

RELIABILITY

DETAILED OVERVIEW

In military testing and operational analysis *reliability* generally refers to *how, why, and when* system hardware and software failures occur: a system is *reliable* if malfunctions or failures occur infrequently, and have small impact on mission success. Otherwise it is *unreliable*. The *degree* of reliability, and types of failure are important, and should be *quantified*.

Formal Definition: *mission reliability* is the ability of an item to perform its required functions for the duration of a specified mission. This definition applies directly to items of continuous use, such as engines, communication devices, and sensors.

Ability is usually assessed as a *probability* or a *mean value*, for example, mean operating time between failures (MOTBF or MTBF; MTBOMF = “mean time between operational mission failures” is also common terminology, but longer). An alternative definition, applicable to single-use items such as missiles or ammunition, is the probability of single-shot success. In the T&E process the above probabilities must be compared to requirement numbers. Such measures are always conditional on conditions like environment (heat, cold) and transport shocks. They are subject to uncertainties that require control and understanding.

A primary function of T&E is to *discover* and *manage* or *control failure modes*; these are identifiable specific weaknesses in system design, manufacture, and typical field usage that can reduce system reliability. At some point this process must end: formally, if the item reliability meets a specific requirement it is acceptable for field usage *on that dimension*, and for specified missions. A further T&E task is to quantify maintenance, repair, and logistics needs that must be satisfied to enable the specified reliability to be achieved and maintained in field use.

Note that failures caused by enemy action (hits in combat, Electronic Warfare) are not usually classed with ordinary mechanical or software failures, or failure modes. *But*, for example, it is important to discover that an aircraft ejection seat does not operate if its activation mechanism is quite vulnerable to enemy fire.

- **Questions:**

(a) In the planned T&E process, how can failure modes be found, removed, or adequately reduced in effect, early, safely, and economically? This is a re-design or modification issue. Search for failure event *precursors* is important: there may be corrosion, fatigue cracks, local heating or vibration.

(b) Are there suitable workarounds in the field, when and if failures do occur?

(c) What are (cost-effective) alternatives for achieving reliable, if somewhat lower, performance in the event of failure? Can the system gracefully degrade? Are there particularly damaging

common-mode failure possibilities that are operationally and economically highly costly? Can re-decision help?

(d) After how much time, or how many operational actions, do failures tend to happen? Have high “infant mortality” effects been removed (by burn-in); is wear and aging understood and controllable? Can testing be done with components/subsystems of different ages and previous stress histories? How do failure types (from different modes) affect repair/restoration times to some useful level of operational utility?

(e) What are logistics support requirements for spare parts, maintenance personnel capability, instrumentation, documentation? These may change (increase) over system life.

(f) Is there adequate evidence that ORD stated requirements are satisfied? These under defined operational conditions?

- **Methods for Control of Failure Sources:**

- (1) *re-design* of the system

- (2) *redundancy*; this may be designed in, but its operation must be checked. *Software* (programs) cannot be made redundant.

- (3) *prevention*: on-line diagnosis to automatically predict and forestall catastrophic failures (Examples: vibration, heating, crack-growth monitoring to forestall helicopter rotor blade breakage, main engine breakdown)

Note: automated systems can fail and give false or no alarms. These must be tested.

RELIABILITY: TESTING MANAGEMENT ISSUES

- *Reliability Requirements*: define early in program.
 - ~ Include all likely/meaningful failure types, causes.
 - ~ Spotlight possible problems early, using DT experience, examination of similar systems.
- *Confront* reliability testing requirements:
 - ~ investment: facilities, personnel, and time-on-test, under different conditions vs. uncertainty concerning faults remaining, age/wearout effects; parameter values;
 - ~ justify appropriateness/timeliness of next testing state;
 - ~ balance testing against costs of field failures.
- Use “*reliability growth*” predictions cautiously:
 - ~ apply appropriate models (use alternatives);
 - ~ assess uncertainty of fitted model;
 - ~ use appropriate analogous-system experience.
- *Software and computer-related faults* often occur and can be highly detrimental to field success, and should be anticipated.
 - ~ Check for consequences (use M&S);
 - ~ check for occurrence in analogous systems.

QUANTIFICATION OF RELIABILITY

Measurements

- The reliability of *continuously-operating* systems, such as ships, land vehicles, aircraft, and certain communication and sensor systems, is typically measured by *operating times to failure* (sometimes called system *lifetimes*); these times are measured *from* a moment at which the system becomes operative or up, for example, just following a successful repair, *to* the next failure (repair requirement); *time* is a simple, but incomplete, measure of exposure to failure; another starting point is at turn-on: power switch (attempted) activation or engine start attempt may coincide with a failure. The type or *mode* of failure should be recorded in detail. Another measure of exposure to failure is *number* of stressful events, such as aircraft landings, particularly on carriers. For systems such as trucks, tanks, and other vehicles, a better measure of exposure may be *distance* (e.g. miles) *between failures*. For land-operating platforms, *terrain*, *vehicle speed*, and *loading* all influence distance between failures, or number of aircraft landings between failures.

- Times or distances to failure typically *vary* considerably and unpredictably; an average of measured times (or distances) is *one* common summary, but simplistic overall averages may obscure evidence of learning or “reliability growth” that often occurs when testing new items under different operating conditions; even carefully constructed averages often involve small numbers of times or events to failure and, hence are uncertain estimates and predictors of unknown quality. Graphical methods are useful for picking up changes or trends in actual data.

Note: it is important to distinguish item *removals* from item *failures* or preventive repairs. Removals may be in anticipation of failures, and be operationally sensible, *but* also be subjectively driven.

- Successive system failures can occur in different subsystems, or from different failure sources or modes; multiple component or subsystem failures may occur on an occasion of system failure, possibly from a *common cause*, such as an environmental surge. These must be recorded. Both the individual times, and the failure types, will typically depend on system environment and handling. A system failure that occurs as a result of an external or internal shock (mechanical, electrical, heat ...) can involve several components or subsystems (modes), causing them to fail simultaneously, or to produce a failure cascade.

- Premature failures may be the result of inappropriate and/or incomplete repair. This event type should be part of the database accumulated.

Note: quick and partial repairs to partial performance may be an operationally suitable and effective strategy.

- The reliability of on-demand or one-shot systems such as guns, missile launchers, and ammunition or missiles is typically quantified by the *number of* successes (good shots, but not necessarily hits) out of a *specified number of trials* (trigger-pulls). If the chosen quantification is

the *number of trials until/to failure* then there is an analogy with the operating-time-to-failure concept.

- Averages and probabilities (estimated) are often used as summaries: fraction of trials (shots, or missions) that are successes from the perspective of hardware and software designed operation. Such fractions may be unreliable estimates and predictors if they haphazardly combine many different operational conditions, and be uncertain (non-predictive) if they are made up of very small numbers. Field testing must be designed to obtain meaningful basic measurements of such *responses* (“dependent variables”) as times, or trials, to failure *and* influential factors (“independent variables”) likely to affect such responses. Details in *Statistics for Testers*.

Note: it is first important to discover and remove or manage physical design faults and vulnerabilities, then to measure or quantify their probability of causing mission failure.

Data Collection

- Meaningful failure data collection on systems in an OT environment is challenging and requires care. The aim of collecting Effectiveness data (e.g. on kills, detections, range, etc.) may dominate the need for adequate Suitability data. This tendency should be resisted. It is desirable to collect actual operating times to failure or trials to failure, and the identities of subsystems involved plus failure modes that have been identified previously, or come as a surprise. These are important Suitability test responses. Summaries of these responses can be related to (“explained by” in statistics-talk) possible causative influences such as environment, heat/cold, fuel quality (engines), operation and maintenance. Study of behavior of analogous systems can suggest the magnitudes of a to-be-tested system’s times to failure and failure modes; the latter must be reported as possible candidates for re-design. It is also important to record and study as many as possible causative influences on the basic reliability responses. In some cases data is made available as counts of failures over some time period (failures per month). These data are not straightforwardly interpreted unless the exposure-to-failure time is known, e.g. running or operating time. In many cases operating or especially flying time is less a measure of stress, and predictor of failure, than is the number of landings (possibly takeoffs); a carrier landing, especially under heavy weather conditions, may be stressful, and the stresses may accumulate to cause failure.

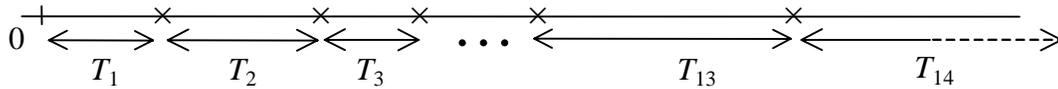
- Test data may be “dirty”: complete times to failure may not be observed (censored) since allowed test time ran out, observations are not recorded or incorrectly recorded, etc. For some procedures, see *Statistics for Testers*.

CAUTION: No foolproof statistical “tricks” are available to compensate for improper data element definition, collection and recording practices that produce unreliable, incomplete, inaccurate, or invalid data entries.

Data Presentation and Exploration

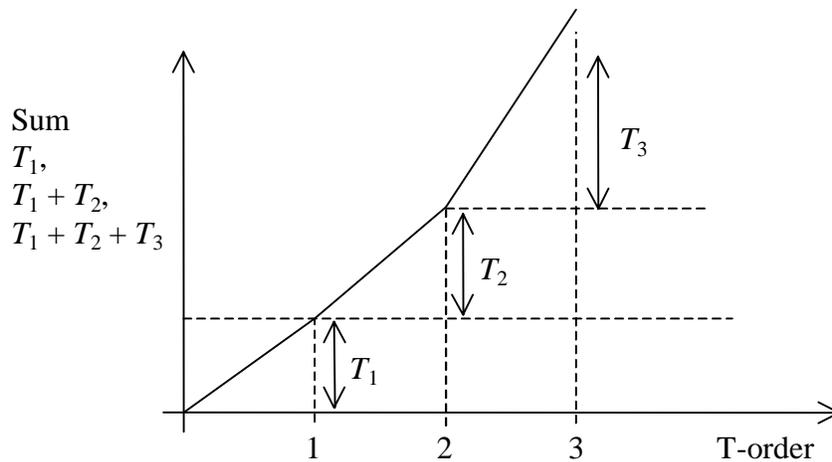
Initial data examination, especially of times to between failure, symbolically $T_1, T_2, T_3 \dots$ where T_1 is the (operating) time to first failure, T_2 is the (operating) time from the moment of system repair completion to the second failure, etc., can be informatively examined by a

Time-Line Plot



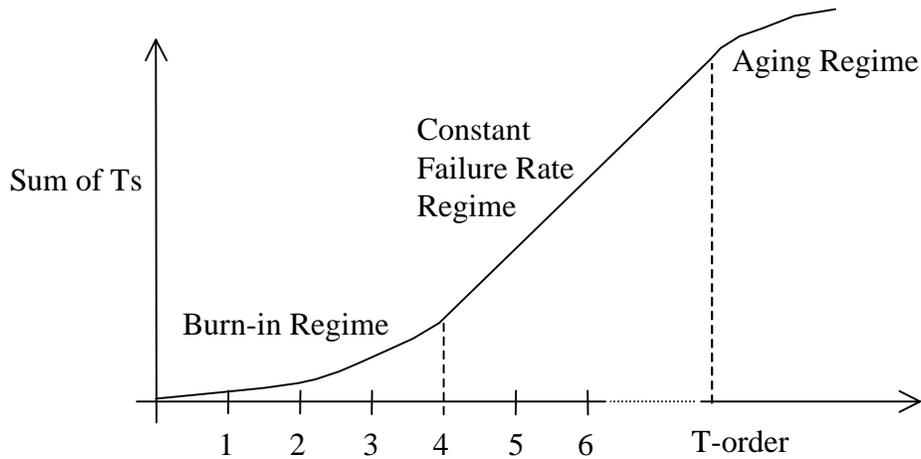
Further information is conveyed if the *location* or *source* of failure is recorded: e.g. if the points refer to a truck, \bullet for engine, \times for tire, \square for transmission, \dots . One can look for bunching of points at the beginning, with stretchout occurring later: this indicates that “infant mortality” effects occurred, and that repairs and effective modifications or re-designs were made. In some systems bunching occurs much later; this suggests that system elements are *aging*, i.e. susceptible to wearout (together, unless the labels are all of one subsystem or component). Any bunching pattern signals that a stressful environment may have occurred. An attempt to understand the source of stress is called for.

An alternative graphical presentation that is vivid and sometimes preferred is the *Cumulative Sum Plot* (CuSum):



The general stretchout of the example Time-Line Plot and the hollow rise of the CuSum Plot both show that times between successive failures are increasing: “reliability growth” is said to occur (not to be expected to continue forever).

A curve like this:



shows that “reliability growth” is occurring, i.e. successive times to failure are getting longer, at first, next is (temporarily) linearly increasing, and finally flattens out. If the plot becomes roughly linear in time the *indication* is that the system’s failure propensity is, for the present, “stationary in time” (in the environment considered). This does not rule out random fluctuations in the times between failure, or in the numbers of failures in mission times of the same length. If and when the CuSum flattens, aging has set in; the item or system may need to be replaced.

- **Strong recommendation:** Do *not* begin (and, especially, end) data analysis immediately with the computation of an overall average, or estimated probability of completing a mission time, e.g. by computing the total number of times to failure that exceed a mission time, divided by the total observed times to failure. Make graphs, look for systematic changes, breaks or jumps. The first objective is to *look for causes of failures, or evidence of improper design, or sensitivity to environment stresses*. It is first important to understand, and if possible correct, a system’s weak points, *then* to measure accurately what has been done.

QUANTIFICATION OF RELIABILITY

Measures of Reliability

These are useful summaries, but omit some details. Awareness of *measure-vulnerability* is essential.

- **MEAN (average) system time to mission failure (MTTF) = m .**

- ~ Appropriate only if system stable (reliability growth has ended)

- ~ Estimate from data: Average n observed times to failure, $(t_1 + t_2 + \dots + t_n)/n = \hat{m}$ (\hat{m} = estimated m). Model needed otherwise (if data is censored)

- (~ **CAUTION:** The formula that states:

$$\begin{array}{l} \text{Mean Time Between} \\ \text{Operational} \\ \text{Mission Failures} \end{array} = \frac{\text{Total operating time (e.g. driving time, fly time, system - on time)}}{\text{Total number of operational mission failures}}$$

is *not generally accurate*, so avoid. It works in theory if the data are exponentially distributed or if the operating time is long, meaning that at least 5 failures are observed during operating time.)

- ~ Error, $\hat{m} - m$, reduced if observation number increased.

- ~ Time to failure is the *active* or “*up*” time from moment system turned on to moment of mission-affecting failure. Do *not* include off times.

NOTE: length of off-times can influence/affect (possibly shorten) active time to failure. Example: an engine, e.g. auto, that is inactive for a long time period may be difficult to start and keep running.

NOTE: observed time to failure is *zero* if system won't “turn on”.

(Do *not* ignore or throw out such “times”. They may tend to occur after the system has been turned off for some time, or goes down for repair, and the unit is returned to inventory.)

- ~ A system MTTF depends on *conditions*: environment (heat, cold, movement shock, time and treatment since last usage). It is important to *test* under all possible mission conditions.

- ***PROBABILITY of mission completion without mission-critical failure***

~ As appropriate:

Probability (Time to mission failure exceeds mission duration),

~ Depends on environment, handling by personnel, previous maintenance

or

Probability (Miles traveled/flown/sailed exceeds mission distance),

~ Depends on environment, speed, maintenance

or

Probability (Rounds/missiles fired adequate for mission success),

~ May depend on time since last burst fired (gun must cool)

or

Probability (Sensor or communication package does not fail during mission)

~ Appropriate only if system stable (reliability growth phase effectively ended)

~ Estimate from data: \hat{p}_{mc}

$$\hat{p}_{mc} = \frac{\text{Number of missions survived without system failure}}{\text{Number of missions (of some duration under some conditions)}}$$

= Estimated probability that time to system failure exceeds mission time

~ Example: sensor

QUANTIFICATION OF RELIABILITY

Models for Reliability

- *Reliability models* mathematically represent or summarize the probability of reliability-related events, such as time or distance, etc., to failure, or failure on demand.
 - ~ They are based on subject-matter facts and assumptions (physics, engineering, psychology, operational environment, etc.);
 - ~ they conveniently summarize or compress observational data;
 - ~ they describe the behavior of future data, given needed *parameter* values; *some* uncertainty can be represented, by *standard errors* or *confidence intervals*, but these do not reflect errors of the model, or changes in conditions. Caution!
 - ~ Understanding of reliability models requires mathematics: algebra, calculus, probability theory. The inclusiveness and appropriateness of phenomena represented (or omitted) is more important than the precise model. Consult an experienced expert.

The Exponential Distribution Model

If an item's chance of failure in any operating period is *not age dependent*, or does not change with number of operating hours since last failure, its random time to failure, T , has the *exponential distribution*:

$$\underbrace{P\{T \leq t\}}_{\text{Probability}\{T \text{ less than } t\}} \equiv \underbrace{F_T(t) = 1 - e^{-\lambda t}}_{\text{Distribution function of } T}$$

- ~ graphical illustrations
- ~ The parameter λ is called a *failure or hazard rate*.
- ~ The MTTF, mean time to failure, $E[T] = 1/\lambda$.
- ~ The probability of mission success using the exponential model is

$$\begin{aligned} \text{Prob}\{\text{Mission Time exceeds } t (= \text{mission duration})\} &= P\{T \geq t | \lambda\} = e^{-\lambda t} \\ &= \exp\left(-\frac{t}{\text{MTTF}}\right) \end{aligned}$$

- ~ The parameter λ must usually be estimated from data.

Result: If the Exponential model holds (is assumed), an estimate of λ , $\hat{\lambda}$ is:

$$\begin{aligned} \hat{\lambda} &= \frac{\text{Number of failures during an operating time, } t_0}{\text{Operating time, } t_0} \\ &\equiv \frac{N(t_0)}{t_0} \end{aligned}$$

- If individual times to failure are t_1, t_2, \dots, t_n , then

$$\hat{\lambda} = \frac{1}{\text{AVE (of times)}} = \frac{n}{t_1 + t_2 + \dots + t_n} = \frac{1}{\bar{t}}$$

- ~ If the item is a truck or tank, operating time may be replaced by operating distance, d_0 .

$$\hat{\lambda} = \frac{\text{Number of failures during operating / travel distance, } d_0}{\text{Operating distance, } d_0}$$

- **Note:** λ , thus, $\hat{\lambda}$, often depends on environmental factors (heat, cold, terrain, handling ...). Observations from different environments should *not* be *uncritically* combined (“pooled”).

- Observations from different environments *can* be combined statistically by use of *nonlinear statistical regression techniques*.

CAUTION: This is a specialized topic that requires expert attention.

The results can be used for prediction, e.g. under specified field conditions, *but* there are several possible error sources so the predictions may be untrustworthy.

~ Series Systems:

- A system consists of two or more subsystems (missile system launch, propulsion, guidance, detonation subsystems)
- Any subsystem failure means system failure
- If subsystems fail exponentially and independently, with rates $\lambda_1, \lambda_2, \dots, \lambda_k$ (k is number of subsystems)

$$\text{Probability}(\text{System does not fail mission}) = e^{-\lambda_1 t_1} e^{-\lambda_2 t_2} \dots e^{-\lambda_k t_k} = e^{-(\lambda_1 t_1 + \lambda_2 t_2 + \dots + \lambda_k t_k)}$$

t_1 = exposure duration, subsystem 1, ... t_j = same, subsystem j ($j = 1, 2, \dots, k$)

- **CAUTION:** environmental effects (shocks to entire system) can induce dependence. Often *reduces* probability of failure avoidance.

The Weibull Distribution Model

If an item's chance of failure in any operating period either (i) steadily *increases with age* (operating time since last failure), or (ii) decreases with age, then its random time to failure, T , is often modeled by the Weibull distribution family:

$$P\{T \leq t\} \equiv F_T(t) = \underbrace{1 - e^{-(\lambda t)^\beta}}_{\substack{\text{Weibull} \\ \text{distribution form}}}$$

- ~ β is the shape parameter:
 - if $\beta = 1$, T is the exponential distribution model;
 - if $\beta > 1$ the model represents situations in which age *increases* the probability of failure (in a time period); (this can be the result of wearout, or accumulated damage, e.g. from a succession of environmental shocks);
 - if $\beta < 1$ the model represents situations in which the probability of failure *decreases* with age. (This effect can be the result of *mixing* data: some from items with short lives, some with long (caused by manufacturing or environmental variations)).
- ~ The *mean* of the general (β not 1) Weibull model is *not* $1/\lambda$ as for the simple exponential (it depends upon β , and is $E[T] = \Gamma(1/\beta)/\lambda\beta$; if $\beta = 2$, $E[T] = \sqrt{\pi}/2\lambda = 0.89$).
- (**Note:** $\Gamma(x)$ is the gamma function, equal to the *factorial* $(x - 1)!$ for integer x . Go to an expert or get a book!)
- ~ $1/\lambda$ is *always* the 63.2nd *percentile* of the Weibull (provides rough estimate of λ from data)
- ~ **Summary.** The Weibull is
 - convenient, frequently used, but
 - *not* the only, or always, best model to use to represent age-dependence effects. For other options, based on particular scientific or operational conditions, see references.

Counting Models

An alternative to observing and recording, then *modeling* the times/distances between failures is to observe and model *counts* of failure-type events during mission periods.

- A basic *simple* model is the **Poisson**: probability of number of “rare events” in fixed time period, when chance of an event (failure) is *equal but small* in any short time interval, and is uninfluenced by past history.
 - ~ Parametric: probability of event (e.g. failure) in $(t, t + h)$ is λh for small h ; λ is the *event rate*.
 - ~ A possible *first* model for describing numbers of failures in given operating times, excluding downtimes of complex systems that are treated as mature.
 - Simple statistical test for Poisson: estimate *mean* number of events per equal intervals, and variance of same (it will be necessary to have counts over several such intervals); if these are close to *equal* the Poisson *may* be a useful interim assumption. It is often true that the variance considerably *exceeds* the mean. Such overvariability has many possible causes, *one* being an irregular background environment that raises the rate temporarily.
 - ~ In general, if $N(t)$ is the number of events in time t ,

$$P\{N(t) = k\} = e^{-\lambda t} \frac{(\lambda t)^k}{k!}.$$

- ~ **Special Case:** Under Poisson model, the probability of *no/zero failures in time t* is $e^{-\lambda t}$.

The parameter λ is the same as that for the Exponential time-to-failure model: $1/\lambda$ is the mean time between failures in this model as well.

- Alternative: **Bernoulli-trials** (“dishonest coin flip”) *model*: if time axis is divided into hours or days (discrete time), *or* into similar missions, e.g. aircraft sorties, *and* either *zero* or *one* failure can occur in a time period, independently of past (as if generated by the flip of a weighted coin with success probability f). The number of failures in n time periods, $N(n)$, has probability

$$P\{N(n) = k\} = \frac{n!}{k!(n-k)!} f^k (1-f)^{n-k}$$

(where $k! = (k)(k-1)(k-2)\dots(1)$). The probability of no/zero failures in n time periods is $(1-f)^n = e^{-\ln(1-f)n}$. If we replace $\ln(1-f)n$ by λt the present model resembles the Exponential. It should! Reasons elsewhere.

- Environment/condition-adjusted (“accelerated”) counting models:

The failure rate/probability parameters, λ and f above, can systematically vary with environmental conditions. For instance, if A is absolute temperature, the Poisson rate at A is $\lambda(A) = \lambda_0 e^{-\gamma/A}$ so as absolute temperature ($A = 273.2 + C^\circ$ Centigrade) becomes large the failure rate increases. The particular model is the Arrhenins model sometimes used in accelerated life testing.

Counting Models and Reliability Growth

- Reliability Growth models, based on Non-homogeneous Poisson.

In many situations, system design or execution faults are gradually removed during development and early operational testing. In this case the number of failure events declines.

~ A model:

$$P\{N(t) = n\} = e^{-H(t)} [H(t)]^n / n!,$$

$$\text{with } H(t) = \int_0^t \lambda(t') dt'.$$

$H(t)$ is called the *hazard*, $\lambda(t)$ the *hazard/failure rate*.

~ Example: Weibull-like hazard rate

$$\lambda(t; \theta, \gamma) = (\theta t)^\beta, \quad \theta \text{ and } \beta \text{ positive.}$$

Note: time to *first* event has Weibull distribution, but not afterwards.

$$H(t) = \int_0^t \lambda(t'; \theta, \gamma) dt' = \frac{\theta^\beta t^{\beta+1}}{\beta+1}.$$

Note: if reliability growth is occurring, β is *less than unity/one*.