Maximum Likelihood Estimators

Learning Outcomes

- Maximum Likelihood Estimators
- Log-Likelihood Functions

Background Information

Estimators

An *estimator* is an operation computing the value of an estimate, that targets the parameter, using measurements from a sample.

$$\mathbf{X} = \{ x, \dots, x_n \}$$

Population

$F(\theta)$

$$\hat{\theta} = g(\mathbf{X})$$

Estimator

$$\hat{\theta} \to \theta$$

Data

Let $X_1, \ldots, X_n \stackrel{iid}{\sim} F(\boldsymbol{\theta})$ where $F(\cdot)$ is a known distribution function and $\boldsymbol{\theta}$ is a vector of parameters. Let $\boldsymbol{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$, be the sample collected.

$\underset{\sim}{X}$

$F(\theta) \longrightarrow \{X_1, \ldots, X_n\}$

$F(\theta) = N(\mu, \sigma^2)$

MLE

Likelihood Function

Using the joint pdf or pmf of the sample $\boldsymbol{X}$, the likelihood function is a function of $\boldsymbol{\theta}$, given the observed data $\boldsymbol{X} = \boldsymbol{x}$, defined as

pdf

$$L(\boldsymbol{\theta}|\boldsymbol{x}) = f(\boldsymbol{x}|\boldsymbol{\theta})$$

$\sum_{i=1}^{n}$ add

$\prod_{i=1}^{n}$ multiply all

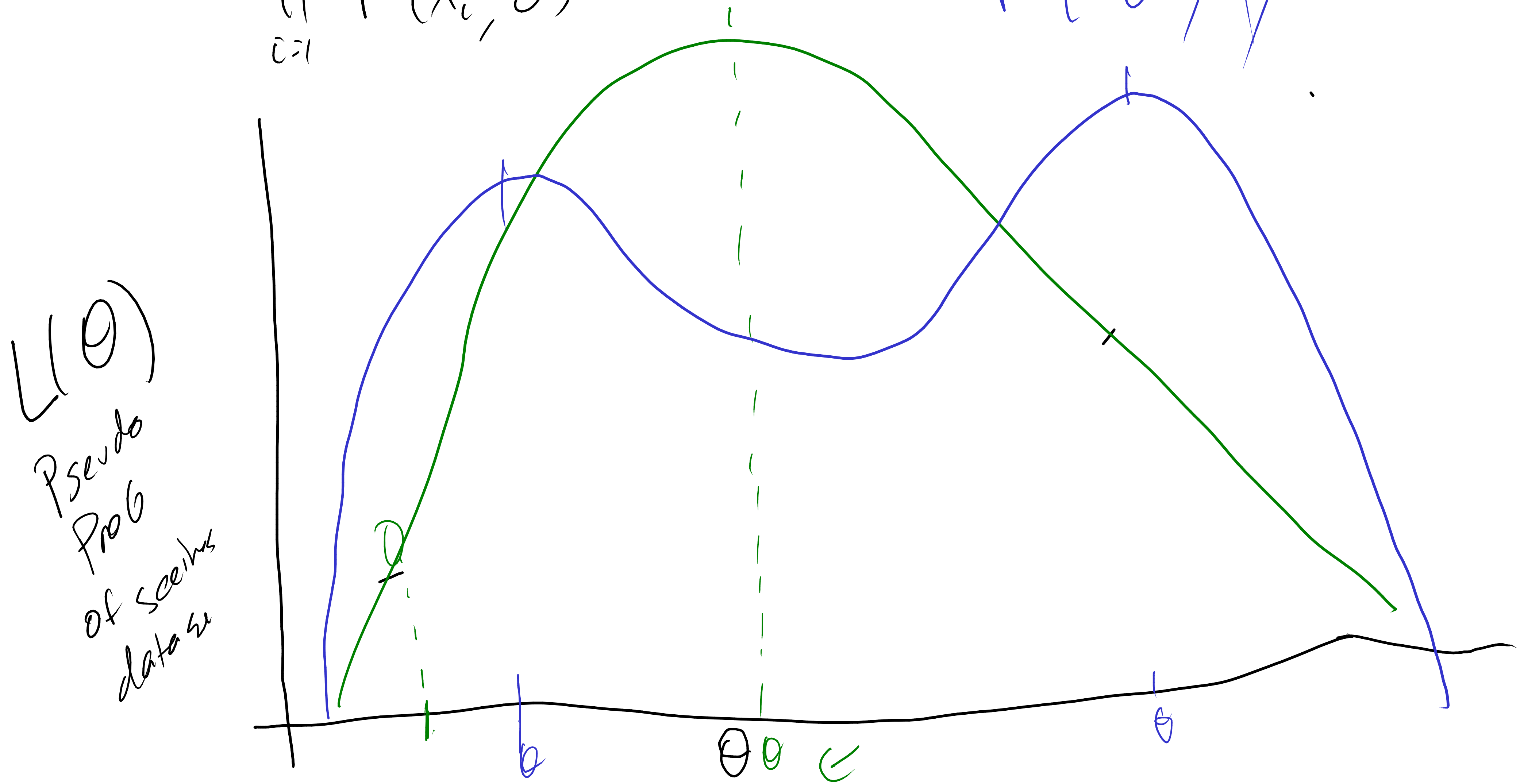
If the data is iid, then

$$f(\boldsymbol{x}|\boldsymbol{\theta}) = \prod_{i=1}^{n} f(x_i|\boldsymbol{\theta})$$

$X \sim F(\theta)$ $f(\theta)$ $\underline{X} = \{1, 2, 3, 4\}$ $f(x;\theta)$

$$L(\theta) = f(1;\theta)\, f(2;\theta)\, f(3;\theta)\, f(4;\theta)$$

$$\prod_{i=1}^{4} f(x_i;\theta)$$ $f(\theta, \theta)$



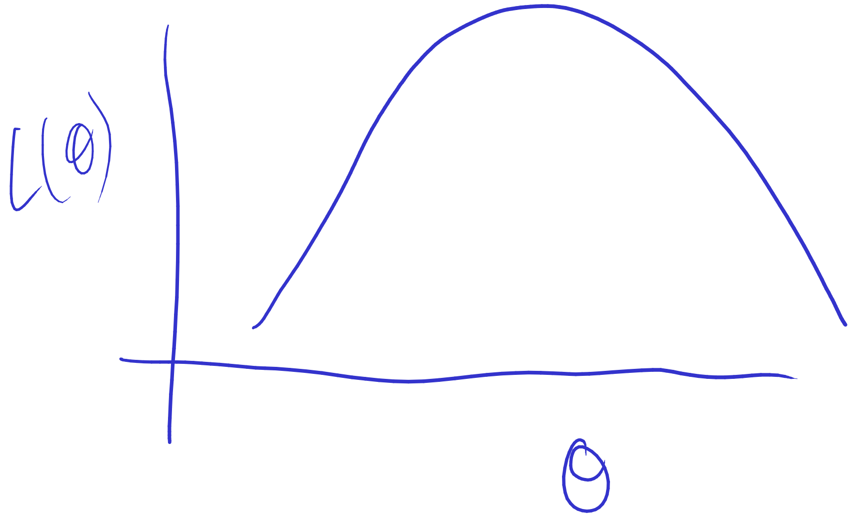
Log-Likelihood Function

If $\ln\{L(\boldsymbol{\theta})\}$ is monotone of $\boldsymbol{\theta}$, then maximizing $\ell(\boldsymbol{\theta}) = \ln\{L(\boldsymbol{\theta})\}$ will yield the maximum likelihood estimators.

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

$$L'(\theta) = 0$$

$$\theta$$



$$\ell(\theta) = \ln\{L(\theta)\}$$

$$\ell'(\theta) = 0$$

$$\ell(\theta) = \sum_{i=1}^{n} \ln(f(x_i;\theta))$$

Maximum log-Likelihood Estimator

The maximum likelihood estimator are the estimates of $\boldsymbol{\theta}$ that maximize $\ell(\boldsymbol{\theta})$.

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}}\ \ell(\theta)$$

hat indicates estimator

Example

Poisson Distribution

λ = rate

Let $\underbrace{X_1, \ldots, X_n}$ $\overset{iid}{\sim}$ $\text{Pois}(\lambda)$, show that the MLE of λ is $\bar{x}$.

$$f(x_i;\lambda) = \frac{e^{-\lambda}\lambda^{x_i}}{x_i!}$$

$$L(\lambda) = \prod_{i=1}^{n} f(x_i;\lambda)$$

$$L(\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!}$$

$$\ell(\lambda) = \ln\left[\prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!}\right]$$

$$= \sum_{i=1}^{n} \ln\left[\frac{e^{-\lambda}\lambda^{x_i}}{x_i!}\right]$$

$$= \sum_{i=1}^{n} \ln(e^{-\lambda}) + \ln(\lambda^{x_i}) - \ln(x_i!)$$

$$= \sum_{i=1}^{n} -\lambda + x_i \ln(\lambda) - \ln(x_i!)$$

$$\ell(\lambda) = -n\lambda + \ln(\lambda)\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \ln(x_i!)$$

$$\ell'(\lambda) = -n + \frac{1}{\lambda}\sum_{i=1}^{n} x_i = 0$$

$$\frac{1}{\lambda}\sum_{i=1}^{n} x_i = n$$

$$\sum_{i=1}^{n} x_i = \lambda n$$

$$\frac{\sum_{i=1}^{n} x_i}{n} = \lambda$$

$$\hat{\lambda} = \bar{x}$$

Normal Distribution

$\sigma^2 = \tau$

Let $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$. Show that the MLE's of μ and σ^2 are $\bar{x}$ and $\frac{n-1}{n}s^2$, respectively.

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$L(\mu,\sigma^2) = \prod_{i=1}^{n} f(x_i;\mu,\sigma^2)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$\ell(\mu,\sigma^2) = \ln\left[\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}}\right]$$

$$= \sum_{i=1}^{n} \ln \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \right]$$

$$= \sum_{i=1}^{n} \ln \left[\frac{1}{\sqrt{2\pi\sigma^2}} \right] + \ln \left(e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \right)$$

$$= \sum_{i=1}^{n} -\ln \left((2\pi\sigma^2)^{1/2} \right) - \frac{(x_i-\mu)^2}{2\sigma^2}$$

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

$$\underbrace{\frac{1}{2\sigma^2}(x_1-\mu)^2} + \cdots$$

$$l(\mu,\sigma^2) = \frac{-n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} \underbrace{(x_i-\mu)^2}$$

$$\frac{d\, l(.)}{d\mu} = 0 - \frac{1}{2\sigma^2} \sum 2(x_i-\mu)(-1)$$

$$= \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) = 0$$

$$\sum_{i=1}^{n} (x_i - \mu) = 0$$

$$\sum_{i=1}^{n} x_i - n\mu = 0$$

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

$$\frac{\sum x_i}{n} = \mu$$

$$\hat{\mu} = \bar{x}$$

$$\frac{1}{x} = x^{-1}$$

$$-1x^{-2}$$

$$l(\mu, \sigma^2) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \left(\sum_{i=1}^{n} (x_i - \mu)^2 \right)$$

$$\frac{d\, l(\mu, \sigma^2)}{d\sigma^2} = -\frac{n}{2} \frac{1 \cancel{2\pi}}{\cancel{2\pi}\sigma^2} + \frac{1}{2(\sigma^2)^2} \left(\sum_{i=1}^{n} (x_i - \mu)^2 \right)$$

$$= \frac{-n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^{n} (x_i - \mu)^2 = 0$$

$$\frac{n}{2\sigma^2} = \frac{1}{2(\sigma^2)^2} \sum (x_i - \mu)^2$$

$$n\sigma^2 = \sum (x_i - \mu)^2$$

$$\hat{\sigma}^2 = \frac{\sum (x_i - \hat{\mu})^2}{n}$$

$$\frac{n-1}{n} S^2$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Exponential Distribution

Let $X_1, \ldots, X_n \stackrel{iid}{\sim} Exp(\lambda)$. Find the MLE of λ