

Chapter 2

Principles of Maximum Likelihood Estimation and The Analysis of Censored Data

Part of the **Iowa State University** NSF/ILI project

Beyond Traditional Statistical Methods

Copyright 1999 D. Cook, W. M. Duckworth, M. S. Kaiser,
W. Q. Meeker, and W. R. Stephenson.

Developed as part of NSF/ILI grant DUE9751644.

January 22, 2003

11h 20min

Chapter 2

Principles of Maximum Likelihood Estimation and The Analysis of Censored Data Objectives

- Motivation for the use of parametric likelihood as a tool for data analysis and inference.
- Principles of likelihood and how likelihood is related to the probability of the observed data.
- Likelihood for a parametric models.
- The use of likelihood confidence intervals for model parameters and other quantities of interest.
- Censoring mechanisms that restrict one's ability to observe actual response values.
- Likelihood for samples containing right and left censored observations.

Importance of Maximum Likelihood

- Versatile method for fitting statistical models to data.
- Use a parametric statistical model to describe a set of data or a process that generated a set of data.
- Much more general than least squares. Allows, in a relatively simple, coherent fashion:
 - ▶ Censoring and truncation.
 - ▶ Non-normal distributions.
 - ▶ Non-independent observations.
 - ▶ Nonlinear regression models.
 - ▶ Multiple sources of variability.
 - ▶ Time-dependent covariates.

Example: Time Between α -Particle Emissions of Americium-241 (Berkson 1966)

Berkson (1966) investigates the randomness of α -particle emissions of Americium-241, which has a half-life of about 458 years.

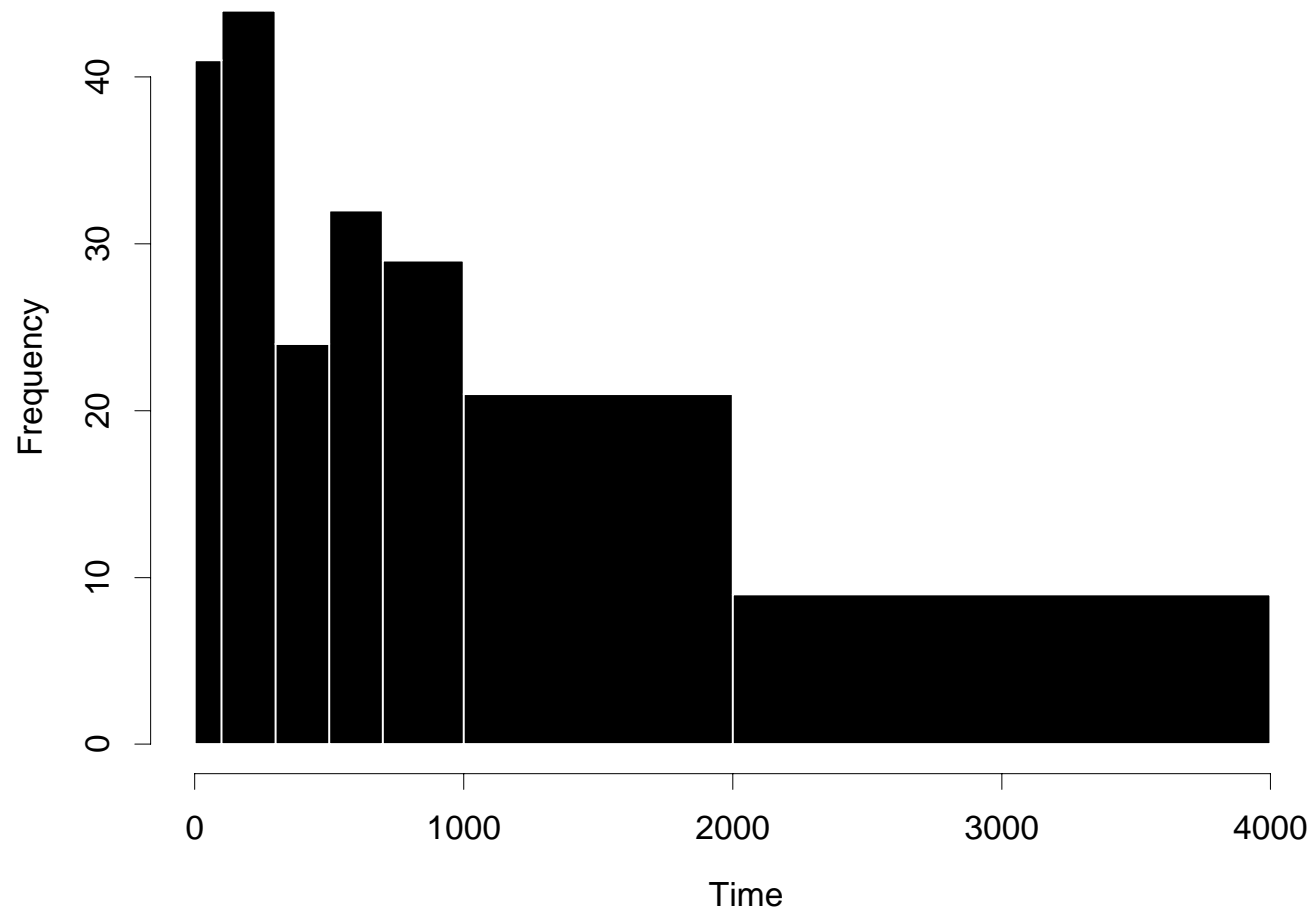
Data: Interarrival times (units: 1/5000 seconds).

- $n = 10,220$ observations.
- Data binned into intervals from 0 to 4000 time units. Interval sizes ranging from 25 to 100 units. Additional interval for observed times exceeding 4,000 time units.
- Smaller samples analyzed here to illustrate sample size effect. We start the analysis with $n = 200$.

Data for α -Particle Emissions of Americium-241

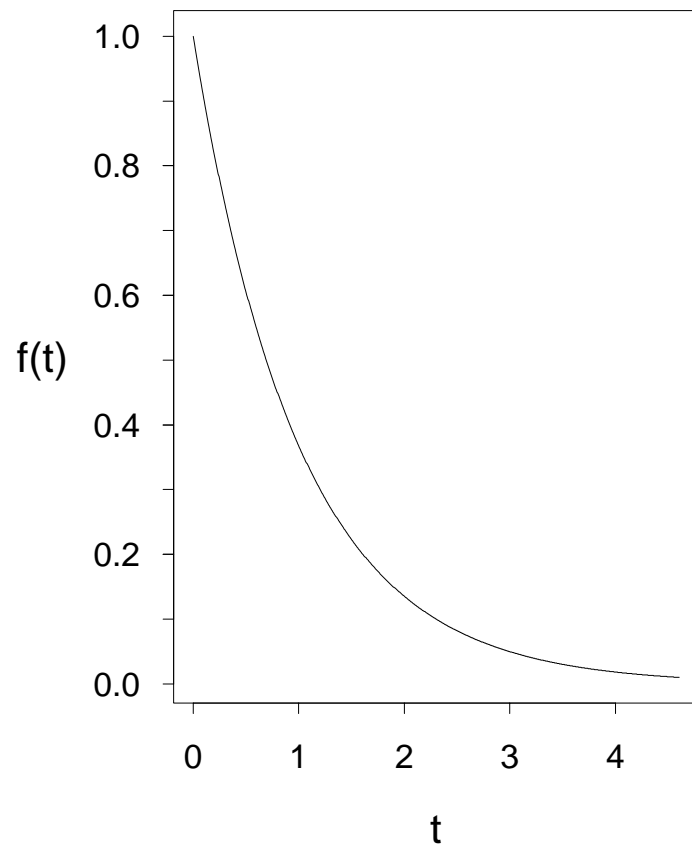
Time		Interarrival Times Frequency of Occurrence	
Interval	Endpoint	All Times	Random Sample of Times
lower	upper	$n = 10220$	$n = 200$
t_{j-1}	t_j		d_j
0	100	1609	41
100	300	2424	44
300	500	1770	24
500	700	1306	32
700	1000	1213	29
1000	2000	1528	21
2000	4000	354	9
4000	∞	16	0
		10220	200

Histogram of the $n = 200$ Sample of α -Particle Interarrival Time Data

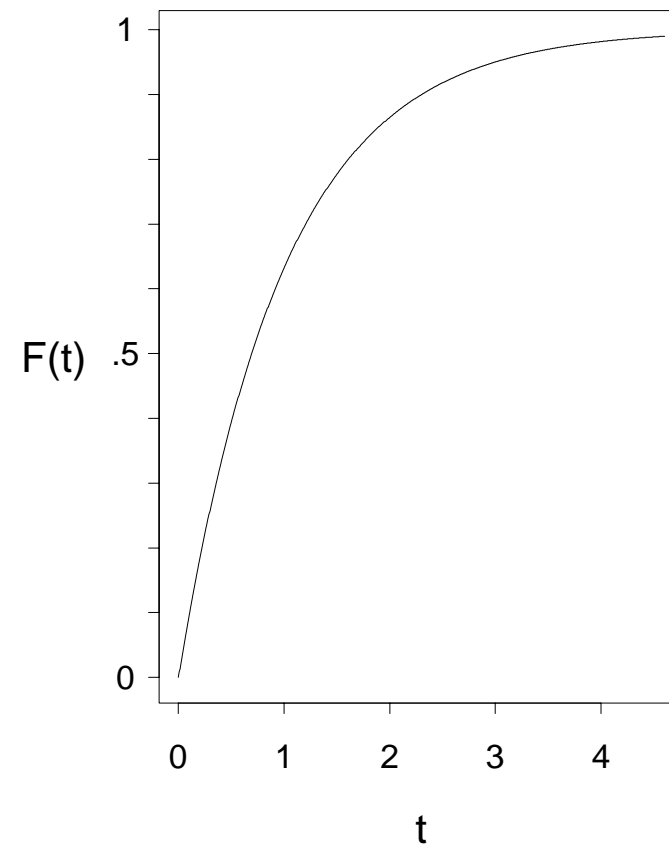


Exponential Distributions PDF and CDF

Probability Density Function



Cumulative Distribution Function



Exponential Distribution

- The exponential cumulative distribution function (cdf) is

$$\Pr(T \leq t) = F(t; \theta) = 1 - \exp\left(-\frac{t}{\theta}\right), \quad t > 0,$$

where θ is the single parameter of the distribution (equal to the first moment or mean, in this example).

- The exponential distribution probability density function (pdf) is

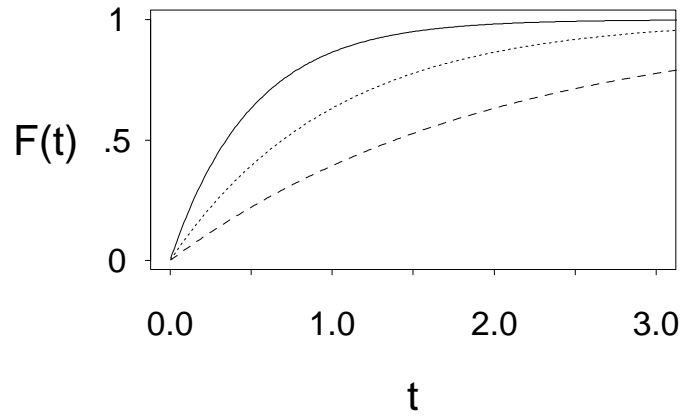
$$f(t; \theta) = \frac{dF(t; \theta)}{dt} = \frac{1}{\theta} \exp\left(-\frac{t}{\theta}\right).$$

- The exponential distribution quantile function is

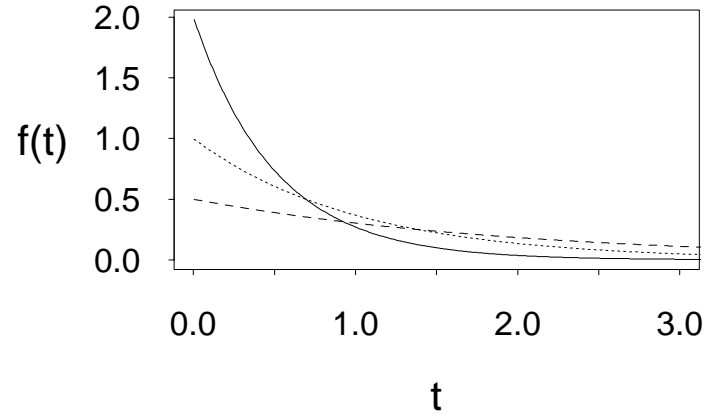
$$t_p = -\theta \log(1 - p).$$

Examples of Exponential Distributions

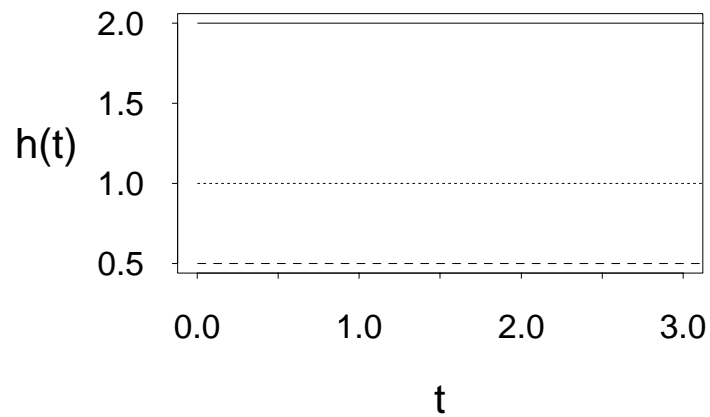
Cumulative Distribution Function



Probability Density Function



Hazard Function



	θ	γ
—	0.5	0
⋯	1.0	0
- - -	2.0	0

Parameters and Functions of Parameters

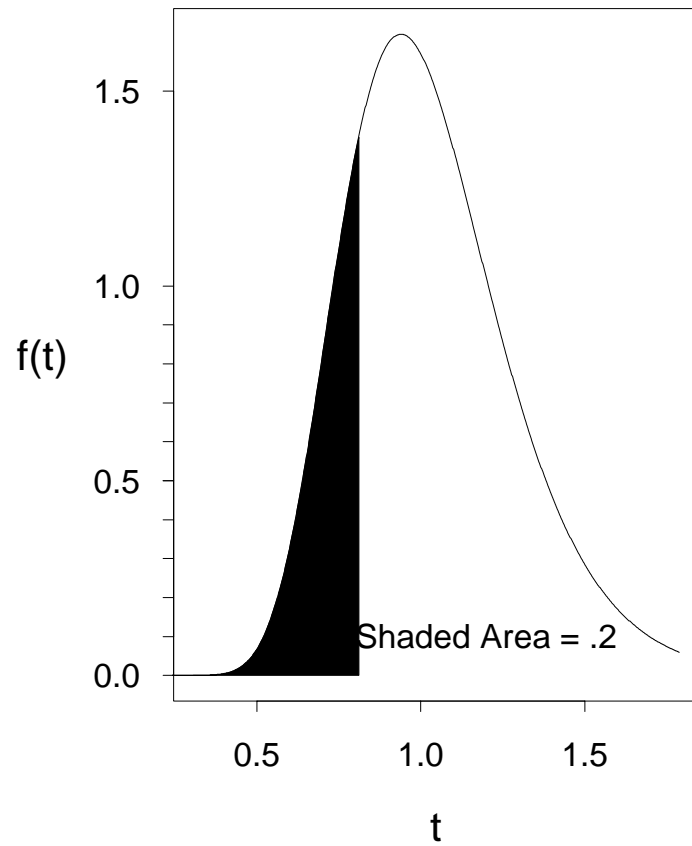
In practical applications, interest often centers on quantities that are functions of θ like

- Probability $p = \Pr(T \leq t) = F(t; \theta)$ for a specified t .
- The p quantile of the distribution of T [value of t_p such that $F(t_p; \theta) = p$].
- The mean of T

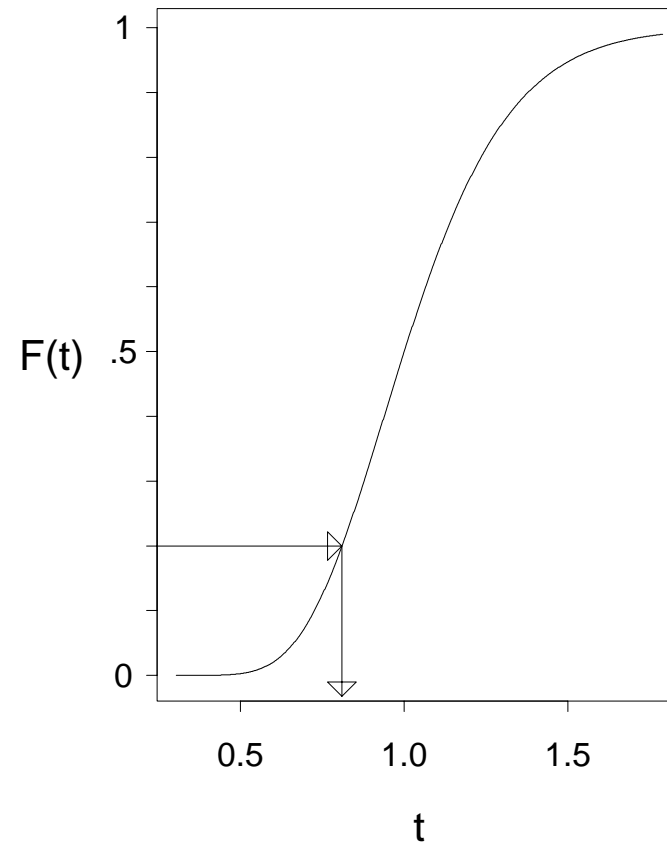
$$E(T) = \int_{-\infty}^{\infty} t f(t) dt$$

Lognormal Distributions PDF and CDF

Probability Density Function



Cumulative Distribution Function



Lognormal Distribution

- The Lognormal cumulative distribution function (cdf) is

$$F(t; \mu, \sigma) = \Phi_{\text{nor}} \left[\frac{\log(t) - \mu}{\sigma} \right], \quad t > 0,$$

where μ is the mean of the logarithms and σ is the standard deviation of the logarithms.

- The Lognormal distribution probability density function (pdf) is

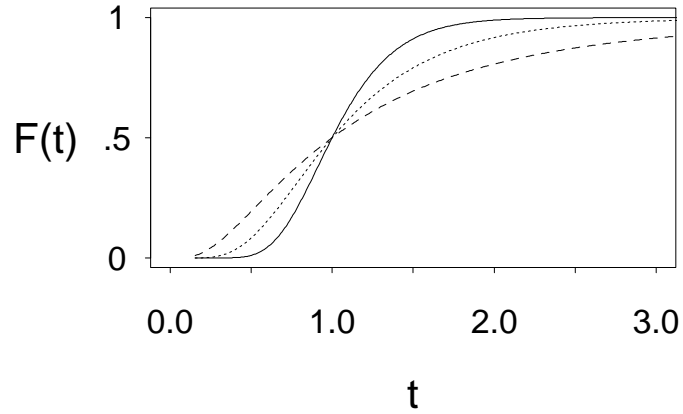
$$f(t; \mu, \sigma) = \frac{dF(t; \mu, \sigma)}{dt} = \frac{1}{\sigma t} \phi_{\text{nor}} \left[\frac{\log(t) - \mu}{\sigma} \right].$$

- The Lognormal distribution quantile function is

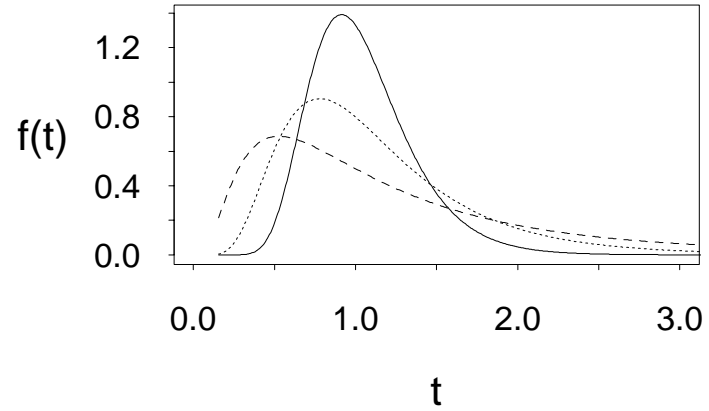
$$t_p = \exp \left[\mu + \sigma \Phi_{\text{nor}}^{-1}(p) \right].$$

Examples of Lognormal Distributions

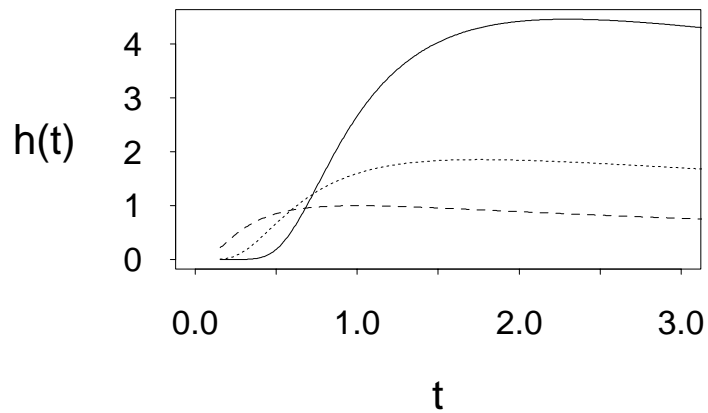
Cumulative Distribution Function



Probability Density Function



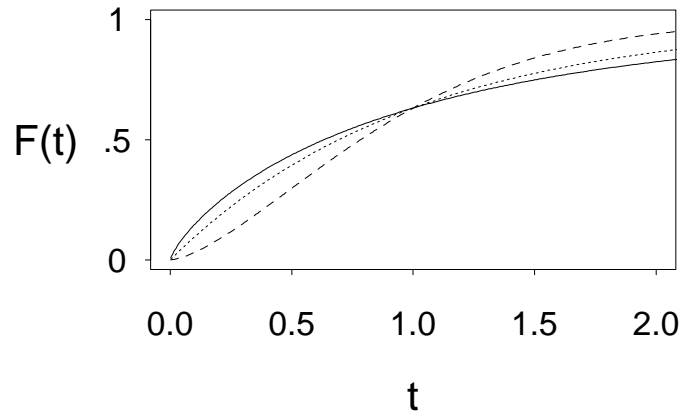
Hazard Function



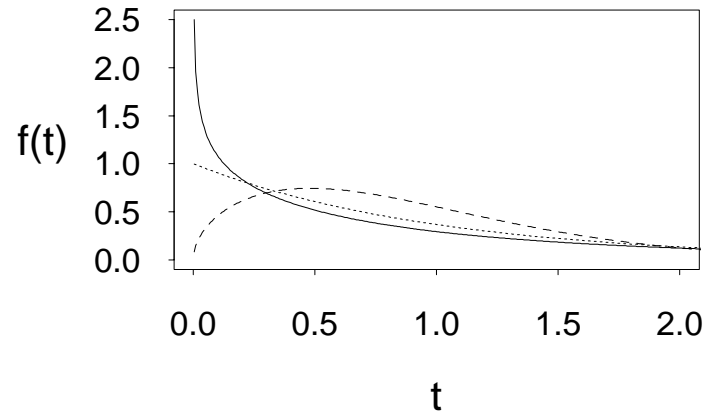
	σ	μ
—	0.3	0
⋯	0.5	0
- - -	0.8	0

Examples of Weibull Distributions

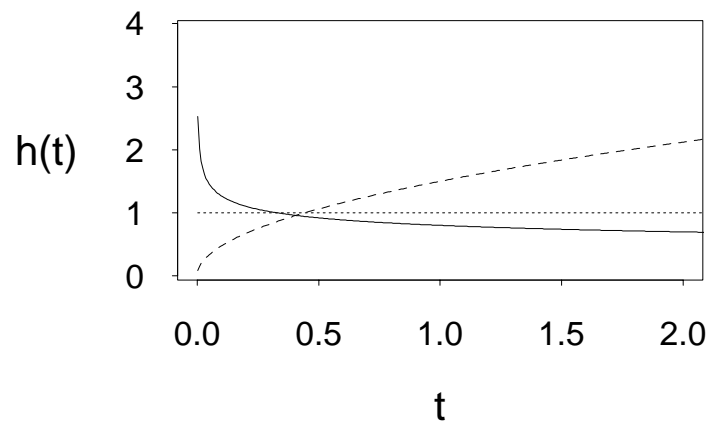
Cumulative Distribution Function



Probability Density Function

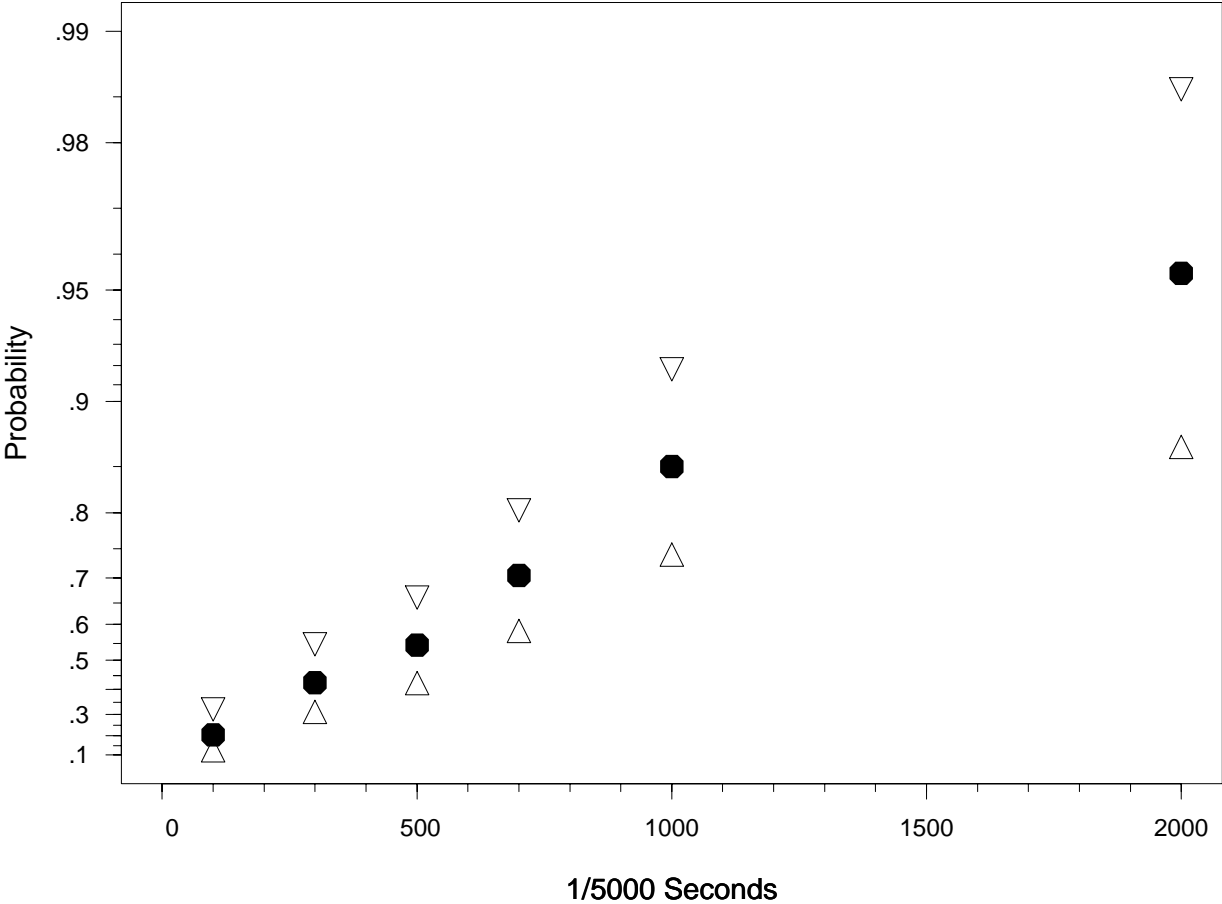


Hazard Function



	β	η
—	0.8	1
⋯	1.0	1
- - -	1.5	1

Exponential Probability Plot of the $n = 200$ Sample of α -Particle Interarrival Time Data. The Plot also Shows Approximate 95% Simultaneous Nonparametric Confidence Bands.



Parametric Likelihood Probability of the Data

- Using the model $\Pr(T \leq t) = F(t; \boldsymbol{\theta})$ for continuous T , the likelihood (probability) for a single observation in the interval $(t_{i-1}, t_i]$ is

$$L_i(\boldsymbol{\theta}; \text{data}_i) = \Pr(t_{i-1} < T \leq t_i) = F(t_i; \boldsymbol{\theta}) - F(t_{i-1}; \boldsymbol{\theta}).$$

Can be generalized to allow for explanatory variables, multiple sources of variability, and other model features.

- The total likelihood is the joint probability of the data. Assuming n independent observations

$$L(\boldsymbol{\theta}) = L(\boldsymbol{\theta}; \text{DATA}) = \mathcal{C} \prod_{i=1}^n L_i(\boldsymbol{\theta}; \text{data}_i).$$

- Want to estimate $\boldsymbol{\theta}$. Find values of $\boldsymbol{\theta}$ for which $L(\boldsymbol{\theta})$ is relatively large.

Exponential Distribution and Likelihood for Interval Data

Data: α -particle emissions of americium-241

- The exponential distribution is

$$F(t; \theta) = 1 - \exp\left(-\frac{t}{\theta}\right), \quad t > 0.$$

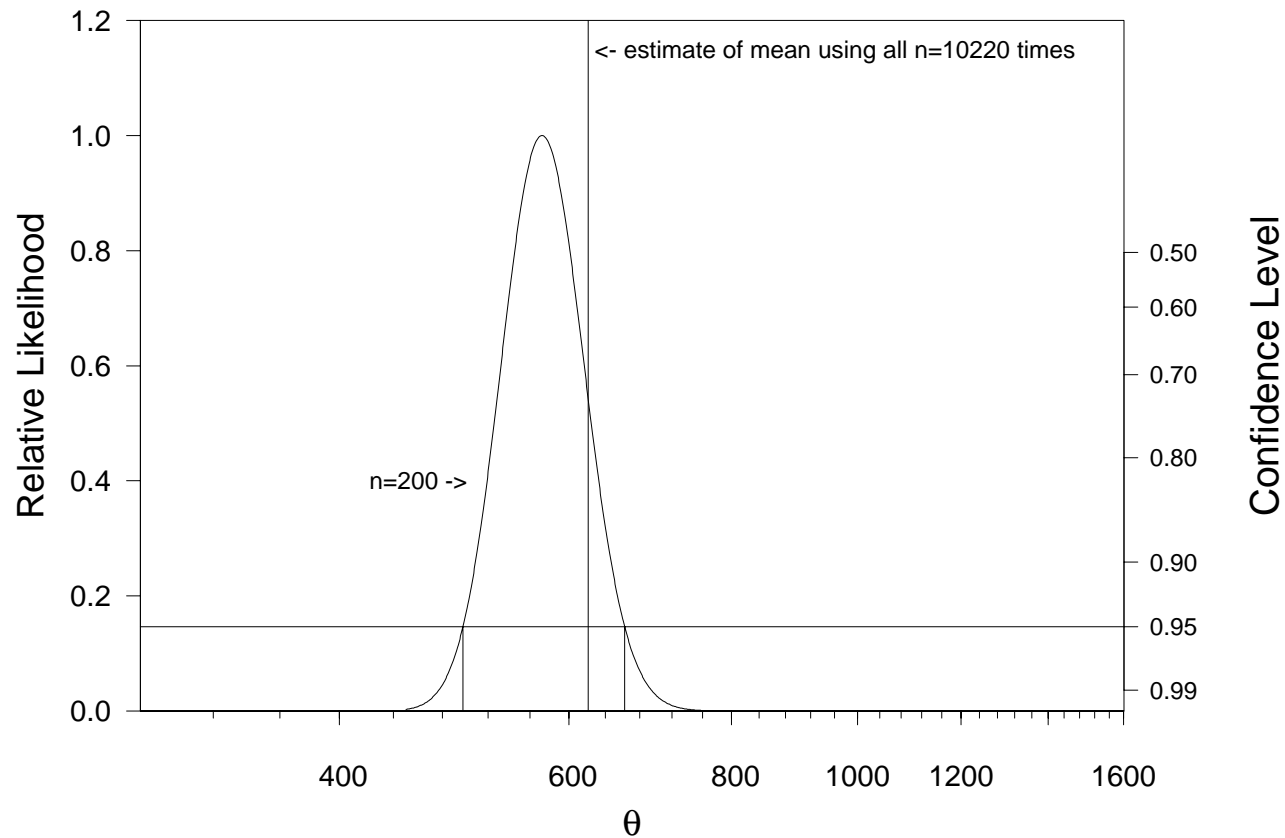
$\theta = E(T)$, the mean time between arrivals.

- The interval-data likelihood has the form

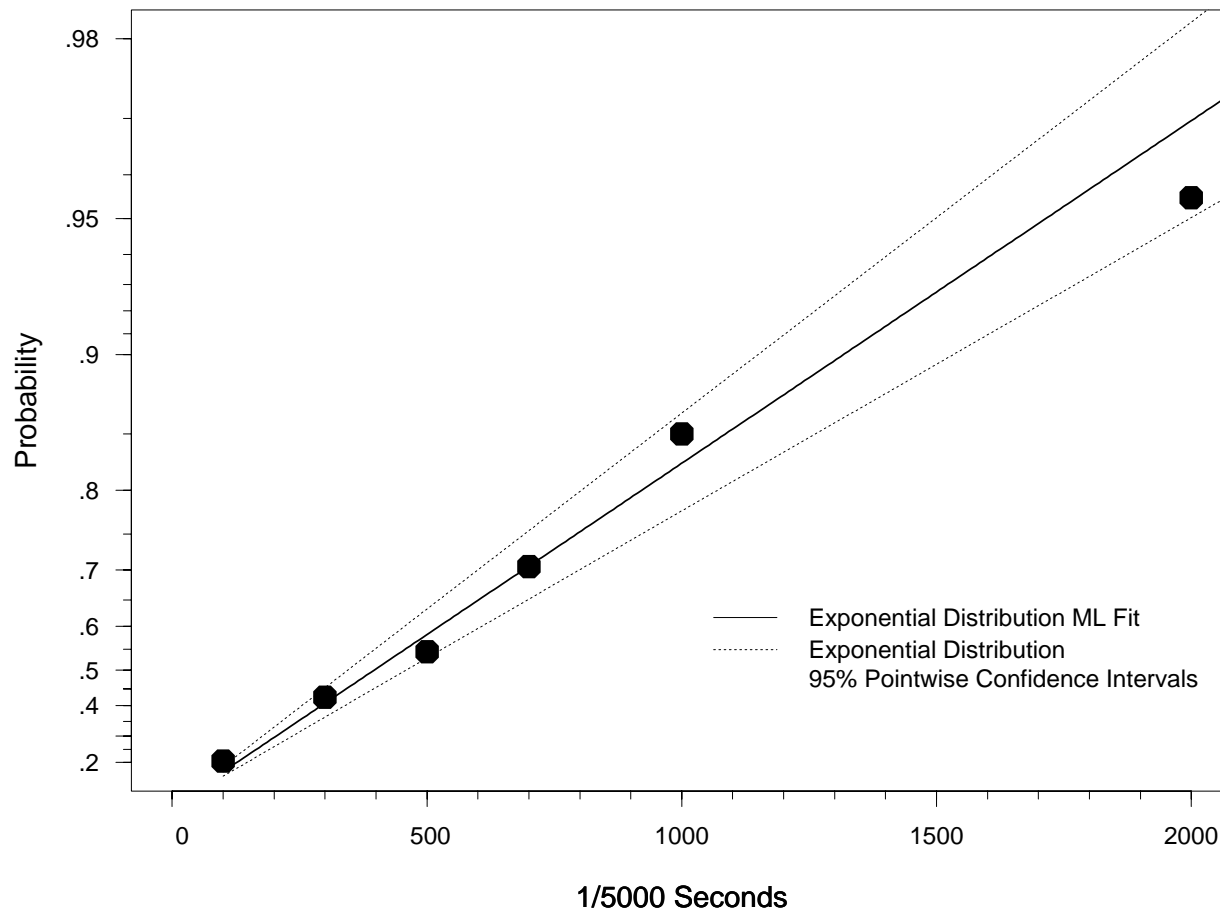
$$\begin{aligned} L(\theta) &= \prod_{i=1}^n L_i(\theta) = \prod_{j=1}^8 \left[F(t_j; \theta) - F(t_{j-1}; \theta) \right]^{d_j} \\ &= \prod_{j=1}^8 \left[\exp\left(-\frac{t_{j-1}}{\theta}\right) - \exp\left(-\frac{t_j}{\theta}\right) \right]^{d_j} \end{aligned}$$

where d_j is the number of interarrival times in the j th interval (i.e., times between t_{j-1} and t_j).

$R(\theta) = L(\theta)/L(\hat{\theta})$ for the $n = 200$ α -Particle Interarrival Time Data. Vertical Lines Give an Approximate 95% Likelihood-Based Confidence Interval for θ



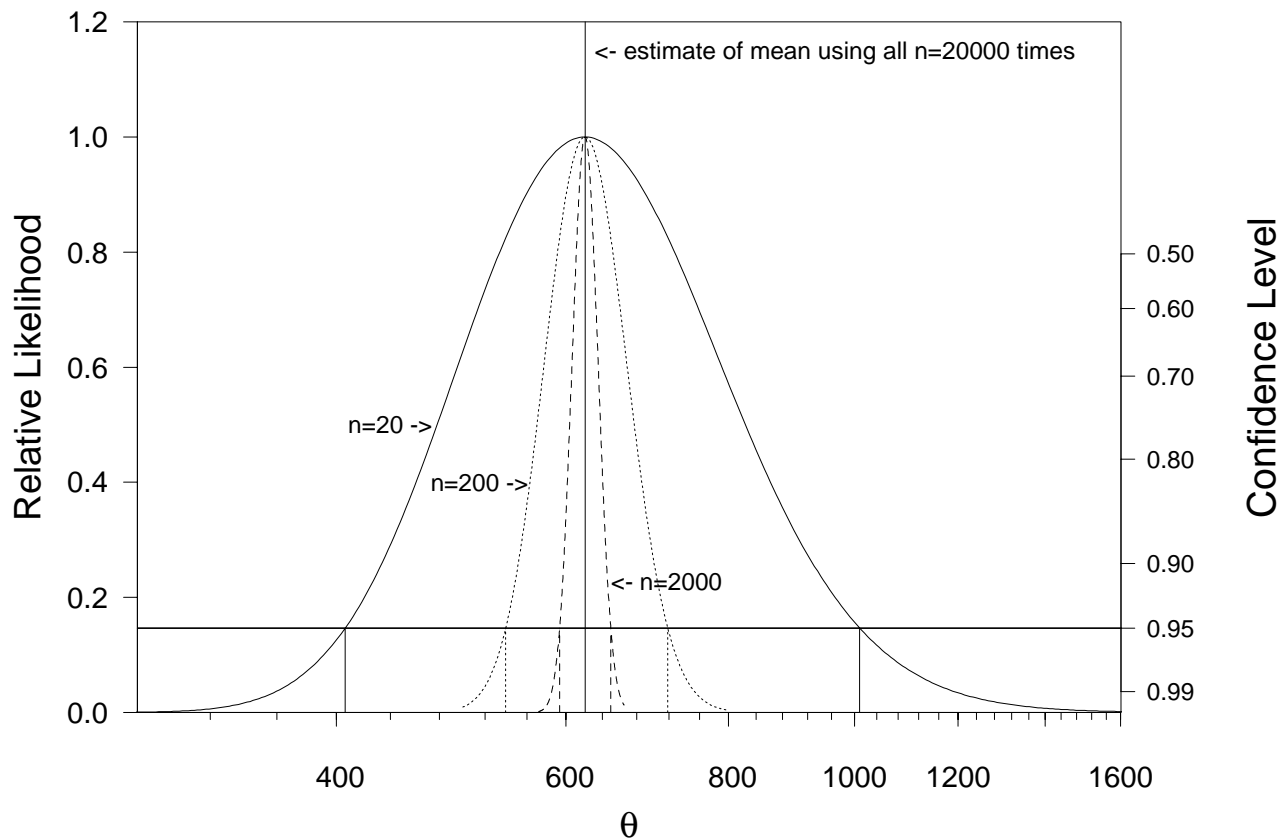
Exponential Probability Plot for the $n = 200$ Sample of α -Particle Interarrival Time Data. The Plot also Shows Parametric Exponential ML Estimate and 95% Confidence Intervals for $F(t)$.



**Example. α -Particle Pseudo Data Constructed
with Constant Proportion within Each Bin**

Time		Interarrival Times Frequency of Occurrence			
Interval Endpoint		Samples of Times			
lower	upper	$n=20000$	$n=2000$	$n=200$	$n=20$
t_{j-1}	t_j	d_j			
0	100	3000	300	30	3
100	300	5000	500	50	5
300	500	3000	300	30	3
500	700	3000	300	30	3
700	1000	2000	200	20	2
1000	2000	3000	300	30	3
2000	4000	1000	100	10	1
4000	∞	0000	000	0	0
		20000	2000	200	20

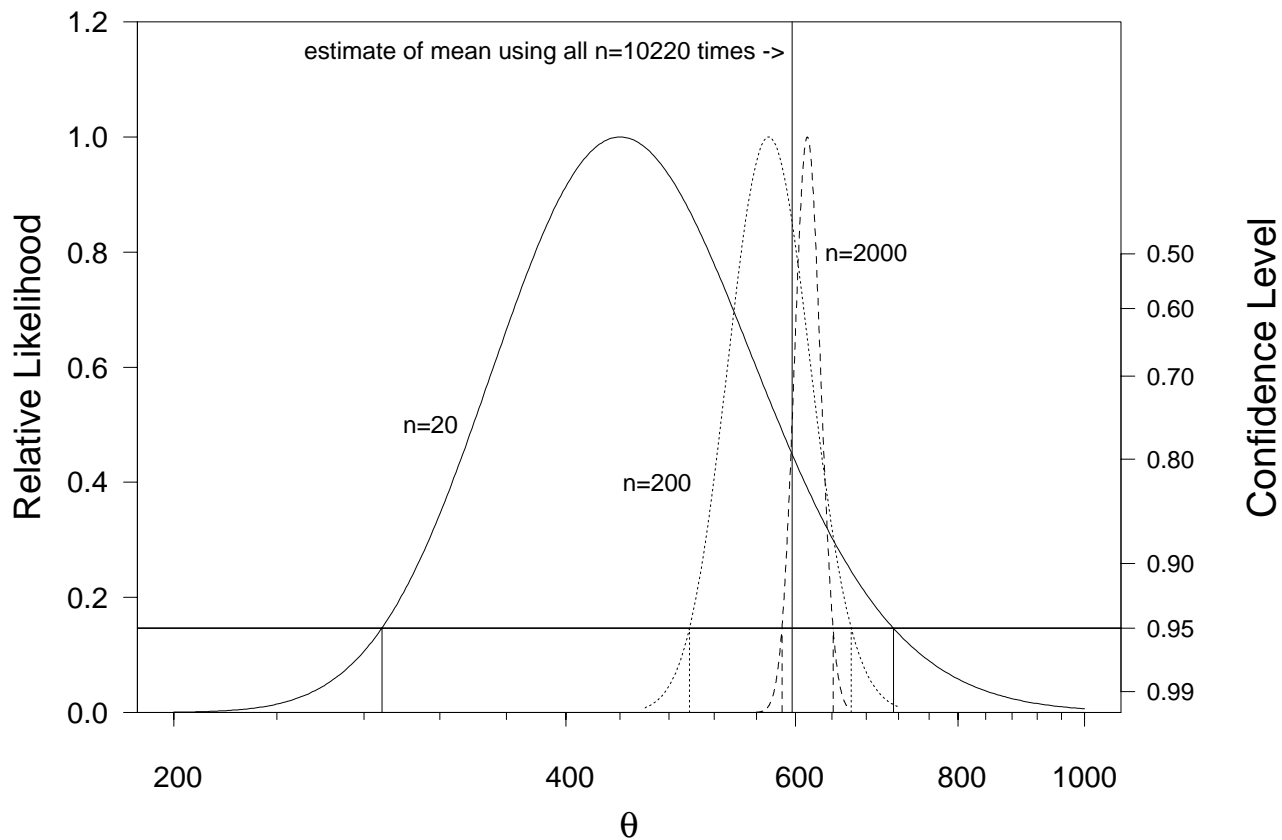
$R(\theta) = L(\theta)/L(\hat{\theta})$ for the $n = 20, 200, \text{ and } 2000$
**Pseudo Data. Vertical Lines Give Corresponding
Approximate 95% Likelihood-Based Confidence
Intervals**



Example. α -Particle Random Samples

Time		Interarrival Times Frequency of Occurrence			
Interval	Endpoint	All Times	Random Samples of Times		
lower	upper	$n = 10220$	$n = 2000$	$n = 200$	$n = 20$
t_{j-1}	t_j		d_j		
0	100	1609	292	41	3
100	300	2424	494	44	7
300	500	1770	332	24	4
500	700	1306	236	32	1
700	1000	1213	261	29	3
1000	2000	1528	308	21	2
2000	4000	354	73	9	0
4000	∞	16	4	0	0
		10220	2000	200	20

$R(\theta) = L(\theta)/L(\hat{\theta})$ for the $n = 20, 200,$ and 2000 Samples from the α -Particle Interarrival Time Data. Vertical Lines Give Corresponding Approximate 95% Likelihood-Based Confidence Intervals.



Likelihood as a Tool for Modeling/Inference

What can we do with the (log) likelihood?

$$\mathcal{L}(\boldsymbol{\theta}) = \log[L(\boldsymbol{\theta})] = \sum_{i=1}^n \mathcal{L}_i(\boldsymbol{\theta}).$$

- Study the surface.
- Maximize with respect to $\boldsymbol{\theta}$ (ML point estimates).
- Look at curvature at maximum (gives estimate of Fisher information and asymptotic variance).

Likelihood as a Tool for Modeling/Inference (Cont.)

- Regions of high likelihood are credible; regions of low likelihood are not credible (suggests confidence regions for parameters).
- If the length of θ is > 1 or 2 and interest centers on subset of θ (need to get rid of nuisance parameters), look at “profiles”
(suggests confidence regions/intervals for parameter subsets).
- Calibrate confidence regions/intervals with χ^2 or simulation (or parametric bootstrap).
- Use “reparameterization” to study functions of θ .

Likelihood Ratio

- The likelihood ratio to test for $H_0 : \theta = \theta_0$ is

$$L(\theta_0)/L(\hat{\theta}).$$

- The log likelihood ratio statistic for a particular θ_0 is

$$X^2 = -2 \log [L(\theta_0)/L(\hat{\theta})]$$

- Reject H_0 at the 5% level of significance if $X^2 > \chi^2_{(1-\alpha;1)}$.

- For example, using the $n = 200$ α -particle data,

$$-2 \log [L(650)/L(572.3)] = 2.94 < \chi^2_{(.95;1)} = 3.84$$

- Likelihood ratio confidence interval: the set of all values of θ such that H_0 at the α level of significance is not rejected.

Large-Sample Approximate Theory for Likelihood Ratios for a Scalar Parameter

- Relative likelihood for θ is

$$R(\theta) = \frac{L(\theta)}{L(\hat{\theta})}.$$

- If evaluated at the true θ , then, asymptotically (in large samples), $-2 \log[R(\theta)]$ follows, a chisquare distribution with 1 degrees of freedom.
- An approximate $100(1 - \alpha)\%$ likelihood-based confidence region for θ is the set of all values of θ such that

$$-2 \log[R(\theta)] < \chi_{(1-\alpha;1)}^2$$

or, equivalently, the set defined by

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

Normal-Approximation Confidence Intervals for θ

- A $100(1 - \alpha)\%$ normal-approximation (or Wald) confidence interval for θ is

$$[\underline{\tilde{\theta}}, \quad \tilde{\theta}] = \hat{\theta} \pm z_{(1-\alpha/2)} \widehat{\text{se}}_{\hat{\theta}}.$$

where $\widehat{\text{se}}_{\hat{\theta}} = \sqrt{[-d^2\mathcal{L}(\theta)/d\theta^2]^{-1}}$ is evaluated at $\hat{\theta}$.

- Based on

$$Z_{\hat{\theta}} = \frac{\hat{\theta} - \theta}{\widehat{\text{se}}_{\hat{\theta}}} \sim \text{NOR}(0, 1)$$

- From the definition of $\text{NOR}(0, 1)$ quantiles

$$\Pr [z_{(\alpha/2)} < Z_{\hat{\theta}} \leq z_{(1-\alpha/2)}] \approx 1 - \alpha$$

implies that

$$\Pr [\hat{\theta} - z_{(1-\alpha/2)} \widehat{\text{se}}_{\hat{\theta}} < \theta \leq \hat{\theta} + z_{(1-\alpha/2)} \widehat{\text{se}}_{\hat{\theta}}] \approx 1 - \alpha.$$

Normal-Approximation Confidence Intervals for θ (contd.)

- A $100(1 - \alpha)\%$ normal-approximation (or Wald) confidence interval for θ is

$$[\underline{\theta}, \tilde{\theta}] = [\hat{\theta}/w, \hat{\theta} \times w]$$

where $w = \exp[z_{(1-\alpha/2)} \widehat{\text{se}}_{\hat{\theta}}/\hat{\theta}]$. This follows after transforming (by exponentiation) the confidence interval

$$[\log(\underline{\theta}), \log(\tilde{\theta})] = \log(\hat{\theta}) \pm z_{(1-\alpha/2)} \widehat{\text{se}}_{\log(\hat{\theta})}$$

which is based on

$$Z_{\log(\hat{\theta})} = \frac{\log(\hat{\theta}) - \log(\theta)}{\widehat{\text{se}}_{\log(\hat{\theta})}} \sim \text{NOR}(0, 1)$$

- Because $\log(\hat{\theta})$ is unrestricted in sign, generally $Z_{\log(\hat{\theta})}$ is closer to an $\text{NOR}(0, 1)$ distribution than is $Z_{\hat{\theta}}$.

Comparisons for α -Particle Data

	All Times $n = 10,220$	Sample of Times	
		$n = 200$	$n = 20$
ML Estimate $\hat{\theta}$	596	572	440
Standard Error $\widehat{se}_{\hat{\theta}}$	6.1	42.7	101
95% Confidence Intervals for θ Based on			
Likelihood	[585, 608]	[498, 662]	[289, 713]
$Z_{\log(\hat{\theta})} \sim \text{NOR}(0, 1)$	[585, 608]	[496, 660]	[281, 690]
$Z_{\hat{\theta}} \sim \text{NOR}(0, 1)$	[585, 608]	[491, 654]	[242, 638]
ML Estimate $\hat{\lambda} \times 10^5$	168	175	227
Standard Error $\widehat{se}_{\hat{\lambda} \times 10^5}$	1.7	13	52
95% Confidence Intervals for $\lambda \times 10^5$ Based on			
Likelihood	[164, 171]	[151, 201]	[140, 346]
$Z_{\log(\hat{\lambda})} \sim \text{NOR}(0, 1)$	[164, 171]	[152, 202]	[145, 356]
$Z_{\hat{\lambda}} \sim \text{NOR}(0, 1)$	[164, 171]	[149, 200]	[125, 329]

Confidence Intervals for Functions of θ

- For one-parameter distributions, confidence intervals for θ can be translated directly into confidence intervals for monotone functions of θ .

- The arrival rate $\lambda = 1/\theta$ is a decreasing function of θ .

$$[\underline{\lambda}, \quad \tilde{\lambda}] = [1/\tilde{\theta}, \quad 1/\underline{\theta}] = [.00151, \quad .00201].$$

- $F(t; \theta)$ is a decreasing function of θ

$$[\underline{F}(t_e), \quad \tilde{F}(t_e)] = [F(t_e; \tilde{\theta}), \quad F(t_e; \underline{\theta})]$$

for any specified value of t_e .

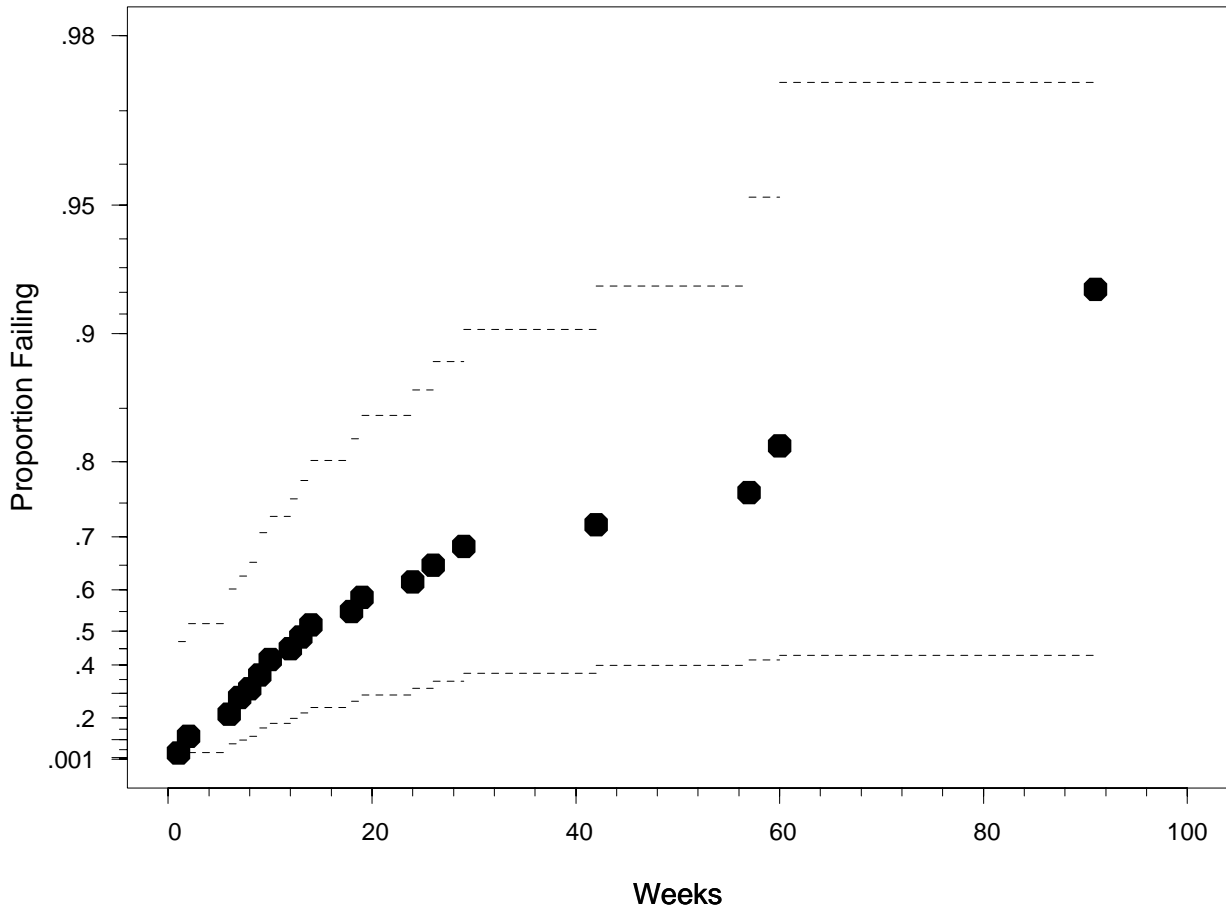
- Quantile $t_p = -\theta \log(1 - p)$ is a increasing function of θ

$$[\underline{t}_p, \quad \tilde{t}_p] = [-\underline{\theta} \log(1 - p), \quad -\tilde{\theta} \log(1 - p)].$$

Leukemia Patient Remission Times (from Lawless (1982))

- The observed times were 1, 1, 2, 2, 2, 6, 6, 6, 7, 8, 9, 9, 10, 12, 13, 14, 18, 19, 24, 26, 29, 31*, 42, 45*, 50*, 57, 60, 71*, 85*, 91 weeks.
- The times marked with a * indicate patients that were still in remission at the time the data were analyzed (right-censored observations).
- At the time the data were analyzed
 - ▶ Twenty-five patients had come out of remission.
 - ▶ Five patients were still in remission.
- Analysts wanted to estimate the cdf of the remission-time distribution.

Exponential Probability Plot of the Remission Times with 95% Simultaneous Confidence Bands.



Density Approximation for Exact Observations

- If $t_{i-1} = t_i - \Delta_i$, $\Delta_i > 0$, and the correct likelihood

$$F(t_i; \boldsymbol{\theta}) - F(t_{i-1}; \boldsymbol{\theta}) = F(t_i; \boldsymbol{\theta}) - F(t_i - \Delta_i; \boldsymbol{\theta})$$

can be approximated with the density $f(t)$ as

$$[F(t_i; \boldsymbol{\theta}) - F(t_i - \Delta_i; \boldsymbol{\theta})] = \int_{(t_i - \Delta_i)}^{t_i} f(t) dt \approx f(t_i; \boldsymbol{\theta}) \Delta_i$$

then the density approximation for exact observations

$$L_i(\boldsymbol{\theta}; \text{data}_i) = f(t_i; \boldsymbol{\theta})$$

may be appropriate.

- For most common models, the density approximation is adequate for small Δ_i .
- There are, however, situations where the approximation breaks down as $\Delta_i \rightarrow 0$.

Likelihood with Exact and Right-Censored Observations

- The probability of a right-censored the observation is

$$L_i(\boldsymbol{\theta}) = \Pr(T > t_i) = F(\infty; \boldsymbol{\theta}) - F(t_i; \boldsymbol{\theta}) = 1 - F(t_i; \boldsymbol{\theta}).$$

- With n independent exact and right-censored observations, the likelihood is:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n \{f(t_i; \boldsymbol{\theta})\}^{\delta_i} \{1 - F(t_i; \boldsymbol{\theta})\}^{1-\delta_i},$$

where $\delta_i = 1$ for an “exact” observation and $\delta_i = 0$ for a right-censored observation.

- For the exponential distribution,

$$L(\theta) = \prod_{i=1}^n \left\{ \frac{1}{\theta} \exp \left[-\frac{t_i}{\theta} \right] \right\}^{\delta_i} \left\{ \exp \left[-\frac{t_i}{\theta} \right] \right\}^{1-\delta_i}.$$

- It is easy to show that the value of θ that maximizes $L(\theta)$ is $\hat{\theta} = \sum_{i=1}^n t_i / r$, where $r = \sum_{i=1}^n \delta_i$ is the number of failures.

ML Estimates for the Exponential Distribution Mean Based on the Density Approximation

- With r exact failures and $n - r$ right-censored observations the ML estimate of θ is

$$\hat{\theta} = \frac{TTT}{r} = \frac{\sum_{i=1}^n t_i}{r}$$

$TTT = \sum_{i=1}^n t_i$, total time on test, is the sum of the failure times plus the censoring time of the units that are censored.

- Using the observed curvature in the log likelihood:

$$\widehat{se}_{\hat{\theta}} = \sqrt{\left[-\frac{d^2 \mathcal{L}(\theta)}{d\theta^2} \right]_{\hat{\theta}}^{-1}} = \sqrt{\frac{\hat{\theta}^2}{r}} = \frac{\hat{\theta}}{\sqrt{r}}.$$

- If the data are complete or failure censored, $2TTT/\theta \sim \chi_{2r}^2$. Then an exact $100(1 - \alpha)\%$ confidence interval for θ is

$$[\underline{\theta}, \tilde{\theta}] = \left[\frac{2(TTT)}{\chi_{(1-\alpha/2; 2r)}^2}, \frac{2(TTT)}{\chi_{(\alpha/2; 2r)}^2} \right].$$

Confidence Interval for the Mean Time in Remission

- The “total time on test” for the 30 patients is
 $TTT = 1 + 1 + 2 + 2 + 2 + 6 \dots + 85 + 91 = 756$ weeks.
 $r = 25$ patients came out of remission.

- The ML estimate of θ is

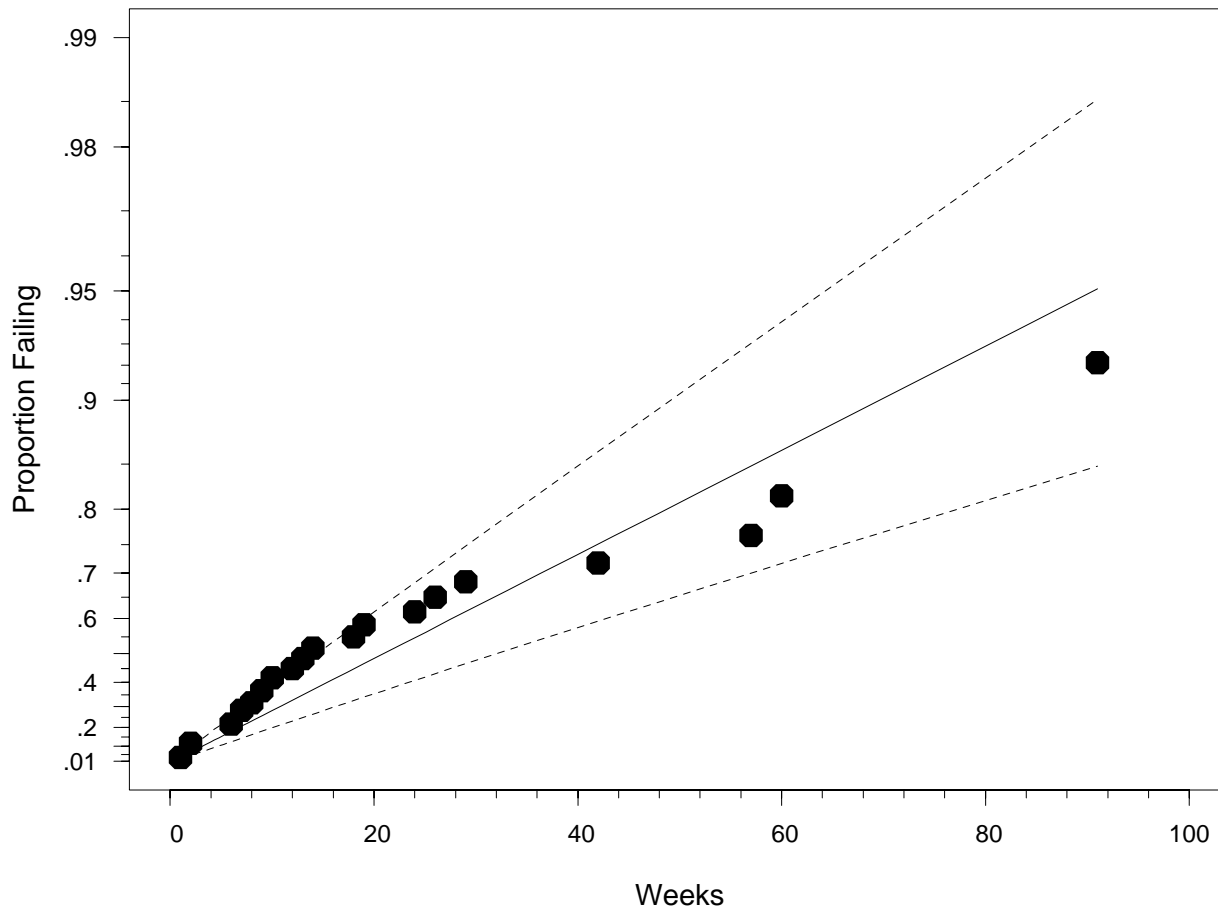
$$\hat{\theta} = 756/25 = 30.24 \text{ weeks.}$$

- An approximate 95% confidence interval for θ is

$$\begin{aligned} [\tilde{\theta}, \hat{\theta}] &= \left[\frac{2(756)}{\chi_{(.975;50)}^2}, \frac{2(756)}{\chi_{(.025;50)}^2} \right] = \left[\frac{1901.76}{46.98}, \frac{1901.76}{16.79} \right] \\ &= [21.17, 46.73] \end{aligned}$$

(approximation due to the random censoring mechanism).

Exponential Probability Plot of Remission Times with Maximum Likelihood Estimates and 95% Pointwise Confidence Intervals for $F(t; \theta)$



Likelihood with Exact and Right-Censored Observations

- The probability of a right-censored the observation is

$$L_i(\boldsymbol{\theta}) = \Pr(T \leq t_i) = F(\infty; \boldsymbol{\theta}) - F(t_i; \boldsymbol{\theta}) = 1 - F(t_i; \boldsymbol{\theta}).$$

- With exact and right-censored observations, the likelihood is:

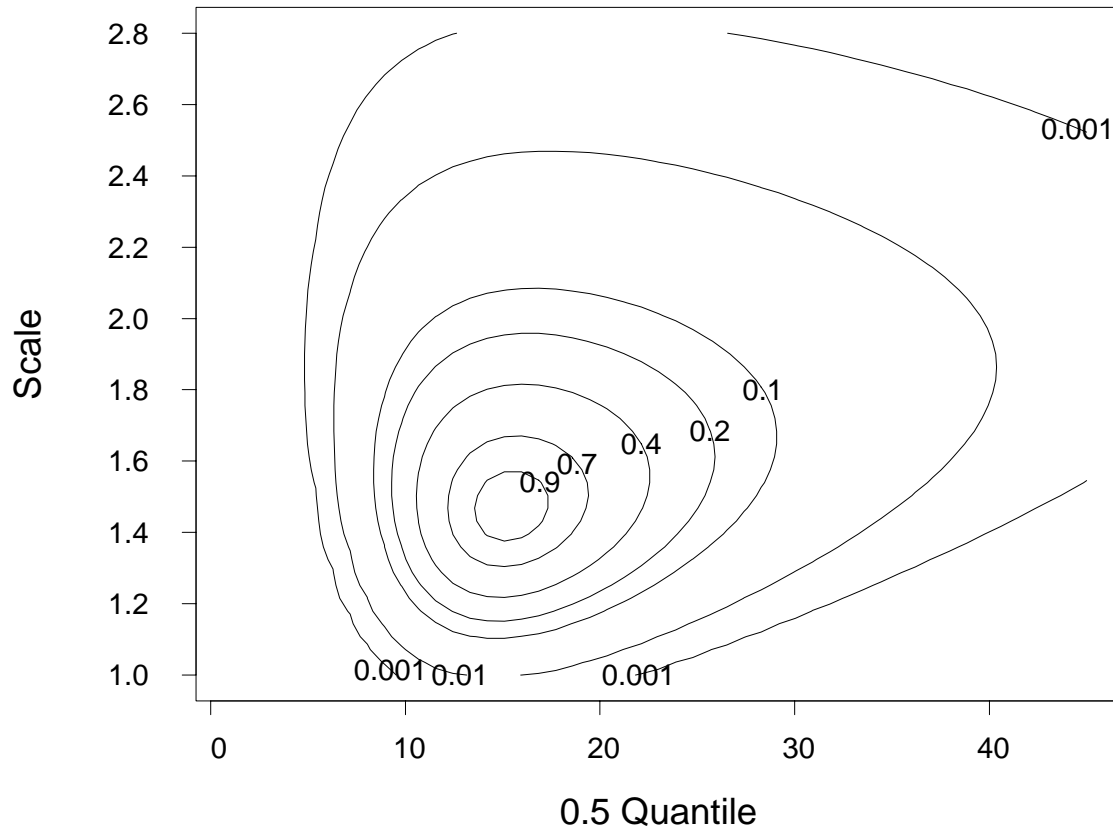
$$L(\boldsymbol{\theta}) = \prod_{i=1}^n \{f(t_i; \boldsymbol{\theta})\}^{\delta_i} \{1 - F(t_i; \boldsymbol{\theta})\}^{1-\delta_i},$$

where $\delta_i = 1$ for an “exact” observation and $\delta_i = 0$ for a right-censored observation.

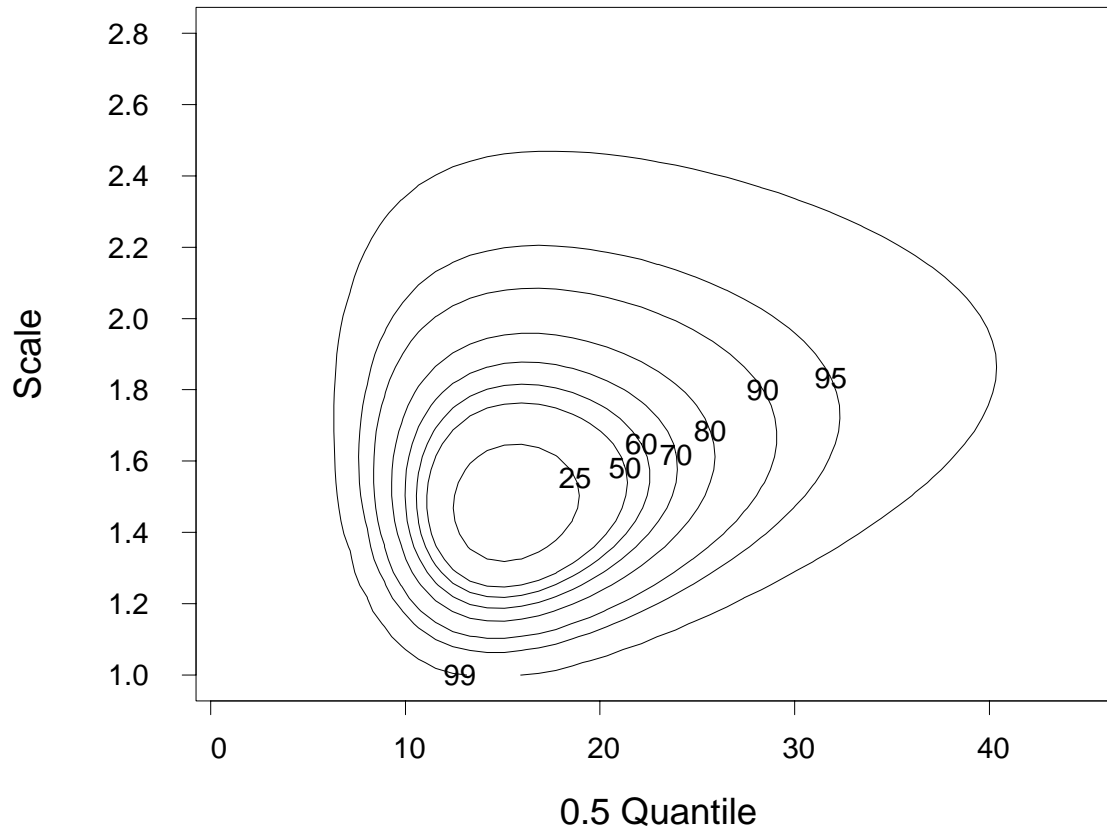
- For the lognormal distribution, $\boldsymbol{\theta} = (\mu, \sigma)$ and

$$L(\mu, \sigma) = \prod_{i=1}^n \left\{ \frac{1}{\sigma t} \phi_{\text{nor}} \left[\frac{\log(t_i) - \mu}{\sigma} \right] \right\}^{\delta_i} \left\{ 1 - \Phi_{\text{nor}} \left[\frac{\log(t_i) - \mu}{\sigma} \right] \right\}^{1-\delta_i}.$$

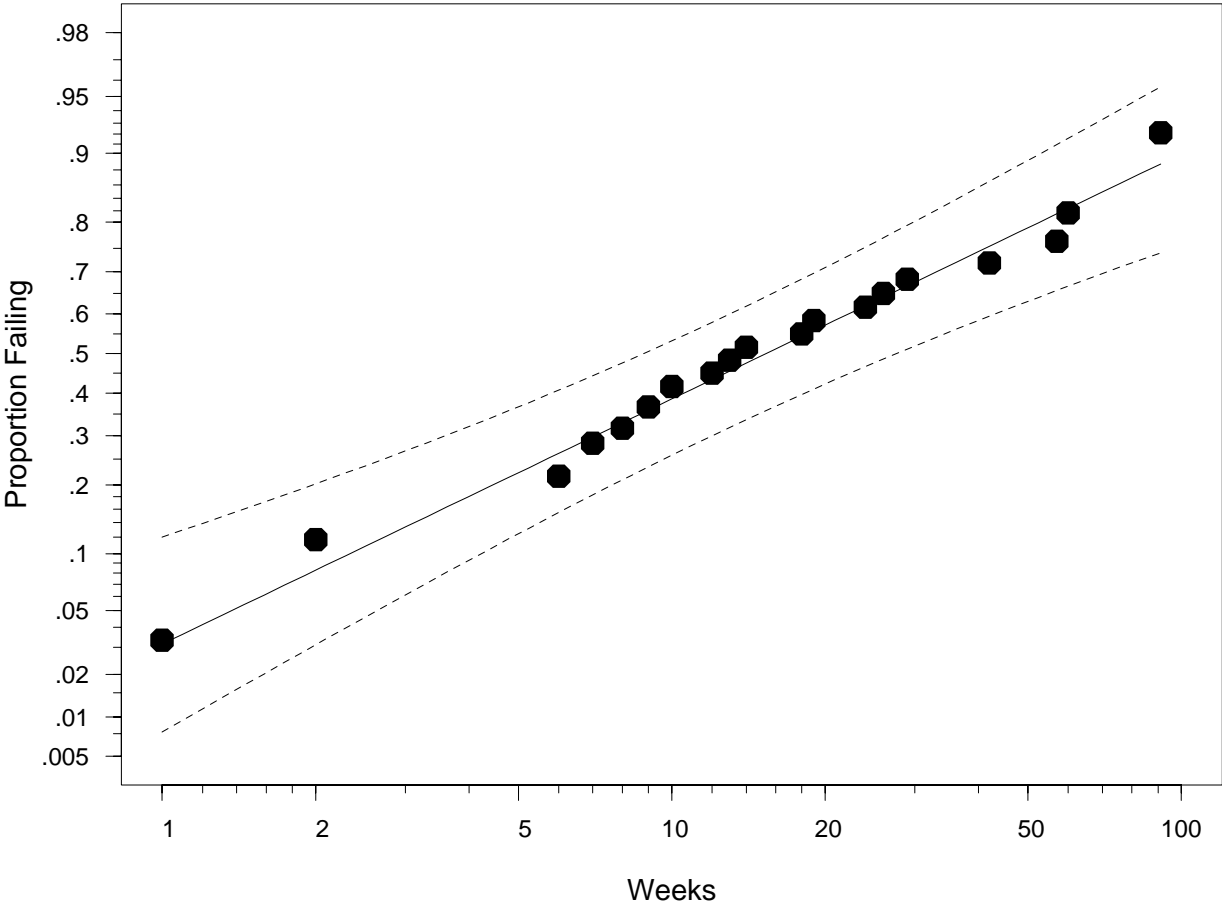
Lognormal Relative Likelihood Plot for the Remission Times



Lognormal Joint Confidence Region for the Remission Times

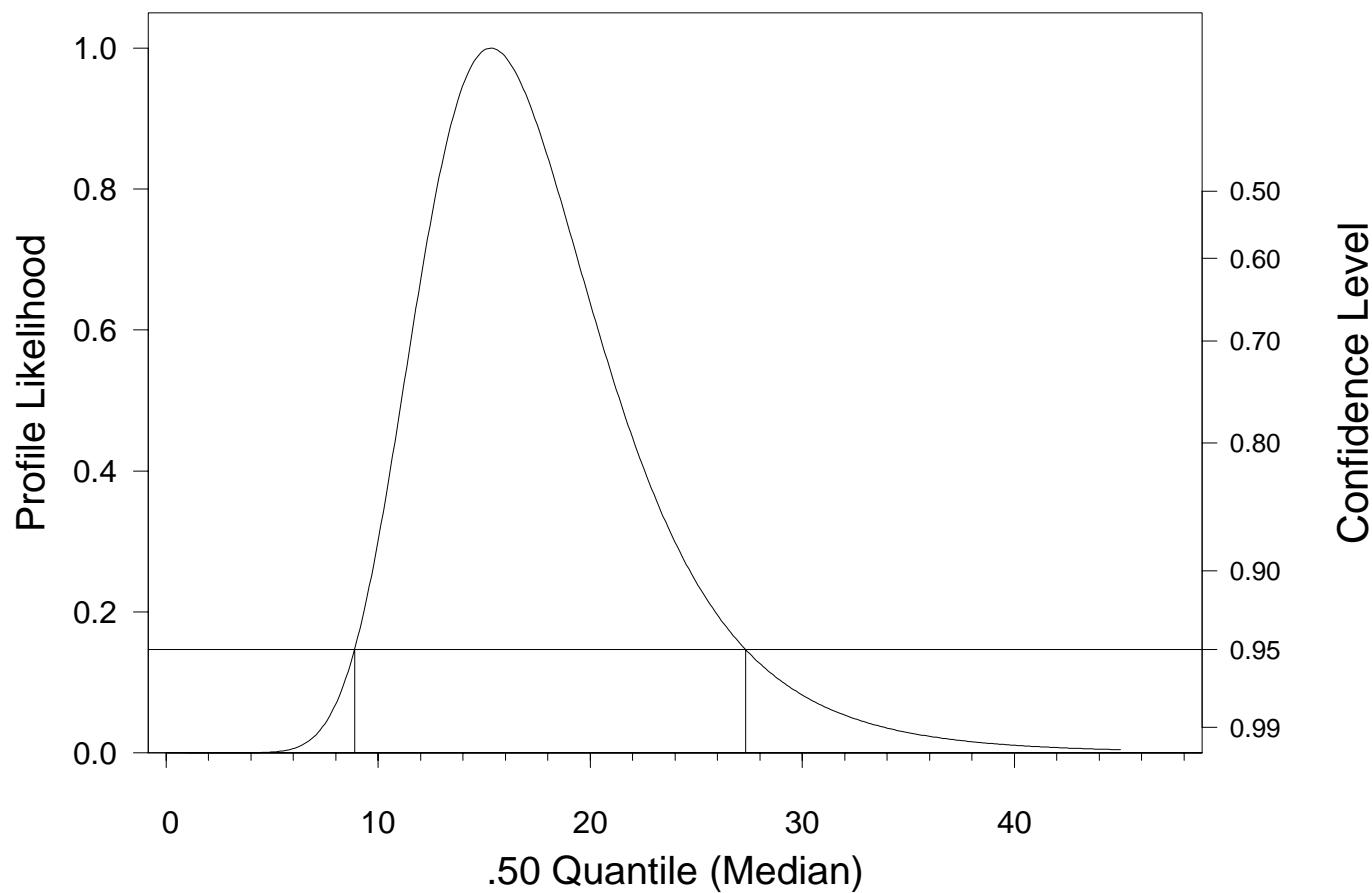


Lognormal Probability Plot of Remission Times with Maximum Likelihood Estimates and Approximate 95% Pointwise Confidence Intervals for $F(t; \mu, \sigma)$



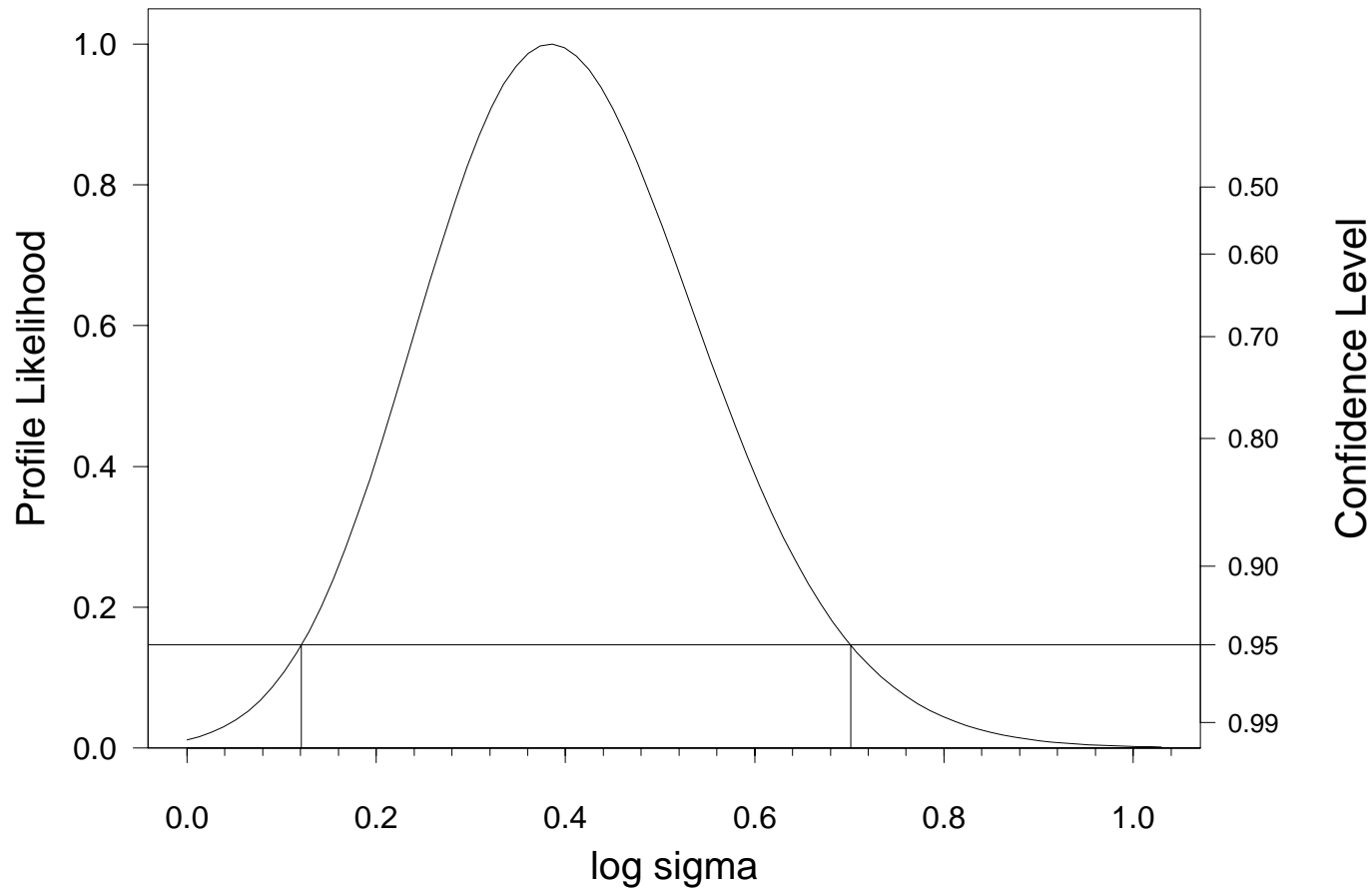
Lognormal Profile Likelihood $R[\exp(\mu)]$ ($\exp(\mu) = t_{.5}$) for the Remission Time Data

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



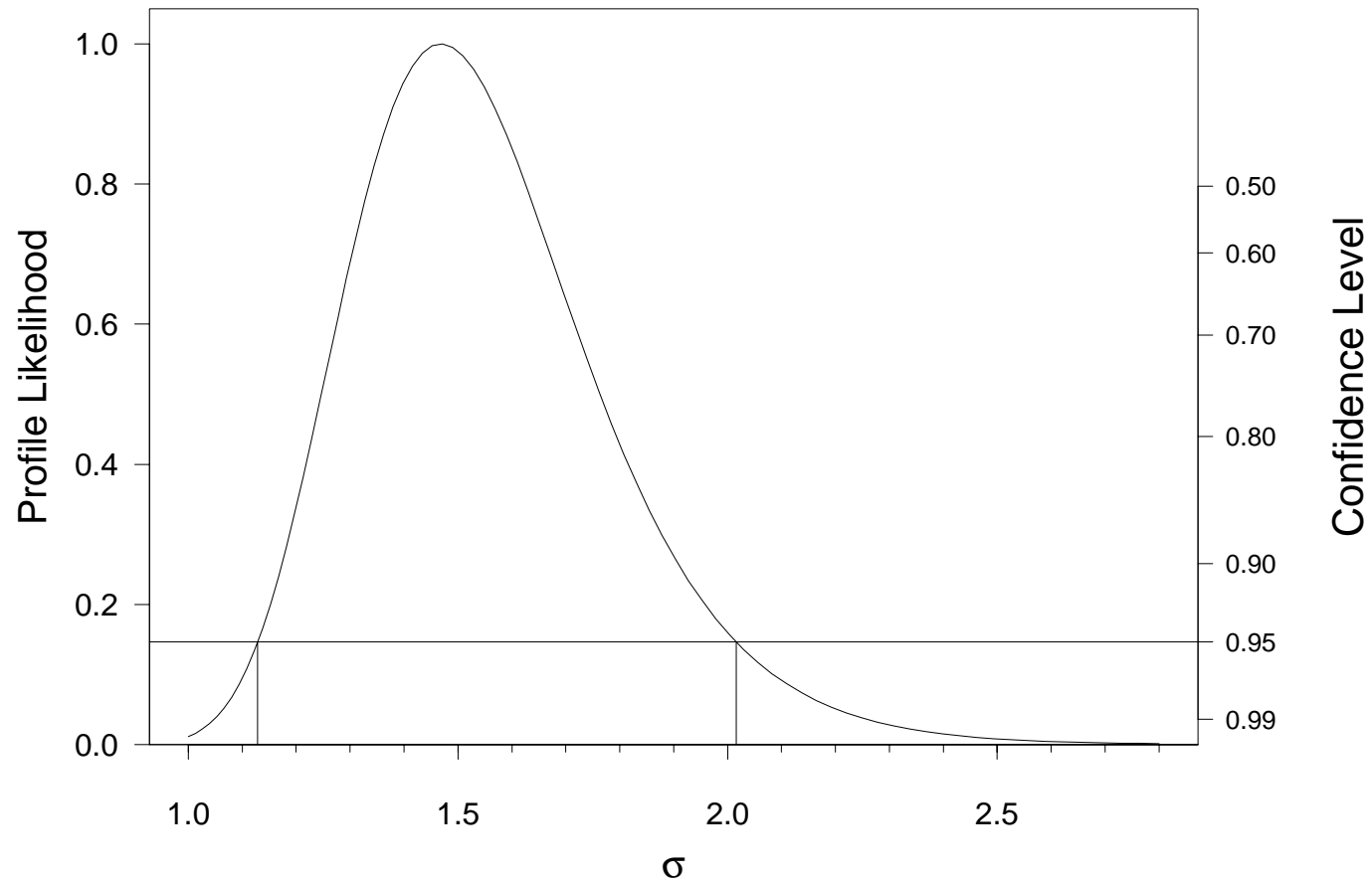
Lognormal Profile Likelihood $R(\log(\sigma))$ for the Remission Time Data

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



Lognormal Profile Likelihood $R(\sigma)$ for the Remission Time Data

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



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 of Meeker and Escobar (1998).

X-Ray Flux-Ratio Data from a Sample of Active Galaxies

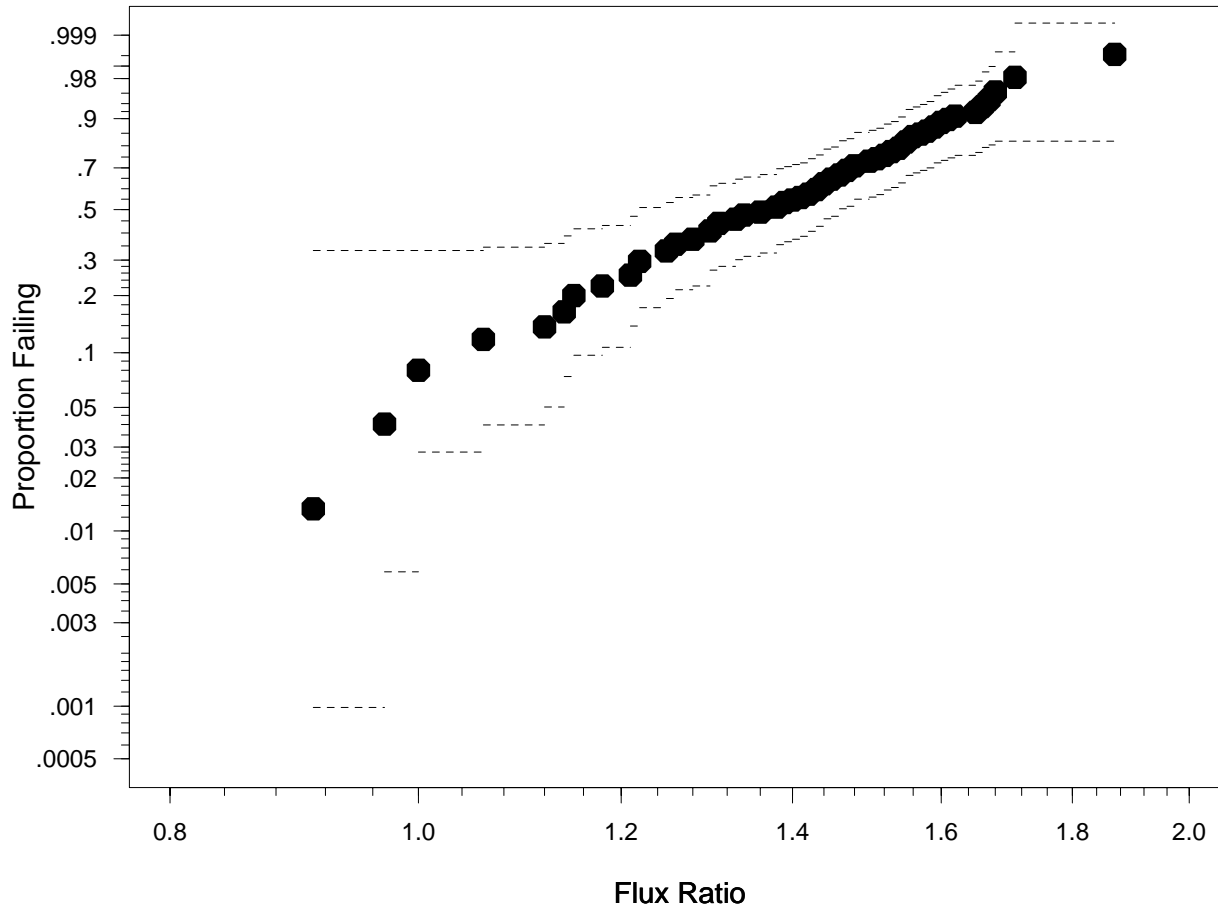
- Ratio of optical to X-ray flux for 107 active galaxies.
- For some galaxies, the flux values were too small to measure. This limitation in observation sensitivity causes left censoring.
- Investigators were interested in quantifying the distribution of this ratio among galaxies in a larger population, in order to gain potential insight into possible physical causes for the different types of radiation.

X-Ray to Optical Flux-Ratio Data

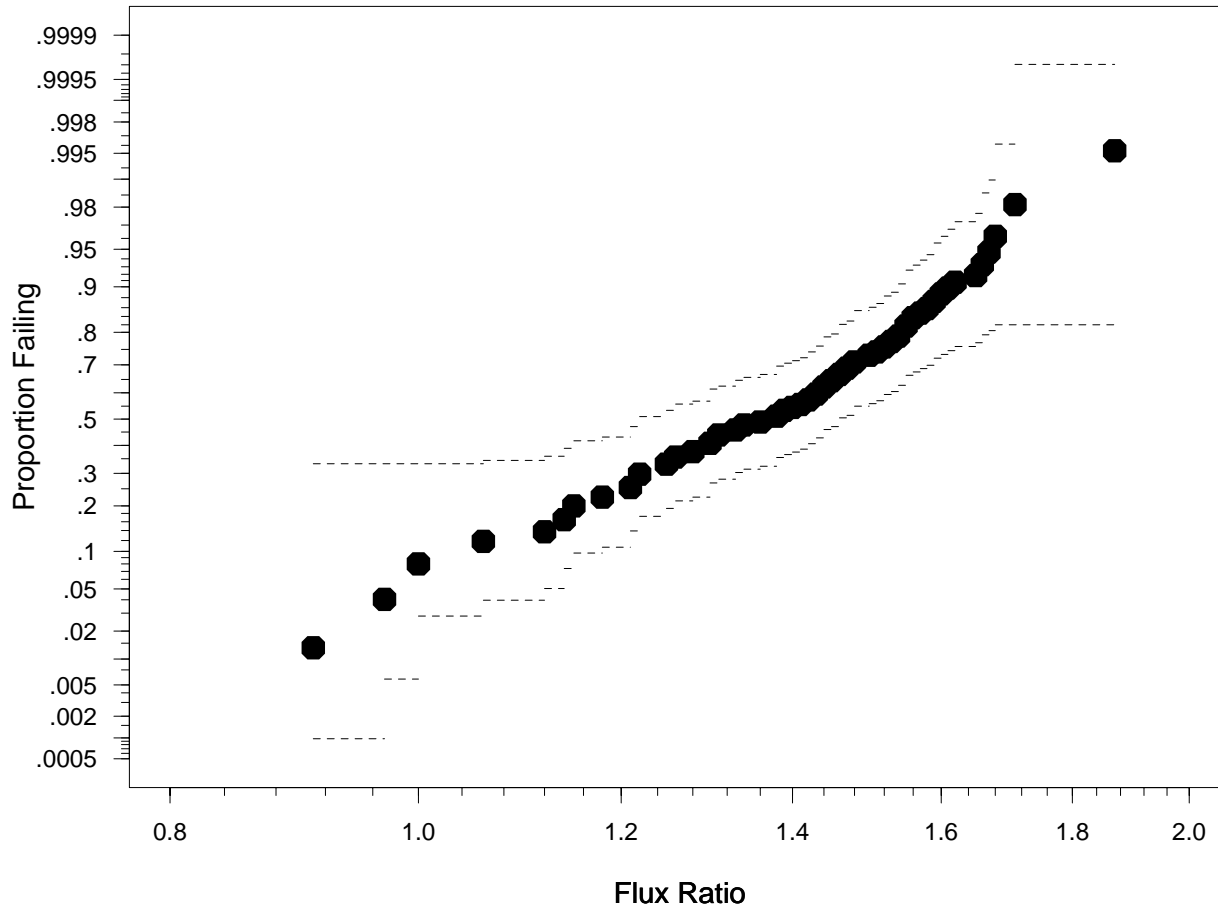
1.52	1.39	1.50	<1.50	<1.57	1.21	1.38	<1.63
1.41	1.06	1.00	1.54	1.68	1.48	<1.59	<1.68
<1.82	1.60	1.58	1.26	1.48	<1.63	1.31	1.47
1.56	1.66	1.15	1.28	1.34	1.57	1.00	1.46
1.25	1.44	1.67	1.38	1.43	1.26	1.46	1.22
1.42	1.15	1.14	<1.38	<1.82	<1.38	1.12	1.71
1.44	1.30	1.30	1.51	1.14	1.25	1.53	1.62
1.22	1.40	<1.25	<1.24	<1.17	<1.23	<1.56	<1.50
<1.15	<1.18	<1.40	<1.33	1.38	1.70	1.43	0.97
<1.47	1.54	1.55	<1.71	1.71	<1.32	1.48	<1.57
1.87	1.59	1.52	1.33	1.38	1.45	<1.04	1.21
<1.12	1.42	1.65	1.21	1.46	1.66	1.55	1.36
1.30	1.22	1.30	1.68	1.61	0.91	<1.36	1.18
1.55	1.44	1.59					

Observations marked with a < are left censored.

Weibull probability plot of X-Ray Flux-Ratio Data with 95% Simultaneous Confidence Bands for $F(t)$



Lognormal probability plot of X-Ray Flux-Ratio Data with 95% Simultaneous Confidence Bands for $F(t)$



Weibull Likelihood with Exact and Left-Censored Observations

- The probability of a left-censored the observation is

$$L_i(\boldsymbol{\theta}) = \Pr(-\infty < T \leq t_i) = F(t_i; \boldsymbol{\theta}) - F(-\infty; \boldsymbol{\theta}) = F(t_i; \boldsymbol{\theta}).$$

- With exact and left-censored observations, the likelihood is:

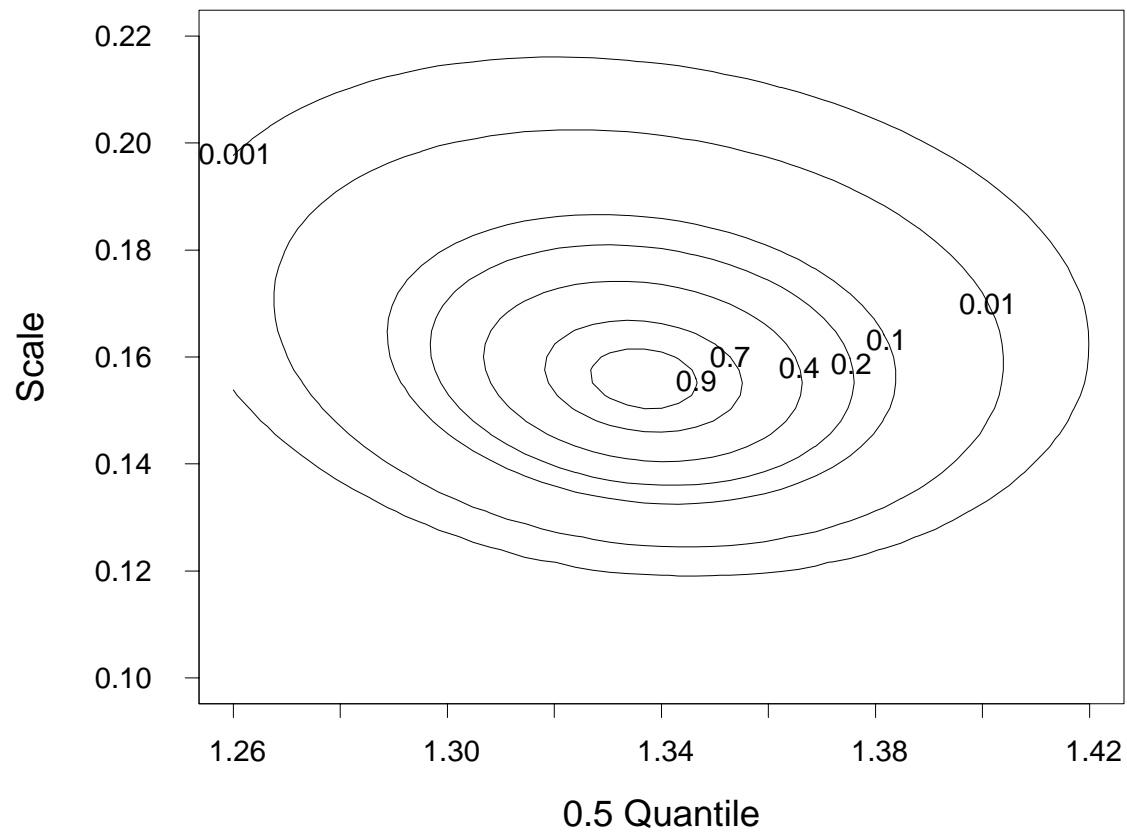
$$L(\boldsymbol{\theta}) = \prod_{i=1}^n \{F(t_i; \boldsymbol{\theta})\}^{1-\delta_i} \{f(t_i; \boldsymbol{\theta})\}^{\delta_i},$$

where $\delta_i = 1$ for an “exact” observation and $\delta_i = 0$ for a left-censored observation.

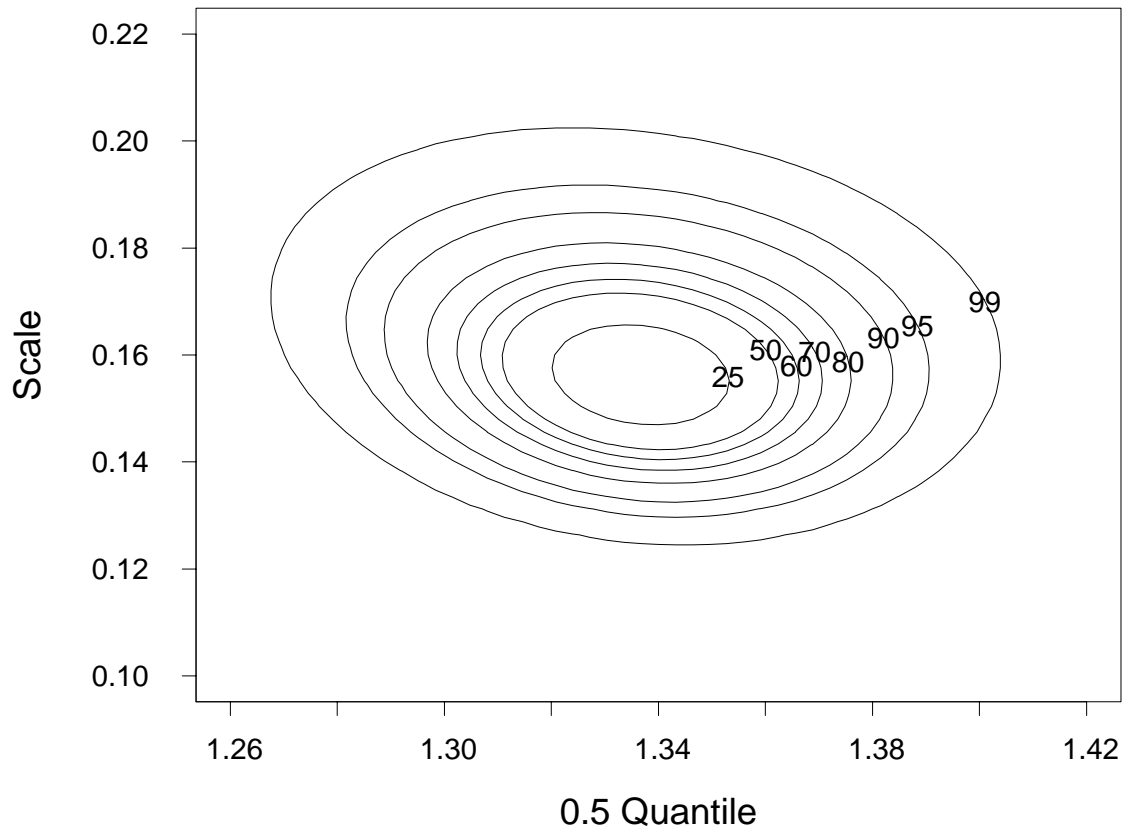
- For the Weibull distribution,

$$L(\mu, \sigma) = \prod_{i=1}^n \left\{ \Phi_{\text{sev}} \left[\frac{\log(t_i) - \mu}{\sigma} \right] \right\}^{1-\delta_i} \left\{ \frac{1}{\sigma t} \phi_{\text{sev}} \left[\frac{\log(t_i) - \mu}{\sigma} \right] \right\}^{\delta_i}.$$

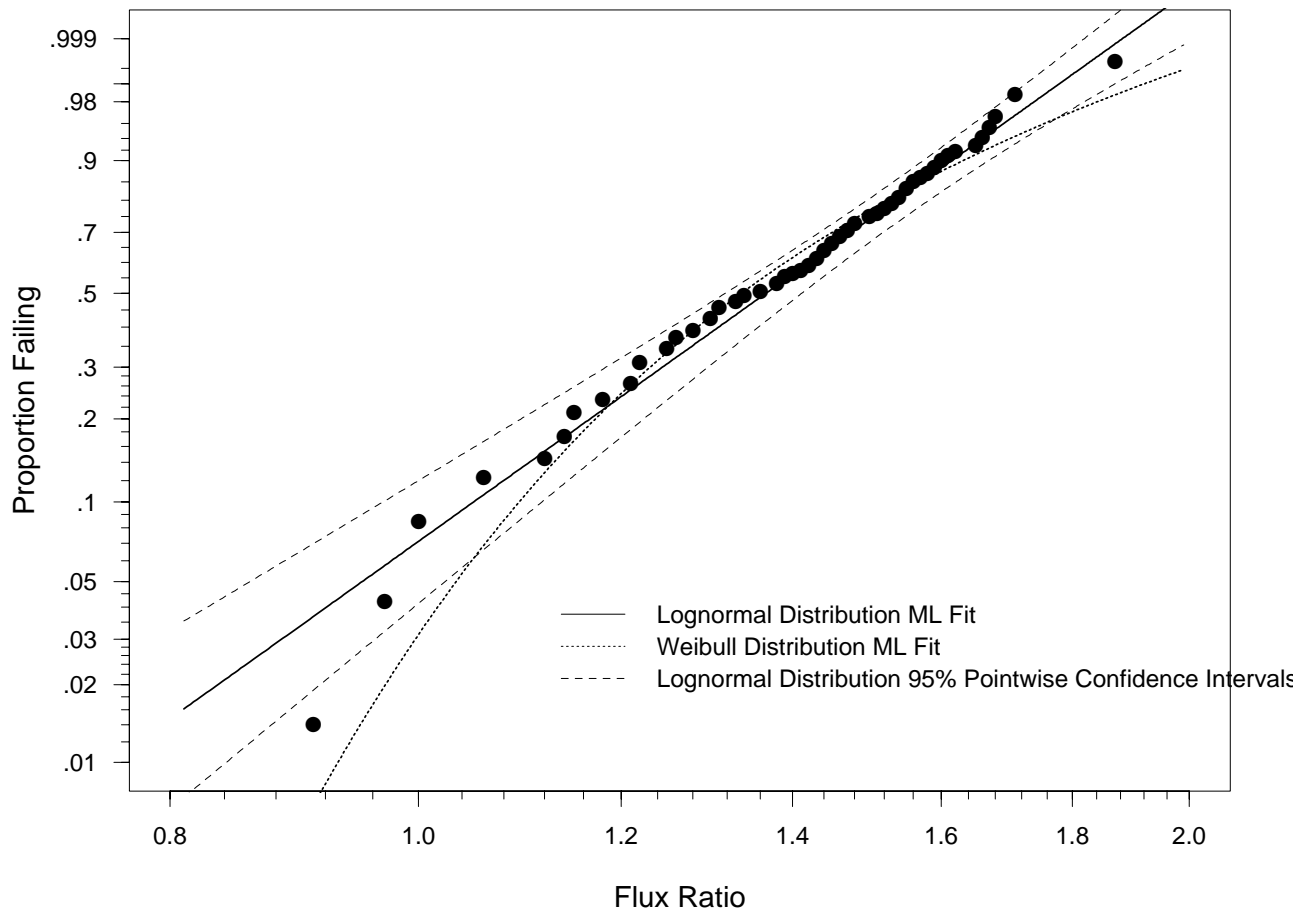
Relative Likelihood for the X-Ray Flux-Ratio Data and the Lognormal Distribution



Lognormal Joint Confidence Region for the X-Ray Flux-Ratio Data



Weibull Probability Plot of X-Ray Flux-Ratio Data with Weibull and Lognormal Maximum Likelihood Estimates and 95% Pointwise Confidence Intervals for the Weibull $F(t; \mu, \sigma)$



Comparison Profile Likelihoods for the Median of the X-Ray Flux-Ratio Distribution Using Weibull and Lognormal Distributions

