

### Censored Data

#### Table of Contents

- Introduction
- Types of Censoring
- Time-Censored (Type I) Experiments
- Failure-Censored (Type II) Experiments
- Summary and Conclusions
- Bibliography
- About the Author
- Other START Sheets Available

#### Introduction

Assume we are dealing with an observation, in time, of a sample of “n” entities placed on test (be these, devices or humans). The experimental observation period is defined as the time elapsed since the experiment begins (time zero) until it is terminated (time  $T_0$ ). However, it often occurs that we need to discontinue our experiment before all the elements in the sample experience the “event of interest” (e.g., failure or death). In such cases, we say that the experiment has been “suspended,” “censored,” or “truncated”.

“Truncation” may not be the most efficient way to conduct an experiment, from the theoretical standpoint. But, due to time, economic or practical considerations, it happens so frequently that statistics had to find ways to deal with it in a successful manner. In this START sheet we overview some of these statistical procedures, we illustrate them via several practical, numerical examples and we provide some references for further reading.

#### Types of Censoring

To better analyze this complex issue, we begin with a characterization of the censoring mechanisms. Such characterization can be based on several elements, among them, the status of the entity observed, both at the time we start and at the time we finish our observation. Censoring mechanisms can also be characterized based on whether or not the experiment is terminated at the time of the “event of interest” (e.g., failure or death).

With respect to the status of the entity observed, censoring can occur at either extreme (or at both ends) of the entity life. That is, we may not know exactly at what time the life of the

entity started or finished. This happens because the entity in question may have already started operating at the time we begin our observation. Or the life may have not yet finished (e.g., failed) by the time we complete our observation period.

Figure 1 illustrates censoring situations. Line “a” shows an entity that has already been “operating” for some unknown period of time, before we start monitoring it. This case is called “left-censoring.” The “X” symbols in Figure 1 represent the points in time when we actually start or finish monitoring the censored entities, other than the beginning (of entity life, at time zero) or the end of the experimental observation period (time  $T_0$ ).

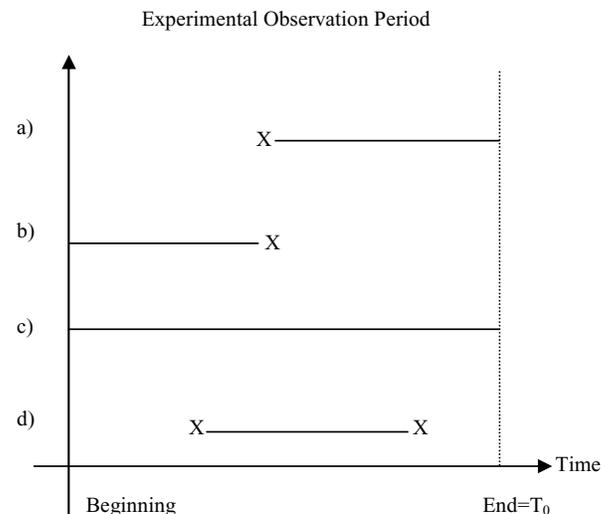


Figure 1. Type I (Time-Truncated) Censoring Cases

Similarly, Line “b” shows an entity that has been monitored since the beginning of its life (i.e., at the start of the experiment) but which we have ceased to observe before the experiment ends (time  $T_0$ ) or it fails. That is, we observe the entity for some time, after which we are not able to monitor it any more. This other type of truncation is known as “right-censoring.”

We can stop monitoring all the entities, putting an end to the experiment, at some pre-specified time  $T_0$ , which is independent of the event of interest (e.g., death). The entity in Line “c” has been monitored all along the experiment. Finally, a more complex example is presented in Line “d”.

START 2004-3, CENSOR

Here, both the beginning and end of the entity “life” are now unknown (interval censored). We can only monitor such entity for some intermediate part of its “life” span. Censoring schemes, where the end of the observation period is not determined by an event of interest (e.g., failure), are referred to as time censoring, time truncation, or suspension in time. Such censoring schemes are not event-driven and are known as Type I. In these schemes, the experiment stopping time ( $T_0$ ) is pre-established and the number of failures observed ( $i$ ) during the period of experimentation is random.

On the other hand, we may elect to observe a sample of “ $n$ ” entities until the time of occurrence of some pre-specified event of interest, such as the time of the  $i^{\text{th}}$  failure or death ( $i \leq n$ ) denoted by the  $X_i$  in Figure 2. That is:

$$0 < X_1 < X_2 < X_3 < \dots < X_i < \infty$$

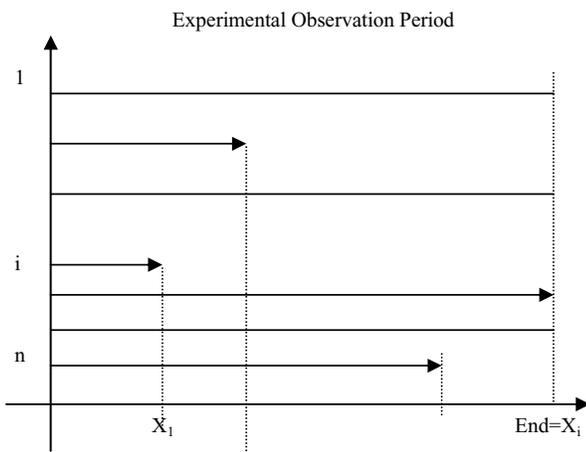


Figure 2. Type II (Event-Driven) Censoring Case

At the time of the  $i^{\text{th}}$  failure (failure times  $X_i$  are denoted in the graph by an arrowhead) we discontinue our observation of the  $(n-i)$  sample elements remaining in operation. This other censoring scheme is often referred to as “failure” or “event” truncation and is known as Type II censoring. In these cases, the experiment stopping time ( $X_i$ ) is random and the number of failures ( $i$ ) occurred during experimentation is pre-established.

In either censoring scheme (Type I or II) the number “ $i$ ” of “events” of interest (e.g., death) observed during the experiment is less than the total “ $n$ ” entities on trial. Some times the distribution of the “lives” of the entities is known. Other times, the probability “ $p$ ” of occurrence of an event during the observation period (time  $T_0$ ), can be calculated. In such cases, we may be able to model the underlying life ( $X$ ) distribution and estimate the parameters of interest such as Mean Time to Failure (MTTF or  $\mu$ ), failure rate (FR or  $\theta$ ), tenth percentile of device life (L-10) and calculate confidence intervals (CI) for them.

Other times, the problem of modeling “life” is further complicated and, thus, approached differently than we do here. Some examples of such complications include when failures are (or are not) replaced at the time they occur, or when the distribution of the “lives” is not Exponential. In such cases, the hazard function (instantaneous probability of failure) is time-dependent and there are several additional parameters than we now need to estimate from the data. In addition, having more complex censoring mechanisms, in conjunction with a time-dependent hazard rate, creates many more theoretical difficulties.

In the rest of this START sheet, we discuss some of the issues involved in estimating reliability parameters from Exponentially distributed censored data and present several numerical examples. We first present the case for time-censored experiments. Then, we discuss failure censored ones, of which experiments developed until the first failure occurs, constitute a special case. We end by giving a short bibliography for further study of time and failure censored experiments.

## Time-Censored (Type I) Experiments

Time censored experiments (or data collection efforts) take place if a test is terminated at a pre-specified time (say  $T_0$ ) as opposed to at the time of a failure. In them, we know the total operating time “ $T$ ” of all “ $n$ ” devices placed in operation, as well as the total number of failures “ $k$ ”. However, we may not know all individual device failure times.

Time-censored experiments occur frequently in practice. For example when say, “ $n$ ” aircraft, carrying a given device on board, simultaneously operate for a total of  $T_0$  hours (Figure 3) and “ $k$ ” failed device are detected and replaced. However, we don’t know the exact times when these devices failed.

In such cases it is convenient that the distribution of the lives are Exponential. Then, the FR  $\theta$  is not dependent on operating time (the device life). In such cases we can afford to ignore the exact moment, during the life of the device that a failure has occurred. Since the FR is constant, the instantaneous probability of a failure is always the same. This allows us to ignore the exact device failure times and still estimate the parameters of interest such as MTTF, FR, L-10, etc., as well as to obtain CI for them.

But time censored estimation is approached in different ways, depending on the nature of the data and on the experimental conditions. We now examine some of these cases.

### Case 1: devices fail and are “instantaneously” replaced

Assume that the distribution of “ $n$ ” entity lives ( $X$ ) is Exponential, with FR  $\theta$  and MTTF (or mean life)  $\mu = 1/\theta$ . Assume that all the “ $n$ ” devices are working under very similar environmental and user profile conditions. Also assume that all “failures” (occurring at unknown times) are “immediately” replaced by identical entities. Finally, assume that we know the

length (total number of hours “ $T_0$ ” of operation) of such experiment or test, and the total number of failures “ $k$ ” observed during this time (Figure 3).

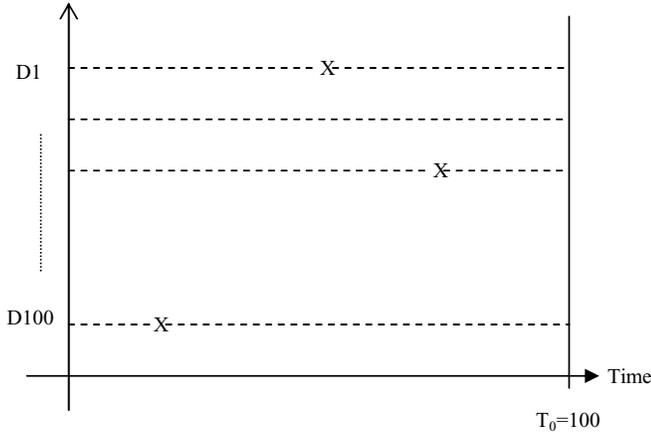


Figure 3. Representation of Type I Censoring:  $n = 100$  Devices Simultaneously on Test

All “ $n$ ” devices on test are independent, identically distributed and operate continuously, being replaced as soon as they fail. We can then consider two statistically equivalent situations. First, consider “ $n$ ” superimposed, identical processes, running for a time  $T_0$ . Then, concatenate one after the other, all the “ $n$ ” independent, identical processes, now running for a test time  $T = n \times T_0$ . In either case, the probability of observing “ $k$ ” failures, in their respective experimental times ( $T_0$  or  $T$ ) is the same. Such probability is obtained via the Poisson distribution (but using FR  $\lambda = n\theta$  or  $\lambda = \theta$ , accordingly).

The statistical formulation of the Poisson Process probability, for a single device having a FR  $\theta$ , and yielding “ $k$ ” failures, during an operating time  $T_0$ , is as follows:

$$P_{\theta} \{N(T_0) = k\} = \frac{e^{-\theta T_0} (\theta T_0)^k}{k!}; k = 0, 1, 2, 3, \dots$$

From the preceding, we obtain that “ $n$ ” independent and identically distributed devices, each one following the Poisson distribution with rate  $\theta$ , operating simultaneously, will observe “ $k$ ” failures (during time  $T_0$ ) with rate  $\lambda = n\theta$ . The Poisson model holds because all the FRs remain the same throughout the entire experiment of length  $T_0$ .

For example, an aircraft operates for a Time  $T_0 = 100$  hours, with a radio FR per hour of  $\theta = 0.0005$  (hence,  $MTTF = \mu = 1/\theta = 2000$  hours). Now, assume that  $n = 100$  aircraft operate simultaneously, during these  $T_0 = 100$  hours. Then, the overall radio FR (for the  $n = 100$  concurrently operating devices) is  $\lambda = n\theta = 0.05$  per hour. Using the Poisson formula (with rate  $\lambda = n\theta$  per unit time) we obtain the probability of observing say more than four

radio failures (denoted  $P\{N(T_0) > 4\}$ ) during the experimental time  $T_0$  (Figure 3):

$$\begin{aligned} P_{\lambda} \{N(T_0) > 4\} &= 1 - P_{\lambda} \{N(T_0) \leq 4\} = 1 - \sum_{k=0}^4 P_{\lambda} \{N(T_0) = k\} \\ &= 1 - \sum_{k=0}^4 \frac{e^{-\lambda T_0} (\lambda T_0)^k}{k!} \end{aligned}$$

For ease in looking up the probability in the Poisson table (instead of calculating it via the Poisson formula) we multiply the hourly rate  $\lambda = n\theta = 0.05$  by 100 hours, obtaining the new rate  $\lambda' = 5$  failures per  $T_0 = 100$  hours (the new unit time) yielding the same results:

$$P_{\lambda'=5} \{N(T_0) = N > 4\} = 1 - \sum_{k=0}^4 \frac{e^{-5} 5^k}{k!} = 1 - 0.4404 = 0.5596$$

For example, assume we detect and replace, say  $k = 4$  failed devices (e.g., radios) during  $T_0 = 100$  hours of simultaneous operation of  $n = 100$  aircraft that carry these. A sample point estimate ( $\theta^*$ ) of an individual radio FR, per unit time, is obtained as:

$$\theta^* = (\text{Total Failures} / \text{Total Time}) = k/n \times T_0 = 4 / (100 \times 100) = 0.0004$$

We can use these results to obtain an approximate 90% CI for the true device (radio) FR (or for its MTTF) using the approach just presented. We search which FR ( $\lambda = n\theta T_0$ ) for similar Poisson Processes yielding up to  $k = 4$  failures, produce coverages close to  $1 - \alpha/2 = 0.95$  and  $\alpha/2 = 0.05$ . Such two FR induce approximate upper and lower limits for a 90% CI:

- First, try  $n\theta T_0 = 2$ : this implies that  $\theta = 2/nT_0 = 2/(100 \times 100) = 0.0002$ . Then, searching the Poisson tables for the trial FR parameter ( $\lambda = n\theta T_0 = 2$ ) we obtain the probability:

$$P_{\lambda=2} \{N(T_0) \leq 4\} = 0.9473$$

- Now, try value  $n\theta T_0 = 9$ : this implies that  $\theta = 9/nT_0 = 9/(100 \times 100) = 0.0009$ . Then, searching the Poisson tables for the FR trial parameter ( $\lambda = n\theta T_0 = 9$ ) we obtain that:

$$P_{\lambda=9} \{N(T_0) \leq 4\} = 0.0550$$

Since the error probabilities are  $1 - \alpha/2 = 0.95$  and  $\alpha/2 = 0.05$ , an approximate 90% CI for the unknown FR  $\theta$  is given by: (0.0002, 0.0009). Likewise, an approximate 90% CI for the MTTF  $= \mu =$

$1/\theta$  is given by the corresponding reciprocal values: (1111.11, 5000).

Notice how, in both cases, the approximate CI covers the true FR and MTTF.

Case 2: devices failed but are not replaced

Now, assume that we have “n” devices with lives (X) that are also Exponential, with MTTF  $\mu$  and FR  $\theta$  ( $=1/\mu$ ) placed on test for a pre-specified time  $T_0$ . However, this time we don’t replace the failed devices. Hence, at the end of our experiment (time  $T_0$ ) we find that “k” of them failed at some unspecified time and only (n-k) are still operating. The probability “p” of any one device “failing” before the experiment ends (at the operating time  $T_0$ ) can be obtained by using the definition of the Exponential distribution function, for time  $T_0$ :

$$p = \text{Prob.Device.Fails} = P(X \leq T_0) = 1 - e^{-\frac{T_0}{\mu}} = 1 - e^{-\theta T_0}$$

Since all devices are independent and identical, the total number of failures “k”, out of the possible “n” occurring in the experiment, is distributed Binomial with parameters n and p:

$$P(\text{Failures} = k; \text{Total} = n; \text{Fail Prob} = p) = C_k^n p^k (1-p)^{n-k}$$

$$= B(k; n; p)$$

Assume, as in the previous example, n = 100 aircraft, each operating for a time  $T_0 = 100$ . Let each aircraft carry a radio with the same MTTF = 2000 hours. Assume that we detect, at the end of the operation, say k = 4 failed radios (but have not replaced them). Then, we can obtain the exact probability “p,” that any radio fails this specific “test” of length  $T_0$ :

$$p = 1 - e^{-T_0/\mu} = 1 - e^{-100/2000} = 1 - 0.9512 = 0.0488$$

And the probability of finding more than, say k = 4 failures, in this experiment is:

$$P\{\text{Failures} > 4\} = 1 - P\{\text{Failures} \leq 4\}$$

$$= 1 - \sum_{k=0}^4 C_k^{100} (0.0488)^k (1 - 0.0488)^{100-k}$$

$$= 1 - 0.457836 = 0.5422$$

A point estimate for the radio probability of failure “p,” for mission time  $T_0 = 100$  hours:

$$p = (\text{Total Failures/Total Devices}) = 4/100 = 0.04.$$

We can also obtain an approximate 90% CI for the true radio FR (or its MTTF) by using the previous approach. We search, for n = 100, which values of the proportion “p” of the Binomial probability, yield up to k = 4 failures, with coverage close to 0.95 and 0.05. The resulting two proportions yield approximate upper and lower limits for a 90% CI for “p”:

- Try value p = 0.02: Then, the Binomial result for k ≤ 4; n = 100 and p = 0.02 yields:

$$P(\text{Failures} \leq 4; \text{Total} = 100; \text{Fail Prob} = 0.02) = 0.9492$$

For a mission time of  $T_0 = 100$  hours, such “p” implies that the device FR  $\theta$  is:

$$p = 1 - e^{-T_0\theta} = 1 - e^{-100\theta} = 0.02 \Rightarrow \theta = -\frac{1}{T_0} \ln(1-p)$$

$$= -\frac{1}{100} \ln(1-0.02) = 0.0002$$

Hence, FR  $\theta = 0.0002$  and the corresponding MTTF  $\mu = 1/\theta = 4949.8$  hours.

- Try now p = 0.09: Then, the Binomial result for k ≤ 4; n = 100 and p = 0.09 yields:

$$P(\text{Failures} \leq 4; \text{Total} = 100; \text{Fail Prob} = 0.09) = 0.0474$$

For a mission time of  $T_0 = 100$  hours, such “p” implies that the radio FR  $\theta$  is:

$$p = P(X \leq T_0) = 1 - e^{-T_0\theta} = 1 - e^{-100\theta} = 0.09 \Rightarrow \theta$$

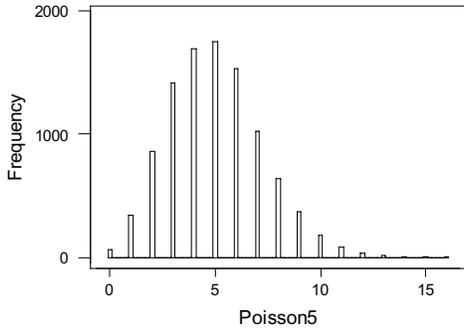
$$= -\frac{1}{100} \ln(1-0.09) = 0.00094$$

Therefore, an approximate 90% CI for the FR  $\theta$  is (0.0002, 0.00094). The MTTF  $\mu$  for a FR  $\theta = 0.00094$  is its reciprocal:  $\mu = 1/\theta = 1060.3$  hours. The corresponding 90% CI for the MTTF is (1060, 4949) hours. These results are comparable to the one for the Poisson.

Following, we show the exact cumulative probabilities for both, the Poisson and Binomial distributions, corresponding to the two examples discussed above, and the histogram of a simulation of

10000 Poisson-5 values. Both distributions are close because the number of devices on test ( $n = 100$ ) is large and the individual device probability of failure ( $p = 0.048$ ) is small. In such cases, the Binomial results can be approximated by the Poisson results:

Failures	Poisson	Binomial
0	0.006738	0.006717
1	0.040428	0.041178
2	0.124652	0.128694
3	0.265026	0.275363
4	0.440493	0.457836
5	0.615961	0.637577
6	0.762183	0.783581
7	0.866628	0.884169
8	0.931906	0.944160
9	0.968172	0.975622
10	0.986305	0.990310



The previous two approaches to deriving CI for the Exponential mean (MTTF) when only the Total Test Time (T) and total number of failures (k) are known, are good, illustrative examples, but are seldom used in real life. Instead, we use more practical procedures.

Moreover, device operation time is often non-overlapping. Devices may have well been working in different periods of time. However, other circumstances being similar, we can reasonably relax, for practical purposes, the preceding assumptions and work as if the time of operation had occurred simultaneously. We discuss such implementations next.

Assume that the situation of interest could be construed as an experiment of the type illustrated in Figure 3. Assume also that there is an undisclosed number of independent devices on test. That is, we only know the operation's total test time "T" and number of failures "k." Then, we may assume that the total operation time T reported is the product  $n \times T_0$  given in Figure 3 (i.e., the number of devices on test, times the experiment length).

Then, if the underlying distribution of the lives is Exponential and the experiment is time terminated (Type I), the distribution of "twice Total Test Time (T) divided by the Mean ( $\mu$ )":

$$2 \times T/\mu$$

is approximately distributed as a Chi Square ( $X^2$ ), with  $\gamma = 2k + 2$  degrees of freedom (DF). We can then use the Chi Square distribution percentiles, with  $DF = 2k + 2$ , to derive a pre-specified CI for the unknown MTTF (or Exponential mean  $\mu$ ) with confidence  $1 - \alpha$ :

$$\left( \frac{2T}{X_{2k+2,1-\alpha/2}^2}; \frac{2T}{X_{2k+2;\alpha/2}^2} \right)$$

where  $X_{2k+2,1-\alpha/2}^2$ ;  $X_{2k+2;\alpha/2}^2$  are the corresponding percentiles of the Chi Square, with  $DF = 2k + 2$ , and  $\alpha$  is the pre-specified CI sampling error that we are willing to absorb.

For example, let some devices operate for  $T = 1700$  hours, with  $k = 3$  failures recorded. Assume that the total number of devices operating is either undisclosed or unknown and assume that a  $100(1 - \alpha)\% = 95\%$  CI for MTTF is sought. From these data we have:

1. Total Time on Test  $T = 1700$ ,
2.  $DF = 2k + 2 = 2 \times 3 + 2 = 8$  and
3. Sampling error  $\alpha = 0.05$  (for,  $1 - \alpha = 0.95$ ).

Hence, the two Chi Square table percentiles, for a 95% CI for the MTTF, are:

$$X_{2k+2,1-\alpha/2}^2 = X_{8,0.975}^2 = 17.54; X_{8,0.025}^2 = 2.18$$

Based on all the preceding data, a 95% CI for the MTTF (or Exponential mean  $\mu$ ) is:

$$(2 \times 1700 / 17.54; 2 \times 1700 / 2.18) = (193.84; 1559.6)$$

Finally, and for comparison with the two procedures developed in the previous sections, we recalculate the corresponding 90% CI for their data:  $T = n \times T_0 = 10000$ ;  $k = 4$  failures and  $X_{2 \times 4 + 2; 0.05}^2 = 3.94$ ;  $X_{10, 0.95}^2 = 18.31$ . Hence, the corresponding CIs are:

For MTTF:  $(2 \times 10000 / 18.31; 2 \times 10000 / 3.94) = (1092.3; 5076.1)$

For Rate: (reciprocals of the above):  $(0.000197; 0.00092)$

In the following table, we summarize the 90% CI values obtained by the three methods. The real parameters used were: MTTF = 2000 hours and Failure Rate = 0.0005:

Method Used	MTTF LwBd	MTTF UpBd	F.Rate LwBd	F.Rate UpBd
Poisson	1111.1	5000	0.000200	0.00090
Binomial	1060	4949	0.000200	0.00094
Practical	1092.3	5076.1	0.000197	0.00092

## Failure-Censored (Type II) Experiments

Failure censoring or truncation occurs when we terminate an experiment of “n” devices at say, the time  $X_k$  of the  $k^{\text{th}}$  failure. At such time,  $(k - 1)$  devices in the experiment have already failed (and we know exactly when) and  $(n - k)$  are still operating (see Figure 2). If device life is distributed Exponential with mean  $\text{MTTF} = \mu$ , we can obtain the sampling distribution of the Total Test Time (T) for the life of the devices in the experiment. From this information, we can obtain the CI for MTTF and all other parameters of interest.

To this effect we analyze first the general case, where failure  $X_k$ ,  $k < n$ , yields the time of truncation. Let  $X_i$  denote the time to failure (i.e., life) of any  $i^{\text{th}}$  device ( $1 \leq i \leq k$ ) in the sample of size “n”. When the experiment is terminated at the time of the  $k^{\text{th}}$  failure, the Total Time on Test “T” of all the “n” devices in the sample is given by:

$$T = \sum_{i=1}^k X_i + (n - k)X_k$$

Since  $k < n$ , time T is the sum of two components: (1) all device failure times (up to  $k^{\text{th}}$  failure) and (2) the product of the truncation time  $X_k$  times the remaining  $(n - k)$  operating devices. The sampling distribution of statistic  $2 \times T/\mu$  is the Chi Square. But now  $DF = 2k$ , twice the number of failures observed during the life test.

Using this distribution we can test (or obtain the CI) for the performance measures of interest (MTTF, FR, L-10, etc.). In particular, we can obtain the  $100(1 - \alpha)\%$  CI for the Exponential mean, MTTF (or  $\mu$ ) by using the formula:

$$\left( \frac{2T}{X_{2k,1-\alpha/2}^2}; \frac{2T}{X_{2k,\alpha/2}^2} \right)$$

where  $X_{2k,1-\alpha/2}^2$ ;  $X_{2k,\alpha/2}^2$  are the corresponding percentiles of the Chi Square distribution, with  $DF = 2k$ , and  $\alpha$  is the sampling error we are willing to accept. The corresponding CI for FR is obtained, as before, via the reciprocals of the CI limits for MTTF.

We illustrate this method via a numerical example. Assume that we place  $n = 45$  devices in a life test and stop testing at the time of the one-but-last failure (denoted  $T_{44} = 313.88$ ). The test is failure truncated at the  $k^{\text{th}} = n - 1 = 44$  failure. Assume that the last failure time ( $T_{45}$ ), had we let this experiment run to its completion, would have occurred at time  $T_{45} = 399.07$ . Assume that the sum of the lives of the  $n = 44$  failed items were 4097.68. In

such truncated life test, the MTTF point estimator is obtained via the statistic:

$$\hat{\mu} = \frac{T}{k} = \frac{\sum_{i=1}^k T_i + (n - k)T_k}{k} = \frac{\sum_{i=1}^{44} T_i + (45 - 44)T_{44}}{44}$$

$$= \frac{4097.68 + 313.88}{44} = 100.26$$

For comparison, had we been able to include the  $45^{\text{th}}$  failure (i.e., time  $T_{45} = 399.07$ ) we would have obtained a point estimator  $\mu = \Sigma T_i/n = 4496.75/45 = 99.92$ , not very different. The additional time ( $85.27 = 399 - 313.8$ ) corresponds to the additional unobserved failure and is compensated by the additional degrees of freedom ( $DF = 2(45 - 44)$ ) that are added.

To develop a CI for  $\mu$ , we now use  $DF = 2k = 2 \times 44 = 88$  (twice the number of the observed failures) for obtaining the two Chi Square table values. Assume that the sampling error is  $\alpha = 0.05$  (for a 95% CI). Then, the percentiles from the Chi Square table are:

$$X_{2 \times 44, 1-\alpha/2}^2 = X_{88, 0.975}^2 = 115.8414; X_{88, 0.025}^2 = 63.9409$$

and the corresponding 95% CI for the mean life  $\mu$  is:

$$(2 \times 4411.56/115.84; 2 \times 4411.56/63.94) = (76.17; 137.99)$$

It is important to emphasize that, if the lives ( $X_k$ ) of the devices follow another statistical distribution than Exponential (say, Weibull) then, obtaining these performance measures becomes much more difficult. The analysis of such cases, due to their larger complexity, will be the topic of separate START sheet.

### The Case of Truncation at the First Failure

We now analyze the case of experiments terminated at the time of the first “failure”. This technique is very useful when, say, the device under test is very expensive, or when there are very few devices and the testing is destructive. Hence, we cannot afford to have many devices fail, because the cost of the experiment can then become prohibitive.

In such cases, the cumulative distribution (F) of the time to first failure (denoted  $X_{(1)}$ ), also called the “Unreliability” of  $X_{(1)}$ , can be obtained by using the fact that all  $n - 1$  independent and identically distributed (Exponential) lives ( $X_2, \dots, X_n$ ) have necessarily outlived this  $X_{(1)}$ . We then calculate the probability

(Reliability) that first failure ( $X_{(1)}$ ) is greater than an arbitrary time (say,  $x$ ) in a sample of size  $n$  (and we denote it,  $(x)$ ):

$$\begin{aligned} \bar{F}_{X_{(1)}}(x) &= 1 - F_{X_{(1)}}(x) = P(X_{(1)} \geq x) = P(X_1 \geq x, \dots, X_n \geq x) \\ &= P(X_1 \geq x) \dots P(X_n \geq x) = \{P(X \geq x)\}^n = e^{-n\theta x} \end{aligned}$$

From here, the distribution of the time to first failure ( $X_{(1)}$ ) using  $\mu = 1/\theta$ , is:

$$F_{X_{(1)}}(x) = 1 - \bar{F}_{X_{(1)}}(x) = 1 - P(X_{(1)} \geq x) = 1 - e^{-n\theta x} = 1 - e^{-\frac{n}{\mu}x}$$

Having the distribution of  $X_{(1)}$  allows us to obtain all the parameters of interest. For all parameters of the distribution of any life  $X$  (our main interest) can be obtained from the parameters of the distribution of  $X_{(1)}$  (the time to first failure).

For example, the MTTF of the first failure is  $\mu/n$  (i.e., the original MTTF “ $\mu$ ” divided by the sample size “ $n$ ”). Hence, the MTTF of any life  $X$  is just “ $n$ ” times the MTTF of the first failure  $X_{(1)}$ . Therefore, by placing as many devices ( $n$ ) as we can afford on test, we will, with high probability, get a first failure (and estimations for all the parameters of interest) much sooner, thus saving calendar time as well as experimental costs.

Assume, for example, that we place  $n = 10$  expensive air conditioning units on a life test and that we observe the first failure after 1575 hours. From the distribution of Time to First Failure above, we know that the “average”  $X_{(1)}$  will occur ten times sooner than the “average” failure of a single unit (its MTTF is 10 smaller). In addition, by using the Total Test Time  $T = n \times X_{(1)}$  we can obtain a 95% CI for an air conditioning unit MTTF. Hence, the standard procedure for deriving a CI for  $\mu$ , with Type II censored data and  $k = 1$  is:

$$\left( \begin{array}{l} \frac{2nxX_{(1)}}{X_{2k,1-\alpha/2}^2} = \frac{2 \times 10 \times 1575}{7.38} = 4268.29; \frac{2nxX_{(1)}}{X_{2k,\alpha/2}^2} \\ = \frac{2 \times 10 \times 1575}{0.051} = 617647.06 \end{array} \right)$$

$$= (4268.29, 617,647.06)$$

## Summary and Conclusions

There are still ways in which the probability of failure, MTTF, L-10, the FR, etc. can be obtained, even when dealing with censored data, as long as we are able to assume that the device life follows the Exponential distribution. However, the degree of difficulty in obtaining such parameters increases as the distribution of the lives of the test data departs from the Exponential, and as the censoring mechanisms implemented become even more convoluted and complex. This START sheet reviews the Exponential case only. For all the other cases, the reader is directed to References 5, 6, 7, 8, and 9 of the bibliography.

## Bibliography

1. Statistical Assumptions of an Exponential Distribution, Romeu, J.L., RAC START: Volume 8, Number 2, <[http://rac.alionscience/pdf/E\\_ASSUME.pdf](http://rac.alionscience/pdf/E_ASSUME.pdf)>.
2. Reliability Estimations for the Exponential Life, Romeu, J.L., RAC START: Volume 10, Number 7, <[http://rac.alionscience.com/pdf/R\\_EXP.pdf](http://rac.alionscience.com/pdf/R_EXP.pdf)>.
3. A Practical Guide to Statistical Analysis of Material Property Data, Romeu, J.L. and C. Grethlein, AMPTIAC, 2000.
4. Probability and Statistics for Engineers and Scientists, Walpole and Myers, Prentice Hall, NJ, 1998.
5. An Introduction to Probability Theory and Mathematical Statistics, Rohatgi, V.K. Wiley, NY, 1976.
6. Methods for Statistical Analysis of Reliability and Life Data, Mann, N., R. Schafer and N. Singpurwalla, John Wiley, NY, 1974.
7. Practical Reliability Engineering, O'Connor, P., John Wiley, NY, 2002.
8. Reliability and Life Testing Handbook, Kececioglu, D., Editor, Volumes 1 and 2, Prentice Hall, NJ, 1993.
9. Weibull Analysis, Dodson, B., Quality Press, 1994.

## About the Author

Dr. Jorge Luis Romeu has over thirty years of statistical and operations research experience in consulting, research, and teaching. He was a consultant for the petrochemical, construction, and agricultural industries. Dr. Romeu has also worked in statistical and simulation modeling and in data analysis of software and hardware reliability, software engineering, and ecological problems.

Dr. Romeu has taught undergraduate and graduate statistics, operations research, and computer science in several American and foreign universities. He teaches short, intensive professional training courses. He is currently an Adjunct Professor of Statistics and Operations Research for Syracuse University and a Practicing Faculty of that school's Institute for Manufacturing Enterprises.

---

For his work in education and research and for his publications and presentations, Dr. Romeu has been elected Chartered Statistician Fellow of the Royal Statistical Society, Full Member of the Operations Research Society of America, and Fellow of the Institute of Statisticians. Romeu has received several international grants and awards, including a Fulbright Senior Lectureship and a Speaker Specialist Grant from the Department of State, in Mexico. He has extensive experience in international assignments in Spain and Latin America and is fluent in Spanish, English, and French.

Romeu is a senior technical advisor for reliability and advanced information technology research with Alion Science and Technology previously IIT Research Institute (IITRI). Since rejoining Alion in 1998, Romeu has provided consulting for several statistical and operations research projects. He has written a State of the Art Report on Statistical Analysis of Materials Data, designed and taught a three-day intensive statistics course for practicing engineers, and written a series of articles on statistics and data analysis for the AMPTIAC Newsletter and RAC Journal.

## Other START Sheets Available

Many Selected Topics in Assurance Related Technologies (START) sheets have been published on subjects of interest in reliability, maintainability, quality, and supportability. START sheets are available on-line in their entirety at <http://rac.alionscience.com/rac/jsp/start/startsheet.jsp>.

---

For further information on RAC START Sheets contact the:

Reliability Analysis Center  
201 Mill Street  
Rome, NY 13440-6916  
Toll Free: (888) RAC-USER  
Fax: (315) 337-9932

or visit our web site at:

<http://rac.alionscience.com>



## About the Reliability Analysis Center

The Reliability Analysis Center is a world-wide focal point for efforts to improve the reliability, maintainability, supportability and quality of manufactured components and systems. To this end, RAC collects, analyzes, archives in computerized databases, and publishes data concerning the quality and reliability of equipments and systems, as well as the microcircuit, discrete semiconductor, electronics, and electromechanical and mechanical components that comprise them. RAC also evaluates and publishes information on engineering techniques and methods. Information is distributed through data compilations, application guides, data products and programs on computer media, public and private training courses, and consulting services. Alion, and its predecessor company IIT Research Institute, have operated the RAC continuously since its creation in 1968.