A SEQUENTIAL STOPPING RULE FOR FIXED-SAMPLE ACCEPTANCE TESTS

Gerald G. Brown

University of California, Los Angeles, California

and

Herbert C. Rutemiller

California State College, Fullerton, California

(Received April 27, 1970)

The occurrence of early failures in a fixed-sample acceptance test, where the sample observations are obtained sequentially, presents an interesting decision problem. It may be desirable to abandon the test at an early stage if the conditional probability of passing is small and the testing cost is high. This paper presents a stopping rule based on the maximum-likelihood estimate of total costs involved in the decision to continue beyond an early failure. A Bernoulli model, an exponential model, and a Weibull model are examined.

SUPPOSE THAT one of the acceptance criteria specified in the design of a new product is that a random sample of n items show r, or fewer, failures in a performance test. If the performance test requires an extensive period of time to complete, an interesting decision problem may arise. The investigator may be able to judge, well before completion, that the probability of passing is low and design modifications are required. By abandoning the partially completed test, the investigator can save the delay cost of waiting until completion, as well as the actual cost of performing the remainder. There is an accompanying risk that the decision to abandon was incorrect, resulting in costs for an unnecessary design change and a complete rerun of the test.

There are two situations where this problem may arise. The first is caused by a limitation in facilities that necessitates sequential item-by-item testing. The observations will be sequential Bernoulli trials.

Second, the performance criterion may be a life test for an extended number of hours. If the n items are tested simultaneously, the sample observations will be order statistics from whatever parametric probability model is the true underlying life distribution.

This paper is concerned with the specification of decision rules for judging whether to continue or abandon performance tests when early failures are encountered. Both the random-sample (sequential Bernoulli trials) and ordered-sample (life testing) cases will be examined. In the life-testing model, we consider cases where the true underlying distribution is either exponential or Weibull.

ASSUMPTIONS

LET C_1 = cost per hour of conducting the test, C_2 = cost per hour of delay in finishing the design (e.g., facilities and personnel devoted to this project that would be otherwise available), C_3 = total cost of redesign (including the time required to implement it).

The cost of redesign C_3 is undoubtedly the most difficult to estimate. This cost is to include whatever redesigns are necessary to make the probability of failure on rerun negligible. To simplify the mathematics, it is assumed that *unnecessary* design changes, caused by incorrectly abandoning the test, will also have a beneficial effect on performance. This assumption appears warranted for many electronic and mechanical systems, where the introduction of redundancies, higher-quality components, etc., can always be expected to improve reliability.

(a) *The Bernoulli model.* Consider a requirement for r or fewer failures in n trials at the point where k trials have been performed and have resulted in i failures. If θ is the probability of a single failure, then the a priori probability of passing is given by the negative binomial probability that the $(r+1)$st failure occurs at trial $(n+1)$, or later:

$$P_0 = \sum_{j=n+1}^{\infty} \binom{j-1}{r} \theta^{r+1}(1-\theta)^{j-r-1}.$$

After i failures are observed in k trials, the conditional probability of passing may be estimated numerically by the method of maximum likelihood,[4] and is

$$\hat{P}_0 = \sum_{j=n-k+1}^{\infty} \binom{j-1}{r-1} \left(\frac{i}{k}\right)^{r-i+1} \left(1-\frac{i}{k}\right)^{j-r+i-1}.$$

Let $\hat{C}_A$ be the estimated expected cost of abandoning the test after k trials. If each individual observation requires h hours, then

$$\hat{C}_A = \hat{P}_0[C_3 + kh(C_1 + C_2)].$$

The second term in the brackets has k as a multiplier, since n additional tests are performed in the rerun, but $(n-k)$ are saved by stopping after trial k.

Let C_0 be the cost of continuing the test at trial k. For its computation, we require an estimate of the expected waiting time for completion of an unsuccessful test. This conditional expectation may be estimated by the method of maximum likelihood as

$$\hat{E}(t)=\frac{h\sum_{j=r-i+1}^{n-k} j\binom{j-1}{r-1}\left(\frac{i}{k}\right)^{r-i+1}\left(1-\frac{i}{k}\right)^{j-r+i-1}}{1-\hat{P}_0}.$$

The estimated cost $\hat{C}_0$ is given by $\hat{C}_0=\hat{E}(t)(C_1+C_2)$. The decision rule will be based on the relative magnitude of $\hat{C}_A$ and $\hat{C}_0$. The simplest rule would be: *If* $\hat{C}_0<\hat{C}_A$, *continue the present test; if* $\hat{C}_0\geqq\hat{C}_A$, *abandon the present test and initiate a redesign.* However, since both $\hat{C}_0$ and $\hat{C}_A$ are estimates subject to sampling error, one might require $\hat{C}_0$ to be substantially higher than $\hat{C}_A$ before initiating redesign. Let D represent a constant, greater than unity, which would require stronger inequality of costs to initiate redesign; then the rule could be: *If* $\hat{C}_0\geqq D\cdot\hat{C}_A$, *abandon the test.* In the examples that follow, D is assumed to be unity for simplicity.

The aforementioned difficulty in estimating the redesign cost C_3 suggests the desirability of a sensitivity analysis on this cost.

(b) *The life-testing model.* It will be assumed in this section that the times of interest to the decision maker are restricted to those where a failure has just occurred.

Let $y_1, y_2, \cdots, y_k$ be the ordered sample observations from a population with life distribution $f(t,\theta)$. Let $\hat{\theta}$ be the maximum-likelihood estimate of θ, based upon the first k order statistics. Let $g(y_1,\cdots,y_k;\theta)$ be the joint density of the k observations, and $g(y_1\cdots y_k, y_{r+1};\theta)$ be the joint density of the first k and $(r+1)$st order statistics.[5]

If t_0 is the life specified as acceptable, then the probability of passing the test after y_k has been observed may be estimated as

$$\hat{P}_0=\int_{t_0}^{\infty}[g(y_1,\cdots,y_k,y_{r+1};\hat{\theta})/g(y_1,\cdots,y_k;\hat{\theta})]\,dy_{r+1}.$$

The estimated cost of abandoning the test is

$$\hat{C}_A=\hat{P}_0[(t_0-y_k)(C_1+C_2)+C_3].$$

To estimate the cost of continuation, we require the expected value of $(y_{r+1}-y_k)$, given that y_{r+1} is less than t_0:

$$\hat{E}(y_{r+1}-y_k)=\int_{yk}^{t_0}[y_{r+1}g(y_1,\cdots,y_k,y_{r+1};\hat{\theta})/g(y_1,\cdots,y_k;\hat{\theta})]\,dy_{r+1}-y_k.$$

We have, then,

$$\hat{C}_0=[\hat{E}(y_{r+1}-y_k)][C_1+C_2].$$

The decision rule is the same as for Bernoulli trials.

EXAMPLES

(a) *Bernoulli trials.*

A computer program designed for a retailer to detect errors in customer's accounts is required to meet the specification that a random sample of 200 accounts shall have no more than 4 undetected errors. The efficacy of the computer program is to be tested by an audit of the 200 accounts. A single individual does the auditing, and he requires one-half hour per account. His charge is $10 per hour. The cost of waiting time for the program to be finished is estimated at $40 per hour. The cost of redesigning the program is estimated as $300.

The test is initiated and the 60th account examined reveals the second error. Should the test be continued?

We have $k=60$, $i=2$, $r=4$, $n=200$, $h=0.5$, $C_1=10$, $C_2=40$, $C_3=300$, $\hat{\theta}=2/60=0.0333$;

$$\hat{P}_0=\sum_{j=141}^{\infty}\binom{j-1}{2}(0.03333)^3(0.96667)^{j-3}=\bar{0}.1511;$$

$$\hat{C}_A=0.153[300+60(0.5)(50)]=\$271.98;$$

$$\hat{E}(t)=45.15; \qquad \hat{C}_0=45.15[50]=\$2257.50.$$

Therefore, abandon the test and initiate redesign of the computer program.

The maximum value of C_3 for which the decision remains unchanged is substantially above 300 dollars: $0.153[C_3(\max)+60(0.5)(50)]=\$2257.50$; $C_3(\max)=\$13254.90$.

(b) *Life testing.* Evaluation of the cost functions for the life-testing model requires, even for relatively simple probability distributions, the evaluation of some complicated integrals that cannot always be obtained in closed form. For example, using the one-parameter exponential model for life distribution, we have

$$f(t,\theta)=1/\theta\exp(-t/\theta), \qquad t\geqq 0,$$

$$F(t,\theta)=1-\exp(-t/\theta).$$

Therefore,

$$g(y_1,\cdots,y_k;\theta)=[n!/(n-k)!](1/\theta)^k\exp[-\textstyle\sum_{i=1}^{i=k}y_i/\theta-(n-k)y_k/\theta];$$

$$g(y_1,\cdots,y_k,y_{r+1};\theta)=[n!/(r-k)!(n-r-1)!](1/\theta)^{k+1}\cdot[\exp(-y_k/\theta)-\exp(-y_{r+1}/\theta)]^{r-k}\cdot[\exp(-\textstyle\sum_{i=1}^{i=k}y_i/\theta)]\cdot\exp[-(n-r)y_{r+1}/\theta].$$

The maximum likelihood estimate for θ[2] is

$$\hat{\theta}=1/k[\textstyle\sum_{i=1}^{i=k}y_i+(n-k)y_k].$$

Replacing θ by $\hat{\theta}$ in the density functions and simplifying, we obtain

$$\hat{P}_0=\int_{t_0}^{\infty}\frac{(n-k)!}{(r-k)!(n-r-1)!}(1/\hat{\theta})\frac{[\exp(-y_k/\hat{\theta})-\exp(-y_{r+1}/\hat{\theta})]^{r-k}}{\exp[-(n-k)y_k/\hat{\theta}]}\cdot\exp[-(n-r)y_{r+1}/\hat{\theta}]\,dy_{r+1}.$$

Calling the integrand in the above expression A, we have

$$\hat{E}(y_{r+1}-y_k)=\int_{y_k}^{t_0}y_{r+1}\cdot A\cdot dy_{r+1}-y_k.$$

It is clear that, even with the simple one-parameter exponential density function, a considerable problem exists in the evaluation of $\hat{P}_0$ and $\hat{E}(y_{r+1}-y_k)$. Numerical integration techniques were used in the two examples that follow. The details of the method are given in the next section.

Example 1. An electronic component is required to pass a performance test of 500 hours. The specification is that 20 randomly selected items shall be placed on test simultaneously, and 5 failures or less shall occur during 500 hours. The cost of performing the test is \$25 per hour. The delay cost of waiting for the design to be completed is \$80 per hour. The cost of redesign is \$5000. Assume that the failure distribution follows a one-parameter exponential model. Three failures are observed at 80, 220, and 310 hours. Should the test be continued?

We have

$$\theta=[80+220+310+17(310)]/3=1960 \text{ hours};$$

$$\hat{P}_0=\int_{500}^{\infty}\frac{17!}{14!2!}(1/1960)\frac{[\exp(-310/1960)-\exp(-y_6/1960)]^2}{\exp[-17(310)/1960]}\cdot\exp(-15y_6/1960)\,dy_6$$
$$=0.79665;$$

$$\hat{C}_A=0.79665[(500-310)(25+80)+5000]=\$19876;$$
$$\hat{E}(y_6-310)=130.05 \text{ hours};\qquad \hat{C}_0=(130.05)(25+80)=\$13655.$$

Therefore, continue the test.

Example 2. A numerically more complicated estimation problem is presented by the two-parameter Weibull distribution,[10] often used as a model in fatigue testing. Let $f(t;\alpha,\beta)=(\beta/\alpha)t^{\beta-1}\exp[-t^\beta/\alpha]$, $t\geqq 0$; $F(t;\alpha,\beta)=1-\exp[-t^\beta/\alpha]$.

Consider the following problem. A specification for an automotive hood latch is that, of 30 items placed on test simultaneously, ten or fewer shall fail during 3000 cycles of operation. The cost of performing the test is \$2.00 per cycle. The delay cost of waiting for the design to be completed is \$0.50 per cycle. The cost of redesign is \$8500. Seven failures are observed at 48, 300, 315, 492, 913, 1108, and 1480 cycles. Shall the test be continued beyond the 1480th cycle? Assume a Weibull density function for failures.

The integral equations for $\hat{P}_0$ and $\hat{E}(y_{11}-y_7)$ may be developed by the methods of part (b) of the second section, using the above density and distribution functions. However, the maximum likelihood estimates of α and β, required in these integrals, can only be obtained by iterative methods.

The appropriate likelihood equations for $y_1,\cdots,y_k$ are

$$\partial L/\partial\alpha=0=-k/\alpha+(1/\alpha^2)\sum_{i=1}^{i=k}y_i^\beta+[(n-k)/\alpha^2]\sum_{i=1}^{i=k}y_i^\beta,$$
$$\partial L/\partial\beta=0=k/\beta+\sum_{i=1}^{i=k}y_i-(1/\alpha)\sum_{i=1}^{i=k}y_i^\beta\ln y_i-[(n-k)/\alpha]y_k^\beta\ln y_k.$$

The parameter α may be eliminated from these equations, and the resulting single equation may be solved iteratively for $\hat{\beta}$.[1] [If iterative determination of these parameters is not desirable, an approximate solution for α and β may be obtained graphically by a method described by NELSON.[6]] Using the Regula-Falsi method,[3] we obtain the maximum likelihood solutions $\hat{\beta}=0.9043$, $\hat{\alpha}=2766.6$.

In turn, these estimates yield $\hat{P}_0=0.25098$, $\hat{E}(y_{11}-y_7)=397.6$. Then

$\hat{C}_A=0.25098[(3000-1480)(2+0.50)+8500]=\$3087.05$; $\hat{C}_0=(397.6)(2+0.50)=\$994.00$.

Therefore, continue the test.

COMPUTATION

IMPLEMENTATION OF the Bernoulli model requires care only in preserving accuracy in the summation operations and providing for problems with large combinatorial terms.

General use of the life-testing model requires an effective numerical integration routine. Several algorithms deserve mention.[7,8]

When the underlying life distribution is the exponential used here, an analytic solution can be augmented by integration of the power series expansion of $\hat{P}_0$:

$$\hat{P}_0=\frac{(n-k)!}{(r-k)!(n-r-1)!}\frac{1}{\hat{\theta}}\left[\sum_{i=0}^{r-k}(-1)^i\binom{r-k}{i}\frac{\hat{\theta}}{n-r+1}\cdot\exp\left\{-\left(\frac{n-r+1}{\hat{\theta}}\right)(t_0-y_k)\right\}\right].$$

This method of solution has the disadvantage of alternate-sign summation, and should be augmented by an ordered addition algorithm.

With life distributions for which the power-series integration fails, such as the Weibull, techniques based upon interpolated polynomial integration were rewarding. Representative of these, Romberg, Simpson, and Gaussian quadrature formulas[3,9] with alteration for interval halving and Richardson's extrapolation were compared for speed and accuracy. With exponential-type distributions, all three techniques functioned adequately (converging to identical solutions), with the Romberg method yielding the least satisfactory results because of the number of evaluations required for a sufficiently accurate piecewise-linear approximation. This is predicted by the error terms for the methods.

The most satisfactory general results came with the use of Gaussian integration, for example, the Gauss-Chebychev two-point formula, which is analogous to and functions at least as well as Simpson, and excels at problems that cause instability in the other techniques. Integral convergence to six decimal places has seldom taken more than five iterations, or 2^5 interval evaluations.

For the numerical examples, integration methods were verified for the exponential model by comparison of power series, interpolated polynomial, and tabulated integration. For the Weibull, application of Simpson, Romberg, and Gauss-Chebychev methods agreed to five decimal places.

REFERENCES

1. S. Dubey, "On Some Statistical Inferences for Weibull Laws," *Naval Res. Log. Quart.* **13,** 228–229 (1966).
2. B. Epstein and M. Sobel, "Life Testing," *J. Amer. Stat. Assoc.* **48,** 488 (1953).
3. P. Henrici, *Elements of Numerical Analysis,* p. 87–88, Wiley, New York, 1964.
4. P. Hoel, *Introduction to Mathematical Statistics,* p. 40–41, Wiley, New York, 1954
5. R. Hogg and A. Craig, *Introduction to Mathematical Statistics* (Second Ed.) Chap. 6, Macmillan, New York, 1965.
6. L. Nelson, "Weibull Probability Paper," *Industrial Quality Control* **23,** 452 (1967).
7. A. Ralston and H. Wilf, *Mathematical Methods for Digital Computers,* Vol. 1, Chap. 22, Wiley, New York, 1960.
8. ——— and ———, *Mathematical Methods for Digital Computers,* Vol. 2, Chap. 6, Wiley, New York, 1967.
9. ———, *A First Course in Numerical Analysis,* Chap. 4, McGraw Hill, New York, 1965.
10. W. Weibull, *Fatigue Testing and Analysis of Results,* Pergamon Press, New York, 1961.