

TRUNCATED SPRT FOR NON-STATIONARY PROCESSES:  
SENSITIVITY OF ASSUMPTIONS

Technical Report 83-21

Damodar Y. Golhar  
Stephen M. Pollock

Industrial and Operations Engineering Department  
The University of Michigan, Ann Arbor

TRUNCATED SPRT FOR NON-STATIONARY PROCESSES:  
SENSITIVITY OF ASSUMPTIONS

Damodar Y. Golhar  
Stephen M. Pollock

Industrial and Operations Engineering Department  
The University Of Michigan, Ann Arbor

Summary:

There is little literature on the truncated SPRT when observations have distribution parameters that change over time. We develop a truncated SPRT when the observations come from a Normal distribution with parameters that linearly increase over time. The truncation point, specified in advance, gives error probabilities within desired limits. The method is developed for two different assumptions about the non-stationary observations: i) the observations are independent and ii) the differences between successive observations are independent. The sensitivity of results to these assumptions is studied.

## 1. Introduction

Most of the literature on the truncated SPRT deals with the situation where the random variables  $X_1, X_2, \dots$ , are independent and identically distributed. In many situations, however (see Golhar (4)), the sequential observations  $X_1, X_2, \dots$ , are continuous random variables whose distribution parameters change with time, and thus form a non-stationary process. These time trends should provide additional information for efficient hypotheses testing.

Anderson (1) and Armitage (2) considered IID Normal random variables when studying the behavior of the operating characteristic and ASN functions of the truncated SPRT. Madsen (5) gave approximate stopping bounds and the truncation point using numerical integration for IID observations. He noted, however, that solving those equations recursively could be hard and, in practice, it might not be possible to obtain both the stopping bounds and the truncation point such that the resulting test would give actual error probabilities less than or equal to the desired error probabilities. Aroian and Robison (3) showed that for a small truncation point, for IID Normal random observations, actual error probabilities can be numerically computed to any desired degree of accuracy. Their method, however, becomes tedious for large truncation point values.

There is also some literature available on the untruncated SPRT for non-stationary processes. Phatarfod (6) considered Markovian dependence among discrete random variables, assuming

the same transition probability matrix at each sampling stage, and derived the expressions for the operating characteristic function and the ASN function. Siegmund (8,9) obtained an expression for the ASN function when independent random variables  $X_1, X_2, \dots$  have means  $\mu_1, \mu_2, \dots$  and variances  $\sigma_1^2, \sigma_2^2, \dots$ . Phatarfod (7) also developed relationships for the ASN and operating characteristic functions for continuous Normal random variables, with Markovian dependence, when testing the hypotheses regarding the correlation between successive observations.

Thus, the literature on the truncated SPRT is for IID random variables and the literature on non-stationary processes deals only with the untruncated SPRT. Here we consider truncation for non-stationary processes.

When using a truncated SPRT it is desirable to specify (in advance) a truncation point such that the resulting test gives the minimum expected number of observations with a constraint on desired error probabilities. We find such a truncation point when the mean and variance of the Normal sequential observations  $X_1, X_2, \dots$  increase linearly over time. This linear trend can be due to one of two possible underlying behaviors: i) the sequential observations are independent or ii) the differences between successive observations are independent. We will find appropriate truncated tests, and investigate the sensitivity of the truncated tests, to these assumptions.

Let  $f_i(X_i/W_j)$  be the density function of a random variable  $X_i$ , at time  $i$ , under the hypothesis  $W_j$ , for  $j = 0, 1$ .

Then the log-likelihood ratio at time  $i$  is

$$Z_i = \ln \left\{ \frac{f_i(X_i/W_1)}{f_i(X_i/W_0)} \right\} \quad (1)$$

If the  $X_i$ 's are independent then the log-likelihood ratio at time  $n$  is

$$\underline{Z}_n = \sum_{i=1}^n Z_i$$

Let  $\alpha_d$  and  $\beta_d$  be the desired error probabilities of type I and type II respectively. Then, Wald's (10) approximate lower and upper stopping bounds are:

$$a \approx \ln \left\{ \frac{\beta_d}{1-\alpha_d} \right\} \quad \text{and} \quad b \approx \ln \left\{ \frac{1-\beta_d}{\alpha_d} \right\} \quad (2)$$

Wald (10) proposed the following decision rule for the SPRT truncated at time  $m$ :

reject  $W_0$  if  $\underline{Z}_n \geq b$  for  $n = 1, 2, \dots, m$

accept  $W_0$  if  $\underline{Z}_n \leq a$  for  $n = 1, 2, \dots, m$

and take one more observation if

$$a < \underline{Z}_n < b \quad \text{for} \quad n = 1, 2, \dots, m-1.$$

If the experiment does not stop at or before  $m$  then

reject  $W_0$  if  $b > \underline{Z}_m > 0$

and accept  $W_0$  if  $a < \underline{Z}_m < 0$  (3)

## 2. Truncation for the SPRT when observations are independent

Assume that the random variables  $X_i$  are independent and normally distributed with unknown mean  $i\mu$  and known variance  $i\sigma^2$ .

Thus,  $X_i \sim N(i\mu, i\sigma^2)$  for all  $i = 1, 2, \dots$ . Then, from relation

(1), the log-likelihood ratio  $Z_i$  at time  $i$  is:

$$Z_i = \ln \frac{(\sigma_0)}{(\sigma_1)} - \frac{(X_i - i\mu_1)^2}{2 i \mu_1^2} + \frac{(X_i - i\mu_0)^2}{2 i \sigma_0^2}$$

We now assume\* that  $\sigma \equiv \sigma_0 = \sigma_1$  to get

$$Z_i = \frac{\mu_1 - \mu_0}{\sigma^2} X_i - i \frac{\mu_1^2 - \mu_0^2}{2 \sigma^2} \quad (4)$$

Since  $Z_i$  is a linear function of only  $X_i$ , it is independent of  $Z_k$ ,  $k \neq i$  and is normally distributed. Taking moments of  $X_i$  in (4) gives

$$Z_{i1} \sim N \left( \frac{id^2}{2}, id^2 \right)$$

and

$$Z_{i0} \sim N \left( -\frac{id^2}{2}, id^2 \right) \quad (5)$$

where  $d = \frac{\mu_0 - \mu_1}{\sigma}$ .

Since the  $Z_i$ 's are independent, the log-likelihood ratio at time  $n$  is

$$\underline{Z}_n = \sum_{i=1}^n Z_i$$

-----

\*If  $\sigma_0 \neq \sigma_1$ , the density function of  $Z_i$  becomes non-Normal and, although the details of this case can be worked out, the analysis becomes complicated and does not contribute to the general conclusions reported here.

In order to find a truncation point that gives actual error probabilities less than or equal to the desired error probabilities, we need to calculate the operating characteristic function  $L(W_j)$  when  $W_j$  is the true hypothesis. Hence we must find the probability density of  $Z_n$  given that  $a < \frac{Z_k}{n-1} < b$  for  $k = 1, 2, \dots, n-1$ . Denote by  $P_j(z, n)$  the prob  $\left\{ (Z_n < z) \bigcap_{k=1}^{n-1} (a < Z_k < b) \right\}$  where  $p_j(z, n)$  is the derivative of  $P_j(z, n)$  with respect to  $z$  for  $j = 0, 1$ .

Successive convolutions are required to calculate  $P_j(z, n)$  namely

$$p_j(z, 1) = f_1(z/W_j) \quad \text{for } j = 0, 1 \quad (6)$$

$$\text{and } p_j(z, n) = \int_a^b p_j(u, n-1) f_n(z-u/W_j) du, \quad n > 1 \text{ and } j = 0, 1 \quad (7)$$

where  $f_n(z/W_j)$  is a Normal density function at time  $n$  with mean  $\frac{nd^2}{2}$  ( $-\frac{nd^2}{2}$ ) and variance  $nd^2$  when the hypothesis  $W_1(W_0)$  is true.

Using these relationships we can calculate

$$L(W_j, m) = \sum_{n=1}^{m-1} \int_{-\infty}^a p_j(z, n) dz + \int_{-\infty}^0 p_j(z, m) dz \quad (8)$$

$$E(N/W_j, m) = \sum_{n=1}^{m-1} n \left\{ \int_{-\infty}^a p_j(z, n) dz + \int_b^{\infty} p_j(z, n) dz \right\} + m \int_a^b p_j(z, m-1) dz. \quad (9)$$

Let  $\alpha_a, \beta_a$ , and  $\alpha_d, \beta_d$  be the actual and desired error probabilities. Also, let  $m^*$  be the non-integer truncation point such that the SPRT truncated at  $m^*$  gives  $\alpha_a = \alpha_d$  and  $\beta_a = \beta_d$ . The integer truncation point  $m^{**}$  will be obtained by rounding  $m^*$

up to the next higher integer.

We can now establish a relationship between  $m^*$ , the desired error probabilities, and the discrimination factor  $d$  for the symmetric case (i.e., for  $\alpha_d = \beta_d$ ), by means of the following procedure:

- i) Given  $\alpha \equiv \alpha_d = \beta_d$ , Wald's constant stopping bounds  $a$  and  $b$  are computed by means of equation (2).
- ii) For a given value of  $d$ , using equations (6) through (8),  $\alpha_a(m)$  ( $=\beta_a(m)$ ) are computed, for different truncations  $m$ , by carrying out the numerical integration.
- iii) The value of  $m$  for which  $\alpha_a(m) = \alpha_d$  is found (by interpolation, if necessary), and is, by definition,  $m^*$

Figure 1, shows  $\ln(m^*)$  vs.  $\ln(d)$ . An approximately linear relationship between  $\ln(m^*)$  and  $\ln(d)$  is immediately apparent. A common slope ( $-1.09$ ) for  $.01 < \alpha < .2$  can be obtained by a linear regression. This suggests that  $m^*$  and  $d$  have the following relationship:

$$\ln(m^*) \cong \ln[k(\alpha)] - 1.09 \ln(d)$$

where  $k(\alpha)$  is a constant and depends upon the value of  $\alpha$ . To obtain  $k(\alpha)$ ,  $m^*$  was plotted against  $\alpha$  for  $d=1$ , as shown in figure 2. This curve is well fit by the equation:

$$k(\alpha) \cong 11.57 - 13 (\alpha)^2$$

Hence, the relationship between  $\ln(m^*)$  and  $\ln(d)$  can be rewritten as:

$$\ln(m^*) \cong \ln(11.57 - 13(\alpha)^2) - 1.09 \ln(d) \quad (10)$$

The smallest integer truncation point  $m^{**}$  is now found by



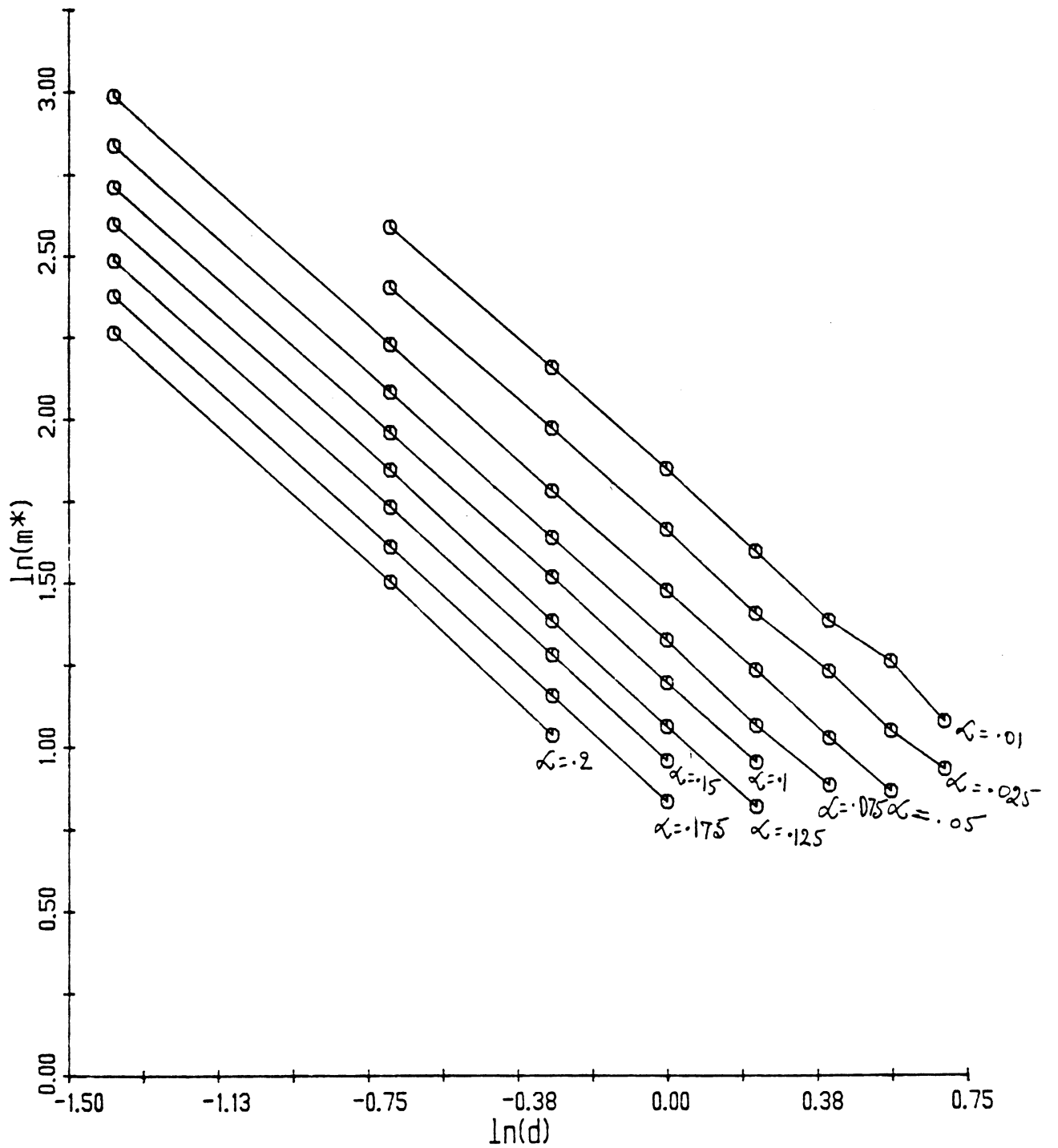


Figure 1 - The relationship between  $\ln(m^*)$  and  $\ln(d)$  for independent observations.

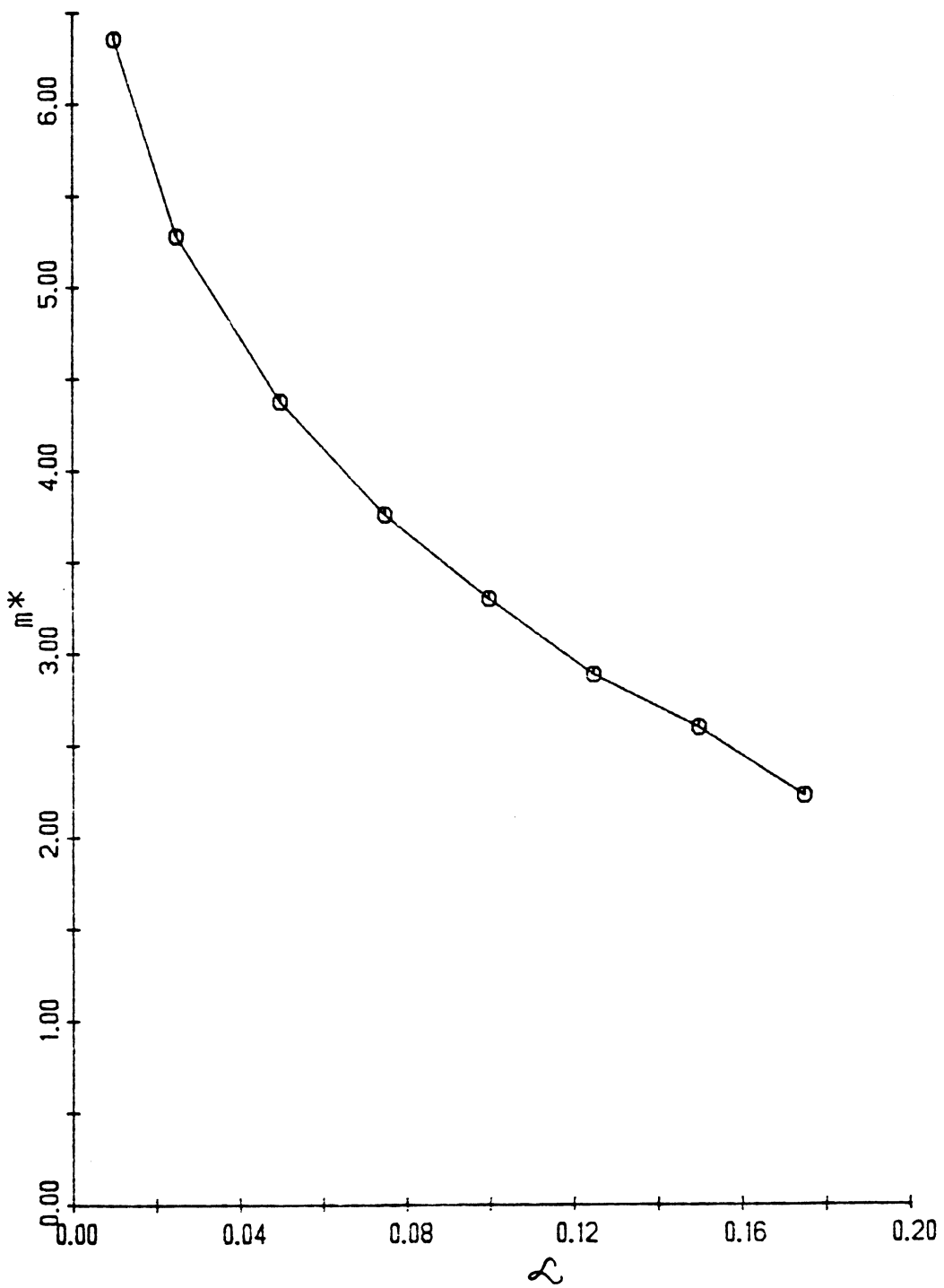


Figure 2 - The relationship between  $m^*$  and  $\alpha$  when  $d=1$  for independent observations.

rounding up the solution to (10). The resulting test will then give actual error probabilities not greater than the desired error probabilities, at the expense of a slight increase in the maximum, and average, sample size.

### 3. Truncation for the SPRT when increments are independent:

In this section we assume that the differences (increments) between successive observations are independent. This underlying behavior can also lead to a linear time dependence of observation means and variances. Thus, we assume that the variables  $X_1, X_2 - X_1, \dots, X_i - X_{i-1}, \dots$  are independent and identically distributed. Let  $g_i(X_1, \dots, X_i / W_j)$  denote the joint density function of observations  $X_1, \dots, X_i$  at time  $i$  when the hypothesis  $W_j$  is true, for  $j = 0, 1$ . Then, the log-likelihood ratio at time  $n$  is

$$Z_n = \ln \left\{ \frac{g_n (X_1, X_2, \dots, X_n / W_1)}{g_n (X_1, X_2, \dots, X_n / W_0)} \right\}$$

By transformation we get,

$$Z_n = \ln \left\{ \frac{g_n (X_1, X_2 - X_1, \dots, X_n - X_{n-1} / W_1)}{g_n (X_1, X_2 - X_1, \dots, X_n - X_{n-1} / W_0)} \right\}$$

since the Jacobian of the transformation is the determinant of an upper triangular matrix with one's along the diagonal. Defining the log-likelihood ratios:

$$Z_1 = \ln \left\{ \frac{g_1 (X_1 / W_1)}{g_1 (X_1 / W_0)} \right\}$$

and  $Z_i = \ln \left\{ \frac{g_2 (X_i - X_{i-1} / W_1)}{g_2 (X_i - X_{i-1} / W_0)} \right\}$  for all  $i > 2$

We get,  $Z_n = Z_1 + \sum_{i=2}^n Z_i$

If we assume that  $X_1 \sim N(\mu_j, \sigma^2)$  for  $j = 0, 1$  and

$X_i - X_{i-1} \sim N(\gamma_j, \delta^2)$  for  $j = 0, 1$  and  $i > 2$  then the mean and variance increase linearly over time with  $\gamma_j$  and  $\delta^2$  respectively.

To find  $m^*$  a numerical integration procedure can be carried out similar to that outlined in section 2. Instead of  $d$ , there are two parameters  $d_1$  and  $d_2$ ,  $d_1 = \frac{\mu_0 - \mu_1}{\sigma}$  and  $d_2 = \frac{\gamma_0 - \gamma_1}{\delta}$ . We chose  $d_1 = .5d_2$ ,  $d_1 = d_2$  and  $d_1 = 2d_2$  to study the relationship between  $m^*$  and  $d_2$  (or, as it turns out, between  $\ln(m^*)$  and  $\ln(d_2)$ ).  $\alpha$  was varied between .01 to .2 and  $d_2$  between .25 to 2.

Note that when  $d_1 = d_2 = d$  the truncated SPRT for IID Normal  $X_i$ 's is a special case of the truncated SPRT for independent increments. In this case we get the linear relationship between  $\ln(m^*)$  and  $\ln(d)$  shown by Golhar (4):

$$\ln(m^*) \cong \ln(-79 + 72(\alpha)^{-.079}) - 2.09 \ln(d) \quad (11)$$

Figure 3 and figure 4 show  $\ln(m^*)$  vs.  $\ln(d_2)$  when  $d_1 = .5d_2$  and  $d_1 = 2d_2$  respectively. It is seen that the relationship between  $\ln(m^*)$  and  $\ln(d_2)$  is non-linear. Some reasons for this are explained by Golhar (4).

Thus, under two different independence assumptions about the sequence of non-stationary random variables  $X_1, X_2, \dots$ , we have obtained truncation points that give actual error probabilities not greater than desired error probabilities.

One question that immediately follows is: how sensitive are these tests, in terms of  $m^{**}$  and the resulting  $\alpha_a, \beta_a$ , and

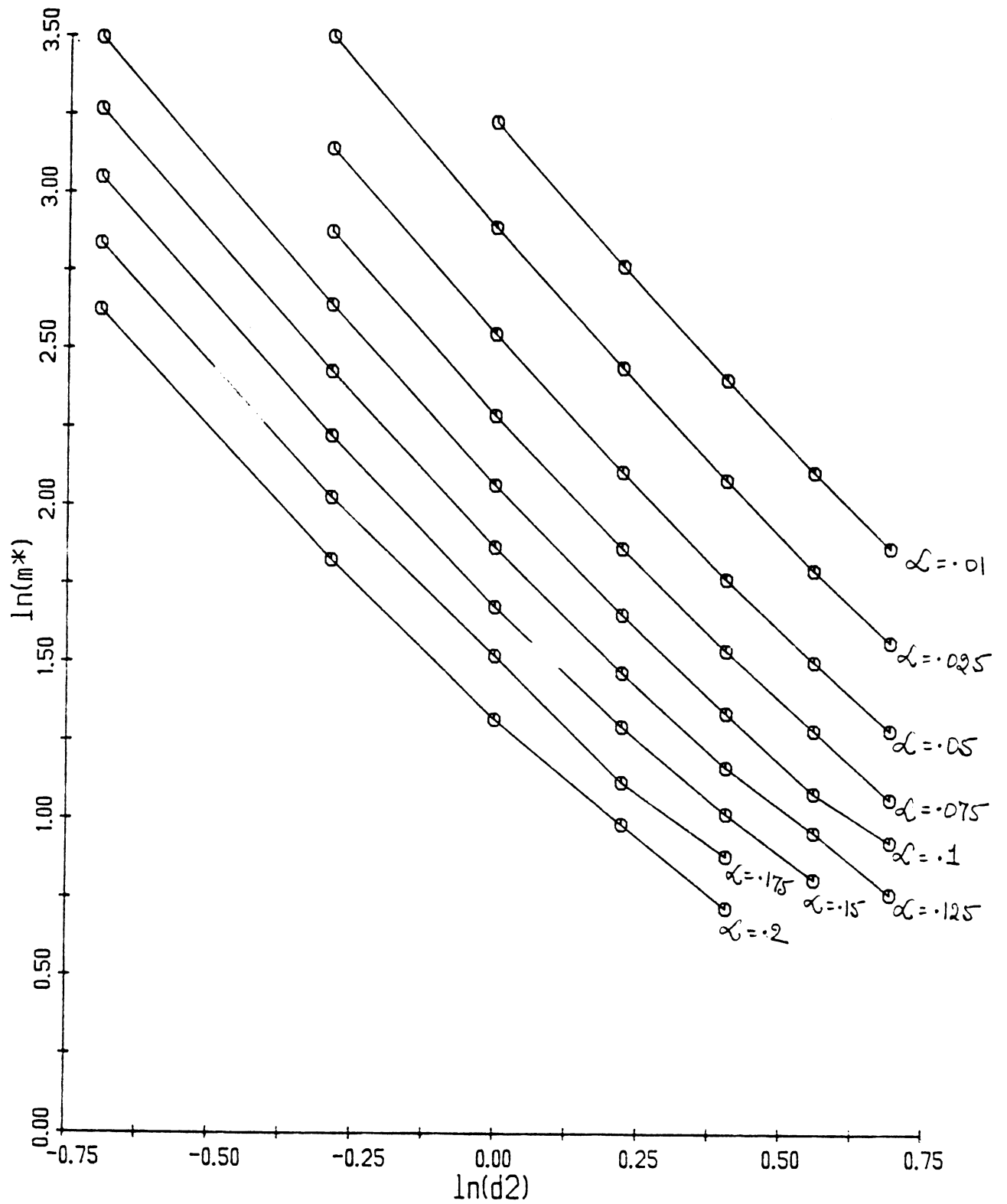


Figure 3 - The relationship between  $\ln(m^*)$  and  $\ln(d_2)$  for independent increments when  $d_1 = .5d_2$

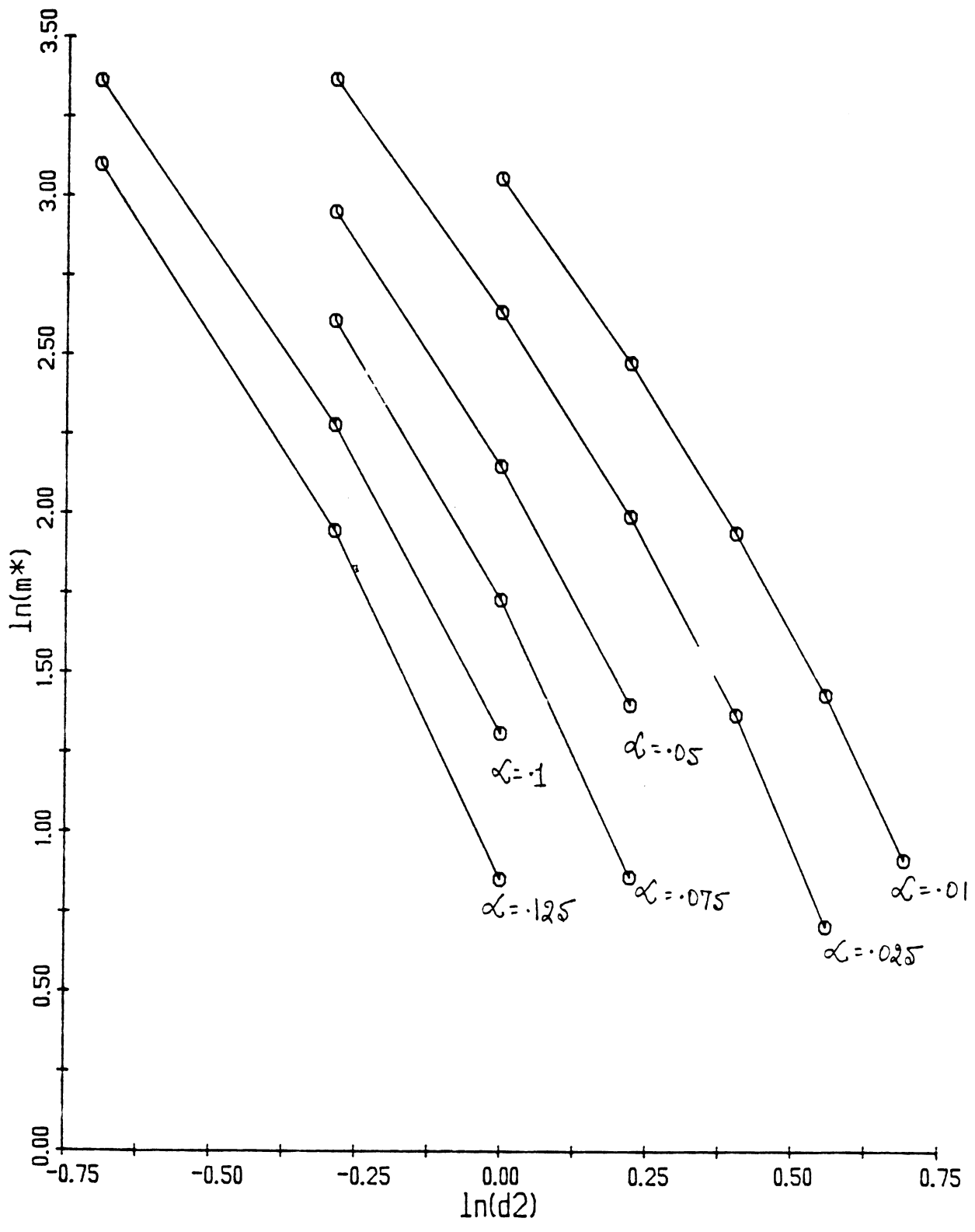


Figure 4 - The relationship between  $\ln(m^*)$  and  $\ln(d_2)$  for independent increments when  $d_1 = 2d_2$

E(N) to the assumptions involved? Since it might be time consuming and/or expensive to verify which model is actually governing the observations, there might exist a range of parameters for which the test derived using the assumptions of one model might be superior to that of the test derived using the other, in the sense that it gives smaller  $\alpha_a = \beta_a$  or E(N) or m\*\*. In the next section we examine this possibility.

#### 4. Sensitivity of independence assumptions:

We have seen that when the marginal mean and variance of a sequence of random variables increases linearly with time, one of the following two assumptions could describe the underlying behavior: i) the observations are independent or ii) increments are independent. The former model we will refer to as IO, the later as II. Thus, for II

$$\begin{aligned} X_j &\sim N(\mu_j, \sigma^2) \quad \text{for } j = 0, 1 \\ X_i - X_{i-1} &\sim N(\mu_j, \sigma^2) \quad \text{for } j = 0, 1 \quad \text{and } i > 2 \end{aligned}$$

and for IO

$$X_i \sim N(i\mu_j, i\sigma^2) \quad \text{for } j = 0, 1.$$

##### A) Assume II but in reality IO:

If II is assumed then the log-likelihood ratio is computed to be:

$$Z_n = \sum_{i=1}^n Z_i = \frac{\mu_1 - \mu_0}{\sigma^2} X_n - \frac{n(\mu_1^2 - \mu_0^2)}{2\sigma^2} \quad (12)$$

and these values of  $Z_n$  will be compared to the thresholds a and b.

However, if the  $X_i$ 's are in reality independent and normally distributed with mean  $i\mu_j$  and variance  $i\sigma^2$ , then,  $Z_1, Z_2, \dots$ ,

$\underline{Z}_n$  will also be independent Normal random variables. Taking moments of (12) we obtain,

$$Z_{n1} \sim N\left(\frac{nd^2}{2}, nd^2\right)$$

$$Z_{n0} \sim N\left(-\frac{nd^2}{2}, nd^2\right)$$

Since we assume II, the truncation point  $m^{**}$  would be obtained from relation (11). However, the values of  $E(N/W_j, m^{**})$  and  $\alpha_a(m^{**}) = \beta_a(m^{**})$  are obtained using relations (8) and (9).

B. Assume IO but in reality II:

Under assumption IO, the computed log-likelihood ratio at time  $i$  is given by relation (4). However, since in reality the increments are independent, we have for  $i > 1$ ,

$$X_i = (X_i - X_{i-1}) + X_{i-1} \quad \text{for any value of } X_{i-1}.$$

Therefore,

$$X_1 \sim N(\mu_j, \sigma^2) \quad \text{for } j = 0, 1$$

$$\text{and } X_i \sim N(\mu_j + X_{i-1}, \sigma^2) \quad \text{for } j = 0, 1 \text{ and } i > 1.$$

Taking moments of equation (4) with the  $X_i$ 's thus distributed we obtain the conditional distributions of  $Z_i$  as:

$$Z_{i1} \sim N\left(\frac{d^2}{2} + z_{i-1}, d^2\right)$$

$$\text{and } Z_{i0} \sim N\left(-\frac{d^2}{2} + z_{i-1}, d^2\right)$$

Since we assumed IO, the truncation point  $m^{**}$  would be obtained from relationship (10). Thus, the values of  $E(N/W_j, m^{**})$  and  $\alpha_a(m^{**}) = \beta_a(m^{**})$  are obtained using relations (8) and (9).

C. Example Results:

To study the effect of wrong assumptions, three values of



the discrimination factor  $d$  were chosen ( $d = .75, 1$  and  $1.5$ ), and  $\alpha_d (= \beta_d)$  was varies between  $.01$  and  $.1$ . For fixed values of  $d$  and  $\alpha_d$ ,  $m^{**}$  was obtained for each model. This  $m^{**}$  was used as the truncation point for that particular assumed model, no matter the reality. For  $m^{**}$  thus known, and fixed  $d$  and  $\alpha_d$ , values of  $\alpha_a$  and  $E(N)$  were obtained.

An example of the results is shown in table 1, where  $d=1$  and  $\alpha_d=.01$ . When we assume II holds, and it does in reality, then  $m^{**} = 25$ ,  $\alpha_a = .0095$ , and  $E(N) = 10.31$ . When we assume II but in reality IO holds, we use the same  $m^{**} = 25$  but get  $\alpha_a = .0035$  and  $E(N) = 6.79$ .

Similarly, when we assume IO then  $m^{**} = 7$ , and if IO actually holds, then  $\alpha_a = .006$ , and  $E(N) = 4.41$ . However, if the same  $m^{**} = 7$  is used because we assume IO, but in reality II holds, we obtain  $\alpha_a = .0353$ , and  $E(N) = 4.2$ .

It can be seen from table 1 that when II holds in reality, using the wrong model gives  $E(N) = 4.2$  which is much less than  $E(N) = 10.31$  obtained by using right model. However, we also obtain  $\alpha_a = .0353$  which is much higher than  $\alpha_d = .01$ . On the other hand, when the observations are independent in reality, the use of a wrong model gives  $\alpha_a = .0035$  which is much less than  $.01$  but it gives  $E(N) = 6.79$  compared to  $4.41$  obtained by using the right model. Since the verification of independence assumptions might be expensive and/or time consuming, an experimenter might prefer such a slight increase in  $E(N)$ , in the event of the underlying assumption being wrong, as long as  $\alpha_a \leq \alpha_d$ .

TABLE 1

SENSITIVITY OF INDEPENDENCE ASSUMPTIONS

WHEN  $d=1$  AND  $\alpha_d = \beta_d = .01$

		REALITY	
		INCREMENTS INDEPENDENT	OBSERVATIONS INDEPENDENT
ASSUME	INCREMENTS ARE INDEPENDENT	$m^{**}=25$ $\alpha_a = \beta_a = .0095$ $E(N) = 10.31$	$m^{**}=25$ $\alpha_a = \beta_a = .0035$ $E(N) = 6.79$
	OBSERVATIONS ARE INDEPENDENT	$m^{**}=7$ $\alpha_a = \beta_a = .0353$ $E(N) = 4.2$	$m^{**}=7$ $\alpha_a = \beta_a = .006$ $E(N) = 4.41$
THAT .			

## 5. Observations:

For other values of  $d$  and  $\alpha_d$  it has been numerically confirmed (Golhar (4)) that the independent increments assumption is marginally superior to that of independent observations, in the sense that the II model gives only a slightly higher  $E(N)$  (in the event when II assumption is wrong) than that given by the right model, but still gives  $\alpha_a \leq \alpha_d$ .

This behavior is due to the fact that, for the IO model, the SPRT is truncated at an early stage assuming that a lot of information will be available. (Note that for IO,  $Z_n = \sum_{i=1}^n Z_i$ ). But when, in reality, the II assumption is true then the IO model gives actual error probabilities much greater than desired error probabilities. On the other hand, the II model makes use of only the most recent information ( $Z_n$  is a function of  $X_n$  only). Hence the truncation point is set high to get  $\alpha_a \leq \alpha_d$ . When in reality IO is true (which uses all the available information) then the independent increments model will give a slightly higher  $E(N)$  than the correct model but still gives actual error probabilities less than the desired error probabilities.

References:

1. Anderson, T.W., "A Modification of the Sequential Probability Ratio Test to Reduce the Sample Size," Ann. Math. Stat., 1960, Vol. 31, pp. 165-196.
2. Armitage, P., "Restricted Sequential Procedures," Biometrika, 1957, Vol. 44, pp. 9-26.
3. Aroian, L.A. and Robison, D.E., "Direct Methods For Exact Truncated Sequential Tests of The Mean of a Normal Distribution," Technometrics, 1969, Vol. 11, no. 4, pp. 661-675.
4. Golhar, D.Y., "Sequential Analysis: Non-Stationary Processes and Truncation," a Ph.D. dissertation submitted to industrial and operations engineering department at the University of Michigan, Ann Arbor, 1983.
5. Madsen, R., "A Procedure For Truncating SPRT'S," Journal of Amer. Stat. Assn., 1974, Vol. 69, no. 36, pp. 403-410.
6. Phatarfod, R.M., "Sequential Analysis of Dependent Observations I," Biometrika, 1965, Vol. 15, pp. 157-165.
7. Phatarfod, R.M., "Sequential Tests For Normal Markov Sequence," Journ. Austrial. Math. Soc., 1971, Vol. 12, pp. 433-440.
8. Siegmund, D.O., "The Variance of One-Sided Stopping Rules," Ann. Math. Stat., 1969, Vol. 40, no.3, pp. 1074-1077.
9. Siegmund, D.O., "On the Asymptotic Normality of One-Sided Stopping Rules," Ann. Math. Stat., 1968, Vol. 39, no. 5, pp. 1439-1497.
10. Wald, A., Sequential Analysis, Dover Publications Inc., New York, 1947.