Chapter 7

Parametric Likelihood Fitting Concepts: Exponential Distribution

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.

January 5, 2006
19h 14min

Chapter 7
Parametric Likelihood Fitting Concepts: Exponential Distribution
Objectives

- Show how to compute a likelihood for a parametric model using discrete data.
- Show how to compute a likelihood for samples containing right censored observations and left censored observations.
- Use a parametric likelihood as a tool for data analysis and inference.
- Illustrate the use of likelihood and normal-approximation methods of computing confidence intervals for model parameters and other quantities of interest.
- Explain the appropriate use of the density approximation for observations reported as exact failures.

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.

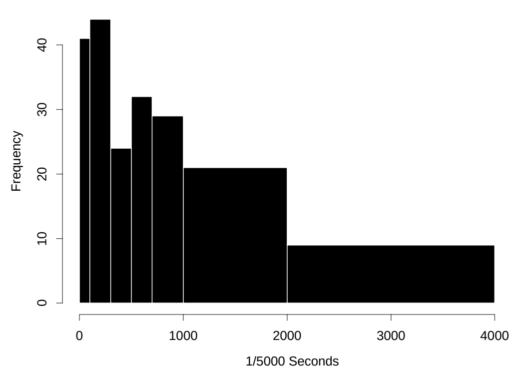
Data: Interarrival times (units: 1/5000 seconds).

- 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.

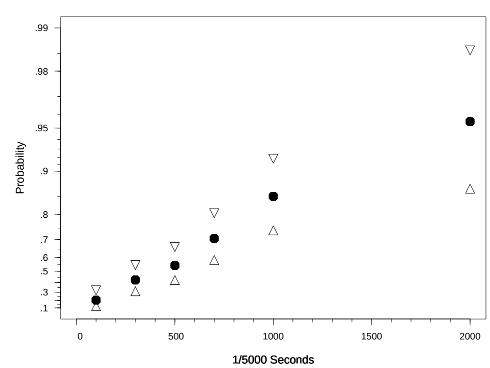
Data for α-Particle Emissions of Americium-241

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

- Want to estimate $\boldsymbol{\theta}$ and $g(\boldsymbol{\theta})$. We will find $\boldsymbol{\theta}$ to make $L(\boldsymbol{\theta})$ large.

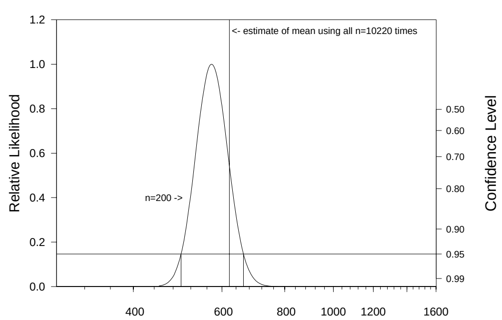
Exponential Distribution and Likelihood for Interval Data

Data: α-particle emissions of americium-241

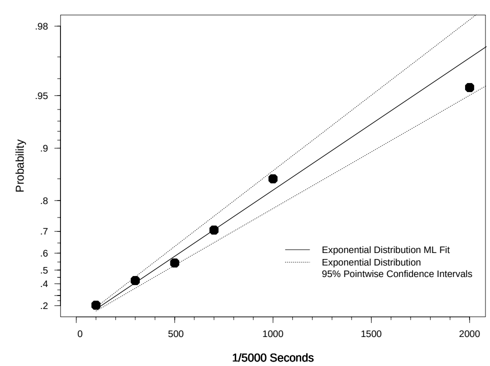
- The exponential distribution is
$$F(t; \theta) = 1 - \exp\left(-\frac{t}{\theta}\right), \quad t > 0.$$
$\theta = \text{E}(T)$, the mean time between arrivals.

- 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 jth interval (i.e., 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



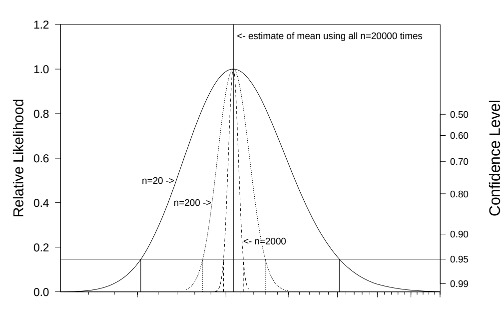
Exponential Probability Plot for the $n = 200$ Sample of α-Particle Interarrival Time Data. The Plot Also Shows Parametric Exponential ML Estimate and 95% Confidence Intervals for $F(t)$.



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

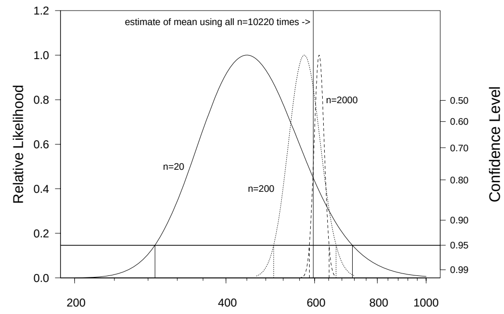
$R(\theta) = L(\theta)/L(\widehat{\theta})$ for the $n = 20$, 200, and 2000 Pseudo Data. Vertical Lines Give Corresponding Approximate 95% Likelihood-Based Confidence Intervals



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

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

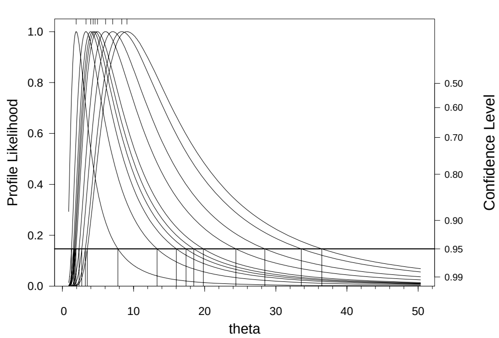
$$\mathcal{L}(\boldsymbol{\theta}) = \log[L(\boldsymbol{\theta})] = \sum_{i=1}^{n} \mathcal{L}_i(\boldsymbol{\theta}).$$

- 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).
- Observe effect of perturbations in data and model on likelihood (sensitivity, influence analysis).

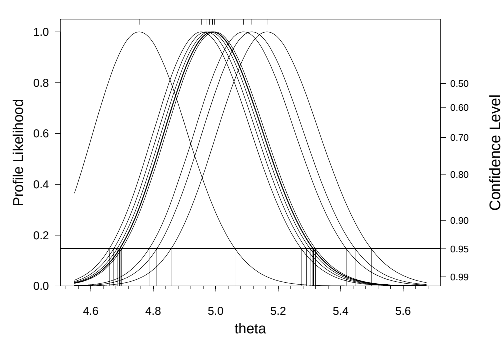
Likelihood as a Tool for Modeling/Inference (Continued)

- 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 confidence regions/intervals with χ^2 or simulation (or parametric bootstrap).
- Use **reparameterization** to study functions of $\boldsymbol{\theta}$.

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



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



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 chisquare 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

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

- General theory in the Appendix.

Normal-Approximation Confidence Intervals for θ

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

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

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

- Based on

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

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

$$\Pr\left[z_{(\alpha/2)} < Z_{\widehat{\theta}} \le z_{(1-\alpha/2)}\right] \approx 1-\alpha$$

implies that

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

Normal-Approximation Confidence Intervals for θ (continued)

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

$$[\underset{\sim}{\theta}, \quad \tilde{\theta}] = [\widehat{\theta}/w, \quad \widehat{\theta}\times w]$$

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

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

which is based on

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

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

Comparisons for α-Particle Data

	All Times	Sample of Times	
	n =10,220	$n=200$	n=20
ML Estimate $\widehat{\theta}$	596	572	440
Standard Error $\widehat{\text{se}}_{\widehat{\theta}}$	6.1	41.7	101
95% Confidence Intervals for θ Based on			
Likelihood	[585, 608]	[498, 662]	[289, 713]
$Z_{\log(\widehat{\theta})} \,\dot{\sim}\, \text{NOR}(0,1)$	[585, 608]	[496, 660]	[281, 690]
$Z_{\widehat{\theta}} \,\dot{\sim}\, \text{NOR}(0,1)$	[585, 608]	[491, 654]	[242, 638]
ML Estimate $\widehat{\lambda}\times 10^5$	168	175	227
Standard Error $\widehat{\text{se}}_{\widehat{\lambda}\times 10^5}$	1.7	13	52
95% Confidence Intervals for $\lambda \times 10^5$ Based on			
Likelihood	[164, 171]	[151, 201]	[140, 346]
$Z_{\log(\widehat{\lambda})} \,\dot{\sim}\, \text{NOR}(0,1)$	[164, 171]	[152, 202]	[145, 356]
$Z_{\widehat{\lambda}} \,\dot{\sim}\, \text{NOR}(0,1)$	[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.

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

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

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

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 $\Delta_i \to 0$.

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:
$$\widehat{\text{se}}_{\widehat{\theta}} = \sqrt{\left[-\frac{d^2\mathcal{L}(\theta)}{d\theta^2}\right]^{-1}\Bigg|_{\widehat{\theta}}} = \sqrt{\frac{\widehat{\theta}^2}{r}} = \frac{\widehat{\theta}}{\sqrt{r}}.$$
- 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
$$[\underset{\sim}{\theta}, \quad \tilde{\theta}] = \left[\frac{2(TTT)}{\chi^2_{(1-\alpha/2;2r)}}, \quad \frac{2(TTT)}{\chi^2_{(\alpha/2;2r)}}\right].$$

Confidence Interval for the Mean Life of a New Insulating Material

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

  1.08, 12.20, 17.80, 19.10, 26.00, 27.90, 28.20, 32.20, 35.90, 43.50, 44.00, 45.20, 45.70, 46.30, 47.80

  Then $TTT = 1.08 + \cdots + 47.80 + 10 \times 47.80 = 950.88$ hours.
- The ML estimate of θ and a 95% confidence interval are:
$$\begin{aligned} \widehat{\theta} &= 950.88/15 = 63.392 \text{ hours} \\ \left[\underset{\sim}{\theta}, \ \tilde{\theta}\right] &= \left[\frac{2(950.88)}{\chi^2_{(.975;30)}}, \frac{2(950.88)}{\chi^2_{(.025;30)}}\right] = \left[\frac{1901.76}{46.98}, \frac{1901.76}{16.79}\right] \\ &= [40.48, \quad 113.26]. \end{aligned}$$

Exponential Analysis With Zero Failures

- 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 conservative $100(1-\alpha)\%$ 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 (Assuming Removal After 200 Hours of Service)

- Here we do the analysis of the fan data after 200 hours of testing when all the fans were still running.
- Thus TTT=14,000 hours. A conservative 95% lower confidence bound on θ is
$$\underset{\sim}{\theta} = \frac{2(TTT)}{\chi^2_{(.95;2)}} = \frac{28000}{5.991} = 4674.$$
- 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.
- A conservative 95% upper confidence bound on $F(10000;\theta)$ is $\widetilde{F}(10000) = F(10000;\underset{\sim}{\theta}) = 1 - \exp(-10000/4674) =$ .882.

  This shows how little information comes from a short test with zero or few failures.

Other Topics in Chapter 7

- Inferences when there are no failures.