Chapter 17

Failure-Time Regression Analysis

William Q. Meeker and Luis A. Escobar
Iowa State University and Louisiana State University

Based on the authors' text *Statistical Methods for Reliability Data*, John Wiley & Sons Inc. 1998.

July 18, 2002
12h 26min

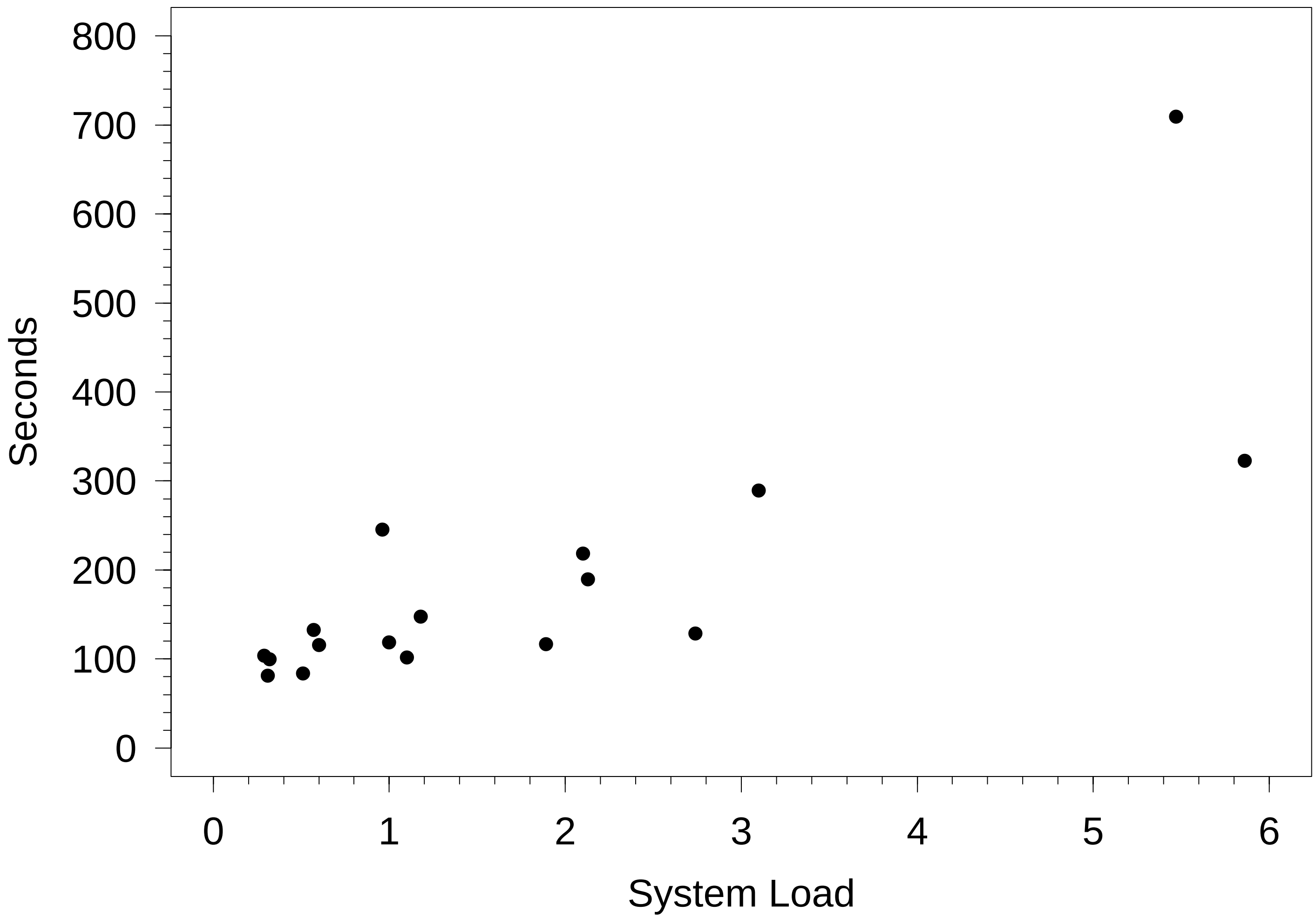
Chapter 17
Failure-Time Regression Analysis
Objectives

- Describe applications of failure-time regression
- Describe graphical methods for displaying censored regression data.
- Introduce time-scaling transformation functions.
- Describe simple regression models to relate life to explanatory variables.
- Illustrate the use of likelihood methods for censored regression data.
- Explain the importance of model diagnostics.
- Describe and illustrate extensions to nonstandard multiple regression models

Computer Program Execution Time Versus System Load

- Time to complete a computationally intensive task.
- Information from the Unix `uptime` command
- Predictions needed for scheduling subsequent steps in a multi-step computational process.

Scatter Plot of Computer Program Execution Time Versus System Load



Explanatory Variables for Failure Times

Useful explanatory variables explain/predict why some units fail quickly and some units survive a long time.

- Continuous variables like stress, temperature, voltage, and pressure.
- Discrete variables like number of hardening treatments or number of simultaneous users of a system.
- Categorical variables like manufacturer, design, and location.

Regression model relates failure time distribution to explanatory variables $\boldsymbol{x} = (x_1, \ldots, x_k)$:

$$\Pr(T \leq t) = F(t) = F(t; \boldsymbol{x}).$$

Failure-Time Regression Analysis

- Material in this chapter is an **extension** of statistical regression analysis with normal distributed data and

$$\text{mean} = \beta_0 + \beta_1 x_1 + \cdots + \beta_s x_k$$

where the x_i are explanatory variables.

- The ideas presented here are more general:
  - ▶ Data not necessarily from a normal distribution.
  - ▶ Data may be censored.
  - ▶ Nonstandard regression models that relate life to explanatory variables.
- Presentation motivated by practical problems in reliability analysis.

Nickel-Base Super-alloy Fatigue Data
26 Observations in Total, 4 Censored
(Nelson 1984, 1990)

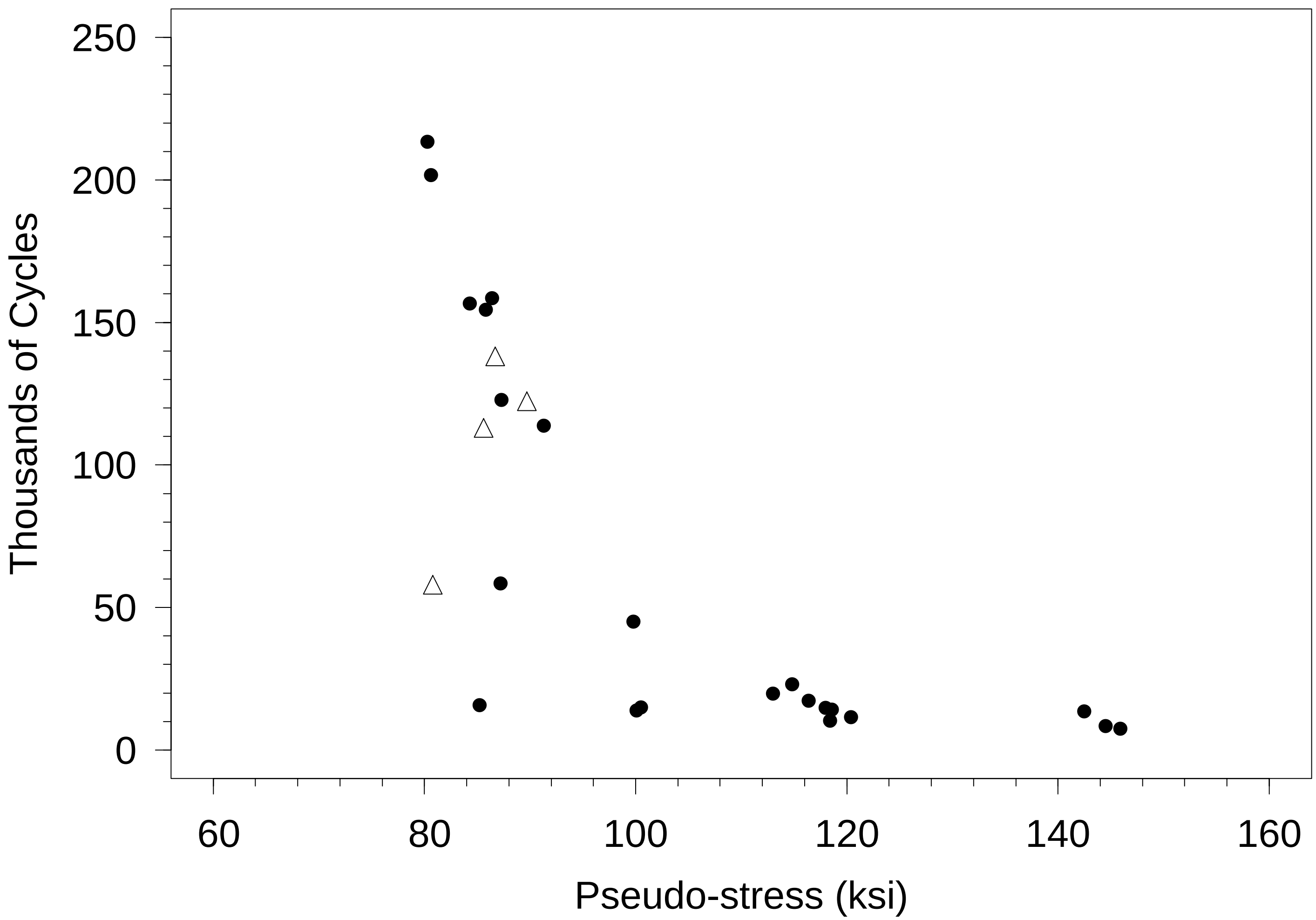
Originally described and analyzed by Nelson (1984 and 1990).

- Thousands of cycles to failure as a function of **pseudo-stress** in ksi.

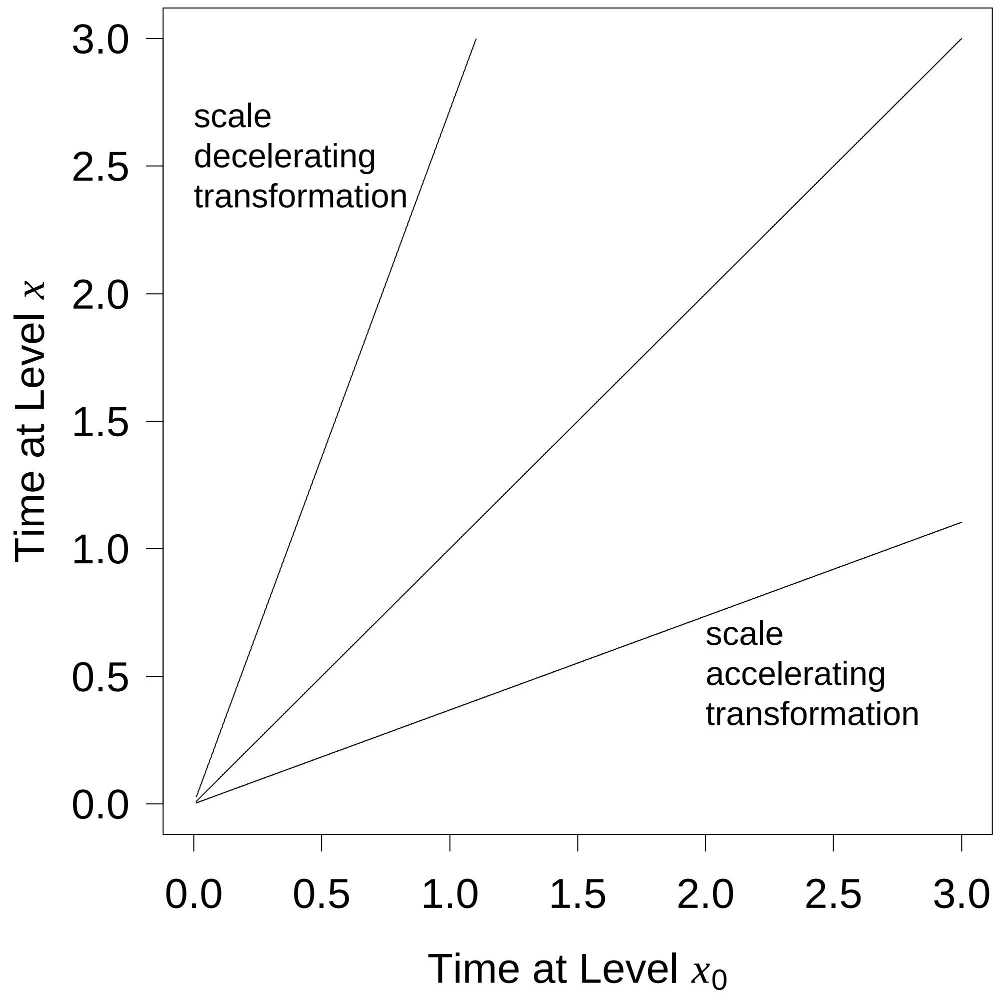
- 26 units tested; 4 units did not fail.

Objective: Explore models that might be used to describe the relationship between life length and the amount of pseudo-stress applied to the tested specimens.

Nickel-Base Super-Alloy Fatigue Data (Nelson 1984, 1990)



SAFT Models
Illustrating Acceleration and Deceleration.



Scale Accelerated Failure Time (SAFT) Model

- Scale Accelerated Failure Time (SAFT) models are defined by

$$T(\boldsymbol{x}) = T(\boldsymbol{x}_0)/\mathcal{AF}(\boldsymbol{x}), \quad \mathcal{AF}(\boldsymbol{x}) > 0, \quad \mathcal{AF}(\boldsymbol{x}_0) = 1$$

  where typical forms for $\mathcal{AF}(\boldsymbol{x})$ are:

  - ▶ $\log[\mathcal{AF}(x)] = \beta_1 x$ with $x_0 = 0$ for a scalar x.
  - ▶ $\log[\mathcal{AF}(\boldsymbol{x})] = \beta_1 x_1 + \ldots + \beta_k x_k$ with $\boldsymbol{x}_0 = \underline{0}$ for a vector $\boldsymbol{x} = (x_1, \ldots, x_k)$.

- When $\mathcal{AF}(\boldsymbol{x}) > 1$, the model accelerates time in the sense that time moves more quickly at $\boldsymbol{x}$ than at $\boldsymbol{x}_0$ so that $T(\boldsymbol{x}) < T(\boldsymbol{x}_0)$.

- When $0 < \mathcal{AF}(\boldsymbol{x}) < 1$, $T(\boldsymbol{x}) > T(\boldsymbol{x}_0)$, and time is decelerated (but we still call this an SAFT model).

The SAFT Transformation Models and Acceleration

In a time transformation plot of $T(\boldsymbol{x}_0)$ vs $T(\boldsymbol{x})$

- The **SAFT** models are straight lines **through** the origin:

- Accelerating SAFT models are straight lines **below** the diagonal.

- Decelerating SAFT models are straight lines **above** the diagonal.

Some Properties of SAFT Models

For a **SAFT** model $T(\boldsymbol{x}) = T(\boldsymbol{x}_0)/\mathcal{AF}(\boldsymbol{x})$, $(\Psi(\boldsymbol{x}) > 0)$, with baseline cdf $F(t;\boldsymbol{x}_0)$ so $\mathcal{AF}(\boldsymbol{x}_0) = 1$

- **Scaled time:** $F(t;\boldsymbol{x}) = F\left[\mathcal{AF}(\boldsymbol{x})\, t;\boldsymbol{x}_0\right]$. Thus the cdfs $F(t;\boldsymbol{x})$ and $F(t;\boldsymbol{x}_0)$ do **not** cross each other.

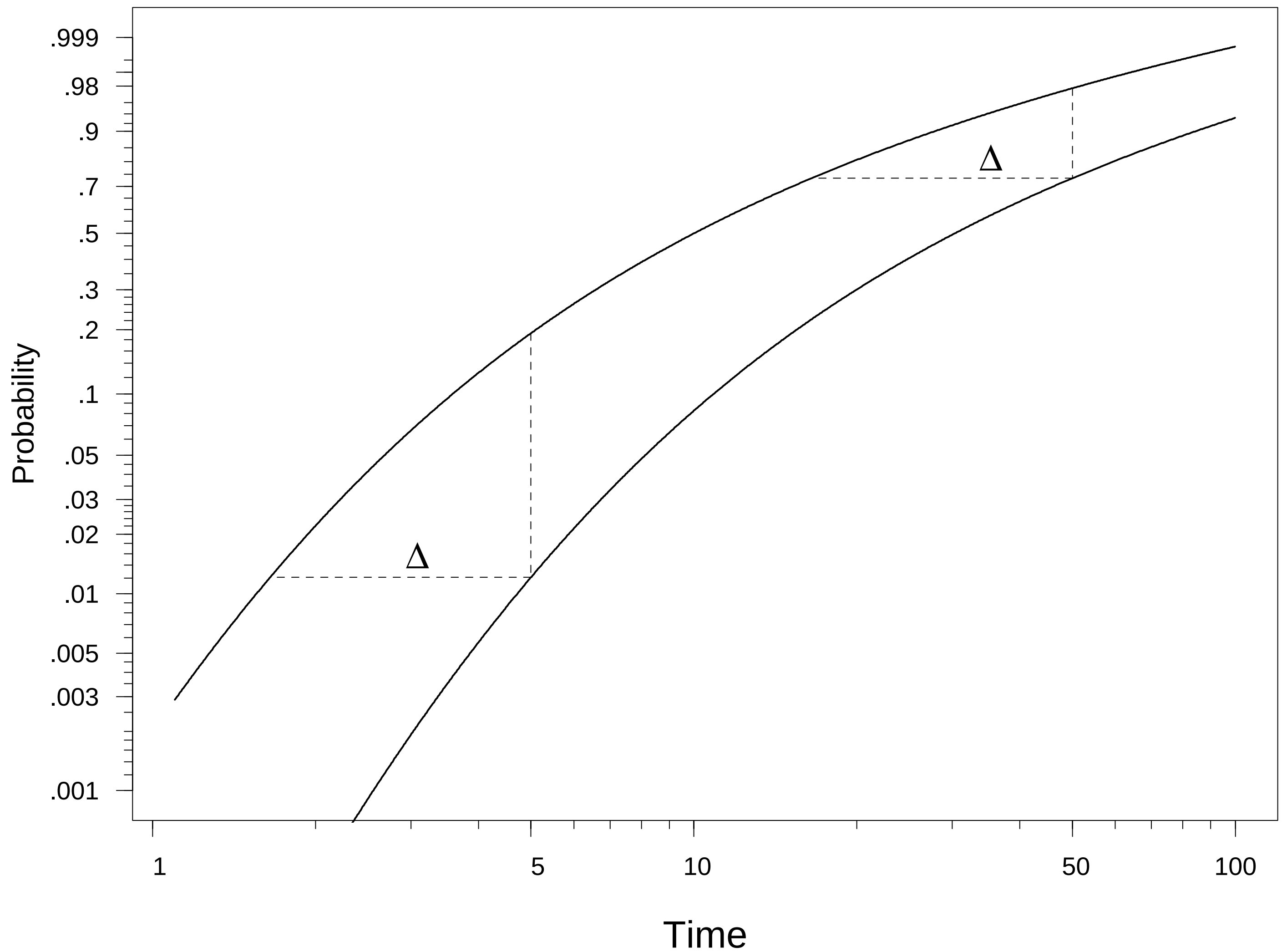
- **Proportional quantiles:** $t_p(\boldsymbol{x}) = t_p(\boldsymbol{x}_0)/\mathcal{AF}(\boldsymbol{x})$. Then taking logs gives

$$\log[t_p(\boldsymbol{x}_0)] - \log[t_p(\boldsymbol{x})] = \log[\mathcal{AF}(\boldsymbol{x})].$$

This shows that in any plot with a log-time scale $t_p(\boldsymbol{x}_0)$ and $t_p(\boldsymbol{x})$ are **equidistant**.

In particular, in a probability plot with a log-time scale, $F(t,\boldsymbol{x})$ is a translation of $F(t,\boldsymbol{x}_0)$ along the $\log(t)$ axis.

Weibull Probability Plot of Two Members from an SAFT Model with a Baseline Lognormal Distribution



Lognormal Distribution Simple Regression Model with Constant Shape Parameter $\beta = 1/\sigma$

- The lognormal simple regression model is

$$\Pr(T \le t) = F(t; \mu, \sigma) = F(t; \beta_0, \beta_1, \sigma) = \Phi_{\text{nor}}\left[\frac{\log(t) - \mu}{\sigma}\right]$$

where $\mu = \mu(x) = \beta_0 + \beta_1 x$ and σ does not depend on x.

- The failure-time log quantile function

$$\log[t_p(x)] = \mu(x) + \Phi_{\text{nor}}^{-1}(p)\sigma$$

is linear in x.

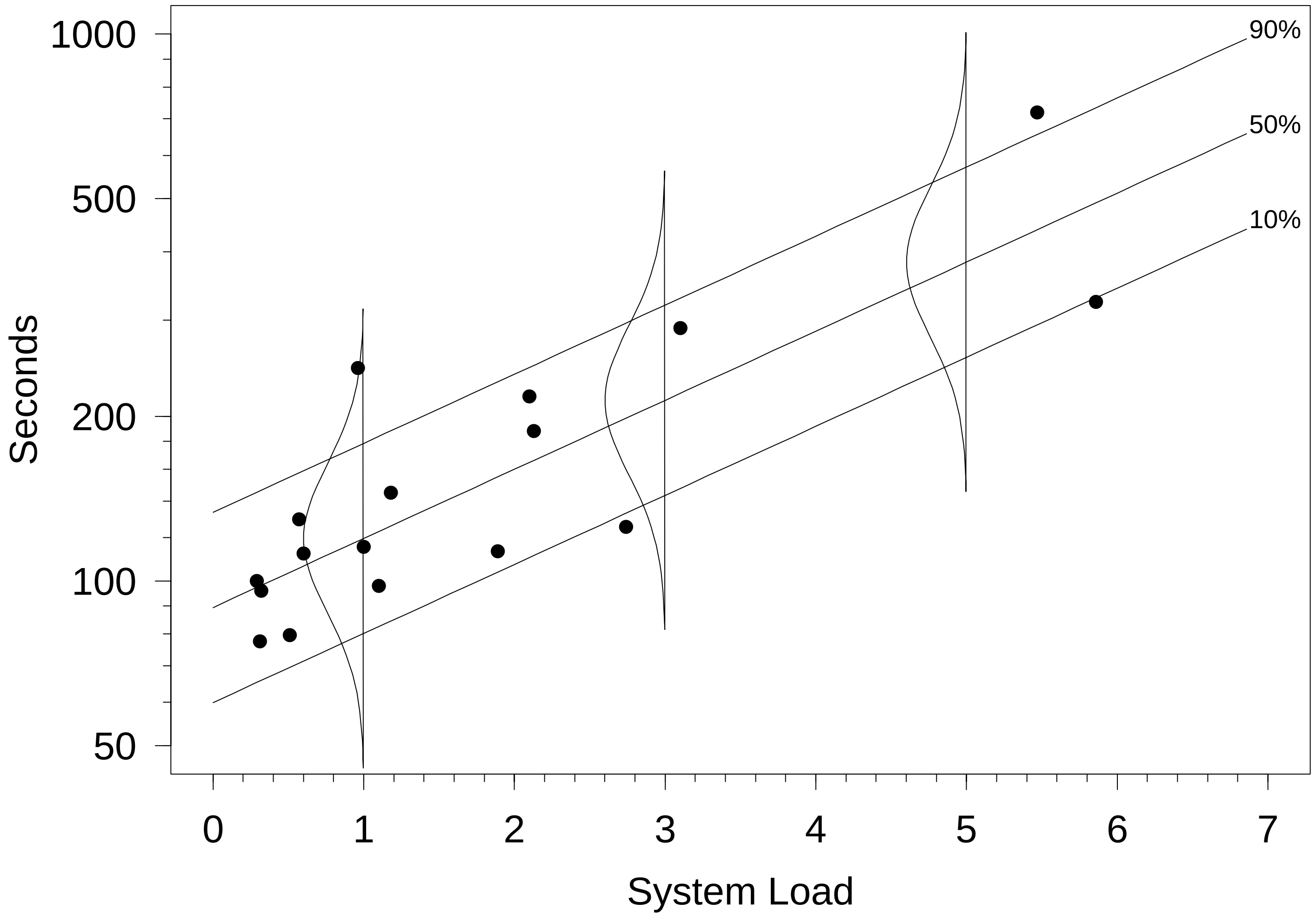
Notice that

$$\frac{t_p(x)}{t_p(0)} = \exp(\beta_1 x)$$

implies that this regression model is a scale accelerated failure time (SAFT) model with $\mathcal{AF}(x) = \exp(-\beta_1 x)$.

Computer Program Execution Time Versus System Load Loglinear Lognormal Regression Model

$$\log[\widehat{t}_p(x)] = \widehat{\mu}(x) + \Phi^{-1}_{\text{nor}}(p)\widehat{\sigma}$$



Likelihood for Lognormal Distribution Simple Regression Model with Right Censored Data

The likelihood for n independent observations has the form

$$
\begin{aligned}
L(\beta_0, \beta_1, \sigma) &= \prod_{i=1}^{n} L_i(\beta_0, \beta_1, \sigma; \mathsf{data}_i) \\
&= \prod_{i=1}^{n} \left\{ \frac{1}{\sigma t_i} \phi_{\mathsf{nor}} \left[\frac{\log(t_i) - \mu_i}{\sigma} \right] \right\}^{\delta_i} \left\{ 1 - \Phi_{\mathsf{nor}} \left[\frac{\log(t_i) - \mu_i}{\sigma} \right] \right\}^{1-\delta_i}
\end{aligned}
$$

where $\mathsf{data}_i = (x_i, t_i, \delta_i)$, $\mu_i = \beta_0 + \beta_1 x_i$,

$$
\delta_i = \begin{cases} 1 & \text{exact observation} \\ 0 & \text{right censored observation} \end{cases}
$$

$\phi_{\mathsf{nor}}(z)$ is the standardized normal pdf and $\Phi_{\mathsf{nor}}(z)$ is the corresponding normal cdf.

The parameters are $\boldsymbol{\theta} = (\beta_0, \beta_1, \sigma)$.

Estimated Parameter Variance-Covariance Matrix

Local (observed information) estimate

$$\hat{\Sigma}_{\widehat{\boldsymbol{\theta}}} = \begin{bmatrix} \widehat{\text{Var}}(\widehat{\beta}_0) & \widehat{\text{Cov}}(\widehat{\beta}_0, \widehat{\beta}_1) & \widehat{\text{Cov}}(\widehat{\beta}_0, \widehat{\sigma}) \\ \widehat{\text{Cov}}(\widehat{\beta}_1, \widehat{\beta}_0) & \widehat{\text{Var}}(\widehat{\beta}_1) & \widehat{\text{Cov}}(\widehat{\beta}_1, \widehat{\sigma}) \\ \widehat{\text{Cov}}(\widehat{\sigma}, \widehat{\beta}_0) & \widehat{\text{Cov}}(\widehat{\sigma}, \widehat{\beta}_1) & \widehat{\text{Var}}(\widehat{\sigma}) \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{\partial^2 \mathcal{L}(\beta_0,\beta_1,\sigma)}{\partial \beta_0^2} & -\frac{\partial^2 \mathcal{L}(\beta_0,\beta_1,\sigma)}{\partial \beta_0 \partial \beta_1} & -\frac{\partial^2 \mathcal{L}(\beta_0,\beta_1,\sigma)}{\partial \beta_0 \partial \sigma} \\ -\frac{\partial^2 \mathcal{L}(\beta_0,\beta_1,\sigma)}{\partial \beta_1 \partial \beta_0} & -\frac{\partial^2 \mathcal{L}(\beta_0,\beta_1,\sigma)}{\partial \beta_1^2} & -\frac{\partial^2 \mathcal{L}(\beta_0,\beta_1,\sigma)}{\partial \beta_1 \partial \sigma} \\ -\frac{\partial^2 \mathcal{L}(\beta_0,\beta_1,\sigma)}{\partial \sigma \partial \beta_0} & -\frac{\partial^2 \mathcal{L}(\beta_0,\beta_1,\sigma)}{\partial \sigma \partial \beta_1} & -\frac{\partial^2 \mathcal{L}(\beta_0,\beta_1,\sigma)}{\partial \sigma^2} \end{bmatrix}^{-1}$$

Partial derivatives are evaluated at $\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\sigma}$.

Standard Errors and Confidence Intervals for Parameters

- Lognormal ML estimates for the computer time experiment were $\widehat{\boldsymbol{\theta}} = (\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\sigma}) = (4.49, .290, .312)$ and an estimate of the variance-covariance matrix for $\widehat{\boldsymbol{\theta}}$ is

$$\widehat{\Sigma}_{\widehat{\boldsymbol{\theta}}} = \begin{bmatrix} .012 & -.0037 & 0 \\ -.0037 & .0021 & 0 \\ 0 & 0 & .0029 \end{bmatrix}.$$

- Normal-approximation confidence interval for the computer execution time regression slope is

$$[\underset{\sim}{\beta_1}, \quad \tilde{\beta}_1] = \widehat{\beta}_1 \pm z_{(.975)} \widehat{\text{se}}_{\widehat{\beta}_1} = .290 \pm 1.96(.046) = [.20, \quad .38]$$

where $\widehat{\text{se}}_{\widehat{\beta}_1} = \sqrt{.0021} = .046$.

Standard Errors and Confidence Intervals for Quantities at Specific Explanatory Variable Conditions

- Unknown values of μ and σ at each level of x
- $\widehat{\mu} = \widehat{\beta}_0 + \widehat{\beta}_1 x$, σ does not depend on x, and

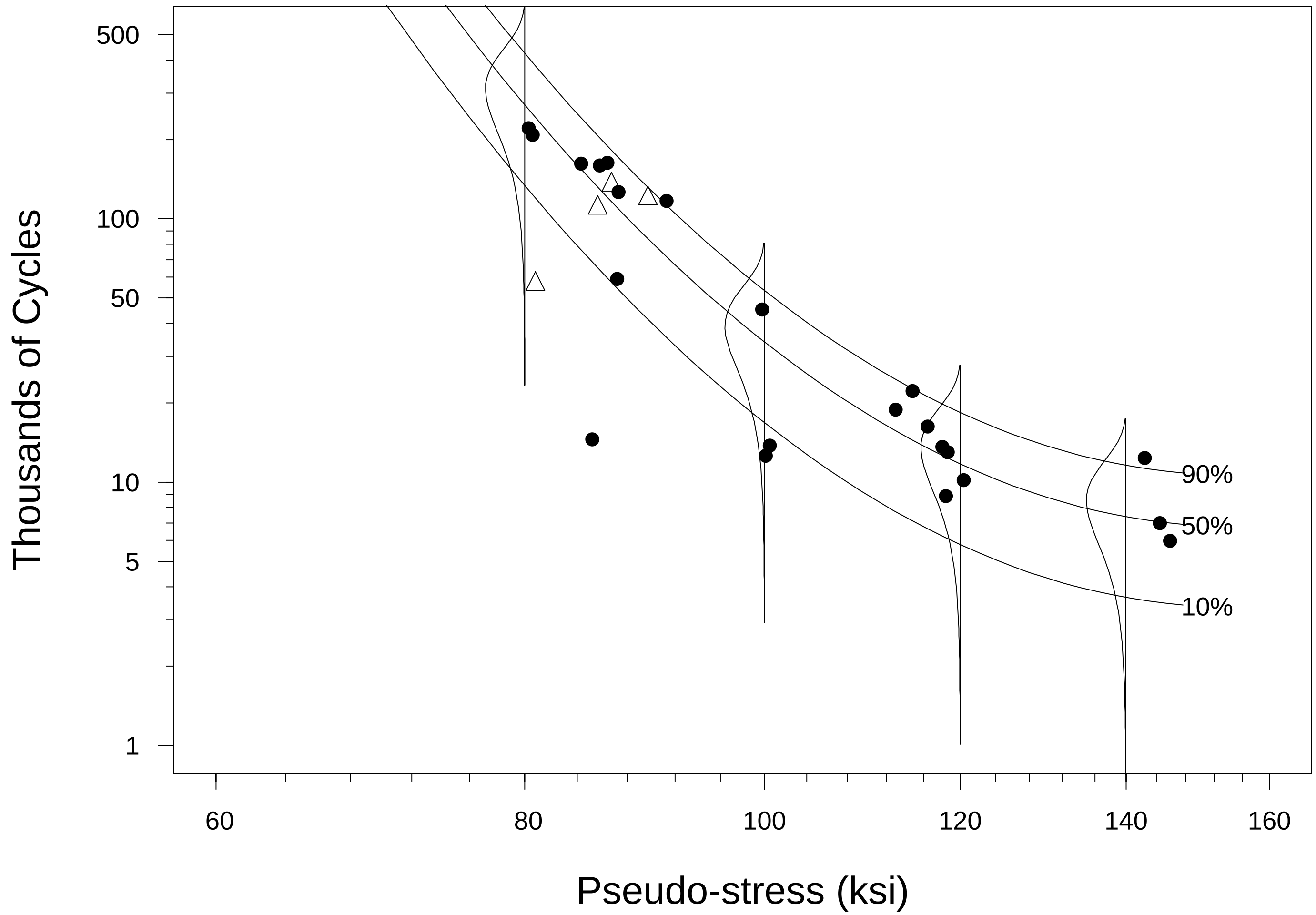
$$\widehat{\Sigma}_{\widehat{\mu},\widehat{\sigma}} = \begin{bmatrix} \widehat{\mathrm{Var}}(\widehat{\mu}) & \widehat{\mathrm{Cov}}(\widehat{\mu},\widehat{\sigma}) \\ \widehat{\mathrm{Cov}}(\widehat{\mu},\widehat{\sigma}) & \widehat{\mathrm{Var}}(\widehat{\sigma}) \end{bmatrix}$$

  is obtained from $\widehat{\mathrm{Var}}(\widehat{\mu}) = \widehat{\mathrm{Var}}(\widehat{\beta}_0) + 2x\widehat{\mathrm{Cov}}(\widehat{\beta}_1,\widehat{\beta}_0) + x^2\widehat{\mathrm{Var}}(\widehat{\beta}_1)$ and $\widehat{\mathrm{Cov}}(\widehat{\mu},\widehat{\sigma}) = \widehat{\mathrm{Cov}}(\widehat{\beta}_0,\widehat{\sigma}) + x\widehat{\mathrm{Cov}}(\widehat{\beta}_1,\widehat{\sigma})$.

- Use the above results with the methods from Chapter 8 to compute normal-approximation confidence intervals for $F(t)$, $h(t)$, and t_p.
- Could also use likelihood or simulation-based confidence intervals.

Log-Quadratic Weibull Regression Model with Constant ($\beta = 1/\sigma$) Fit to the Fatigue Data

$$\log[\widehat{t}_p(x)] = \widehat{\mu}(x) + \Phi_{\text{sev}}^{-1}(p)\widehat{\sigma},\ x = \log(\textbf{pseudo-stress})$$



Weibull Distribution Quadratic Regression Model with Constant Shape Parameter $\beta = 1/\sigma$

This is an AF time model with the following characteristics:

- The Weibull quadratic regression model is

$$\text{Pr}[T \le t] = \Phi_{\text{sev}}\left\{[\log(t) - \mu]/\sigma\right\}$$

where $\mu = \mu(x) = \beta_0 + \beta_1 x + \beta_2 x^2$ and σ does not depend on x.

- The failure-time log quantile function

$$\log[t_p(x)] = \mu(x) + \Phi_{\text{sev}}^{-1}(p)\sigma$$

is quadratic in x. Also

$$t_p(x) = \exp(\beta_1 x + \beta_2 x^2) \times t_p(0)$$

which shows that the model is an SAFT model.

Likelihood for Weibull Distribution Quadratic Regression Model with Right Censored Data

The likelihood for n independent observations has the form

$$
\begin{aligned}
L(\beta_0, \beta_1, \beta_2, \sigma) &= \prod_{i=1}^{n} L_i(\beta_0, \beta_1, \beta_2, \sigma; \text{data}_i) \\
&= \prod_{i=1}^{n} \left\{ \frac{1}{\sigma t_i} \phi_{\text{sev}} \left[\frac{\log(t_i) - \mu_i}{\sigma} \right] \right\}^{\delta_i} \left\{ 1 - \Phi_{\text{sev}} \left[\frac{\log(t_i) - \mu_i}{\sigma} \right] \right\}^{1-\delta_i}.
\end{aligned}
$$

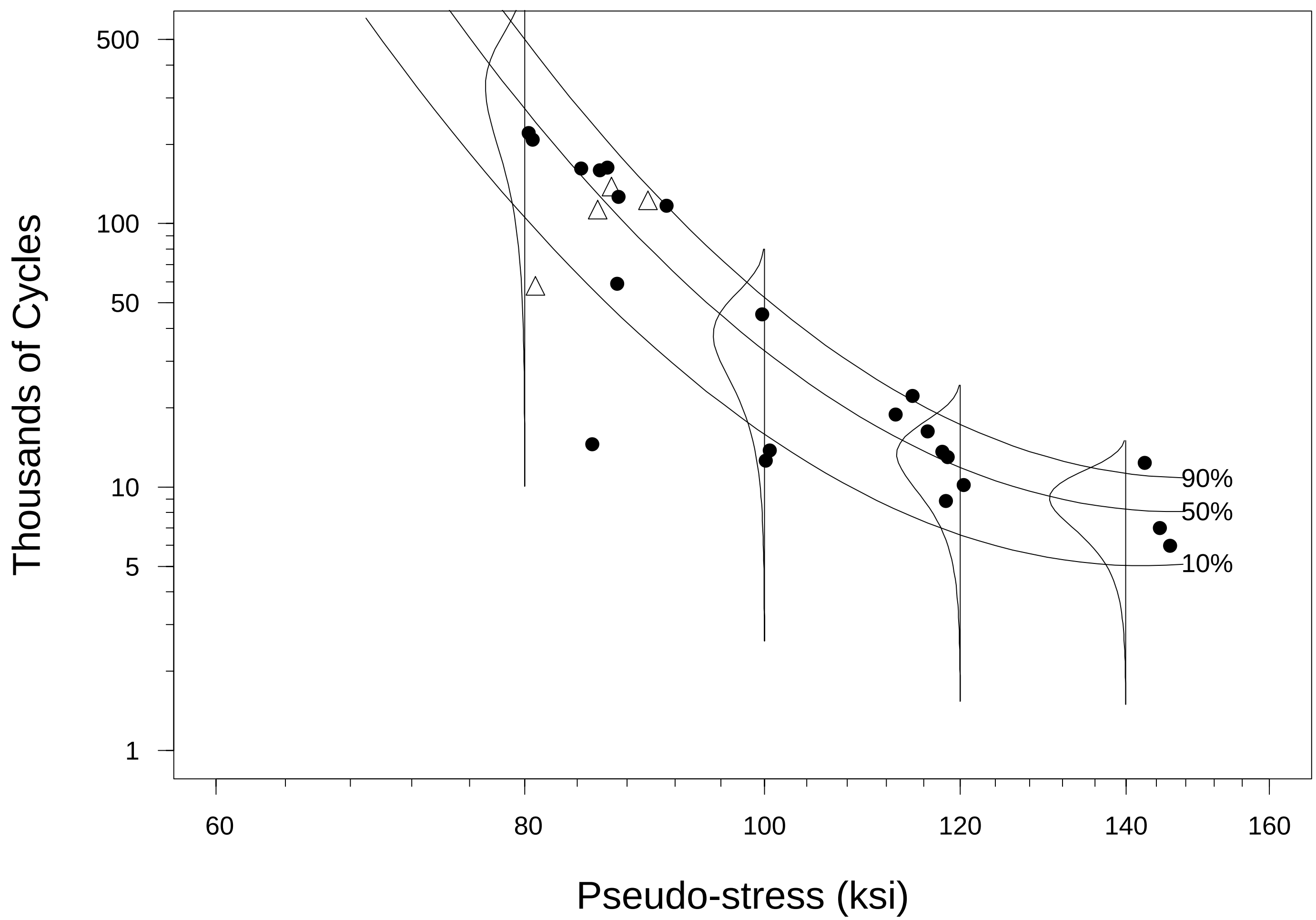
where $\mu_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$,

$$
\delta_i = \begin{cases} 1 & \text{exact observation} \\ 0 & \text{right censored observation} \end{cases}
$$

The parameters are $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, \sigma)$.

Log-quadratic Weibull Regression Model with Nonconstant $\beta = 1/\sigma$ Fit to the Fatigue Data

$$\log[\widehat{t}_p(x)] = \widehat{\mu}(x) + \Phi^{-1}_{\text{sev}}(p)\widehat{\sigma}(x)$$



Weibull Distribution Quadratic Regression Model with Nonconstant $\beta = 1/\sigma$

- The Weibull quadratic regression model is

$$\Pr[T \le t] = \Phi_{\text{sev}}\left\{[\log(t) - \mu]/\sigma\right\},$$

where $\mu = \mu(x) = \beta_0^{[\mu]} + \beta_1^{[\mu]}x + \beta_2^{[\mu]}x^2$ and $\log(\sigma) = \log[\sigma(x)] = \beta_0^{[\sigma]} + \beta_1^{[\sigma]}x$.

- The failure-time log quantile function is

$$\log[t_p(x)] = \mu(x) + \Phi_{\text{sev}}^{-1}(p)\sigma(x)$$

which is **not** quadratic in x.
Also

$$t_p(x) = \exp\left[\mu(x) - \mu(0)\right]\exp\left[\Phi_{\text{sev}}^{-1}(p)\left\{\sigma(x) - \sigma(0)\right\}\right] \times t_p(0).$$

Because the coefficient in front of $t_p(0)$ depends on p this shows that the model is **not** an SAFT model.

Likelihood for Weibull Distribution Quadratic Regression Model with Nonconstant $\beta = 1/\sigma$ and Right Censored Data

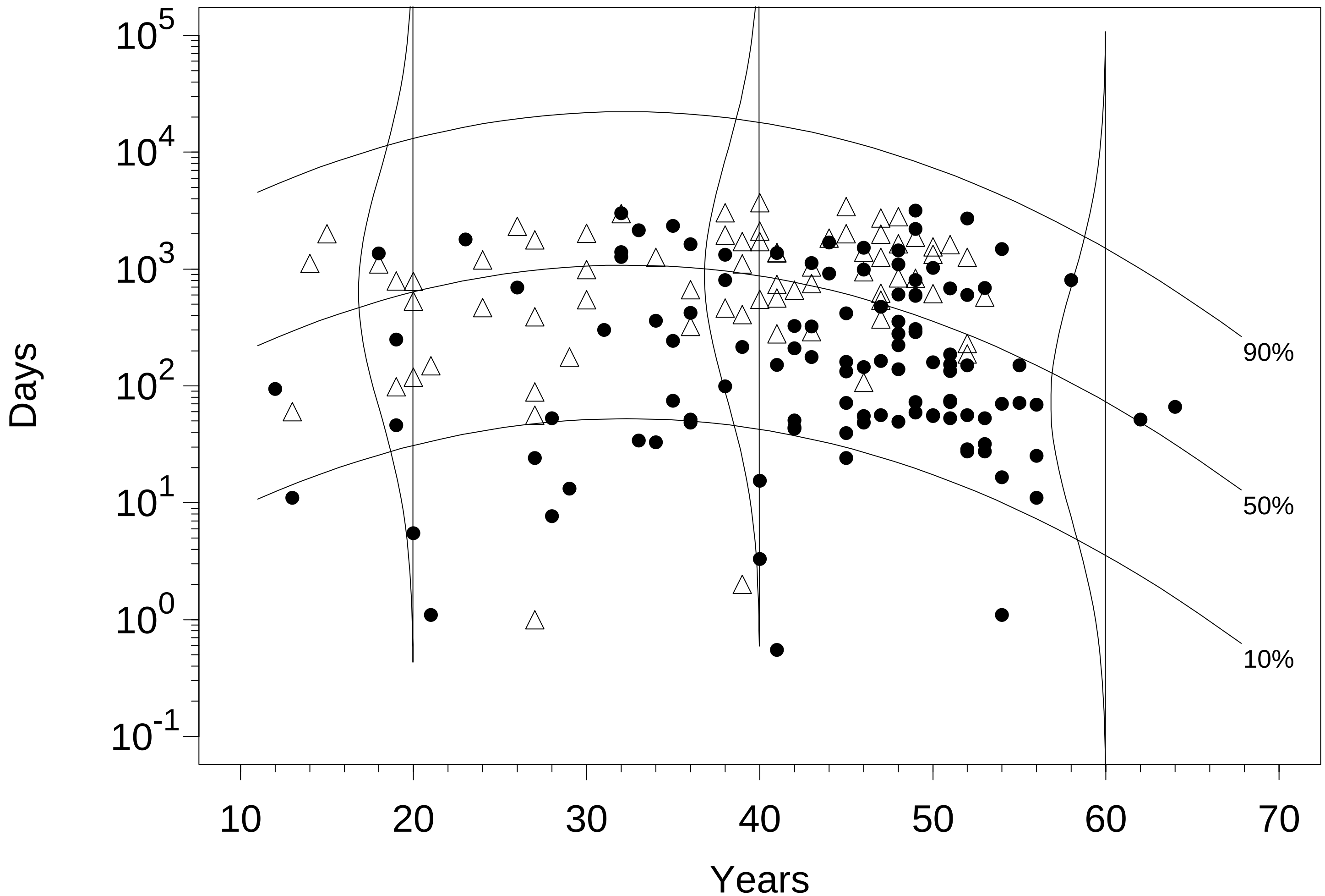
The likelihood for n independent observations has the form

$$
\begin{aligned}
& L(\beta_0^{[\mu]}, \beta_1^{[\mu]}, \beta_2^{[\mu]}, \beta_0^{[\sigma]}, \beta_1^{[\sigma]}) \\
= & \prod_{i=1}^{n} L_i(\beta_0^{[\mu]}, \beta_1^{[\mu]}, \beta_2^{[\mu]}, \beta_0^{[\sigma]}, \beta_1^{[\sigma]}; \text{data}_i) \\
= & \prod_{i=1}^{n} \left\{ \frac{1}{\sigma_i t_i} \phi_{\text{sev}} \left[\frac{\log(t_i) - \mu_i}{\sigma_i} \right] \right\}^{\delta_i} \left\{ 1 - \Phi_{\text{sev}} \left[\frac{\log(t_i) - \mu_i}{\sigma_i} \right] \right\}^{1-\delta_i}.
\end{aligned}
$$

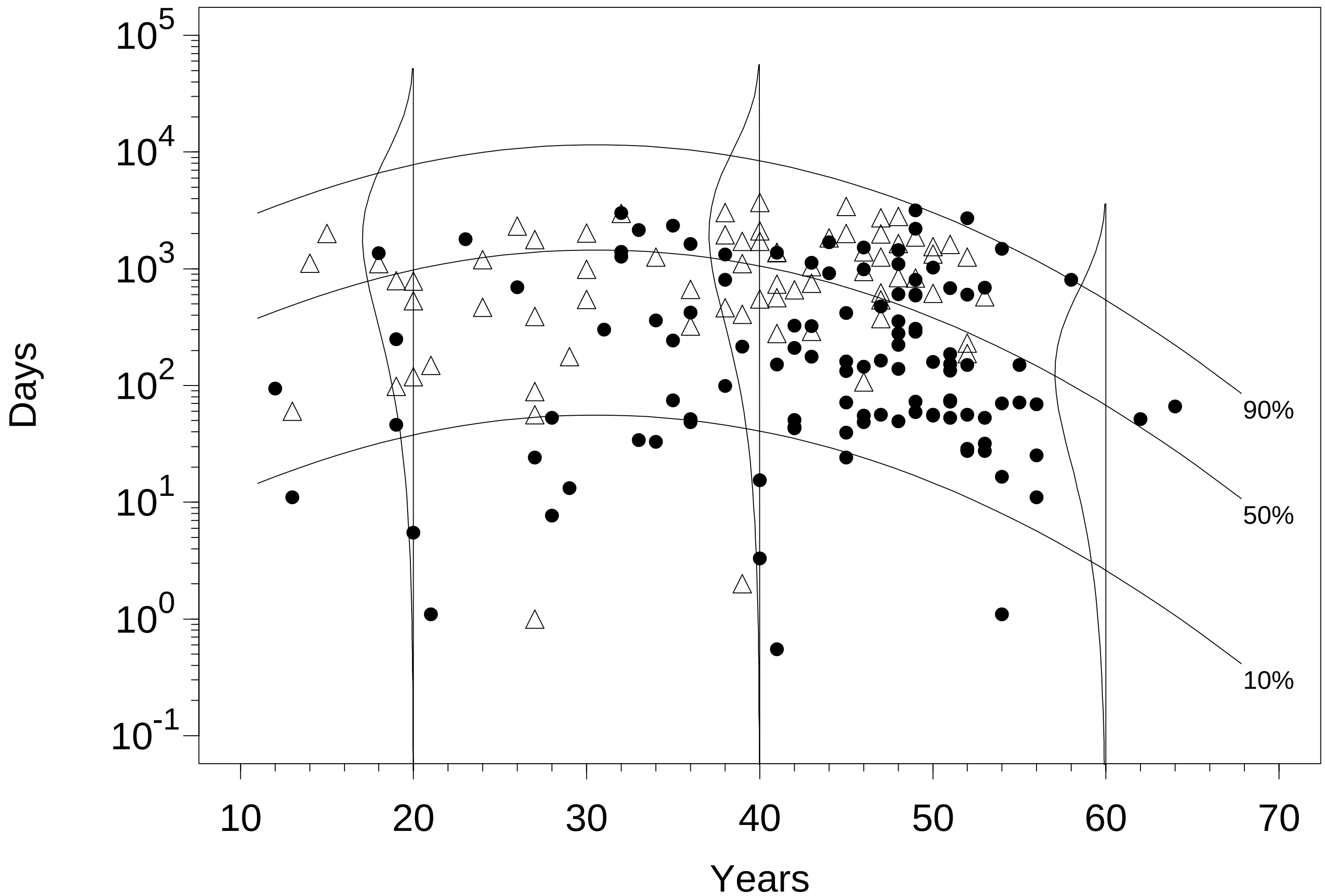
where $\mu_i = \beta_0^{[\mu]} + \beta_1^{[\mu]} x_i + \beta_2^{[\mu]} x_i^2$ and $\sigma_i = \exp\left(\beta_0^{[\sigma]} + \beta_1^{[\sigma]} x_i\right)$.

Parameters are $\boldsymbol{\theta} = (\beta_0^{[\mu]}, \beta_1^{[\mu]}, \beta_2^{[\mu]}, \beta_0^{[\sigma]}, \beta_1^{[\sigma]})$.

Stanford Heart Transplant Data
Quadratic Log-Mean Lognormal Regression Model



Stanford Heart Transplant Data
Quadratic Log-Location Weibull Regression Model



Extrapolation and Empirical Models

- Empirical models can be useful, providing a smooth curve to describe a population or a process.
- Should not be used to extrapolate outside of the range of one's data.
- There are many different kinds of extrapolation
  - In an explanatory variable like stress
  - To the upper tail of a distribution
  - To the lower tail of a distribution
- Need to get the right curve to extrapolate: look toward physical or other process theory

Checking Model Assumptions

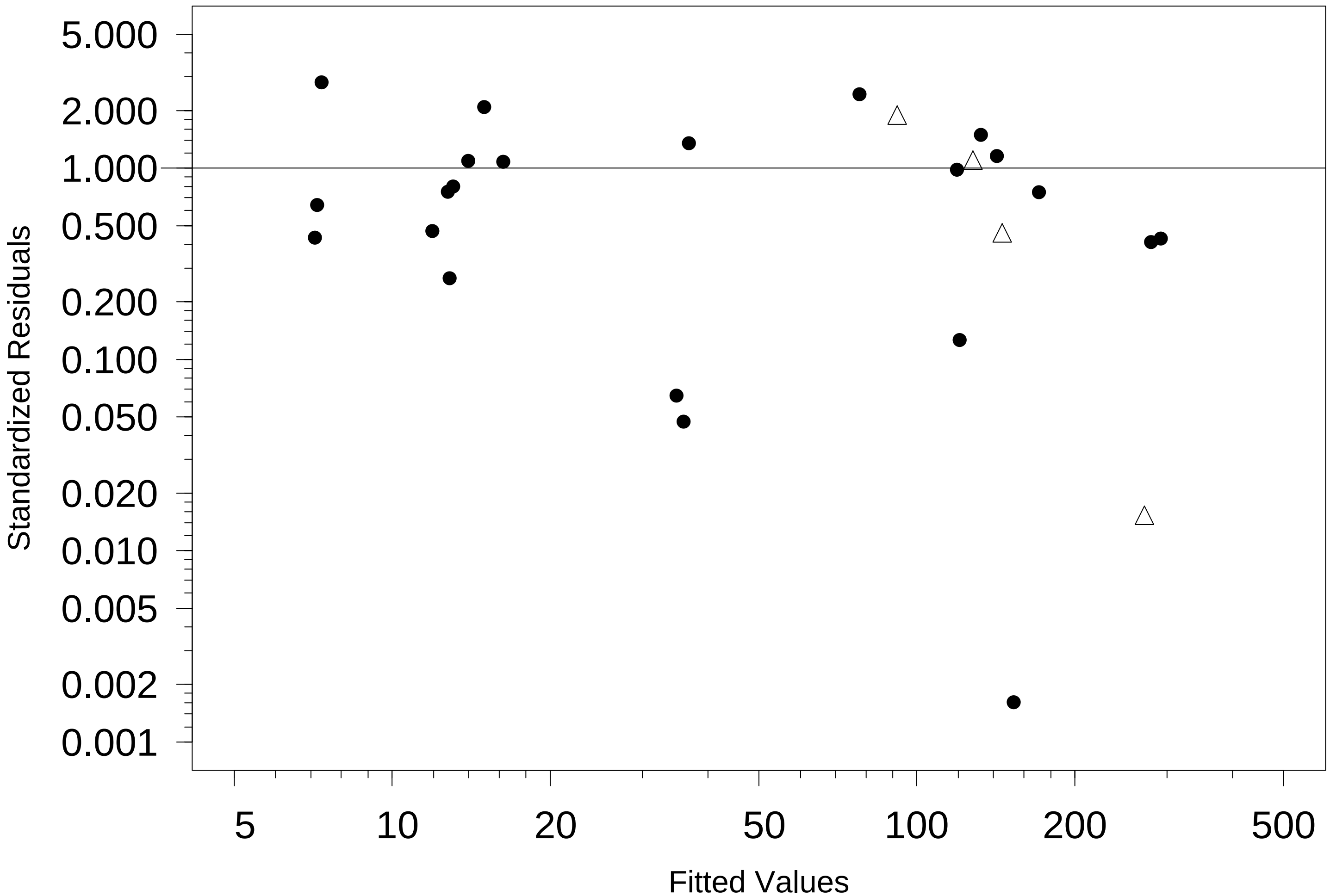
- Graphical checks using generalizations of usual diagnostics (including residual analysis)
  - ▶ Residuals versus fitted values.
  - ▶ Probability plot of residuals.
  - ▶ Other residual plots.
  - ▶ Influence analysis (or sensitivity) analysis. (see Escobar and Meeker 1992 Biometrics).
- Most analytical tests can be suitably generalized, at least approximately, for censored data (especially using likelihood ratio tests).

Definition of Residuals

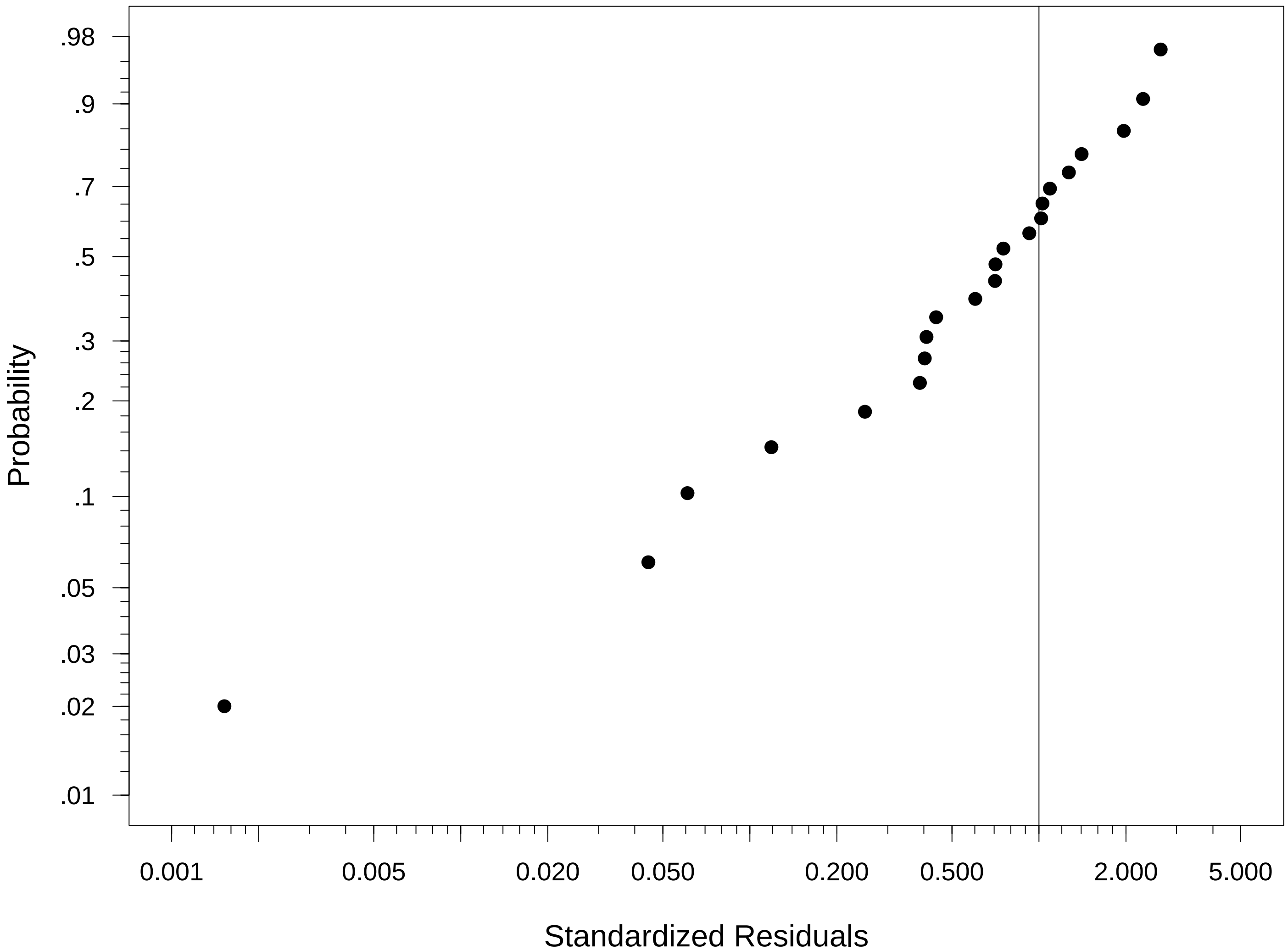
- For location-scale distributions like the normal, logistic, and smallest extreme value,
$$\widehat{\epsilon}_i = \frac{y_i - \widehat{y}_i}{\widehat{\sigma}}$$
where $\widehat{y}_i$ is an appropriately defined fitted value (e.g., $\widehat{y}_i = \widehat{\mu}_i$).

- With models for positive random variables like Weibull, lognormal, and loglogistic, standardized residuals are defined as
$$\exp(\widehat{\epsilon}_i) = \exp\left[\frac{\log(t_i) - \log(\widehat{t}_i)}{\widehat{\sigma}}\right] = \left(\frac{t_i}{\widehat{t}_i}\right)^{\frac{1}{\widehat{\sigma}}}$$
where $\widehat{t}_i = \exp(\widehat{\mu}_i)$ and when t_i is a censored observation, the corresponding residual is also censored.

Plot of Standardized Residuals Versus Fitted Values for the Log-Quadratic Weibull Regression Model Fit to the Super Alloy Data on Log-Log Axes



Probability Plot of the Standardized Residuals from the Log-Quadratic Weibull Regression Model Fit to the Super Alloy Data



Empirical Regression Models and Sensitivity Analysis Objectives

- Describe a class of regression models that can be used to describe the relationship between failure time and explanatory variables.

- Describe and illustrate the use of empirical regression models.

- Illustrate the need for sensitivity analysis.

- Show how to conduct a sensitivity analysis.

Picciotto Data
(Picciotto 1970)

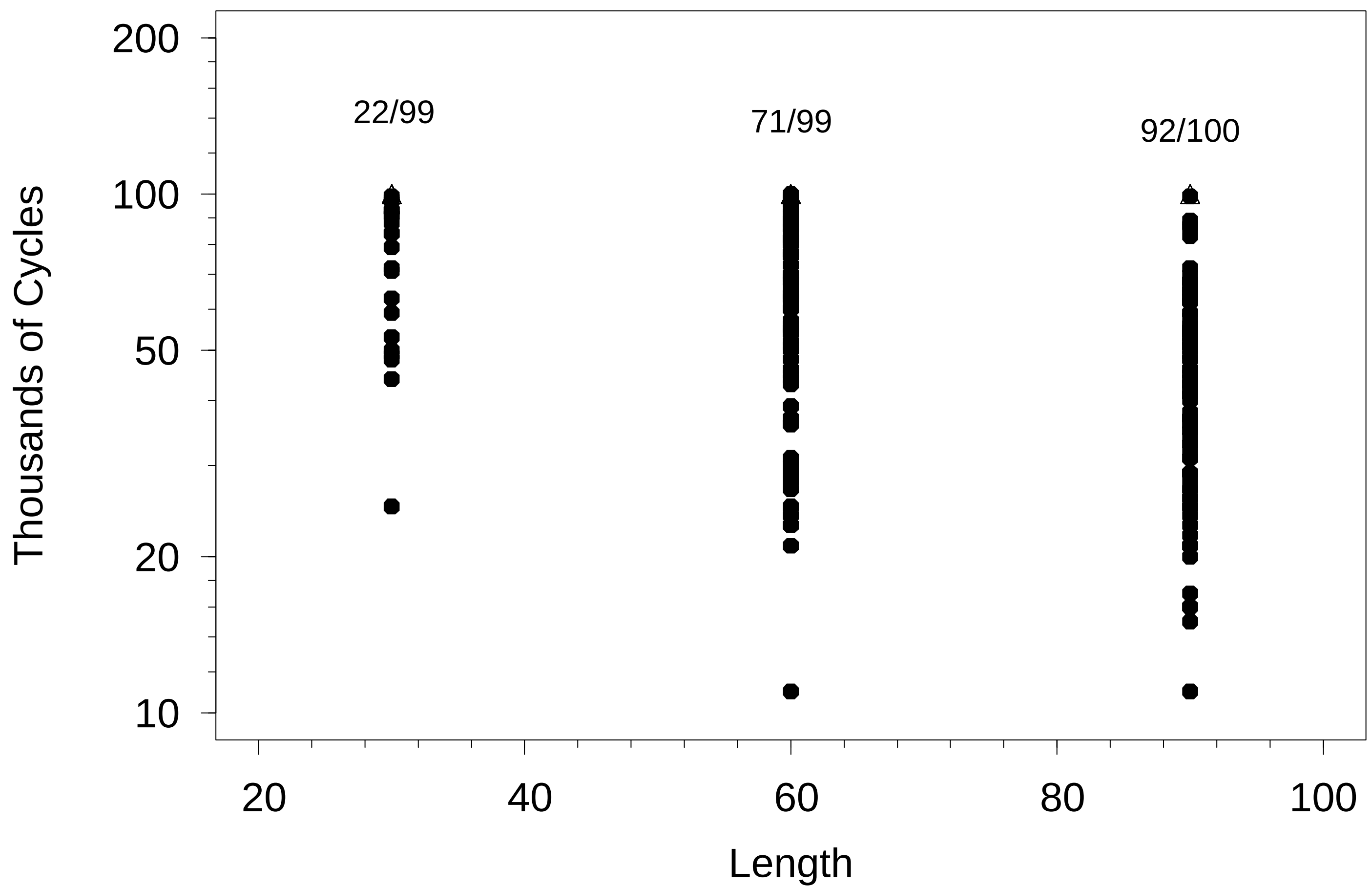
- Cycles to failure on specimens of yarn of different lengths.

- Subset of data at particular level of stress and particular specimen lengths (30mm, 60mm, and 90mm)

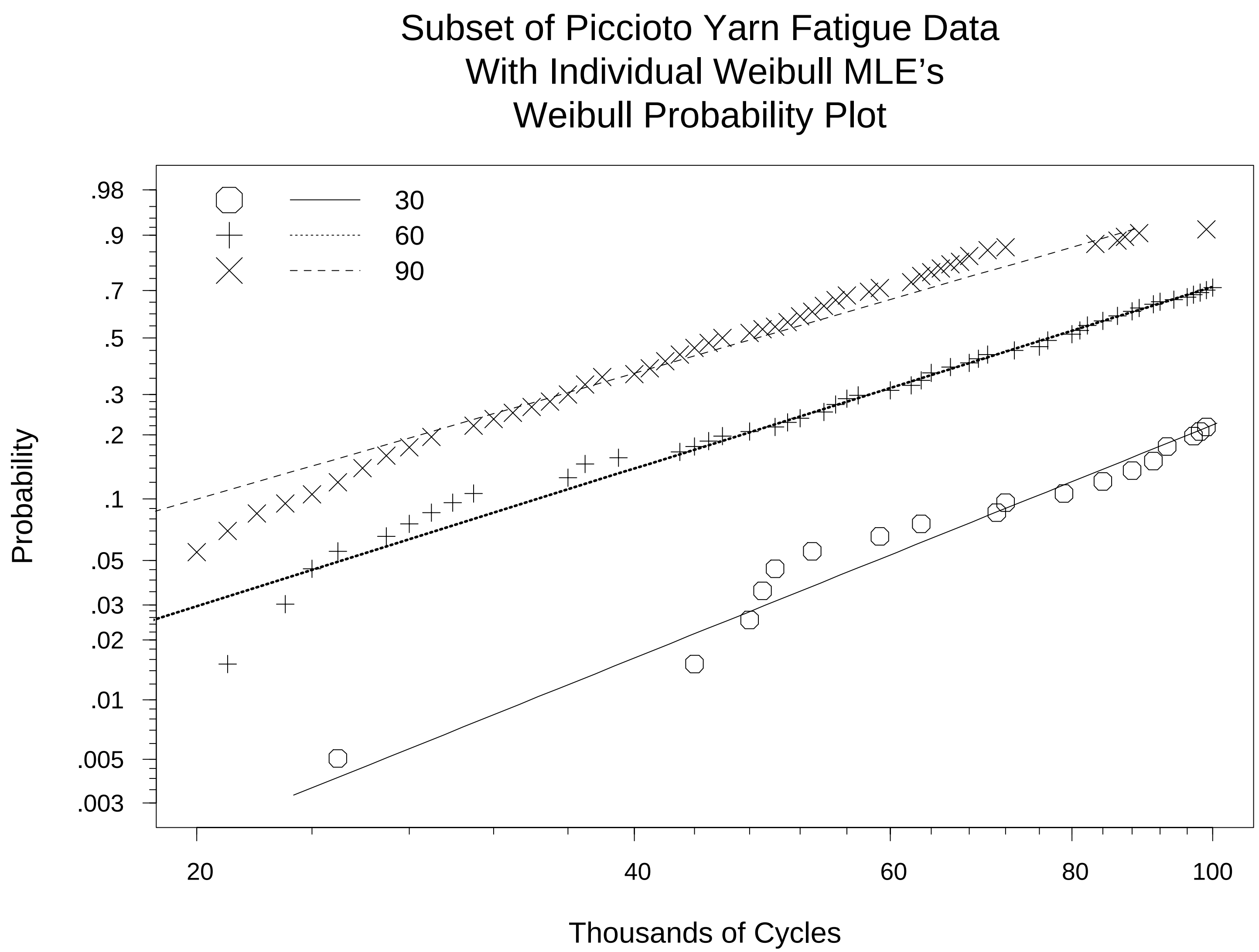
- Original data were uncensored. For purposes of illustration, the data censored at 100 thousand cycles.

- Suppose that the goal is to estimate life for 10 mm units.

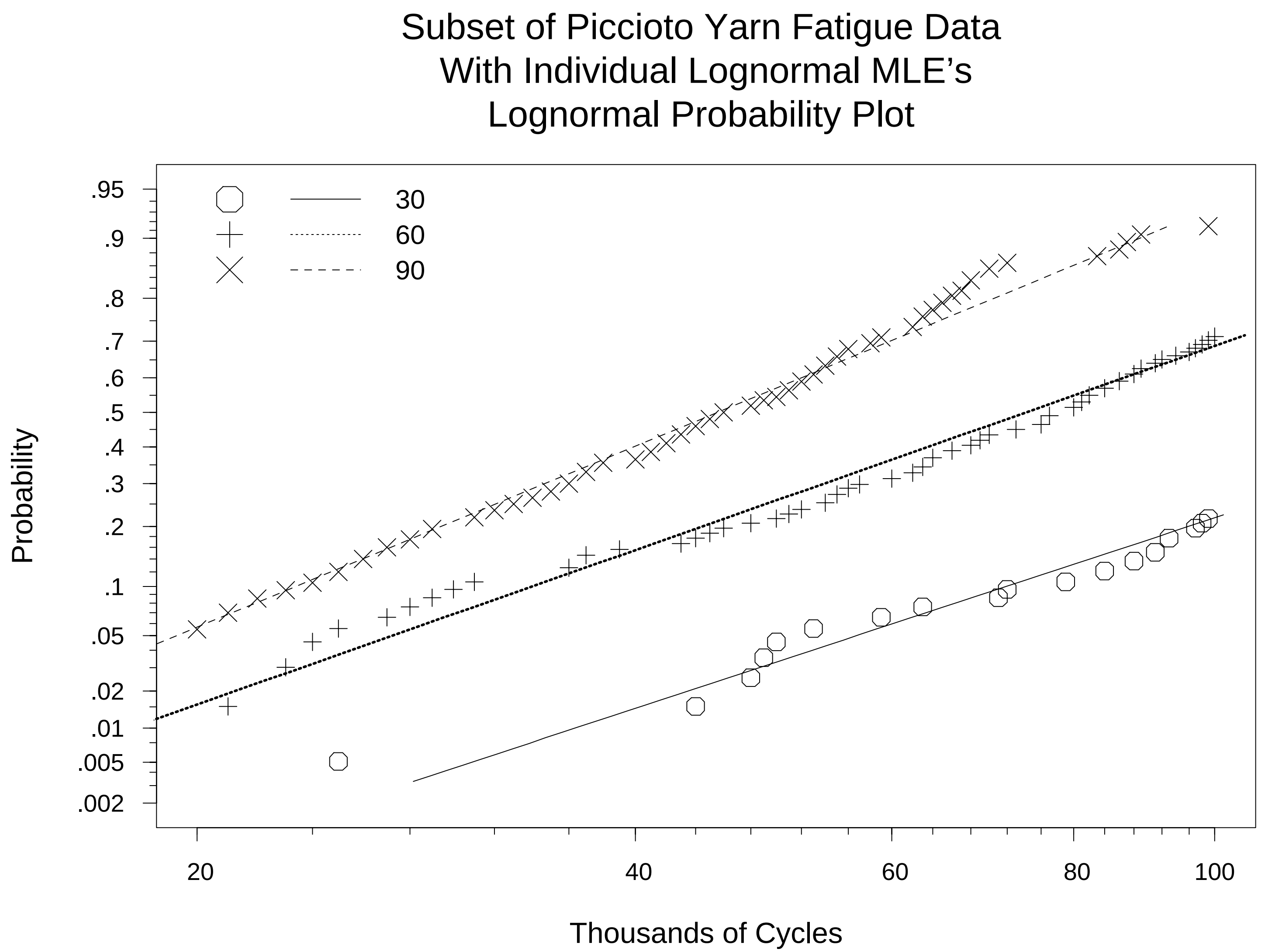
Picciotto Data Showing Fraction Failing at Each Length



Picciotto Data
Multiple Weibull Probability Plot
with Weibull ML Estimates



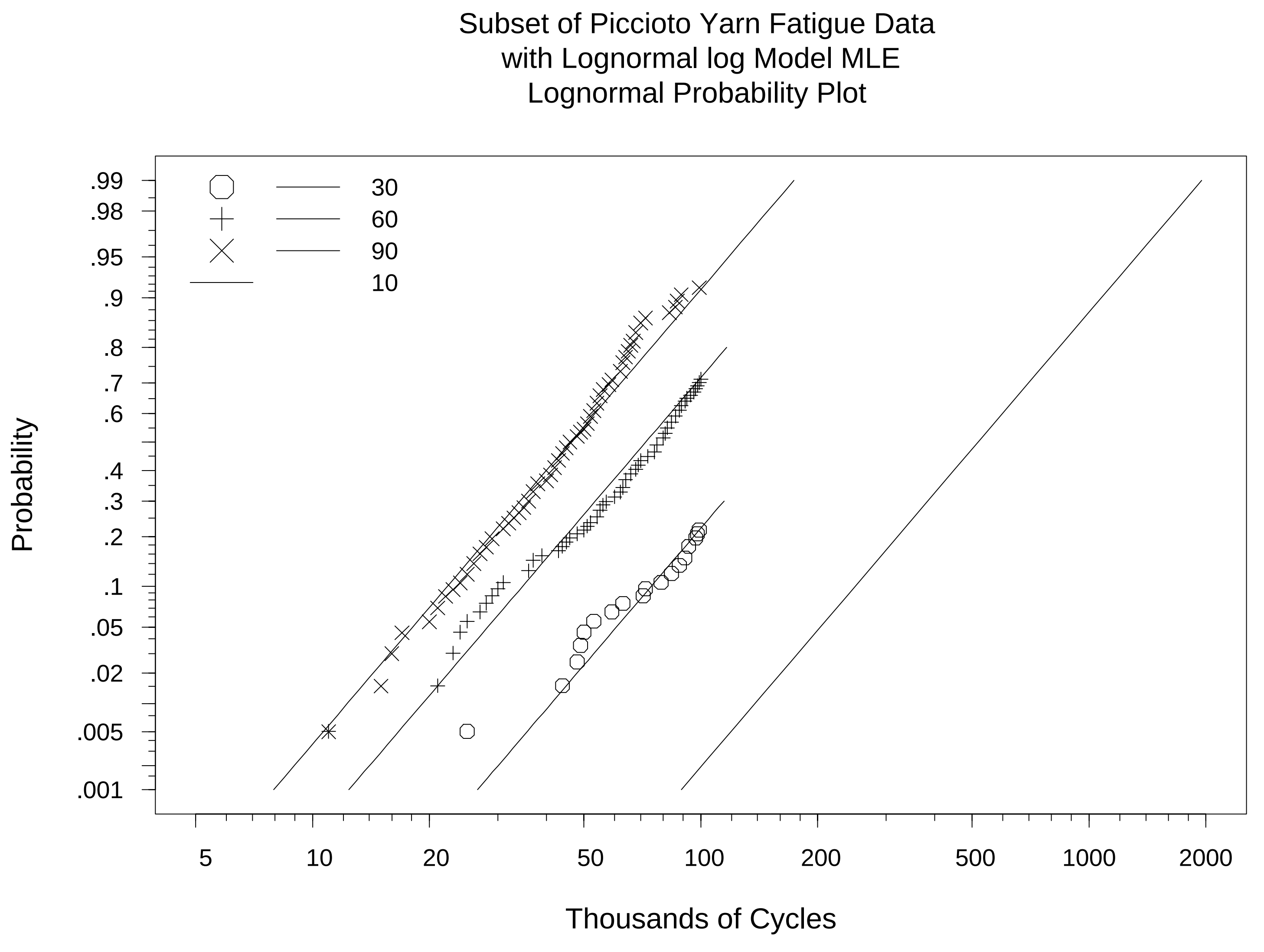
Picciotto Data
Multiple Lognormal Probability Plot
with Lognormal ML Estimates



Suggested Strategy for Fitting Empirical Regression Models and Sensitivity Analysis

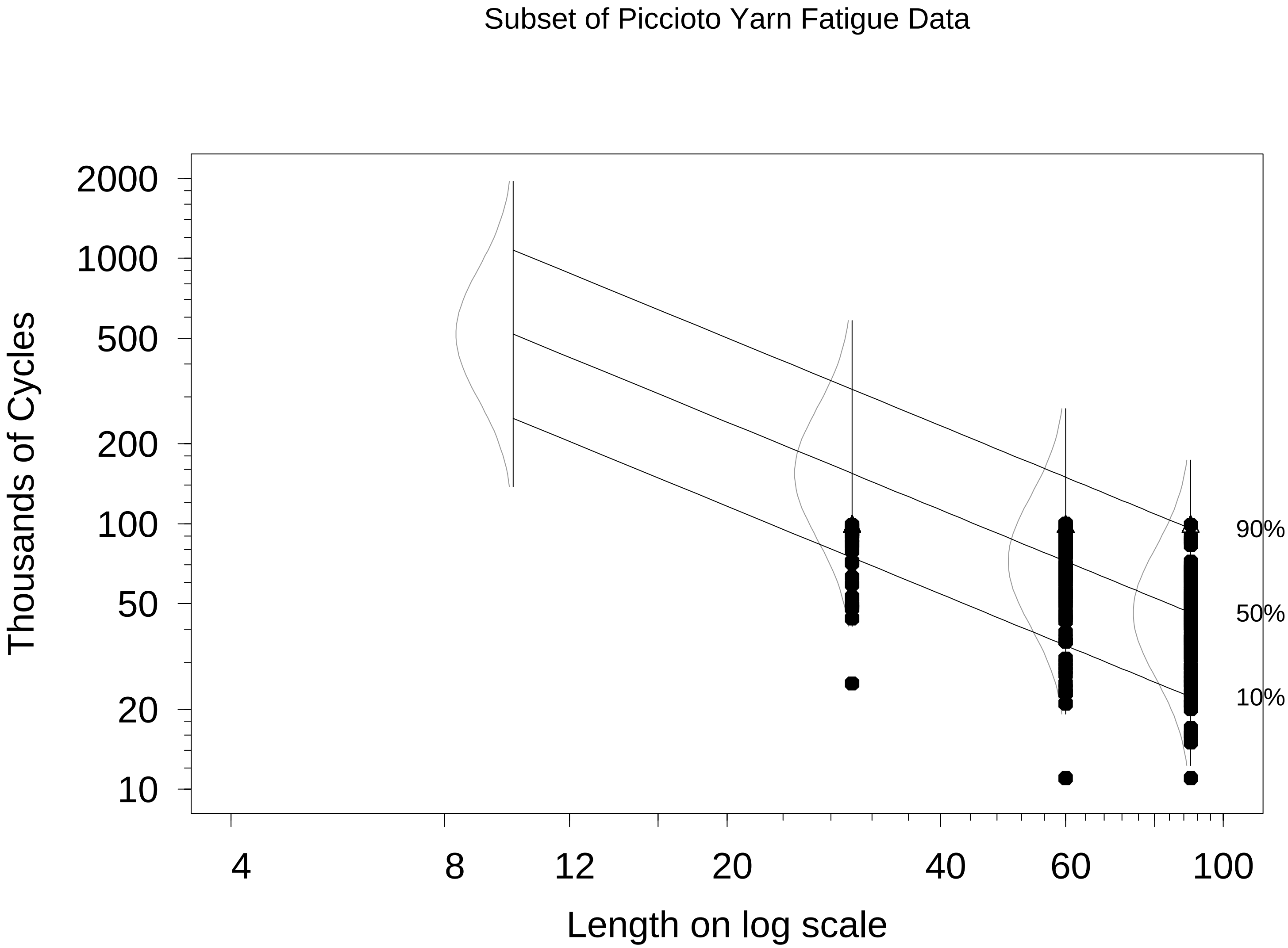
- Use data and previous experience to choose a base-line model:
  - Distribution at individual conditions
  - Relationship between explanatory variables and distributions at individual conditions
- Fit models
- Use diagnostics (e.g., residual analysis) to check models
- Assess uncertainty
  - Confidence intervals quantify statistical uncertainty
  - Perturb and otherwise change the model and reanalyze (sensitivity analysis) to assess model uncertainty.

Picciotto Data Multiple Lognormal Probability Plot with Lognormal ML Log-Linear Regression Model Estimates

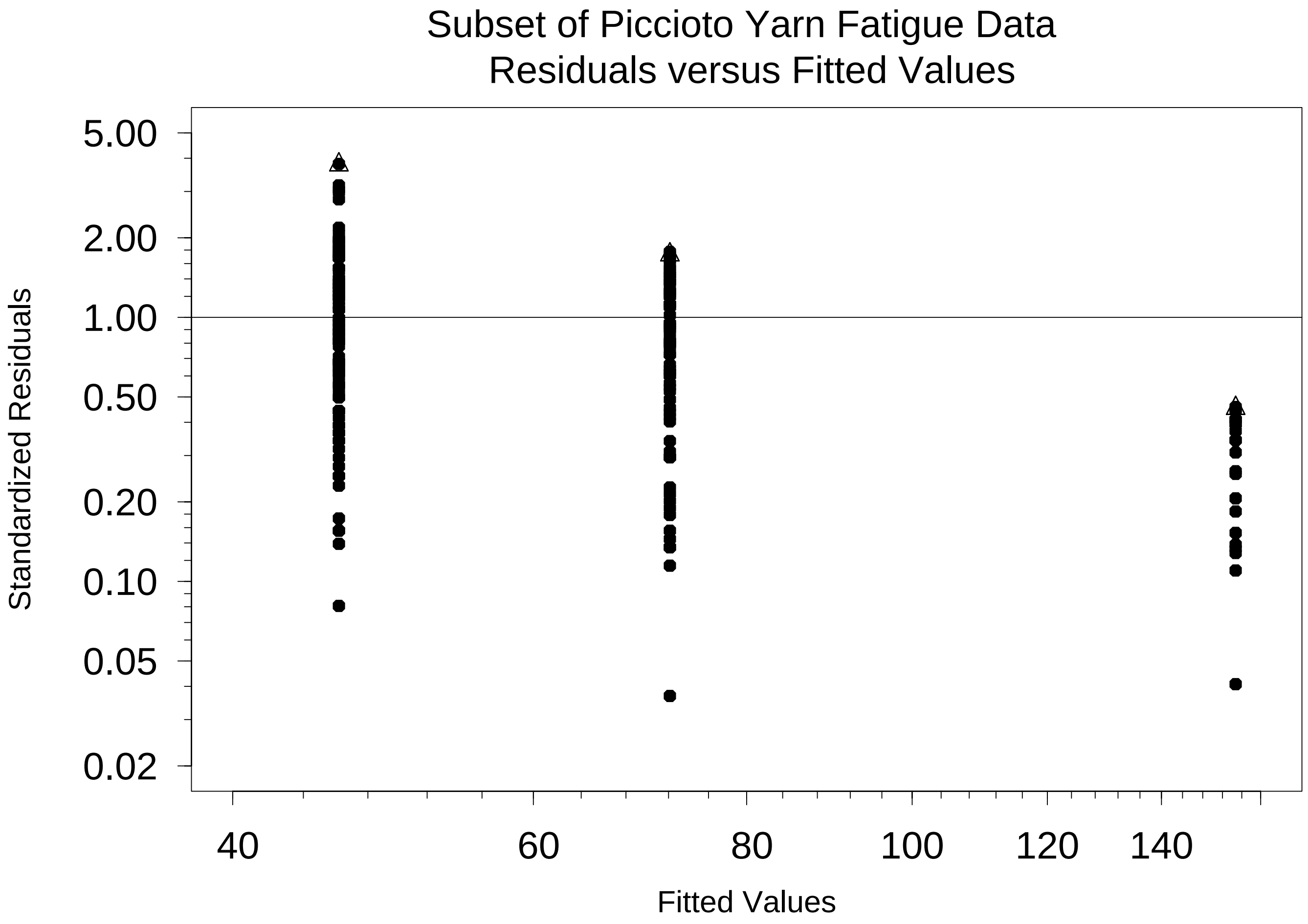


Scatter Plot of the Picciotto Data with Lognormal ML Log-Linear Regression Model Estimate Densities

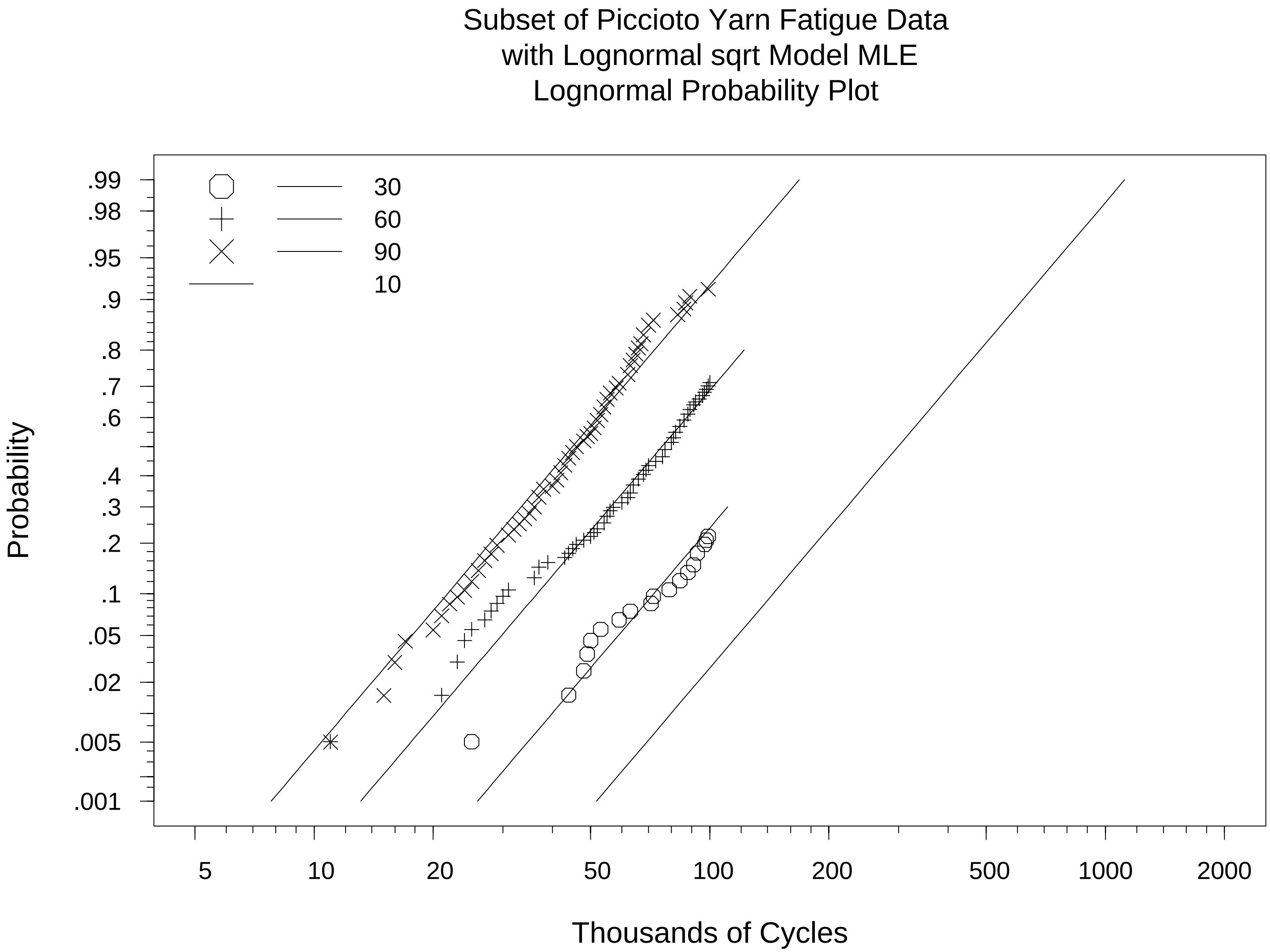
$$\log[\widehat{t}_p(x)] = \widehat{\beta}_0 + \widehat{\beta}_1 \log(\textbf{Length}) + \Phi_{\text{nor}}^{-1}(p)\widehat{\sigma}$$



Picciotto Data
Log-Linear Model Residual vs Fitted Values Plot

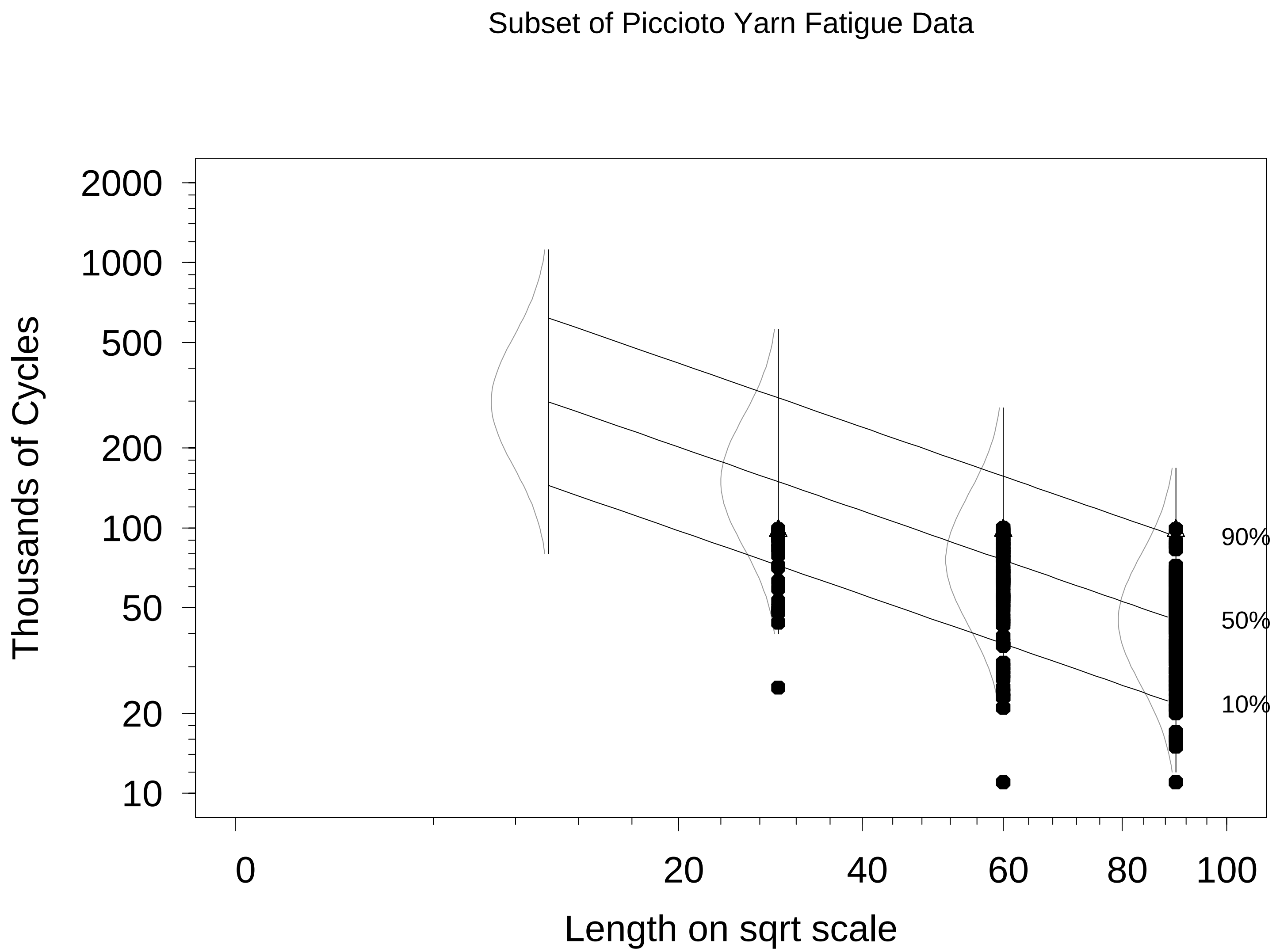


Picciotto Data
Multiple Lognormal Probability Plot
with Lognormal ML Squareroot-Linear Regression
Model Estimates

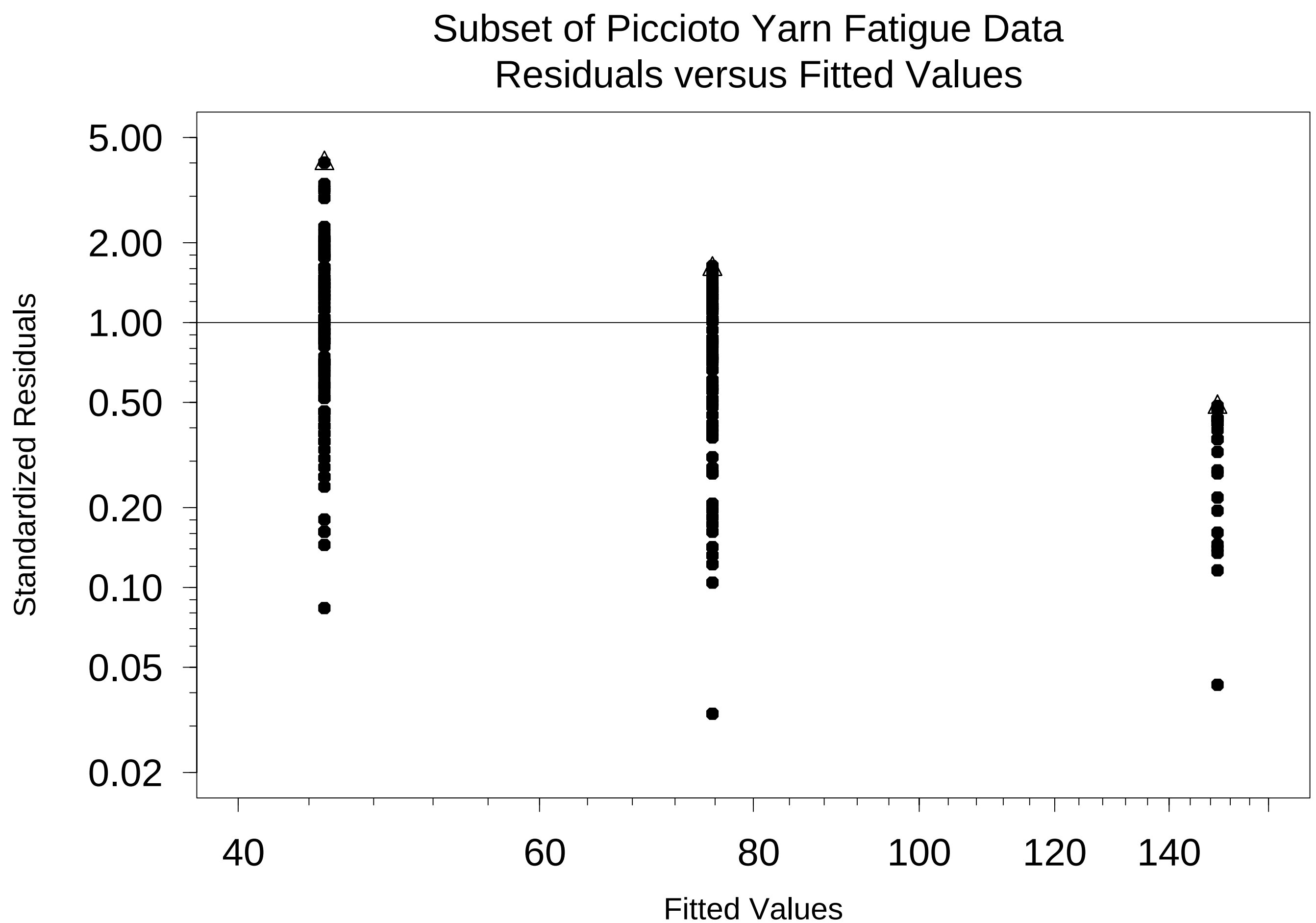


Scatter Plot of the Picciotto Data with Lognormal ML Squareroot-Linear Regression Model Estimate Densities

$$\log[\widehat{t_p}(x)] = \widehat{\beta}_0 + \widehat{\beta}_1\sqrt{\textbf{Length}} + \Phi^{-1}_{\text{nor}}(p)\widehat{\sigma}$$



Picciotto Data
Squareroot-Linear Model Residual vs Fitted Values Plot



Examples of Power Transformations

λ	Transformation
-2	$x_i^* \sim -1/x_i^2$
-1	$x_i^* \sim -1/x_i$
$-.5$	$x_i^* \sim -1/\sqrt{x_i}$
$-.333$	$x_i^* \sim -1/x_i^{1/3}$
0	$x_i^* \sim \log(x_i)$
$.333$	$x_i^* \sim x_i^{1/3}$
$.5$	$x_i^* \sim \sqrt{x_i}$
1	$x_i^* \sim x_i$
2	$x_i^* \sim x_i^2$

Box-Cox Transformation

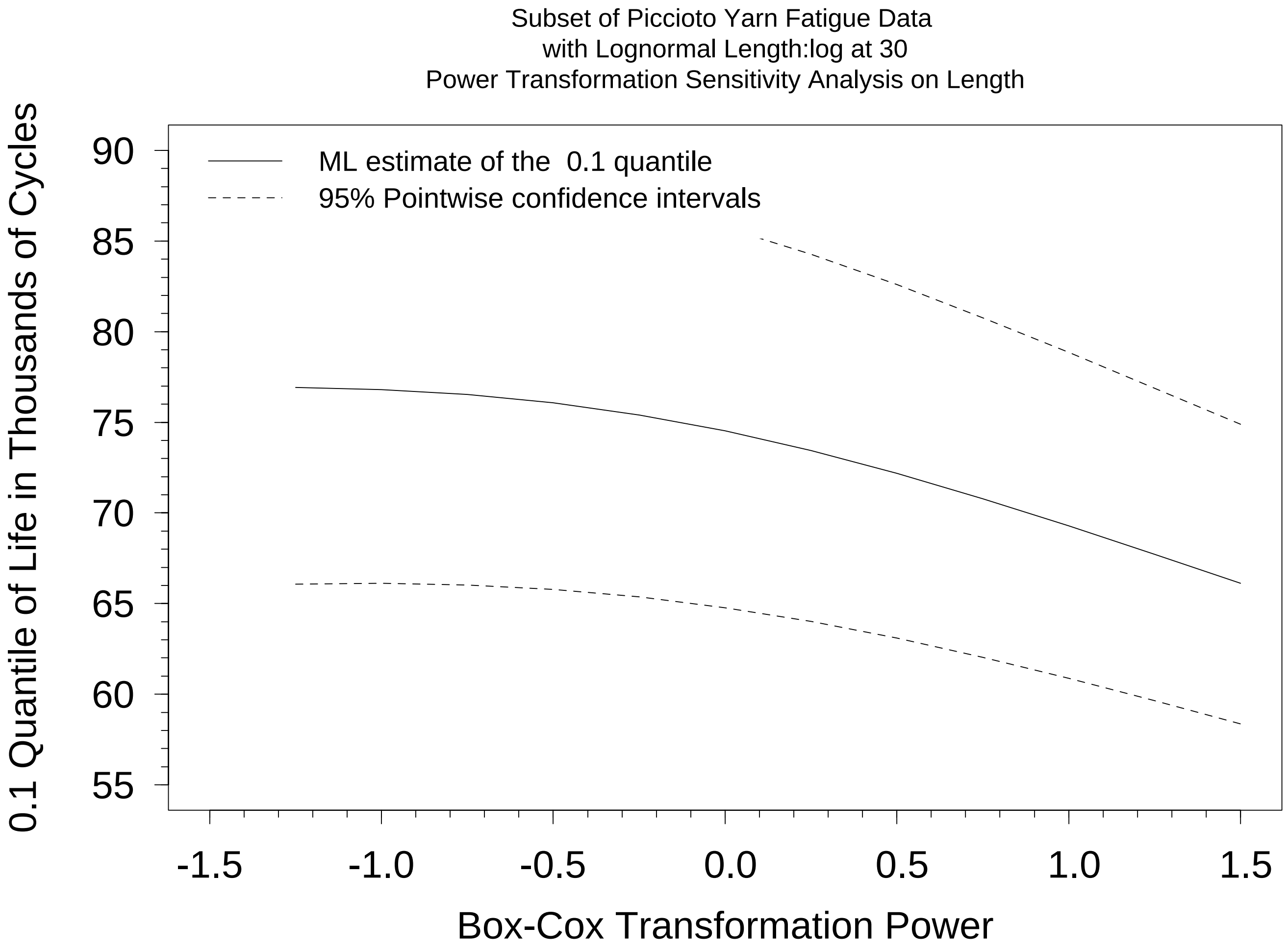
- The Box-Cox family of power transformations is

$$x_i^* = \begin{cases} \frac{x_i^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \log(x_i) & \lambda = 0 \end{cases}$$

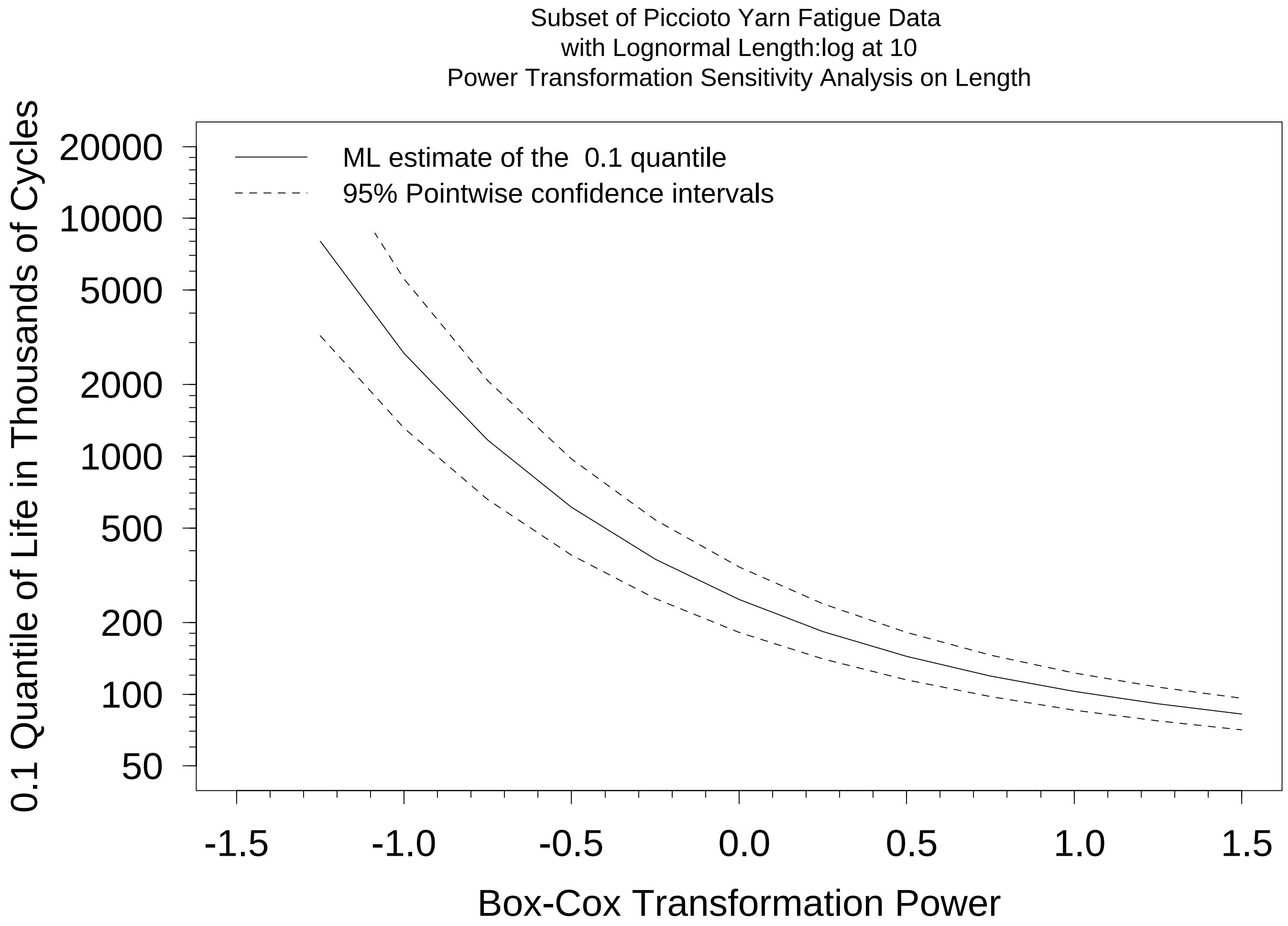
  where x_i is the original, untransformed explanatory variable and for observation i and λ is power transformation parameter.

- The Box-Cox transformation has the following properties:
  - ▶ The transformed value x_i^* is an increasing function of x_i.
  - ▶ For fixed, x_i, x_i^* is a continuous function of λ through 0.

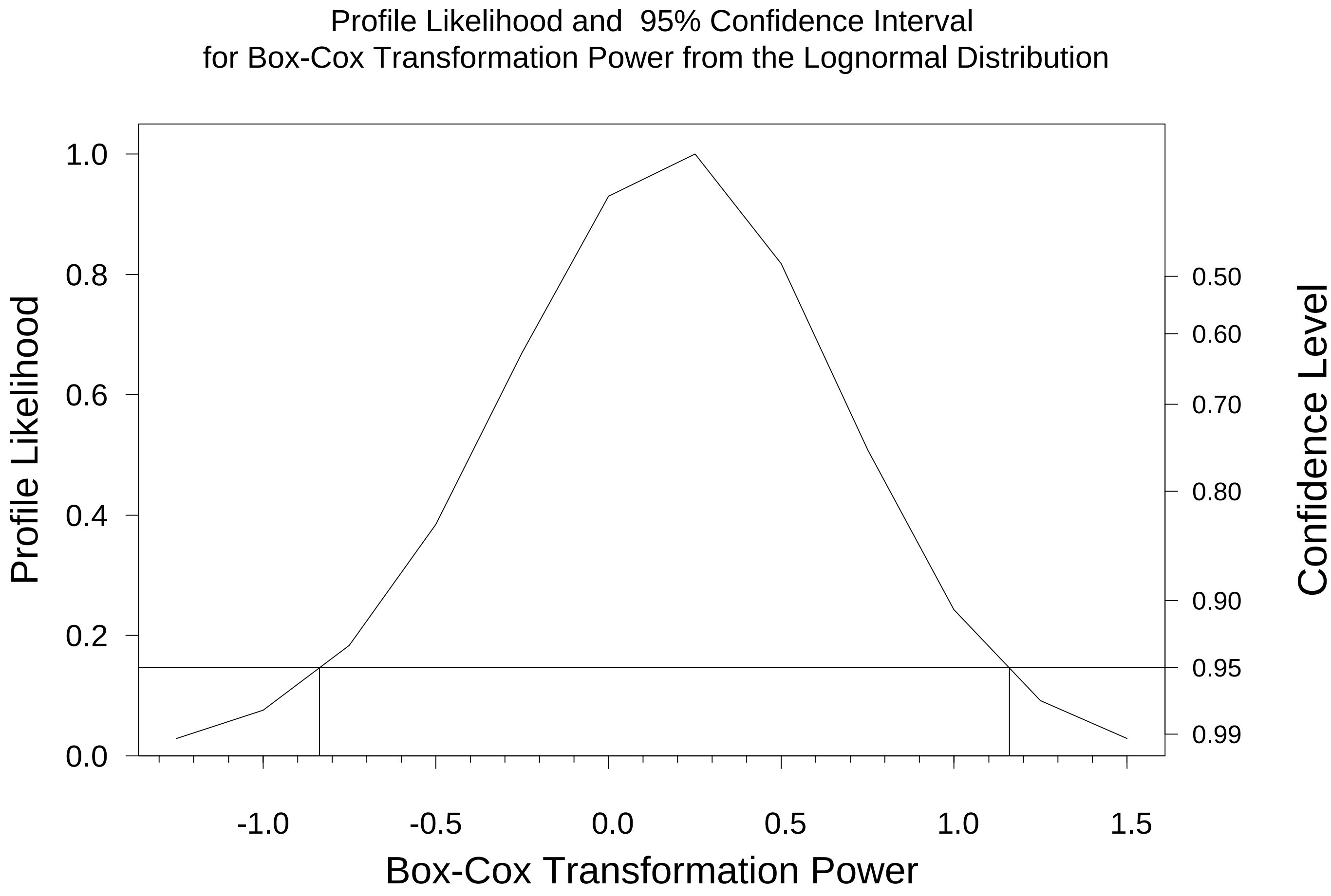
Picciotto Data
Relationship Sensitivity Analysis
for the .1 Quantile at Length 30 mm



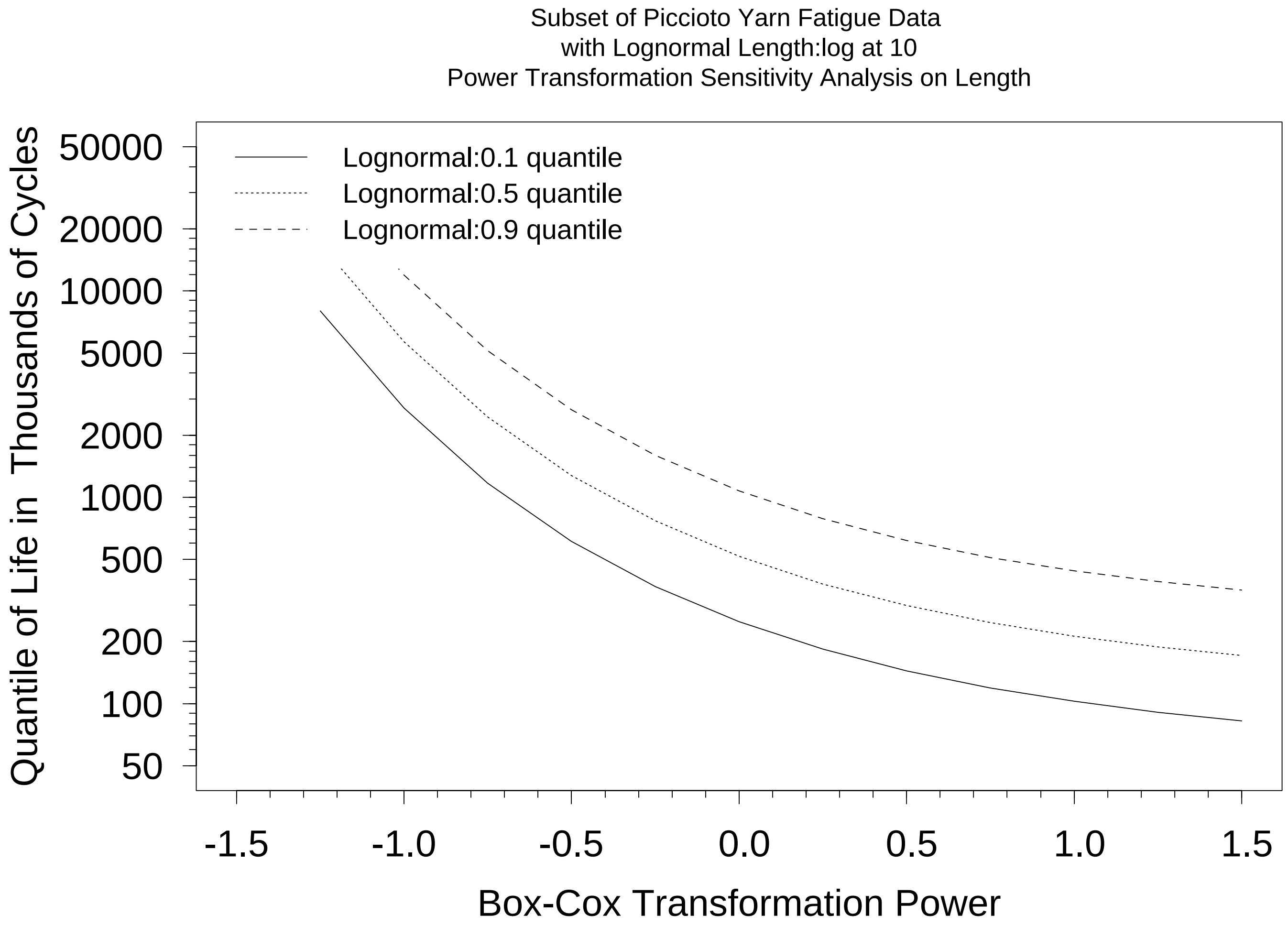
Picciotto Data
Relationship Sensitivity Analysis
for the .1 Quantile at Length 10 mm



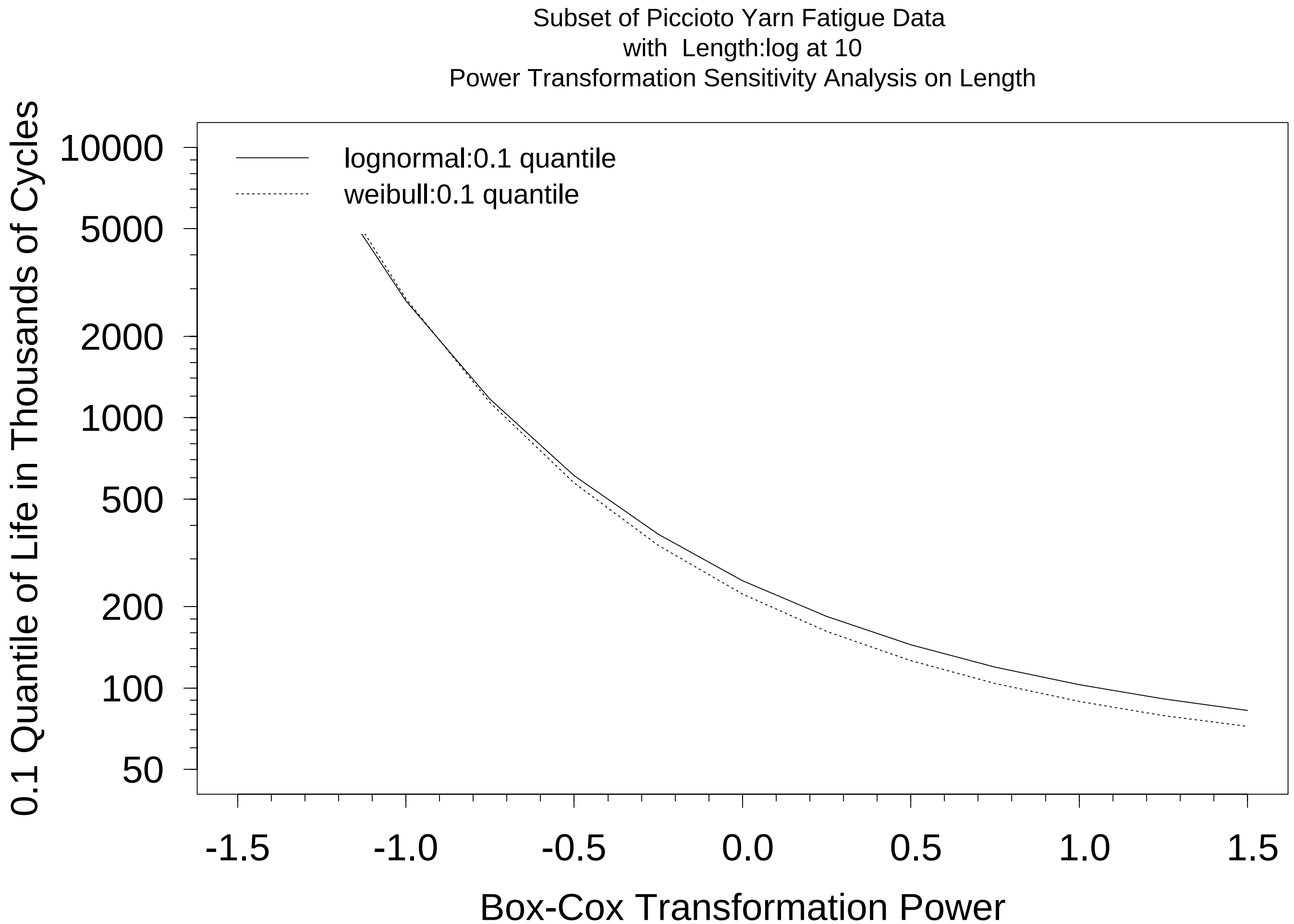
Picciotto Data
Profile Plot for Different Box-Cox Parameters for the Length Relationship



Picciotto Data
Relationship Sensitivity Analysis .1, .5, .9 Quantiles at 10 mm Length



Picciotto Data
Relationship/Distribution Sensitivity Analysis .1 Quantile at 10 mm Length



Glass Capacitor Failure Data

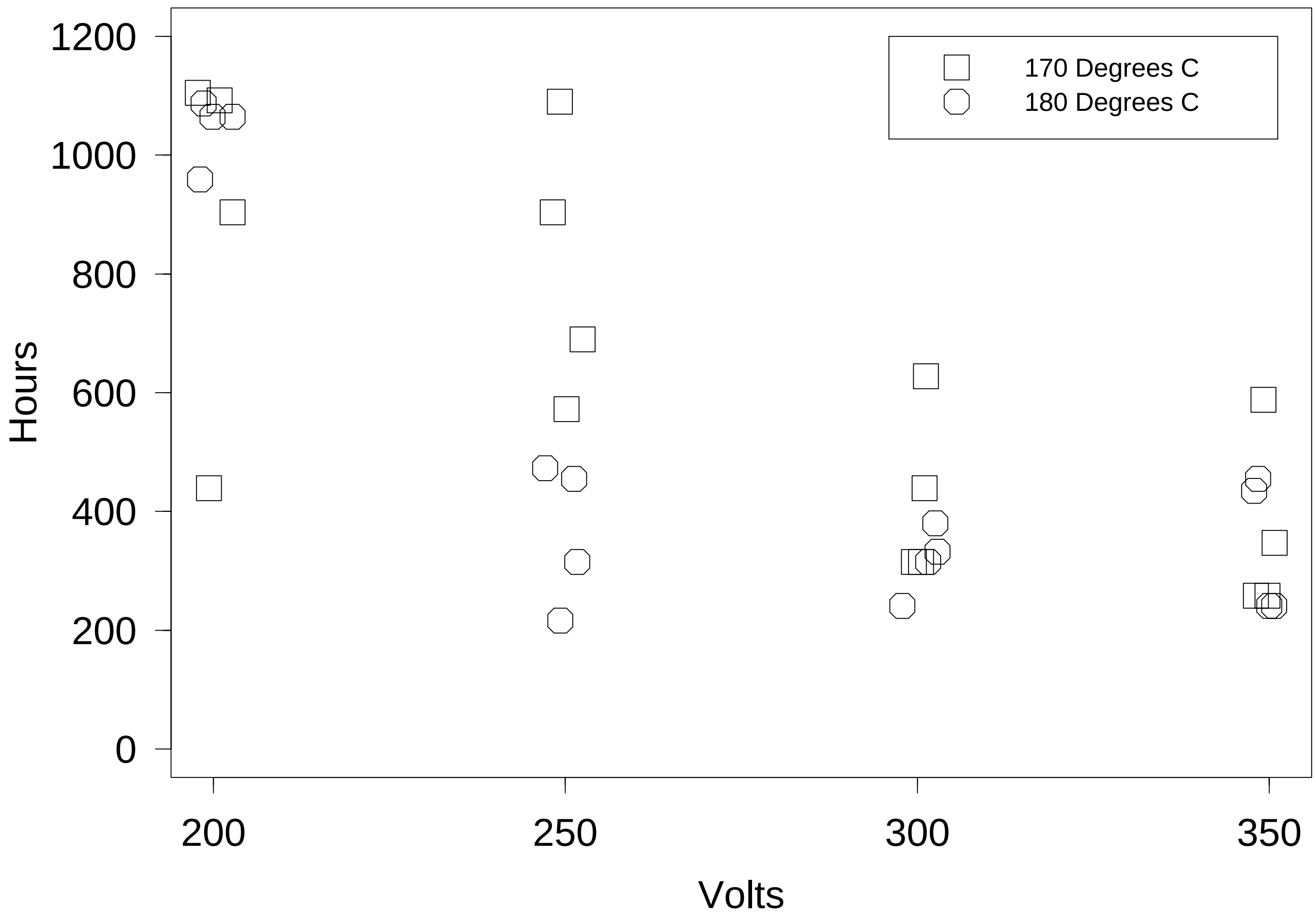
- Experiment designed to determine the effect of voltage and temperature on capacitor life.

- 2×4 factorial, 8 units at each combination.

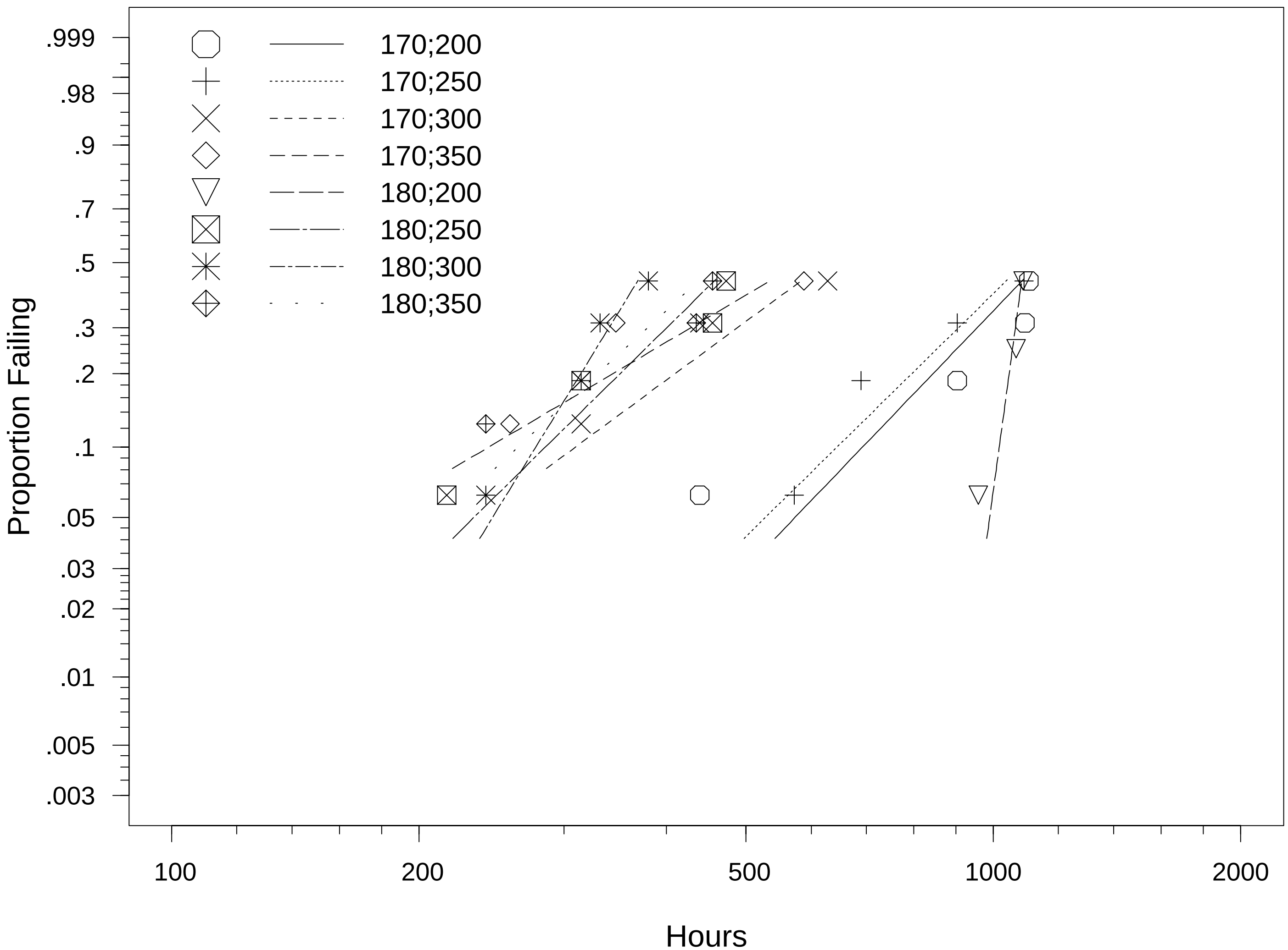
- Test at each combination run until 4 of 8 units failed (Type II censoring).

- Original data from Zelen (1959).

Scatter Plot the Effect of Voltage and Temperature on Glass Capacitor Life (Zelen 1959)



Weibull Probability Plots of Glass Capacitor Life Test Results at Individual Temperature and Voltage Test Conditions



Two-Variable Regression Models

- Additive model

$$\log[t_p(\boldsymbol{x})] = y_p(\boldsymbol{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \Phi^{-1}(p)\sigma.$$

  Because

$$t_p(\boldsymbol{x}) = \exp\left[y_p(\boldsymbol{x})\right] = \exp(\beta_1 x_1 + \beta_2 x_2)\, t_p(\underline{0})$$

  this is an SAFT model

- The interaction model

$$\log[t_p(\boldsymbol{x})] = y_p(\boldsymbol{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \Phi^{-1}(p)\sigma.$$

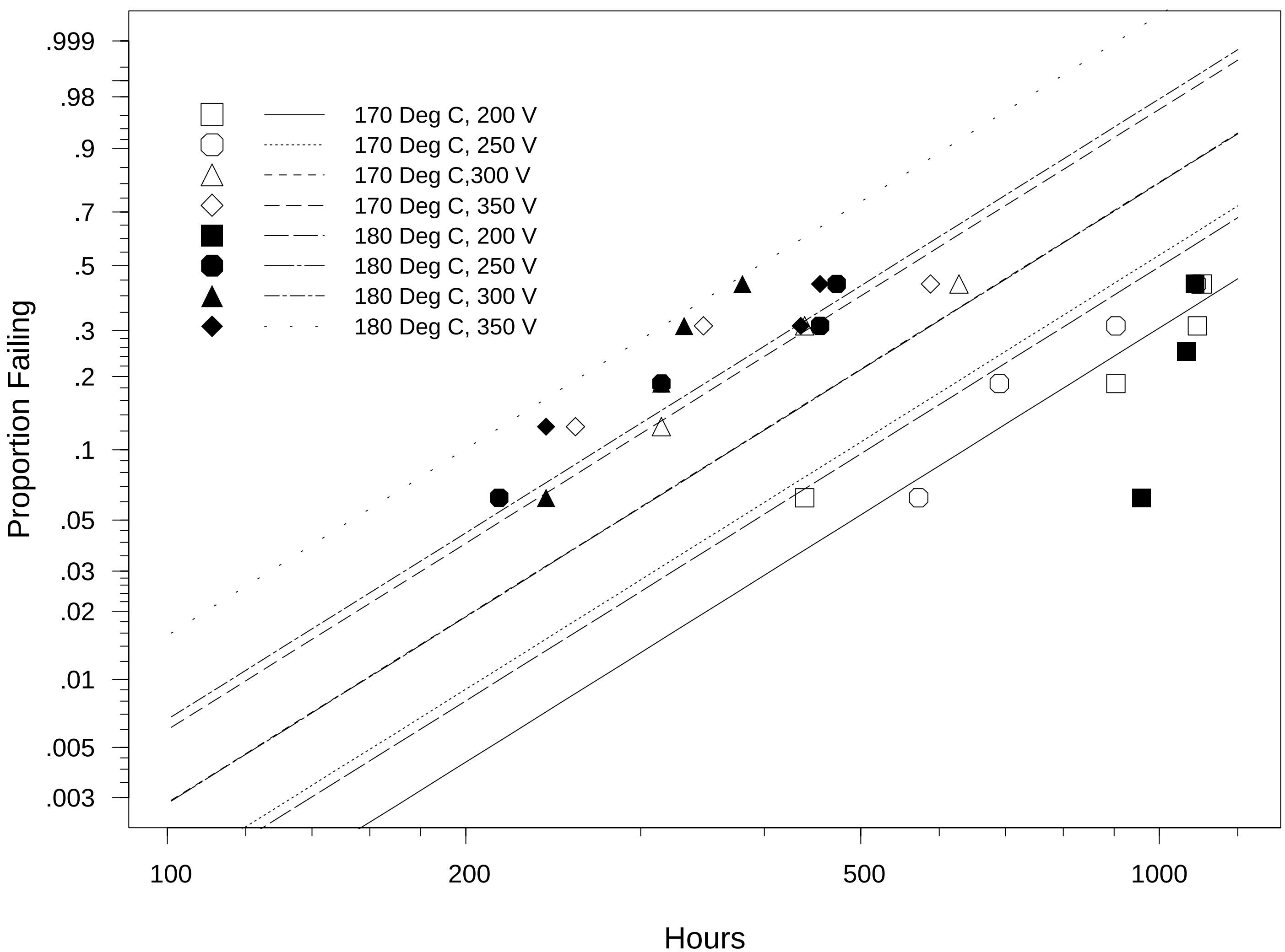
  is also SAFT.

- Comparing the two models gives

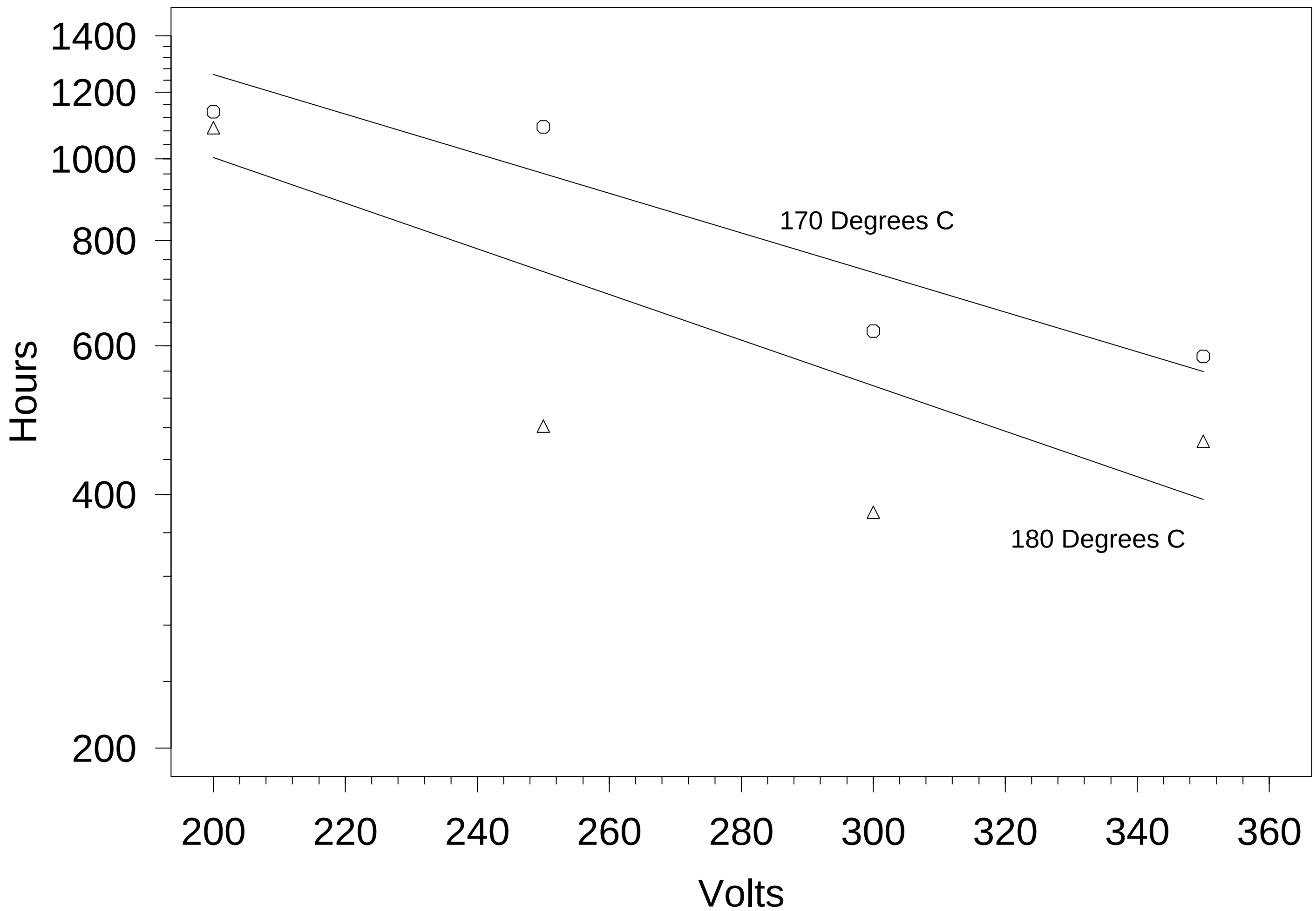
$$-2 \times (\mathcal{L}_1 - \mathcal{L}_2) = -2 \times (-244.24 + 244.17) = .14$$

  which is small relative $\chi^2_{.95,1} = 3.84$

Weibull Probability Plots with Weibull Regression Model ML Estimates of $F(t)$ at each Set of Conditions for the Glass Capacitor Data



Estimates of Weibull $t_{.5}$ Plotted for each Combination of the Glass Capacitor Test Conditions



The Proportional Hazard Failure Time Model

The proportional hazard (PH) model assumes

$$h(t; \boldsymbol{x}) = \Psi(\boldsymbol{x}) h(t; \boldsymbol{x}_0), \quad \text{for all } t > 0.$$

$h(t; \boldsymbol{x}_0)$ and $F(t; \boldsymbol{x}_0)$ denote the baseline hazard function and cdf of the model.

The PH model implies (some details later):

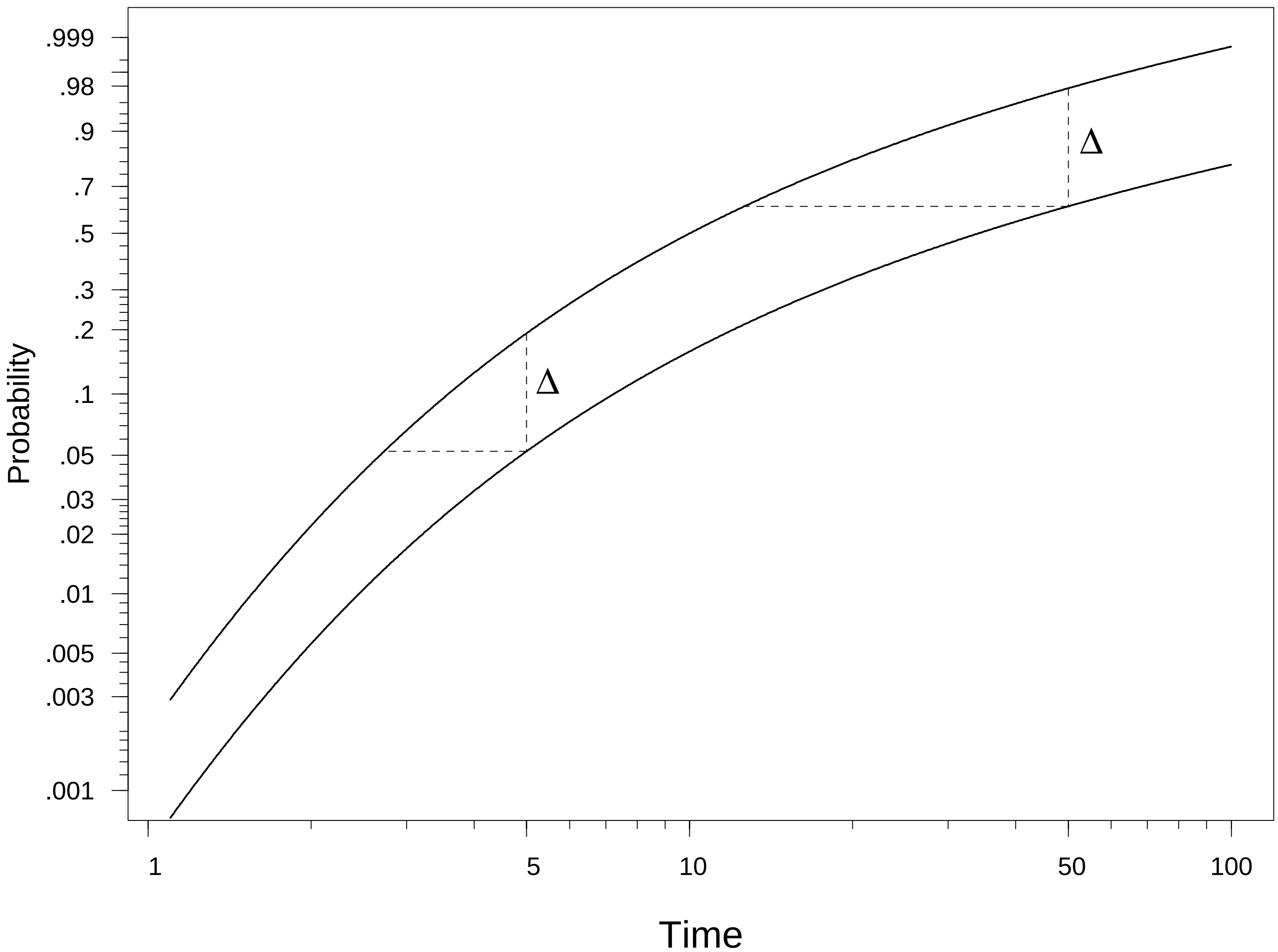
- $S(t; \boldsymbol{x}) = [S(t; \boldsymbol{x}_0)]^{\Psi(\boldsymbol{x})}$ or $1 - F(t; \boldsymbol{x}) = [1 - F(t; \boldsymbol{x}_0)]^{\Psi(\boldsymbol{x})}$. If $\Psi(\boldsymbol{x}) \neq 1$, $F(t; \boldsymbol{x})$ and $F(t; \boldsymbol{x}_0)$ do not cross and the model is accelerating if $\Psi(\boldsymbol{x}) > 1$ and decelerating if $\Psi(\boldsymbol{x}) < 1$.

- From $1 - F(t; \boldsymbol{x}) = [1 - F(t; \boldsymbol{x}_0)]^{\Psi(\boldsymbol{x})}$ and taking logs (twice):

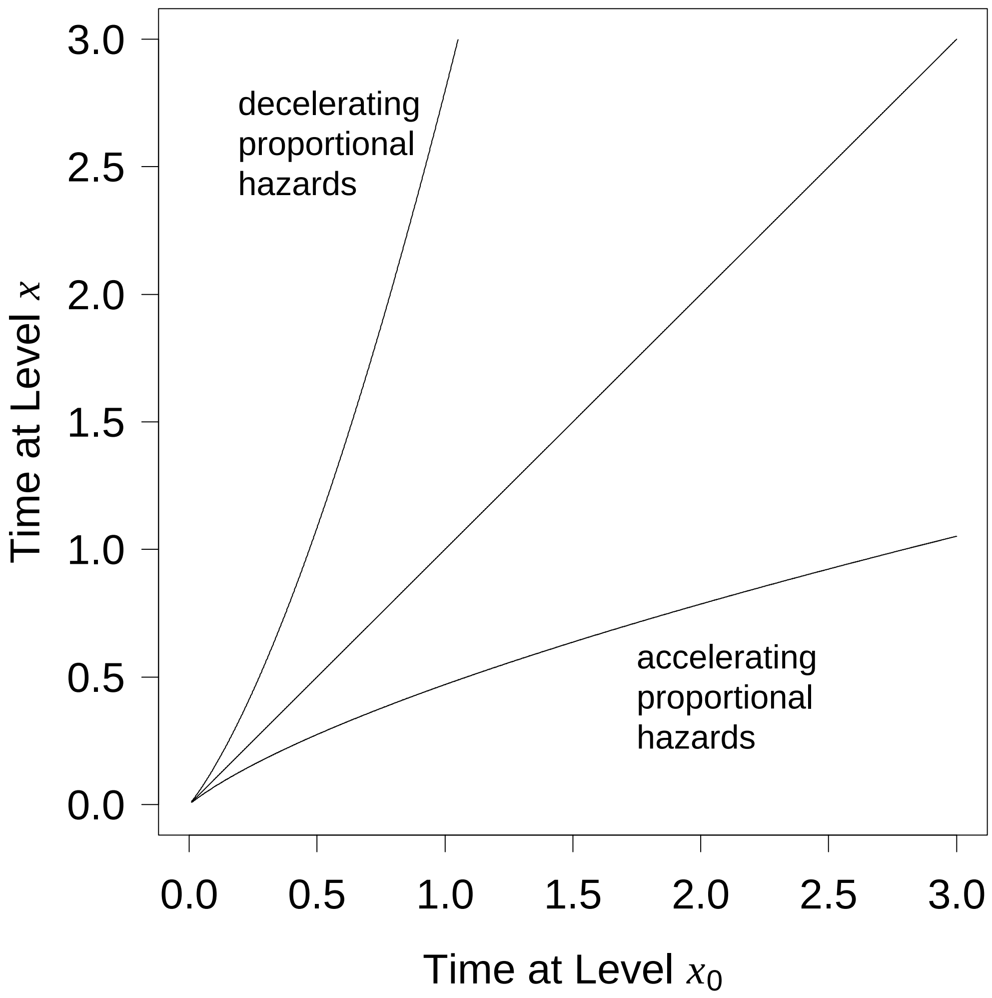
  $$\log\{-\log[1 - F(t; \boldsymbol{x})]\} - \log\{-\log[1 - F(t; \boldsymbol{x}_0)]\} = \log[\Psi(\boldsymbol{x})].$$

  Thus in a Weibull probability scale $F(t; \boldsymbol{x})$ and $F(t; \boldsymbol{x}_0)$ are equidistant. In particular, Weibull plots of $F(t; \boldsymbol{x})$ and $F(t; \boldsymbol{x}_0)$ are translation of each other along the probability axis.

Weibull Probability Plot of Two Members from a PH Model with a Baseline Lognormal Distribution



Proportional Hazard Model (Lognormal Baseline) as a Time Transformation



Interpreting PH Models as a Failure Time Transformation.

Suppose that $T(\boldsymbol{x}_0) \sim F(t; \boldsymbol{x}_0)$ and define the time transformation:

$$T(\boldsymbol{x}) = F^{-1}\left(1 - \{1 - F\left[T(\boldsymbol{x}_0); \boldsymbol{x}_0\right]\}^{\frac{1}{\Psi(\boldsymbol{x})}}; \boldsymbol{x}_0\right).$$

It can be shown that:

- $T(\boldsymbol{x})$ and $T(\boldsymbol{x}_0)$ follow the PH relationship

$$h(t; \boldsymbol{x}) = \Psi(\boldsymbol{x}) h(t; \boldsymbol{x}_0).$$

- $T(\boldsymbol{x})$ is a monotone transformation of $T(\boldsymbol{x}_0)$ such that

$$\begin{array}{lcllll} T(\boldsymbol{x}) & < & T(\boldsymbol{x}_0) & \text{if} & \Psi(\boldsymbol{x}) > 1, & \text{accelerating} \\ T(\boldsymbol{x}) & = & T(\boldsymbol{x}_0) & \text{if} & \Psi(\boldsymbol{x}) = 1, & \text{identity transform} \\ T(\boldsymbol{x}) & > & T(\boldsymbol{x}_0) & \text{if} & \Psi(\boldsymbol{x}) < 1, & \text{decelerating} \end{array}$$

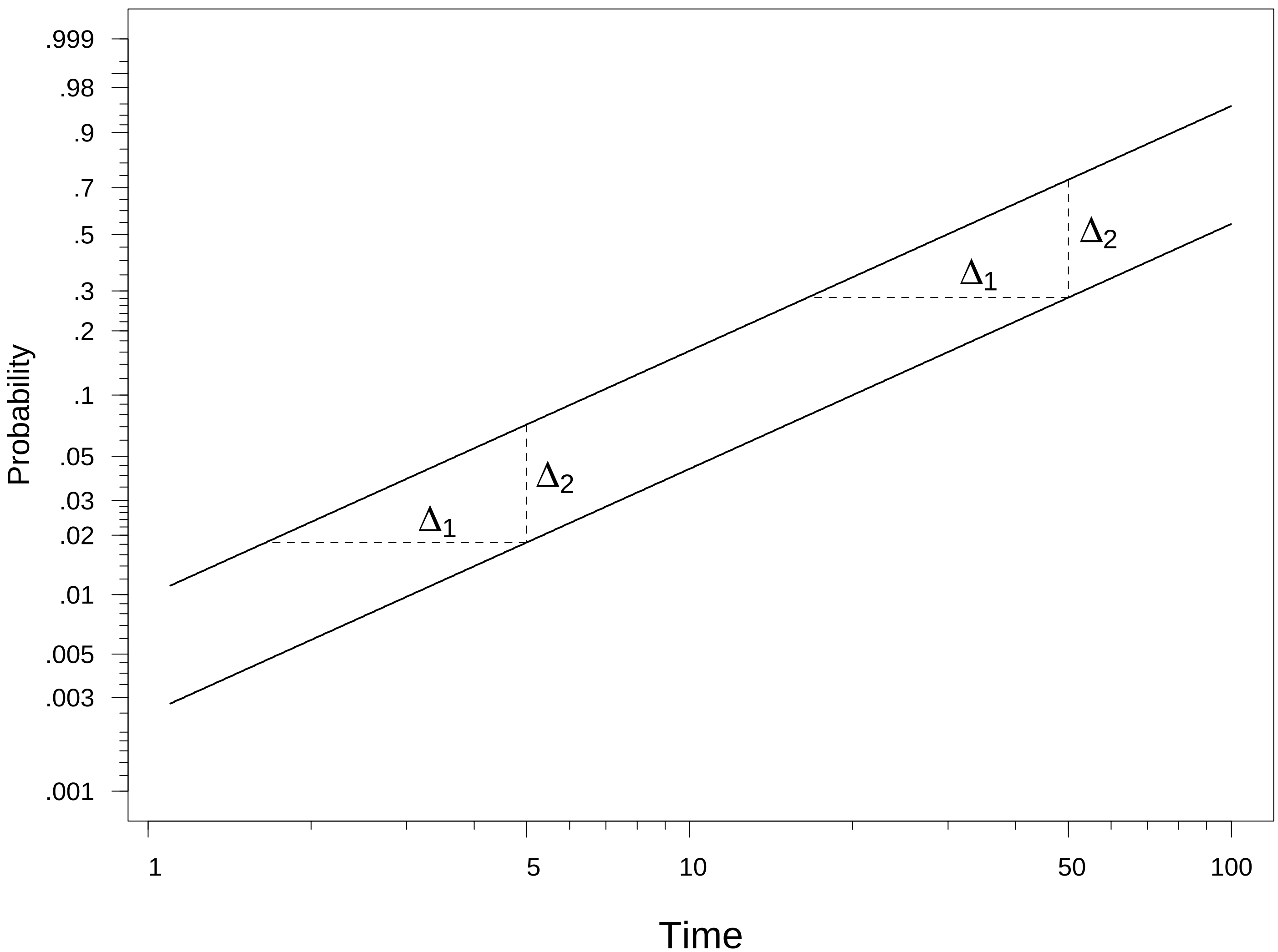
Some Comments on the Proportional Hazard Failure Time Model

Here we consider the proportional hazards, PH, model

$$h(t; \boldsymbol{x}) = \Psi(\boldsymbol{x}) h(t; \boldsymbol{x}_0), \quad \text{for all } t > 0.$$

- In general, the PH model does not preserve the baseline distribution. For example, if $T(\boldsymbol{x}_0)$ has a lognormal distribution then $T(\boldsymbol{x})$ has a power lognormal distribution.
- The semiparametric model without a particular specification of $h(t; \boldsymbol{x}_0)$ is known as **Cox's** proportional hazards model.

Weibull Probability Plot of Two Members from a PH/SAFT Model with a Baseline Weibull Distribution



Contrasting SAFT and PH Models

The **scale** failure time model $t_p(\boldsymbol{x}) = t_p(\boldsymbol{x}_0)/\Psi(\boldsymbol{x})$ and the PH model $h(t;\boldsymbol{x}) = \Psi(\boldsymbol{x})h(t;\boldsymbol{x}_0)$ are equivalent if only if the baseline distribution is Weibull.

In other words, the Weibull distribution is the only baseline distribution for which a SAFT model is also a a PH model, and vice versa.

Heuristic argument: Consider $F(t;\boldsymbol{x}_1)$ and $F(t;\boldsymbol{x}_2)$. The Weibull probability plots of these two cdfs are translation of each other in both the probability and the log(t) scale if and only if the plots are straight parallel lines.

Statistical Methods for the Semiparametric (Cox) PH Model

Data: n units, $t_{(1)}, \ldots, t_{(r)}$ ordered failure times with $n - r$ censored observations (usually multiply censored).

- $\mathcal{RS}_i = \mathcal{RS}(t_{(i)} - \varepsilon)$ is the **risk set** just before the failure at time $t_{(i)}$.

- Each unit has a vector $\boldsymbol{x}_i$ of explanatory variables (often called covariates).

Cox PH Model Likelihood (Probability of the Data)

- The probability that individual i dies in $[t, t + \Delta t]$ is

$$h(t; \boldsymbol{x}_i)\Delta t = h(t; \boldsymbol{x}_0)\Delta t \Psi(\boldsymbol{x}_i) = h(t; \boldsymbol{x}_0)\, \mathrm{exp}(\boldsymbol{x}_i' \boldsymbol{\beta})\Delta t$$

where $\Psi(\boldsymbol{x}_i) = \mathrm{exp}(\boldsymbol{x}_i'\boldsymbol{\beta}) = \mathrm{exp}(\beta_1 x_1 + \cdots + \beta_k x_k)$ and $\boldsymbol{x}_0 = \boldsymbol{0}$

- If a death occurs at time t, the probability that it was individual i is

$$L_i = \frac{h(t; \boldsymbol{x}_i)\Delta t}{\sum_{\ell \in \mathcal{RS}_i} h(t; \boldsymbol{x}_\ell)\Delta t} = \frac{\mathrm{exp}(\boldsymbol{x}_i'\boldsymbol{\beta})}{\sum_{\ell \in \mathcal{RS}_i} \mathrm{exp}(\boldsymbol{x}_\ell'\boldsymbol{\beta})}$$

- Conditional on the observed failure times, the joint probability of the data is

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{r} \frac{\mathrm{exp}(\boldsymbol{x}_i'\boldsymbol{\beta})}{\sum_{\ell \in \mathcal{RS}_i} \mathrm{exp}(\boldsymbol{x}_\ell'\boldsymbol{\beta})}$$

- Need mild conditions on $\boldsymbol{x}_i$ and the censoring mechanism.

Comments on the Cox PH Model Likelihood

- Cox (1972) called $L(\boldsymbol{\beta})$ a conditional likelihood.
- $L(\boldsymbol{\beta})$ is not a true likelihood, but there is justification for treating it as such.
- With no censoring, Kalbfleisch and Prentice (1973) show that $L(\boldsymbol{\beta})$ can be justified as a marginal likelihood based on ranks of the observed failures. Argument can be extended to Type II censoring.
- Cox (1975) provided a **partial likelihood** justification.
- Asymptotic theory shows that the estimates obtained from maximizing $L(\boldsymbol{\beta})$ have high efficiency.
- There are some technical difficulties when there are ties observations. Approximations for $L(\boldsymbol{\beta})$ generally used.

Estimating $h(t; \boldsymbol{x}_0)$ [or $S(t; \boldsymbol{x}_0)$ or $F(t; \boldsymbol{x}_0)$]

- Basic idea is to substitute $\widehat{\boldsymbol{\beta}}$ for $\boldsymbol{\beta}$ and maximize the probability of the data (after adjustment for the explanatory variables), as in the product limit estimator.

- Let $q_i = S(t_i + \varepsilon; \boldsymbol{x}_0)$, $i = 1, \ldots, r$ for the r failure times and let $q_j = S(t_j + \varepsilon; \boldsymbol{x}_0)$, $j = r+1, \ldots, n$ for the $n - r$ censored observations. If there are **no ties** among the failure times,

$$L(\boldsymbol{q}) = \prod_{i=1}^{r} \left(q_{i-1}^{\widehat{\Psi}(\boldsymbol{x}_i)} - q_i^{\widehat{\Psi}(\boldsymbol{x}_i)} \right) \prod_{j=r+1}^{n} q_{j-1}^{\widehat{\Psi}(\boldsymbol{x}_j)}$$

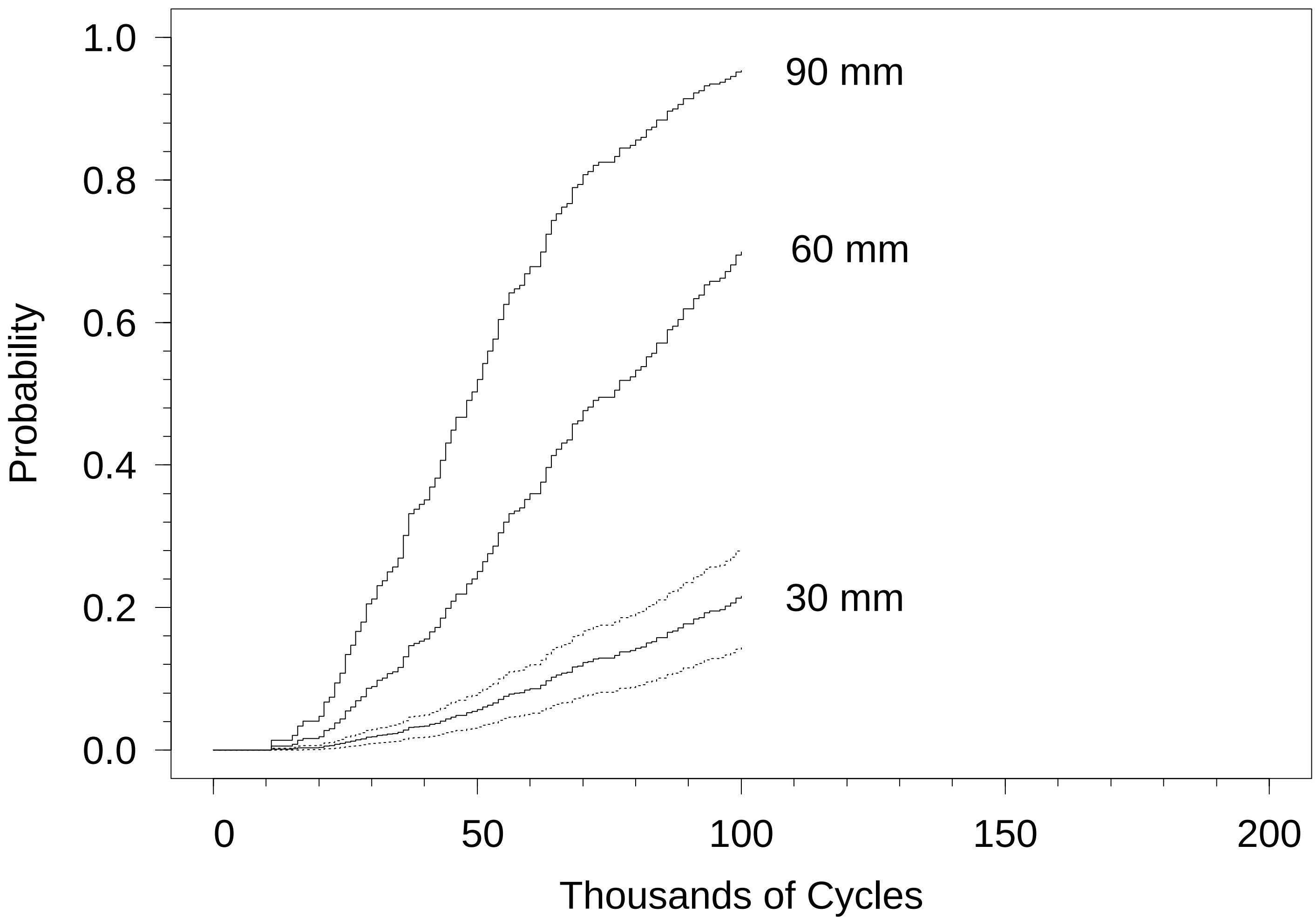
  where $\widehat{\Psi}(\boldsymbol{x}_i) = \exp(\boldsymbol{x}_i' \widehat{\boldsymbol{\beta}})$ and $q_0 \equiv 1$.

- Maximize $L(\boldsymbol{q})$ with respect to q_i, $i = 1, \ldots, r$ to get the step-function estimate of $S(t; \boldsymbol{x}_0)$.

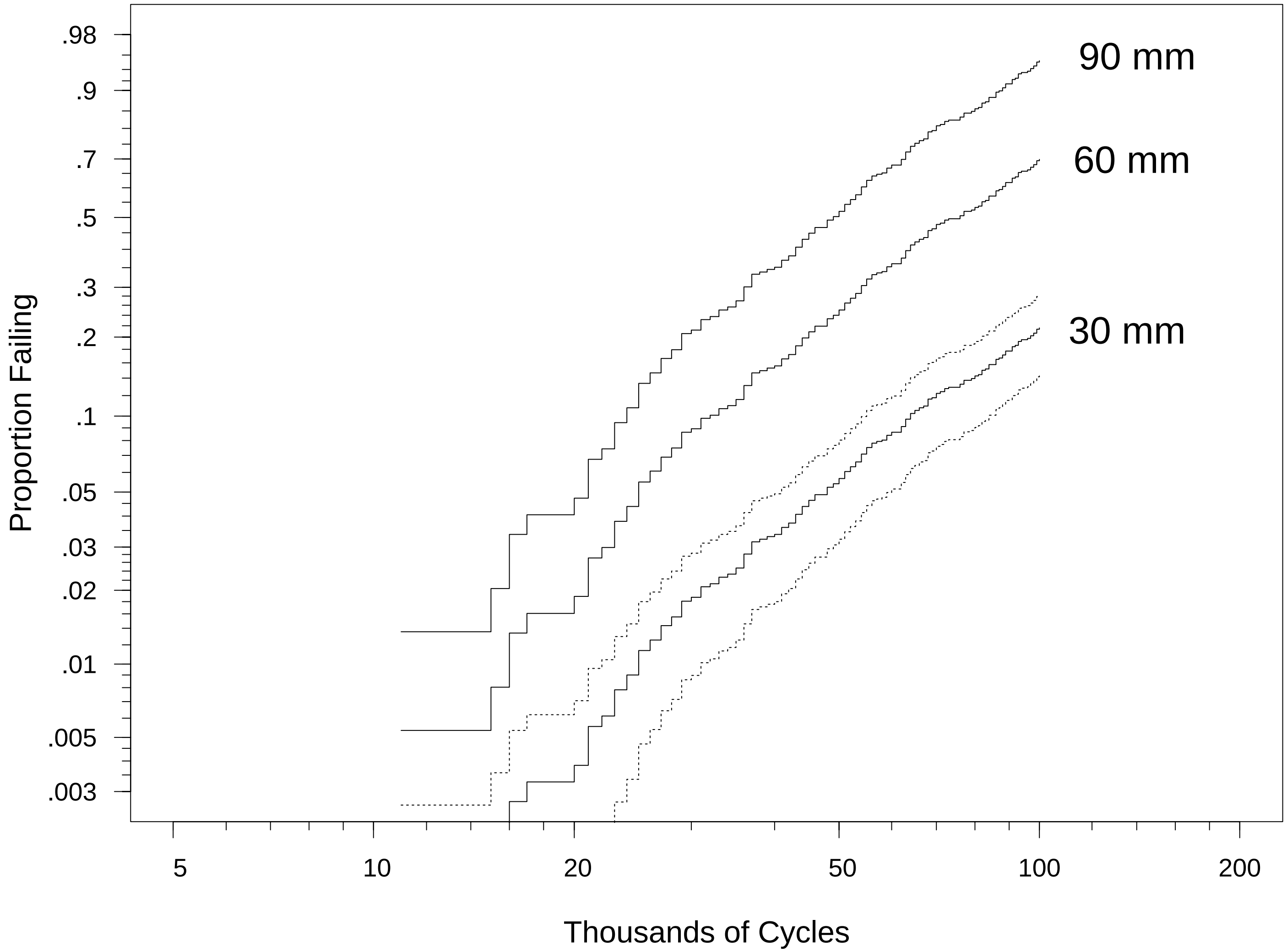
- Estimate of $S(t)$ at $\boldsymbol{x}$ is $\widehat{S}(t; \boldsymbol{x}) = [\widehat{S}(t; \boldsymbol{x}_0)]^{\widehat{\Psi}(\boldsymbol{x})}$.

- Setup more complicated when there **are ties**.

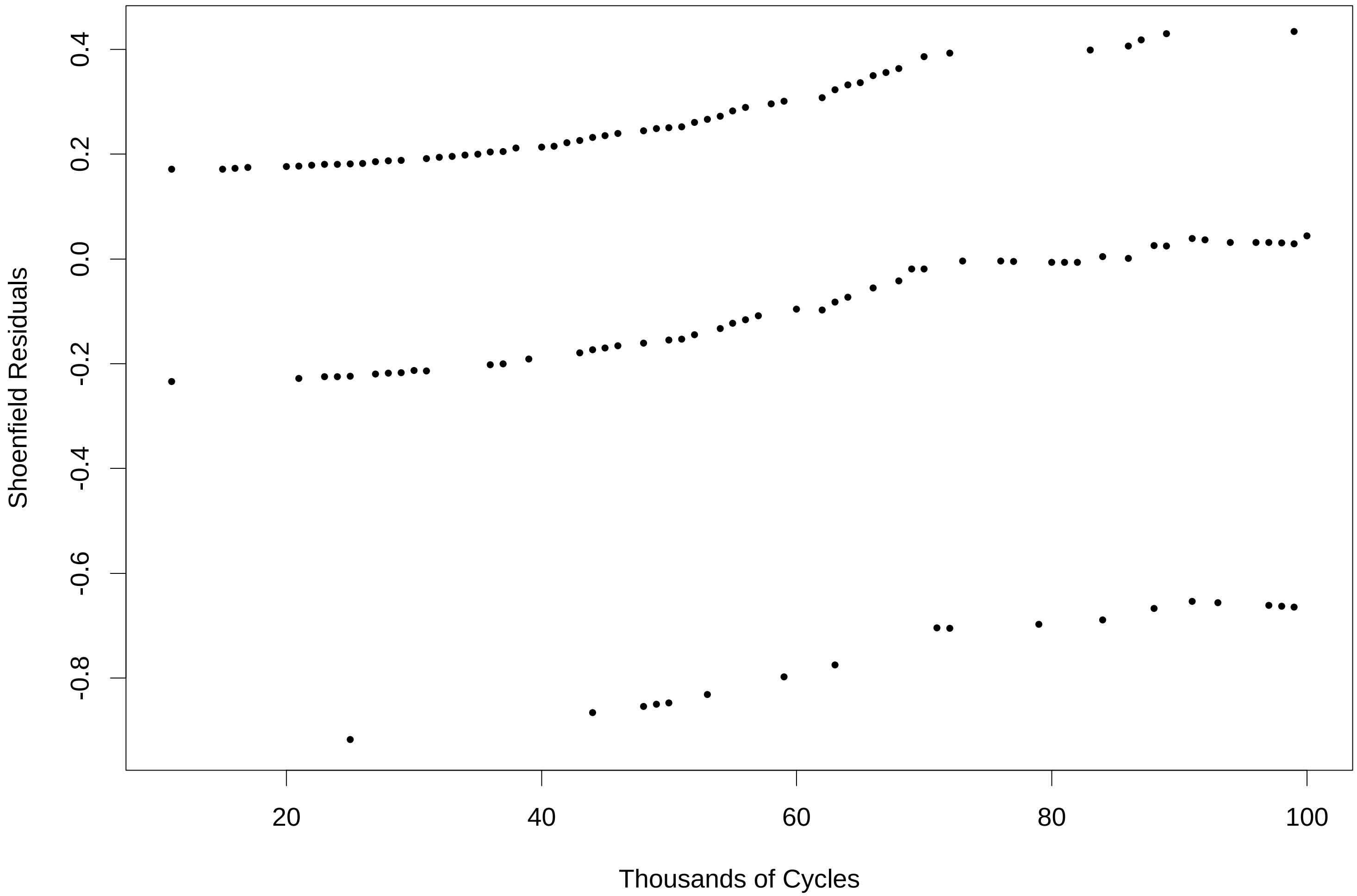
Picciotto Cox PH Model Survival Estimates (Linear Axes)



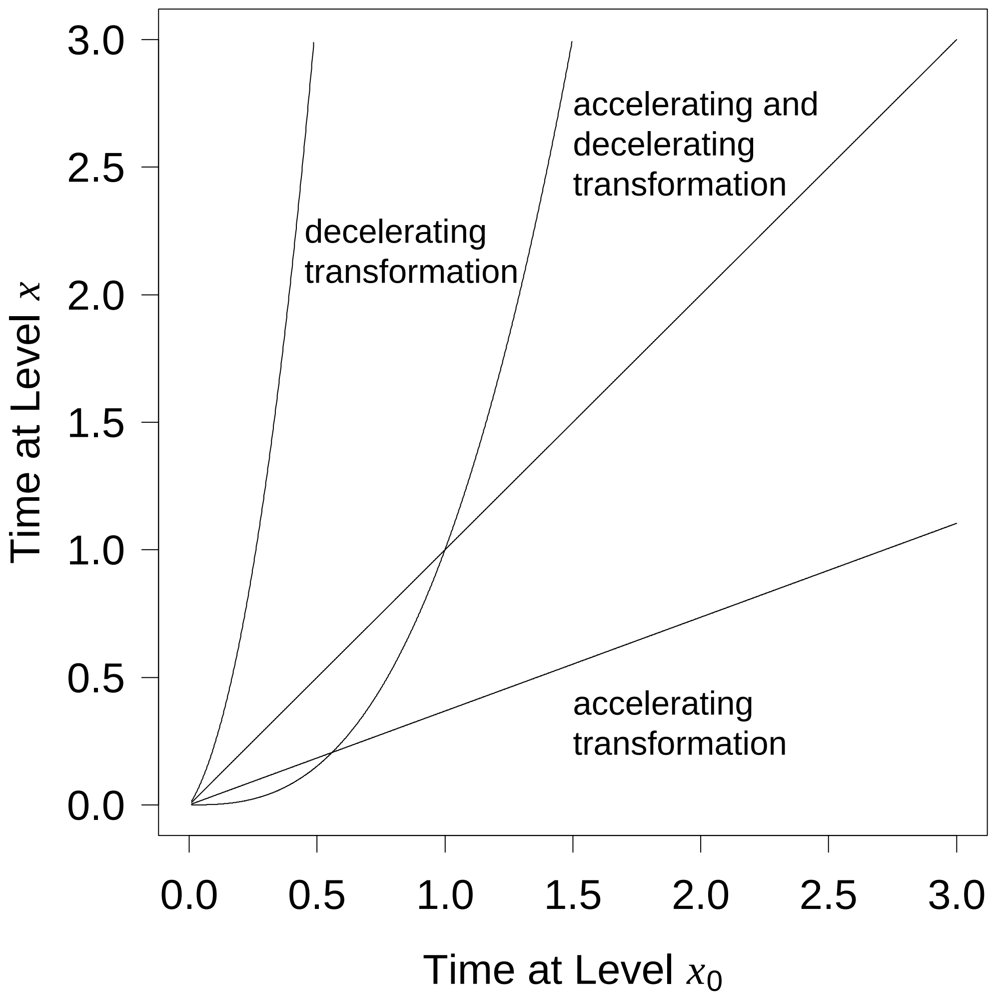
Picciotto Cox PH Model Survival Estimates
(on Weibull Probability Scales)



Picciotto Cox PH Model
Schoenfeld Residuals versus Time



General Failure Time Transformation Graph



General Failure Time Transformation Model

A general time transformation model is

$$T(\boldsymbol{x}) = \Upsilon\left[T(\boldsymbol{x}_0), \boldsymbol{x}\right]$$

where $\boldsymbol{x}_0$ are some **baseline** conditions.

We assume that $\Upsilon(t, \boldsymbol{x})$ satisfies the following conditions:

- $\Upsilon(t, \boldsymbol{x})$ is nonnegative, i.e., $\Upsilon(t, \boldsymbol{x}) \geq 0$ for all t and $\boldsymbol{x}$.
- For fixed $\boldsymbol{x}$, $\Upsilon(t, \boldsymbol{x})$ is monotone increasing in t.
- For all $\boldsymbol{x}$, $\Upsilon(0, \boldsymbol{x}) = 0$.
- For $\boldsymbol{x}_0$ the transformation is the identity transformation, i.e., $\Upsilon(t, \boldsymbol{x}_0) = t$ for all t.

General Failure Time Transformation Model

This general **assumed** transformation model implies:

- The distribution of $T(\boldsymbol{x})$ can be determined from the distribution of $T(\boldsymbol{x}_0)$ and $\boldsymbol{x}$. In particular, $t_p(\boldsymbol{x}) = \Upsilon\left[t_p(\boldsymbol{x}_0), \boldsymbol{x}\right]$ for $0 \leq p \leq 1$.
- In a plot of $T(\boldsymbol{x}_0)$ versus $T(\boldsymbol{x})$
  - ▶ $T(\boldsymbol{x})$ entirely below the diagonal line implies acceleration.
  - ▶ $T(\boldsymbol{x})$ entirely above the diagonal line implies deceleration.
  - ▶ Other $T(x)$s imply acceleration sometimes and deceleration other times. The cdfs of $T(\boldsymbol{x})$ and $T(\boldsymbol{x}_0)$ cross.
  - ▶ Scale time transformations and proportional hazard model are special cases.

Other Topics in Failure-Time Regression

- Models with non-location-scale distributions.
- Nonlinear relationship regression model.
- Fatigue-limit regression model.
- Special topics in regression: random effects models, non-parametric SAFT regression models, parametric PH regression models.