}_\theta(Z_1) \neq 0 \\ \frac{b^2\cdot\mathfrak{L}(\theta)+a^2(1-\mathfrak{L}(\theta))}{V(Z_1)}, & \mathbb{E}_\theta(Z_1) = 0 \end{cases} \tag{4.6.0.1}$$

where, $\mathbb{E}_\theta(Z_1) = \log(\frac{\theta_0}{\theta_1}) + \theta(1/\theta_0 - 1/\theta_1)$ and $V_\theta(Z_1) = \theta^2(1/\theta_0 - 1/\theta_1)^2$.

4.7 Results - OC & ASN Curves:

Here, the results on the *Wald's Approximate OC & ASN Functions* and *corresponding parameters*, as given by 4.5.0.2, 4.5.0.3, 4.5.0.1 and 4.6.0.1 respectively.

$t_0(\theta)$	$\mathfrak{L}(\theta)$	θ	$\mathbb{E}_\theta(Z_1)$	$V_\theta(Z_1)$	$\mathbb{E}_\theta(\mathfrak{N})$
−0.010	0.511	4.458	−0.0002	0.049	424.298
−0.100	0.613	4.413	−0.0025	0.048	419.855
0.200	0.715	4.365	−0.0049	0.047	402.454
0.500	0.909	4.223	−0.0119	0.044	313.038
0.700	0.961	4.132	−0.0165	0.042	256.175
0.800	0.975	4.087	−0.0188	0.041	232.557
1.000	0.990	4.000	−0.0231	0.040	194.577

Table 6: *Results for Wald's OC & ASN Functions*

The OC and ASN curves are illustrated in the following diagram, i.e., Figure 8.

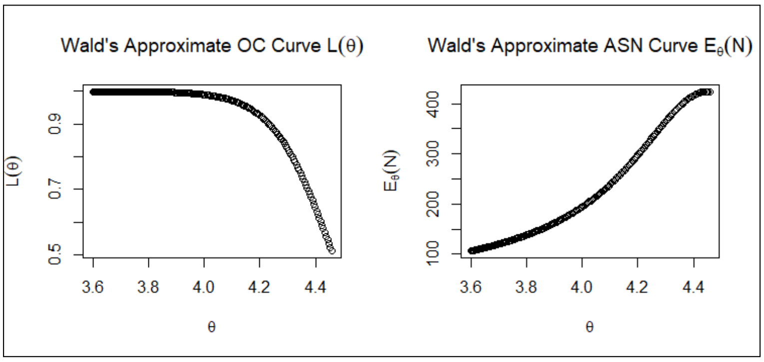

Figure 8: *Wald's Approximate OC and ASN Curves against* θ