

IMPACT OF CONTROLLING THE SUM OF ERROR PROBABILITY IN THE SEQUENTIAL PROBABILITY RATIO TEST

Bijoy Kumar Pradhan

1. INTRODUCTION

Pradhan (1971) has considered a modified sequential probability ratio test (SPRT) with the assumption that the sum of the error probability does not exceed a pre-assigned bound. Patel and Dharmadhikari (1974) found a method on which the average of the two average sample numbers (ASN) is minimized by giving the same mass $\frac{1}{2}$ to the null hypothesis as well as to the alternative hypothesis under the condition that the sum of the error probabilities equals a specified constant. Wald (1974) has compared the efficiency of a sequential test S of strength (α, β) by the ratio

$$\frac{n_0(\alpha, \beta)}{E_{\theta_0}(n|s)}$$

when H_0 is true and by

$$\frac{n_0(\alpha, \beta)}{E_{\theta_1}(n|s)}$$

when H_1 is true and calculated the number of observations that can be saved by adopting sequential probability ratio test instead of using current test procedure for testing mean of a normally distributed variate and found that the sequential test results in an average saving of at least 47 percent in the necessary number of observations as compared with the current test.

In the present context a generalized modified method is proposed to minimized the weighted average of the two average sample numbers $E_{\theta_0}(n)$ and $E_{\theta_1}(n)$ by attaching relative weights a_1 and a_2 respectively with $a_1 + a_2 = 1$ such that the sum of error probabilities in the sequential probability ratio test is a pre-assigned constant i.e., $\alpha + \beta = k$, $0 < k < 1$ with the intention to minimized the expected value of the required number of observations simultaneously for all θ

to get uniformly best test which is admissible. We shall compare the expected number of observations required by the generalized modified sequential procedure of strength (α, β) such that $\alpha + \beta = k$ (a preassigned constant) for testing H_0 against H_1 with the fixed number of observations needed for the current most powerful test to attain the same strength (α, β) . The results are applied to the case when the random variate X follows a normal law as well as Bernoulli law. When the random variate X follows a normal law, we find that for $\alpha + \beta = k$ with $\alpha = \beta$ from 0.01 to 0.05 (the range most frequently employed), the generalized modified sequential procedure results in an average saving of at least 51% in the necessary number of observations as compared with the current most powerful test having same strength (α, β) . But when the random variate X follows Bernoulli law, we find that for $\alpha + \beta = k$ (a preassigned constant) with $\alpha \neq \beta$, the generalized modified sequential procedure results in an average saving of at least 51% in the necessary number of observations as compared with the current most powerful test procedure having same strength. In both cases the gain is a decreasing function of k . Finally, we provide an example on the real data which also shows the same sort of results.

2. THE GENERALIZED MODIFIED METHOD

Let $f(x, \theta)$ denote the probability function (or probability density function) of the random variable X where the parameter θ is unknown. Suppose we have a test procedure of strength (α, β) for testing a simple null hypothesis $H_0 : \theta = \theta_0$ against a simple alternative hypothesis $H_1 : \theta = \theta_1 (> \theta_0)$.

Suppose x_1, x_2, \dots , etc be the successive observations of the random variable X . For any positive integral value m , the probability that a sample x_1, x_2, \dots, x_m is obtained is given by

$$p_{1m} = \prod_{i=1}^m f(x_i, \theta_1) \quad (2.1)$$

when H_1 is true.

$$p_{0m} = \prod_{i=1}^m f(x_i, \theta_0) \quad (2.2)$$

when H_0 is true.

Let

$$Z_i = \log \frac{f(x_i, \theta_1)}{f(x_i, \theta_0)}. \quad (2.3)$$

Hence

$$\sum_{i=1}^m Z_i = \log \frac{\hat{P}_{1m}}{\hat{P}_{0m}}. \quad (2.4)$$

The sequential probability ratio test for testing H_0 against H_1 is described by Wald (1974) as follows:

Two positive constants A and B ($B < A$) are chosen. At each stage of the experiment (at the m^{th} trial for each integral value of m), the sum $\sum Z_i$ is computed.

If $\sum Z_i \geq \log A$, the process is terminated with the rejection of H_0 .

If $\sum Z_i \leq \log B$, the process is terminated with the acceptance of H_0 .

If $\log B < \sum Z_i < \log A$, we continue the experiment by taking an additional observations.

Wald (1974) has shown that boundaries A and B of the sequential probability ratio test have the approximations:

$$A \cong \frac{1-\beta}{\alpha} \quad \text{and} \quad B \cong \frac{\beta}{1-\alpha} \quad (2.5)$$

Suppose n denote the number of observations required by the sequential tests defined by Wald (1974) to reach a decision and further suppose that $E_{\theta}(n|s)$ denote the expected value of n when θ is the true value of the parameter of a sequential test S . Wald (1974) has shown that the following approximations hold:

$$E_{\theta_0}(n|s) = \frac{(1-\alpha) \log B + \alpha \log A}{E_{\theta_0}(Z)} \quad \text{and}$$

$$E_{\theta_1}(n|s) = \frac{\beta \log B + (1-\beta) \log A}{E_{\theta_1}(Z)} \quad (2.6)$$

Wald (1974) has compared the expected number of observations required by the sequential probability ratio test of strength (α, β) for testing H_0 against H_1 with the fixed number of observations $n(\alpha, \beta)$ needed for the current most powerful test to attain the same strength (α, β) . Since the average saving of the sequential test as compared with the current test is

100 $\left[1 - \frac{E_{\theta_1}(n|s)}{n(\alpha, \beta)} \right]$ percent if H_1

is true and 100 $\left[1 - \frac{E_{\theta_0}(n|s)}{n(\alpha, \beta)} \right]$ percent if H_0 is true.

Wald (1974) has found that the sequential test analysis results in an average saving of at least 47 percent in the necessary number of observations as com-

pared with the current test procedure when the observations are taken from a normal population with unknown mean θ and variance unity.

We know that if an admissible sequential test S exists for which the expected value of the number of observations is minimized for all θ , then that test may be regarded a uniformly best test. Keeping view on this point to get an uniformly best test which is admissible, a generalized modified method is proposed to the two average sample numbers $E_{\theta_0}(n)$ and $E_{\theta_1}(n)$ of strength (α, β) by attaching weights a_1 and a_2 respectively with $a_1 + a_2 = 1$ and by controlling the sum of error probabilities in the sequential probability ratio test. Then a comparison is made with the current fixed sample size procedure.

Hence, the present method is to minimize

$$a_1 E_{\theta_0}(n) + a_2 E_{\theta_1}(n) \quad (2.7)$$

subject to constraints

$$a_1 + a_2 = 1; \quad a_i > 0, \quad i = 1, 2;$$

and $\alpha + \beta = k$;

where k is a pre-assigned constant such that $0 < k < 1$.

Here, we assume $\alpha + \beta = k$, $0 < k < 1$, so that $0 < B < 1 < A < \infty$.

Appendix-A shows that the minimization of (2.7) holds good if

$$a_1 = \frac{C_1 \left[\log \frac{A}{B} - (A - B) \right]}{C_1 \left[\log \frac{A}{B} - (A - B) \right] + C_2 \left[\log \frac{A}{B} - \frac{A - B}{AB} \right]}$$

and

$$a_2 = 1 - a_1; \quad (2.8)$$

subject to

$$\frac{(1 - \alpha) \log B + \alpha \log A}{E_{\theta_0}(Z)} = \frac{\beta \log B + (1 - \beta) \log A}{E_{\theta_1}(Z)} \quad (2.9)$$

Calculate $E_{\theta_0}(Z)$ and $E_{\theta_1}(Z)$. Fix α and find β for which (2.9) is satisfied. Substitute this value of α and β in (2.8) which will give the optimum values of a_1 and a_2 i. e. $a_{1(opt)}$ and $a_{2(opt)}$.

$$\text{Hence } E(n) = n^{**}(\text{say}) = a_{1(opt)} E_{\theta_0}(n) + a_{2(opt)} E_{\theta_1}(n) \quad (2.10)$$

3. COMPARISON

It is required to compare the optimum sample size n^{**} given by (2.10) of the sequential probability ratio test of strength (α, β) , $\alpha + \beta = k$ with the corresponding fixed sample size procedure having the same strength (α, β) . The optimum sample size for the fixed sample size procedure is based on uniformly most powerful test. The procedure is as follows.

Suppose sample x_1, x_2, \dots, x_n of n independent observations of the random variable X is available. Then a test statistic is constructed in order to test a simple null hypothesis $H_0 : \theta = \theta_0$ against a simple alternative hypothesis $H_1 : \theta = \theta_1$ based on the corresponding strength (α, β) of the sequential test procedure. If $L(x_1, x_2, \dots, x_n; \theta_1)$ denote the likelihood function, then the procedure is to reject H_0 if

$$L(x_1, x_2, \dots, x_n; \theta_1) > k_0 L(x_1, x_2, \dots, x_n; \theta_0) \quad (3.1)$$

and to accept H_0 , otherwise; see Rao (1952), where k_0 is so determined that the probability of type I error = α . This is same as to find a critical region W_0 given by

$$W_0 = \{(x_1, \dots, x_n) : \bar{x} > k_0\} \quad (3.2)$$

$$\text{such that } \Pr[\bar{x} > k_0 | H_0] = \alpha \quad (3.3)$$

Hence

$$\Pr[\bar{x} > k_0 | H_1] = 1 - \beta \quad (3.4)$$

From (3.3) and (3.4) we can able to find the optimum sample size n^* for the fixed sample size procedure.

Then the average percentage of saving in the number of observations due to the use of the present generalized modified method of the sequential test procedure over the corresponding fixed sample size procedure is

$$100 \left(1 - \frac{n^{**}}{n^*} \right). \quad (3.5)$$

Where n^{**} is the optimum sample size for the sequential probability ratio test of strength (α, β) with $\alpha + \beta = k$ and is given by (2.10).

4. APPLICATIONS

Case I: When the observations follow the normal law.

Let the random variable X follows a normal law with unknown mean θ and

variance unity. Consider the problem of testing a simple null hypothesis $H_0 : \theta = \theta_0$ against a simple alternative hypothesis $H_1 : \theta = \theta_1 (> \theta_0)$.

For this case

$$E_{\theta_1}(Z) = \frac{1}{2}(\theta_0 - \theta_1)^2 \text{ and } E_{\theta_0}(Z) = -\frac{1}{2}(\theta_0 - \theta_1)^2 \quad (4.1)$$

For the fixed sample size procedure given by (3.1), we have

$$n^* = n(\alpha, \beta) = \frac{(\lambda_1 - \lambda_0)^2}{(\theta_0 - \theta_1)^2}, \quad (4.2)$$

where $G(\lambda_0) = 1 - \alpha$ and $G(\lambda_1) = \beta$ and $G(t)$ is the probability that a normally distributed random variable with mean zero and variance unity will take a value less than t i.e.

$$G(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{x^2}{2}} dx \quad (4.3)$$

Note: the probability of Type I error α and the probability of Type II error β for the current test procedure were found firstly by the generalized modified sequential procedure of strength (α, β) with $\alpha + \beta = k$. Then using those (α, β) with $\alpha + \beta = k$, we find the optimum sample size for the current test procedure.

From (4.1) we find $C_1 = C_2$ and hence equation (2.9) is satisfied for $\alpha = \beta$.

The optimum values of a_1 and a_2 are found by substituting $C_1 = C_2$ and $\alpha = \beta$ in (2.8) which gives

$$a_{1(opt)} = a_{2(opt)} = \frac{1}{2} \quad (4.4)$$

Hence, we have

$$\begin{aligned} n^{**} &= \frac{1}{2} E_{\theta_0}(n) + \frac{1}{2} E_{\theta_1}(n) \\ &= \frac{1}{2} \cdot \frac{(1 - \alpha) \log B + \alpha \log A}{-(\theta_0 - \theta_1)^2} + \frac{1}{2} \cdot \frac{\beta \log B + (1 - \beta) \log A}{(\theta_0 - \theta_1)^2} \end{aligned} \quad (4.5)$$

The gain in the average number of observations given by the sequential probability ratio test as compared with current most powerful test is $100 \left(1 - \frac{n^{**}}{n^*} \right)$.

Table 1 shows for the range α and β with $\alpha + \beta = k$, $\alpha = \beta$ from 0.01 to 0.05 (the range most frequently employed), the sequential test results in an average

saving of at least 51% in the necessary number of observations as compared with the current test. For $\alpha = \beta$ with $\alpha + \beta = k$, the gain is a decreasing function of k .

TABLE 1
Average percentage of size of sample with the generalized modified SPRT analysis as compared with current most powerful test for testing mean of the normally distributed variate for $\alpha + \beta = k$ with $\alpha = \beta$

	$\alpha=0.01$ $\beta=0.01$	$\alpha=0.02$ $\beta=0.02$	$\alpha=0.03$ $\beta=0.03$	$\alpha=0.04$ $\beta=0.04$	$\alpha=0.05$ $\beta=0.05$
k	0.02	0.04	0.06	0.08	0.10
Gain (%)	58.454	55.722	53.874	52.319	51.035

Case II: When the observations follow the Bernoullian law.

Let the random variable X follows a Bernoullian law i.e. $X \sim b(1, p)$, where p is unknown. Consider a problem of testing a simple null hypothesis $H_0 : p = p_0$ against a simple alternative hypothesis $H_1 : p = p_1 (> p_0)$.

Suppose a sample (x_1, x_2, \dots, x_n) of n independent observations of the random variable X is available. Then a test statistic may be constructed to test $H_0 : p = p_0$ against $H_1 : p = p_1 (> p_0)$ for α and β with $\alpha + \beta = k$ (a pre-assigned constant), $0 < k < 1$.

We know that for sufficiently large sample, n , the distribution of the sum $\sum_{i=1}^n X_i$ may be regarded as normally distributed when $X_i, i=1,2, \dots, n$ follows a Bernoullian law.

Hence, we have

$$p_r \left[\frac{\sum X_i - np_1}{\sqrt{np_1q_1}} \leq \frac{d - np_1}{\sqrt{np_1q_1}} \right] = \beta \text{ and } p_r \left[\frac{\sum X_i - np_0}{\sqrt{np_0q_0}} \leq \frac{d - np_0}{\sqrt{np_0q_0}} \right] = 1 - \alpha \tag{4.6}$$

$$\text{Let } \lambda_1(n) = \frac{d - np_1}{\sqrt{np_1q_1}} \text{ and } \lambda_0(n) = \frac{d - np_0}{\sqrt{np_0q_0}} \tag{4.7}$$

Firstly find the optimum sample size n^{**} of sequential test procedure of strength α and β with $\alpha + \beta = k$. Then for those α and β find the value of $\lambda_1(n)$ and $\lambda_0(n)$ and hence solving (4.7) for n , we can able to find the value of n , say n^* , the required optimum sample size, for the current most powerful procedure.

The optimum sample size adopting sequential probability ratio test analysis for testing $H_0 : p = p_0$ against $H_1 : p = p_1 (> p_0)$ of strength (α, β) with $\alpha + \beta = k$ is given by (see Appendix – B).

$$n^* = a^*_{1(opt)} E_{p_0}(n) + a^*_{2(opt)} E_{p_1}(n) = \frac{(1 - \alpha) \log \frac{\beta}{1 - \alpha} + \alpha \log \frac{1 - \alpha}{\beta}}{p_0 \log \frac{p_1}{p_0} + (1 - p_0) \log \frac{1 - p_1}{1 - p_0}} \tag{4.8}$$

Hence the average percentage of saving in the number of observations due to the use of the present generalized modified method sequential test procedure over the corresponding fixed sample size procedure is $100\left(1 - \frac{n^{**}}{n^*}\right)$.

Suppose, we are required to test $H_0 : p = p_0 = 0.6$ against $H_1 : p = p_1 = 0.7$. Here,

$$E_{p_0}(\bar{z}) = p_0 \log \frac{p_1}{p_0} + (1 - p_0) \log \frac{1 - p_1}{1 - p_0} = -0.0098074$$

and

$$E_{p_1}(\bar{z}) = p_1 \log \frac{p_1}{p_0} + (1 - p_1) \log \frac{1 - p_1}{1 - p_0} = 0.0093811.$$

For given α , it is required to find the value of β satisfying (19) of Appendix B by which we can find the optimum value of n^{**} given by (4.8).

The optimum value of n^{**} for $\alpha + \beta = k$ is given below in Table 2.

TABLE 2
The optimum value of n^{**} for SPRT procedure for testing $H_0 : p = p_0 = 0.6$ against $H_1 : p = p_1 = 0.7$

	$\alpha = 0.01$ $\beta = 0.007972$	$\alpha = 0.02$ $\beta = 0.0163024$	$\alpha = 0.03$ $\beta = 0.024727$	$\alpha = 0.04$ $\beta = 0.033194$	$\alpha = 0.05$ $\beta = 0.041681$
n^{**}	209	174	153	137	125

From Table 2 we see that the optimum sample size exists for sequential probability ratio test procedure for $\alpha \neq \beta$ and β close to α .

The optimum sample size for the current test procedure for α and β as mentioned in Table 2 is given in Table 3.

TABLE 3
The optimum value of n^* for fixed sample size test procedure for testing $H_0 : p = p_0 = 0.6$ against $H_1 : p = p_1 = 0.7$.
 $H_0 : p = p_0 = 0.6$ against $H_1 : p = p_1 = 0.7$ for $\alpha + \beta = k, \alpha \neq \beta$

	$\alpha = 0.01$ $\beta = 0.007972$	$\alpha = 0.02$ $\beta = 0.0163024$	$\alpha = 0.03$ $\beta = 0.024727$	$\alpha = 0.04$ $\beta = 0.033194$	$\alpha = 0.05$ $\beta = 0.041681$
n^*	514	394	332	288	256

The gain in the average number of observations given by the sequential probability ratio test procedure over the corresponding most powerful test for testing

$H_0 : p = p_0 = 0.6$ against $H_1 : p = p_1 = 0.7$ is $100\left(1 - \frac{n^{**}}{n^*}\right)$ and is given in Table

4 given below.

Table 4 shows that the sequential test results in an average saving of at least 51% in the necessary number of observations as compared with the current most powerful test procedure for testing $H_0 : p = p_0 = 0.6$ against $H_1 : p = p_1 = 0.7$ for $\alpha + \beta = k, \alpha \neq \beta$. Here the gain is a decreasing function of k .

TABLE 4

Average percentage saving of size of sample with the generalized modified SPRT analysis compared with the current most powerful test procedure for testing H_0 against H_1 for $H_0: p = p_0 = 0.6$ against for $H_1: p = p_1 = 0.7$ for $\alpha + \beta = k, \alpha \neq \beta$

	$\alpha = 0.01$ $\beta = 0.007972$	$\alpha = 0.02$ $\beta = 0.0163024$	$\alpha = 0.03$ $\beta = 0.024727$	$\alpha = 0.04$ $\beta = 0.033194$	$\alpha = 0.05$ $\beta = 0.041681$
k	0.017972	0.0363024	0.054727	0.073194	0.091681
Gain (%)	59.338	56.837	53.916	52.430	51.172

EXAMPLE

A manufacturing plant produces articles with average quality characteristic 135 on the basis of the following criterion.

1. The producer is not willing to run the process having the rejection greater than $\alpha\%$.
2. The purchaser is not willing to accept the lot average quality of 150 in more than $\beta\%$ cases.

Prepare a sequential plan for the acceptance inspection on the basis of the following observations having known that $\sigma = 25$ and the characteristic follows a normal law at strength $(\alpha, \beta) : (0.01, 0.01); (0.02, 0.02); (0.03, 0.03); (0.04, 0.04)$ and $(0.05, 0.05)$.

The observations are:

123 144 133 136 148 106 152 125 138 127 130 146 152 141 125 126
129 137 136 138 134 140 157 123 130

Find the average percentage saving in size of sample with sequential procedure as compared with current most powerful test for testing mean of a normally distributed variate.

We want to construct a sequential plan to test $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1 > \theta_0$, when $X_1 \dots X_n$ follows normal law with mean θ and known standard deviation σ .

For strength (α, β) and for each m , we compute the acceptance number

$$a_m = \frac{\sigma^2}{\theta_1 - \theta_0} \log \frac{\beta}{1 - \alpha} + m \frac{\theta_0 + \theta_1}{2},$$

and the rejection number

$$r_m = \frac{\sigma^2}{\theta_1 - \theta_0} \log \frac{1 - \beta}{\alpha} + m \frac{\theta_0 + \theta_1}{2}.$$

If $\sum_{i=1}^m X_i < a_m$, the sequential plan is terminated with the acceptance of H_0 .

If $\sum_{i=1}^m X_i > r_m$, the sequential plan is terminated with the rejection of H_0 .

If $a_m < \sum_{i=1}^m X_i < r_m$, the experiment is continued by taking additional observations.

The average sample number required for the proposed generalized modified sequential procedure subject to $a + \beta = k$ is

$$E(n) = n^{**} = \frac{1}{2} \frac{(1-\alpha) \log B + \alpha \log A}{\frac{(\theta_0 - \theta_1)^2}{2\sigma^2}} + \frac{1}{2} \frac{\beta \log B + (1-\beta) \log A}{\frac{(\theta_0 - \theta_1)^2}{2\sigma^2}}.$$

If the optimum sample size for the current most powerful test is n^* , then the gain in the average number of observations given by the proposed generalized modified method as compared with current most powerful test is $100 \left(1 - \frac{n^{**}}{n^*} \right)$.

Here, $\theta_0 = 135$, $\theta_1 = 150$, $\sigma = 25$.

For strength $(a, \beta) = (0.01, 0.01)$, we find $a_m = -191.4633271 + 142.5m$, and $r_m = 191.4633271 + 142.5m$.

We find for $m = 25$, $\sum_{i=1}^{25} X_i = 3366$, $a_m = 3371.037$ and hence $\sum_{i=1}^{25} X_i < a_m$.

Hence the sequential plan is terminated with the acceptance of the lot after inspecting the 25th item.

We also find $n^{**} = 25.018$ and $n^* = 60.218$.

For the strength $(a, \beta) = (.02, .02)$, we find $a_m = -162.1591791 + 142.5m$, and $r_m = 162.1591791 + 142.5m$.

We find for $m = 21$, $\sum_{i=1}^{21} X_i = 2826$, $a_m = 2830.341$ and hence $\sum_{i=1}^{21} X_i < a_m$.

Hence, sequential sampling plan is terminated with the acceptance of the lot after inspecting the 21st item.

We also find $n^{**} = 20.756$ and $n^* = 46.877$.

For strength $(a, \beta) = (.03, .03)$, we find $a_m = -144.8374454 + 142.5m$, and $r_m = 144.8374454 + 142.5m$.

We find that for $m = 18$, $\sum_{i=1}^{18} X_i = 2418$, $a_m = 2420.163$ and hence $\sum_{i=1}^{18} X_i < a_m$.

Hence, sequential sampling plan is terminated with the acceptance of the lot after inspecting the 18th item.

We also find $n^{**} = 18.153$ and $n^* = 39.355$.

For strength $(a, \beta) = (0.04, 0.04)$, we find $a_m = -132.4189096 + 142.5m$,
and $r_m = 132.4189096 + 142.5m$.

We find that for $m = 17$, $\sum_{i=1}^{17} X_i = 2281$, $a_m = 2290.081$ and hence $\sum_{i=1}^{17} X_i < a_m$.

Hence, sequential sampling plan is terminated with the acceptance of the lot after inspecting the 17th item.

We also find $n^{**} = 16.243$ and $n^* = 34.066$.

For strength $(a, \beta) = (0.05, 0.05)$, we find $a_m = -122.6849575 + 142.5m$,
and $r_m = 122.6849575 + 142.5m$.

We find that for $m=16$, $\sum_{i=1}^{16} X_i = 2051$, $a_m = 2157.315$ and hence $\sum_{i=1}^{16} X_i < a_m$.

Hence, sequential sampling plan is terminated with the acceptance of the lot after inspecting the 16th item.

We also find $n^{**} = 14.722$ and $n^* = 30.066$.

TABLE 5

Sample size required for SPRT procedure and average percentage saving in size of sample with sequential procedure as compared with current most powerful test for testing mean of a normal distributed variate

	$\alpha = .01$ $\beta = .01$	$\alpha = .02$ $\beta = .02$	$\alpha = .03$ $\beta = .03$	$\alpha = .04$ $\beta = .04$	$\alpha = .05$ $\beta = .05$
k	0.02	0.04	0.06	0.08	0.10
Sample size required for S.P.R.T procedure $m = n$	25	21	18	17	16
Average Sample Number required for the proposed generalized modified sequential procedure	25.018	20.756	18.153	16.243	14.722
$E(n) = n^{**}$					
Optimum sample size for the current most powerful test n^*	60.218	46.877	39.355	34.066	30.066
Gain (%)	58.454	55.722	53.874	52.319	51.035

Table 5 shows for the range a and β with $a + \beta = k$, $a = \beta$ from 0.01 to 0.05 the sequential test results in an average saving of at least 51% in the necessary number of observations as compared with the current test. For $a = \beta$ with $a + \beta = k$, the gain is a decreasing function of k .

*Department of Statistics
Utkal University, Bhubaneswar*

BIJOY KUMARR PRADHAN

ACKNOWLEDGEMENTS

Author is thankful to the referee for many useful suggestion for improving the manuscript.

APPENDIX A

Minimize

$$a_1 E_{\theta_0}(n) + a_2 E_{\theta_1}(n) \quad (1)$$

subject to

$$a_1 + a_2 = 1; a_i > 0, i=1,2$$

and

$$\alpha + \beta = k \quad (2)$$

where $E_{\theta}(n)$ is the expected value of n when θ is the true value of the parameter, (α, β) is the required strength of the sequential probability ratio test to test the null hypothesis $H_0 : \theta = \theta_0$ against the alternative hypothesis $H_1 : \theta = \theta_1 (> \theta_0)$ and k is a pre-assigned constant such that $0 < k < 1$. We assume that $0 < k < 1$, so that $0 < B < 1 < A < \infty$ and n , the number of observations required to reach a decision.

Minimization of (1) with respect to the constraints (2) is equivalent to the minimization of

$$F = a_1 E_{\theta_0}(n) + (1 - a_1) E_{\theta_1}(n) - \lambda(\alpha + \beta - k) \quad (3)$$

with respect to α, β, a_1 and λ where λ is a Lagrange's multiplier.

As

$$E_{\theta_0}(n) = \frac{(1 - \alpha) \log B + \alpha \log A}{E_{\theta_0}(Z)}$$

and

$$E_{\theta_1}(n) = \frac{\beta \log B + (1 - \beta) \log A}{E_{\theta_1}(Z)}$$

where A and B are boundaries of sequential probability ratio test and have approximation $A \cong \frac{1 - \beta}{\alpha}, B \cong \frac{\beta}{1 - \alpha}$ and $Z = \log \left[\frac{f(x, \theta_1)}{f(x, \theta_0)} \right]$ and $E_{\theta_i}(n)$ denotes expectation when $\theta = \theta_i$ for $i=0,1$ given by Wald (1974); (3) takes the form

$$F = a_1 \left[\frac{(1 - \alpha) \log B + \alpha \log A}{E_{\theta_0}(Z)} \right] + (1 - a_1) \left[\frac{\beta \log B + (1 - \beta) \log A}{E_{\theta_1}(Z)} \right] - \lambda(\alpha + \beta - k). \quad (4)$$

Differentiating (4) with respect to α, β, a_1 , and λ and equate to zero, we find

$$\frac{\partial F}{\partial \alpha} = a_1 \left[\frac{1}{E_{\theta_0}(Z)} \log \frac{A}{B} \right] + (1 - a_1) \left[-\frac{A - B}{E_{\theta_1}(Z)} \right] - \lambda = 0; \tag{5}$$

$$\frac{\partial F}{\partial \beta} = a_1 \left[\frac{1}{E_{\theta_0}(Z)} \cdot \frac{A - B}{AB} \right] + (1 - a_1) \left[-\frac{1}{E_{\theta_1}(Z)} \log \frac{A}{B} \right] - \lambda = 0; \tag{6}$$

$$\left[\frac{(1 - \alpha) \log B + \alpha \log A}{E_{\theta_0}(Z)} \right] - \left[\frac{\beta \log B + (1 - \beta) \log A}{E_{\theta_1}(Z)} \right] = 0; \tag{7}$$

and

$$\alpha + \beta - k = 0 \tag{8}$$

From (8) we find $\alpha + \beta = k$.

Solving (5) and (6) for λ , we find

$$\begin{aligned} & a_1 \left[\left\{ \frac{1}{E_{\theta_1}(Z)} - \frac{1}{E_{\theta_0}(Z)} \right\} \log \frac{A}{B} \right] + (A - B) \left[\left\{ \frac{1}{E_{\theta_0}(Z)AB} - \frac{1}{E_{\theta_1}(Z)} \right\} \right] \\ & = \frac{1}{E_{\theta_1}(Z)} \cdot \left[\log \frac{A}{B} + (A - B) \right] \end{aligned} \tag{9}$$

Wald [(1974), Appendix A.2] has shown that $E_{\theta_0}(Z) < 0$ and $E_{\theta_1}(Z) > 0$. If we write $E_{\theta_0}(Z) = -C_1$ and $E_{\theta_1}(Z) = C_2$, where $C_1, C_2 > 0$, (9) takes the form

$$a_1 \left[(C_1 + C_2) \log \frac{A}{B} - (A - B) \left(C_1 + \frac{C_2}{AB} \right) \right] = C_1 \left[\log \frac{A}{B} - (A - B) \right], \tag{10}$$

and hence

$$a_1 = \frac{C_1 \left[\log \frac{A}{B} - (A - B) \right]}{C_1 \left[\log \frac{A}{B} - (A - B) \right] + C_2 \left[\log \frac{A}{B} - \frac{A - B}{AB} \right]},$$

and

$$a_2 = 1 - a_1 \text{ as } a_1 + a_2 = 1. \tag{11}$$

The optimum value of a_1 and a_2 are found by (10) subject to the satisfaction of (7) and (8).

Remark 1: As A and B are boundary points of the sequential probability ratio test given by Wald [1974], there is no necessity of differentiating (4) with respect to α and β by substituting the values of A and B in terms of α and β .

Remark 2: Even if substituting the values of A and B in (4) and differentiating w. r. t. α and β and solving, we find the same results as shown in (5), (6), (9), (10) and (11).

APPENDIX B

Minimized

$$a_1^* E_{p_0}(n) + a_2^* E_{p_1}(n) \quad (12)$$

subject to the constraints

$$a_1^* + a_2^* = 1; a_i^* > 0 \text{ for } i=1,2$$

and

$$\alpha + \beta = k; \quad (13)$$

where k is a pre-assigned constant such that $0 < k < 1$.

Here we assume that the random variable X follows a Bernoullian law i. e. $X \sim B(1, p)$ where p is unknown. Hence, we have

$$E_{p_1}(z) = E_{p_1} \left[\log \frac{f(x, p_1)}{f(x, p_0)} \right] = p_1 \log \frac{p_1}{p_0} + (1 - p_1) \log \frac{1 - p_1}{1 - p_0}. \quad (14)$$

$$E_{p_0}(z) = E_{p_0} \left[\log \frac{f(x, p_1)}{f(x, p_0)} \right] = p_0 \log \frac{p_1}{p_0} + (1 - p_0) \log \frac{1 - p_1}{1 - p_0}. \quad (15)$$

Since for $p = p_1$, $L(p) = \beta$, we have

$$E_{p_1}(n) = \frac{L(p) \log B + [1 - L(p)] \log A}{E_{p_1}(Z)} = \frac{\beta \log B + (1 - \beta) \log A}{E_{p_1}(Z)}. \quad (16)$$

Since for $p = p_0$, $L(p) = 1 - \alpha$, we have

$$E_{p_0}(n) = \frac{L(p) \log B + [1 - L(p)] \log A}{E_{p_0}(Z)} = \frac{(1 - \alpha) \log B + \alpha \log A}{E_{p_0}(Z)} \quad (17)$$

Wald (1974), Appendix A.2) has shown that $E_{p_0}(n) < 0$ and $E_{p_1}(n) > 0$. If we write $E_{p_0}(Z) = -C_1^*$ and $E_{p_1}(Z) = C_2^*$ then $C_1^* > 0$ and $C_2^* > 0$.

Proceeding as per Appendix A, the optimum value of a_1^* and a_2^* are found by

$$a_1^* = \frac{C_1^* \left[\log \frac{A}{B} - (A - B) \right]}{C_1^* \left[\log \frac{A}{B} - (A - B) \right] + C_2^* \left[\log \frac{A}{B} - \frac{A - B}{AB} \right]},$$

$$a_2^* = 1 - a_1^*, \quad (18)$$

subject to

$$\alpha + \beta = k$$

$$\text{and } \frac{(1 - \alpha) \log B + \alpha \log A}{\beta \log B + (1 - \beta) \log A} = \frac{p_0 \log \frac{p_1}{p_0} + (1 - p_0) \log \frac{1 - p_1}{1 - p_0}}{p_1 \log \frac{p_1}{p_0} + (1 - p_1) \log \frac{1 - p_1}{1 - p_0}}. \quad (19)$$

For $\alpha = \beta$, L.H.S. of (19) equals to -1. But the R.H.S. of (19) can not be equal to -1 for $p_1 > p_0$ or $p_1 < p_0$.

Hence the optimum values of a_1^* and a_2^* will be found only when $\alpha \neq \beta$. From (12), (13), (14), (15), (16), (17), (18) and (19) we find that

$$E(n) = n^{**}(\text{say}) = a_{1(\text{opt})}^* E_{p_0}(n) + a_{2(\text{opt})}^* E_{p_1}(n) = \frac{(1 - \alpha) \log \frac{\beta}{1 - \alpha} + \alpha \log \frac{1 - \alpha}{\beta}}{p_0 \log \frac{p_1}{p_0} + (1 - p_0) \log \frac{1 - p_1}{1 - p_0}}. \quad (20)$$

REFERENCES

- G.C. PATEL, S. W. DHARMADHIKARI, (1974), *Controlling the sum of error probabilities in the sequential probability ratio test*, "Journal of the Indian statistical Association", 12, pp. 49-57.
- M. PRADHAN, (1971), *Sequential probability ratio test modified for boundaries*, "Journal of the Indian statistical Association", 9, pp. 30-37.
- C.R. RAO, (1952), *Advanced statistical methods in Biometric Research*, Wiley, New York.
- A. WALD, (1974), *Sequential Analysis*, Wiley, New York.

SUMMARY

Impact of controlling the sum of error probability in the sequential probability ratio test

A generalized modified method is proposed to control the sum of error probabilities in sequential probability ratio test to minimize the weighted average of the two average sample numbers under a simple null hypothesis and a simple alternative hypothesis with the restriction that the sum of error probabilities is a pre-assigned constant to find the optimal sample size and finally a comparison is done with the optimal sample size found from fixed sample size procedure. The results are applied to the cases when the random variate follows a normal law as well as Bernoullian law.