Chap: 7

$\boxed{7.1}$

Statistical Estimation and Sampling Distribution

Parameter: In statistical, the term parameter is used to denote a quantity θ, that is a property of an unknown probability distribution. For example it may be the mean, variance or a particular quantile of the probability distribution.

Statistics: In statistical, the term statistic is use to denote a quantity that is a property of a sample. for example, it may be sample varience or a particular sample quantile.

Point of Estimate of Parameter

$\boxed{\text{unknown parameter } \theta}$

$\downarrow$

$\boxed{\text{probability distribution, } f(x, \theta)}$

$\downarrow$

$\boxed{\text{Data observation}}$

$\downarrow$

$\boxed{\text{Point estimation.}}$

$\hat{\theta} = \dfrac{\sum x_i}{n}$ (Sample mean)

$\theta = \mu$ (mean)

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

Population varience $= \sigma^2$

Sample varience $= s^2$

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

Prove $E(s^2) = \sigma^2$

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

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

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

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

Now notice that

$$E(x_i) = \mu$$

So that

$$E((x_i - \mu)^2) = \mathrm{var}(x_i) = \sigma^2$$

Further more

$$E(\bar{x}) = \mu$$

So that

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

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

Example 1 page : 303

$x_0 \sim B(n, p_0)$

$\hat{p}_0 = \frac{x_0}{n}$

$\bar{\mu} = x$

$E(x_i) = \mu \quad 1 \leq i \leq n$

so that

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

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

7.2 Properties of Point Estimate.

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

$P = \frac{x}{n}$

$x \sim B(n, p)$

$E(x) = np$

$E(\hat{p}) = E\left(\frac{x}{n}\right) = \frac{1}{n}E(x) = \frac{1}{n}np = p$

Hence $\hat{p}$ is the unbiased point estimate of the success probability p

Book problem

$\boxed{7.2.1}$

Information

$E(x_1) = \mu$

$Var(x_1) = 10$

$E(x_2) = \mu$ and $Var(x_2) = 15$

$$\hat{\mu}_1 = \frac{x_1}{2} + \frac{x_2}{2}$$

$$\hat{\mu}_2 = \frac{x_1}{4} + \frac{3x_2}{4}$$

$$\hat{\mu}_3 = \frac{x_1}{6} + \frac{x_2}{3} + 9$$

Solution

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

So Here

bias $(\hat{\mu}_1) =$ E

Solution

$E(\hat{\mu_1}) = \mu_1$

(a) $\mu_1 = \frac{x_1}{2} + \frac{x_2}{2}$

$E(X_1) = \mu$ and E(

$E(\hat{\mu_1}) = E\left(\frac{x_1}{2} + \frac{x_2}{2}\right)$

$E(\hat{\mu_1}) = \frac{Ex_1}{2} + \frac{Ex_2}{2}$

$= \frac{\mu}{2} + \frac{\mu}{2}$

$= \mu$ (Ans)

$\mu_2 = \frac{x_1}{4} + \frac{3x_2}{4}$

$E(\hat{\mu_2}) = E\left(\frac{x_1}{4} + \frac{3x_2}{4}\right)$

$= \frac{Ex_1}{4} + \frac{E\,3x^2}{4}$

$= \frac{\mu}{4} + \frac{3\mu}{4}$

$= \mu$

$$\hat{\mu_3} = \frac{x_1}{6} + \frac{x_2}{3} + 9$$

$$E(\hat{\mu_3}) = E\left(\frac{x_1}{6} + \frac{x_2}{3} + 9\right)$$

$$E(\hat{\mu_3}) = \frac{Ex_1}{6} + \frac{Ex_2}{3} + 9$$

$$= \frac{\mu}{6} + \frac{\mu}{3} + 9$$

$$= \frac{\mu}{2} + 9$$

$$\text{bias} = E(\hat{\mu_3}) - \mu$$

$$= \frac{\mu}{2} + 9 - \mu$$

$$= 9 - \frac{\mu}{2} \text{ (Ans)}$$

(b) $\hat{\mu_1} = \frac{x_1}{2} + \frac{x_2}{2}$

$Var(x_1) = 10$ and

$$Var(\hat{\mu_1}) = Var\left(\frac{x_1}{2} + \frac{x_2}{2}\right)$$

$$= \frac{Var\, x_1}{2^2} + \frac{Var\, x_2}{2^2}$$

$$= \frac{10}{2^2} + \frac{15}{2^2}$$

$$= \frac{10}{4} + \frac{15}{4}$$

$$= \frac{25}{4} \text{ (Ans)}$$

$$\hat{\mu_2} = \frac{x_1}{4} + \frac{3x_2}{4}$$

$$Var(\hat{\mu_2}) = Var\left(\frac{x_1}{4} + \frac{3x^2}{4}\right)$$

$$= \frac{Var(x_1)}{4^2} + \frac{3\, Var(x^2)}{4^2}$$

$$= \frac{10}{4^2} + \frac{3^2 \times 15}{4^2}$$

$$= \frac{145}{16} \text{ (Ans)}$$

$$\hat{\mu}_3 = \frac{x_1}{6} + \frac{x_2}{3} + 9$$

$$Var(\hat{M}_3) = Var\left(\frac{x_1}{6} + \frac{x_2}{3}\right)$$

$$= \frac{Var(x_1)}{6^2} + \frac{Var(x_2)}{3^2}$$

$$= \frac{10}{6^2} + \frac{15}{3^2}$$

$$= \frac{70}{36}$$

$\hat{M}_3$ is the smallest (Ans)

(e) $\therefore$ $MSE(\hat{M_1}) = var(\hat{M_1}) + bias^2$ $\boxed{M=8}$

$$= \frac{25}{4} + 0$$

$$= \frac{25}{4}$$

$$MSE(\hat{M_2}) = var(\hat{M_2}) + bias^2$$

$$= \frac{145}{4} + 0$$

$$= \frac{145}{4}$$

$$MSE(\hat{M_3}) = var(\hat{M_3}) + bias^2$$

$$= \frac{70}{36} + \left(9 - \frac{M}{2}\right)^2$$

$$= \frac{70}{36} + \left(9 - \frac{8}{2}\right)^2$$

$$= 26.944 \text{ (Ans)}$$

7.3 Sampling Distribution

Sample proportion

$x \sim B(n, p)$

$$\hat{p} = \frac{x}{n}$$

This estimate can be referred to as a sample proportion since it represent the proportion of sucesses observed in a sample of n trials.

$$Var(\hat{p}) = \frac{p(1-p)}{n}$$

if $x \sim B(n,p)$, then the sample proportion $\hat{p} = \frac{x}{n}$

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$$

Sample Mean

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

Sample mean is the subset of given population.

The sample mean is defined as the avg of n observation from the sample.

Sample Variance

For a sample $x \dots x_n$ obtained from a population with a mean μ and a variance σ^2 consider the variance estimate

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

7.4 constructing Parameter Estimate

Method of Moments point Estimate for one Parameter

If a data set consists of observation $x_1 \dots x_n$ from a probability distribution that depends upon one unknown parameter θ, the method of moments point estimate θ of the Parameter is found by Solving the equation.

$$\bar{X} = E(x)$$

Method of Moments Point Estimate for two parameter

$$\bar{x} = E(x) \text{ and } s^2 = \text{Var}(x)$$

Maximum Likelihood Estimate

Maximum Likelihood Estimate for one Parameter

$$L(x_1 \dots x_n, \theta) = f(x_1 \theta) \times \dots \times f(x_n, \theta)$$

therefore

$$L(\theta) = \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x} = p^x (1-p)^{n-x}$$

$$\boxed{\# \quad L(\theta) = \prod_{x=1}^{n} f(x, \theta)}$$

Where $x = x_1 + \dots + x_n$ and the maximum likelihood estimate $\hat{p}$ is the value that maximizes this.

The log-likehood is

$$\ln(L) = x \ln(p) + (n-x) \ln(1-p)$$

and

$$\frac{d \ln(L)}{dp} = \frac{x}{p} - \frac{n-x}{1-p}$$

$$\hat{p} = \frac{x}{n}$$

Maximum Likehood Estimate for Parameters.

$$L(x_1 \dots x_n\ \theta_1, \theta_2) = f(x_1\ \theta_1, \theta_2) \times \dots \times f(x_n, \theta_1, \theta_2)$$

Example 1: class, work

Mid Mark : 14, 13, 12.5, 13.5, 13.5 $\sim N(12, 4)$

$$f(14, \theta) \times f(13, \theta) \times f(12.5, \theta) \times f(13.5, \theta) \times f(13.5, \theta)$$

$$= \prod_{i=1}^{5} f(x_i, \theta)$$

$$L(\theta) = \sum_{i=1}^{n} (x_i, \theta)$$

$$\Rightarrow \ln L(\theta) = \sum_{i=1}^{n} \ln f(x, \theta)$$

$$L(\theta) = \mu$$

$$\frac{d \ln(\theta)}{d(\theta)} = 0$$

7.4.3

We want to prove that the same estimate is obtained by using method of moments as well as maximum likelihood estimation.

Method of moments consist of equalizing expectation with mean value:

$$EX = \frac{1}{\lambda} = \overline{x}$$

$$\Longrightarrow \hat{\lambda} = \frac{1}{\overline{x}}$$

Maximum likelihood estimation consists of finding $\hat{\lambda}$ that will maximize L:

$$L(\lambda; x_1, x_2, ..., x_n) = \lambda^n \cdot e^{-\lambda \sum_{i=1}^{n} x_i}$$

It's complicated to find maximum of function L, therefore we will make function lnL. Since ln is increasing function maximum remains the same as in function L.

$$lnL(\lambda; x_1, x_2, ..., x_n) = nln\lambda - \lambda \sum_{i=1}^{n} x_i$$

Now it's easy to find value $\hat{\lambda}$ by equalizing first derivate with zero.

$$\frac{\partial lnL(\lambda; x_1, x_2, ..., x_n)}{\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i$$

$$\Longrightarrow \hat{\lambda} = \frac{1}{\overline{x}}$$

$$E(S^2) = \sigma^2$$

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

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

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

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

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

$\times$ $\boxed{E(x_i) = \mu = \operatorname{var}(x) = \sigma^2}$

$E(x_i) = \mu$

$E(x_i - \mu)^2 = \operatorname{var}(x) = \sigma^2$

$E(\bar{x}) = \mu$

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

$$= \frac{1}{n-1}\left(\sum_{i=1}^{n} \sigma^2 - n \cdot \left(\frac{\sigma^2}{n}\right)\right)$$

$$= \sigma^2$$