The Fallacy of Reliability Prediction in Automotive Engineering

by

T P Davis,
Small & Medium Car Vehicle Centre,
Ford Werke AG,
50742, Köln, Germany

Introduction.

In this short article, I discuss some problems in automotive engineering with regard to setting reliability targets and trying to make predictions about reliability in the field. It is argued that improving reliability is more important than predicting it.

Definition of reliability.

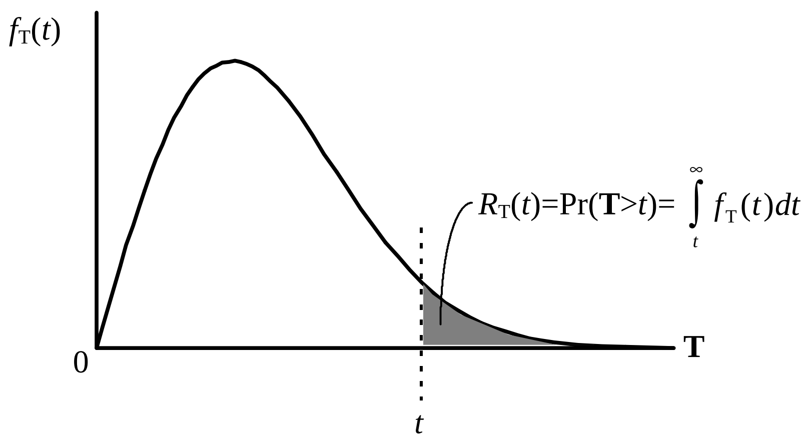
Define a random variable, **T**, which is the time (or mileage) to some critical event (e.g. total failure of function, or a specified degradation of function) of an automobile component *in its field of operation*. The random variable **T** has an associated probability density, denoted $f_T(t)$, with $t>0$ necessarily. A common choice for $f_T(.)$ in automotive engineering is the Weibull density, but what follows is true for any probability distribution.

The *reliability* of the system at time t, which we will denote by $R_T(t)$, is defined as

$$R_T(t)=\Pr(\mathbf{T}>t)=\int_t^\infty f_T(t)dt \qquad (1)$$

where Pr(.) means "the probability of" the expression in brackets. So in words, reliability at time t is defined as the probability that the failure time (or mileage), **T**, of the component in its field of operation exceeds the value t. For example t might be 10 years or 150,000 miles.

Expression (1) can be illustrated graphically as follows, the reliability at t being equivalent to the shaded area under the probability density curve to the right of t. It should be obvious that $R_T(t)$ is a decreasing function of t.



Reliability *prediction* is concerned with estimating the probability $R_T(t)$ given by (1). The perceived need to make these predictions is often driven by the setting of so-called reliability targets for the component in the field. By contrast, reliability *improvement* is concerned only with maximizing $R_T(t)$. Of course, we cannot do better than to maximize $R_T(t)$, and we show later how we can know we have maximized it, even though we show that we cannot estimate or predict it.

The mathematics of prediction.

Unfortunately, the expression given by (1) is not nearly as innocent as it looks, because in practice the distribution of **T** is not homogeneous across the entire population of components in the field - it depends on other variables. These can be classified as *a*) *design variables* (or control factors in robustness-speak), and *b*) *error variables* (or noise factors). Because values for the design factors are fixed prior to the product being released to the field, we need not consider these any further here. On the other hand, noise factors and their associated effects on the component are very much in evidence after the product is released to the field, and therefore have a major impact on $R_T(t)$ and on our ability to estimate it.

To see this, define the collection of all possible noise factors that impact the component by **N**; then **N** is a multi-dimensional random variable made up of contributions from various noise factor sources, for example piece-to-piece variation in dimensions, wear-out over time, customer duty cycles, climate & road conditions, and interactions with neighbouring components in the system.

If we denote:-

> the joint probability density of **T** and **N** by $f_{T,N}(t,n)$,
> the conditional density of **T** *given* a fixed point in the noise space, $n \in \mathbf{N}$ by $f_{T|N}(t|n)$,
> and the marginal density of **N** alone by $f_N(n)$,

then by standard results in distribution theory, (1) can be re-expressed as

$$R_T(t) = \int_t^\infty \int_{n\in N} f_{T,N}(t,n)\,dn\,dt = \int_t^\infty \int_{n\in N} f_{T|N}(t|n)\, f_N(n)\,dn\,dt = \int_{n\in N} R_{T|N}(t|n)\, f_N(n)\,dn \qquad (2)$$

where $\int_{n\in N}$ is a multiple integral over the entire noise space, and

$R_{T|N}(t|n) = \int_t^\infty f_{T|N}(t|n)dt$ denotes the reliability at t conditional on a given point in the noise space. In probability notation, $R_{T|N}(t|n) = \Pr(\mathbf{T}>t|n \in \mathbf{N})$, the vertical bar "|" being read as "conditional on".

Hence, we can see from (2) that the reliability value $R_T(t)$ requires information about both $f_{T|N}(t|n)$ and $f_N(n)$. Some information may be available about $f_{T|N}(t|n)$ - for example, a rig or vehicle test used for "signing off" the component represents a particular, specific value in the noise space of **N**, n_R say, so results from the test (if the sample size were large enough) could be used to estimate (rather than confirm) the reliability of **T** *conditional* on this n_R value. We can denote this reliability by $R_{T|N}(t|n_R)$. Extensions to cover the range of **N**, not just n_R specifically, have to rely on knowing the relationship between $R_{T|N}(t|n)$ and *all* values of n contained in **N**. In some simple cases, where the noise space is small, this might be achieved by application of known theory (for example *Miner's rule* which states that the failure rate of a component is directly proportional to the applied stress, or the *Arrhenius model* which says that the average log-failure time decreases with increasing temperature), or by empirical deduction from simple experiments conducted over different noise factor levels. In practice however, because **N** is typically multi-dimensional, the technical problems associated with finding the relationship between $R_{T|N}(t|n)$ and n are considerable.

Unfortunately, an even more serious problem in evaluating (2) is estimating $f_N(n)$ - the marginal multivariate density of the noise factors. This we do not, and cannot know, because of the complexity of **N** - think about the different driving styles of automobile users, and the different climatic and road conditions across the world where vehicles are in use.

REMARK: Other industries, particularly nuclear and aerospace, have much simpler noise spaces and smaller populations in the field about which to worry. This coupled with more thorough field data that are not contaminated with the effects of a warranty policy, explains why reliability target setting and reliability estimation through (2) may be more viable in these industries.

Reliability improvement.

Fortunately, all is not lost. We can see directly from (2) that to maximize $R_T(t)$ we should maximize $R_{T|N}(t|n)$ over all possible n-values, which, assuming that the relationship between **T** and **N** is monotonic (i.e. the more severe the noise condition, the lower the reliability, which is exactly what Miner's rule, and the Arrhenius theory, says), is equivalent to maximizing just $R_{T|N}(t|n_R)$. That is, to maximize $R_T(t)$, we should aim at maximizing our performance on tests with a known "representative" point, n_R, in the noise space.

Techniques to achieve this improvement include using the testing facility (e.g. rig or CAE model) to run statistically designed experiments in the already mentioned design variables against the appropriate noise factor combination which represents n_R, and selecting the design configuration which has the best performance (i.e. least degradation). These types of experiments are sometimes called *Parameter Design* experiments (for example, see Taguchi (1986)).

Note that $R_T(t)$ is *not the same thing* as $R_{T|N}(t|n_R)$, although this important distinction often gets blurred, particularly when using associated statistical concepts such as confidence intervals to assess reliability from test results. In other words, we cannot translate estimates of $R_{T|N}(t|n_R)$ into numerical predictions about $R_T(t)$. Essentially, this is the point made by W Edwards Deming regarding the difference between enumerative and analytical studies (see pages 131-133 of Deming's classic 1986 book *Out of the Crisis*). The quantity $R_{T|N}(t|n_R)$ is an enumerative evaluation of some existing test results, whereas numerical prediction about

$R_T(t)$ is analytical, because it involves unknown and unknowable properties of the noise-space. The only prediction (hence inference) that should be made about $R_T(t)$ is whether a particular design action on the component will improve it or not.

Summary.

In summary, trying to make reliability predictions and the setting of associated targets, by estimating and making inference about $R_T(t)$ as given by (1) is not tractable because of expression (2).

It is worth finishing this short article with quotes from two relevant sources, which sum up the main arguments.

Firstly, Hahn and Meeker (1994) say, in discussing the differences between enumerative and analytical studies,

> "... a reliability engineer should question the assumption that the results of a laboratory life test will adequately predict field failure rates" (pages 8&9),

and Clausing (1994) says

> "The traditional product development objective is to quantify the expected reliability in the field. A much better objective is to maximize the improvement in reliability. The two are incompatible; we must choose between them. Precise quantification of the expected reliability requires extensive testing of one product configuration. Improvement is achieved by many systematic changes to the configuration. The traditional quantification of reliability has developed sophisticated analyses and methodology.... specialized brilliance in answering the wrong question" (pages 11&12).

References.

Clausing, D. (1994). *Total quality development*. ASME Press, NY.

Deming, W.E. (1986). *Out of the crisis*. Massachusetts Institute of Technology, Center for Advanced Engineering Studies, Cambridge, MA.

Hahn, G.J. & Meeker, W.Q. (1994). "Statistical intervals for quality assurance: the realities". *ASQC, Statistics Division Newsletter*, Vol. 14, #4. Part 1, pages 6-10. Part 2, pages 20-25.

Taguchi, G. (1986). *System of experimental design (Vol. 2).* Unipub, Kraus International Publications, White Plains, NY.

Dr. Timothy P. Davis is a Chartered Statistician and a quality manager in the automotive industry. He is a member of ASQ.