A PUBLICATION OF THE RELIABILITY INFORMATION ANALYSIS CENTER

VOLUME 1, NUMBER 2

RELTIQUES

POWERFUL TECHNIQUES IN RMQSI FROM THE RIAC

UNDERSTANDING EXPONENTIAL SEQUENTIAL TESTS

By: Jorge Luis Romeu, Ph.D., Senior Science Advisor; Quanterion Solutions

TABLE OF CONTENTS

INTRODUCTION

The actual sample size "n" required in testing and confidence interval (CI) derivation is of tremendous importance for practitioners. For example, sample size carries with it a price tag in time, resources or both. In industrial applications, neither of these is plentiful.

When samples are taken all at one time, it is called single sampling. The problem of calculating the sample size for deriving a general CI has been discussed in Reference 2. Samples for acceptance testing were presented in Reference 3. Censored samples have been discussed in Reference 4. All of these cases, however, only treat the situation where fixed samples of pre-determined sizes are taken, all at one time.

An alternative consists in taking the samples in multiple stages and assessing their results at each stage. This allows the possibility of stopping the process and reaching an early decision, if certain conditions are met. For example, if the start data are clear-cut in favor of (or against) the hypothesis, then curtailing the test can save significant time and resources. Such is the case where samples are taken in successive stages, according to the assessment of results obtained from the previous sampling stages. This is known as "multiple sampling".

In Reference 1, double sampling plans were discussed and then extended to higher dimension sampling plans, namely sequential tests. However, only the case of dichotomous variables (i.e. those that have only two outcomes, e.g. pass or fail) was discussed. Such types of sequential plans are based on the discrete Binomial distribution. In this paper, the discussion is extended to sequential testing plans and continuous random variables (r.v.), especially those distributed exponentially. The same approach as that used in Reference 1 is applied.

In the remainder of this paper, several numerical examples of Sequential Probability Ratio Tests (SPRTs) for the Exponential case are presented, using the example of risks $\alpha = \beta = 0.128$ and a "discrimination ratio" (DR) = 2. The cumulative failure times of the device under test are first used to build the SPRT. The cumulative number of failures, and the Poisson distribution (which is related to the Exponential distribution), are then used to construct an equivalent SPRT. Some issues related to the calculation of the Average Sample Number (ASN) and Expected Test Times (ETT), two performance measures that evaluate the efficiency of SPRTs, are then discussed and the reader is introduced to the problem of test truncation. Finally, the results of the presented sequential testing approaches are discussed: (1) SPRTs for the continuous case developed in this paper, and (2) SPRTs for the discrete case, discussed in Reference 1, then contrasting their respective advantages and disadvantages.

SEQUENTIAL TESTS (SPRTS) FOR CONTINUOUS RANDOM VARIABLES

A good way to introduce and illustrate the theory behind SPRTs for continuous r.v. is via a numerical example. The same SPRT example given in Reference 1 is re-developed in this paper for exponentially-distributed life tests.

Assume that we want to test that the acceptable mean life of a device (MTTF or μ) is 200 hours or more (i.e., the null hypothesis H_0: $\mu \geq 200$) versus that it is 100 hours or less (the alternative hypothesis is H_1: $\mu \leq 100$). The same example presented in Reference 1 is reused here to establish comparisons and stress the differences in SPRT performance and efficiency when these hypotheses are used with discrete distributions.

Assume that the r.v. "life of a device", denoted "X", follows a continuous distribution with mean μ. Its cumulative density (CDF) and probability density (pdf) functions under the two hypotheses H_i: $\mu = \mu_i$, for i = 0,1 are given by:

$$P\{X \leq T\} = F_{\mu_i}(T) = \int_0^T f_i(x)dx$$

Therefore:

$$P\{Life.Under.H_i\} = P_{\mu_i}\{Device.Lives.up.to.Time : T\} = P_{\mu_i}\{X \leq T\} = F_{\mu_i}(T); i = 0,1$$

In the case where the distribution of life "X" is exponential, the above reduces to:

$$F_{\mu_i}(t)=\int_0^t f_{\mu_i}(x)dx=1-e^{\frac{-t}{\mu_i}};f_{\mu_i}(t)=\frac{1}{\mu_i}e^{\frac{-t}{\mu_i}};i=0,1;\mu_0=200;\mu_1=100$$

Now, place "n" independent devices sequentially on test, one at a time, until they all fail. Then, the Probability Ratio (PR) that the sequence of such "n" life tests, r.v. $X_1,\ldots,X_n$, have actually experienced failure times of $t_1,\ldots,t_n$, under hypotheses H_0 and H_1, is defined by:

$$\frac{P\{UnderH_1\}}{P\{UnderH_0\}}=\frac{P_{\mu_1}(X_1\le t_1,\ldots,X_n\le t_n)}{P_{\mu_0}(X_1\le t_1,\ldots,X_n\le t_n)}=\frac{P_{\mu_1}(X_1\le t_1)\ldots P_{\mu_1}(X_n\le t_n)}{P_{\mu_0}(X_1\le t_1)\ldots P_{\mu_0}(X_n\le t_n)}=\frac{F_{\mu_1}(t_1)\cdots F_{\mu_1}(t_n)}{F_{\mu_0}(t_1)\cdots F_{\mu_0}(t_n)}$$

Since the X's are now continuous r.v., the substitution of CDFs with pdf's yields:

$$q_n(t_1,\cdots,t_n)=\frac{f_{\mu_1}(t_1)\cdots f_{\mu_1}(t_n)}{f_{\mu_0}(t_1)\ldots f_{\mu_0}(t_n)}$$

In the specific case where device "lives" (X) are exponentially distributed, this becomes:

$$q_n(t_1,\cdots,t_n)=\frac{\frac{1}{\mu_1}\exp(-\frac{t_1}{\mu_1})\cdots\frac{1}{\mu_1}\exp(-\frac{t_n}{\mu_1})}{\frac{1}{\mu_0}\exp(-\frac{t_1}{\mu_0})\cdots\frac{1}{\mu_0}\exp(-\frac{t_n}{\mu_0})}=\frac{\left(\frac{1}{\mu_1}\right)^n\exp\left\{-\sum_{i=1}^n\frac{t_i}{\mu_1}\right\}}{\left(\frac{1}{\mu_0}\right)^n\exp\left\{-\sum_{i=1}^n\frac{t_i}{\mu_0}\right\}}$$

As done in Reference 1, both hypothesis test errors, α (Producer's Risk, or the probability of rejecting a device having an acceptable life), and β (Consumer's Risk, or the probability of accepting a device with unacceptable life) are defined as: α = β = 0.128. Then, two values, "A" and "B" can be found such that, at any SPRT stage "n" (having tested "n" devices sequentially (one at a time) and having obtained failure times ($t_1,\ldots,t_n$), the probability of the "q_n" function can be determined, just as occurred in the discrete distribution case in Reference 1, fulfills:

$$P\{q_n>A\}=P\left\{\frac{\left(\frac{1}{\mu_1}\right)^n\exp\left\{-\sum_{i=1}^n\frac{t_i}{\mu_1}\right\}}{\left(\frac{1}{\mu_0}\right)^n\exp\left\{-\sum_{i=1}^n\frac{t_i}{\mu_0}\right\}}>A\right\}=\beta;$$

$$P\{q_n<B\}=P\left\{\frac{\left(\frac{1}{\mu_1}\right)^n\exp\left\{-\sum_{i=1}^n\frac{t_i}{\mu_1}\right\}}{\left(\frac{1}{\mu_0}\right)^n\exp\left\{-\sum_{i=1}^n\frac{t_i}{\mu_0}\right\}}<B\right\}=1-\alpha$$

Hence, we again define S (B, A) is again defined as the SPRT, via the above equations, that compares q_n with A and B at every stage "n", and supports a decision to either (1) accept H_0, if q_n < B; (2) accept H_1, if q_n > A; or (3) continue testing, if B < q_n < A. However, as with the discrete SPRT, such a formulation is not easy to work with. It can be improved by defining the natural logarithm (ln) of the qn above, denoted "z_n", as:

$$z_n=\ln(q_n)=\ln\left[\left(\frac{\mu_0}{\mu_1}\right)^n Exp\left\{-\left(\frac{1}{\mu_1}-\frac{1}{\mu_0}\right)\sum_{i=1}^n t_i\right\}\right]=n\ln\left(\frac{\mu_0}{\mu_1}\right)-\left(\frac{1}{\mu_1}-\frac{1}{\mu_0}\right)\sum_{i=1}^n t_i$$

The "z_n" is a linear function that depends on the number of stages (n), or devices tested, on the two parameters (or MTTFs) μ_0, μ_1 on test, and on the actual failure times: $t_1,\ldots,t_n$. The "continuation region", defined by z_n, is now bounded by the logarithms of A and B:

$$\ln(B)<z_n=an+b\sum_{i=1}^n t_i<\ln(A);with:a=\ln\left(\frac{\mu_0}{\mu_1}\right);b=-\left(\frac{1}{\mu_1}-\frac{1}{\mu_0}\right)$$

As in Reference 1, coefficients "a" and "b" are just functions of test parameters μ_i, i = 0, 1, and constants A and B can be approximated by (References 5 through 9):

$$A\cong\frac{(1-\beta)}{\alpha};B\cong\frac{\beta}{1-\alpha}$$

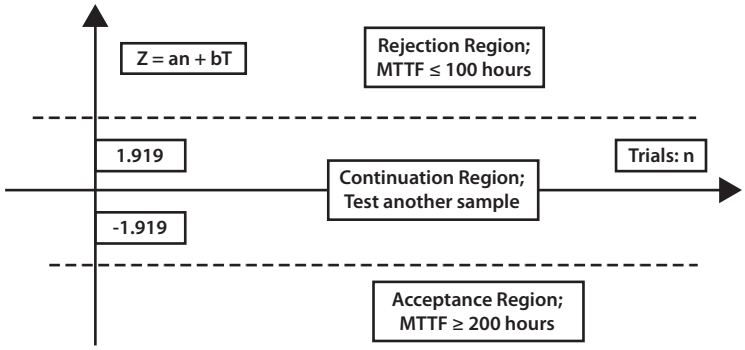
Hence, for any SPRT stage "n" (corresponding to having tested "n" devices sequentially, one at a time, and having obtained failure times: $t_1,\ldots,t_{n-1}$) the function "z_n" will meet:

$$P\{z_n>\ln(A)\}=P\left\{n\ln\left(\frac{\mu_0}{\mu_1}\right)-\left(\frac{1}{\mu_1}-\frac{1}{\mu_0}\right)\sum_{i=1}^n t_i>\ln(A)\right\}=\beta;$$

$$P\{z_n<\ln(B)\}=P\left\{n\ln\left(\frac{\mu_0}{\mu_1}\right)-\left(\frac{1}{\mu_1}-\frac{1}{\mu_0}\right)\sum_{i=1}^n t_i<\ln(B)\right\}=1-\alpha$$

Thus, the SPRT denoted S (B,A) and defined by the above equations, compares "z_n" with the logarithms of A and B, at every stage "n", and supports a decision to either (1) accept H_0 if z_n < ln(B); (2) accept H_1if z_n > ln(A); or (3) continue testing if ln(B) < z_n < ln(A). The discrimination ratio is = μ_0 / μ_1 = 200 / 100 = 2.0.

For example, assume we are at stage "ten", having placed the tenth (n = 10) device on test. Also, assume that failures occurred at times 99.9, 210.8, 166.2, 77.8, 105.3, 193.7, 170.1, 256.3, 327.4 and 219.3 hours with a Total Test Time (T) = ΣT_i = 1827.2. Since μ_0 = 200, μ_1 = 100, n =10, α = β = 0.128, and since the SPRT coefficients "a" and "b" can be calculated and values "A"



and "B" can be approximated, we obtain the following results:

$$a = \ln\left(\frac{\mu_0}{\mu_1}\right) = \ln\frac{200}{100} = \ln(2.0) = 0.693147$$

$$b = -\left(\frac{1}{\mu_1} - \frac{1}{\mu_0}\right) = -\left(\frac{1}{100} - \frac{1}{200}\right) = -(0.01 - 0.005) = -0.005$$

$$A \cong \frac{(1-\beta)}{\alpha} = \frac{(1-0.128)}{0.128} = 6.813; \Rightarrow \ln(A) = \ln(6.813) = 1.919$$

$$B \cong \frac{\beta}{1-\alpha} = \frac{0.128}{1-0.128} = 0.147; \Rightarrow \ln(B) = \ln(0.147) = -1.919$$

In the example, the value z_n = -2.204 (Table A) obtained at the tenth stage is compared with the values for ln(A) = 1.919 and ln(B) = -1.919, which also yield the plot boundaries of the SPRT continuation region. Since z_n is smaller than ln (B), and falls below the lower bound of the continuation region, the decision is made to stop testing and to accept the null hypothesis (H_0) that the exponential MTTF (μ) is 200 hours or more. A plot of this example is shown below. Notice how the path of z_n for all stages up to the tenth falls inside the plot continuation region delimited by ±1.919. As soon as z_n falls outside the continuation region (stage 10, or n = 10), testing is stopped and the corresponding decision regarding the test hypotheses is made:

$$z_{10} = a \times 10 + b\sum_{i=1}^{10} t_i = 0.693147 \times 10\text{-}0.005 \times 1827.18 = -2.2044 < -1.919$$

Sequential Test Example (n=10)

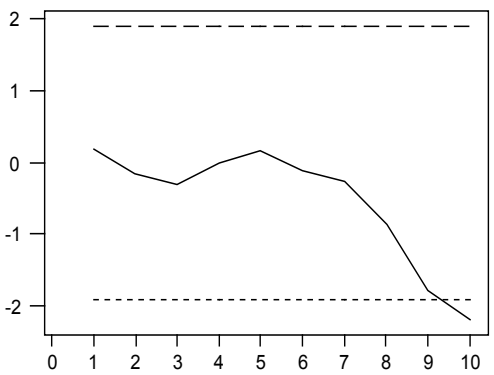


Table A

Stage	FailTime	Z(n)
1	99.929	0.19350
2	210.820	-0.16745
3	166.270	-0.30565
4	77.855	-0.00178
5	105.325	0.16474
6	193.720	-0.11071
7	170.170	-0.26841
8	256.300	-0.85677
9	327.400	-1.80062
10	219.380	-2.20437

IMPROVING THE SPRT CALCULATIONS

As was done with the discrete SPRT (Reference 1), the implementation of the test procedure can be improved upon by finding some convenient boundaries to compare $T = \Sigma T_i$ (Total Test Time) directly. This is obtained, as before, by isolating T in the above inequalities:

$$\ln(B) < z_n = an + b\sum_{i=1}^{n} t_i < \ln(A); but: b < 0$$

$$h_0 + sn = \frac{\ln(B)}{b} - \frac{a}{b}n > \sum_{i=1}^{n} t_i > \frac{\ln(A)}{b} - \frac{a}{b}n = h_1 + sn$$

$$with: h_0 = \frac{\ln(B)}{b}; h_1 = \frac{\ln(A)}{b}; s = -\frac{a}{b}$$

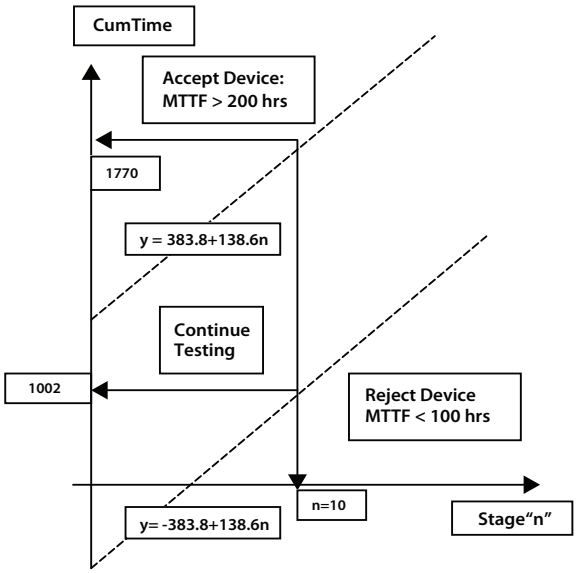
The above results yield, in the numerical example, the following SPRT boundaries:

$$h_0 = \frac{\ln(B)}{b} = \frac{-1.919}{-0.005} = 383.8; h_1 = \frac{\ln(A)}{b} = \frac{1.919}{-0.005} = -383.8;$$

$$s = -\frac{a}{b} = -\frac{0.6931}{-0.005} = 138.63; T = \sum_{i=1}^{n} t_i; \text{for.the.n}^{\text{th}}\text{.stage;}$$

$$r_n = h_0 + sn = 383.8 + 138.63n > T > -383.8 + 138.63n = h_1 + sn = a_n$$

By letting the number of "stages" "n", run from 1, 2, … , we obtain the SPRT decision "boundaries" (or "acceptance and rejection numbers" ($a_n; r_n$)) for Total Time on Test T:



Parameters for our current example are:
$h_0 = 383.8$; $h_1 = -383.8$; $s = 138.6$

Stage	Reject	Accept	CumTime
1	-245.2	522.4	99.9
2	-106.5	661.1	310.7
3	32.1	799.7	477.0
4	170.7	938.3	554.9
5	309.3	1077.0	660.2
6	448.0	1215.6	853.9
7	586.6	1354.2	1024.1
8	725.2	1492.8	1280.4
9	863.9	1631.5	1607.8
10	**1002.5**	**1770.1**	**1827.2**
11	1141.1	1908.7	**Accept**
12	1279.8	2047.4	
13	1418.4	2186.0	
14	1557.0	2324.6	
15	1695.6	2463.3	
16	1834.3	2601.9	
17	1972.9	2740.5	
18	2111.5	2879.1	
19	2250.2	3017.8	
20	2388.8	3156.4	
21	2527.4	3295.0	
22	2666.1	3433.7	
23	2804.7	3572.3	
24	2943.3	3710.9	

Notice how it is impossible to reject the device (H_1: MTTF $\mu_1 \le 100$ hrs.) until at least n = 3 devices have been tested with a cumulative or Total Test Time of less than 32.1 hrs. After that, a device at any stage "n" can be rejected if the Total Test Time is less than its Rejection Bound "r_n". We can accept the device (H_0: MTTF $\mu_0 \ge 200$ hrs.) at any stage "n" if its Total Test Time is larger than its Acceptance Bound "a_n". The probability of an erroneous decision is, at most, $\alpha = \beta = 0.128$. In the example, an accept decision at stage 10 is made because cumulative time for 10th device is $T = 1827.2 \ge a_{10} = 1770.1$.

As in the discrete case discussed in Reference 1, both boundary line equation parameters (the slope and the intercepts) depend only on the SPRT errors α, β, and on the test MTTFs μ_0, μ_1:

$$h_0 = \frac{\ln\left(\frac{\beta}{1-\alpha}\right)}{-\left(\frac{1}{\mu_1}-\frac{1}{\mu_0}\right)}; h_1 = \frac{\ln\left(\frac{(1-\beta)}{\alpha}\right)}{-\left(\frac{1}{\mu_1}-\frac{1}{\mu_0}\right)}; s = -\frac{\ln\left(\frac{\mu_0}{\mu_1}\right)}{-\left(\frac{1}{\mu_1}-\frac{1}{\mu_0}\right)}$$

Also discussed in Reference 1, an SPRT can be applied to an SPC/Quality Control problem. For example, assume that there is interest in assessing a batch of incoming devices, and that the acceptable quality level (AQL) is defined by some minimum performance (say, $\mu \ge 200$ hrs). Also, assume there is an unacceptable quality (Lot Tolerance Percent Defective - LTPD) below which the batch will be rejected (say, if the mean life is less than 100 hrs).

Now assume that, for procedural ease, cost, or any other practical reason, it is decided to test the lot by taking one item at a time and testing each one sequentially, instead of drawing a single sample of fixed size "n" at one time, and placing them all on a life test.

Then, for the acceptance sampling problem described above, all of the previously described SPRT derivations and results are also applicable, with the pertinent modifications.

THE RELATIONSHIP BETWEEN THE EXPONENTIAL AND POISSON DISTRIBUTIONS

Readers familiar with MIL-HDBK-781A and sequential testing may have seen the equations above derived from the Poisson distribution using "r" (cumulative number of failures) instead of cumulative test time "T", as was done here. Both derivations are equivalent, given the statistical relationship between the Exponential and Poisson distributions:

Assume the device life "X" follows the exponential distribution. Then, if it has not failed by time T:

$$P_\mu\{Device.Outlives.T\} = P_\mu\{X > T\} = e^{\frac{-T}{\mu}}$$

This is equivalent to stating that the r.v. "number of failure events by time T", or N(T), (which follows the Poisson Distribution) has undergone zero events by time T:

$$P_\mu\{No.Events.By.T = 0\} = P_\mu\{N(T) = 0\} = \frac{e^{\frac{-T}{\mu}}\left(\frac{T}{\mu}\right)^{N(T)}}{N(T)!} = \frac{e^{\frac{-T}{\mu}}\left(\frac{T}{\mu}\right)^{0}}{0!} = e^{\frac{-T}{\mu}} = P_\mu\{X > T\}$$

Hence, using the above formulation the Probability Ratio (PR) can alternately be found that the sequence of "n" independent Poisson life tests has actually experienced "r" failures, in a cumulative test time "T" = Σt_i, under both SPRT hypothesis H_0 and H_1:

$$\frac{P\{N(T).Events.Under H_1\}}{P\{N(T).Events.Under H_0\}} = \frac{P_{\mu_1}(N(T)=r)}{P_{\mu_0}(N(T)=r)} = \frac{\frac{e^{\frac{-T}{\mu_1}}\left(\frac{T}{\mu_1}\right)^{r}}{r!}}{\frac{e^{\frac{-T}{\mu_0}}\left(\frac{T}{\mu_0}\right)^{r}}{r!}} = \left(\frac{\mu_0}{\mu_1}\right)^{r} Exp\left\{-T\left(\frac{1}{\mu_1}-\frac{1}{\mu_0}\right)\right\}$$

Taking logarithms, as before, but re-arranging now for the "number of events" N(T) = r:

$$\ln(B) < \ln\left[\left(\frac{\mu_0}{\mu_1}\right)^r Exp\left\{-T\left(\frac{1}{\mu_1}-\frac{1}{\mu_0}\right)\right\}\right] = r\ln\left(\frac{\mu_0}{\mu_1}\right) - T\left(\frac{1}{\mu_1}-\frac{1}{\mu_0}\right) < \ln(A)$$

$$h_0 + sT = \frac{\ln(B)}{b} + \frac{a}{b}T < r < \frac{\ln(A)}{b} + \frac{a}{b}T = h_1 + sT$$

$$with: h_0 = \frac{\ln(B)}{b}; h_1 = \frac{\ln(A)}{b}; s = \frac{a}{b}; a = \frac{1}{\mu_1} - \frac{1}{\mu_0}; b = \ln\left(\frac{\mu_0}{\mu_1}\right)$$

This is equivalent to the previous derivation, now assessed based on the number of failures "r" instead of on the cumulative test time "T". This is the familiar form used in MIL-HDBK-781, where time "T" is on the horizontal axis and failures, "r", on the vertical. Here, the upper and lower bounds are defined by equations $h_0 + sT$ and h1 + sT, both functions of time T. The rest of the derivations are all carried out in the same manner as done in the previous section.

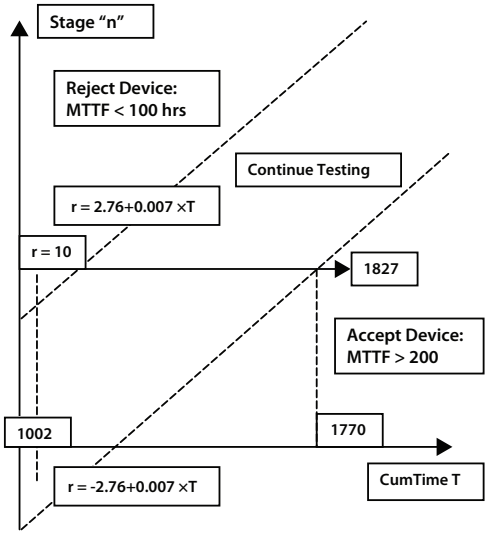
As an illustration, the acceptance and rejection ($^{a_n;r_n}$) number of failures "r" corresponding to the numerical example given previously, with Total Time T = 1827.2, are calculated:

$$h_0 = \frac{\ln(0.147)}{0.693} = -2.76; h_1 = \frac{\ln(6.81)}{0.693} = 2.76; s = \frac{a}{b} = \frac{0.005}{0.693} = 0.007;$$

$$a = \frac{1}{200} - \frac{1}{100} = 0.005; b = \ln\left(\frac{200}{100}\right) = 0.693; Hence:$$

$$a_{10} = 10.02 = -2.76 + 0.007 \times T < r < 2.76 + 0.007 \times T = 15.56 = r_{10}$$

The Figure shows how, for n = 10, the cumulative time T = 1770 is the maximum time acceptable, for a failure to occur and still continue testing.



Since the tenth (r =10) failure occurs at time "T" = 1827.2 **we accept H_0**. Such cumulative time fails to fall inside the Continuation Region. It would require a minimum number of failures r = 10.02. But the result "r = 10" is smaller: T falls in the Acceptance region.

AVERAGE SAMPLE NUMBER, AVERAGE TEST TIME, AND TEST TRUNCATION

The main advantage of sequential tests over fixed time tests is the reduction of "long run" or average sample size, and of the test time required to arrive at an accept/reject decision. But "SN" (the "sample number") is a probabilistic outcome (r.v.). Hence, its "Expected Value" (or "Average Sample Number" – ASN) depends on the true (but unknown) value of the underlying test parameter, be it MTTF "μ", percent defective "p", or any other.

The ASN is obtained following the traditional definition of Expected Value, i.e., by multiplying each SPRT stage "n" (the number of samples taken so far) by the probability of arriving at a decision (be it H_0 or H_1) at that stage. The ASN depends on both the true parameter "μ", as well as on not having made a decision earlier in the test process. That is, "T" has followed a path inside the SPRT "continuation region", up to the present stage:

$$ASN(\mu) = E_\mu\{SN\} = \sum_{SN \geq 1} SN \times P_\mu\{Decision.at.Stage.SN.But.not.before\}$$

Once the ASN is calculated, the Expected Test Time (ETT) can also be obtained, since it is known that the distribution of life "X" is Exponential. Calculation of values for ASN and ETT for SPRT tests fall beyond the scope of the present paper. Readers interested in such advanced topics may consult References 5, 6, 7, 8 and 9.

Some SPRTs may require many stages. In such cases, their main advantage (saving time and sample size) vanishes. Hence, it is often of interest to establish a maximum number of stages and Total Test Time at which the SPRT is terminated and a decision is made based on the results available at the appropriate time and stage. This is known as "test truncation".

An efficient method to establish the SPRT truncation stage and time is via Type I and II errors (α and β) and the DR = μ_0 / μ_1. For exponentially-distributed lives, the statistic 2T/μ is distributed as a Chi Square with "2r" degrees of freedom (DF). Hence, the Chi Square table is searched for the smallest DF such that the ratio corresponding to its two Chi Square percentiles, at probabilities (1 - α) and β, fulfills:

$$\underset{2r}{Minimum}\left\{\frac{X^2_{1-\alpha,2r}}{X^2_{\beta,2r}}\right\} \geq \frac{1}{DR} = \frac{\mu_1}{\mu_0}$$

In the current example, the SPRT parameters are α = β = 0.128 and DR = 200 / 100 = 2:

D. F. "2r"	22	23	24	25
$X^2_{0.872,\ r}$	14.77	15.59	16.43	17.26
$X^2_{0.128,\ r}$	29.62	30.79	31.96	33.12
Ratio	0.499	0.505	0.51	0.52

Hence, the Chi Square percentiles ratio meeting these conditions corresponds to DF = 23:

$$\frac{X^2_{1-0.128,2r}}{X^2_{0.128,2r}} = \frac{X^2_{0.872,23}}{X^2_{0.128,23}} = \frac{15.59}{30.79} = 0.505 \geq \frac{100}{200} = 0.5$$

The Truncation Stage "r" is then: 2×r = 23 => r ≈ 12 stages. From here, the Truncation Time T* can also be obtained, since the distribution of 2T/μ is Chi Square, with DF = 2r = 24:

$$T^* = \frac{1}{2} \times \mu_0 \times X^2_{1-\alpha,2r} = 0.5 \times 200 \times 16.428 = 1642.8$$

Therefore, the SPRT should terminate (be truncated) at the 12th stage (failure) or at a Total Test Time of 1642.8 hours, whichever comes first. If the last stage results are such that the 12th failure occurs before 1642.8 hours, the null hypothesis H_0 is rejected, but if the Total Test Time reaches 1642.8 hours before the 12th failure occurs, then the null hypothesis that MTBF ≥ 200 is accepted.

COMPARING SPRT APPROACHES

In this and the Reference 1 paper, sequential methods for testing the "life" of a device, assuming such life is distributed exponentially, have been developed. Life, errors α and β, and the hypotheses H_0, H_1were used in both papers. Their corresponding SPRT test performances and characteristics are now discussed.

In Reference 1, a test time of duration T was pre-defined. Then, devices were tested, one at a time, for that time. Finally, an investigation was made into whether each device survived or failed its test, and the number of successes (or failures) out of the total "n" devices placed on test was counted. This process represented the binomial SPRT.

In the example in this paper, the devices were allowed to run until they failed. Then, either the Cumulative Test Time, or at the cumulative number of failures that occurred during the testing were evaluated. This process represents the continuous time, exponential (or equivalently, the Poisson) SPRT. The continuous time test is always more efficient, since no information is lost.

The efficiency of the binomial SPRT depends on the length of time allotted for each trial. The longer this time is, the better the SPRT. Test time also affects the average number of "stages" required to arrive at a decision (the ASN), as well as the ETT.

For example, the binomial SPRT in Reference 1 used T = 20 hours, for α = β = 0.128, and MTTF μ_0 = 200, μ_1 = 100. If T increases, the SPRT performance improves. Shown below are the binomial SPRT slopes, intercepts, DR and proportions p_0, p_1, obtained for increasing values of the pass/fail test length T. The MTTFs μ_0 = 200, μ_1 = 100, remain fixed:

P/F Test Time	T = 20	T = 30	T = 40	T = 50
Proportion p_0	0.905	0.861	0.819	0.779
Proportion p_1	0.819	0.741	0.670	0.607
Discr. Ratio	1.105	1.162	1.222	1.283
Intercepts	±2.578	±2.489	±2.404	±2.323
Slope	0.866	0.805	0.749	0.697

Notice how the SPRT parameters depend on the pass/fail (P/F) test time. As T increases, intercepts are closer, the slope becomes smaller and the DR = p_0 / p_1 increases. Letting the test run for a longer time provides more information and, hence, better discrimination.

CONCLUSIONS

An overview of sequential tests for the continuous case have been provided for the exponential distribution. SPRTs are widely used in practice, and are included in MIL-HDBK-781A, "Reliability Testing for Engineering Development". Sequential tests are also used in Quality Control and acceptance sampling, when batch items are tested, one at a time, to assess the quality of a lot, instead of testing a fixed sample size all at one time. The problems of assessment of SPRT test plans, via the ASN and the ETT, as well as of establishing test truncation procedures, were also highlighted.

SPRTs are very useful tools. By the same token, they are theoretically complex. In this paper, and in Reference 1, the discussion has been kept simple and focused. Hence, there are several important SPRT variants that were not covered. Some will be mentioned below. The interested reader can pursue more detail in References 5 through 9.

For example, the equation for the SPRT constant "A" differs from that used in MIL-HDBK-781A, which uses a correction factor of (DR+1)/2*DR; this paper does not. The purpose of such a correction factor is to reduce the differences between "nominal" and "achieved" risks, α and β, that appear when an SPRT is truncated. In the presented non-truncated example of this paper, the true SPRT risks are 0.128. Once the SPRT is truncated, these risk values change.

In the truncated SPRT Plan III-D of MIL-HDBK-781 (which is comparable to the presented example) the nominal risks α and β are 0.1. The reader will notice that both Plans have identical slopes (0.007213). However, the intercepts, truncation stages and maximum test times differ somewhat. MIL-HDBK-781 usage of a "correction" factor for the constant "A", is the reason for these differences.

SPRT stages can also include more than one unit on test. Moreover, all of

the units can be placed on test at the start. In such cases, the individual sequential times to failures are monitored, but their respective Cumulative Times to size-adjusted SPRT acceptance and rejection boundaries should be compared. The derivations are similar to the ones presented here, with the pertinent modifications.

In addition, the lives of many repairable systems are approximately exponentially distributed. For this reason, SPRTs are widely used to test repairable systems. In such cases, MTBF instead of MTTF becomes the SPRT performance measure of interest.

Finally, using the same numerical examples as in binomial SPRTs (Reference 1) allowed a comparison between the efficiencies of two options: (1) letting each device reach its natural end-of-life, versus (2) limiting the binomial P/F test to a maximum time, T_0, and observing whether each device passed or failed the test. Letting a device reach its natural life always yields more information and, hence, improves the test efficiency.

BIBLIOGRAPHY

1. Romeu, J. L., "Understanding Binomial Sequential Testing", RIAC START, Volume 12, Number 2, http://www.theriac.org/pdfs/startsheets/binseq.pdf
2. Romeu, J. L., "Estimating the Sample Size in Reliability Experiments", SPES/QP ASA Section Newsletter, Vol. 19, No. 1 (Summer 2011), pp. 8—11
3. Romeu, J. L., "Operational Characteristic (OC) Functions and Acceptance Sampling Plans", RIAC START, Volume 11, Number 8, http://www.theriac.org/pdfs/startsheets/oc_curves.pdf
4. Romeu, J. L., "Censored Data", RIAC START, Volume 11, Number 3, http://www.theriac.org/pdfs/startsheets/censor.pdf
5. Duncan, A. J., "Quality Control and Industrial Statistics", Fourth Ed., R.D. Irwin, 1974
6. MIL-HDBK-781D and MIL-STD-781D, "Reliability Testing for Engineering Development, Qualification and Production", U.S. Department of Defense, October 1986
7. Wald, A., "Sequential Analysis", John Wiley & Sons, 1947
8. Mann, N., R. Schafer and N. Singpurwalla, "Methods for Statistical Analysis of Reliability and Life Data", John Wiley & Sons, 1974.
9. Kececioglu, D., "Reliability and Life Testing", Volume 2, Prentice Hall, 1993

ABOUT THE AUTHOR

Jorge Luis Romeu is an Associate Professor Emeritus of Statistics and Computers, formerly with the Department of Mathematics at SUNY-Cortland.

Romeu is a Research Professor with Syracuse University, a Senior Science Advisor with the RIAC (Reliability Information Analysis Center), and an industrial statistics consultant.

Romeu earned a Ph.D. in Industrial Engineering and Operations Research from Syracuse University (1990) and a "Licenciado en Matematicas, Especialista en Estadistica" degree from the University of Havana, Cuba (1973). He is a Chartered Statistician Fellow of the Royal Statistical Society, a Senior Member of the American Society for Quality, and a member of the American Statistical Association. He holds both ASQ Quality and Reliability certifications, and the RSS professional (Chartered Statistician Fellow) accreditation. Dr. Romeu retired Emeritus in December of 1998, after fourteen years of teaching at SUNY Cortland.