Statistical Inference

Properties of Estimation:

Suppose we have a random sample $x_1, x_2, \cdots, x_n$ taken from the population $f(\theta)$, θ = unknown population parameter.

Obj: Find a "reasonable" estimator of θ, using the random sample $x_1, x_2, \cdots, x_n$ at hand.

We want construct a reasonable estimator of θ, say $T = T(x_1, x_2, \cdots, x_n)$

Criteria for a Good Estimator:

(i) Unbiaseness:

The estimator T is said to be an unbiased estimator of θ, if $\boxed{E(T) = \theta}$

Eg: r.s $\overset{\mu}{x_1}, \overset{\mu}{x_2}, \cdots, \overset{\mu}{x_n}$ from $N(\mu, 1)$ population.

Then sample mean $\bar{x}$ is an unbiased estimator of μ.

(T) $\leftarrow \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

To show $\bar{x}$ is unbiaseness: To prove: $E(\bar{x}) = \mu$

$$E(\bar{x}) = E\left[\frac{1}{n}\sum x_i\right] = \frac{1}{n}\left[\sum_{i=1}^{n} \underbrace{E(x_i)}\right] = \frac{1}{n}\sum \mu = \frac{1}{n}\cdot n\mu = \mu$$

$x_1, x_2, \cdots, x_n \overset{iid}{\sim} f(\mu, \sigma^2)$

Q. Let $\overset{\mu}{x_1}, \overset{\mu}{x_2}, \cdots, \overset{\mu}{x_n}$ be a random sample of size 'n' from popln with mean 'μ' & variance 'σ^2' [suppose both are unknown] popln with mean 'μ' & variance 'σ^2' [suppose both are unknown].

Then $E(\bar{x}) = \mu$, where $\bar{x} = \frac{1}{n}\sum x_i$ [Sample Mean]

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

[sample var]

$E(\bar{x}) = \mu$ [Proof is same as above].

To prove: $E(s'^2) = \sigma^2$

$$E(s'^2) = E\left[\frac{1}{n-1}\sum(x_i - \bar{x})^2\right]$$

$$= \frac{1}{n-1} E\left[\sum(x_i - \bar{x})^2\right] \rightarrow \text{expand.}$$

$$= \frac{1}{n-1} E\left[\sum x_i^2 - n\bar{x}^2\right]$$

$$= \frac{1}{n-1}\left[\sum \underbrace{E(x_i^2)}_{(\sigma^2+\mu^2)} - n\,\underbrace{E(\bar{x}^2)}_{\left(\mu^2 + \frac{\sigma^2}{n}\right)}\right].$$

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

$$= \frac{1}{n-1}\left[n\sigma^2 - \sigma^2\right]$$

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

$Var(x_i) = \sigma^2$.

$E(x_i^2) - [E(x_i)]^2 = \sigma^2$

$E(x_i^2) = \sigma^2 + \mu^2$.

$Var(\bar{x}) = \sigma^2/n$

$E(\bar{x}^2) - [\underbrace{E(\bar{x})}_{\mu}]^2 = \frac{\sigma^2}{n}$

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

$Var(\bar{x}) = Var\left(\frac{1}{n}\sum x_i\right)^2$

$= \frac{1}{n^2} Var\left(\sum x_i\right)^2$

$= \frac{1}{n^2}\sum Var(x_i)$

$= \frac{1}{n^2}\sum\sigma^2 = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$

Note: For estimating μ, $T = \bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$

Eg: $T' = \frac{2x_1 + x_2 + \cdots + x_{n-1} - x_n}{n}$. Is $E(T) = \mu$?

$$E(T') = \frac{1}{n}\left[2E(x_1) + E(x_2) + \cdots + E(x_{n-1}) - E(x_n)\right]$$

$$= \frac{1}{n}\left[2\mu + \mu + \cdots + \mu - \mu\right] = \frac{1}{n}\cdot\cancel{n}\mu = \mu.$$

$$= \frac{1}{n}\Big[\underbrace{\widetilde{2\mu}}_{1} + \underbrace{\mu + \cdots\cdots + \mu}_{(n-2)} - \mu\Big] = \frac{1}{\cancel{n}}\cdot \cancel{n}\mu = \mu.$$

If we have 2 unbiased estimators T & T', for the unknown popln parameter θ, then, choose the one with lesser variance.

(i) If $Var(T) < Var(T') \Rightarrow$ Choose T

(ii) If $Var(T') < Var(T) \Rightarrow$ Choose T'

$\therefore$ We should always the Minimum Variance Unbiased Estimator (MVUE).

Eg: $T' = \dfrac{2x_1 + x_2 + \cdots\cdots - x_n}{n}$

$$Var(T') = \frac{1}{n^2} Var\left(2x_1 + x_2 + \cdots - x_n\right)$$

$$= \frac{1}{n^2}\left[4\,Var(x_1) + Var(x_2) + \cdots + Var(x_n)\right]$$

$$= \frac{1}{n^2}\cdot\left[4\sigma^2 + (n-1)\sigma^2\right] = \frac{(n+3)\sigma^2}{n^2}.$$

$$= \underbrace{\frac{\sigma^2}{n}}_{Var(T)} + \frac{3\sigma^2}{n^2}$$

$\therefore$ $Var(T') > Var(T) \Rightarrow$ Out of $T, T' \Rightarrow$ Choose $T = \bar{x}$

Out of all possible unbiased estimators, for the unknown population parameter θ, we want to choosen the one with the minimum possible variance. This estimator is known as Uniformly Min Var. Unbi

This estimator is known as Uniformly Min Var Unbiased Estimator (UMVUE)

Cramer Rao Lower Bound (CRLB):

CRLB gives the lowest limit on the variance of any unbiased estimator of the unknown popln parameter.

If $E(T) = \theta$, Then $Var(T) \geq c$ where $c = CRLB$ and T is any unbiased estimator.

Suppose $x_1, x_2, \ldots, x_n \overset{iid}{\sim} f(\theta)$, θ = unknown popln parameter. Find CRLB for an unbiased estimator of θ.

$x_1, x_2, \ldots, x_n \overset{iid}{\sim} f(\theta)$

(i) Likelihood fn: $L(\theta) = \prod_{i=1}^{n} f_\theta(x_i)$

(ii) Log-likelihood fn: $\ell(\theta) = \log L(\theta)$

(iii) Find: $g(\theta) = \frac{\partial}{\partial\theta} \ell(\theta)$.

(iv) Find $I(\theta) = E[g(\theta)]^2$.

(v) $CRLB = \frac{1}{I(\theta)}$.