Planning Reliability Assessment

by **William Q. Meeker, Gerald J. Hahn and Necip Doganaksoy**

Let's say you have designed a new metal spring and want to estimate the time by which 10% of such springs will fail. Many reliability tests require estimation of a percentile or quantile t_p of the distribution for time to failure—$t_{0.10}$ in this example. But how many units do you need to test and for how long?

The Basic Approach

The life test will provide an estimate, $\hat{t}_p$ of t_p, and a 95% confidence interval to contain t_p. A lower confidence bound on t_p is $\hat{t}_p/\hat{R}$, and an upper confidence bound is $\hat{t}_p \times \hat{R}$, where the precision factor, $\hat{R} > 1$, is estimated from the data. For example, if $\hat{R} = 2$, the upper confidence bound on t_p exceeds $\hat{t}_p$ by a factor of 2, and the lower confidence bound is half of $\hat{t}_p$. $\hat{R} = 1.3$ implies a much narrower confidence interval and greater precision in estimating t_p. $\hat{R}$ will depend on the sample size, n, and the test duration, t_c—the time at which unfailed units are removed from the test.

In planning a life test, you need to specify a target precision factor, R^*, to obtain a reasonably sized confidence interval. $\hat{R}$ is random, varying from one test to the next. Therefore, select R^* so the $\hat{R}$ attained exceeds R^* about half the time. Then, find a combination of n and t_c to estimate t_p with a precision factor close to R^*.

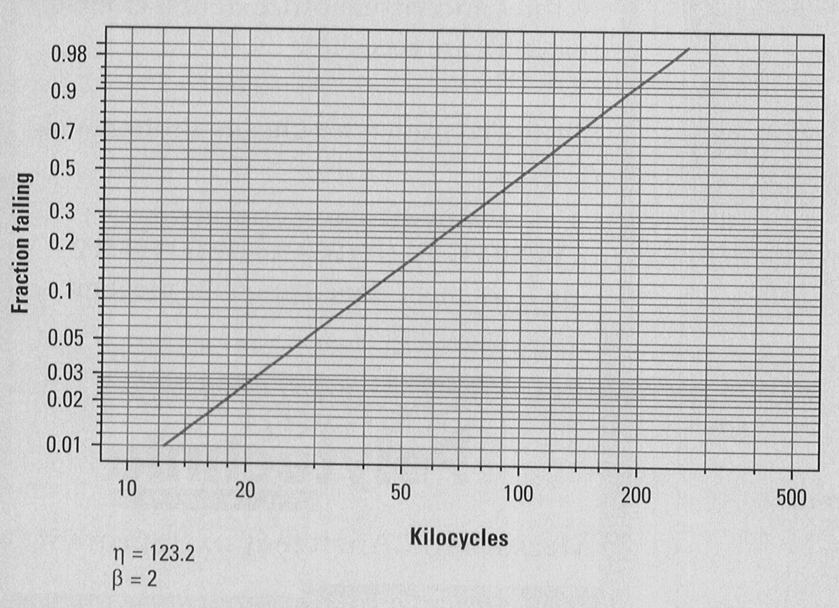
We suggest you use simulation to do this. The basic idea is for the computer to generate many samples of size n for test duration t_c to resemble the data expected from the life test and analyze the results for each such sample. Then repeat for different n and t_c to compare the resulting statistical uncertainties.

Simulation can help determine how many units to test and for how long.

The specific procedure, which from now on will focus on $t_{0.10}$, is:

- **Step one:** From past experience and engineering judgment, assume a statistical distribution for time to failure—say a Weibull distribution with initially specified values for the shape parameter β—and a planning value for $t_{0.10}$. Then, determine the assumed Weibull distribution scale parameter η, from β and $t_{0.10}$.[1]
- **Step two:** Specify an initially proposed n and t_c.
- **Step three:** Randomly generate n times to failure from the assumed distribution. Many of these randomly generated times will exceed the time t_c at which the test is terminated and are, therefore, taken as unfailed or censored at time t_c.
- **Step four:** Apply the maximum likelihood (ML) method to the simulated data to compute estimates of the parameters of the time to failure distribution, the estimate $\hat{t}_{0.10}$ of $t_{0.10}$, a two-sided confidence interval for $t_{0.10}$ and the resulting $\hat{R}$.
- **Step five:** Repeat steps three and four many times to obtain a distribution of $\hat{R}$. Then, find the distribution's geometric mean, $\bar{R}_G$, which is an estimate of the median of the distribution of $\hat{R}$. This characterizes the precision you can expect in estimating $t_{0.10}$ for the chosen n and t_c. Compare $\bar{R}_G$ with the target R^*.
- **Step six:** Repeat steps three, four and five for different n and t_c, and assess their impact on $\bar{R}_G$. From this, select n and t_c for the life test.

FIGURE 1 Assumed Weibull Distribution for Metal Spring Life



The Metal Spring Example

You have 45 representative metal springs available for testing and five machines to test the springs under cyclic compressive stress with the displacement encountered in application. You will, therefore, test nine randomly selected groups of five springs for up to t_c hours.

A cycling rate of three cycles per minute can be safely used without creating new failure modes to accelerate the test. To end the test after slightly more than two months, with each group running for a week, you determine the unfailed units run for t_c = 30 kilocycles.

Say you then want to estimate $t_{0.10}$ with a 95% confidence interval and a precision factor $R^* = 1.5$. You first must ask, "Will a life test with $n = 45$ and $t_{0.10}$ = 30 kilocycles satisfy this requirement? If not, what combination of n and t_c will?"

FIGURE 2 First Simulation of Spring Life Test

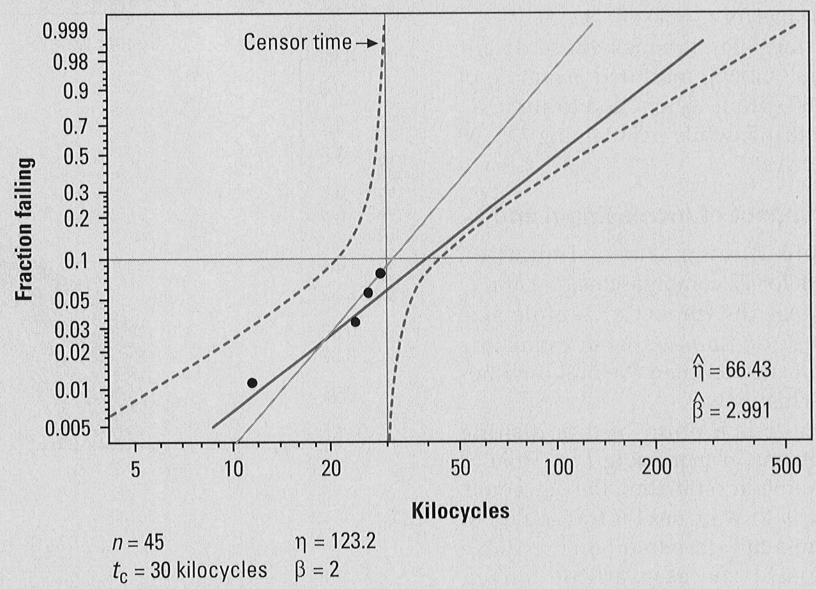
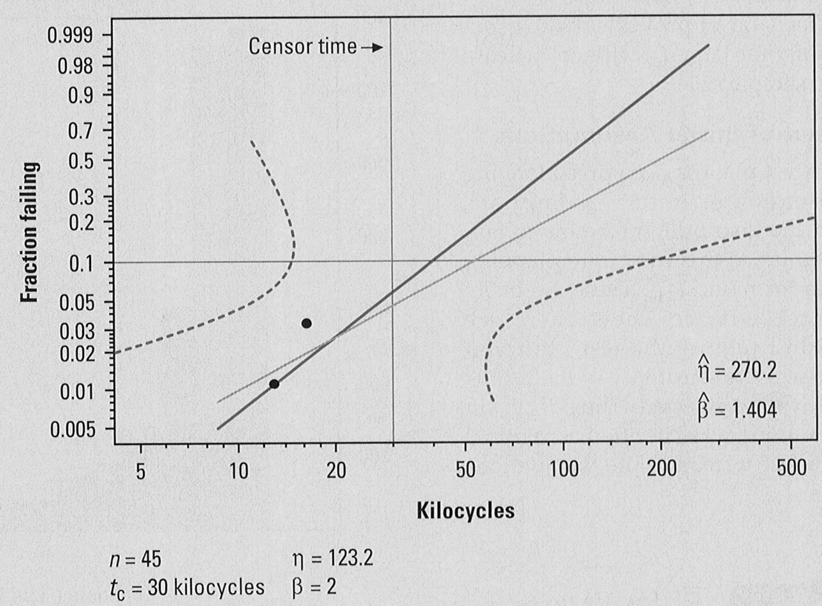


FIGURE 3 Second Simulation of Spring Life Test



Simulation Results

Determine n and t_c to achieve R^* = 1.5 by following these steps:

- **Step one:** Time to failure is assumed to follow a Weibull distribution with $\beta = 2$ and $t_{0.10} = 40$ kilocycles, implying $\eta = 123.2$ kilocycles (see Figure 1).
- **Step two:** Consider $n = 45$ units and $t_c = 30$ kilocycles.
- **Step three:** The computer randomly generates 45 times to failure from the assumed Weibull distribution. Four of these times (11.5, 24.0, 26.3 and 28.7 kilocycles) were less than 30 kilocycles. The remaining 41 values exceeded 30 kilocycles and were taken to represent unfailed springs at 30 kilocycles.
- **Step four:** The data yielded the ML estimates $\hat{\eta} = 66.43$ kilocycles, $\hat{\beta} = 2.991$ and $\hat{t}_{0.10} = 31.30$ kilocycles. An approximate 95% confidence interval to contain $t_{0.10}$ is [22.48, 43.59]. Thus, $\hat{R} = 43.59/31.30 = 1.39$. The solid gray and brown lines in Figure 2 show the resulting fitted time to failure distribution and the distribution implied by the planning values, respectively. The dashed lines are (pointwise) 95% confidence intervals on t_p for different values of p.

  A second simulation is shown in Figure 3. It resulted in only two failures, the estimate $\hat{t}_{0.10} = 54.37$ kilocycles and a much wider 95% confidence interval [14.93, 197.95], giving a value of $\hat{R} = 3.64$.
- **Step five:** Steps three and four are repeated to obtain a total of 5,000 simulations. The Weibull distribution fits for the first 50 simulations are shown in Figure 4 (p. 92). The large variability in the $\hat{t}_{0.10}$ estimates is evidenced by the spread in the fitted lines crossing the horizontal axis at 0.10—showing values of $\hat{t}_{0.10}$ ranging from about 28 to more than 500 kilocycles for the 50 simulations, with $n = 45$ and t_c = 30 kilocycles.

  Figure 5 (p. 92) shows the 5,000 values of $\hat{R}$. The geometric mean of these values is $\bar{R}_G = 2.5$, appreciably exceeding the targeted $R^* = 1.5$.
- **Step six:** To achieve the specified R^*, either n or t_c—or both—need to be increased. In this application, it is difficult to increase the sample size beyond $n = 45$, but it is feasible to test unfailed springs to t_c = 50 kilocycles by extending the test duration to nearly four months. Steps three to five were, therefore, repeated for $n = 45$ and $t_c = 50$.

The results are shown in Figures 6 and 7 (p. 93). These indicate much less spread and reduced $\bar{R}_G$ to 1.55. This was close enough to $R^* = 1.5$. The life test was, therefore, conducted with n = 45 springs for $t_C = 50$ kilocycles.

Further simulations, whose details are not shown, indicated a sample of $n = 180$ springs is required to attain $\hat{R}_G$ close to 1.5, while maintaining $t_C = 30$ kilocycles.

The Impact of Increasing n and t_C

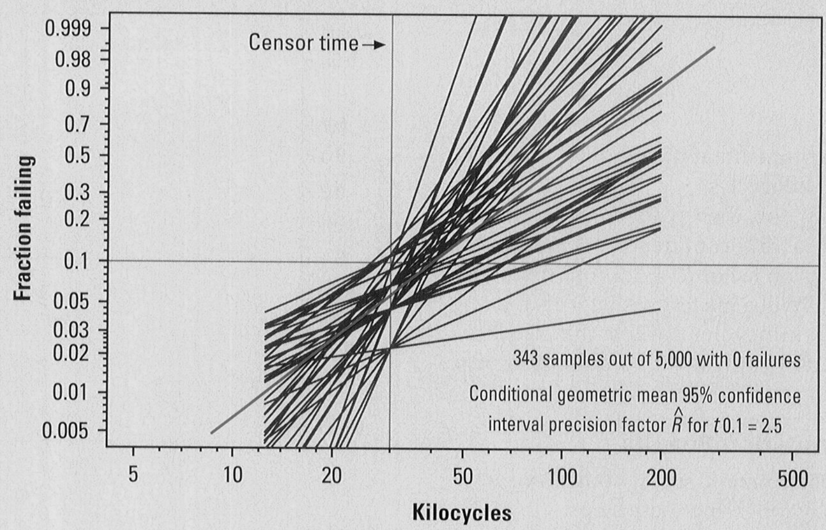
Table 1 summarizes simulation results for 12 combinations of n and t_C to assess the impact of sample size and test duration on $\overline{R}_G$ in estimating $t_{0.10}$ for the assumed Weibull distribution. This shows:

- There is a point of diminishing returns in increasing t_C. A rule of thumb in estimating the percentile, t_p, is to wait until a fraction, p, of units fail (in parentheses in Table 1) or, preferably, a little longer. Waiting much longer provides little added information unless you also need to estimate t_p for a larger value of p, say 50%.
- Precision improves slowly, especially for large t_C, with an increase in sample size.

Impact of Initial Assumptions

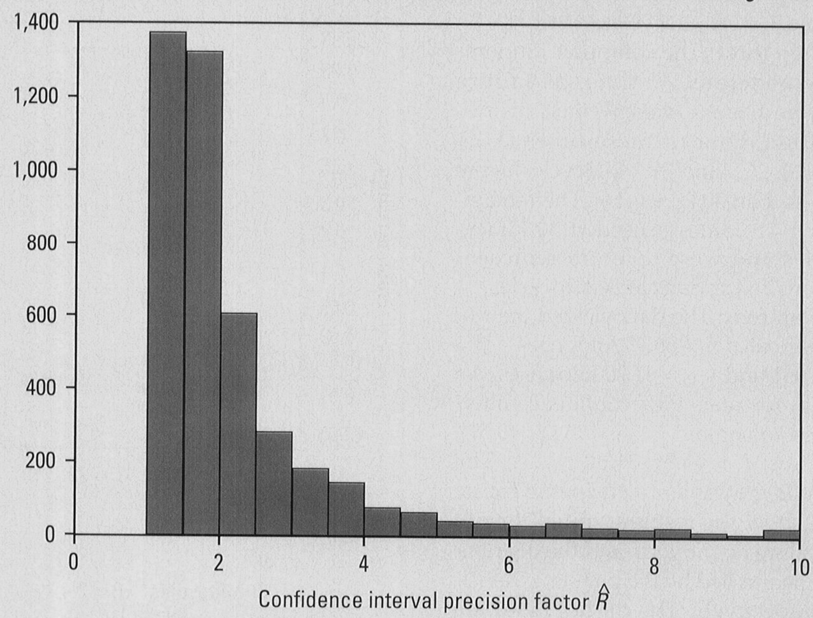
Life test planning requires planning information or initial assumptions about the distribution for time to failure, which in this case were planning values for η and $t_{0.10}$. These are likely incorrect—if we knew them we wouldn't require the test—but you can assess the sensitivity of the results to the values selected. Thus, if $t_{0.10}$ in our example was specified as 20 or 60 kilocycles (instead of 40), simulations indicate 45 springs need to be tested for 25 and 75 kilocycles, respectively.

FIGURE 4 Summary of 50 Simulations With n = 45 and t_c = 30



n = 45 η = 123.2
t_C = 30 kilocycles β = 2 with E(r) = 2.59 (expected number of failures)

FIGURE 5 $\hat{R}$ From 5,000 Simulations With n = 45 and t_c = 30



Six values of $\hat{R}$ exceeding 10 were omitted.

TABLE 1 $\overline{R}_G$ for Various Combinations Of n and t_c

Test duration (t_c) in kilocycles	Sample size (n) 45	90	180
30 (0.06)	2.50	1.87	1.49
50 (0.15)	1.55	1.34	1.23
100 (0.48)	1.47	1.32	1.21
200 (0.93)	1.41	1.28	1.19

The expected proportion failing is in parenthesis.

Implementation Issues And Extensions

We discussed only the technical question of determining n and t_C, but there are many other things to consider in planning a life test. For example, test units should be representative of the population of interest, and the test must closely resemble the applications environment.

Also, the methods for determining sample size and test duration can be readily modified to plan test programs for other situations, such as:

- Unequal test duration—for example,

testing one-third of the units for 30, 50 and 70 kilocycles, respectively.

- A different time to failure distribution, such as the lognormal.
- Accelerated testing[2] or degradation testing.[3]
- Demonstrating high reliability over a defined lifetime.[4]

Statistical Analysis And an Alternative Method

Once the life test has been conducted, you can analyze the data to obtain an ML estimate and a confidence interval for t_p.[5,6]

A mathematical formula[7] provides an alternative for determining n and t_c, which can be used instead of simulation or to suggest starting values for n and t_c in the simulation.

NOTE

The approach described here was conducted using the SPLIDA package (available at www.public.iastate.edu/~splida) of S-Plus 6 (available from Insightful Corp.). The procedures have also been programmed in the JMP scripting language for release 6 of JMP (available from the SAS Institute). The simulations can also be implemented using other commercially available packages, but some programming skill is required.

REFERENCES

1. W.Q. Meeker and L.A. Escobar, *Statistical Methods for Reliability Data*, Section 2.8, John Wiley and Sons, 1998.
2. G.J. Hahn, W.Q. Meeker and Necip Doganaksoy, "Speedier Reliability Analysis," *Quality Progress*, June 2003, pp. 58-64.
3. W.Q. Meeker, Necip Doganaksoy and G.J. Hahn, "Using Degradation Data for Reliability Analysis," *Quality Progress*, June 2001, pp. 60-65.
4. W.Q. Meeker, G.J. Hahn and N. Doganaksoy, "Planning Life Tests for Reliability Demonstration," *Quality Progress*, August 2004, pp. 80-82.
5. Meeker and Escobar, *Statistical Methods for Reliability Data*, Section 8.4.3, see reference 1.
6. Necip Doganaksoy, G.J. Hahn and W.Q. Meeker, "Product Life Data Analysis: A Case Study," *Quality Progress*, June 2000, pp. 115-122.
7. Meeker and Escobar, *Statistical Methods for Reliability Data*, Section 10.5.4, see reference 1.

FIGURE 6 Summary of 50 Simulations With $n = 45$ and $t_c = 50$

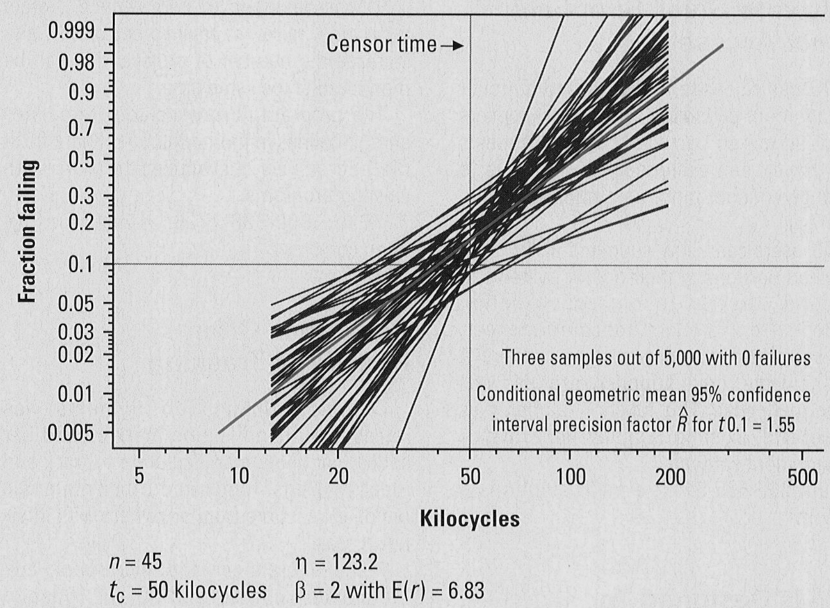
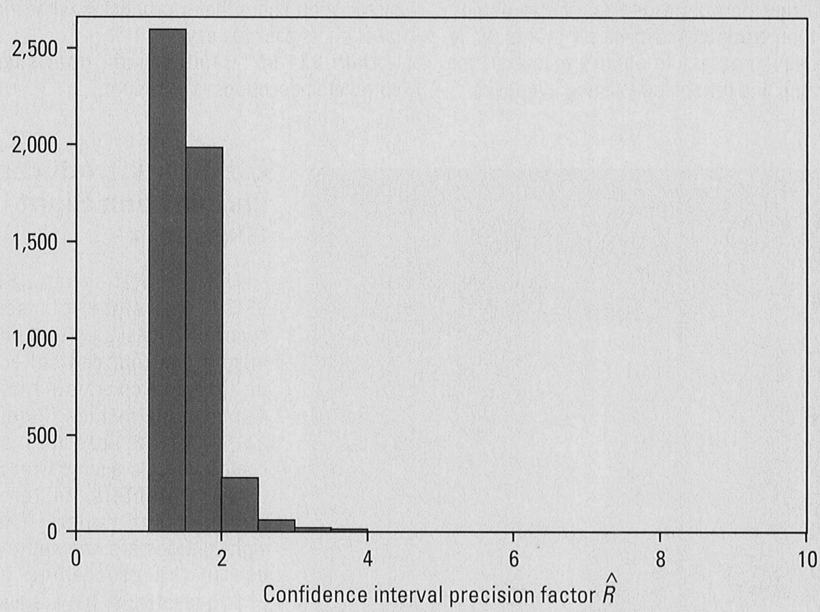


FIGURE 7 $\hat{R}$ From 5,000 Simulations With $n = 45$ and $t_c = 50$



WILLIAM Q. MEEKER *is a professor of statistics and distinguished professor of liberal arts and sciences at Iowa State University in Ames. He has a doctorate in administrative and engineering systems from Union College in Schenectady, NY. Meeker is a fellow of the American Statistical Assn. and a Senior Member of ASQ.*

GERALD J. HAHN *is a retired manager of applied statistics at the GE Global Research Center in Schenectady, NY. He has a doctorate in statistics and operations research from Rensselaer Polytechnic Institute in Troy, NY, where he is also an adjunct faculty member. Hahn is a Fellow of the American Statistical Assn. and ASQ.*

NECIP DOGANAKSOY *is a statistician and a Six Sigma Master Black Belt at the GE Global Research Center in Schenectady, NY. He has a doctorate in administrative and engineering systems from Union College in Schenectady, NY. Doganaksoy is a fellow of the American Statistical Assn. and a Senior Member of ASQ.*

Please comment

If you would like to comment on this article, please post your remarks on the *Quality Progress* Discussion Board at www.asq.org, or e-mail them to editor@asq.org.