Debunking Stress Rupture Theories Using Weibull Regression Plots

Anne Ryan Driscoll

DataWorks 2021

Virginia Tech adriscoll@vt.edu

NASA Strand and Vessel Testing

- NASA's Engineering Safety Center (NESC) project to assess safety of Composite Overwrapped Pressure Vessels (COPVs)
- COPVs
  - Transport gasses under high pressure
  - Metal Liner
  - Wrapped by a Series of Carbon Strands
- Research Question: **Reliability of COPVs at Use Conditions for the Expected Mission Life**
  - Primary Focus on Strands
  - Secondary Focus on Relationship to Vessels
  - Strands Less Expensive to Test
- https://www.nasa.gov/offices/nesc/home/Feature_COPVs_Jan-2012.html

Analysis: Ordinary Least Squares

Example: Jet Turbine Engine Thrust

- A researcher is tasked with creating a model to explain jet turbine engine thrust* with the following predictors
  - x_1: primary speed of rotation
  - x_2: fuel flow rate
  - x_3: exhaust temperature
  - x_4: ambient temperature at time of test

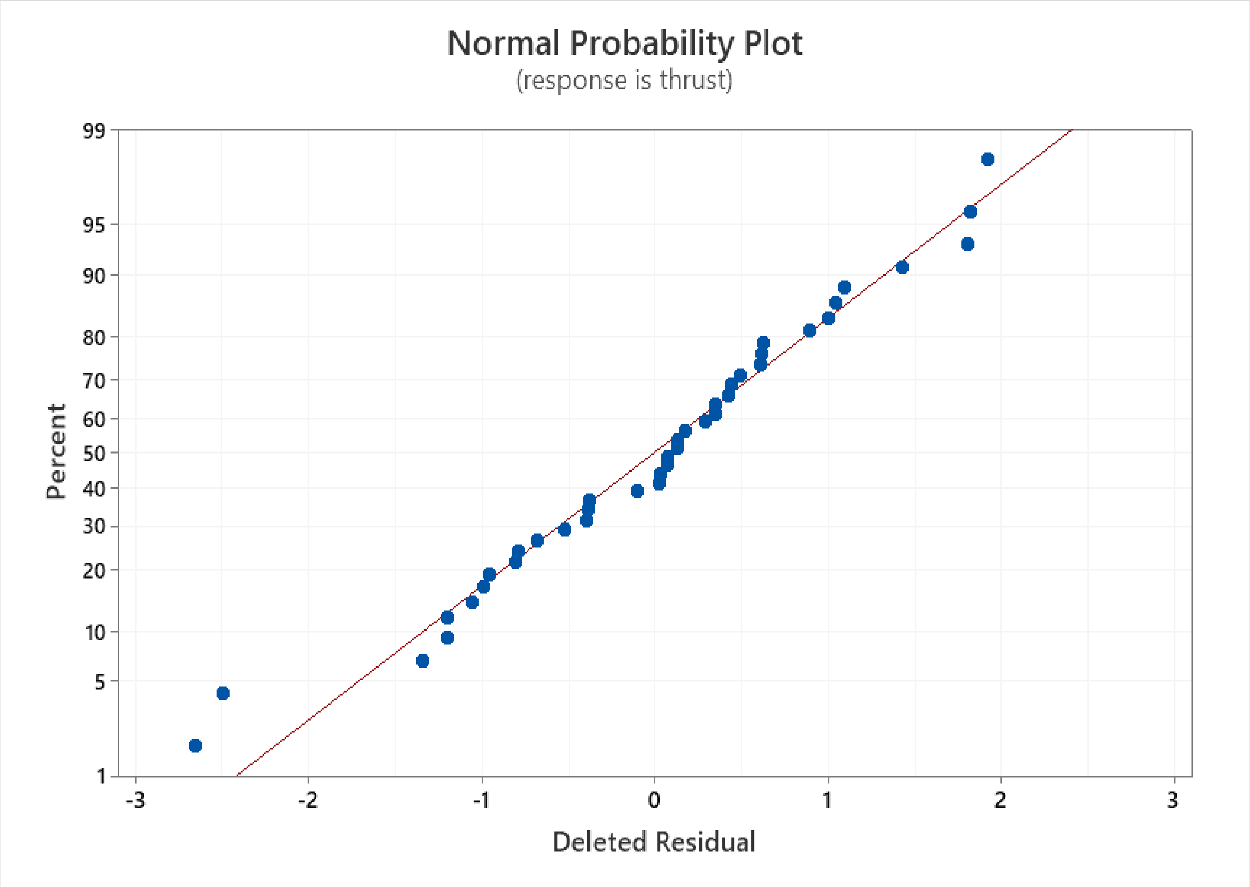
- Begin analysis with ordinary least squares (OLS) regression

*Jet Turbine data in Table B.13, page 566 of Montgomery, Peck, and Vining (2012)

Checking Assumptions

- Proper OLS regression requires checking the basic assumptions of the modeling technique
- These assumptions can be explored through the **residuals** of the analysis
- Basic Assumptions
    1. constant error variance
    2. independent data
    3. roughly normal distribution

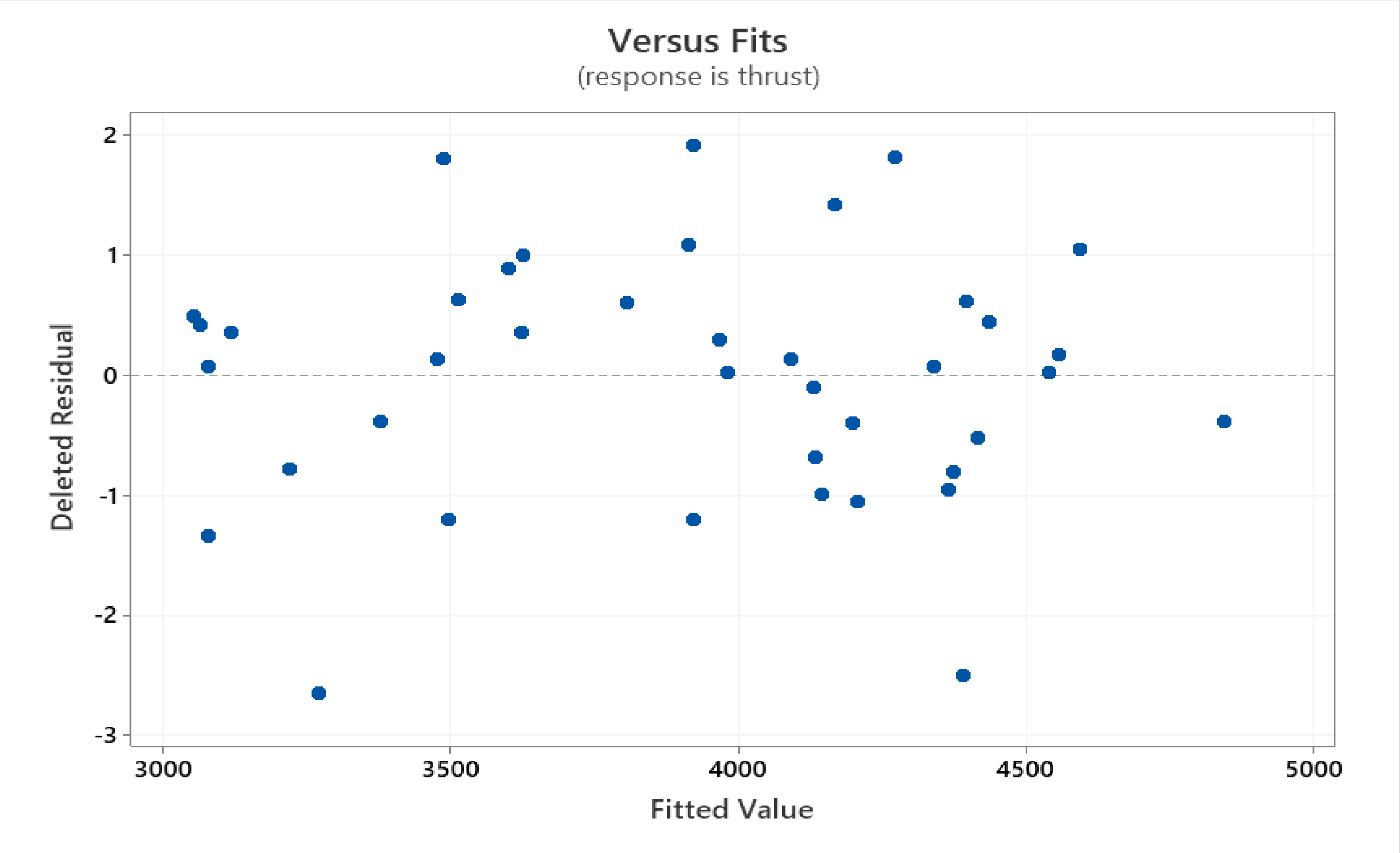
Normal Probability Plot: Jet Turbine Data



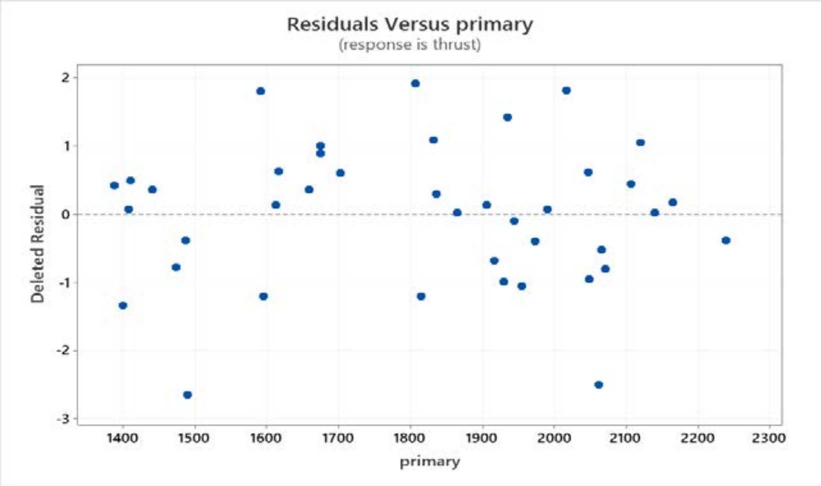
- Adequate fit to the data to assume a normal distribution

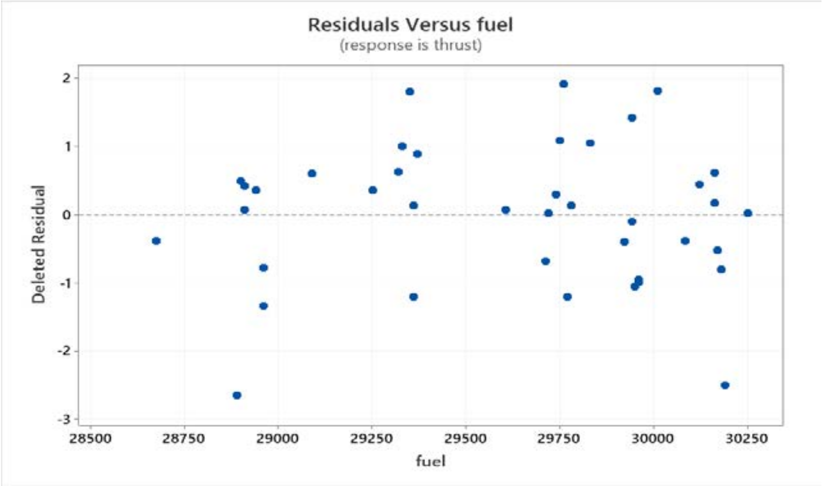
Residuals vs. Fits Plot: Jet Turbine Data

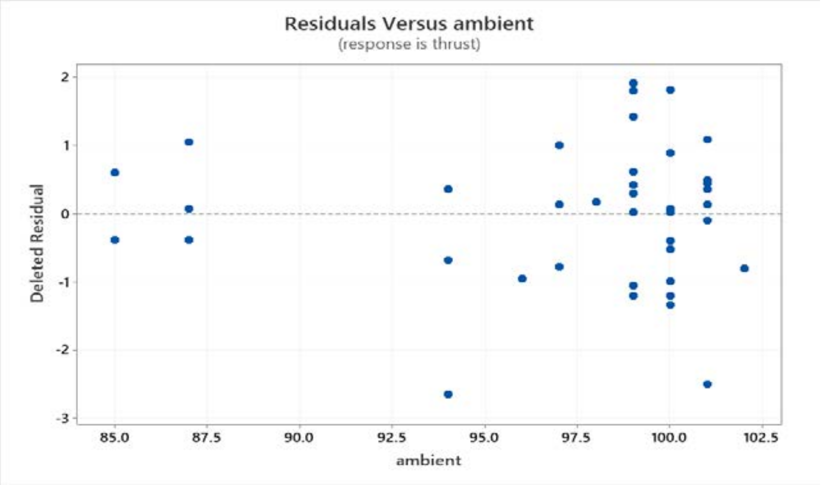
- Goal: To check for any special patterns
- Is the variance constant?
- No major issues in plot to suggest transforming data

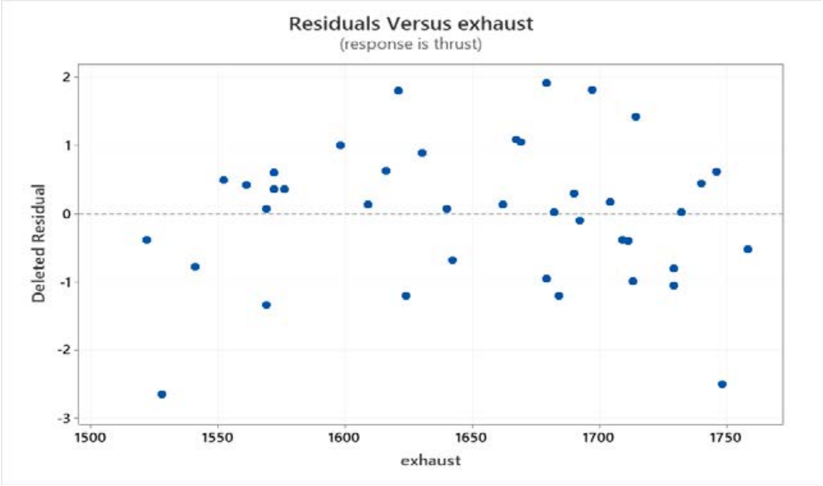


Residuals vs. Explanatory Variables







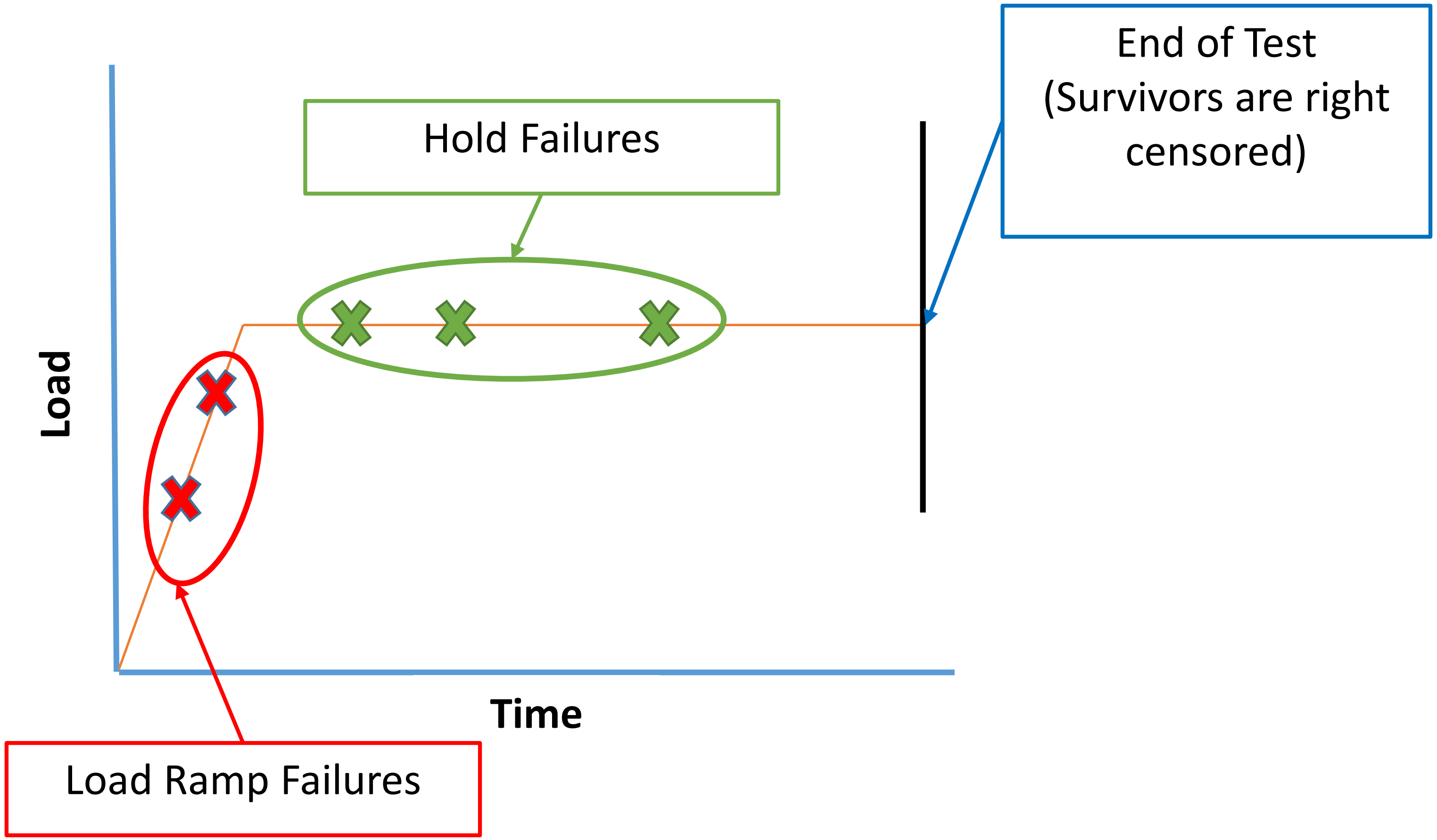


- Check for unusual patterns. Not major issues in these plots.

Analysis: Stress Rupture Data

Description of Stress Rupture Test

- Stress Rupture
  - Failures occur after a period of time where there is no increase in load
- Failures are needed to determine reliability
- Must extrapolate from where test is performed versus where reliability predictions are made
- Test strands at higher loads and then extrapolate
- Need a model to make predictions



Basic Weibull Distribution

- Probability Density Function

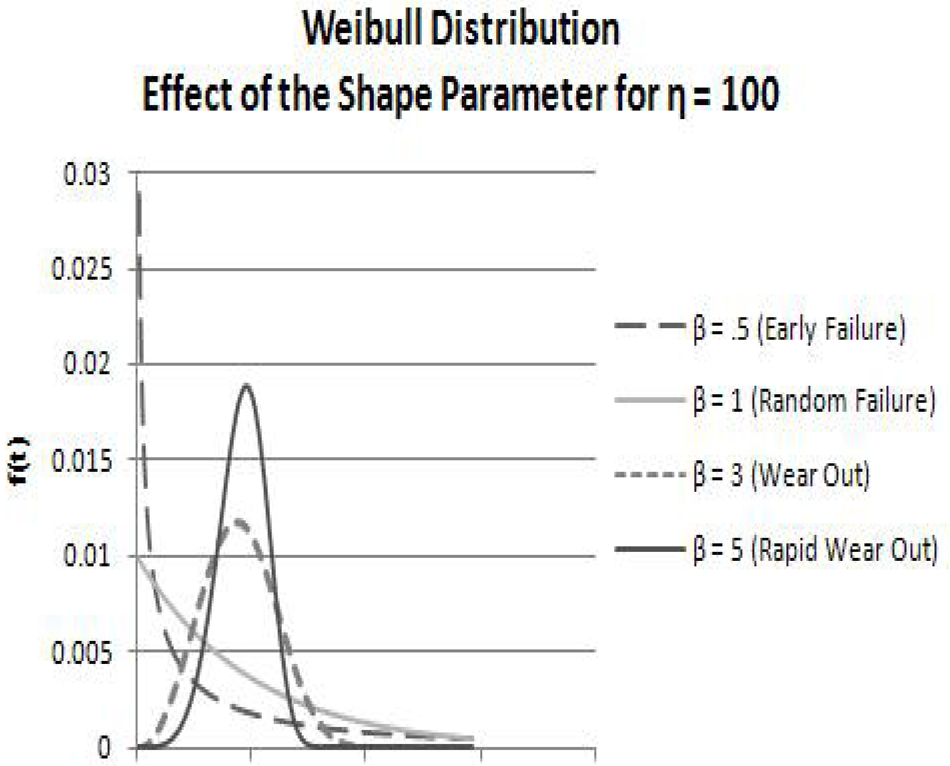
$$f(t, \beta, \eta) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} \mathrm{e}^{-\left(\frac{\mathrm{t}}{\eta}\right)^{\beta}}$$

- Survivor Function = 1 – F(t)

$$S(t) = e^{-\left(\frac{t}{\eta}\right)^{\beta}}$$

β: shape parameter

η: scale parameter (characteristic life, time where 63.2% of units will fail)



Smallest Extreme Value Distribution

- Smallest extreme value (SEV) distribution is an alternate parameterization of the Weibull distribution
- SEV represents Weibull as a log-location-scale distribution
  - If t_i is Weibull, then $\log(t_i)$ is SEV
- Weibull/SEV relationship mimics the Normal/Lognormal relationship
- Parameters
  - Log-location: $\mu = \log(\eta)$ Scale: $\sigma = \frac{1}{\beta}$
- Residuals
  - $e_i = \log t_i - \mu$
  - Scaled: $z_i = \beta e_i = \beta(\log t_i - \mu) = \frac{\log t_i - \mu}{\sigma}$
- Survivor Function
  - $S(t_i) = P(T > t_i) = e^{-e^{z_i}}$

Classic Stress Rupture Model: Weibull

- Classic Weibull Survival Function

$$S(t_i) = P(T > t_i) = e^{-\left(\frac{t_i}{t_{ref}} SR^{\rho}\right)^{\beta}}$$

Note: $\eta = t_{ref} SR^{-\rho}$

- Observed Life Time: t_i
- SR: Stress Ratio, ratio of stress level to strength scale parameter
- Critical Parameters:
  - ρ: controls the relationship between the failure time and stress ratio (SR)
  - β: Shape parameter for time to Failure
  - t_{ref}: Reference time to Failure

Classic Stress Rupture Model: SEV

- SEV Survival Function

$$S(t_i) = e^{-\left(\frac{t_i}{t_{ref}} SR^{\rho}\right)^{\beta}} = e^{-e^{\beta(\log t_i - \theta + \rho\ \ln(SR))}}$$

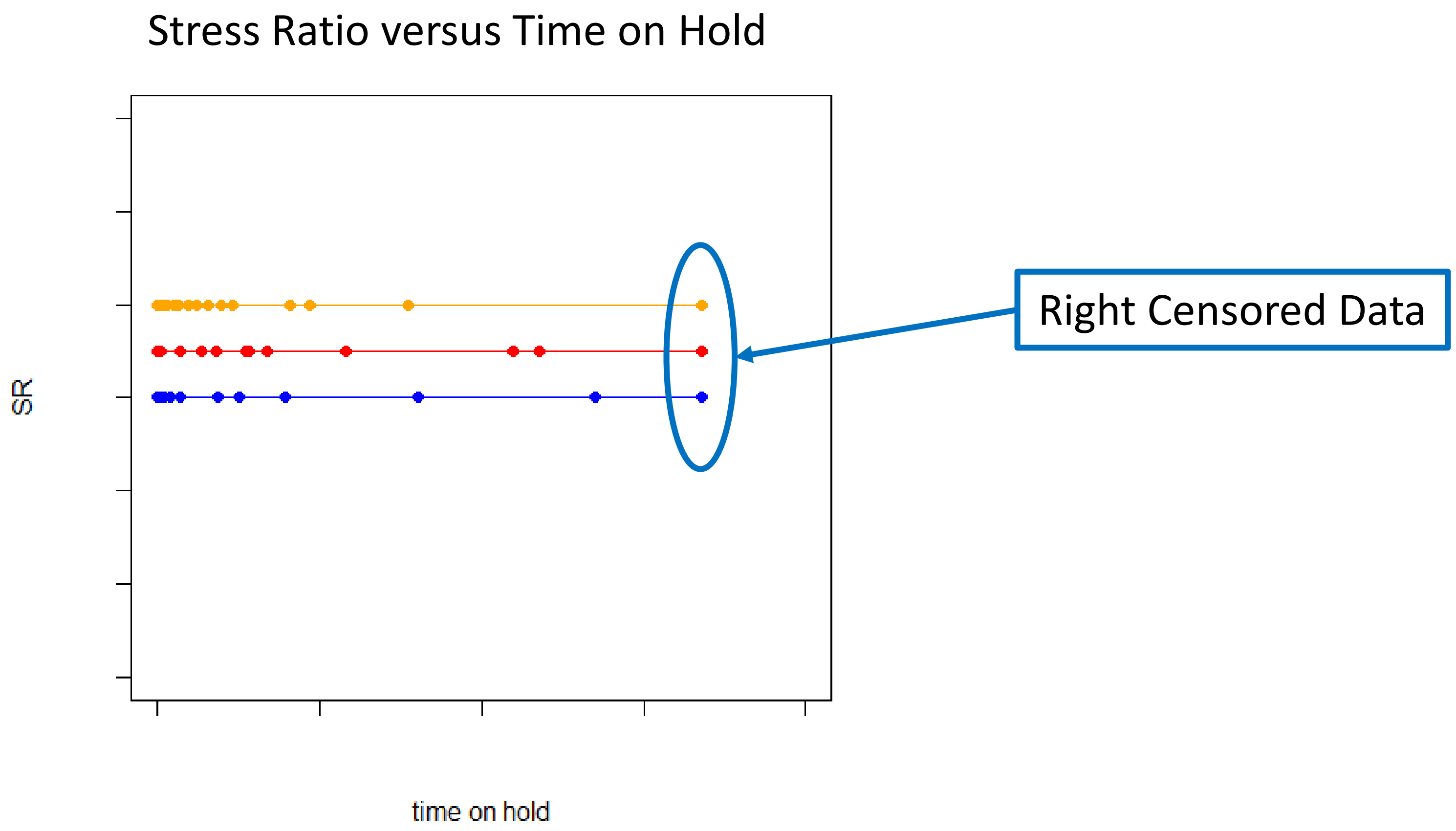
where $\theta = \log(t_{ref})$ and $\mu = \log(\eta) = \theta - \rho\ \ln(SR)$

Now working with a linear model, similar to simple linear regression

- Scaled Residuals
  - $z_i = \beta e_i = \beta(\log t_i - \mu) = \beta(\log t_i - \theta + \rho \ln(SR))$
  - Used for predictions of the log probability for specific observations

Weibull Regression: Fitting the Model

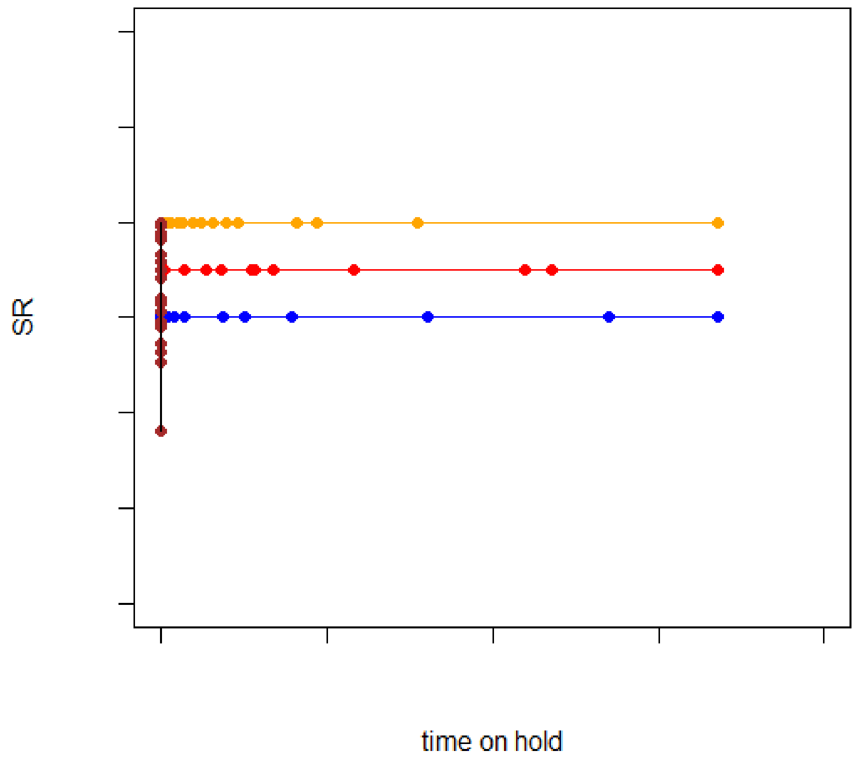
Stress Rupture Data



Stress Rupture Data with Ramp Failures

- Statisticians on team suggest a conditional analysis using only data that makes it to stress rupture hold.
- All ramp failures are dropped.
- In the literature, a procedure termed “effective time” has been suggested that incorporate ramp data as stress rupture data
- Recent work has shown the errors and limitations of this “effective time” approach

Stress Ratio versus Time on Hold



Note: X-Axis is ***Time on Hold***

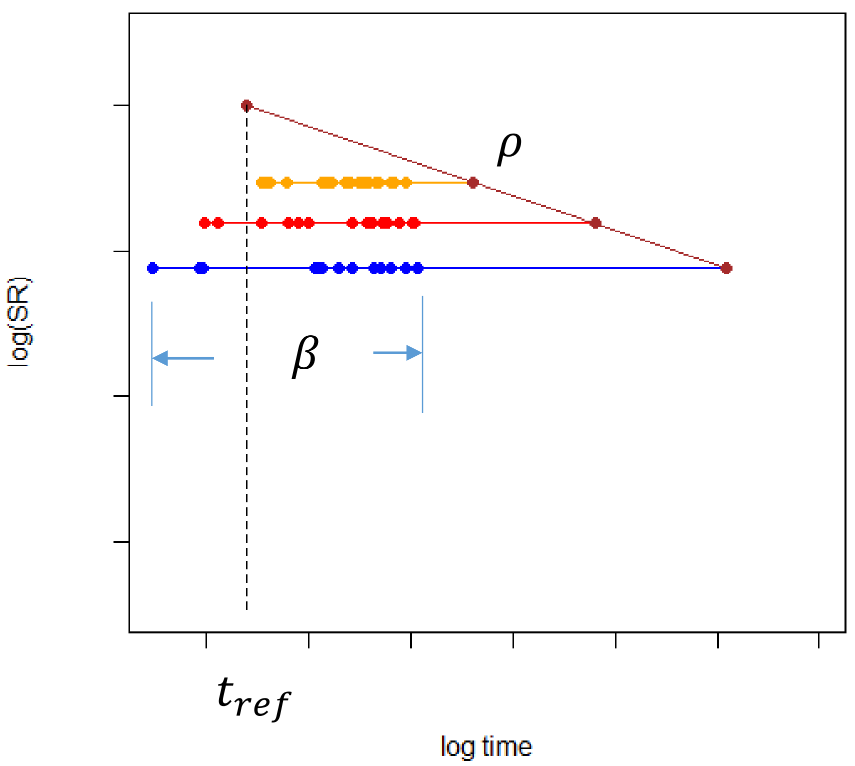
Failures on Ramp Have No Time On Hold!

True Structure of the Stress Rupture Model

$$\mu = \log(\eta) = \theta - \rho \ln(SR)$$

- ρ: controls the relationship between the failure time and stress ratio (SR)
- β: Shape parameter for time to Failure
- t_{ref}: Reference time to Failure

Log Stress Ratio versus Log Time



Stress Rupture model explains the behavior of the items *on hold*.

- Weibull regression gives us estimates for ρ, β and t_{ref}

Weibull Regression: Residuals

Weibull Residuals

- The Maximum Likelihood Estimates of the Model Depends upon:

$$e_i = \log t_i - \mu_i$$

$$= \log t_i - \log t_{ref} + \rho \ln SR_i$$

- e_i Is the "Raw Residual"
- Equally Important is the "Scaled Residual," βe_i.
- These Residuals Contain All of the Relevant Information on t_{ref}, ρ, β.

Proper Basis for Constructing Probability Plots

- Estimate the Model
- Construct the Scaled Residuals
- Calculate the Median Ranks for these Residuals (Overall not by SR!)
- Plot $\ln[-\ln(1 - mr_i)]$ versus the βe_i
- Method Extends Easily to More Complicated Models

Analogs to Standard Regression Residual Plots

- Proper Standard Regression Residual Plots Use Standardized Residuals
  - Let y_i Represent the Observed Value for the i^{th} Response
  - Let $\hat{y}_i$ Represent the Predicted Value for y_i Based on the Assumed Model
- Raw Residual:

$$e_i = y_i - \hat{y}_i$$

- Let s_i Be an Appropriate Estimate of the Standard Error for e_i.

Analogs to Standard Regression Residual Plots

- Standardized Regression Residual:

$$t_i = \frac{e_i}{s_i}$$

  - Follows a true t-Distribution
  - Standard Error Properly Accounts for the Variability in the Raw Residual
  - Proper Basis for Tentative Outlier Detection: ± 3
  - Can Still Flag More Outliers than Are "Real"
  - Issue: Multiple Comparisons
- Proper Basis for Evaluating Standard Regression Model Adequacy.
- Routinely Provided by Standard Software

Analogs to Standard Regression Residual Plots

- Extension Requires a New Residual: The "Probability Residual"
- Builds Off the Standard Probability Plot
- Let $r_{p,i}$ denote the "probability plot residual" defined by

$$r_{p,i} = \ln[-\ln(1 - mr_i)] - \beta e_i$$

$$mr_i = \hat{F}(t_i) = \frac{AR_i - 0.3}{n + 0.4}$$

where AR_i is the adjusted rank and n is the sample size.

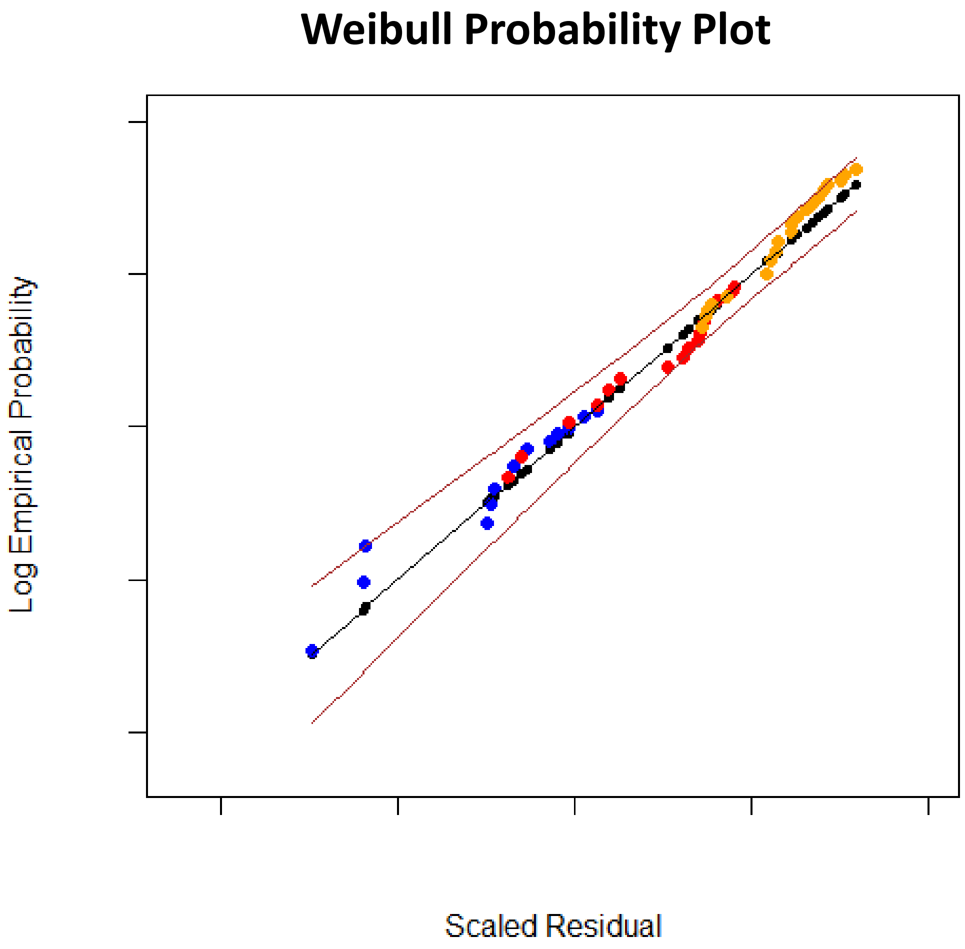
Analogs to Standard Regression Residual Plots

- Standardized Regression Residual:

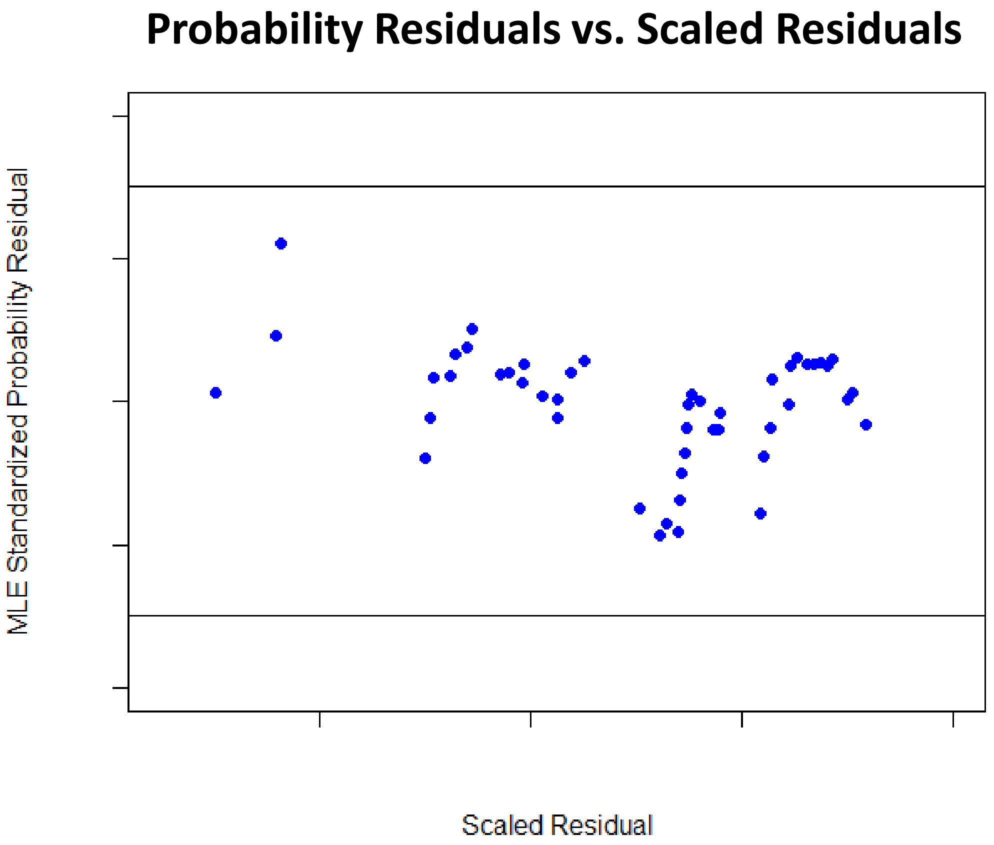
$$r_i = \frac{r_{p,i}}{s_i}$$

- Asymptotically, the r_i:
  - Follow a Standard Normal Distribution
  - Need many failures to see the Asymptotic Behavior
  - Using ± 3 as Cut-Off is Even More Prone to Indicate Outliers
- The Analog Plots Are the Best Basis for Evaluating Model Adequacy.

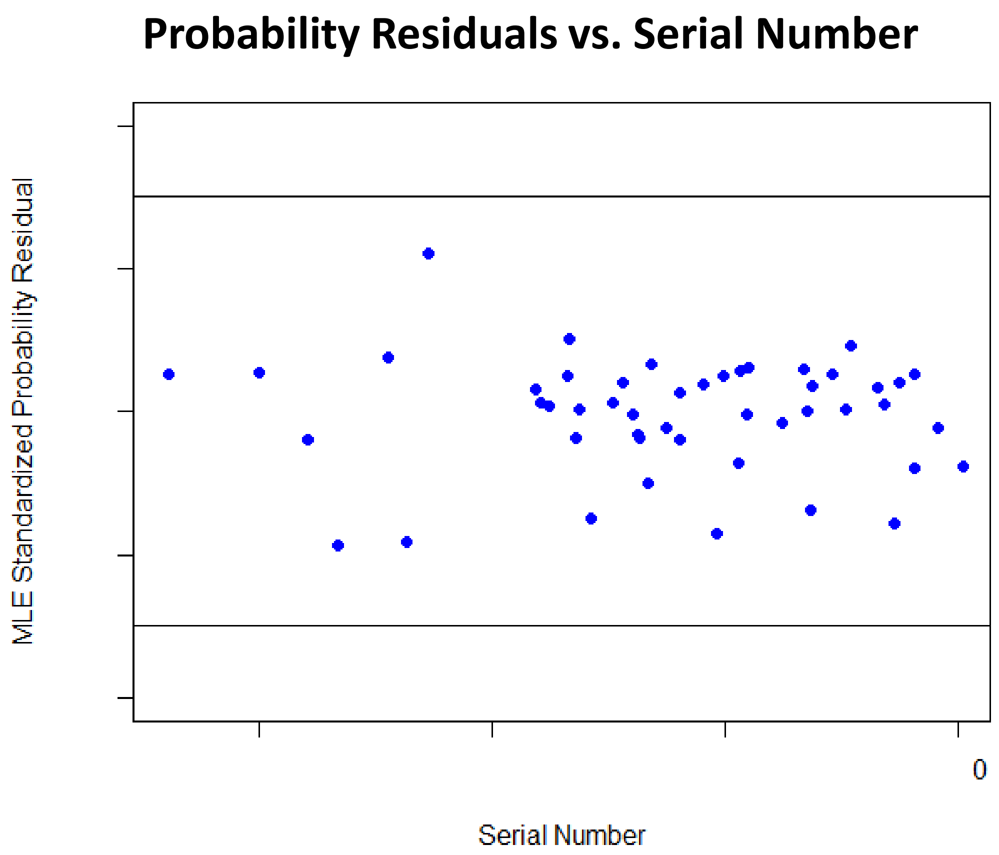
Standard Probability Plot



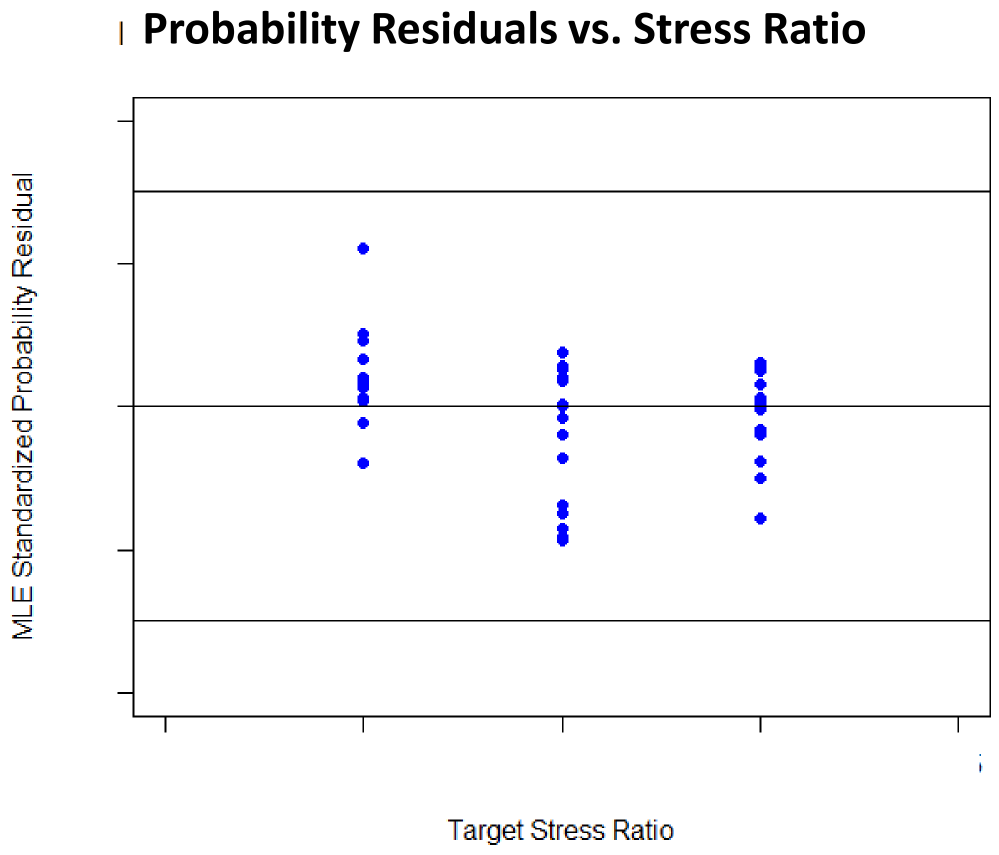
Probability Residuals versus Scaled



Probability Residuals versus Serial Number



Probability Residuals versus Stress Ratio

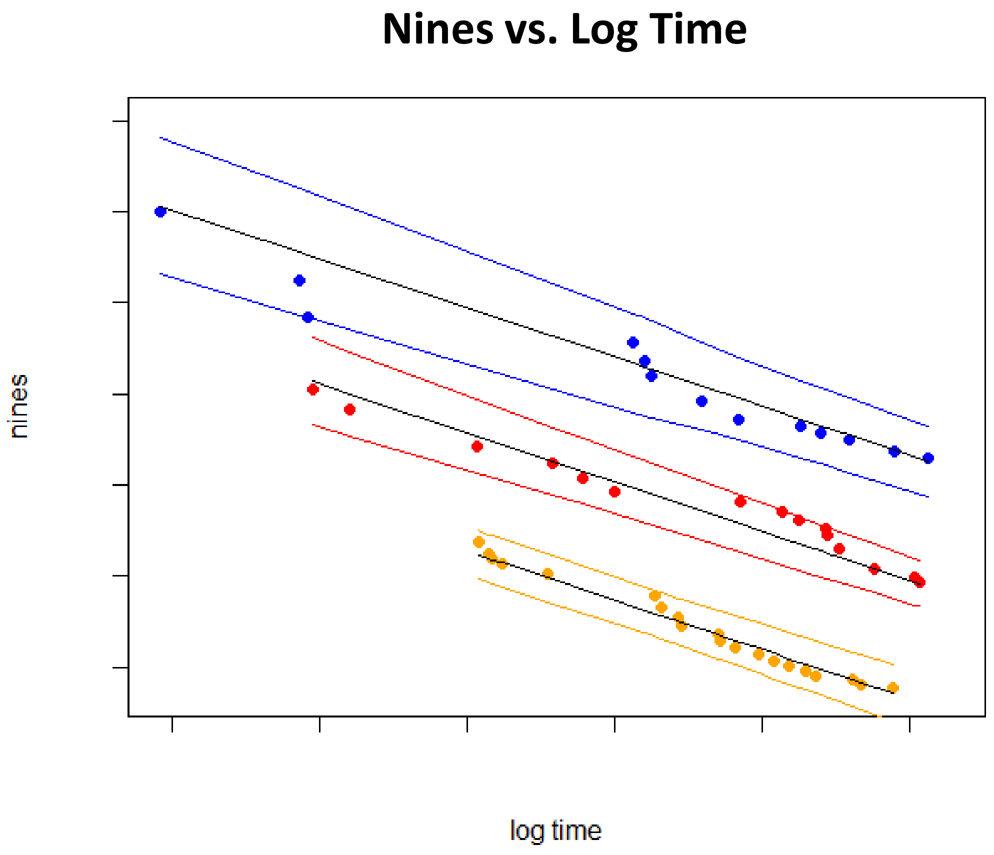


Nines in Reliability

- A statistic in reliability that is just another way of reporting the reliability of a product, system, etc.
- Literally just a count of the number of nines for reliability
- Ex: 99% reliability is 2 nines of reliability
    - Interpretation: One in 100
- 99,999% is 5 nines of reliability
    - Interpretation: One in 100,000

- Nines Calculation: $-\log_{10}\left(1 - F(t)\right)$

Nines versus Log Time



Note: Slope of Each Prediction Line Is $-\beta$

Prediction Line Is $-\beta e_i$ (Negative Scaled Residual)