<u>Zero Failures</u>: Calculate Reliability with No Recorded Failures

Tim Adams, ASQ CRE
NASA Kennedy Space Center
Engineering Directorate

ASQ Reliability and Risk Division Webinar
January 11, 2024

ASQ

Prospectus

Goal

- For the practitioner, provide methods with an illustrated Microsoft Excel tool to calculate reliability when the life-data set contains no recorded failures (zero failures).

Instructional Objective

- Given life data with no failures and the zero-failure Excel tool, calculate reliability, interpret the result, and describe the method(s) used.
- The methods will address both types of data (discrete and continuous) and both types of statistics (classical and Bayesian).

Prerequisites

- Familiarity (not mastery) with the following concepts (jargon): discrete data, continuous data, point estimate, interval estimate, confidence limit, one-sided confidence bound, probability distribution, probability density function, Bayesian inference, and Bayesian's prior distribution and likelihood function that make the posterior distribution.

ASQ

Tim.Adams@NASA.gov

Life-Data Types

1. Life-data types use different methods and math models to convert the data to a probability of success (reliability).
2. And these data types (e.g., Zero-Failure Data) can use more than one math model.
3. It is the analyst's job to select the method and math model(s) to estimate reliability.

	Data Sets in Reliability		
Number of Failures:	All	Some	None
Name of Data Set:	Complete Data	Censored Data	Zero-Failure Data
	Covariate Data (from an Accelerated Test)		Not Applicable

Tim.Adams@NASA.gov

Summary: Methods Used (details - slide 12)

Type of Statistics	Demand-Based Data	Time-Based Data
Interval Estimates via Classical Statistics	Binomial distribution via the F distribution	Exponential distribution via the chi-square distribution
	Inverse of the beta distribution	
	Informal-Bayesian Weibull distribution	
Point Estimates via Bayesian Statistics	Jeffreys beta prior distribution with binomial likelihood function	Jeffreys gamma prior distribution with Poisson likelihood function

1. What is "no failures" or "zero failures"?

- The life data sets with no failures appears in two forms:
  1. **No times-to-failure** since the data contains only run time or uptime durations.
  2. **No failure events** since the data contains only successes in all demands, tests, or attempts.

2. When do data sets with no failures occur?

- Data sets with no failures occur for two reasons:
  1. By design (via meeting specific test conditions without failing).
  2. By chance (via censored data).

- By design
  - A zero-failure test is used to infer a level of reliability. Thus, no failures are expected or desired to occur in the test.
- By chance
  - The historical operating data that is collected for all units of the item under study are still operating
  - Or did not incur a failure at the end of the previous use (i.e., “turned off” by choice and not “failed off” being the start of an undesired down state which will have an associated downtime or repair time).

Tim.Adams@NASA.gov

3. Why instruction for zero failures? (1 of 2)

- Classical statistics makes reliability equal to one ($R = 1$) for data sets with no failures (i.e., failure count, $f = 0$). This estimate for reliability is a point estimate that is overly optimistic and typically viewed as not true.

- For time-based (continuous) data, when the failure count equals zero, the **failure rate** (λ) equals zero
  - $\lambda = f/\Sigma t_i = 0/\Sigma t_i = 0$, where Σt_i is the sum of each unit's test or run time).

- For demand- or event-based (discrete) data, when the failure count equals zero, the **failure probability** (p_F) equals zero
  - $p_F = f/\Sigma d_i = 0$, where Σd_i is the sum of each unit's number of demands from each test or operation.
  - This makes the **success probability** (p_S) equal to one.
  - That is, since $p_F + p_S = 1$, then $p_s = 1 - p_F = 1 - 0 = 1$.

3. Why instruction for zero failures? (2 of 2)

- Bayesian statistics addresses the limitations of classical statistics in the mentioned cases by assuming an initial reliability level for the item.
  - Typically, this is done using a non-informative prior and updating the initial reliability estimate with the accumulated test data.

- In brief, **Bayesian Statistics** …
  - … quantitatively combines human belief (a subjective-based probability distribution, called the prior distribution, denoted **π**) with operational or test data that is evaluated using a likelihood function (an objective-based probability distribution, denoted **f**) to make a third distribution (called the posterior distribution, denoted **π_1**).
  - The Bayes' Theorem equation can be written as $\pi_1 = (\pi * f) / \sum (\pi * f)$.

- Slide 11 provides additional information on the Bayesian method.

4. How is the Excel Tool organized?

- Select a calculation method for data sets with zero failures based on three attributes:
    1. **Type of input** (continuous time-based or discrete demand-based)
    2. **Type of statistics** (classical or Bayesian)
    3. **Type of output** (point estimate or interval estimate)

For Zero Failures, Type of Output	**Excel Tool Provides**	**Excel Tool Does Not Provide**
Classical Statistics	Lower-Bound Confidence Interval for Reliability	Reliability Point Estimate less than one when f = 0
Bayesian Statistics	Point Estimate	Credible Interval

Tim.Adams@NASA.gov

5. Ref: Point Estimate vs. Interval Estimate

- Recall from 3, for life data with no failures, classical statistics does not provide a meaningful **point estimate**, a single number as the “best estimate” with no specification as to how much in error that number is likely to be (Lewis, p. 107).
- However, for life data with no failures, classical statistics does provide a meaningful but conservative **interval estimate or confidence interval**, a means by which the precision of the point estimate can be determined.
  - That is, the lower and upper confidence limits indicate how tightly the life data’s histogram (sampling distribution) is compressed around the true value of the estimated quantity. For example, the 95% confidence interval means that for 95% of the data sets, the true value of failure rate λ or failure probability p_F will lie between the calculated confidence limits (Lewis, pp. 28 & 30).
  - Reliability typically is concerned only with the lower limit—thus, the one-sided or one-tailed confidence limit is used.
- Bayesian Statistics does provide both a point estimate and interval estimate (**credible interval**).
  - For data sets with no failures and the distributions used in the Bayesian method, the Bayesian point estimate for reliability is more optimistic than classical lower-bound estimate for reliability.

6. Ref: Classical Statistics vs. Bayesian Statistics

- **Classical statistics** is based upon an experiment repeated an indefinite number of times (N) that determines the relative frequencies of occurrence in repetitions (f/N, where f is the frequency of occurrence; f/N is erratic with small N and stabilizes with large N). The interpretation of the probability (p) is called **objective** or **frequentist**.
  - Example: See page 2 at Comparison: Bayesian Statistics vs. Classical Statistics
- **Bayesian Statistics** interprets probability as an amount of epistemic confidence (strength of beliefs, hypotheses) rather than a frequency. Bayesian Statistics method quantitatively combines human belief (a subjective-based probability distribution, called the **prior distribution**) with operational or test data that is evaluated using a **likelihood function** (an objective-based probability distribution) to make a third distribution (called the **posterior distribution**). The posterior distribution is then used to provide the point estimate (e.g., as a failure rate or failure probability), expected value (mean), medium, and other distribution values.
- The Excel tool uses a special type of prior distribution called the Jeffreys prior. A **Jeffreys prior** is used when there is insufficient information to form an informed prior distribution about the failure rate or failure probability since the data currently on hand in this case has no failures.
  - The Jeffreys prior is referred to as a noninformative prior and is intended to convey little prior belief or information. A noninformative prior allows the data (described by the likelihood function) to speak for themselves.

7. Formulas used in the Excel Tool

<table>
<tr><th>For Zero Failures, Formulas Used</th><th>Demand-Based Data</th><th>Time-Based Data</th></tr>
<tr><td rowspan="3">Interval Estimates via Classical Statistics

This method works the demand-based and time-based methods in reverse that uses reliability to find sample size</td><td>• Finds the lower-bound reliability (R_L) for the binomial distribution via the F distribution. Statistical confidence is 1-α, number of attempts or tests is n, and number of successes is y. With no failures (y = n), then $R_L = \alpha^{1/n}$. (Ireson, Coombs, and Moss, pp. 25.34 & 25.35):

$$R_L = \frac{y}{y + (n - y + 1)F_{\alpha,2(n-y+1),2y}}$$
<u>Note</u>: (Modarres, p. 140) provides a failure-space version (i.e., uses f instead of y) that provides the upper-bound for failure probability instead of R_L.</td><td rowspan="2">• Finds the lower-bound mean time between failure ($MTTF_L$) and upper-bound failure rate (λ_U) for the exponential distribution via the chi-square distribution. Statistical confidence is 1-α, total test time is T, and no failures (f = 0) make (2f+2), the chi-square's degrees of freedom for a time-terminated (Type I) test, as 2. (Ebeling, p. 382):

$$MTTF_L = \frac{2T}{\chi^2_{\alpha,2f+2}} = \frac{2T}{\chi^2_{\alpha,2}} = \frac{T}{-\ln\alpha}$$
$$\lambda_U = \frac{1}{MTTF_L} = \frac{-\ln\alpha}{T}$$
$R_L = e^{-\lambda_U t}$ where <u>t</u> is mission time.

Since this method requires no failures (f = 0) and when T = n*t where n is number of items, then:
$R_L = \alpha^{1/n}$</td></tr>
<tr><td>• Finds the lower-bound reliability via the inverse of the beta distribution. With no failures (y successes = n attempts), then $R_L = \alpha^{1/n}$. (Microsoft Excel):

$$R_L = BETA.INV(\alpha, y, n - y + 1)$$</td></tr>
<tr><td colspan="2">• Finds the sample size n via the Informal-Bayesian Weibull distribution. Test duration is T with no failures, lower-bound reliability is R_L, mission time is t, statistical confidence is 1-α, and Weibull shape parameter is β (a priori information). When T = t and for any β > 0, then $R_L = \alpha^{1/n}$. (Dodson and Schwab, p. 115):

$$n = \frac{-\ln\alpha}{\left[\frac{T(-\ln R_L)^{\frac{1}{\beta}}}{t}\right]^{\beta}}$$</td></tr>
<tr><td>Point Estimates via Bayesian Statistics</td><td>• Finds the point-estimate-failure probability (p_F) a via Jeffreys beta prior distribution with binomial likelihood function. (<u>NASA Bayesian Inference Handbook</u>, p. 34):

$$p_F = \frac{\text{failure count} + 0.5}{\text{total number of attempts} + 1}$$</td><td>• Finds the point-estimate-failure rate (λ) via Jeffreys gamma prior distribution with Poisson likelihood function. (<u>NASA Bayesian Inference Handbook</u>, p. 40):

$$\lambda = \frac{\text{failure count} + 0.5}{\text{total run time}}$$</td></tr>
</table>

7. References for Formulas

- Dodson and Schwab, Accelerated Testing, 2006
- Ebeling, An Introduction to Reliability and Maintainability Engineering, 2005 reissue
- Ireson, Coombs, and Moss, Handbook of Reliability Engineering and Management, 2nd Edition, 1996
- Lewis, Introduction to Reliability Engineering, 2nd Edition, 1996
- NASA Bayesian Inference Handbook, 2009

Note: The formulas and references are provided in the Excel's "Read Me" worksheet.

Non-Example

of Zero-Failure Data Analysis

Given: Complete Data for 75 Times to Repair

Find: Probability the item will be repaired at time t

Analysis: See next three slides

Source: Quantitative RMA For The Practitioner, ASQ RRD Webinar, Jan 2018, slides 8-10

Tim.Adams@NASA.gov

Maintainability at time t, M(t_i) = ?

- What is the probability an item will be repaired in 2 hours (t_i = 120 minutes)? The item's repair-time history and frequency distribution of the repair times (histogram) are provided in the next two pages.

Answer:

- Since it is postulated the repair history fits a lognormal probability density function (PDF), transform each repair time with the natural logarithm [i.e., $t \rightarrow \ln(t)$]. For this transformed set, the mean (μ_n) is 5.295134 minutes and the population standard deviation (σ_n)* is 0.653912 minutes.

- M(t_i) = LOGNORM.DIST(t_i, μ_n, σ_n, True)

  M(120) =LOGNORM.DIST(120,5.295134,0.653912,True) = 0.2188 ≈ **22%**

- Thus, based on the history, as a point estimate there is a 22% chance the item will be repaired in 2 or less hours (or 78% chance after 2 hours).

* Note: The population standard deviation (not the sample standard deviation) is used when the maximum likelihood estimator (MLE) is used with complete data (no censored times) to estimate the parameters of the lognormal probability distribution.

75 historical times (minutes) to repair item x

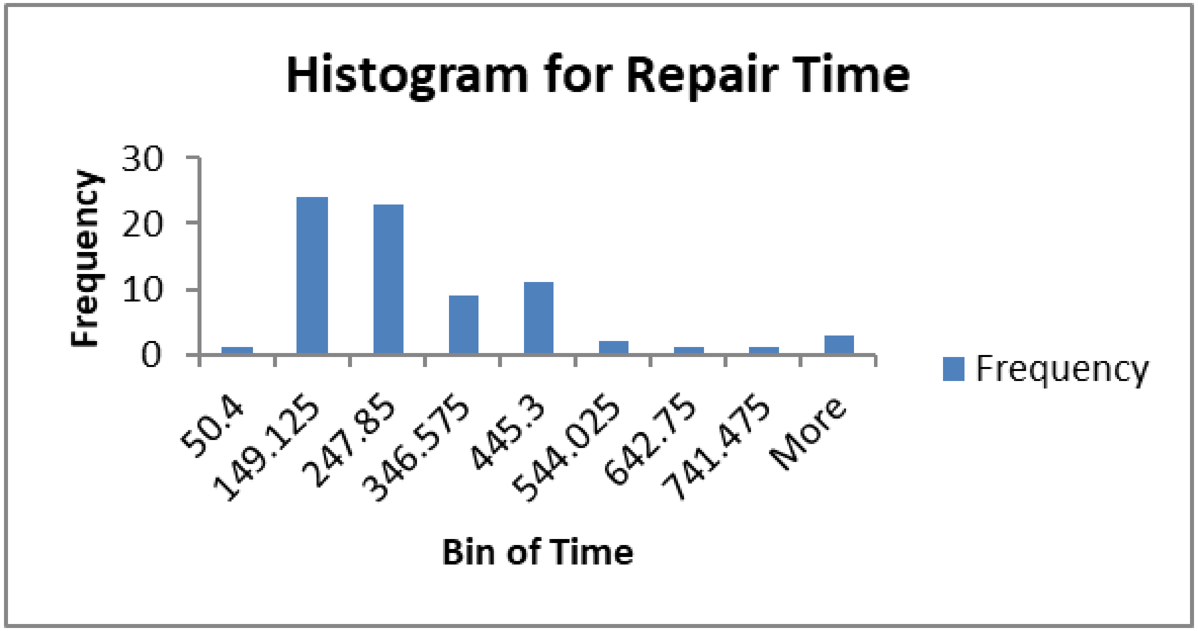
50.4	92.2	115.3	178.4	216.4	253.2	371.4	559.1
56.2	96.1	122.7	184.5	222.6	263.6	382.1	643.8
72.6	99.7	128.3	187.2	231	273.3	383.5	789.3
73.3	102.5	131.1	189.7	231.4	295.1	385	830.7
76.1	103.7	141.3	193.4	237.3	305.2	414	840.2
78.5	104.8	166	203.5	238.6	310.4	420.5	
80.6	105	166.1	204.1	243.7	340.7	426.5	
83.4	106.8	168	204.4	244.7	349.4	431	
84.6	107.3	170.6	215.3	252.1	355.8	457.4	
89	109.2	174.4	215.8	252.2	363.6	462.9	

- Statistics for **untransformed data**: Median = 204.40, mean (average) time to repair (MTTR) = 246.80, and sample standard deviation = 173.24.
- Statistics for **transformed data**: Median = 5.326816, mean (average) time to repair (MTTR) = 5.295134, sample standard deviation = 0.658316, and population standard deviation = 0.653912.
- Statistics via **lognormal relationships** based on transformed data: Median = 199.36 via =EXP(μ_n) or =LOGNORM.INV(0.5,μ_n,σ_n) and mean = 246.89 via =EXP(μ_n+0.5*(σ_n^2)).

Tim.Adams@NASA.gov

75 repairs times as a histogram

- The **histogram** below is made by Excel's "Analysis ToolPak" add-in.
- The shape of this histogram suggests the lognormal probability distribution's probability density function (PDF).
- The **lognormal PDF** "is often used to model usage data, such as … repair time of a maintained system." (Practical Reliability Engineering, 5th Edition, O'Connor, 2012, pp. 35 & 410)



Bio – Timothy C Adams

- Tim is a Senior Engineer in the Engineering Directorate at the NASA John F. Kennedy Space Center's (KSC). He serves as a technical resource in engineering assurance with a specialty in quantitative Reliability Engineering and Technical Risk -- and he is the founder and Technical Editor of the "KSC Reliability" website.

- Tim has 30 years of experience in Reliability Engineering and related engineering assurance disciplines. At Johnson Space Center (JSC), he was a Flight Systems Safety Engineer, Reliability Engineer, and Lead of the Office of Safety, Reliability, and Quality Assurance's Analysis and Assessment Methodology Group.

- For the American Society for Quality (ASQ), Tim is a senior member, a Certified Reliability Engineer (CRE), and was a member of an ASQ team that reviewed the CRE exam.

- Contact Information: Tim.Adams@NASA.gov 321-258-2415

ASQ