

Point Estimation by MLE

Lesson 5

Review

- Defined Likelihood
- Maximum Likelihood Estimation
 - Step 1: Construct Likelihood
 - Step 2: Maximize function
 - Take Log of likelihood function
 - Take derivative of function
 - Set derivative = 0
 - Solve for parameter

Point estimation

- **Goal: 'fit' a model to data**
- Single “best” value of a parameter that finds most support in the data set
- Contrasts with interval estimation (e.g. 95% confidence interval)

- Examples:
$$\mu_{SS} = \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

$$\sigma_{SS}^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$$

Maximum Likelihood

- Step 1: Write a likelihood function describing the likelihood of the observation
- Step 2: Find the value of the model parameter that maximized the likelihood

$$\frac{dL}{d\rho} = 0$$

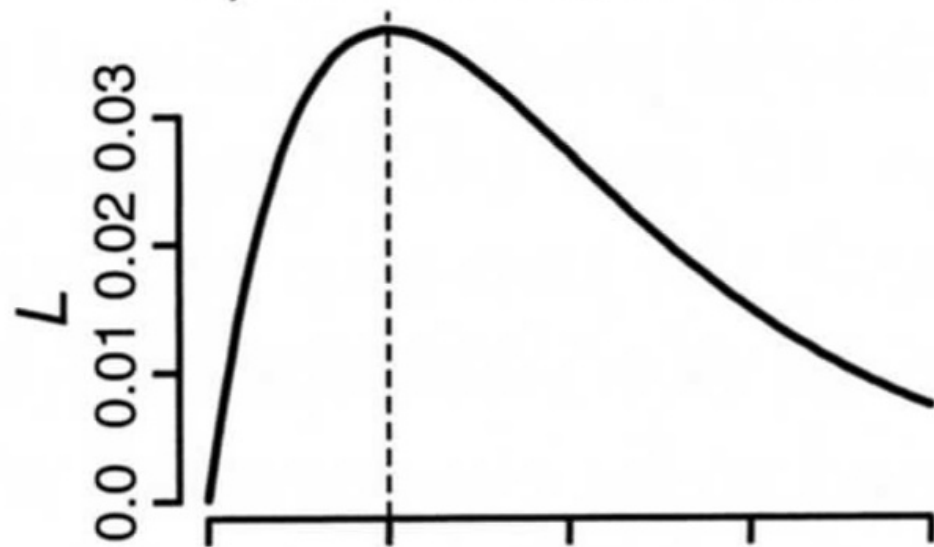
$$L = \rho e^{-\rho a}$$

$$\ln L = \ln \rho - \rho a$$

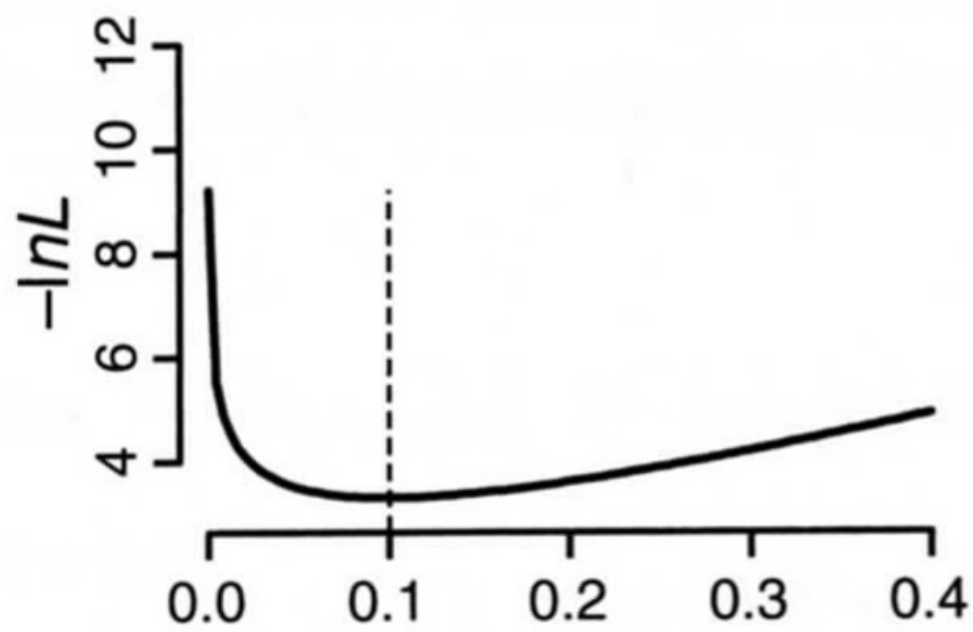
$$\frac{\partial \ln L}{\partial \rho} = \frac{1}{\rho} - a = 0$$

$$\rho_{ML} = \frac{1}{a} = 0.1 \text{ day}^{-1}$$

a) Likelihood function



b) $-\ln L$



A second data point

- Suppose a second plant dies at day 14
- Step 1: Define the likelihood

$$\begin{aligned} L &= \Pr(a_1, a_2 | \rho) \\ &= \Pr(a_2 | a_1, \rho) \Pr(a_1 | \rho) \\ &= \Pr(a_2 | \rho) \Pr(a_1 | \rho) \\ &\propto \text{Exp}(a_2 | \rho) \text{Exp}(a_1 | \rho) \end{aligned}$$

Assume measurements are independent

- Step 2: Find the maximum

$$L = \rho e^{-\rho a_1} \cdot \rho e^{-\rho a_2}$$

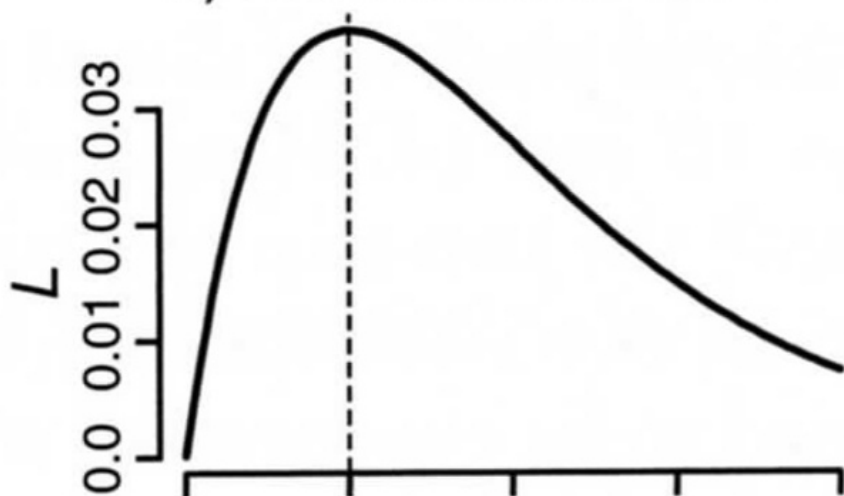
$$\ln L = 2 \ln \rho - \rho a_1 - \rho a_2$$

$$\frac{\partial \ln L}{\partial \rho} = \frac{2}{\rho} - (a_1 + a_2) = 0$$

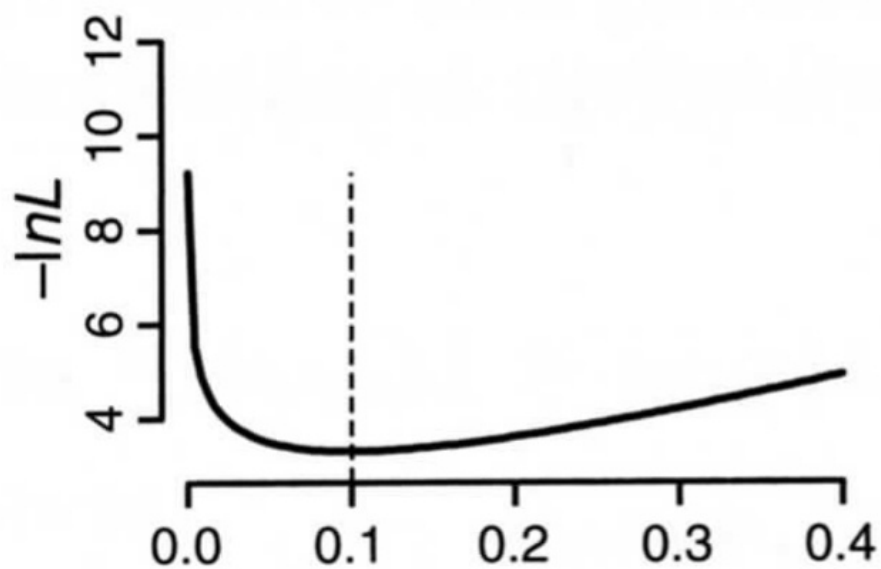
$$\rho_{ML} = \frac{2}{a_1 + a_2} = 0.0833 \text{ day}^{-1}$$

$n = 1$

a) Likelihood function

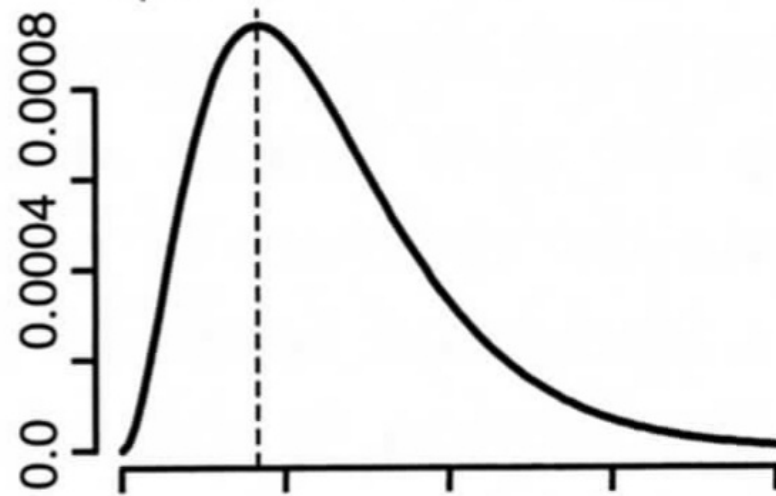


b) $-\text{Log likelihood}$

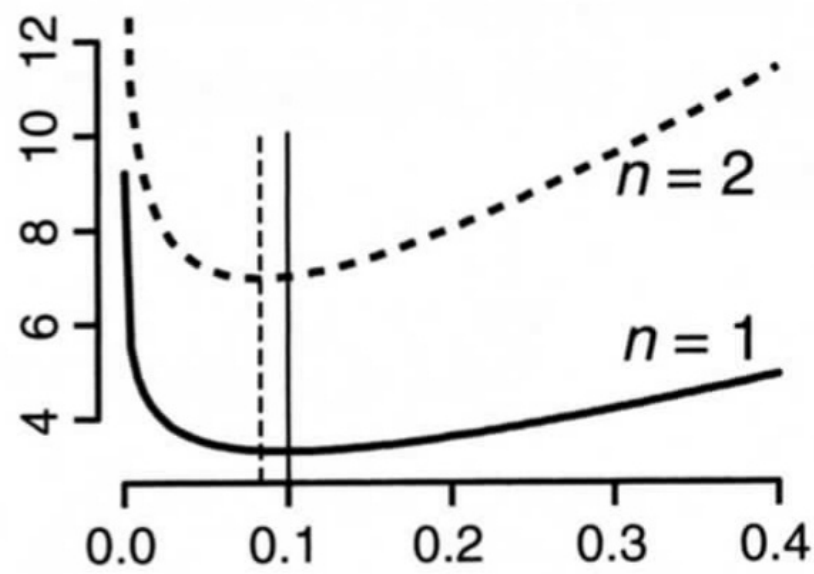


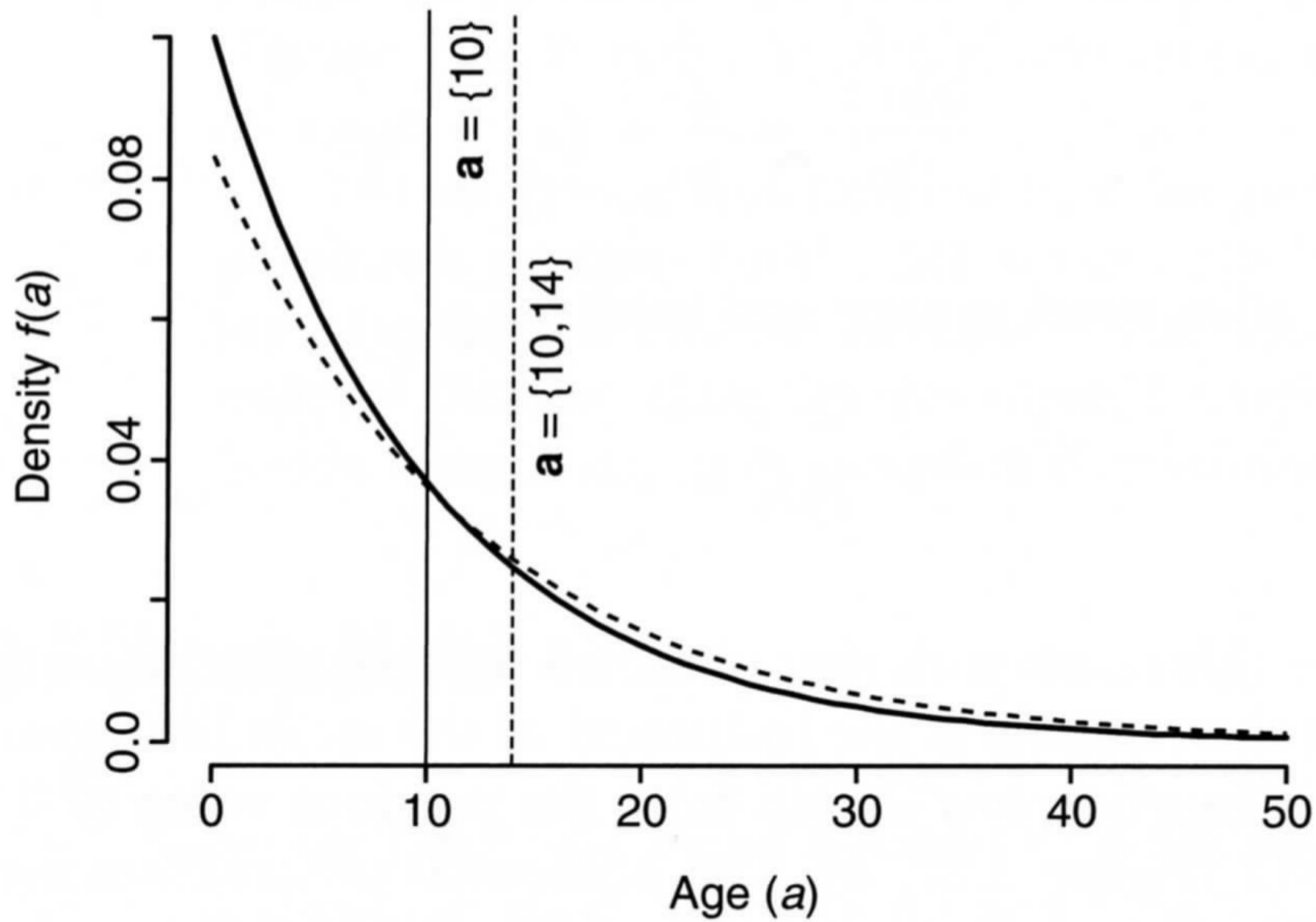
$n = 2$

c) Likelihood function



d) $-\text{Log likelihood}$





A whole data set

- Step 1: Define Likelihood

$$L = Pr(a_1, a_2, \dots, a_n | \rho)$$

Assume measurements are independent

$$= \prod_{i=1}^n Pr(a_i | \rho)$$

$$= \prod_{i=1}^n \text{Exp}(a_i | \rho)$$

- Step 2:
Find the maximum

$$L = \prod_{i=1}^n \rho e^{-\rho a_i}$$

$$\ln L = \sum_{i=1}^n (\ln \rho - \rho a_i)$$

$$= n \ln \rho - \rho \sum_{i=1}^n a_i$$

$$\frac{\partial \ln L}{\partial \rho} = \frac{n}{\rho} - \sum_{i=1}^n a_i = 0$$

$$\rho_{ML} = \frac{n}{\sum_{i=1}^n a_i} = 1/\bar{a}$$

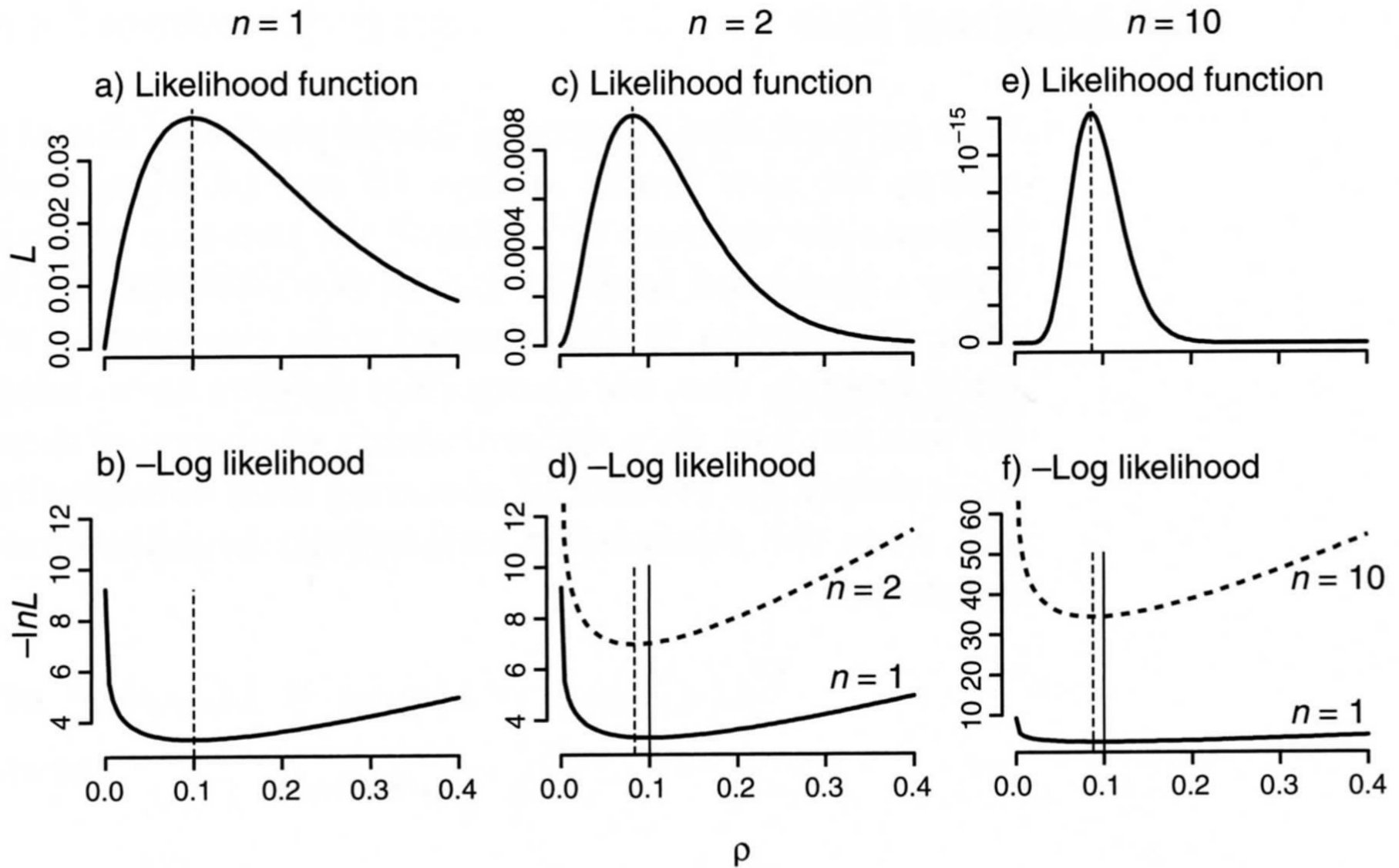


FIGURE 3.2. Likelihood functions for the exponential model with three different sample sizes. Note the different scales on the vertical axes.

Mortality – Alternate data

- First example based on knowing times individuals died
- What if instead we only knew the final outcome
 - N = number of individual plants
 - y = number that survived
- Want to find the MLE for
 - θ = probability of surviving to the end of the expm't

- Step 1: Write the likelihood

$$L = \text{Binom}(y|N, \theta) \propto \theta^y (1 - \theta)^{N-y}$$

- Step 1: Write the likelihood

$$L = \text{Binom}(y|N, \theta) \propto \theta^y (1 - \theta)^{N-y}$$

- Step 2: Find the maximum

$$\ln L = y \ln \theta + (N - y) \ln(1 - \theta)$$

- Step 1: Write the likelihood

$$L = \text{Binom}(y|N, \theta) \propto \theta^y (1 - \theta)^{N-y}$$

- Step 2: Find the maximum

$$\ln L = y \ln \theta + (N - y) \ln(1 - \theta)$$

$$\frac{\partial \ln L}{\partial \theta} = \frac{y}{\theta} - \frac{(N - y)}{(1 - \theta)}$$

- Step 1: Write the likelihood

$$L = \text{Binom}(y|N, \theta) \propto \theta^y (1 - \theta)^{N-y}$$

- Step 2: Find the maximum

$$\ln L = y \ln \theta + (N - y) \ln(1 - \theta)$$

$$\frac{\partial \ln L}{\partial \theta} = \frac{y}{\theta} - \frac{(N - y)}{(1 - \theta)}$$

$$(1 - \theta) y = \theta (N - y)$$

- Step 1: Write the likelihood

$$L = \text{Binom}(y|N, \theta) \propto \theta^y (1 - \theta)^{N-y}$$

- Step 2: Find the maximum

$$\ln L = y \ln \theta + (N - y) \ln(1 - \theta)$$

$$\frac{\partial \ln L}{\partial \theta} = \frac{y}{\theta} - \frac{(N - y)}{(1 - \theta)}$$

$$(1 - \theta) y = \theta (N - y)$$

$$\theta_{ML} = \frac{y}{N}$$

Mortality – combining age of death and survival till end

- Step 1 – Write the likelihood

$$L = \text{Binom}(y|N, \theta) \propto \theta^y (1 - \theta)^{N-y}$$

Mortality – combining age of death and survival till end

- Step 1 – Write the likelihood

$$L = \text{Binom}(y|N, \theta) \propto \theta^y (1 - \theta)^{N-y}$$

$$\theta = P(\text{survive to } T|\rho) = 1 - P(\text{dead by } T|\rho)$$

Mortality – combining age of death and survival till end

- Step 1 – Write the likelihood

$$L = \text{Binom}(y|N, \theta) \propto \theta^y (1 - \theta)^{N-y}$$

$$\theta = P(\text{survive to } T|\rho) = 1 - P(\text{dead by } T|\rho)$$

$$= 1 - (1 - e^{-\rho T}) = e^{-\rho T}$$



Exponential CDF

Mortality – combining age of death and survival till end

- Step 1 – Write the likelihood

$$L = \text{Binom}(y|N, \theta) \propto \theta^y (1 - \theta)^{N-y}$$

$$\theta = P(\text{survive to } T|\rho) = 1 - P(\text{dead by } T|\rho)$$

$$= 1 - (1 - e^{-\rho T}) = e^{-\rho T}$$

$$L = \text{Binom}(y|N, \rho) \propto e^{-\rho T y} (1 - e^{-\rho T})^{N-y}$$

- Step 2: Find the maximum

$$L = e^{-\rho T y} (1 - e^{-\rho T})^{N-y}$$

- Step 2: Find the maximum

$$L = e^{-\rho T y} (1 - e^{-\rho T})^{N-y}$$

$$\ln L = -y \rho T + (N - y) \ln(1 - e^{-\rho T})$$

- Step 2: Find the maximum

$$L = e^{-\rho T y} (1 - e^{-\rho T})^{N-y}$$

$$\ln L = -y \rho T + (N - y) \ln(1 - e^{-\rho T})$$

$$\frac{\partial \ln L}{\partial \rho} = -y T - (N - y) \frac{T e^{-\rho T}}{1 - e^{-\rho T}}$$

- Step 2: Find the maximum

$$L = e^{-\rho T y} (1 - e^{-\rho T})^{N-y}$$

$$\ln L = -y \rho T + (N - y) \ln(1 - e^{-\rho T})$$

$$\frac{\partial \ln L}{\partial \rho} = -y T - (N - y) \frac{T e^{-\rho T}}{1 - e^{-\rho T}}$$

$$\rho_{ML} = \frac{-\ln(y/N)}{T}$$

Survival Analysis

$$L = \text{Binom}(y|N, \theta) \propto \theta^y (1 - \theta)^{N-y}$$

Died



Survived

Survival Analysis

$$L = \text{Binom}(y|N, \theta) \propto \theta^y (1 - \theta)^{N-y}$$

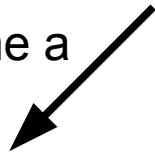
$$L = \prod_{i=1}^y \theta_i \times \prod_{i=1}^{n-y} (1 - \theta_i)$$

Survival Analysis

$$L = \text{Binom}(y|N, \theta) \propto \theta^y (1 - \theta)^{N-y}$$

$$L = \prod_{i=1}^y \theta_i \times \prod_{i=1}^{n-y} (1 - \theta_i)$$

Mortality at time a



$$L = \prod_{i=1}^y \rho e^{-\rho a_i} \times \dots$$

Survival Analysis

$$L = \text{Binom}(y|N, \theta) \propto \theta^y (1 - \theta)^{N-y}$$

$$L = \prod_{i=1}^y \theta_i \times \prod_{i=1}^{n-y} (1 - \theta_i)$$

Mortality at time a

Censored
at time c

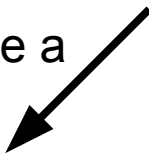
$$L = \prod_{i=1}^y \rho e^{-\rho a_i} \times \prod_{i=1}^k e^{-\rho c_i} \times \dots$$

Survival Analysis

$$L = \text{Binom}(y|N, \theta) \propto \theta^y (1 - \theta)^{N-y}$$

$$L = \prod_{i=1}^y \theta_i \times \prod_{i=1}^{n-y} (1 - \theta_i)$$

Mortality at time a



Censored
at time c



Survived to
time T



$$L = \prod_{i=1}^y \rho e^{-\rho a_i} \times \prod_{i=1}^k e^{-\rho c_i} \times e^{-\rho T(n-y-k)}$$

$$L = \prod_{i=1}^y \rho e^{-\rho a_i} \times \prod_{i=1}^k e^{-\rho c_i} \times e^{-\rho T(n-y-k)}$$

$$L = \prod_{i=1}^y \rho e^{-\rho a_i} \times \prod_{i=1}^k e^{-\rho c_i} \times e^{-\rho T(n-y-k)}$$

$$\ln L = y \ln \rho - \rho \sum_{i=1}^y a_i - \rho \sum_{i=1}^k c_i - \rho T(n-y-k)$$

$$L = \prod_{i=1}^y \rho e^{-\rho a_i} \times \prod_{i=1}^k e^{-\rho c_i} \times e^{-\rho T(n-y-k)}$$

$$\ln L = y \ln \rho - \rho \sum_{i=1}^y a_i - \rho \sum_{i=1}^k c_i - \rho T(n-y-k)$$

$$\frac{\partial \ln L}{\partial \rho} = \frac{y}{\rho} - \sum_{i=1}^y a_i - \sum_{i=1}^k c_i - T(n-y-k) = 0$$

$$L = \prod_{i=1}^y \rho e^{-\rho a_i} \times \prod_{i=1}^k e^{-\rho c_i} \times e^{-\rho T(n-y-k)}$$

$$\ln L = y \ln \rho - \rho \sum_{i=1}^y a_i - \rho \sum_{i=1}^k c_i - \rho T(n-y-k)$$

$$\frac{\partial \ln L}{\partial \rho} = \frac{y}{\rho} - \sum_{i=1}^y a_i - \sum_{i=1}^k c_i - T(n-y-k) = 0$$

$$\rho_{ML} = \frac{y}{\sum_{i=1}^y a_i - \sum_{i=1}^k c_i - T(n-y-k)}$$

MLE for the Normal

$$\begin{aligned} L &= \prod_{i=1}^n \mathbf{N}(y_i | \mu, \sigma^2) \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(y_i - \mu)^2}{2\sigma^2}\right] \end{aligned}$$

$$L = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp\left[\frac{-(y_i - \mu)^2}{2\sigma^2}\right]$$

$$\ln L = -\frac{n}{2} \ln(2\pi) - n \ln \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2$$

$$\begin{aligned} L &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(y_i - \mu)^2}{2\sigma^2}\right] \\ \ln L &= -\frac{n}{2} \ln(2\pi) - n \ln \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2 \\ \frac{\partial \ln L}{\partial \mu} &= \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \mu) = 0 \end{aligned}$$

$$\begin{aligned} L &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(y_i - \mu)^2}{2\sigma^2}\right] \\ \ln L &= -n \ln \sigma - \frac{n}{2} \ln(2\pi) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2 \\ \frac{\partial \ln L}{\partial \mu} &= \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \mu) = 0 \end{aligned}$$

$$\sum_{i=1}^n y_i = n\mu$$

$$\begin{aligned} L &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-(y_i - \mu)^2}{2\sigma^2}\right] \\ \ln L &= -n \ln \sigma - \frac{n}{2} \ln(2\pi) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2 \\ \frac{\partial \ln L}{\partial \mu} &= \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \mu) = 0 \end{aligned}$$

$$\sum_{i=1}^n y_i = n\mu$$

$$\mu_{ML} = \bar{y}$$

Normal Variance

$$\ln L = -n \ln \sigma - \frac{n}{2} \ln(2\pi) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2$$

Normal Variance

$$\ln L = -n \ln \sigma - \frac{n}{2} \ln(2\pi) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2$$

$$\frac{\partial \ln L}{\partial \sigma} = \frac{-n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (y_i - \mu)^2 = 0$$

Normal Variance

$$\ln L = -n \ln \sigma - \frac{n}{2} \ln(2\pi) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2$$

$$\frac{\partial \ln L}{\partial \sigma} = \frac{-n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (y_i - \mu)^2 = 0$$

$$\frac{n}{\sigma} = \frac{1}{\sigma^3} \sum_{i=1}^n (y_i - \mu)^2$$

Normal Variance

$$\ln L = -n \ln \sigma - \frac{n}{2} \ln(2\pi) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2$$

$$\frac{\partial \ln L}{\partial \sigma} = \frac{-n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (y_i - \mu)^2 = 0$$

$$\frac{n}{\sigma} = \frac{1}{\sigma^3} \sum_{i=1}^n (y_i - \mu)^2$$

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \mu)^2$$

Population Growth Rate

$$n_{t+1} = n_t \lambda = n_t e^r$$

$$n_{t+1} = n_t e^{r+\epsilon_t}$$

$$\ln n_{t+1} = \ln n_t + r + \epsilon_t$$

$$x_{t+1} = x_t + r + \epsilon_t$$

$$\epsilon_t \sim N(0, \sigma^2)$$

$$x_{t+1} \sim N(x_t + r, \sigma^2)$$

- Step 1: Write Likelihood

$$L = \prod_{t=1}^T N(x_{t+1} | x_t + r, \sigma^2)$$
$$= \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^T \exp \left[\frac{-1}{2\sigma^2} \sum_{t=1}^T (x_{t+1} - x_t - r)^2 \right]$$

- Step 1: Write Likelihood

$$L = \prod_{t=1}^T N(x_{t+1} | x_t + r, \sigma^2)$$
$$= \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^T \exp \left[\frac{-1}{2\sigma^2} \sum_{t=1}^T (x_{t+1} - x_t - r)^2 \right]$$

- Step 2: Maximize Likelihood
 - Take logs

$$\ln L = -T \ln \sigma - \frac{1}{2\sigma^2} \sum_{t=1}^T (x_{t+1} - x_t - r)^2$$

Solve for mean

$$\ln L = -T \ln \sigma - \frac{1}{2\sigma^2} \sum_{t=1}^T (x_{t+1} - x_t - r)^2$$

$$\frac{\partial \ln L}{\partial r} = \frac{1}{\sigma} \sum_{t=1}^T (x_{t+1} - x_t - r) = 0$$

Solve for mean

$$\ln L = -T \ln \sigma - \frac{1}{2\sigma^2} \sum_{t=1}^T (x_{t+1} - x_t - r)^2$$

$$\frac{\partial \ln L}{\partial r} = \frac{1}{\sigma} \sum_{t=1}^T (x_{t+1} - x_t - r) = 0$$

$$Tr = \sum_{t=1}^T (x_{t+1} - x_t)$$

Solve for mean

$$\ln L = -T \ln \sigma - \frac{1}{2\sigma^2} \sum_{t=1}^T (x_{t+1} - x_t - r)^2$$

$$\frac{\partial \ln L}{\partial r} = \frac{1}{\sigma} \sum_{t=1}^T (x_{t+1} - x_t - r) = 0$$

$$Tr = \sum_{t=1}^T (x_{t+1} - x_t)$$

$$r_{ML} = \frac{x_T - x_0}{T}$$

Solve for variance

$$\ln L = -T \ln \sigma - \frac{1}{2\sigma^2} \sum_{t=1}^T (x_{t+1} - x_t - r)^2$$

Solve for variance

$$\ln L = -T \ln \sigma - \frac{1}{2\sigma^2} \sum_{t=1}^T (x_{t+1} - x_t - r)^2$$

$$\frac{\partial \ln L}{\partial \sigma} = \frac{-T}{\sigma} + \frac{1}{\sigma^3} \sum_{t=1}^T (x_{t+1} - x_t - r)^2 = 0$$

Solve for variance

$$\ln L = -T \ln \sigma - \frac{1}{2\sigma^2} \sum_{t=1}^T (x_{t+1} - x_t - r)^2$$

$$\frac{\partial \ln L}{\partial \sigma} = \frac{-T}{\sigma} + \frac{1}{\sigma^3} \sum_{t=1}^T (x_{t+1} - x_t - r)^2 = 0$$

$$\sigma_{ML}^2 = \frac{1}{T} \sum_{t=1}^T (x_{t+1} - x_t - r)^2$$

Homework

- You want to know the density of fish in a set of experimental ponds
- You observe the following counts in ten ponds:
5,6,7,3,6,5,8,4,4,3
- What is your process model?
- What is your data model?
- Solve for the analytical MLE
- What is the estimate for this population?