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.

Population

$F(\theta)$

$\mathbf{X} = \{X_1, \ldots, X_n\}$

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

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

Data

$F(\cdot)$ ~ known distribution

Let $X_1, \ldots, X_n \overset{iid}{\sim} F(\boldsymbol{\theta})$ 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.

$\underline{X}$

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

What is $g(\cdot)$?

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

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

If the data is iid, then

$\sim P(X = x)$

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

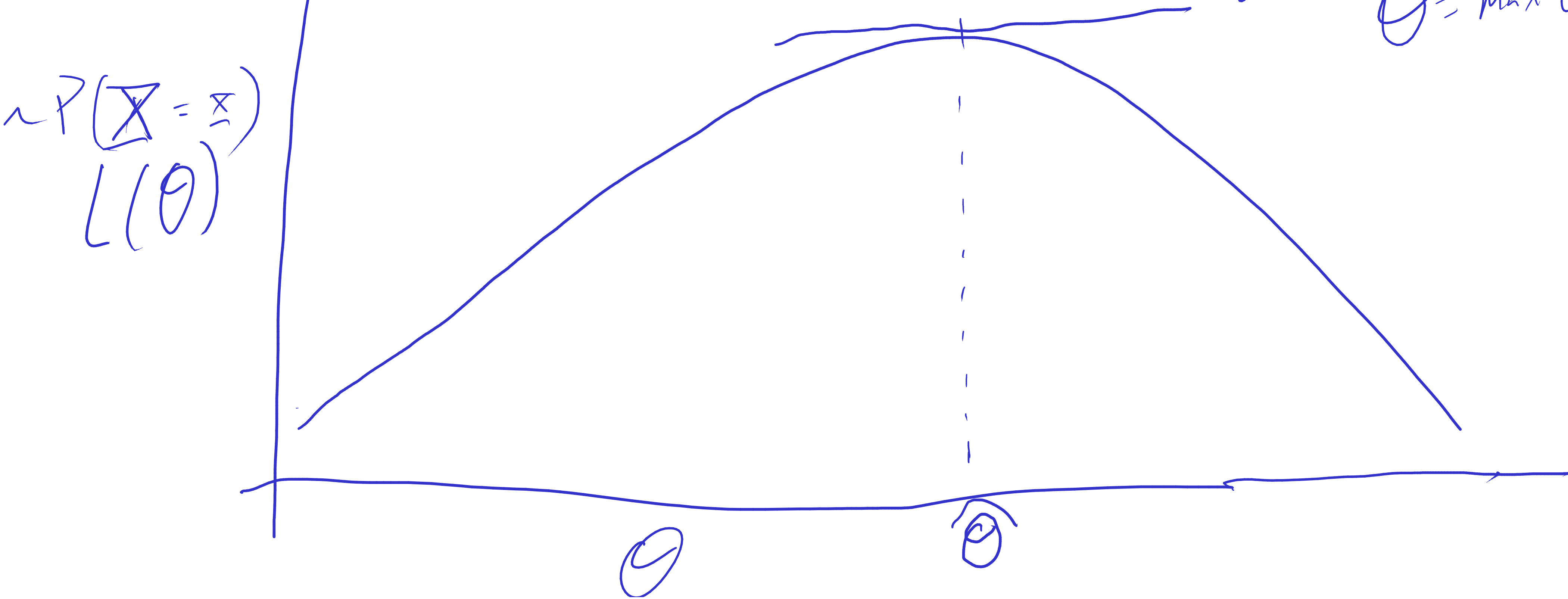
$f(x|\theta) = f(x_i|\theta)$ known (a number)

$-\infty \le \theta \le \infty$

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

$L'(\theta) = 0$

$\hat{\theta} = \max L(\theta)$



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.

$$\hat{\theta} = \arg\max_{\theta} L(\theta)$$

$$\hat{\theta} = \arg\max_{\theta} \ln(L(\theta))$$

Maximum log-Likelihood Estimator

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

Example

Poisson Distribution

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

$$P(X=x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$\underline{X} = \{x_1, \ldots, x_n\}$$

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

$$\ell(\lambda \mid \underline{X} = \underline{x}) = \ln\left(\prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!}\right)$$

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

$$= \sum_{i=1}^{n} \ln(e^{-\lambda}) + \ln(\lambda^{x_i}) - \ln(x_i!) \qquad ax^2 + bx + c$$

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

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

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

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

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

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

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

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.

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

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

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

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

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

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

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

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

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

$$\sum(x_i-\mu) = 0$$

$$S^2 = \frac{1}{n-1}\sum(x_i-\bar{x})^2 \qquad \sum x_i - n\mu = 0$$

$$\sum X_i = n\mu$$

$$\frac{\sum X_i}{n} = \hat{\mu}$$

$$\frac{dl}{d\sigma^2} = \frac{-n}{2}\frac{1 \cancel{2\pi}}{\cancel{2\pi}\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum(x_i-\mu)^2$$

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

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

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

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

Exponential Distribution

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