Report Date: 2026-02-02

Lifetime Analysis System - Statistical Reliability Assessment

TECHNICAL ANALYSIS REPORT

1) Analysis title:	Weibull Distribution Lifetime Analysis
2) Data source file:	dataset-1.csv

3) Main objectives:

- Perform statistical distribution fitting to identify optimal failure model
- Estimate Weibull parameters (shape and scale) with confidence intervals
- Generate reliability predictions and lifetime percentiles (B10, B50, B95)
- Validate model accuracy through goodness-of-fit testing
- Provide engineering recommendations for maintenance and design decisions

4) Statistical methods:

- Maximum Likelihood Estimation (MLE) for parameter fitting
- Bootstrap resampling (1000 iterations) for uncertainty quantification
- Monte Carlo simulation (10,000 samples) for lifetime prediction
- Kolmogorov-Smirnov test for model validation

5) Summary of analysis results:

Parameter	Estimated Value	Confidence Interval
Shape Parameter (Beta)	3.01	Bootstrap 95% CI
Scale Parameter (Eta)	24538.97	Bootstrap 95% CI
B10 Life (10% failure)	7999.1 cycles	Monte Carlo
B50 Life (median)	21661.22 cycles	Monte Carlo
B95 Life (95% failure)	35008.37 cycles	Monte Carlo
Mean Time To Failure	21783.01 cycles	Calculated
Model Validation	Kolmogorov-Smirnov	p-value: 0.6762

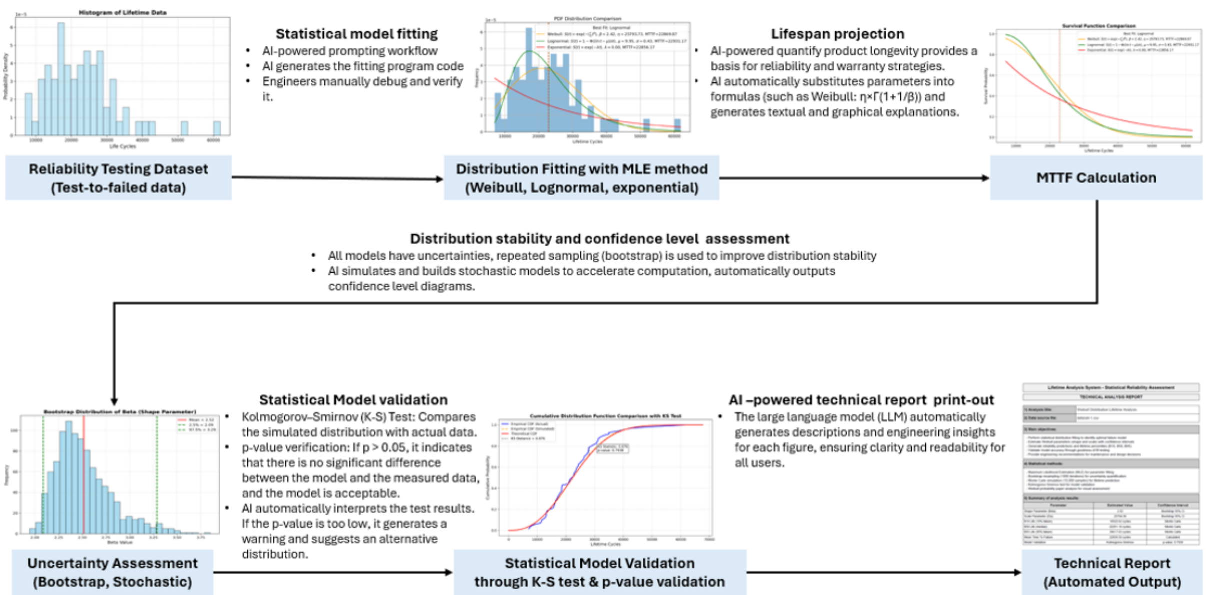
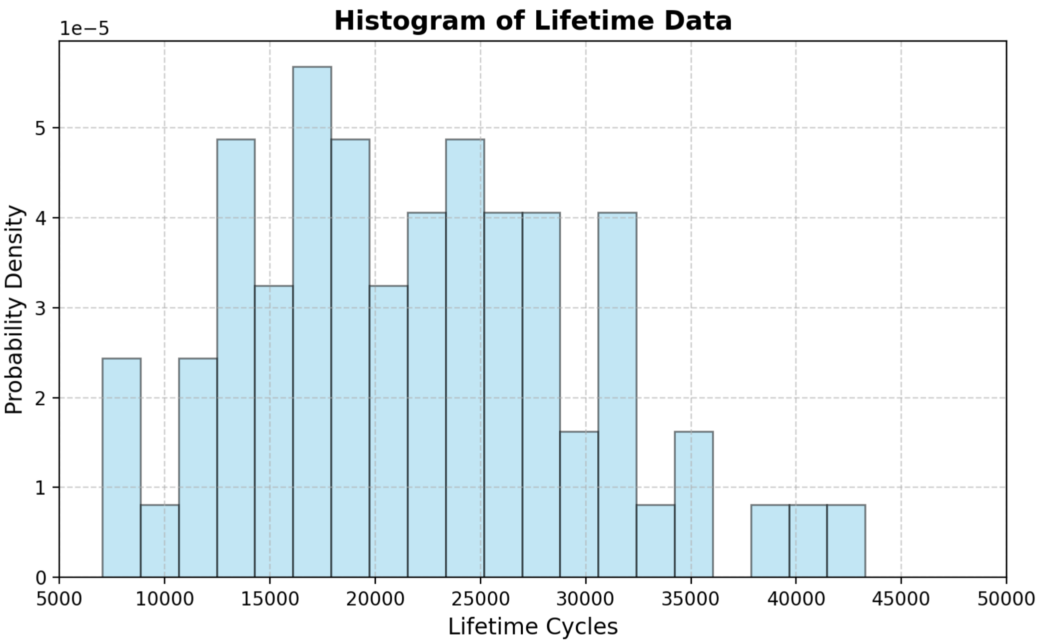


Figure 1: Histogram of Lifetime Data



Description:

This figure illustrates the distribution of the observed failure times collected from the test samples. Each bar in the histogram represents the frequency of failures occurring within a specific range of lifetime cycles. By examining the overall shape of the histogram, we can identify how the failures are distributed over time - for instance, whether most components fail early (indicating early-life issues), steadily (indicating random failures), or predominantly at later stages (indicating wear-out mechanisms).

Engineering Significance:

Understanding the failure-time distribution allows engineers to:

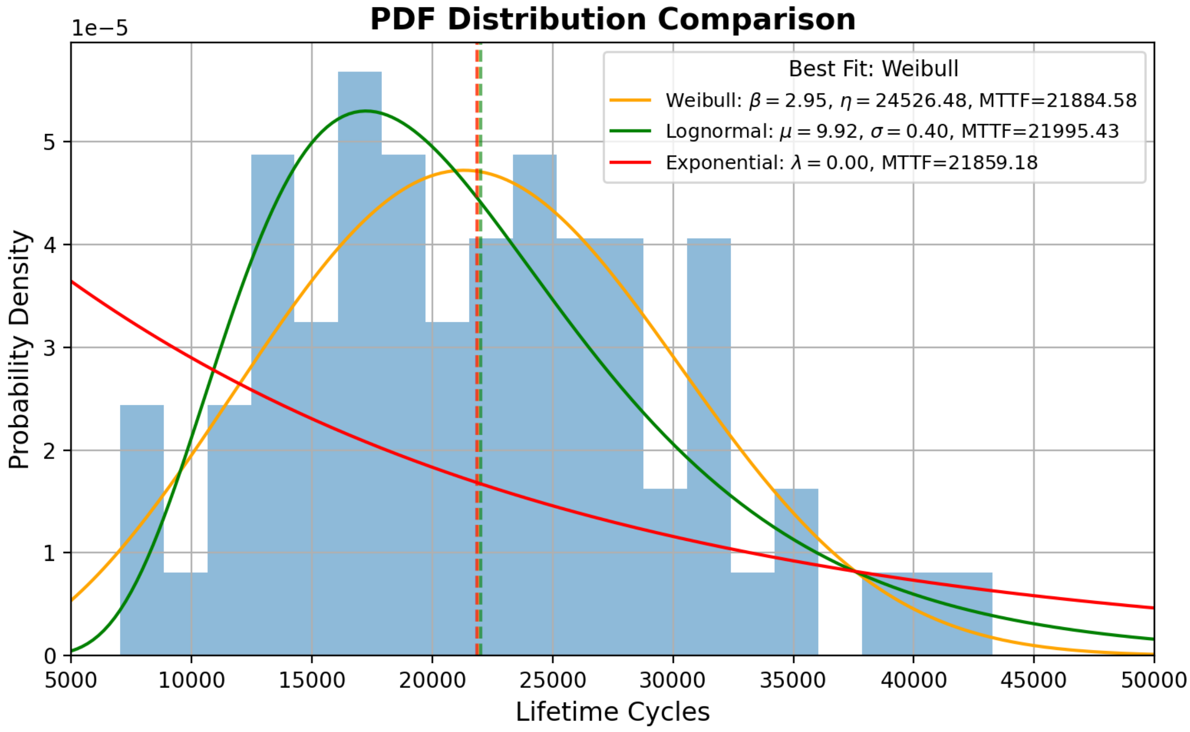
Assess failure behavior - For example, a right-skewed distribution with a long tail may suggest a small portion of components last significantly longer than average, while a steep early peak could imply process or material inconsistency.

Choose an appropriate statistical model - The histogram's shape provides visual cues for model selection (e.g., Weibull for wear-out failures, Exponential for constant failure rates).

Guide reliability improvements - If failures are concentrated in early cycles, process control or burn-in testing may be needed; if they appear in later cycles, focus should shift toward material fatigue, corrosion, or long-term degradation mechanisms.

This figure thus establishes the empirical baseline of the product's lifetime characteristics, forming the foundation for quantitative modeling and predictive reliability assessment in later stages.

Figure 2: Probability Density Function (PDF) Comparison



Description:

This figure compares three lifetime models, Weibull, Lognormal, and Exponential. Each fitted to the observed failure data using the Maximum Likelihood Estimation (MLE) method. MLE is a statistical approach that finds the most plausible model parameters so that the fitted curve best represents how failures occur over time.

Each colored curve represents a probability density function (PDF), which describes the likelihood that a component will fail at a given point in its lifetime. The Weibull distribution is widely used for mechanical parts because it can represent early-life, random, or wear-out failures depending on its shape. The Lognormal distribution often models chemical or environmental degradation, where aging accumulates gradually. The Exponential distribution assumes a constant failure rate and is typical for electronic or random-stress failures.

Once the models are fitted, the Mean Time To Failure (MTTF) can be calculated directly from their estimated parameters.

Engineering Significance:

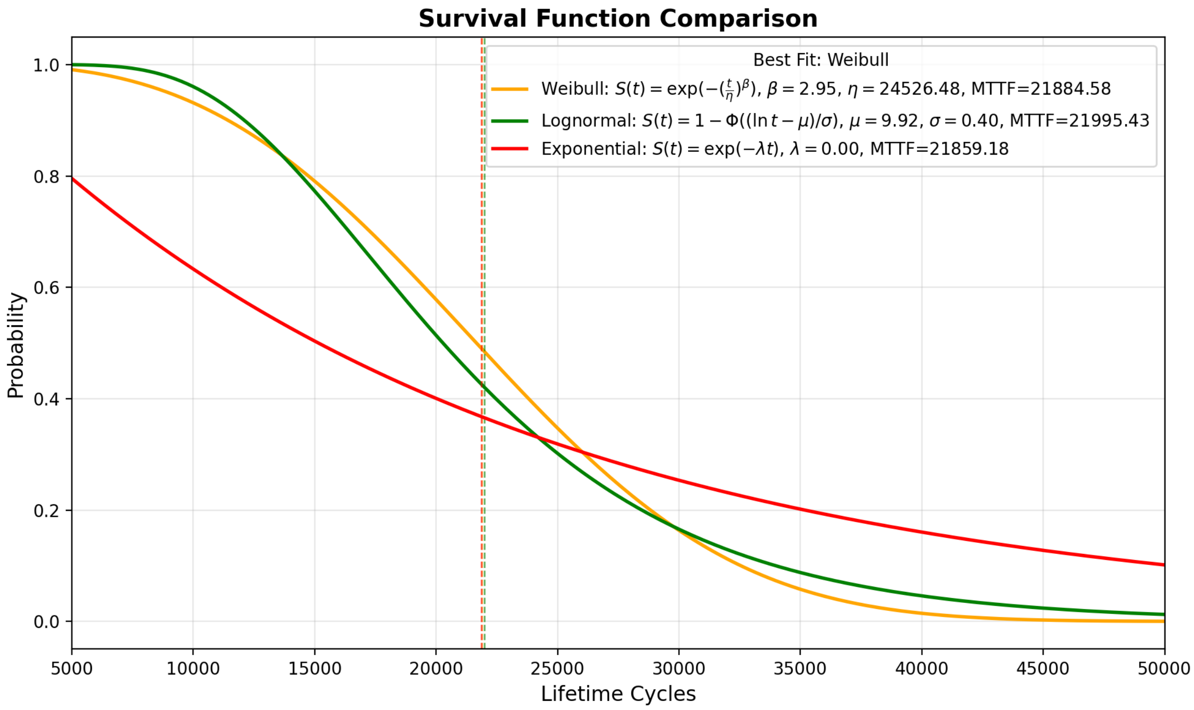
This comparison highlights how different assumptions lead to different reliability interpretations. By observing which model best fits the data:

Model Selection Insight - Identifying which distribution best fits the data helps engineers infer whether failures are due to material aging, process variability, or random stress.

Quantitative Lifetime Estimation - The fitted model provides a consistent way to compute MTTF, transforming test results into measurable reliability metrics.

Design & Clarity Guidance - Comparing these model curves visually demonstrates how mathematical modeling translates raw test data into practical engineering understanding.

Figure 3: Survival Function Comparison



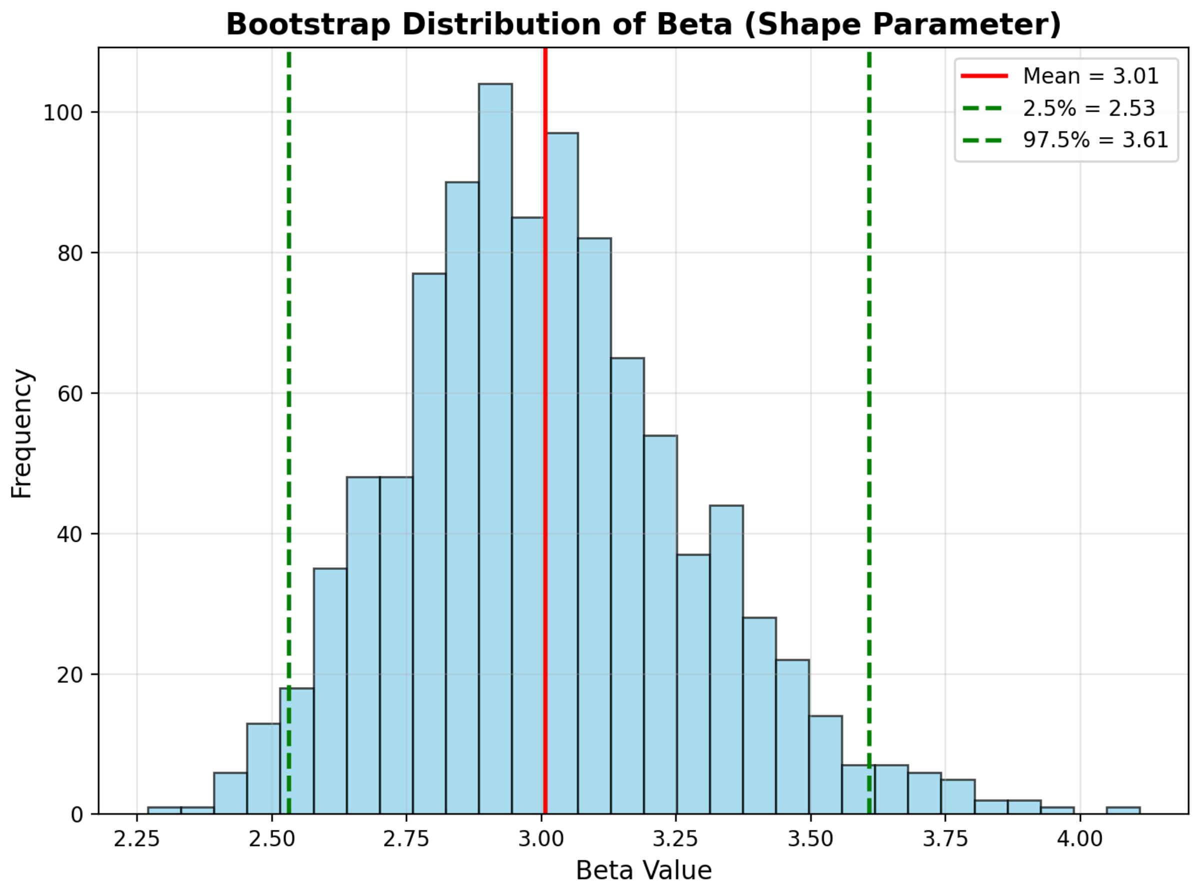
Description:

This figure presents the Survival Function, denoted as S(t), for the analyzed failure data compared against the fitted probability distributions (e.g., Weibull, Lognormal, and Exponential). The curve represents the probability that a component will perform its required function without failure for a specific period. Starting at a probability of 1.0 (100%), the curve monotonically decreases over time, visually depicting the reliability decay of the product. The model that best aligns with the empirical data provides the most accurate baseline for predicting the remaining population at any given cycle count.

Engineering Significance:

The Survival Function is the primary tool for communicating reliability to stakeholders and guides several critical decisions:
Determine Warranty Periods - By pinpointing the time at which the survival probability remains high (e.g., 90% reliability), engineers can set warranty limits that minimize financial risk from returns.
Plan Maintenance Schedules - The slope of the curve indicates how quickly reliability is lost. A steep drop suggests a need for preventive maintenance before the "knee" of the curve is reached.
Predict Fleet Availability - For large-scale deployments, this function estimates the percentage of units that will remain operational at a future date, aiding in spare parts inventory planning.

Figure 4: Bootstrap Distribution of Beta (Shape Parameter)



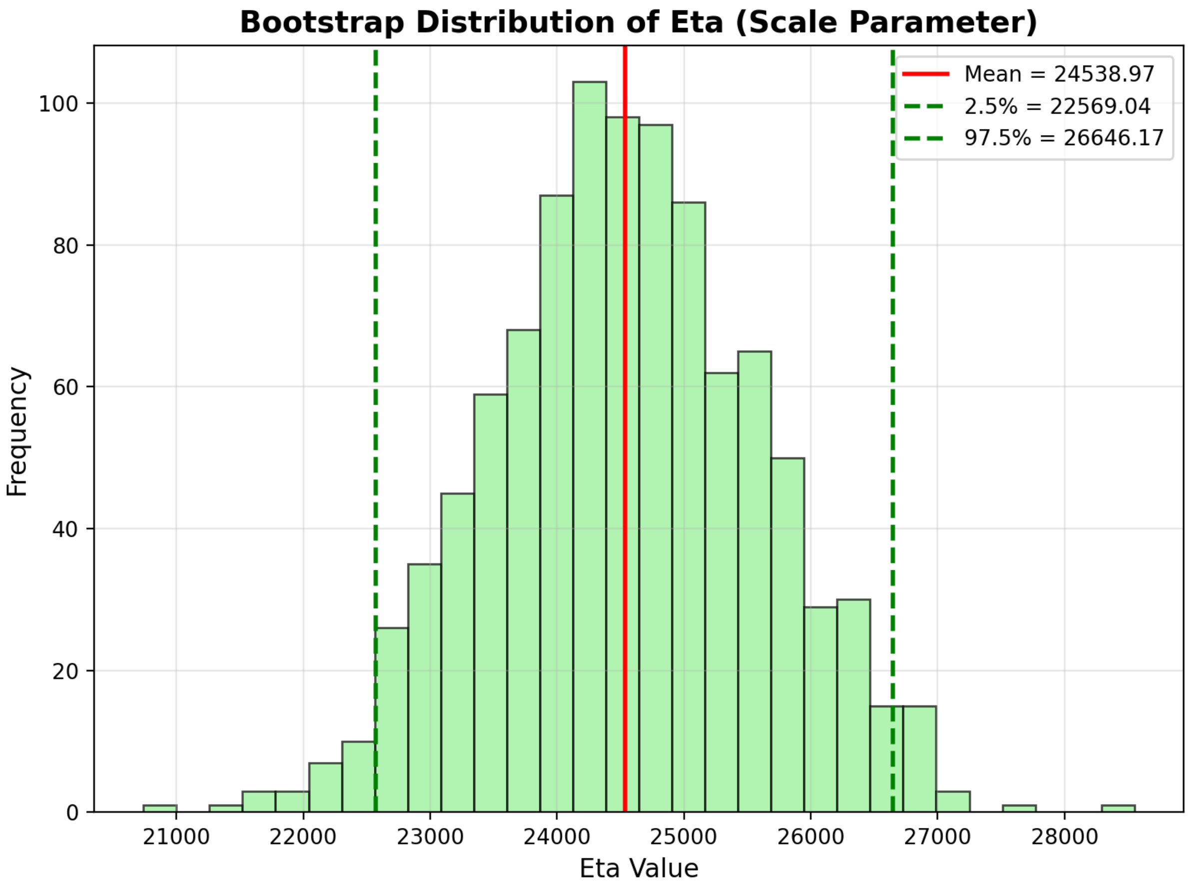
Description:

This histogram illustrates the uncertainty associated with the Weibull Shape Parameter (Beta), generated using the Bootstrap Resampling method (e.g., 1000 iterations). Instead of calculating a single static value, this technique simulates hundreds of "virtual experiments" by resampling the original data with replacement. The resulting distribution shows the range of probable values for Beta. The red vertical line indicates the mean estimated Beta, while the dashed green lines define the 95% Confidence Interval (CI) (spanning from the 2.5th to the 97.5th percentile). This visualization transforms a single point estimate into a probability distribution, revealing the statistical stability of the calculated failure mode.

Engineering Significance:

Quantifying the uncertainty of the shape parameter is critical for validating the physics of failure:
Confirm Failure Mode Statistically - A Beta value greater than 1.0 implies wear-out. However, this chart proves it rigorously. If the lower bound of the 95% CI is also strictly greater than 1.0, engineers can claim with 95% statistical confidence that the failure mechanism is indeed wear-out, ruling out random variation.
Assess Data Quality - The width of the distribution reflects the "tightness" of the data. A narrow, sharp peak indicates high consistency and sufficient sample size, whereas a wide, flat distribution suggests noisy data or too few samples, warning that the model parameters may be volatile.
Support Robust Decision Making - By understanding the potential range of Beta, engineering teams can base their maintenance strategies not just on the "average" scenario, but on conservative estimates that account for sampling error.

Figure 5: Bootstrap Distribution of Eta (Scale Parameter)



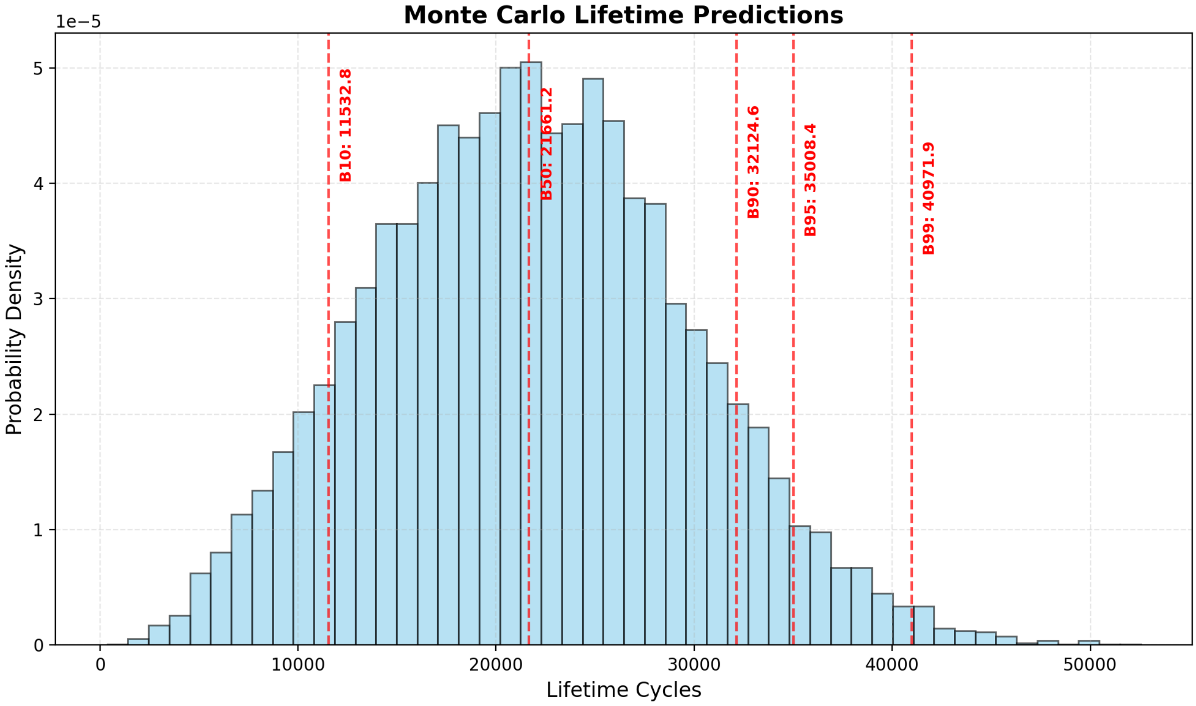
Description:

This histogram depicts the uncertainty distribution of the Weibull Scale Parameter (Eta), also known as the Characteristic Life. Mathematically, Eta represents the time at which 63.2% of the population is expected to fail. By applying bootstrap resampling, this chart reveals the potential variability in the product's longevity due to sampling error. The central red line marks the mean estimated Characteristic Life, while the green dashed lines delineate the 95% Confidence Interval. This interval provides a realistic "best-case" and "worst-case" scenario for the product's lifespan.

Engineering Significance:

Understanding the range of the Scale Parameter is essential for logistics and financial planning:
Plan Spare Parts Inventory - The lower bound of the confidence interval (2.5% percentile) is critical for supply chain management. It indicates the earliest timeframe where a significant volume of replacements might be needed, preventing stockouts.
Assess Manufacturing Consistency - A narrow distribution for Eta implies a highly controlled manufacturing process where unit-to-unit variation is minimal. A wide spread suggests process instability or inconsistent raw materials, leading to unpredictable product lifetimes.
Compare Supplier Performance - When evaluating components from different vendors, comparing their Eta distributions allows for a statistical decision. Even if two suppliers have the same "average" life, the one with the narrower confidence interval is the superior choice due to its predictability.

Figure 6: Monte Carlo Lifetime Predictions



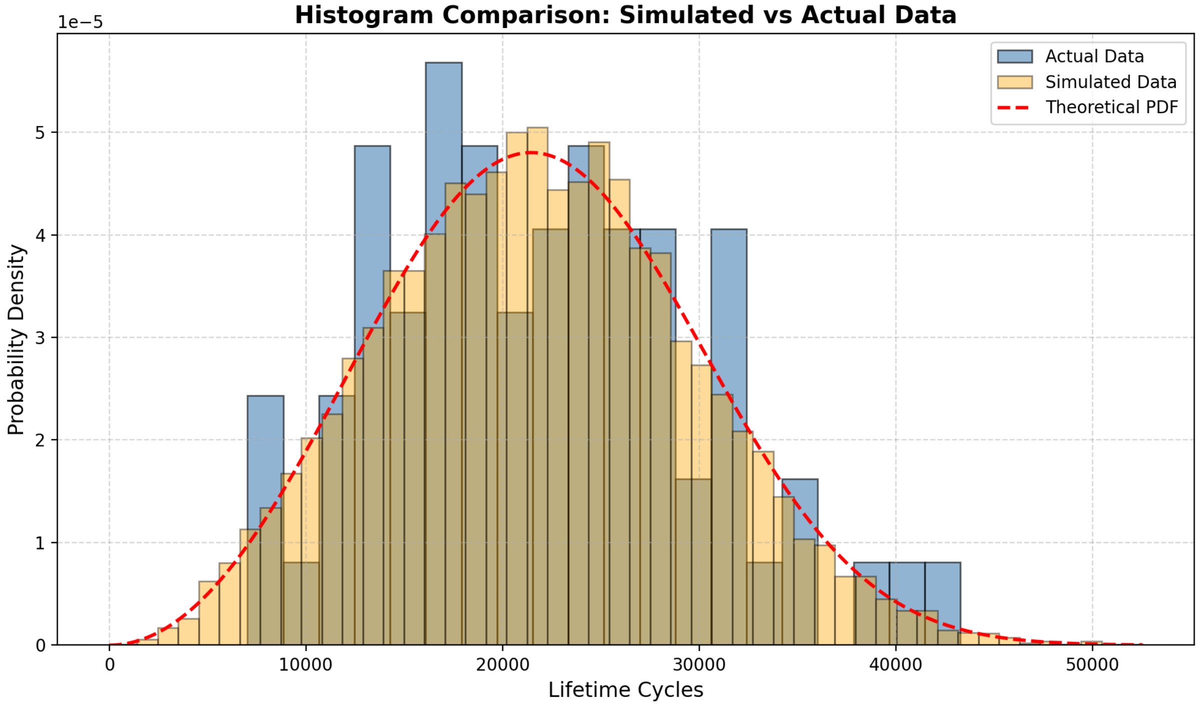
Description:

This figure presents the results of a Monte Carlo Simulation (e.g., 10,000 runs) used to forecast the expected lifetime distribution of the population. Unlike simple calculations that rely on single parameter estimates, this simulation inputs the range of probable Weibull parameters (Beta and Eta) derived from the bootstrap analysis to generate thousands of hypothetical failure times. The resulting histogram shows the probability density of these predicted outcomes. Key reliability milestones are explicitly marked: B10 Life (the time by which 10% of units are expected to fail) and B50 Life (the median life, where 50% of units have failed).

Engineering Significance:

This predictive model translates statistical data into actionable business and engineering metrics:
Define Warranty Risks (B10 Focus) - The B10 value is the industry standard for defining warranty periods. Setting a warranty term below this threshold ensures that fewer than 10% of products will fail in the field, keeping replacement costs within the budget.
Schedule Preventive Maintenance (B50 Focus) - The B50 value indicates the "average" life expectancy. Maintenance teams can use this metric to plan fleet-wide overhauls or end-of-life replacements before the bulk of the population reaches wear-out.
Quantify Tail Risk - By visualizing the full spread of outcomes, engineers can assess the risk of "early outliers" (left tail) versus "long-surviving units" (right tail), ensuring the system design accommodates the full variability of the manufacturing process.

Figure 7: Histogram Comparison: Simulated vs Actual Data



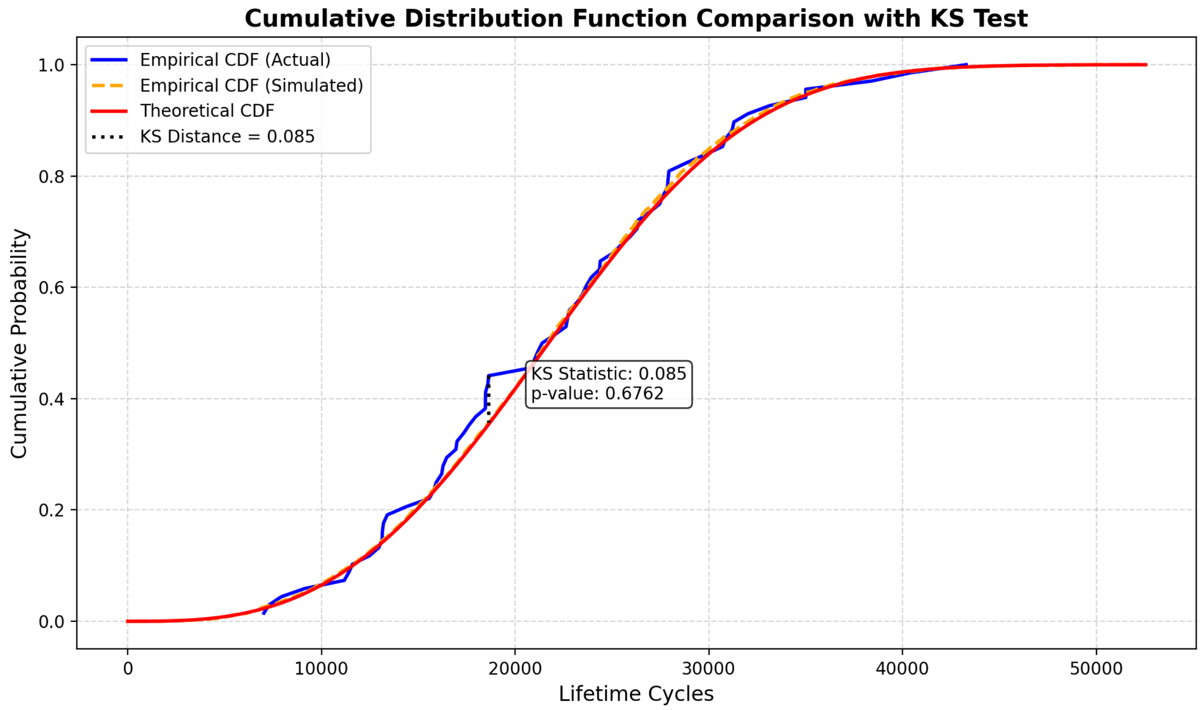
Description:

This figure presents a visual validation of the fitted model by overlaying the histogram of the Actual Observed Data (blue bars) with the Simulated Data (yellow bars) generated from the Weibull model. The red dashed line represents the theoretical Probability Density Function (PDF). The height of the bars corresponds to the probability density, allowing for a direct comparison of the distribution shapes. A high degree of overlap between the blue and yellow areas indicates that the simulation accurately replicates the physical failure behavior observed in reality.

Engineering Significance:

Visualizing the fit is a crucial first step in model validation:
Verify Distribution Shape - It confirms whether the model correctly captures key characteristics such as the peak failure time and the spread (variance). For instance, if the simulated data (yellow) accurately tracks the tail of the actual data (blue), the model is suitable for predicting long-term wear-out.
Detect Systemic Biases - Significant gaps between the actual and simulated histograms would reveal where the model under- or over-estimates risk (e.g., missing early-life failures), prompting a review of the chosen distribution.

Figure 8: CDF Comparison & KS Test



Description:

This figure illustrates the quantitative Goodness-of-Fit using the Cumulative Distribution Function (CDF). It compares the Empirical CDF (blue stepped line, representing actual data) against the Theoretical CDF (red smooth curve, representing the fitted model). The plot includes the results of the Kolmogorov-Smirnov (KS) Test, a non-parametric statistical method. The "KS Distance" marks the maximum vertical divergence between the two curves, quantifying the error. The displayed p-value indicates the statistical significance of the fit; a higher p-value suggests that the model is consistent with the observed data.

Engineering Significance:

The CDF comparison provides the statistical evidence required for final model acceptance:
Statistical Acceptance Criteria - The p-value serves as the objective pass/fail metric. Typically, a p-value > 0.05 implies there is no significant evidence to reject the model, validating its use for reliability predictions.
Assess Prediction Accuracy - The proximity of the blue and red curves, particularly in the upper percentiles, assures engineers that the model is reliable for calculating critical metrics like warranty limits (B10 life) and mean life (MTTF).