Chapter 7

Parametric Likelihood Fitting Concepts: Exponential Distribution

W. Q. Meeker, L. A. Escobar, and F. G. Pascual
Iowa State University, Louisiana State University, and Washington State University.

Based on Meeker, Escobar, and Pascual (2021): *Statistical Methods for Reliability Data, Second Edition*, John Wiley & Sons Inc.

May 24, 2021
10h 53min

Chapter 7
Parametric Likelihood Fitting Concepts: Exponential Distribution
Objectives

Topics discussed in this chapter are:

- How to compute a likelihood for a parametric model using interval-censored and right-censored data.
- The use of likelihood and Wald methods of computing confidence intervals for model parameters and other quantities of interest.
- The appropriate use of the density approximation for observations reported as exact failures.
- The effect that sample size has on confidence interval width and the likelihood shape.
- How to make exponential distribution inferences with zero-failures.

Chapter 7

Segment 1

Likelihood for Interval-Censored Data Times Between α-Particle Arrivals

Example: Times Between α-Particle Emissions of Americium-241

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

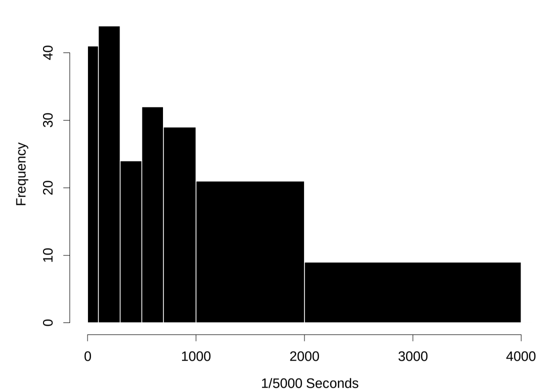
Data: Interarrival times (units: 1/5000 seconds).

- 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 4000 time units.
- Smaller samples analyzed here to illustrate sample size effect. We start the analysis with n =200.

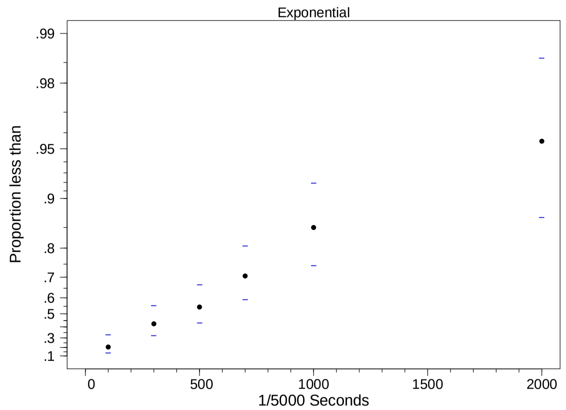
Data for α-Particle Emissions of Americium-241

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

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



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

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

  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

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

- As explained in Chapter 2, we take $\mathcal{C} = 1$.

- We will find values of $\boldsymbol{\theta}$ to make $L(\boldsymbol{\theta})$ large.

Exponential Distribution and Likelihood for Interval Data

Data: α-particle emissions of americium-241

- The exponential distribution cdf is

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

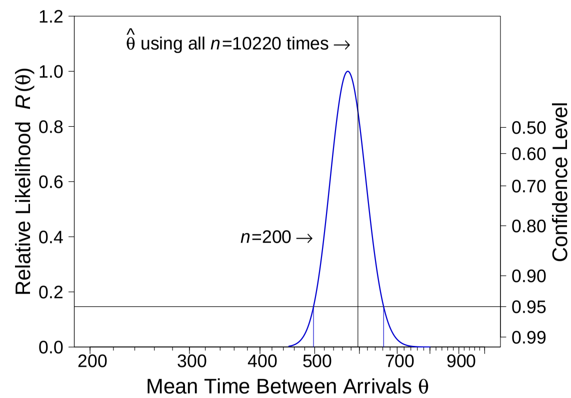
  where $\theta = \mathrm{E}(T)$ is the mean time between arrivals.

- The interval-data likelihood for the α-particle data is

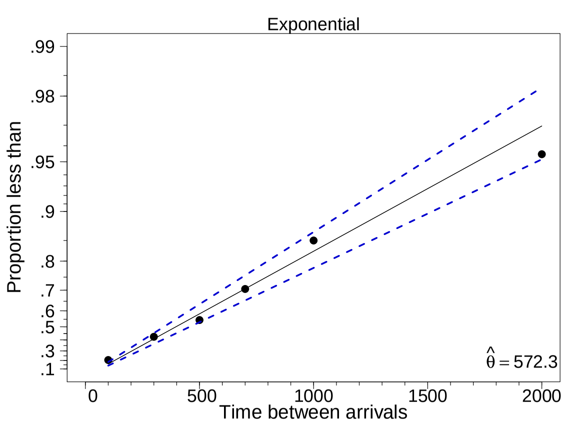
$$L(\theta) = \prod_{i=1}^{200} L_i(\theta) = \prod_{i=1}^{200} [F(t_i; \theta) - F(t_{i-1}; \theta)]$$
$$= \prod_{j=1}^{8} [F(t_j; \theta) - F(t_{j-1}; \theta)]^{d_j} = \prod_{j=1}^{8} \left[\exp\left(-\frac{t_{j-1}}{\theta}\right) - \exp\left(-\frac{t_j}{\theta}\right)\right]^{d_j},$$

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

$R(\theta) = L(\theta)/L(\widehat{\theta})$ 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 $F(t)$



Chapter 7

Segment 2

Likelihood as a Tool for Modeling and Inference and Methods for Confidence Intervals

Likelihood as a Tool for Modeling and 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).
- Observe effect of perturbations in data and model on likelihood (sensitivity, influence analysis).

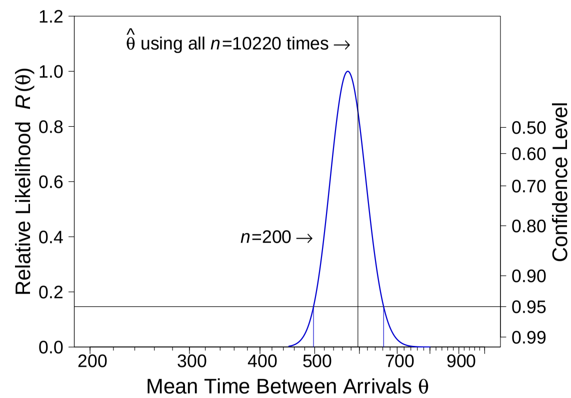
Likelihood as a Tool for Modeling and Inference (Continued)

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

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

- Relative likelihood for θ is
$$R(\theta) = \frac{L(\theta)}{L(\widehat{\theta})}.$$
- If evaluated at the true θ, then, asymptotically, $-2\log[R(\theta)]$ follows, a chi-square distribution with 1 degree of freedom.
- An approximate $100(1-\alpha)\%$ likelihood-based confidence region for θ is the set of all values of θ such that
$$-2\log[R(\theta)] < \chi^2_{(1-\alpha;1)}$$
or, equivalently, the set defined by
$$R(\theta) > \exp\left[-\chi^2_{(1-\alpha;1)}/2\right].$$
- A one-sided approximate $100(1-\alpha)\%$ lower or upper confidence bound is obtained by replacing $1-\alpha$ with $1-2\alpha$ and using the appropriate endpoint.
- General theory in the Appendix.

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



Wald Confidence Intervals for θ

- A $100(1-\alpha)\%$ Wald (or normal-approximation) confidence interval for θ is
$$[\underset{\sim}{\theta}, \quad \widetilde{\theta}] = \widehat{\theta} \pm z_{(1-\alpha/2)}\text{se}_{\widehat{\theta}}.$$
where $\text{se}_{\widehat{\theta}} = \sqrt{\left[-d^2\mathcal{L}(\theta)/d\theta^2\right]^{-1}}$ is evaluated at $\widehat{\boldsymbol{\theta}}$.
- Based on
$$Z_{\widehat{\theta}} = \frac{\widehat{\theta} - \theta}{\text{se}_{\widehat{\theta}}} \,\dot{\sim}\, \text{NORM}(0,1)$$
- From the definition of NORM(0, 1) quantiles,
$$\Pr\left[z_{(\alpha/2)} < Z_{\widehat{\theta}} \le z_{(1-\alpha/2)}\right] \approx 1-\alpha$$
implies that
$$\Pr\left[\widehat{\theta} - z_{(1-\alpha/2)}\text{se}_{\widehat{\theta}} < \theta \le \widehat{\theta} + z_{(1-\alpha/2)}\text{se}_{\widehat{\theta}}\right] \approx 1-\alpha.$$

Wald Confidence Intervals for θ (continued)

- A $100(1-\alpha)\%$ Wald (or normal-approximation) confidence interval for θ is
$$[\underset{\sim}{\theta}, \quad \widetilde{\theta}] = [\widehat{\theta}/w, \quad \widehat{\theta} \times w]$$
where $w = \exp[z_{(1-\alpha/2)}\text{se}_{\widehat{\theta}}/\widehat{\theta}]$. This follows after transforming (by exponentiation) the confidence interval
$$[\underset{\sim}{\log(\theta)}, \quad \widetilde{\log(\theta)}] = \log(\widehat{\theta}) \pm z_{(1-\alpha/2)}\text{se}_{\log(\widehat{\theta})}$$
which is based on
$$Z_{\log(\widehat{\theta})} = \frac{\log(\widehat{\theta}) - \log(\theta)}{\text{se}_{\log(\widehat{\theta})}} \,\dot{\sim}\, \text{NORM}(0,1)$$
- Because $\log(\widehat{\theta})$ is unrestricted in sign, $Z_{\log(\widehat{\theta})}$ is usually closer to a NORM(0, 1) distribution than is $Z_{\widehat{\theta}}$.

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.

$$[\underset{\sim}{\lambda}, \quad \widetilde{\lambda}] = [1/\widetilde{\theta}, \quad 1/\underset{\sim}{\theta}] = [1/662, \quad 1/498] = [0.00151, \quad 0.00201].$$

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

$$[\underset{\sim}{F}(t_e), \quad \widetilde{F}(t_e)] = [F(t_e;\widetilde{\theta}), \quad F(t_e;\underset{\sim}{\theta})]$$

$$[\underset{\sim}{F}(1000), \quad \widetilde{F}(1000)] = \left[1-\exp\left(\frac{-1000}{662}\right), \quad 1-\exp\left(\frac{-1000}{498}\right)\right] = [0.779, \quad 0.866].$$

Comparison of Confidence Intervals for the α-Particle Data

	All Times	Sample of Times	
	n =10,220	$n=200$	n=20
Mean Time Between Arrivals θ			
ML Estimate $\widehat{\theta}$	596	572	440
Standard Error $\text{se}_{\widehat{\theta}}$	6.1	41.7	101
95% Confidence Intervals for θ Based on			
Likelihood	[585, 608]	[498, 662]	[289, 713]
$Z_{\log(\widehat{\theta})} \stackrel{\cdot}{\sim}$ NORM$(0,1)$	[585, 608]	[496, 660]	[281, 690]
$Z_{\widehat{\theta}} \stackrel{\cdot}{\sim}$ NORM$(0,1)$	[585, 608]	[491, 654]	[242, 638]
Arrival Rate $\lambda \times 10^5$			
ML Estimate $\widehat{\lambda} \times 10^5$	168	175	227
Standard Error $\text{se}_{\widehat{\lambda}\times 10^5}$	1.7	13	52
95% Confidence Intervals for $\lambda \times 10^5$ Based on			
Likelihood	[164, 171]	[151, 201]	[140, 346]
$Z_{\log(\widehat{\lambda})} \stackrel{\cdot}{\sim}$ NORM$(0,1)$	[164, 171]	[152, 202]	[145, 356]
$Z_{\widehat{\lambda}} \stackrel{\cdot}{\sim}$ NORM$(0,1)$	[164, 171]	[149, 200]	[125, 329]

Chapter 7

Segment 3

Density Approximation for "Exact" Failures

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 $\Delta_i \to 0$ (Chapter 10).

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

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

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

- Using the observed curvature in the likelihood:

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

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

$$[\underset{\sim}{\theta}, \quad \widetilde{\theta}] = \left[\frac{2(TTT)}{\chi^2_{(1-\alpha/2;2r)}}, \quad \frac{2(TTT)}{\chi^2_{(\alpha/2;2r)}}\right].$$

Confidence Interval for the Mean Life of a New Insulating Material

- A life test for a new insulating material used $n = 25$ specimens which were tested simultaneously at a high voltage of 30 kV.
- The test was run until $r = 15$ of the specimens failed.
- The 15 failure times (hours) were recorded as:

1.15, 3.16, 10.38, 10.75, 12.53, 16.74, 22.54, 25.01, 33.02, 33.93, 36.17, 39.06, 44.56, 46.65, 55.93

Then $TTT = 1.15 + \cdots + 55.93 + 10 \times 55.93 = 950.88$ hours.

- The ML estimate of θ and a 95% confidence interval are:

$$\widehat{\theta} = 950.88/15 = 63.392 \text{ hours}$$

$$[\underset{\sim}{\theta}, \widetilde{\theta}] = \left[\frac{2(950.88)}{\chi^2_{(0.975;30)}}, \frac{2(950.88)}{\chi^2_{(0.025;30)}}\right] = \left[\frac{1901.76}{46.98}, \frac{1901.76}{16.79}\right] = [40.48, \quad 113.26].$$

Chapter 7

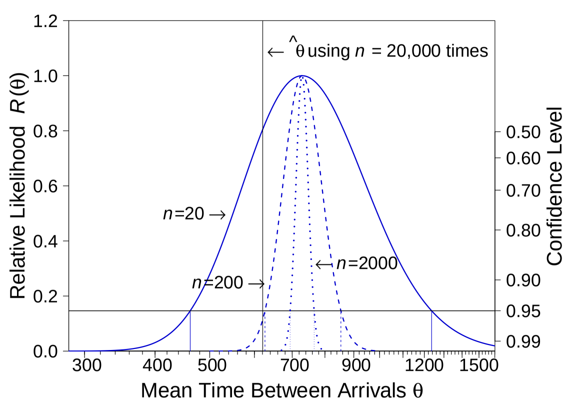
Segment 4

Effect of Sample Size on Confidence Interval Width and the Likelihood Shape

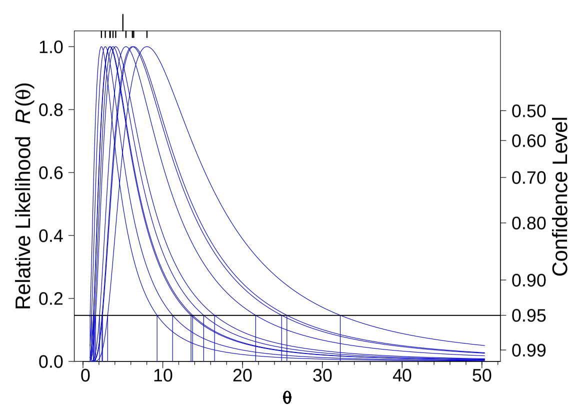
Example. α-Particle Pseudo Data Constructed with Constant Proportion within Each Bin

Time		Interarrival Times Frequency of Occurrence			
Interval Endpoint		Samples of Times			
lower t_{j-1}	upper t_j	n=20000	n=2000 d_j	n=200	n=20
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

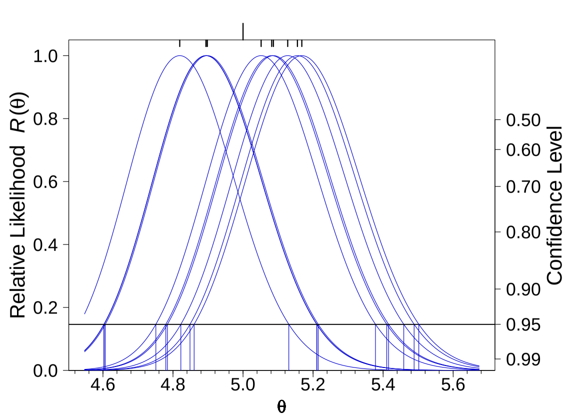
$R(\theta) = L(\theta)/L(\widehat{\theta})$ for the $n = 20$, 200, and 2000 Pseudo Data. Vertical Lines Give Corresponding Approximate 95% Likelihood-Based Confidence Intervals



Relative Likelihood for Simulated Exponential ($\theta = 5$) Samples of Size $n = 3$



Relative Likelihood for Simulated Exponential ($\theta = 5$) Samples of Size $n = 1000$



Effect of Sample Size on the Likelihood

- In large samples, the curvature at the maximum of the likelihood will be large, resulting in narrow confidence intervals.
- In large samples, the log-likelihood can be approximated well by a quadratic function.
- In small samples, the likelihood for the exponential mean can be skewed to the right.
- In small samples, there can be much variability of the width of confidence intervals.

Chapter 7

Segment 5

Exponential Distribution Inferences with Zero-Failures

Exponential Distribution Inferences with Zero Failures

- An ML estimate for the exponential distribution mean θ cannot be computed unless the available data contains one or more failures.
- For a sample of n units with running times $t_1, \ldots, t_n$ and an assumed exponential distribution, a <u>conservative</u> $100(1-\alpha)\%$ lower confidence bound for θ is

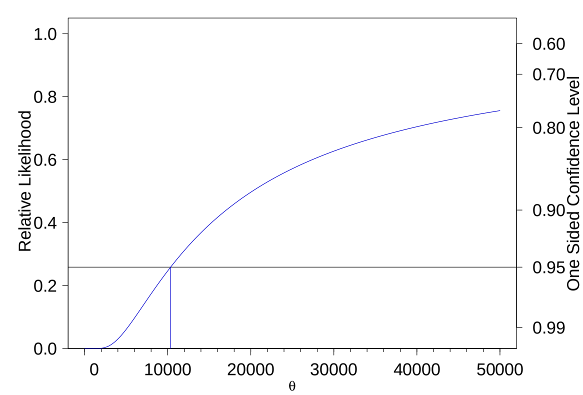
$$\underset{\sim}{\theta} = \frac{2(TTT)}{\chi^2_{(1-\alpha;2)}} = \frac{2(TTT)}{-2\log(\alpha)} = \frac{TTT}{-\log(\alpha)}.$$

- The lower bound $\underset{\sim}{\theta}$ 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.
- This bound is based on the fact that under the exponential failure-time distribution, with immediate replacement of failed units, the number of failures observed in a life test with a fixed total time on test has a Poisson distribution.

Analysis of the Diesel Generator Fan Data After 200 Hours of Service

- Suppose that all fans were removed from service after 200 hours of operation at which time all of the 70 fans were still running (zero failures).
- Then $TTT = 70 \times 200$=14,000 hours.
- The likelihood and relative likelihood functions (they are identical) are monotone increasing in θ but bounded above by 1.
- The ML estimate does not exist, but it is possible to find a lower confidence bound on θ.
- The likelihood method could be used, but because it depends on a large-sample approximation (and there are zero failures), the procedure might not be trustworthy.
- Fortunately, the conservative method is available.

Relative Likelihood for the Diesel Generator Fan Data After 200 Hours of Service and Zero Failures



Conservative Confidence Bound Based on the Diesel Generator Fan Data After 200 Hours of Service

- Again, with TTT=14,000 hours, a <u>conservative</u> 95% lower confidence bound on θ is

$$\underset{\sim}{\theta} = \frac{2(TTT)}{\chi^2_{(0.95;2)}} = \frac{28000}{5.991} = 4674.$$

- A conservative 95% upper confidence bound on $F(10000;\theta)$ is $\widetilde{F}(10000) = F(10000;\underset{\sim}{\theta}) = 1 - \exp(-10000/4674) = 0.882$.
- Using the entire data set, $\widehat{\theta} = 28{,}701$ and a likelihood-based approximate 95% lower confidence bound is $\underset{\sim}{\theta} =$ 18,485 hours.
- Again, using the entire data set, the upper confidence bound on $F(10000;\theta)$ is $\widetilde{F}(10000) = F(10000;\underset{\sim}{\theta}) = 1 - \exp(-10000/18485) = 0.4178$. This shows how little information is available from a short test with few or zero failures.

References

Berkson, J. (1966). Examination of randomness of α-particle emissions. In F. N. David (Editor), *Festschrift for J. Neyman, Research Papers in Statistics*. Wiley. []

Meeker, W. Q., L. A. Escobar, and F. G. Pascual (2021). *Statistical Methods for Reliability Data* (Second Edition). Wiley. [1]