Chapter 11

Parametric Maximum Likelihood: Other Models

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

Copyright 1998-2001 W. Q. Meeker and L. A. Escobar.
Based on the authors' text *Statistical Methods for Reliability Data*, John Wiley & Sons Inc. 1998.

July 18, 2002
12h 25min

Parametric Maximum Likelihood: Other Models
Chapter 11 Objectives

- ML estimation for the gamma and the extended generalized gamma (EGENG) distributions.
- ML estimation for the BISA, IGAU, and GOMA distributions.
- ML estimation for the limited failure population model.
- ML estimation for truncated data (or data from truncated distributions)
- ML estimation for threshold-parameter distributions like the 3-parameter lognormal and the 3-parameter Weibull distributions (using generalized threshold-scale or GETS distribution)

Fitting the Other Distributions and Models

- Likelihood principles similar to location-scale distributions.

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} L_i(\boldsymbol{\theta}; \text{data}_i) = \prod_{i=1}^{n} \left[f(t_i; \boldsymbol{\theta})\right]^{\delta_i} \left[1 - F(t_i; \boldsymbol{\theta})\right]^{1-\delta_i}$$

where $\text{data}_i = (t_i, \delta_i)$,

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

and $F(t_i; \boldsymbol{\theta})$ and $f(t_i; \boldsymbol{\theta})$ are the specified distribution's cdf and pdf, respectively.

- For some non-location-scale distributions (e.g. GETS) the density approximation breaks down and one should use the actual interval probability instead.

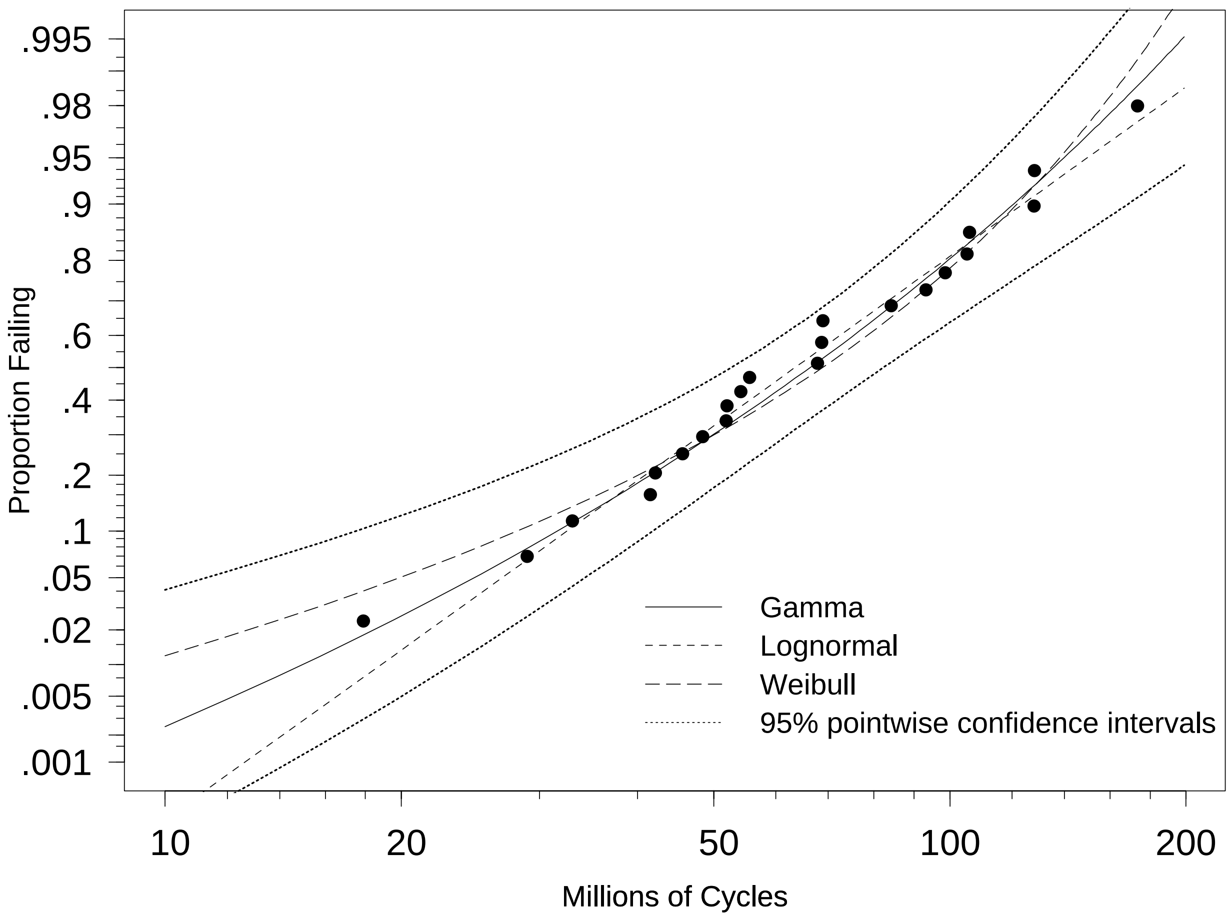
- Left censored and interval censored observations also could be included, as described in Chapter 2.

Confidence Intervals for Other Distributions and Models

Confidence intervals and regions similar to location-scale distributions.

- Normal approximation confidence intervals (using the delta method and appropriate transformations) are simple and are adequate in large samples or for rough approximations.
- Profile likelihood and corresponding intervals provide useful insight into the information available about a particular parameter or functions of parameters.
- Bootstrap and simulation-based intervals will generally provide confidence intervals with excellent approximations to nominal coverage probabilities, but will require more computer time (and may not be available in commercial software).

Lognormal Probability Plot of the Bearing Failure Data, Comparing ML Estimates of the Gamma, Lognormal and Weibull Distributions. Approximate 95% Pointwise Confidence Intervals for the Gamma Cdf are Also Shown



Fitting the Gamma Distribution

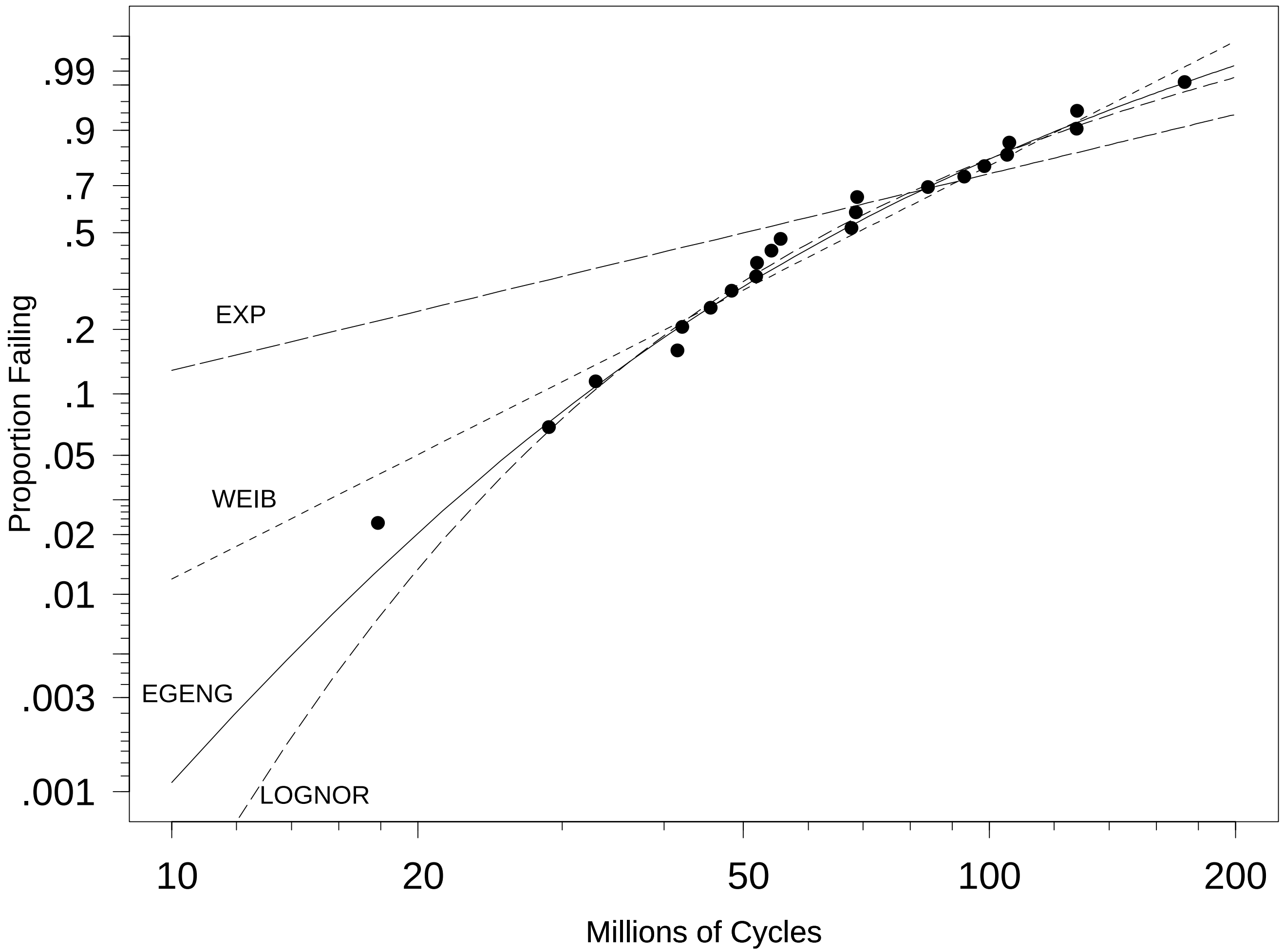
- Scale and shape parameter estimated.

- For the bearing data, the gamma, lognormal and Weibull distributions are similar over the range of the data.

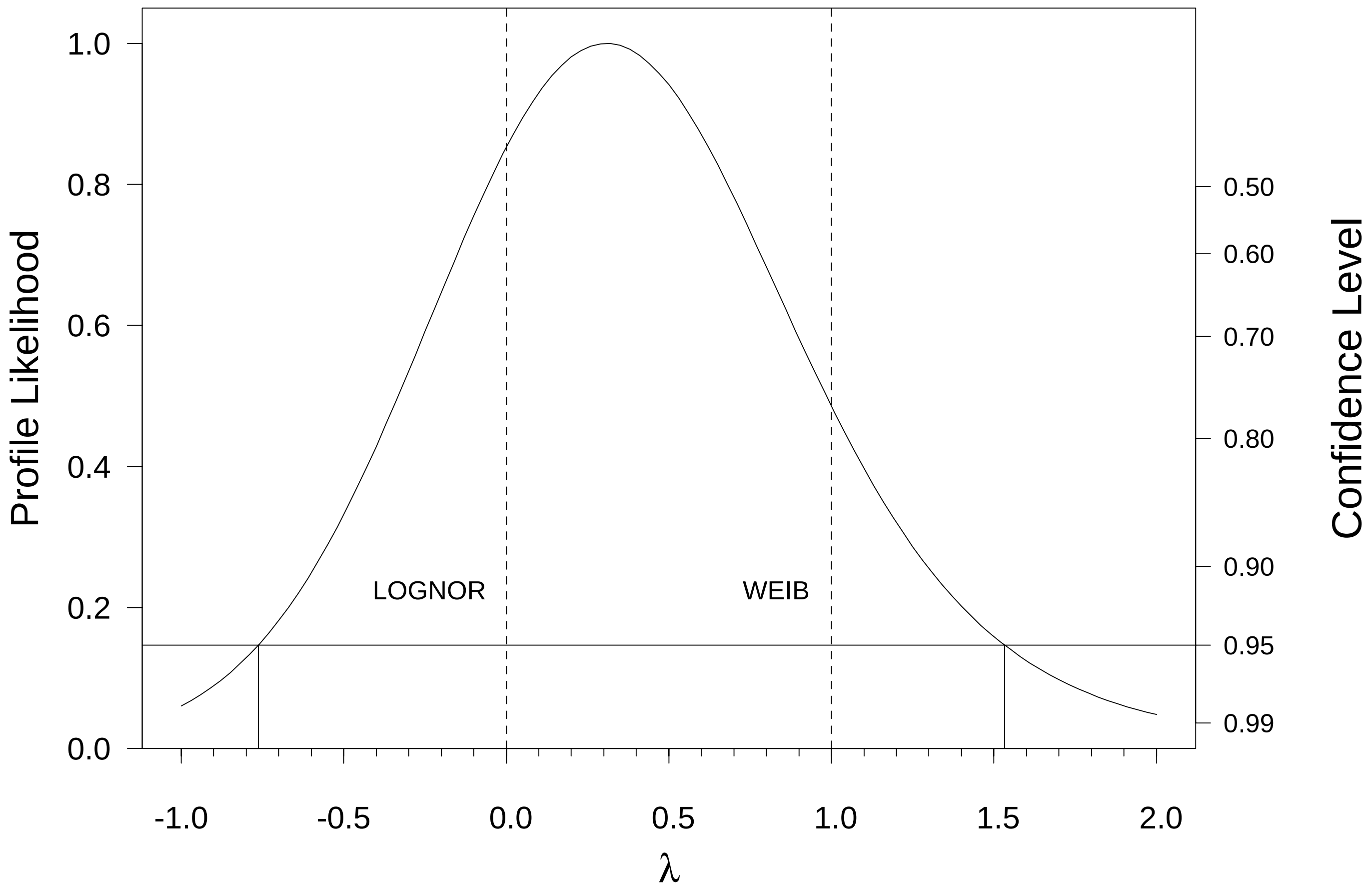
Fitting the Extended Generalized Gamma (EGENG) Distribution

- $T \sim \text{EGENG}(\mu, \sigma, \lambda)$
- Special cases: Weibull ($\lambda = 1$), Lognormal ($\lambda = 0$), and Gamma ($\theta = \lambda^2 \exp(\mu), \sigma = \lambda, \kappa = 1/(\lambda)^2$.
- A more flexible curve for the data
- Can use EGENG to see if there is evidence for one distribution over the other
- For the bearing data, the EGENG distribution provides a compromise between lognormal and Weibull.
- The EGENG, lognormal and Weibull agree well within the range of the data. Important deviations in the lower tail of the distribution illustrate the danger or extrapolation.
- The profile likelihood shows that the data do not, in this case, provide strong evidence for one distribution over the other.

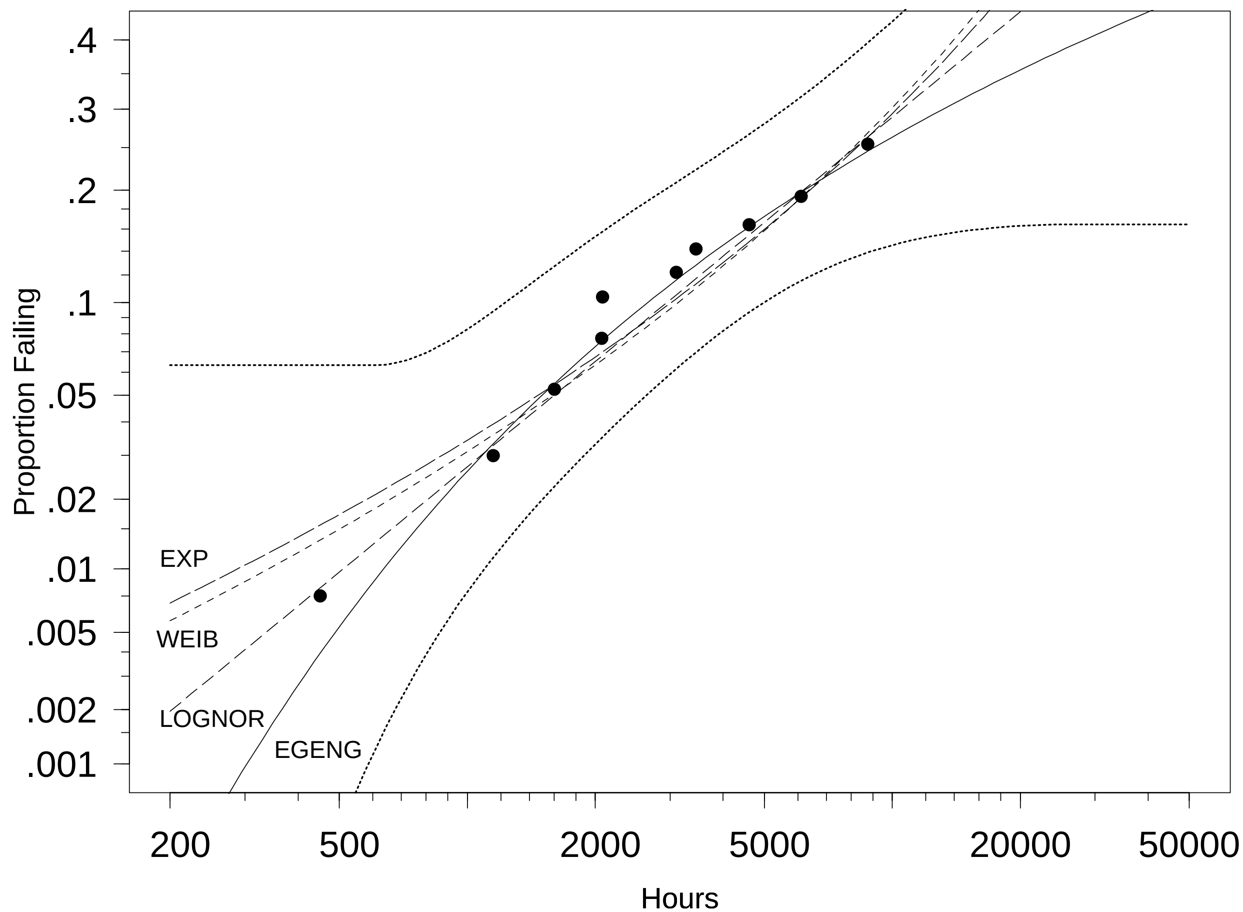
Weibull Probability Plot of the Bearing Failure Data Showing Exponential, Weibull, Lognormal, and Generalized Gamma ML Estimates of $F(t)$



Profile Likelihood Plot for EGENG λ for the Bearing Failure Data Showing Weibull and Lognormal Distributions as Special Cases



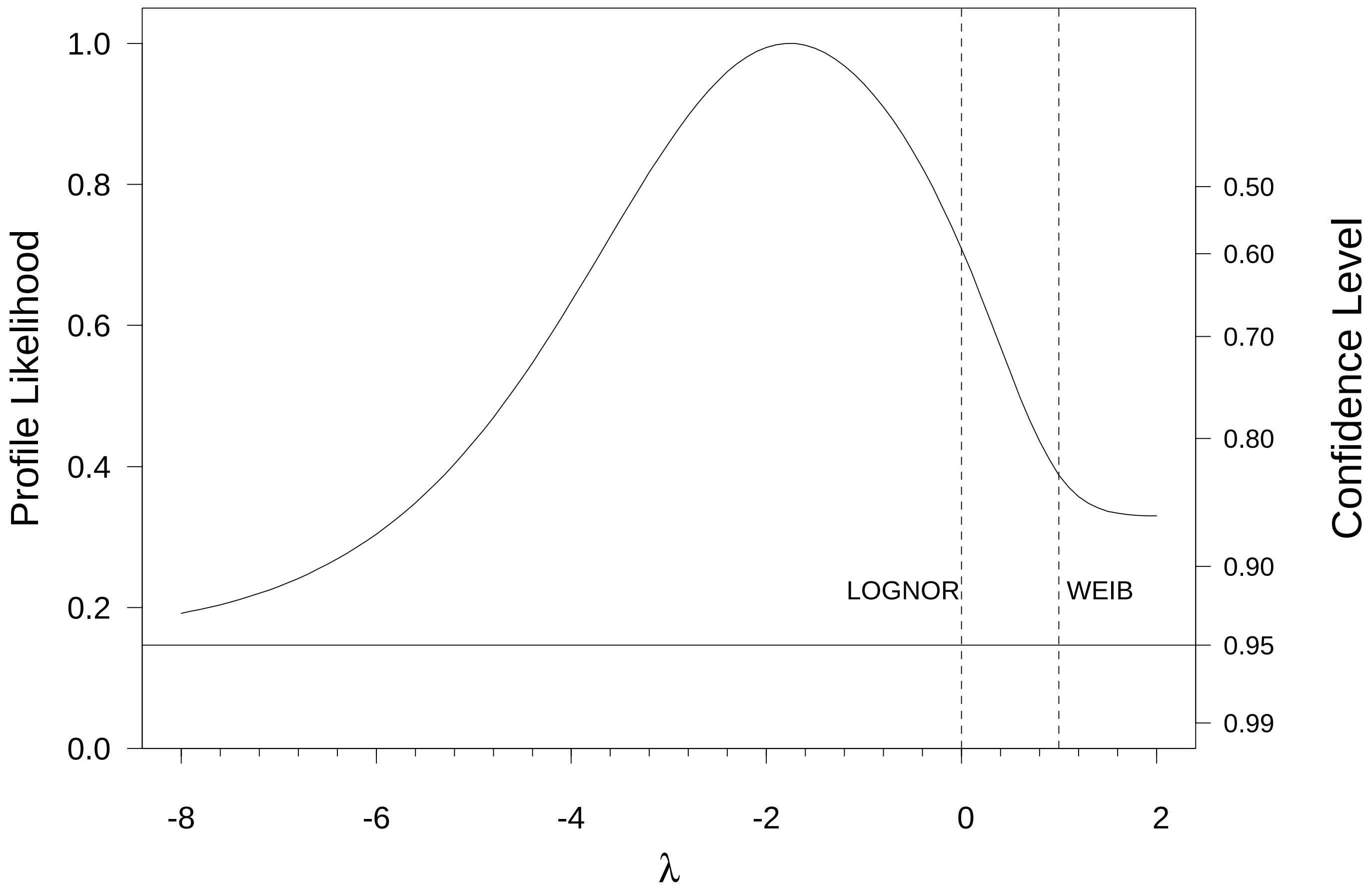
Lognormal Probability Plot of the Fan Failure Data Showing Generalized Gamma ML Estimates and Corresponding Approximate 95% Pointwise Confidence Intervals for $F(t)$ Along with Exponential, Weibull, and Lognormal ML Estimates of $F(t)$



Fitting the EGENG Distribution to the Fan Data

- Only 12 failures out of 70 units (multiple censoring).
- Lognormal fits the data well. Weibull and exponential also fit the data reasonably well. Can EGENG do better?
- The EGENG has a larger likelihood than the other distributions, but the difference is statistically unimportant.
- Comparison shows that the position of the smallest observation does not have much influence on the fit (small order statistics have a large amount of variability).
- Fitting a 3-parameter distribution to 12 failures is **overfitting**.

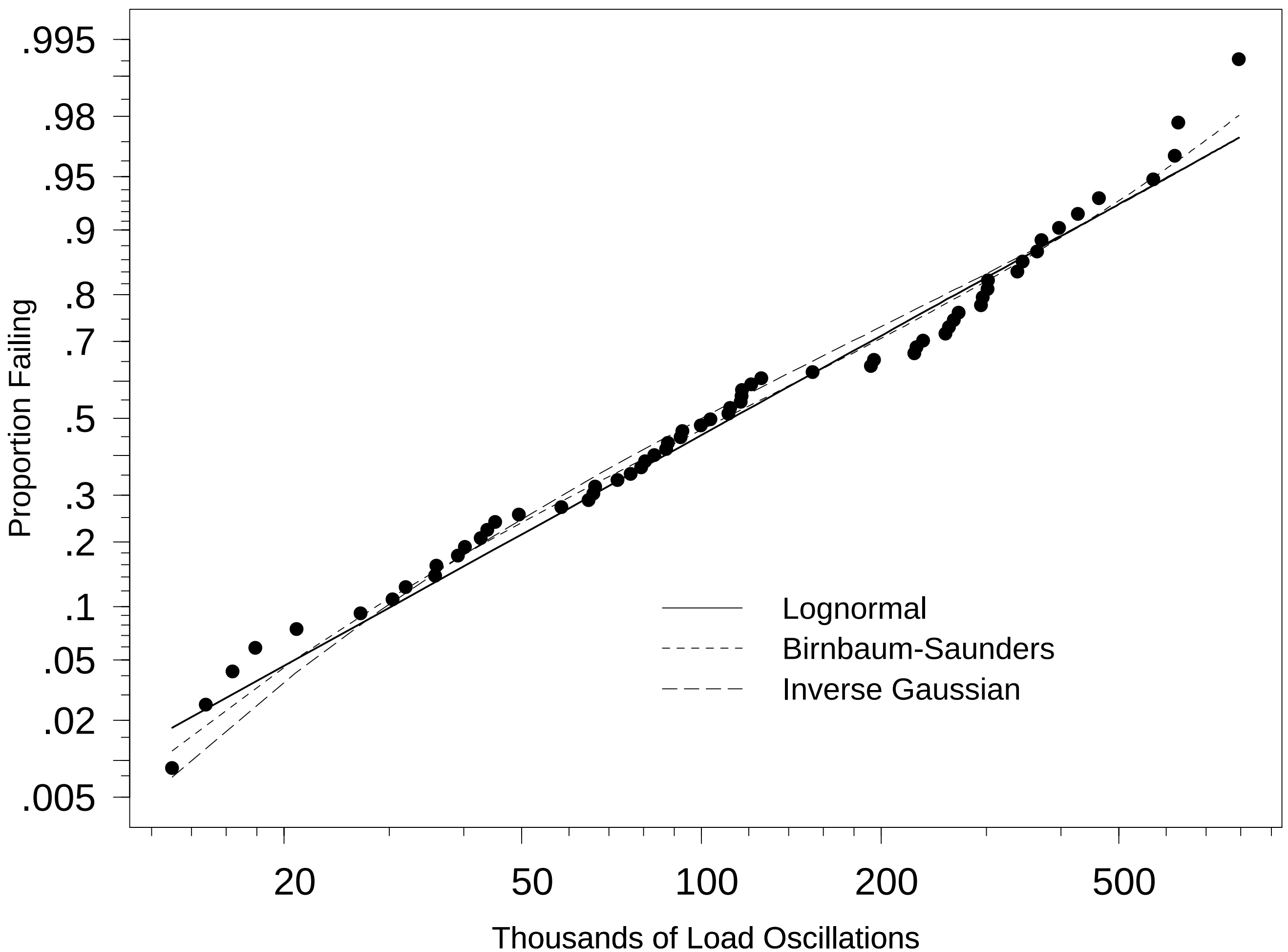
Profile Likelihood Plot for EGENG λ for the Fan Failure Data Showing Weibull and Lognormal Distributions as Special Cases



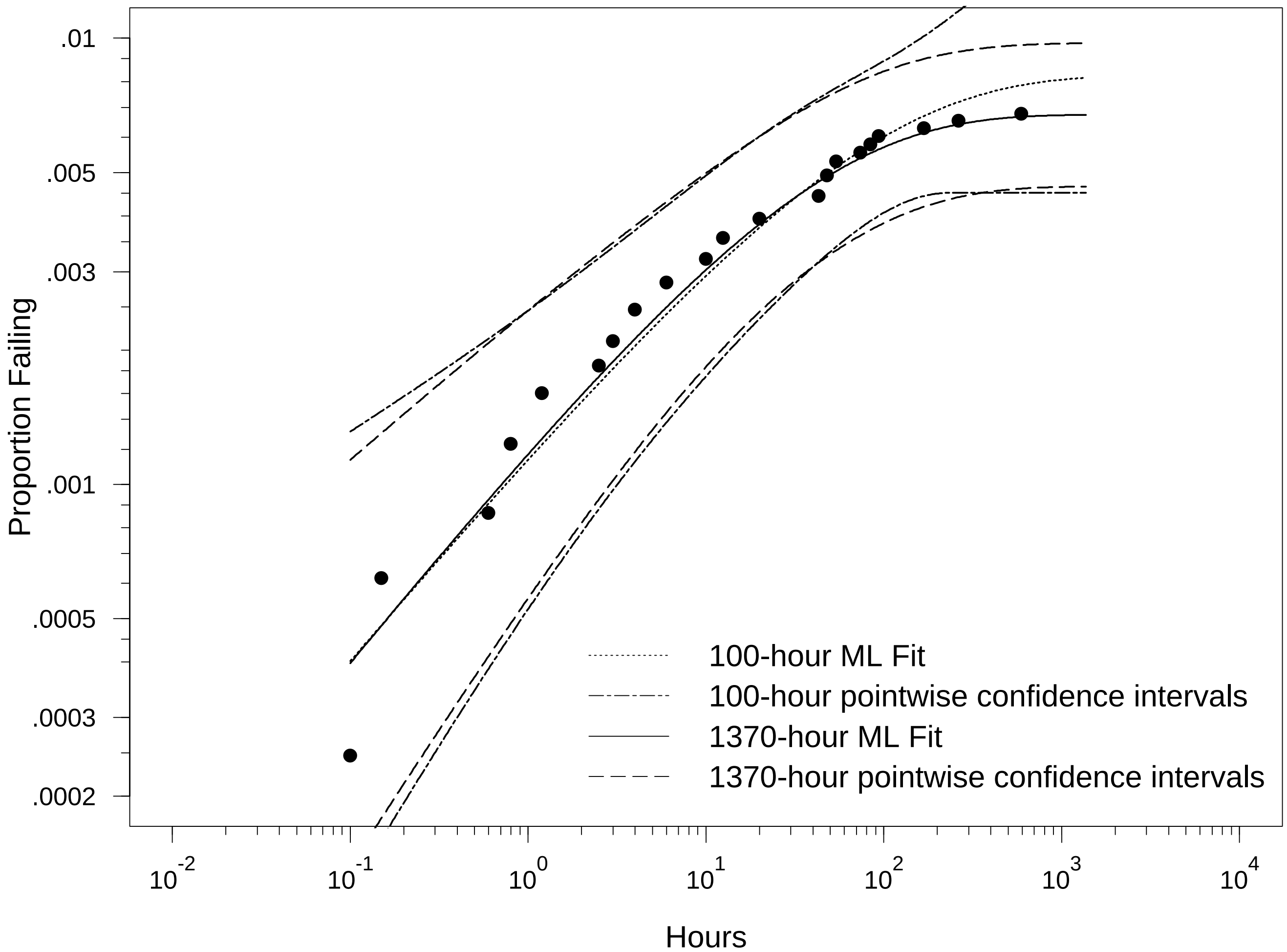
Fitting the BISA and IGAU Distributions

- Distributions motivated by similar degradation models
  - Inverse-Gaussian (IGAU) based on time of first crossing of a threshold for a **continuous-time** Brownian motion process with drift.
  - Birnbaum-Saunders (BISA) based on **discrete-time** growth of fatigue cracks until fracture.
- For some values of their parameters, these distributions are very similar to each other and to the lognormal distribution.

Lognormal Probability Plot of Yokobori's Fatigue Failure Data on Cylindrical Specimens at 52.658 ksi Showing Lognormal, BISA and IGAU Distribution ML Estimates



Weibull Probability Plot of Integrated Circuit Failure-Time Data with ML Estimates of the Weibull/LFP Model After 1370 Hours and 100 Hours of Testing



Limited Failure Population (LFP) Model

- Only a small proportion (p) of the population is susceptible to failure.
- The Weibull/LFP model is

$$\Pr(T \leq t) = pF(t;\mu,\sigma) = p\Phi_{\text{sev}}\left[\frac{\log(t)-\mu}{\sigma}\right].$$

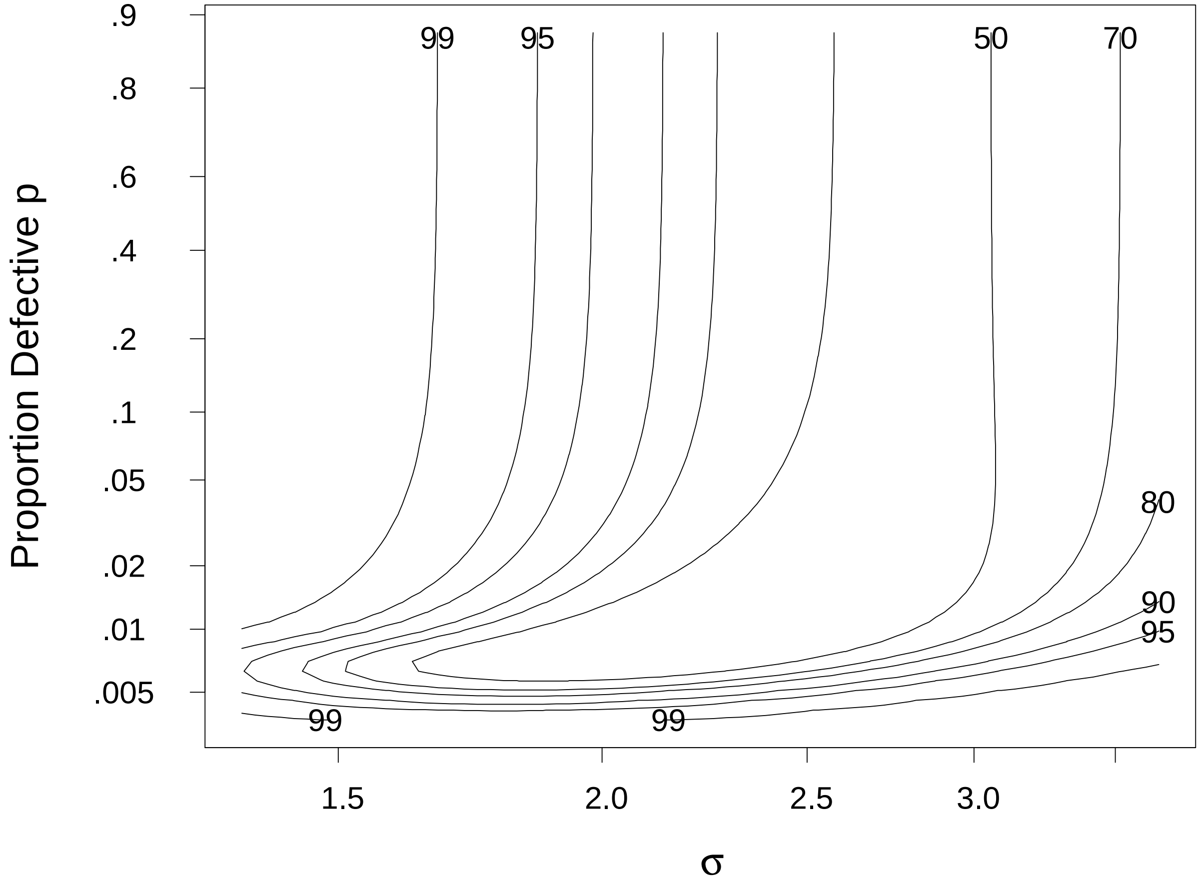
  Similar for lognormal or other distributions.

- ML methods work. The likelihood has the form

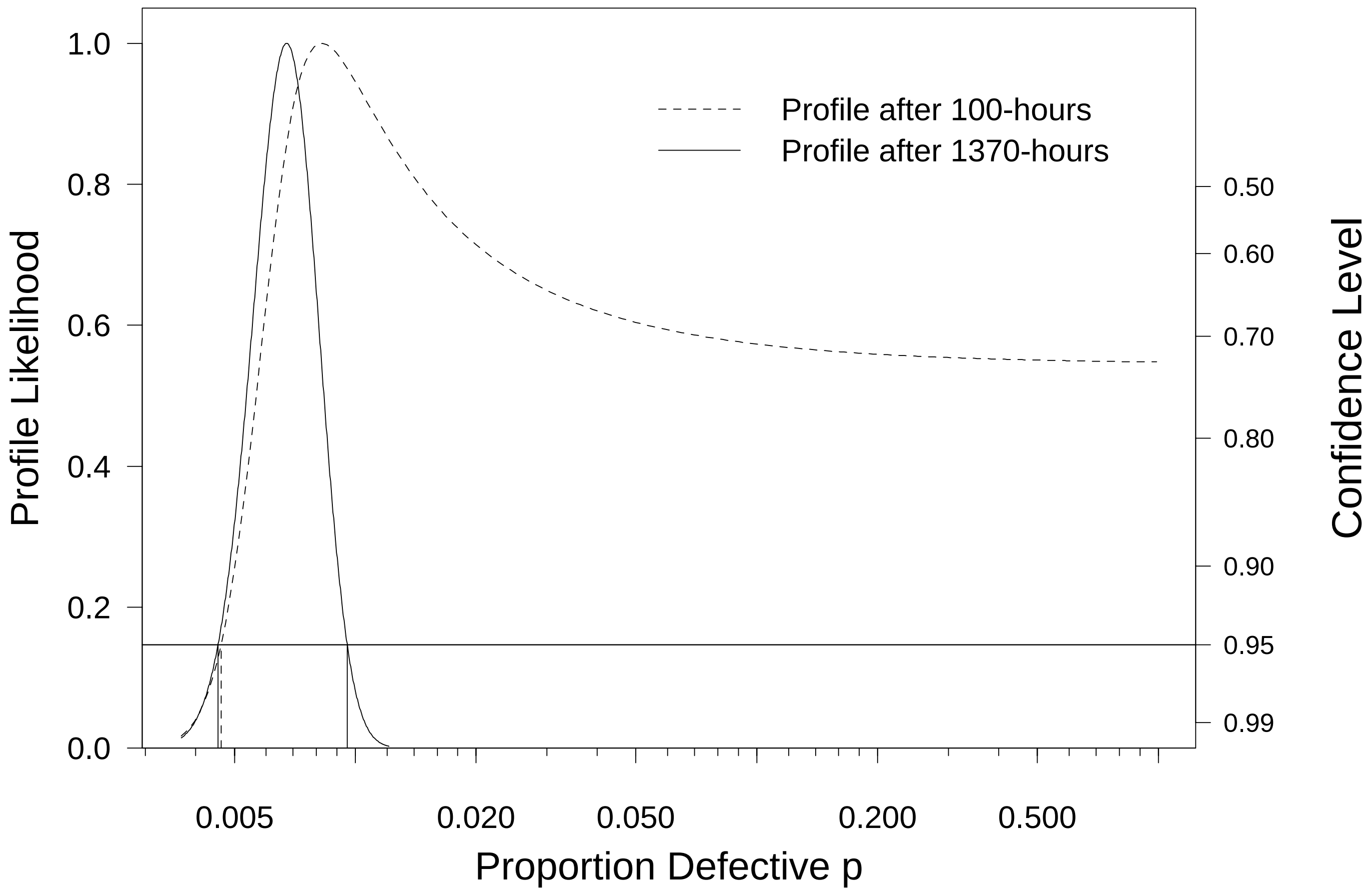
$$\begin{aligned} L(\mu,\sigma,p) \; &= \; \prod_{i=1}^{n}\left\{\frac{p}{\sigma}\,\phi_{\text{sev}}\left[\frac{\log(t_i)-\mu}{\sigma}\right]\right\}^{\delta_i} \times \\ & \quad\; \left\{1 - p\Phi_{\text{sev}}\left[\frac{\log(t_i)-\mu}{\sigma}\right]\right\}^{1-\delta_i}. \end{aligned}$$

  Need to test until a high proportion (e.g. 90% or more) of the susceptible subpopulation has failed. See Meeker (1987) for more details.

Approximate Joint Confidence Regions For the LFP Parameters p and $\log(\sigma)$ Based on a Two-Dimensional Profile Likelihood After 100 Hours of Testing



Comparison of Profile Likelihoods for p, the LFP Proportion Defective, After 1370 and 100 Hours of Testing



Relationship Between Wald and Profile Likelihood-Based Confidence Regions/Intervals

Result: Using the Wald (normal-theory) based interval is equivalent to using a quadratic approximation to the loglikelihood profile.

- See Meeker and Escobar (1995) for proof.
- Likelihood-based interval does not depend on transformation.
- Simulation and some theory suggests that the likelihood-based interval provides a better asymptotic approximation.
- Interval for the LFP parameter provides an extreme example of where the quadratic approximation breaks down.

Analysis of Truncated Data

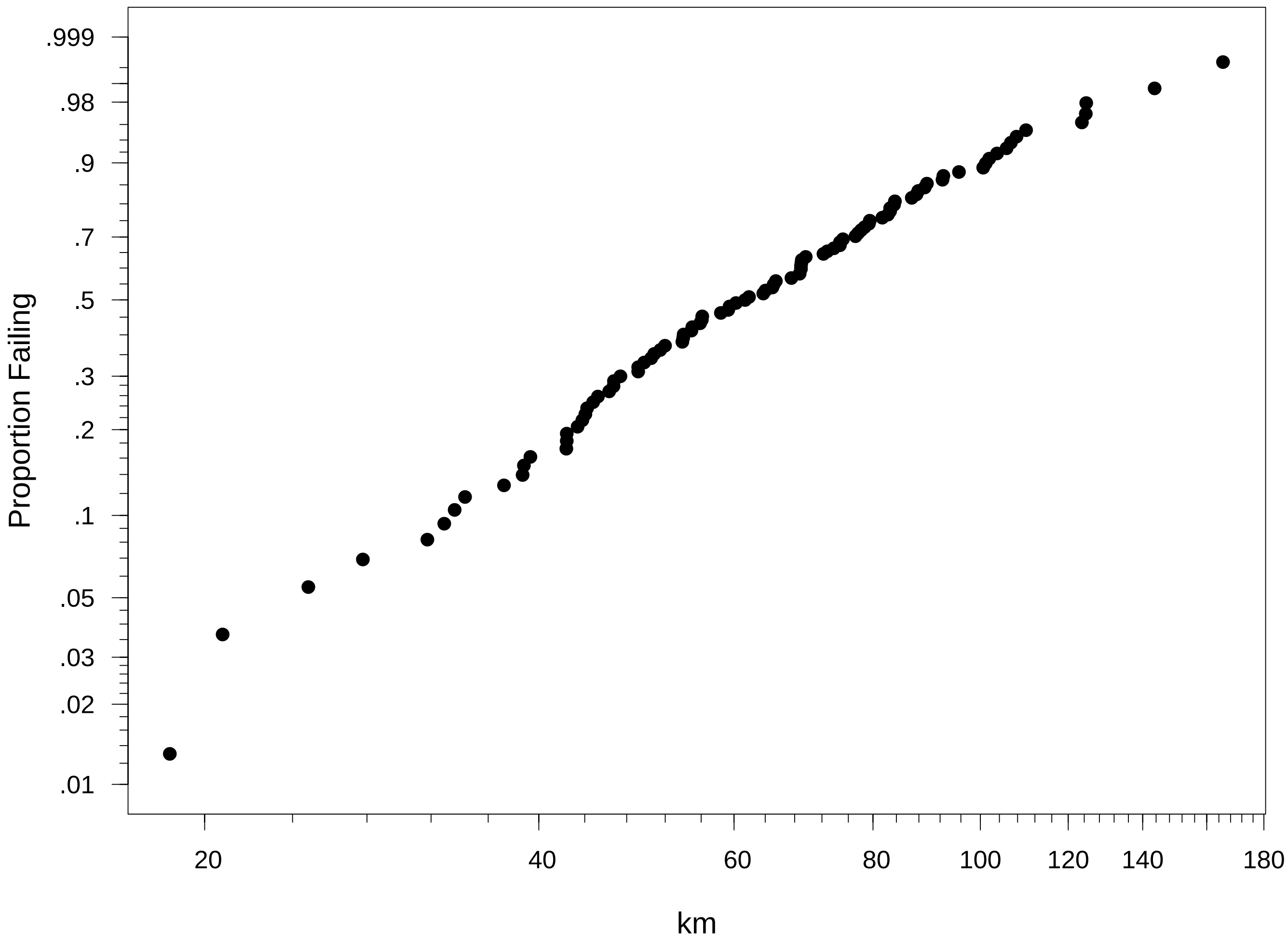
Some relevant topics in the analysis of truncated data include:

- Importance of distinguishing between truncated data and censored data.
- Nonparametric estimation with left-truncated data.
- ML estimation with left-truncated data.
- Nonparametric estimation and ML estimation with right (and left) truncated data.

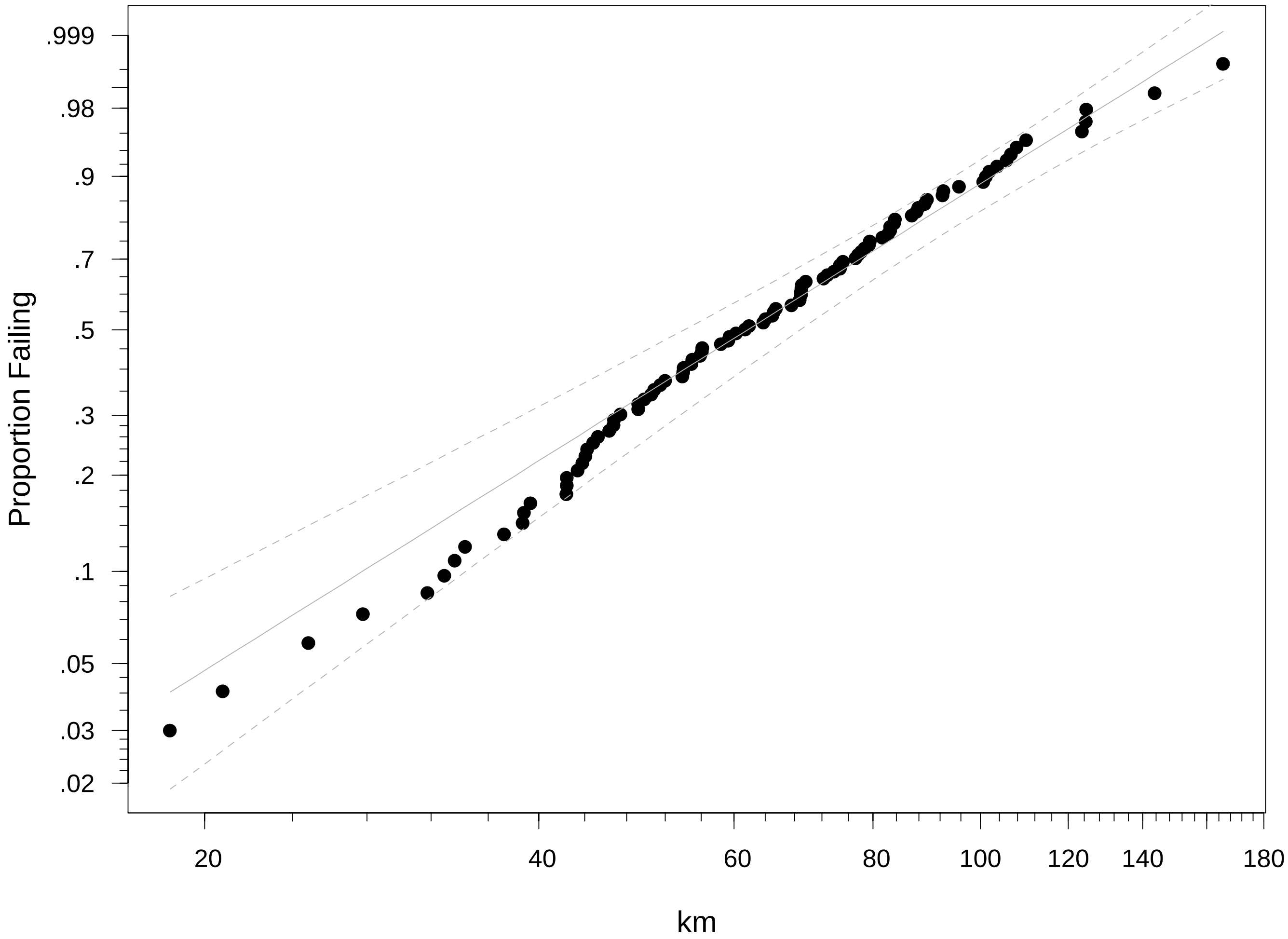
Distribution of Brake Pad Life from Observational Data

- Pad wear (W as a proportion of wear at the end of life) was measured and distance driven (V in thousands of km) was recorded on automobiles that came in for service. Data from Kalbfleisch and Lawless (1992).
- Time of failure for each pad was imputed from the observed wear rate as $Y = V/W$.
- Units having already had a pad replacement were omitted from the data. Thus, high-rate units are under represented in the sample.
- To analyze the data, each unit can be viewed as having been left-truncated at its observation time (if it had failed before its observation time, we would not know of the unit's existence because it would have been omitted from the sample).

Weibull Probability Plot of the Nonparametric Estimate of Brake Pad Life, Conditional on Failure After 6.951 Thousand km



Weibull Probability Plot of the Weibull-Adjusted Nonparametric Estimate of Brake Pad Life Distribution



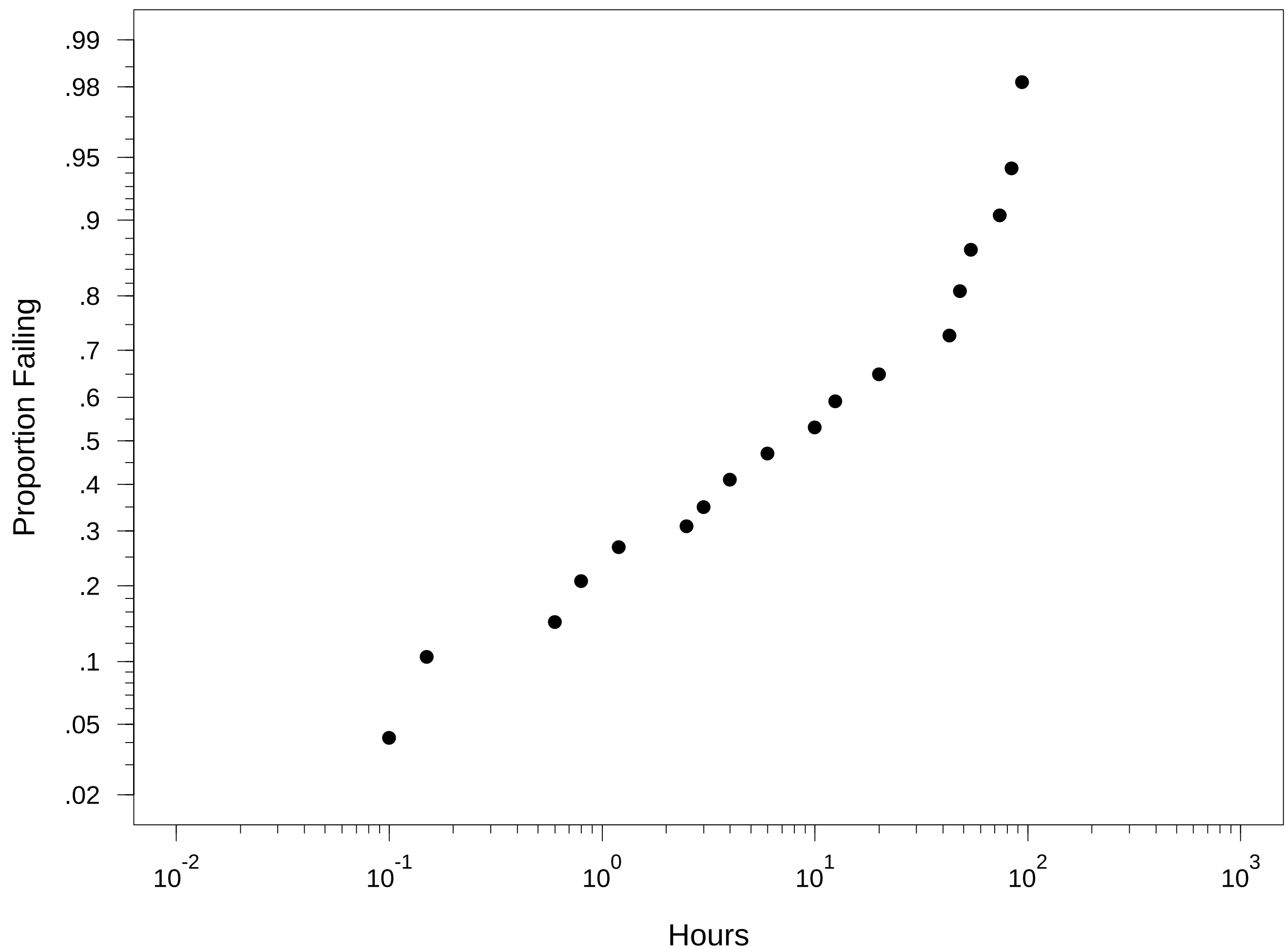
IC Failure Data from a Limited Failure Population

- Of the n =4,156 integrated circuits tested, there were 25 failures in the first 100 hours of testing.

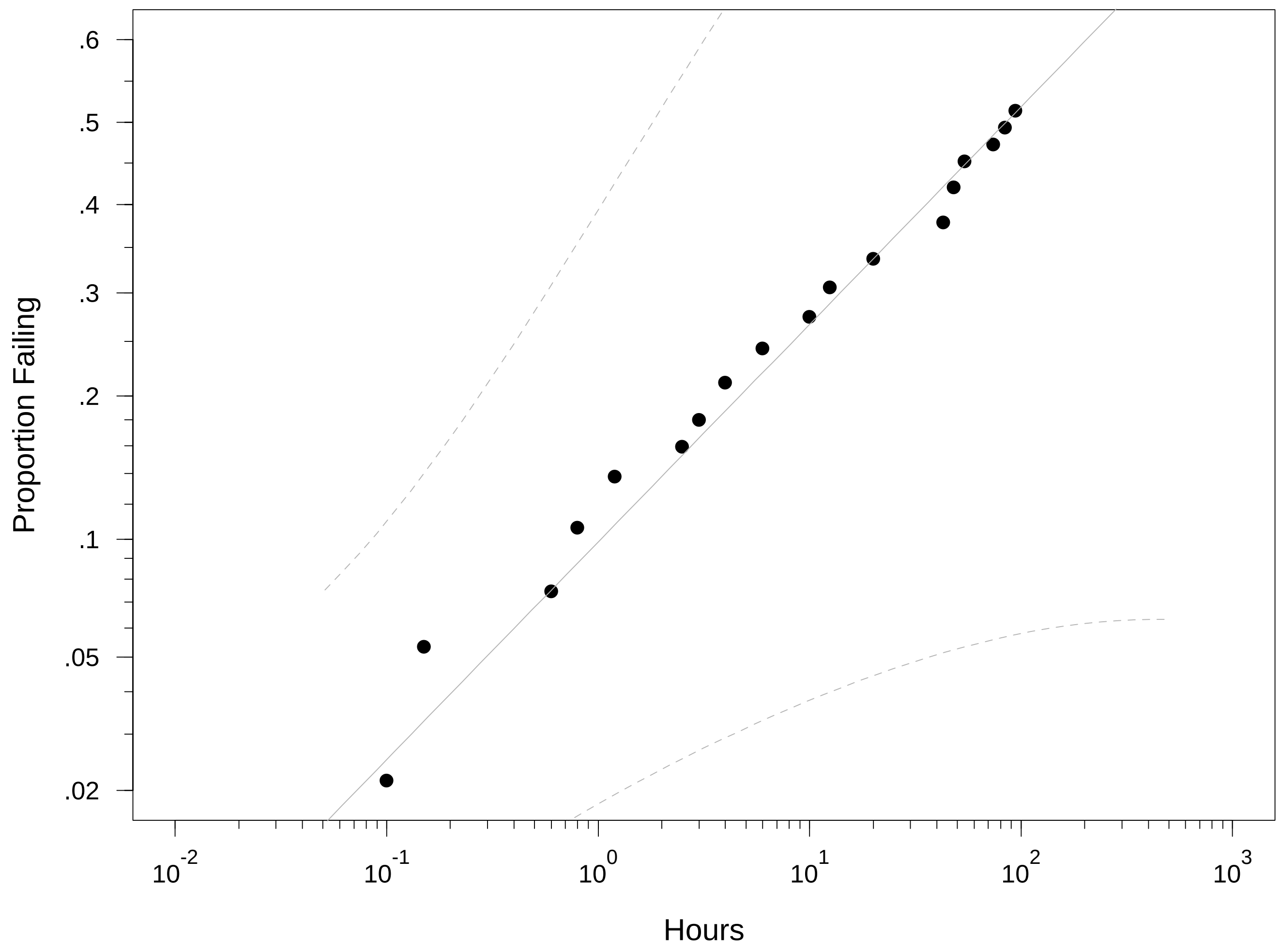
- The number of susceptible units in the sample is unknown.

- The 25 failures can be viewed as a sample from a distribution truncated on the right at 100 hours.

Lognormal Probability Plot of the Nonparametric Estimate of the IC Failure-Time Distribution Conditional on Failure Before 100 Hours



Lognormal Probability Plot of the Lognormal-Adjusted (Unconditional) Nonparametric Estimate of the IC Failure-Time Distribution



Three-Parameter Weibull and Lognormal Distributions

- Let γ be the threshold parameter. Then

$$\begin{aligned} F(t;\mu,\sigma,\gamma) &= \Phi\left[\frac{\log(t-\gamma)-\mu}{\sigma}\right] \\ f(t;\mu,\sigma,\gamma) &= \frac{1}{\sigma(t-\gamma)}\,\phi\left[\frac{\log(t-\gamma)-\mu}{\sigma}\right] \end{aligned}$$

  for $t > \gamma$. Φ_{sev} and ϕ_{sev} are used for the Weibull distribution and Φ_{nor} and ϕ_{nor} are used for the lognormal distribution.

- Similarly, a threshold parameter can be added to other distributions for positive random variables.

Inferences for 3-Parameter Weibull or Lognormal Distributions Assuming that Threshold γ is Known

- If γ can be assumed to be known, we can subtract γ from all times and fit the two-parameter Weibull distribution to estimate μ and σ.

- Need to adjust inferences accordingly (e.g., add γ back into estimates of quantiles or subtract γ from times before computing probabilities).

- Similar methods can be used for other distributions for positive random variables.

- ML works if used correctly.

Fitting the Three-Parameter Weibull Distribution Likelihood with Unknown Threshold γ

Two possible problems with ML that need to be avoided

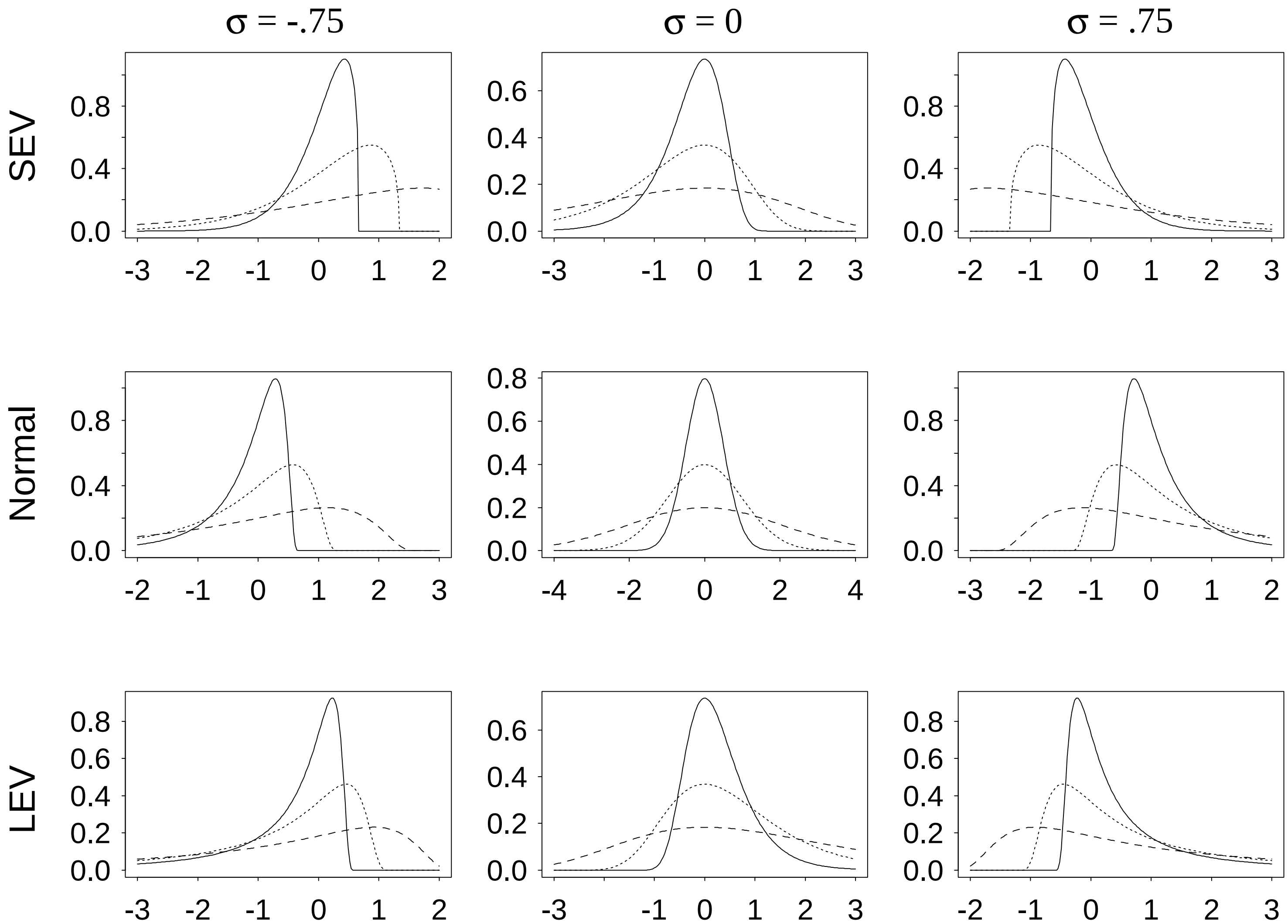
- If the smallest observation is an exact failure, there may be paths in the parameter space leading to infinite likelihood when the **density approximation likelihood** is used.

  Using the **correct** likelihood will allow one to avoid this problem.

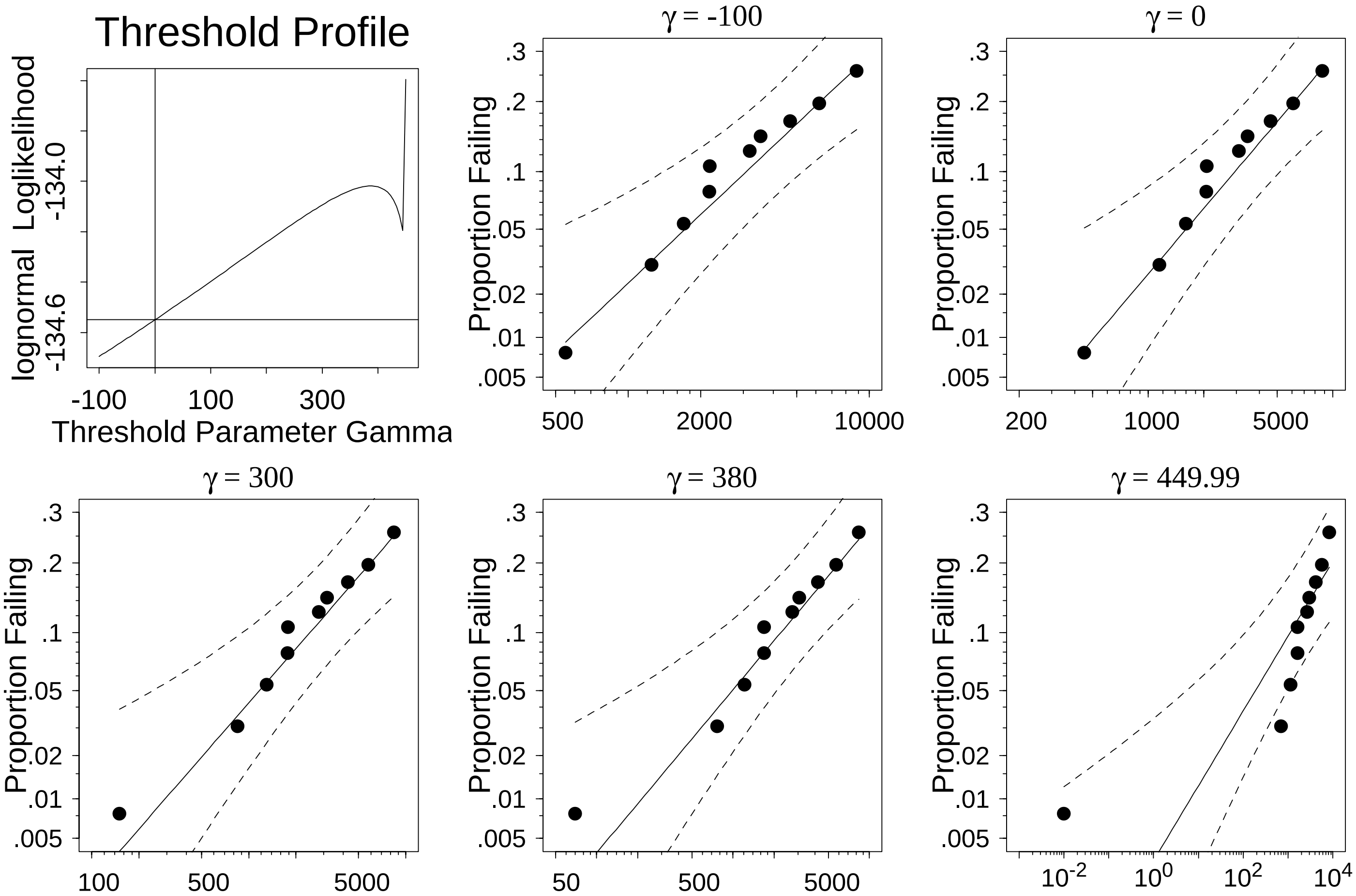
- For some data sets the ML estimate of σ will approach 0 (on the boundary of the parameter space).

  This problem can be avoided by extending to the parameter space to allow values of $\sigma \leq 0$. For the 3-parameter Weibull (lognormal), use the SEV (NORM) GETS distribution.

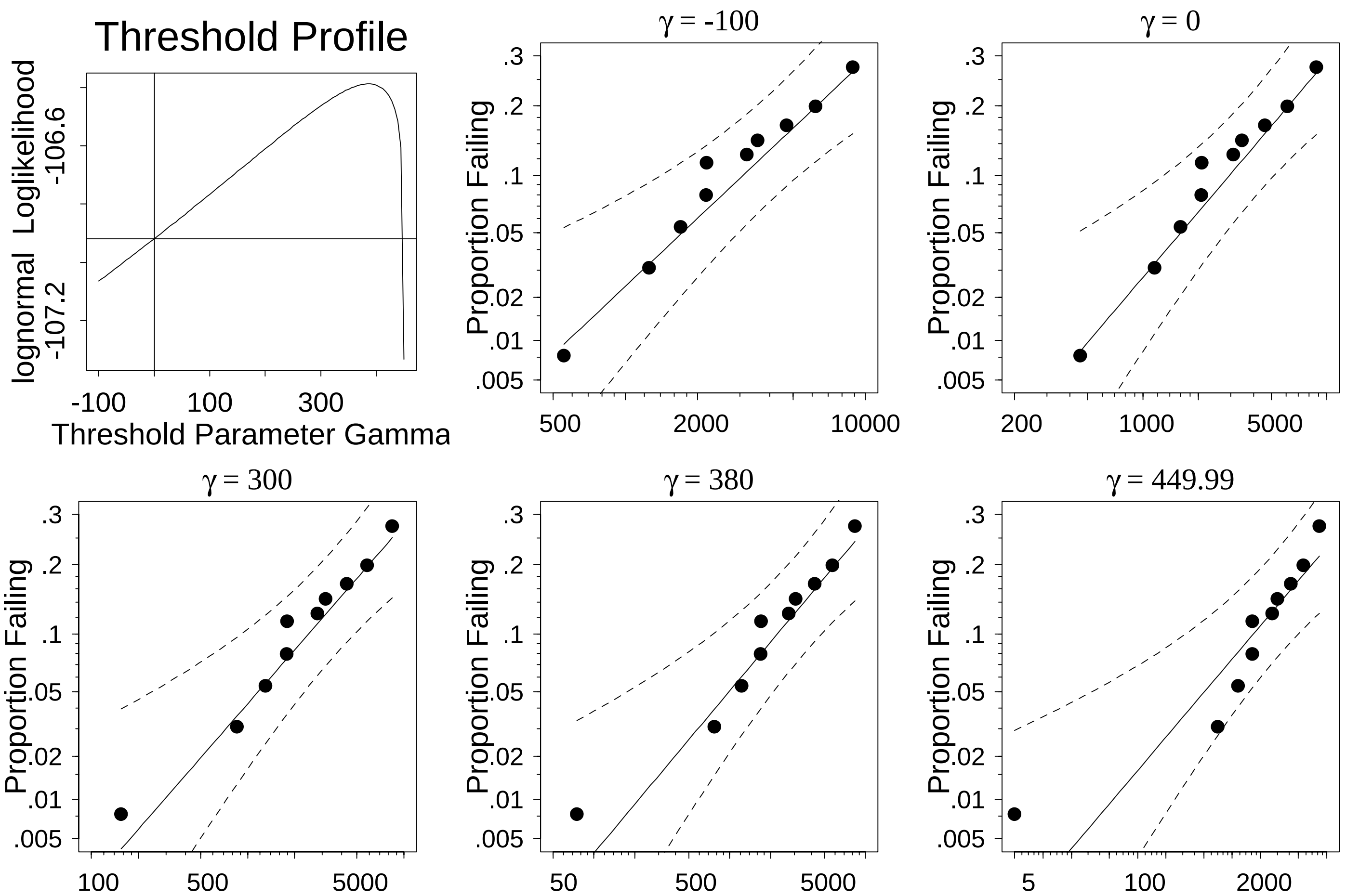
SEV-GETS, NOR-GETS, and LEV-GETS pdfs with $\alpha = 0$, $\sigma = -.75, 0, .75$, and $\varsigma = .5$ (Least Disperse), 1, and 2 (Most Disperse)



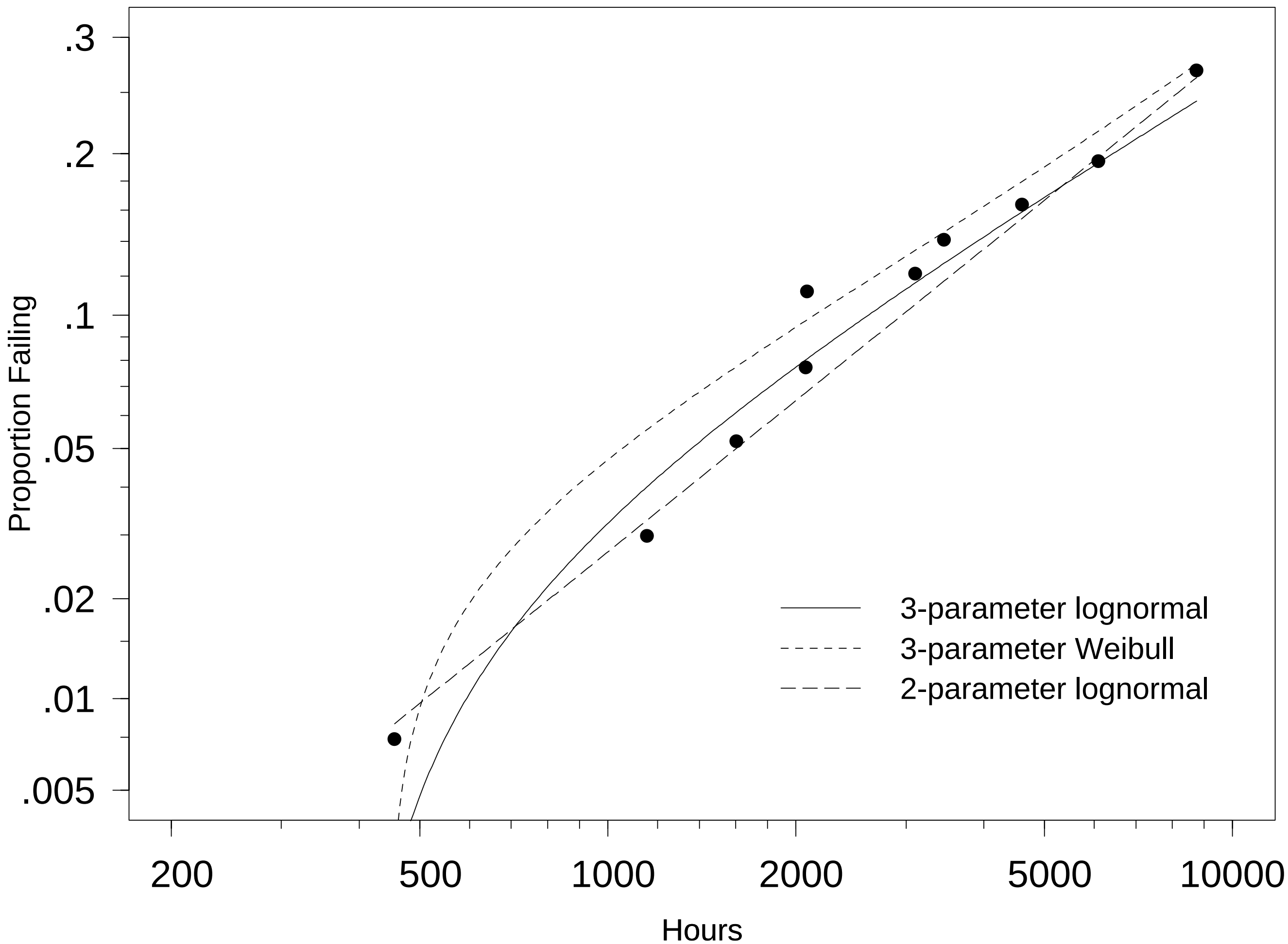
Density Approximation Profile Likelihood for γ and 3-Parameter Lognormal Probability Plots of the Fan Data with γ Varying Between -100 and 449.999



Correct Likelihood ($\Delta = .01$) Profile Likelihood for γ and 3-Parameter Lognormal Probability Plots Of The Fan Data with γ Varying Between -100 and 449.999



Lognormal Probability Plot Comparing ML Estimates Three-Parameter Lognormal and Three-Parameter Weibull Distributions for the Turbine Fan Data



Fitting the 3-Parameter Lognormal and 3-Parameter Weibull Distributions to the Fan Data

- Only 12 failures out of 70 units (multiple censoring).

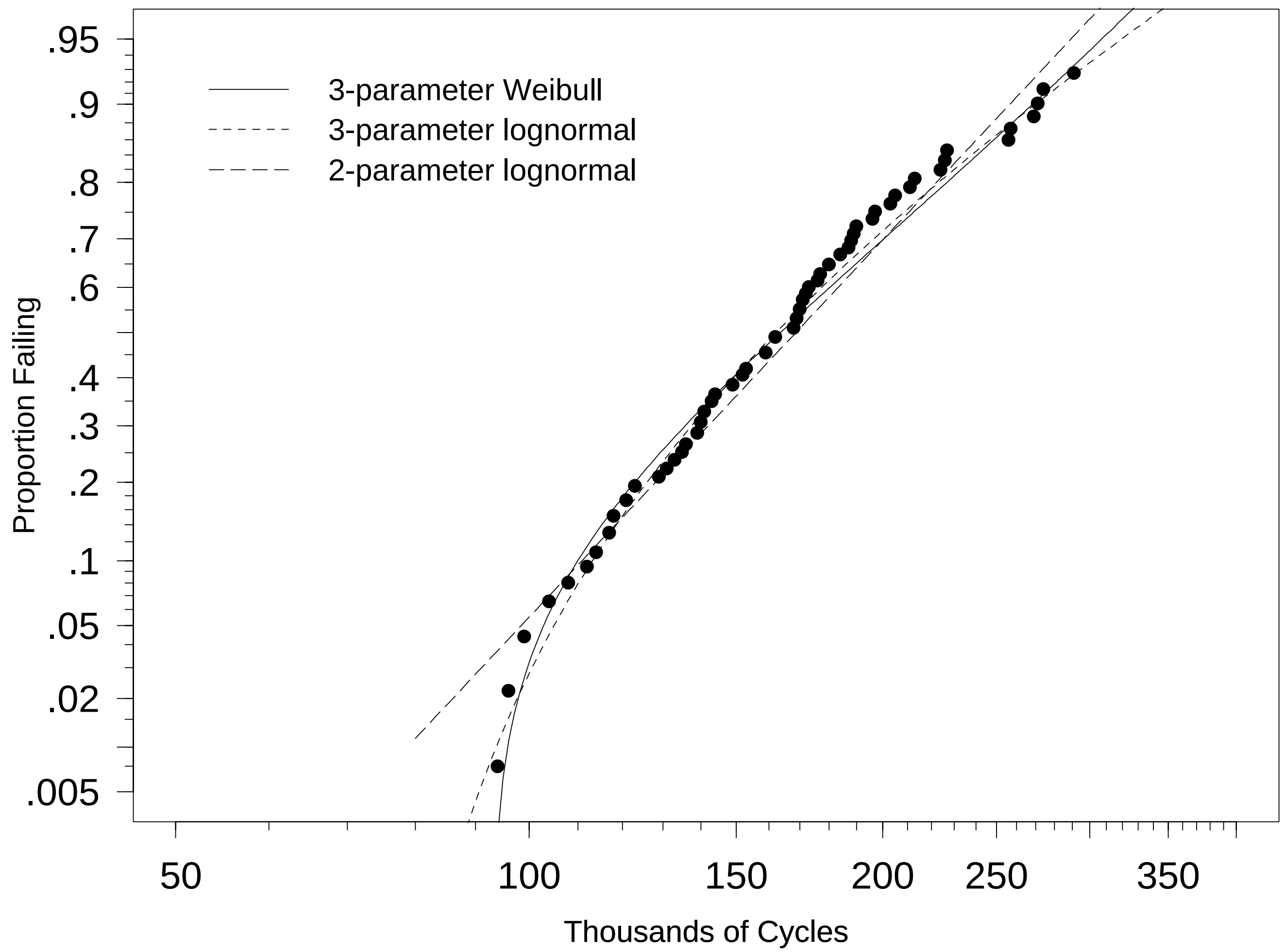
- 2-parameter Lognormal fits the data well. Weibull and exponential also fit the data reasonably well. Can 3-parameter distributions do better?

- For the 3-parameter distributions, the density approximation breaks down. One should use the **correct** likelihood.

- ML suggests that there is a positive threshold, but the level of improvement is statistically unimportant.

- Fitting a 3-parameter distribution to 12 failures is **overfitting**.

Lognormal Probability Plot Comparing Three-Parameter Lognormal and Three-Parameter Weibull Distributions for the Alloy T7987 Data



Three-Parameter Weibull Distribution

- Let γ be the threshold parameter. Then

$$
\begin{aligned}
F(t;\mu,\sigma,\gamma) &= \Phi_{\text{sev}}\left[\frac{\log(t-\gamma)-\mu}{\sigma}\right] \\
f(t;\mu,\sigma,\gamma) &= \frac{1}{\sigma(t-\gamma)}\,\phi_{\text{sev}}\left[\frac{\log(t-\gamma)-\mu}{\sigma}\right]
\end{aligned}
$$

  Both functions equal 0 for $t < 0$.

- Similar for the two-parameter exponential and three-parameter lognormal.

Inferences for 3-Parameter Weibull Assuming that γ is Known

- If γ can be assumed to be known, we can subtract γ from all times and fit the two-parameter Weibull distribution.

- Need to adjust inferences accordingly (e.g. add γ back into estimates of quantiles or subtract γ from times before computing probabilities).

- Similar for 3-parameter lognormal and 2-parameter exponential.

The Three-Parameter Weibull Distribution Likelihood for Right Censored Data

- The likelihood has the form

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

- Problem: when $\gamma \to t_{(1)}$ and $\sigma \to 0$, $L(\mu,\sigma,\gamma) \to \infty$.

- **Solution:** Do not use the density approximation; use the correct likelihood (based on small intervals).