File: /General/MLAB-Text/Papers/degrade/degrade.tex

Failure Probability Estimation with Accelerated Degradation Testing

Gary D. Knott, Ph.D.
Civilized Software, Inc.
12109 Heritage Park Circle
Silver Spring, MD 20906 USA
Tel. (301) 962-3711
URL: www.civilized.com

The problem of estimating the probability of a failure event arises in many diverse situations. In quality control for example, we may have a batch of identical electrical components such as a transducer of some kind, or a batch of identical mechanical components such as a shock absorber or a compressor. We wish to assess the reliability of these components in terms of the *load* that can be handled without substantial loss of normal response. Time may also play a role; the response may *degrade* over time in a manner which is *accelarated* in the presence of a high load.

Another example arises where a population of patients are being treated with a drug or other treatment which produces undesiable side-effects over time. The "amount" of treatment becomes the load, and the level of side-effects becomes the response. In this case, we may numerically measure response as 100-(% disfunctionality-due-to-side-effects), so that a low response indicates a loss of function.

Let the response at time t of a random subject subjected to the load x be specified by $f(t, x, \beta) + \epsilon$ where β is a vector of unknown parameters occurring in the response function, and ϵ is a mean 0, variance σ^2 normal random variable representing the error in response measurment and the effect of inter-subject variability.

One commonly applicable response model function that we shall use here is the four parameter logistic form: $f(t, x, \beta) =$ the value y such that $\beta_3 g(y)^{\beta_1} t + \beta_4 g(y)^{\beta_2} x = 1$, where $g(y) = y/(1-y)$. Usually, $\beta_1 > 0, \beta_2 > 0, \beta_3 > 0$ and $\beta_4 > 0$.

Now suppose we have measured the degradation responses $y_1, y_2, \ldots, y_n$ of a set of n subjects at the times $t_1, t_2, \ldots, t_n$, after having been subjected to the loads $x_1, x_2, \ldots, x_n$ respectively. There are various complications that can arise, such as the case where the measured responses are not independent samples, in which case, we may have a so-called *repeated measures* study. If the loads are higher than usual, we have an *accelarated* study which, of course, raises the question as to the adequacy of our model for such loads. Finally, the times $t_1, t_2, \ldots, t_n$ may be such that only modest response changes are seen. When this happens, our model must be adequate to extrapolate to longer unobserved times.

With due attention to the above issues, we may write:

$$\begin{aligned} y_1 &= f(t_1, x_1, \beta) + \epsilon_1 \\ y_2 &= f(t_2, x_2, \beta) + \epsilon_2 \\ &\vdots \\ y_n &= f(t_n, x_n, \beta) + \epsilon_n \end{aligned}$$

where $\epsilon_1, \ldots, \epsilon_n$ are independent samples of the mean 0, variance σ^2 random variable ϵ.

We can then estimate the unknown parameters β by curve-fitting the model $f(t, x, \beta)$ to the data $(t_1, x_1, y_1), (t_2, x_2, y_2), \ldots, (t_n, x_n, y_n)$. We may then estimate σ as the standard deviation $\hat{\sigma}$ of the deviations $y_i - f(t_i, x_i, \hat{\beta})$ where $\hat{\beta}$ is our obtained estimator for β.

Let Y_{tx} be the random variable $f(t, x, \beta) + \epsilon$. Now, after estimating β, we may approximate the probability $P(Y_{tx} \le c)$ by $P(f(t, x, \hat{\beta}) + \epsilon \le c)$. But then $P(Y_{tx} \le c) \approx P(\epsilon/\hat{\sigma} \le (c - f(t, x, \hat{\beta}))/\hat{\sigma})$, and $\epsilon/\hat{\sigma}$ is approximately a mean 0, variance 1 normal random variable with the distribution function $\Phi(v) = \frac{1}{\sqrt{2\pi}\hat{\sigma}} \int_{-\infty}^{v} e^{-u^2/2} du$ and the density function $\phi(v) = \frac{d\Phi(v)}{dv}$.

Thus, $P(Y_{tx} \le c) \approx \Phi(c - f(t, x, \hat{\beta})/\hat{\sigma})$, and the density function of Y_{tx} is approximatedy $p(t, x, c) := \phi(c - f(t, x, \hat{\beta})/\hat{\sigma})$. The graph of the function $p(t, x, c)$ is also called the *response probability surface*. In particular, the graph of $p(t, x_0, c)$ is the response probability surface for a given load x_0. The volume under this surface over any (t, c)-region R is the probability that the load-x_0 response Y_{tx} will equal c at time t for some (t, c)-pair in R. Similarly the graph of $p(t_0, x, c)$ is the response probability surface at a given time t_0.

Let the response value c_0 or less be considered a *failure*. Then $P(Y_{tx} \le c_0)$ is the probability that the failure state is exhibited at time t, subject to load x. We are often interested in the distribution of the failure time T_x which is the earliest time for which $Y_{tx} \le c_0$. When f is a monotonically-decreasing function of t, the random variable T_x can be defined as the value of t such that $f(t, x, \beta) = c_0 - \epsilon$. Sometimes we can explicitly write T_x

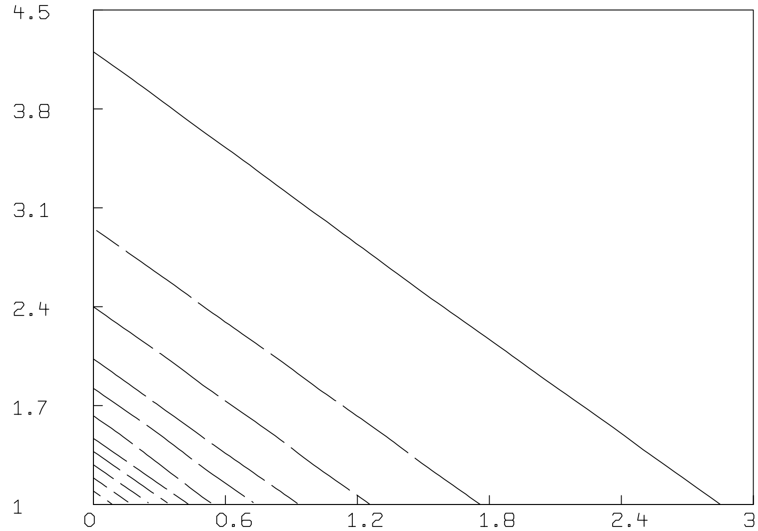
in terms of the random variable ϵ; however in all cases we can estimate the distribution of T_x by means of a Monte-Carlo simulation. We shall demonstrate this in the example below.

An Example

The *MLAB* mathematical and statistical modeling system is well-suited to perform the parameter-estimation, simulation, and graphics required for a failure probability analysis as described above.

To begin, let us read-in the (t, x, y) data points.

```
m = read(data,120,3)
/* degrade.do -- Accelarated Degradation Testing */
reset
echodo  = 3

fct g(y) = 1/(B1-B2*y)-1
fct h(t,x,y) = B5*g(y)^B3*t+B6*g(y)^B4*x-1

fct f(t,x) = root(y, 1.001, 1.999, B5*g(y)^B3*t+B6*g(y)^B4*x-1)

constraints q = (B1>1, B2>0, B3<0, B4<0, B5>0, B6>0)

data = cross((0:10:.5),(1:10:.5))

B1=2; B2=1; B3=-1/2; B4=-1/2; B5=3/2; B6=4/3

m = points(f,data)

draw contour(m) lt svmarker

view
```


4.5
3.8
3.1
2.4
1.7
1
0
0.6
1.2
1.8
2.4
3