MODULE - 5 :- 1

1) Prove that for a random sample of size n (i.e) $(x_1, x_2, \ldots x_n)$ taken from the population $s^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2$ is not unbaised estimation of the parameter σ^2. But $\frac{n}{n-1} s^2$ is unbiased estimation of σ^2

For a sample variance s^2

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

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

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

To make s^2 an unbaised estimator, we use

$$s^2_{\text{unbaised}} = \frac{n}{n-1} s^2$$

Hence s^2 is baised and $\frac{n}{n-1} s^2$ is an unbaised estimator of σ^2

2) Show that if T is an unbaised estimator of θ then T^2 and $T^{1/2}$ are baised estimators of θ^2 and $\theta^{1/2}$ respectively

Given $E(T) = \theta$

for T^2:

$$E(T^2) \neq \theta^2$$

This is because

$$E(T^2) = var(T) + (E(T))^2 = \sigma_T^2 + \theta^2$$

where σ_T^2 is the variance of T, Thus T^2 is biased for θ^2

for $T^{1/2}$

Using Jensen's inequality for the concave function $f(x) = \sqrt{x}$:

$$E(T^{1/2}) < \sqrt{E(T)} = \sqrt{\theta}$$

Therefore, $T^{1/2}$ is baised for $\theta^{1/2}$

3) Find the Maximum Likelihood Estimator for the parameter λ of poisson distribution

The likelihood function for a poisson distribution with parameter λ is given by:

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

Taking the natural logarithm

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

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

To find MLE, differentiate with respect to λ and set to zero:

$$\frac{d}{d\lambda} \ln L(\lambda) = \sum_{i=1}^{n} \frac{x_i}{\lambda} - n = 0$$

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

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

Hence, the MLE for λ is $\hat{\lambda} = \bar{x}$

TUTORIAL - II

1) Find the maximum likelihood estimators for random sampling from a normal population for

i) population mean the population variance is know ;

Given a normal distribution $X \sim N(\mu, \sigma^2)$:

The likelihood function is

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

Taking the natural logrithm

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

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

To find the MLE, we take the derivative of $\ln L(\mu)$ with respect to μ and set it to zero :

$$\frac{d}{d\mu} \ln L(\mu) = \frac{d}{d\mu}\left(-\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i-\mu)^2\right) = 0$$

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

$$\sum_{i=1}^{n} x_i - n\mu = 0 \qquad \mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

ii) Maximum likelihood Estimator for population variance

Let $x_1, x_2, x_3 \cdots x_n$ be random sample from a normal distribution with mean μ (known) and variance σ^2

The likelihood function:

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

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

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

Taking the derivative with respect to σ^2 and setting it to 0

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

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

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

so the MLE for σ^2 is

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

iii) Simultanous Estimation of Both Population Mean and variance

for the joint estimation of μ and σ^2, We maximize the joint likelihood function

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

The log-likelihood function is :

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

$$\frac{\partial}{\partial u} \ln L(\mu, \sigma^2) = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) = 0$$

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

$$\frac{\partial}{\partial u} \ln L(\mu, \sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_{i=1}^{n} (x_i - \mu)^2 = 0$$

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

so the MLES are:

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

These are the maximum likelihood estimators for the population mean and variance when both are unknown.

~~This completes the solution to the first question~~ ~~iii~~

2) Derive normal equations of straight line fit, $y = ?$ when $x = 1.5$

x	0	1	2	3	4
y	1.8	3.3	4.5	6.3	8.0

$$\Sigma y = na + b\,\Sigma x$$

$$\Sigma xy = a\,\Sigma x + b\,\Sigma x^2$$

using given data

$$\Sigma x = 0 + 1 + 2 + 3 + 4 = 10$$

$$\Sigma y = 1.8 + 3.3 + 4.5 + 6.3 + 8.0 = 23.9$$

$$\Sigma x^2 = 0^2 + 1^2 + 2^2 + 3^2 + 4^2 = 30$$

$$\Sigma xy = 0 \times 1.8 + 1 \cdot 3 \times 3 + 2 \times 4.5 + 3 \times 6.3$$
$$+ 4 \times 8.0 = 54.4$$

substitute these values in equation

$$23.9 = 5a + 10b \qquad 54.4 = 10a + 30b$$

$$5a + 10b = 23.9 \Rightarrow a + 2b = 4.78$$

$a = 4.78 - 2b$ in equation 2

$$54.4 = 10(4.78 - 2b) + 30b$$

$$54.4 = 47.8 + 10b$$

$$10b = 6.6 \qquad \boxed{b = 0.66}$$

$$a = 4.78 - 2b = 4.78 - 1.32 = 3.46$$

so the fitted line is :

$$y = 3.46 + 0.66x \qquad x = 1.5$$

$$y = 3.46 + 0.66(1.5) = 3.46 + 0.99 = 4.45$$

3) Fit a second degree parabola to the following data taking x as the independent variable

x	1	2	3	4	5	6	7	8	9
y	2	7	6	11	10	11	10	10	6

$$\Sigma y = na + b\Sigma x + c\Sigma x^2$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3$$

$$\Sigma x^2 y = a\Sigma x^2 + b\Sigma x^2 + c\Sigma x^4$$

$$\Sigma x = 1+2+3+4+5+6+7+8+9 = 45$$

$\Sigma y = 2+7+8+11+10+11+10+10+6 = 73$

$\Sigma x^2 = 1^2+2^2+3^2+4^2+5^2+6^2+7^2+8^2+9^2 = 285$

$\Sigma xy = 1\cdot2+2\cdot7+3\cdot6+4\cdot11+5\cdot10+6\cdot11+7\cdot10+$
$8\cdot10+9\cdot6 = 444$

$\Sigma x^3 = 1^3+2^3+3^3+4^3+5^3+6^3+7^3+8^3+9^3 = 2025$

$\Sigma x^4 = 1^4+2^4+3^4+4^4+5^4+6^4+7^4+8^4+9^4 = 15333$

$73 = 9a+45b+285c$

$444 = 45a+285b+2025c$

$3506 = 258a+2025b+1533c$

Solving these simultanous equations :

$$\begin{bmatrix} 9 & 45 & 285 \\ 45 & 285 & 2025 \\ 285 & 2025 & 15333 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 73 \\ 444 \\ 3506 \end{bmatrix}$$

$a \simeq 0.1 \quad b \simeq 0.5 \quad c \simeq 1$:

Quatratic equation :

$y \simeq 0.1x^2 + 0.5x + 1$