Maximum Likelihood Estimators

Learning Outcomes

- Maximum Likelihood Estimators
- Properties

Background Information

Estimators

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

$\overline{\underline{X}} = (x_1, \dots, x_n)$

$F(\theta)$

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

Data

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.

MLE Properties

Unbiased Estimators

Let $X_1, \ldots, X_n$ be a random sample from a distribution with parameter θ. Let $\hat{\theta}$ be an estimator for a parameter θ. Then $\hat{\theta}$ is an unbiased estimator if $E(\hat{\theta}) = \theta$. Otherwise, $\hat{\theta}$ is considered biased.

$E(\hat{\theta}) = \theta$ unbiased

$E(\hat{\theta}) \neq \theta$ biased

Consistent Estimators

Let $X_1, \ldots, X_n$ be a random sample from a distribution with parameter θ. The estimator $\hat{\theta}$ is a consistent estimator of the θ if

1. $E\{(\hat{\theta} - \theta)^2\} \to 0$ as $n \to \infty$
2. $P(|\hat{\theta} - \theta| \geq \epsilon) \to 0$ as $n \to \infty$ for every $\epsilon > 0$

Invariance Property

If $\hat{\theta}$ is an ML estimator of θ, then for any one-to-one function g, the ML estimator for $g(\theta)$ is $g(\hat{\theta})$.

Large Sample Theory

Let $X_1, \ldots, X_n$ be a random sample from a distribution with parameter θ. Let $\hat{\theta}$ be the MLE estimator for a parameter θ. As $n \rightarrow \infty$, then $\hat{\theta}$ has a normal distribution with mean θ and variance $1/nI(\theta)$, where

iid

$$I(\theta) = E\left[-\frac{\partial^2}{\partial\theta^2}\log\{f(X;\theta)\}\right]$$

$$\hat{\theta} \sim N\left(\theta, \frac{1}{nI(\theta)}\right)$$

Example

Exponential Distribution

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

$f(x) = \frac{1}{\lambda} e^{-x/\lambda}$

1. Likelihood
2. log-likelihood
3. Derivative
4. Set to $\emptyset$
5. Solve for the parameter

$\hat{\lambda} \sim N\left(\lambda, \frac{1}{n \frac{1}{\lambda^2}}\right) = \hat{\lambda} \sim N\left(\lambda, \frac{\lambda^2}{n}\right)$

$I(\lambda) = E\left(-\frac{d^2}{d\lambda^2} \log(f(x))\right)$

$\ln(f(x)) = -\ln(\lambda) - \frac{x}{\lambda}$

$= -\frac{1}{\lambda} + \frac{x}{\lambda^2}$

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

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

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

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

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

$$\ell(\lambda) = -n\ln(\lambda) - \frac{1}{\lambda} n\bar{X}$$

$$E\left(-\frac{1}{\lambda^2} + \frac{2X}{\lambda^3}\right)$$

$$-\frac{1}{\lambda^2} + \frac{2}{\lambda^3}E(X)$$

$$-\frac{1}{\lambda^2} + \frac{2\cancel{\lambda}}{\lambda^{\cancel{3}\,2}}$$

$$\frac{1}{\lambda^2} = I(\lambda)$$

$$n\bar{X} = \sum_{i=1}^{n} x_i$$

$$\frac{d\,\ell(\lambda)}{d\lambda} = -\frac{n}{\lambda} + \frac{1}{\lambda^2}\, n\bar{x}$$

$$\cancel{\lambda^2}\,\frac{n}{\cancel{\lambda}} = \frac{n\bar{x}}{\cancel{\lambda^2}}\,\cancel{\lambda^2}$$

$$\frac{\lambda n}{n} = \frac{n\bar{x}}{n}$$

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

Poisson Distribution

Let $X_1, \ldots, X_n \overset{iid}{\sim} \text{Pois}(\lambda)$, Find the sampling distribution of the MLE.

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

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

$$\frac{1}{n \frac{1}{\lambda}} = \frac{\lambda}{n}$$

$$\hat{\lambda} \sim N\left(\lambda, \frac{1}{n I(\lambda)}\right)$$

$$\ln\left(\frac{\lambda^x e^{-\lambda}}{x!}\right) = x \ln(\lambda) - \cancel{\lambda} - \cancel{\ln(x!)}$$

$$\ell' = \frac{x}{\lambda}$$

$$\ell'' = -\frac{x}{\lambda^2}$$

$$E\left(\frac{X}{\lambda^2}\right) = \frac{1}{\lambda^2} E(X) = \frac{1}{\lambda^2}\lambda = \frac{1}{\lambda} = I(\lambda)$$

Normal Distribution

$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

Let $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$. Are the MLE's of μ and σ^2 unbiased?

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

$$E(\hat{\sigma}^2) = E\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)$$

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

$$= \frac{1}{n} E\left(\sum_{i=1}^{n}(x_i^2 - 2x_i\bar{x} + \bar{x}^2)\right)$$

$$= \frac{1}{n} E\left(\sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum x_i + n\bar{x}^2\right)$$

$\bar{X} = \frac{\sum X_i}{n}$

$n\bar{X} = \sum_{i=1}^{n} X_i$

$$= \frac{1}{n} E\left(\sum X_i^2 - 2\bar{X} n \bar{X} + n\bar{X}^2\right)$$

$$= \frac{1}{n} E\left(\sum X_i^2 - n\bar{X}^2\right)$$

$$= \frac{1}{n}\left[\sum_{i=1}^{n} E(X_i^2) - n E(\bar{X}^2)\right]$$

$$= \frac{1}{n}\left[\sum (\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)\right]$$

$$\frac{1}{n}\left[n\sigma^2 \cancel{+ n\mu^2} - \sigma^2 \cancel{- n\mu^2}\right]$$

$$\frac{1}{n}\left(n\sigma^2 - \sigma^2\right)$$

$$X_i \sim N(\mu, \sigma^2)$$

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$E(X_i^2) = \sigma^2 + \mu^2$$

$$Var(X) = E(X^2) - E(X)^2$$

σ^2 μ^2

$$E(\bar{X}^2) = \frac{\sigma^2}{n} + \mu^2$$

$$\frac{\sigma^2 (n-1)}{n} = E(\hat{\sigma}^2)$$

$$E\left(\frac{n}{n-1}\hat{\sigma}^2\right) = \frac{n}{n-1} E(\hat{\sigma}^2)$$

$$\cancel{\frac{n}{n-1}\,\frac{(n-1)}{n}}\,\sigma^2$$

$$\sigma^2$$

$$\frac{n\,\hat{\sigma}^2}{n-1} = \frac{\sum (x_i-\bar{x})^2}{\cancel{n}}\,\frac{\cancel{n}}{n-1}$$

$$= \frac{\sum (x_i-\bar{x})^2}{n-1} = s^2$$

$$E(\hat{\mu}) = E(\overline{X}) = E\left(\frac{1}{n}\sum X_i\right) \qquad X_i \sim N(\mu, \sigma^2)$$

$$\mu$$

$$\frac{1}{n} E\left(\sum X_i\right) = \frac{1}{n}\sum E(X_i)$$

$$= \frac{1}{n}\sum \mu = \frac{n\mu}{n} = \mu$$