

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

Copyright 1998-2008 W. Q. Meeker and L. A. Escobar.
 Based on the authors' text *Statistical Methods for Reliability Data*, John Wiley & Sons Inc. 1998.

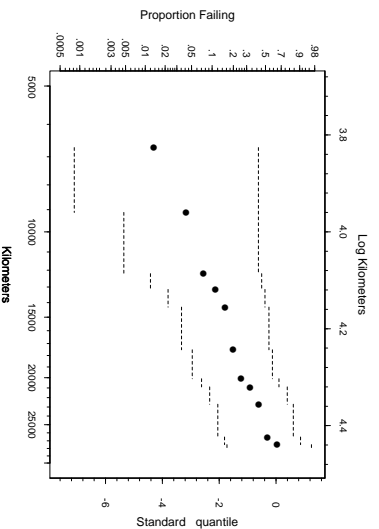
July 30, 2009
 8h 3min

8-1

- Illustrate likelihood-based methods for parametric models based on log-location-scale distributions (especially Weibull and Lognormal).
- Construct and interpret likelihood-ratio-based confidence intervals/regions for model parameters and for **functions** of model parameters.
- Construct and interpret normal-approximation confidence intervals/regions.
- Describe the advantages and pitfalls of assuming that the log-location-scale distribution shape parameter is known.

8-2

Weibull Probability Plot of the Shock Absorber Data



8-3

Weibull Distribution Model Likelihood for Right Censored Data

- The Weibull distribution model is

$$\Pr(T \leq t) = F(t; \mu, \sigma) = \Phi_{\text{sev}} \{[\log(t) - \mu]/\sigma\}.$$

- The likelihood has the form

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

$$\delta_i = \begin{cases} 1 & \text{if } t_i \text{ is an exact observation} \\ 0 & \text{if } t_i \text{ is a right censored observation} \end{cases}$$

$\phi_{\text{sev}}(z)$ is the standardized smallest extreme value density.

8-4

Lognormal Distribution Model Likelihood for Right Censored Data

- The lognormal distribution model is

$$\Pr(T \leq t) = F(t; \mu, \sigma) = \Phi_{\text{nor}} \{[\log(t) - \mu]/\sigma\}.$$

- The likelihood has the form

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

$$\delta_i = \begin{cases} 1 & \text{if } t_i \text{ is an exact observation} \\ 0 & \text{if } t_i \text{ is a right censored observation} \end{cases}$$

$\phi_{\text{nor}}(z)$ is the standardized normal density.

8-5

- Illustrate likelihood-based methods for parametric models based on log-location-scale distributions (especially Weibull and Lognormal).
- Construct and interpret likelihood-ratio-based confidence intervals/regions for model parameters and for **functions** of model parameters.
- Construct and interpret normal-approximation confidence intervals/regions.
- Describe the advantages and pitfalls of assuming that the log-location-scale distribution shape parameter is known.

8-2

Weibull Distribution Model Likelihood for Right Censored Data

- The Weibull distribution model is

$$\Pr(T \leq t) = F(t; \mu, \sigma) = \Phi_{\text{sev}} \{[\log(t) - \mu]/\sigma\}.$$

- The likelihood has the form

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

$$\delta_i = \begin{cases} 1 & \text{if } t_i \text{ is an exact observation} \\ 0 & \text{if } t_i \text{ is a right censored observation} \end{cases}$$

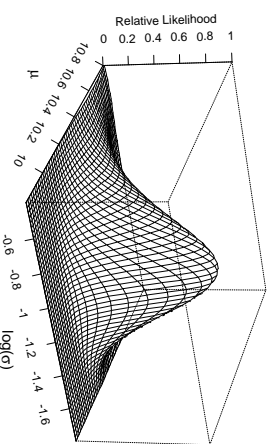
$\phi_{\text{sev}}(z)$ is the standardized smallest extreme value density.

8-4

Weibull Relative Likelihood for the Shock Absorber Data

ML Estimates: $\hat{\mu} = 10.23$ and $\hat{\sigma} = .3164$

$$R(\mu, \log(\sigma)) = L(\mu, \log(\sigma)) / L(\hat{\mu}, \log(\hat{\sigma}))$$

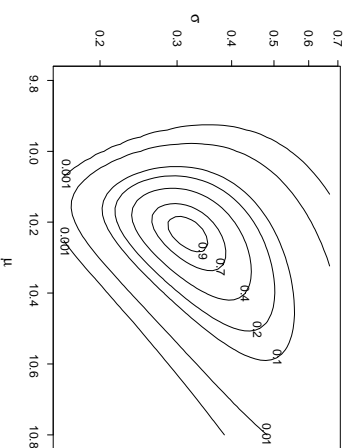


8-6

Weibull Relative Likelihood for the Shock Absorber Data

ML Estimates: $\hat{\mu} = 10.23$ and $\hat{\sigma} = .3164$

$$R(\mu, \sigma) = L(\mu, \sigma) / L(\hat{\mu}, \hat{\sigma})$$

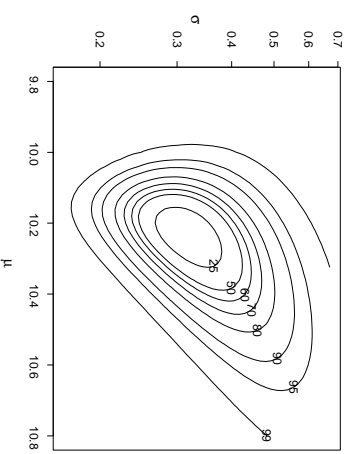


8 - 7

Weibull Likelihood-Based Joint Confidence Regions for μ and σ for the Shock Absorber Data

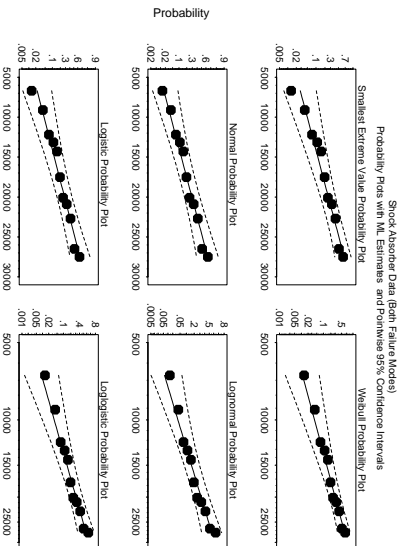
ML Estimates: $\hat{\mu} = 10.23$ and $\hat{\sigma} = .3164$

$$R(\mu, \sigma) > \exp \left[-\chi^2_{(1-\alpha; 2)} / 2 \right] = 100\alpha\%$$



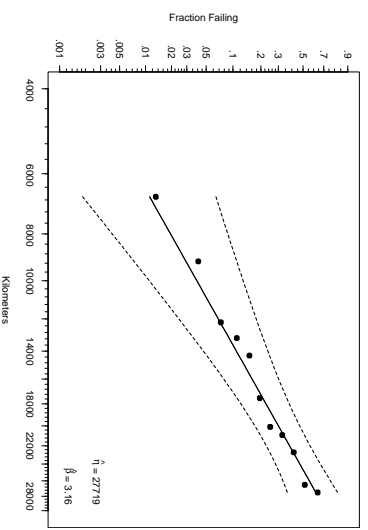
8 - 8

Six-Distribution ML Probability Plot of the Shock Absorber Data



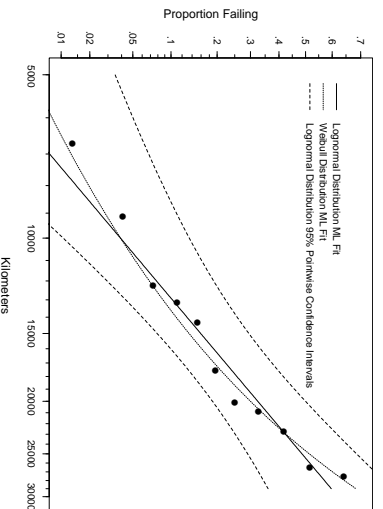
8 - 9

Weibull Probability Plot of Shock Absorber Failure Times (Both Failure Modes) with Maximum Likelihood Estimates and Normal-Approximation 95% Pointwise Confidence Intervals for $F^*(t)$



8 - 10

Lognormal Probability Plots of Shock Absorber Data with ML Estimates and Normal-Approximation 95% Pointwise Confidence Intervals for $F^*(t)$. The Curved Line is the Weibull ML Estimate.



8 - 11

Large-Sample Approximate Theory for Likelihood Ratios for Parameter Vector

- Relative likelihood for (μ, σ) is

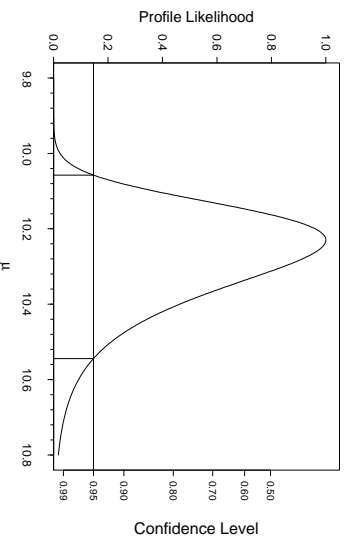
$$R(\mu, \sigma) = \frac{L(\mu, \sigma)}{L(\hat{\mu}, \hat{\sigma})}$$

- If evaluated at the true (μ, σ) , then, asymptotically, $-2\log[R(\mu, \sigma)]$ follows, a chisquare distribution with 2 degrees of freedom.
- General theory in the Appendix.

8 - 12

Weibull Profile Likelihood $R(\mu)$ ($\exp(\mu) \approx t_{.63}$)
for the Shock Absorber Data

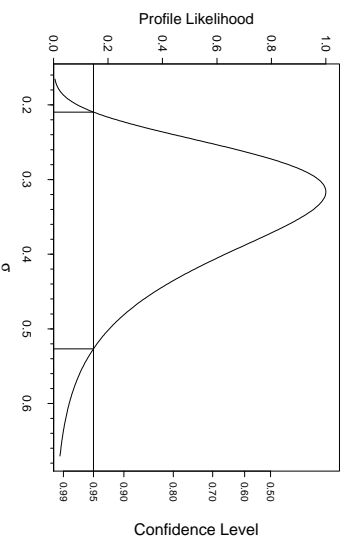
$$R(\mu) = \max_{\sigma} \left[\frac{L(\mu, \sigma)}{L(\hat{\mu}, \hat{\sigma})} \right]$$



8 - 13

Weibull Profile Likelihood $R(\sigma)$ ($\sigma = 1/\beta$)
for the Shock Absorber Data

$$R(\sigma) = \max_{\mu} \left[\frac{L(\mu, \sigma)}{L(\hat{\mu}, \hat{\sigma})} \right]$$



8 - 14

Large-Sample Approximate Theory for Likelihood Ratios for Parameter Vector Subset

Need: Inferences on subset θ_1 , from the partition $\theta = (\theta_1, \theta_2)'$.

- $k_1 = \text{length}(\theta_1)$.
- When $(\theta_1, \theta_2)' = (\mu, \sigma)$, profile likelihood for $\theta_1 = \mu$ is

$$R(\mu) = \max_{\sigma} \left[\frac{L(\mu, \sigma)}{L(\hat{\mu}, \hat{\sigma})} \right].$$
- If evaluated at the true $\theta_1 = \mu$, then, asymptotically, $-2 \log[R(\mu)]$ follows, a chisquare distribution with $k_1 = 1$ degrees of freedom.
- General theory in the Appendix.

8 - 15

Asymptotic Theory of Likelihood Ratios – Continued

- An approximate $100(1 - \alpha)\%$ likelihood-based confidence region for θ_1 is the set of all values of θ_1 such that

$$-2 \log[R(\theta_1)] < \chi_{1-\alpha, k_1}^2$$
 or, equivalently, the set defined by

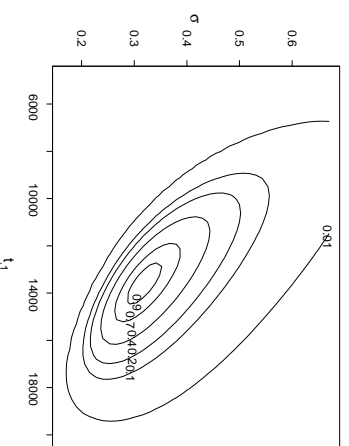
$$R(\theta_1) > \exp \left[-\chi_{1-\alpha, k_1}^2 / 2 \right].$$

- Transformation of θ_1 will not affect the confidence statement.
- Can improve the asymptotic approximation with simulation (only small effect except in very small samples).

8 - 16

Contour Plot of Weibull Relative Likelihood $R(t_1, \sigma)$ for the Shock Absorber Data (Parameterized with t_1 and σ)

$$R(t_1, \sigma) = L(t_1, \sigma) / L(\hat{t}_1, \hat{\sigma})$$



8 - 17

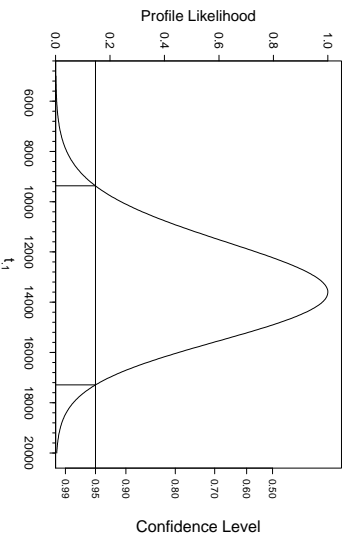
Confidence Regions and Intervals for Functions of μ and σ

- Likelihood approach can be applied to functions of parameters.
- Define the function of interest as one of the parameters, replacing one of the original parameters giving one-to-one reparameterization $g(\mu, \sigma) = [g_1(\mu, \sigma), g_2(\mu, \sigma)]$.
- Then follow previous procedure.
- Simple to implement if function and its inverse are easy to compute.

8 - 18

**Weibull Profile Likelihood $R(t_1)$
for the Shock Absorber Data**

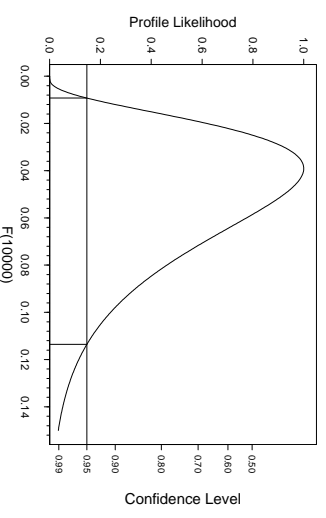
$$R(t_1) = \max_{\sigma} \left[\frac{L(t_1, \hat{\sigma})}{L(t_1, \hat{\sigma})} \right]$$



8 - 19

**Weibull Profile Likelihood $R[F(10000)]$
for the Shock Absorber Data**

$$R[F(10000)] = \max_{\sigma} \left\{ \frac{L[F(10000), \hat{\sigma}]}{L[F(10000), \hat{\sigma}]} \right\}$$



8 - 20

Asymptotic Theory of ML Estimation

Let $\hat{\theta}$ denote the ML estimator of θ .

- If evaluated at the true value of θ , then asymptotically, (large samples) $\hat{\theta}$ has a MVN($\theta, \Sigma_{\hat{\theta}}$) and thus the **Wald** statistic

$$(\hat{\theta} - \theta)' \left[\Sigma_{\hat{\theta}} \right]^{-1} (\hat{\theta} - \theta)$$

has a chisquare distribution with k degrees of freedom, where k is the length of θ .

- Here, $\Sigma_{\hat{\theta}} = I_{\hat{\theta}}^{-1}$ is the large sample approximate covariance matrix where the Fisher information matrix for θ is

$$I_{\theta} = E \left[- \frac{\partial^2 \mathcal{L}(\theta)}{\partial \theta \partial \theta'} \right].$$

8 - 21

Asymptotic Theory for Wald's Statistic

- Alternative asymptotic theory is based on the large-sample distribution of quadratic forms (Wald's statistic).

- Let $\Sigma_{\hat{\theta}}$ be a consistent estimator of $\Sigma_{\hat{\theta}}$, the asymptotic covariance matrix of $\hat{\theta}$. For example,

$$\Sigma_{\hat{\theta}} = \left[- \frac{\partial^2 \mathcal{L}(\theta)}{\partial \theta \partial \theta'} \right]^{-1}$$

where the derivatives are evaluated at $\hat{\theta}$.

- Asymptotically, the Wald statistic

$$w(\theta) = (\hat{\theta} - \theta)' \left[\Sigma_{\hat{\theta}} \right]^{-1} (\hat{\theta} - \theta)$$

when evaluated at the true θ , follows a chisquare distribution with k degrees of freedom, where k is the length of θ .

8 - 22

Asymptotic Theory for Wald's Statistic – Continued

- An approximate 100(1 - α)% confidence region for θ is the set of all values of θ in the ellipsoid

$$(\hat{\theta} - \theta)' \left[\Sigma_{\hat{\theta}} \right]^{-1} (\hat{\theta} - \theta) \leq \chi^2_{(1-\alpha, k)}.$$

- This is sometimes known as the normal-theory confidence region.

- Can specialize to functions or subsets of θ .

- Can transform to improve asymptotic approximation. Try to get a log likelihood with approximate quadratic shape.

8 - 23

**Normal-Approximation Confidence Intervals
for Model Parameters**

- Estimated variance matrix for the shock absorber data

$$\Sigma_{\hat{\mu}, \hat{\sigma}} = \begin{bmatrix} \text{Var}(\hat{\mu}) & \text{Cov}(\hat{\mu}, \hat{\sigma}) \\ \text{Cov}(\hat{\mu}, \hat{\sigma}) & \text{Var}(\hat{\sigma}) \end{bmatrix} = \begin{bmatrix} .01208 & .00399 \\ .00399 & .00535 \end{bmatrix}$$

- Assuming that $Z_{\hat{\mu}} = (\hat{\mu} - \mu) / \widehat{SE}_{\hat{\mu}} \sim \text{NOR}(0, 1)$ distribution, an approximate 100(1 - α)% confidence interval for μ is

$$[\hat{\mu}_L, \hat{\mu}_U] = \hat{\mu} \pm z_{(1-\alpha/2)} \widehat{SE}_{\hat{\mu}}$$

where $\widehat{SE}_{\hat{\mu}} = \sqrt{\text{Var}(\hat{\mu})}$.

- Assuming that $Z_{\log(\hat{\sigma})} = [\log(\hat{\sigma}) - \log(\sigma)] / \widehat{SE}_{\log(\hat{\sigma})} \sim \text{NOR}(0, 1)$ an approximate 100(1 - α)% confidence interval for σ is

$$[\hat{\sigma}_L, \hat{\sigma}_U] = [\hat{\sigma}^w/w, \hat{\sigma} \times w]$$

where $w = \exp \left[z_{(1-\alpha/2)} \widehat{SE}_{\log(\hat{\sigma})} \right]$ and $\widehat{SE}_{\log(\hat{\sigma})} = \sqrt{\text{Var}(\hat{\sigma})}$.

8 - 24

Normal-Approximation Confidence Intervals for Function $g_1 = g_1(\mu, \sigma)$

- ML estimate $\hat{g}_1 = g_1(\hat{\mu}, \hat{\sigma})$.
- Assuming $Z_{\hat{g}_1} = (\hat{g}_1 - g_1)/\widehat{SE}_{\hat{g}_1} \sim \text{NOR}(0, 1)$, an approximate $100(1 - \alpha)\%$ confidence interval for g_1 is

$$[\hat{g}_1, \tilde{g}_1] = \hat{g}_1 \pm z_{(1-\alpha/2)} \widehat{SE}_{\hat{g}_1},$$

where

$$\widehat{SE}_{\hat{g}_1} = \sqrt{\widehat{\text{Var}}(\hat{g}_1)} = \left[\left(\frac{\partial g_1}{\partial \mu} \right)^2 \widehat{\text{Var}}(\hat{\mu}) + \left(\frac{\partial g_1}{\partial \sigma} \right)^2 \widehat{\text{Var}}(\hat{\sigma}) + 2 \left(\frac{\partial g_1}{\partial \mu} \right) \left(\frac{\partial g_1}{\partial \sigma} \right) \widehat{\text{Cov}}(\hat{\mu}, \hat{\sigma}) \right]^{1/2}$$

- Partial derivatives evaluated at $\hat{\mu}, \hat{\sigma}$.
- General theory in the appendix.

8-25

Normal-Approximation Confidence Interval for $F(t_e; \mu, \sigma)$

Objective: Obtain a point estimate and a confidence interval for $\Pr(T \leq t_e) = F(t_e; \mu, \sigma)$ at a fixed and known point t_e .

- The ML estimates $\hat{\theta} = (\hat{\mu}, \hat{\sigma})$ and $\hat{\Sigma}_{\hat{\theta}}$ are available.
- The ML estimate for $F(t_e; \mu, \sigma)$ is

$$\hat{F} = F(t_e; \hat{\mu}, \hat{\sigma}) = \Phi(\hat{\zeta}_e)$$

where $\hat{\zeta}_e = [\log(t_e) - \hat{\mu}]/\hat{\sigma}$.

- In the context of Wald's theory, however, there are many ways to obtain a confidence interval for $F(t_e; \mu, \sigma)$.

8-26

Confidence Interval for $F(t_e; \mu, \sigma)$ —Continued

Note: Wald's confidence intervals depend on the parameterization used to derive the intervals.

For example, an approximate $100(1 - \alpha)\%$ confidence interval for $F(t_e; \mu, \sigma)$ can be obtained using:

- The asymptotic normality of $Z_{\hat{F}} = (\hat{F} - F)/\widehat{SE}_{\hat{F}}$
- $$[\hat{F}, \tilde{F}] = \hat{F}(t_e) \pm z_{(1-\alpha/2)} \widehat{SE}_{\hat{F}}.$$

- The asymptotic normality of $Z_{\logit(\hat{F})} = [\text{logit}(\hat{F}) - \text{logit}(F)]/\widehat{SE}_{\logit(\hat{F})}$

$$[\hat{F}, \tilde{F}] = \left[\frac{\hat{F}(t_e)}{\hat{F}(t_e) + (1 - \hat{F}(t_e)) \times w}, \frac{\hat{F}(t_e)}{\hat{F}(t_e) + (1 - \hat{F}(t_e))/w} \right]$$

where $w = \exp\{z_{(1-\alpha/2)} \widehat{SE}_{\hat{F}} / [\hat{F}(t_e)(1 - \hat{F}(t_e))]\}$.

8-27

Confidence Interval for $F(t_e; \mu, \sigma)$ —Continued

Comments:

- Often the confidence interval based on the asymptotic normality of $Z_{\hat{F}}$ has poor statistical properties caused by the slow convergence toward normality of $Z_{\hat{F}}$.
- The confidence interval based on the transformation $Z_{\logit(\hat{F})}$ can have better statistical properties if $Z_{\logit(\hat{F})}$ converges to normality faster than $Z_{\hat{F}}$.

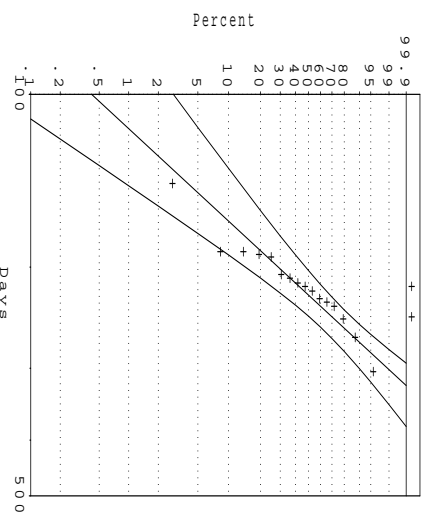
8-28

ML Estimates for Biomedical Data

Here we display SAS[®] Proc Reliability ML estimates (Weibull and lognormal) for the DMBA and the IUD data.

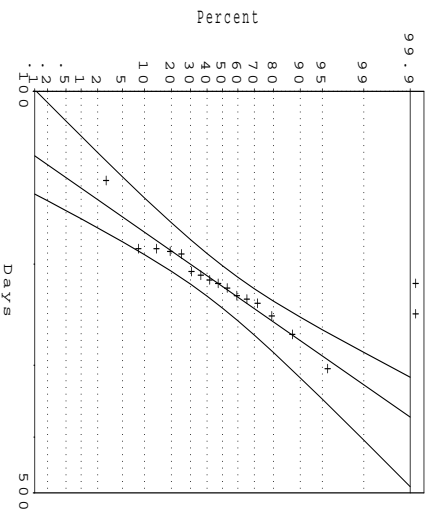
8-29

SAS[®] Proc Reliability Nonparametric and Weibull ML Estimate for DMBA Data with Parametric Pointwise Approximate 95% Confidence Intervals



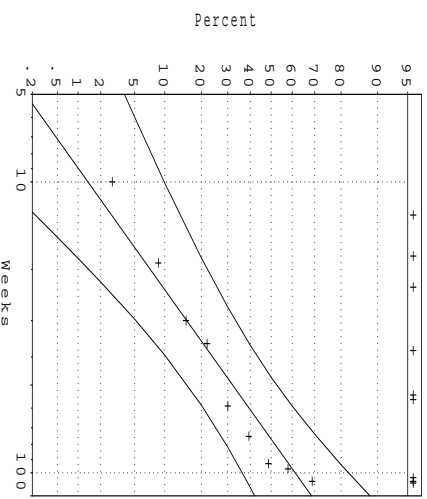
8-30

SAS® Proc Reliability
Nonparametric and Lognormal ML Estimate for
DWBA Data with Parametric Pointwise Approximate
95% Confidence Intervals



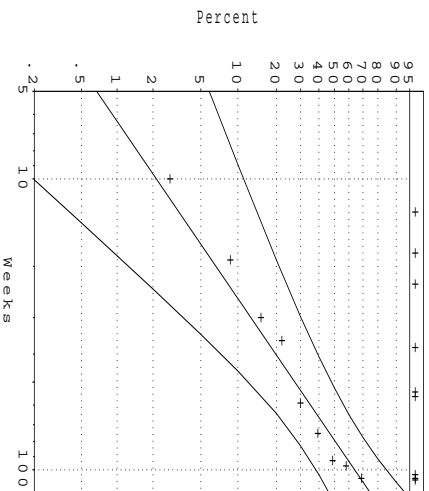
8-31

SAS® Proc Reliability
Lognormal ML Estimate for IUD Data with a set of
Pointwise Approximate 95% Confidence Intervals



8-32

SAS® Proc Reliability
Weibull ML Estimate for IUD Data with a set of
Pointwise Approximate 95% Confidence Intervals



8-33

Inference when σ (or Weibull β) is Given

- Simplifies problem. Only one parameter with r failures and t_1, \dots, t_n failures and censor times

$$\hat{\eta} = \left(\frac{\sum_{i=1}^n t_i^\beta}{r} \right)^{1/\beta}, \quad \text{se}_{\hat{\eta}} = \frac{\hat{\eta}}{\beta} \sqrt{\frac{1}{r}}$$

- Provides much more precision, especially with small r .
- If 0 failures can provide
 - ▶ Upper confidence bound on $F(t)$.
 - ▶ Lower confidence bound on t_p .
- Requires sensitivity analysis because β is in doubt.
- Danger of misleading inferences.

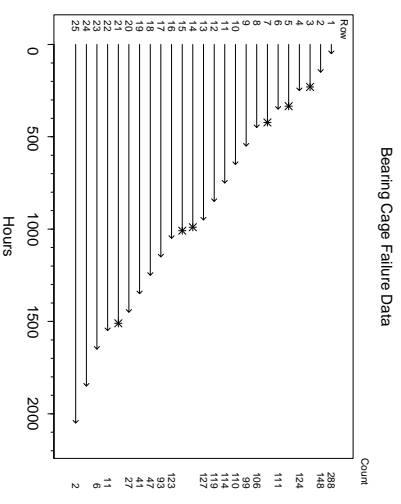
8-34

Bearing-Cage Fracture Field Data

- A population of $n = 1703$ units had been introduced into service over time and 6 failures have been observed.
- There is concern that the B10 design life specification of $t_{.1} = 8$ thousand hours was not being met.
- ML estimate is $\hat{t}_{.1} = 3.903$ thousand hours and an approximate 95% likelihood-ratio confidence interval for $t_{.1}$ is [2.093, 22.144] thousand hours.
- Management also wanted to know how many additional failures could be expected in the next year.

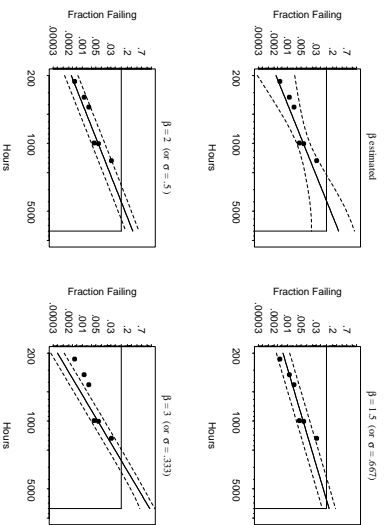
8-35

Bearing-Cage Fracture Data Event Plot



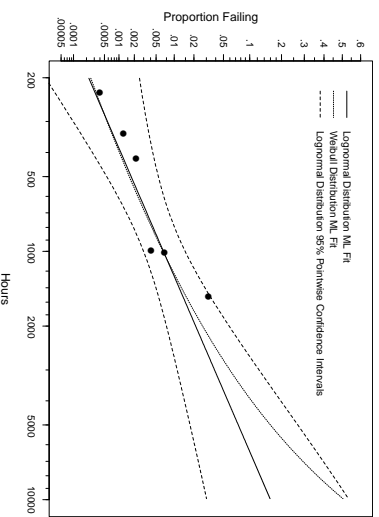
8-36

Weibull Probability Plots Bearing-Cage Fracture Data with Weibull ML Estimates and Sets of 95% Pointwise Confidence Intervals for $F(t)$ with β Estimated, and Assumed Known Values of $\beta = 1.5, 2, \text{ and } 3$.



8-37

Lognormal and Weibull Comparison Bearing-Cage Fracture Field Data Lognormal Probability Paper



8-38

Weibull/SEV Distribution with Given $\beta = 1/\sigma$ and Zero Failures

- ML Estimate for the Weibull Scale Parameter η Cannot be Computed Unless the Available Data Contains One or More Failures.
- For a sample of n units with running times t_1, \dots, t_n and no failures, a conservative $100(1-\alpha)\%$ lower confidence bound for η is

$$\eta = \left(\frac{2 \sum_{i=1}^n t_i^\beta}{\chi^2_{1-\alpha; 2}} \right)^{1/\beta}$$

- The lower bound η can be translated into an lower confidence bound for functions like t_p for specified p or a upper confidence bound for $F(t_e)$ for a specified t_e .

8-39

Component A Safe Data

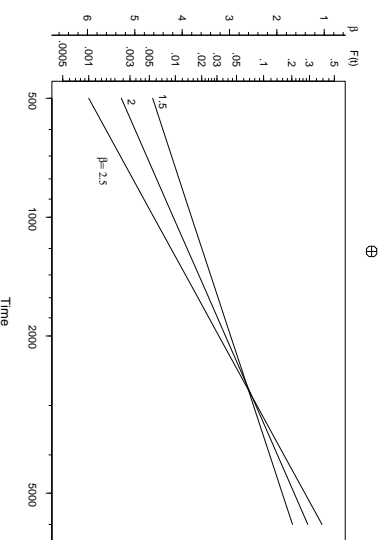
- A metal component in a ship's propulsion system fails from fatigue-caused fracture.
- Because of persistent reliability problems, the component was redesigned to have a longer service life.
- Previous experience suggests that the Weibull shape parameter is near $\beta = 2$, and almost certainly between 1.5 and 2.5.
- Newly designed components were put into service during the past year and no failures have been reported.

Hours:	500	1000	1500	2000	2500	3000	3500	4000
Number of Units:	10	12	8	9	7	9	6	3

Staggered entry data, with no reported failures.

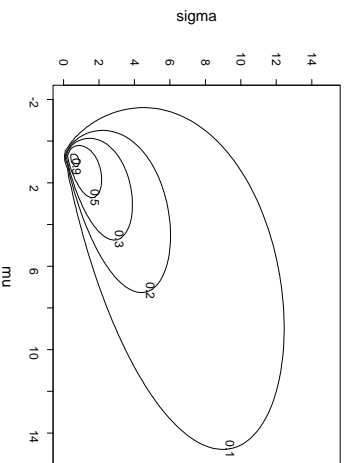
8-40

Weibull Model 95% Upper Confidence Bounds on $F(t)$ for Component-A with Different Fixed Values for the Weibull Shape Parameter



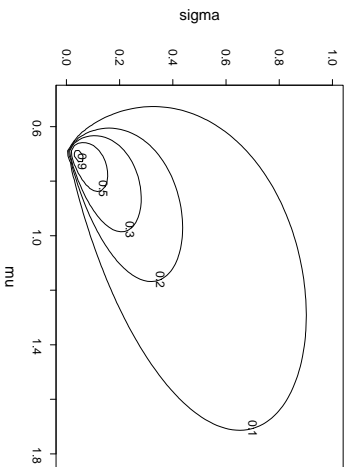
8-41

Relative Weibull Likelihood with One Failure at 1 and One Survivor at 2



8-42

Relative Weibull Likelihood with One Failure at 1.9 and One Survivor at 2



8-43

Regularity Conditions

- Each technical result (e.g., asymptotic distribution of an estimator) has its own set of conditions on the model (see Lehmann 1983, Rao 1973).

- Frequent reference to Regularity Conditions which give rise to simple results.

- For special cases the regularity conditions are easy to state and check. For example, for some location-scale distributions the needed conditions are:

$$\lim_{z \rightarrow -\infty} \frac{z^2 \phi^2(z)}{\Phi(z)} = 0$$

$$\lim_{z \rightarrow +\infty} \frac{z^2 \phi^2(z)}{1 - \Phi(z)} = 0.$$

- In **non-regular** models, asymptotic behavior is more complicated (e.g., behavior depends on θ), but there are still useful asymptotic results.

8-44

Regularity Conditions – Continued

Some **typical** regularity conditions include:

- Support does not depend on unknown parameters.
- Number of parameters does not grow too fast with n .
- Continuous derivatives of log likelihood (w.r.t. θ).
- Bounded derivatives of likelihood.
- Can exchange the order of differentiation of log likelihood w.r.t. θ and integration w.r.t. data.
- Identifiability.

8-45

Other Topics Related to Parametric Likelihood Covered in the Book

- Truncated data (Chapter 11).
- Threshold parameters (Chapter 11).
- Other distributions (e.g., gamma) (Chapter 11).
- Bayesian methods (Chapter 14).
- Multiple failure modes (Chapter 15).

8-46