Estimation

Point estimation:
A sample-based statistics is used to approximate or estimate a population parameter θ, is called a point estimator for θ and is denoted by $\hat{\theta}$.

Desirable properties of a point estimate:
- $\hat{\theta}$ to be unbiased for θ;
- $\hat{\theta}$ to have a small variance for large sample size;

Unbiased estimate:
An estimator $\hat{\theta}$ is an unbiased estimator for parameter θ if and only if $E\left[\hat{\theta}\right] = \theta$

Estimation

Example: Let $X_1, X_2, X_3, \dots, X_n$ be a random sample of size n from a distribution X with mean μ. The sample mean, $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is an unbiased estimator for μ.

We have, $E[X_i] = E[X] = \mu, i = 1,2, \dots, n$. Therefore,

$$E[\bar{X}] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n} n\mu = \mu$$

Example: Let $X_1, X_2, X_3, \dots, X_n$ be a random sample of size n from a distribution X with mean μ and variance σ^2. The sample statistic, $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is a biased estimator for σ^2, where , $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is the sample mean of X .

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

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

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

Example: Let $X_1, X_2, X_3, \dots, X_n$ be a random sample of size n from a distribution X with mean μ and variance σ^2. The sample variance, $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is an unbiased estimator for σ^2, where , $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is the sample mean of X .

Solution:

It is given that $E[X_i] = \mu; VarX_i = \sigma^2; E\left[X_i^2\right] = E[X^2] = \sigma^2 + \mu^2$

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

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

$$\Rightarrow \frac{n-1}{n}E[S^2] = E[X^2] - E[\bar{X}^2] = \sigma^2 + \mu^2 - E[\bar{X}^2] \qquad (1)$$

To find $E[\bar{X}^2]$, recall

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

$$= \frac{1}{n^2}\sum_{i=1}^{n} VarX_i + \mu^2 = \frac{1}{n^2}(n\sigma^2) + \mu^2 = \frac{\sigma^2}{n} + \mu^2 \qquad (2)$$

Substitute (2) into (1), we have,

$$\frac{n-1}{n}E[S^2] = \sigma^2 + \mu^2 - \frac{\sigma^2}{n} - \mu^2 = \frac{n-1}{n}\sigma^2 \Rightarrow E[S^2] = \sigma^2$$

Estimation

Point estimation:
A sample-based statistics is used to approximate or estimate a population parameter θ, is called a point estimator for θ and is denoted by $\hat{\theta}$.

Desirable properties of a point estimate:
- $\hat{\theta}$ to be unbiased for θ;
- $\hat{\theta}$ to have a small variance for large sample size;

Example: Let $\bar{X}$ be the sample mean based on a random sample of size n from a distribution with mean μ and variance σ^2. If we use $\bar{X}$ as the estimator of μ, let's evaluate the variance of this estimator.

$$Var\bar{X} = Var\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} VarX_i = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

Estimation

Point estimation:

A sample-based statistics is used to approximate or estimate a population parameter θ, is called a point estimator for θ and is denoted by $\hat{\theta}$.

Desirable properties of a point estimate:

- $\hat{\theta}$ to be unbiased for θ;
- $\hat{\theta}$ to have a small variance for large sample size;

Example: Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ be the sample mean based on a random sample, $X_1, X_2, \ldots, X_n$ of size n from a distribution with mean μ and variance σ^2. we use $\bar{X}$ as the estimator of μ, then $E[\bar{X}] = \mu, Var\bar{X} = \frac{\sigma^2}{n}$

Point estimation:
A sample-based statistics is used to approximate or estimate a population parameter θ, is called a point estimator for θ and is denoted by $\hat{\theta}$.

1. The moments-based estimator

Consider a random variable X with n samples $x_i, i = 1,2,3, \dots, n$. Recall that $E[X^k](k = 1,2,3, \dots)$ are the k^{th} moments for X. The estimator M_k for $E[X^k]$ based on sample values is given as:

$M_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k$, e.g., $M_1 = \frac{1}{n}\sum_{i=1}^{n} x_i$, $M_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$, $M_3 = \frac{1}{n}\sum_{i=1}^{n} x_i^3$, ...

Example: Let $x_1, x_2, \dots, x_n$ be a random sample from a *Gamma* distribution with parameters α *and* β. Find the moments-based estimator of α *and* β.

Recall the following properties of a Gamma random variable X: (relating the moments and the parameters)

$$E[X] = \alpha\beta$$
$$VarX = \alpha\beta^2 \Rightarrow E[X^2] - (E[X])^2 = \alpha\beta^2$$

Replace the quantities in the above equations by their estimators, we have,

$$M_1 = \hat{\alpha}\hat{\beta}$$
$$M_2 - M_1^2 = \hat{\alpha}\hat{\beta}^2$$

Solving these two equations gives:

$$\hat{\alpha} = \frac{{M_1}^2}{M_2 - {M_1}^2}; \qquad \hat{\beta} = \frac{M_2 - {M_1}^2}{M_1};$$

Estimation

Example: The following is a random sample from a Gamma distribution with parameters $\alpha\ and\ \beta$. Estimate $\alpha\ and\ \beta$ from these sample values using the moments-based estimator.

x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
45.3	42.4	86.8	84.5	25.1	91.5	72.8	40.4	31.6	26.0

$$M_1 = \frac{1}{10}\sum_{i=1}^{10} x_i = 54.64, \quad M_2 = \frac{1}{10}\sum_{i=1}^{10} {x_i}^2 = 3613.3$$

$$\hat{\alpha} = \frac{{M_1}^2}{M_2 - {M_1}^2} = \frac{54.64^2}{3613.3 - 54.64^2} = 4.76$$

$$\hat{\beta} = \frac{M_2 - {M_1}^2}{M_1} = \frac{3613.3 - 54.64^2}{54.64} = 11.49$$

This random sample is generated using Matlab
x = gamrnd(5,10,[1 10])

The true values of $\alpha\ \ and\ \beta$ are 5 and 10

Estimation

1. The moments-based estimator

Consider a random variable X with n samples $x_i, i = 1,2,3, \dots, n$. Recall that $E[X^k](k = 1,2,3, \dots)$ are the k^{th} moments for X. The estimator M_k for $E[X^k]$ based on sample values is given as:

$M_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k$ e.g., $M_1 = \frac{1}{n}\sum_{i=1}^{n} x_i$, $M_2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$, $M_3 = \frac{1}{n}\sum_{i=1}^{n} x_i^3$, …

Practice Problem: Let $x_1, x_2, \dots, x_n$ be a random sample from a *normal* distribution with parameters $\mu\ and\ \sigma^2$. Find the moments-based estimator of $\mu\ and\ \sigma^2$.

We already know that for a normal random variable X, we have,

$$E[X] = \mu; \quad VarX = E[X^2] - (E[X])^2 = \sigma^2$$

Replacing both sides of these equations by its estimators gives

$$M_1 = \hat{\mu}; \ \ M_2 - M_1^2 = \widehat{\sigma^2}$$

Therefore, we have,

$$\hat{\mu} = M_1 = \frac{1}{n}\sum_{i=1}^{n} x_i \, ; \qquad \widehat{\sigma^2} = M_2 - {M_1}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i)^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})^2$$

Estimation

2. Maximum likelihood estimate:

Example 1 (discrete case): Let $x_1, x_2, \dots, x_n$ be a random sample from a Poisson distribution with parameter k. We want to find the value of k that given the maximum probability of observing this sample.

Solution:

Recall the Poisson density function: $P[X = x] = f(x) = \frac{e^{-k}k^x}{x!},\ x = 0,1,2,\dots$

The probability of obtaining the given sample is:

$$P[X_1 = x_1, X_2 = x_2, \dots, X_n = x_n] = P[X_1 = x_1]P[X_2 = x_2]\cdots P[X_n = x_n]$$

$$= \prod_{i=1}^{n} P[X_i = x_i] = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \frac{e^{-k}k^{x_i}}{x_i!} = L(k)$$

This probability is a function of k and is called the ***likelihood function***.

$$L(k) = \prod_{i=1}^{n} \frac{e^{-k}k^{x_i}}{x_i!} = \frac{e^{-nk}k^{\sum x_i}}{\prod x_i!}$$

We want to find the value of k that maximize this likelihood function!!!

Estimation

2. **Maximum likelihood estimate:**

Solution: (continue)

$$L(k) = \prod_{i=1}^{n} \frac{e^{-k} k^{x_i}}{x_i!} = \frac{e^{-nk} k^{\sum x_i}}{\prod x_i!}$$

We first take the natural logarithm of $L(k)$ to obtain the ***log likelihood function***:

$$\ln L(k) = -nk + \ln k \sum_{i=1}^{n} x_i - \ln\left(\prod_{i=1}^{n} x_i!\right)$$

Notice that the value of k that maximizes $\ln L(k)$ also maximize $L(k)$ (***why?***)

$$\frac{d}{dk} \ln L(k) = -n + \frac{1}{k} \sum_{i=1}^{n} x_i = 0 \Rightarrow k = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x} = \hat{k}$$