UNIT - 3

1) Let $x_1, x_2, x_3 \dots x_n$ be a random sample from Normal distribution $N(\mu, 1)$ population. Show that $t = \frac{1}{N}\sum_{i=1}^{n} x_i^2$ is an unbaised estimator of μ^2+1.

A) $E[\hat{\theta}] = E\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2 - 1\right]$

$\Rightarrow E[\hat{\theta}] = \frac{1}{n}E\left[\sum_{i=1}^{n} x_i^2\right] - 1$

$\Rightarrow E[\hat{\theta}] = \frac{1}{n}\sum_{i=1}^{n} E[x_i^2] - 1$

$E[x_i^2] = Var(x_i) + (E[x_i])^2 = 1+\mu^2$

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

$\Rightarrow \mu^2$

$\Rightarrow$ Hence $\frac{1}{n}\sum_{i=1}^{n} x_i^2 - 1$ is unbaised of μ^2.

2) Let t_1 & t_2 be two unbaised estimator of θ. show that estimator $t = at_1 + (1-a)t_2$ is an unbaised estimator of θ

A) $E(t_1) = \theta$ & $E(t_2) = \theta$

define a new estimator: $t = at_1 + (1-a)t_2$ (a is a constant)

$E(t) = E[at_1 + (1-a)t_2]$

$E(t) = aE[t_1] + (1-a)E[t_2]$

$\Rightarrow E(t) = a\theta + (1-a)\theta$

$\Rightarrow E(t) = \theta(a+1-a)$

$\Rightarrow E(t) = \theta$

$\Rightarrow$ Since $E[t] = \theta$, the estimator $t = at_1 + (1-a)t_2$ is also unbaised estimator of θ.

3) Let $T1$ & $T2$ be two consistent estimator of $\mu 1$ & $\mu 2$ respectively. Prove that $aT1 + bT2$ is a consistent estimator of $a\mu 1 + b\mu 2$, where a and b are constant and independent of population.

A) T_1 is a consistent estimator of μ_1:

$$T_1 \xrightarrow{P} \mu_1 \quad \text{as } n \to \infty$$

T_2 is a consistent estimator of μ_2:

$$T_2 \xrightarrow{P} \mu \quad \text{as } n \to \infty$$

a & b are constants

$$\lim_{x \to \infty} P(|T_n - \theta| > \epsilon) = 0$$

$$\Rightarrow \lim_{n \to \infty} P\left(\left|aT_1 + bT_2 - (a\mu_1 + b\mu_2)\right| > \epsilon\right) = 0$$

$$\Rightarrow \left|aT_1 + bT_2 - (a\mu_1 + b\mu_2)\right| = \left|a(T_1 - \mu_1) + b(T_2 - \mu_2)\right|$$

By triangle inequality

$$\left|a(T_1 - \mu_1) + b(T_2 - \mu_2)\right| \leq |a||T_1 - \mu_1| + |b||T_2 - \mu_2|$$

since $T_1 \xrightarrow{P} \mu_1$ and $T_2 \xrightarrow{P} \mu_2$:

$$|T_1 - \mu| \xrightarrow{P} 0 \quad \& \quad |T_2 - \mu_2| \xrightarrow{P} 0$$

multiplying by constants.

$$|a||T_1 - \mu_1| \xrightarrow{P} 0 \quad \text{and} \quad |b||T_2 - \mu_2| \xrightarrow{P} 0$$

Adding two quantities converging in probability to 0 also converges to 0:

$$\left|a(T_1 - \mu_1) + b(T_2 - \mu_2)\right| \xrightarrow{P} 0$$

conclusion

$$aT_1 + bT_2 \xrightarrow{P} a\mu_1 + b\mu_2$$

Hence, $aT_1 + bT_2$ is a consistent estimator of $a\mu_1 + b\mu_2$

4) If x_1, x_2 and x_3 constitute a random sample of size 3 from normal population with mean μ and variance σ^2. Find the most efficient estimator among the three statistics $t_1 = \frac{x_1+x_2+x_3}{3}$, $t_2 = \frac{x_1+2x_2+x_3}{4}$ and $t_3 = x + \frac{x_2+x_3}{2}$

A) $T = \sum w_i x_i$.

i) $\frac{x_1+x_2+x_3}{3}$: $\sum w_i^2 = 3 \cdot \left(\frac{1}{3}\right)^2 = \frac{1}{3}$

ii) $\frac{x_1+2x_2+3x_3}{6}$: $\sum w_i^2 = \left(\frac{1}{6}\right)^2 + \left(\frac{2}{6}\right)^2 + \left(\frac{3}{6}\right)^2 = \frac{14}{36} = \frac{7}{18}$

iii) $\frac{x_1+x_2+3x_3}{5}$: $\sum w_i^2 = \left(\frac{1}{5}\right)^2 + \left(\frac{1}{5}\right)^2 + \left(\frac{3}{5}\right)^2 = \frac{11}{25}$

Smallest variance is (i)

most efficient: $\frac{x_1+x_2+x_3}{3}$

5) Let $x_1, x_2, x_3, \ldots, x_n$ be a random sample from Normal distribution $N(\mu, \sigma^2)$ population. Prove that $t = \frac{\sum_{i=1}^{n} x_i}{n}$ is a good estimator of μ.

A) $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

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

$\Rightarrow$ Due to the linearity of expectation, we can bring the constant $\frac{1}{n}$ outside.

$\Rightarrow$ $\cancel{\frac{1}{n}}$ $E(\bar{x}) = \frac{1}{n}\sum_{i=1}^{n} E(x_i)$

$\Rightarrow$ Since each x_i is from a normal distribution with mean μ, the expected value of each x_i is $E(x_i) = \mu$

$\Rightarrow$ $E(\bar{x}) = \frac{1}{n}\sum_{i=1}^{n} \mu$.

$\Rightarrow$ $E(\bar{x}) = \frac{1}{n}(n\mu)$

$\Rightarrow$ This simplifies to $E(\bar{x}) = \mu$ proving that the sample mean is an unbaised estimator of μ.

6) Let $x_1, x_2, x_3, \ldots, x_n$ be a random from a population with population density function $f(x,\theta) = \theta x^{\theta-1}$; $0 < x < 1, \theta > 0$. Find the sufficient estimator θ.

A) $L(\theta | x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i, \theta)$

$L(\theta | x_1, \ldots, x_n) = \prod_{i=1}^{n} (\theta x_i^{\theta-1})$

$L(\theta | x_1, \ldots, x_n) = \theta^n \prod_{i=1}^{n} x_i^{\theta-1}$

$L(\theta | x_1, x_2, \ldots, x_n) = \theta^n \left(\prod_{i=1}^{n} x_i\right)^{\theta-1}$

$\Rightarrow \left(\prod_{i=1}^{n} (x_i)\right)^{\theta-1} = e^{\ln\left(\left(\prod_{i=1}^{n} x_i\right)^{\theta-1}\right)} = e^{(\theta-1)\sum_{i=1}^{n} \ln(x_i)}$

$\Rightarrow L(\theta | x_1, \ldots, x_n) = \theta^n e^{(\theta-1)\sum_{i=1}^{n} \ln(x_i)}$

$\Rightarrow L(\theta | x_1, \ldots, x_n) = \underbrace{\left(\theta^n e^{\theta \sum_{i=1}^{n} \ln(x_i)}\right)}_{g(T(x), \theta)} \underbrace{\left(e^{-\sum_{i=1}^{n} \ln(x_i)}\right)}_{h(x)}$

$\Rightarrow$ The sufficient estimator for θ is $T(x) = \sum_{i=1}^{n} \ln(x_i)$

7) The mean and variance of a random sample of 64 observation were computed as 160 and 100 respectively. Compute the 95% confidence limit of population mean.

A) $n = 64$

$s^2 = 100 \Rightarrow s = 10$

$S.E = \frac{s}{\sqrt{n}} \Rightarrow \frac{10}{\sqrt{64}} = \frac{10}{8} = 1.25$

margin of error $= Z_{\alpha/2} \times S.E$

$= 1.96 \times 1.25 = 2.45$

confidence limit : $\bar{x} \pm$ margin of Error $= 160 \pm 2.45$

$\Rightarrow$ The 95% confidence limit for the population mean are $(157.55, 162.45)$.

10) A random sample of 700 units from a large consignment and in that 200 were damaged. Find 95% confidence limit for the ~~por~~proportion of damage units in the consignment.

A) $p = \frac{200}{700} \simeq 0.2857$

$S.E = \sqrt{\frac{P(1-P)}{n}} \Rightarrow \sqrt{\frac{0.2857(1-0.2857)}{700}} \Rightarrow \sqrt{0.0002914} \simeq 0.01707$

margin of Error = $Z \cdot SE$

$= 1.96 \times 0.01707$

$= 0.03346$

lower limit = $0.2857 - 0.03346 = 0.25224$

Upper limit = $0.2857 + 0.03346 = 0.31916$

$\Rightarrow$ The 95% confidence limit for the proportion of damaged unit is between 25.2% to 31.9%.

11) out of 2000 customer saving account a sample of 600 accounts was taken to rest the accuracy of hosting and balancing there in 45 mistakes were found assign limits with in no of defective case can be expected 95% level. Find out the confidential limit for 95% significant level.

A) $p = \frac{45}{600} = 0.075$

$= p \times$ population size $\Rightarrow 0.075 \times 2000 \Rightarrow 1500$

$\Rightarrow$ The standard deviation of the proportion is

$\sqrt{\frac{P(1-P)}{n}} \Rightarrow \sqrt{\frac{0.075(1-0.075)}{600}} \simeq 0.0107$

$\Rightarrow$ margin of error $\Rightarrow Z \times \sigma_p = 1.96 \times 0.0107 \simeq 0.021$

$\Rightarrow$ margin of error of total population $\Rightarrow 0.021 \times 2000 \Rightarrow 420$

$\Rightarrow$ lower limit $\Rightarrow 1500 - 420 = 1080$

$\Rightarrow$ higher limit $\Rightarrow 1500 + 420 \Rightarrow 1920$

13) A research worker wishes to estimate the mean of population by using sufficiently large sample. The probability is 0.95 that the sample mean will ~~not~~ not differ from the true mean by more than 25% of the standard deviation. How large a sample should be taken?

A)
$$\varepsilon = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$0.25\cancel{\sigma} = 1.96 \times \frac{\cancel{\sigma}}{\sqrt{n}}$$

$$\Rightarrow \sqrt{n} = \frac{1.96}{0.25} \Rightarrow 7.84$$

$$\Rightarrow n = 61.47 \simeq 62$$

15) A manufacturing concern to estimate the average amount purchase of its product in a month by the customers. If the standard deviation is Rs 10, find sample size, if the maximum error is not to exceed Rs 3 with probability of 0.99.

A) $\sigma = 10$

maximum error = 3

$$\varepsilon = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow 3 = 2.576 \times \frac{\sigma}{\sqrt{n}} \Rightarrow 3 = 2.576 \times \frac{10}{\sqrt{n}}$$

$$\Rightarrow \sqrt{n} = 8.5867 \Rightarrow n = 73.73$$

rounding off $n = 74$

17) A random sample of 100 articles selected from a large batch of articles contain 5 defective articles. Set up 99 percent confidence limits for the ~~proo~~ proportion of defective items in the batch.

A) $n = 100$

no. defective $x = 5$

$$p = \frac{5}{100} = 0.05$$

$$z_{0.005} = 2.576$$

$$SE = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.05 \times 0.95}{100}} = \sqrt{0.000475} = 0.0217$$

margin of error

$$\mathcal{E} = z_{0.005}\, SE = 2.576 \times 0.02179 \simeq 0.05617$$

So, the normal approximation 99% CI is

$$\hat{p} \pm \mathcal{E} = 0.05 \pm 0.05617$$

Lower limit $= 0.05 - (2.576 \times 0.0218) \simeq -0.0062$

Upper limit $= 0.05 + (2.576 \times 0.0218) \simeq 0.1062$

⇒ The 99% confidence limit for proportion of defective items in the batch are approximately $(0.0011, 0.1062)$.

14) A random sample of 500 pineapple was taken from a large consignment and 65 of them were found to be bad. Show that the standard error of the proportion of bad ones in a sample of the size is 0.015.

A) $\hat{p} = \frac{65}{500} = 0.13$

$$SE = \sqrt{\frac{p(1-p)}{n}} \Rightarrow \sqrt{\frac{0.13(1-0.13)}{500}} \Rightarrow 0.0150$$

Rounding off the standard error of the proportion of bad pineapple is 0.015.

20) A random sample of 100 article selected from a large batch of articles contain 5 defective articles. If the batch contains 2669 items, set up 95% confidence interval for the population of defective items.

A) $\hat{p} = \frac{x}{n} \Rightarrow \frac{5}{100} = 0.05$

Margin of error $= Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$\Rightarrow 1.96\sqrt{\frac{0.05(1-0.05)}{100}} \Rightarrow 1.96\sqrt{\frac{0.0475}{100}} \simeq 0.0427$

CI for p. $\hat{p} \pm ME = 0.05 \pm 0.0427$

lower bound $\hat{p} \Rightarrow$ $0.05 - 0.0427 = 0.0073$

upper bound $\hat{p} \Rightarrow$ $0.05 + 0.0427 = 0.0927$

Lower bound $\Rightarrow 2669 \times 0.0073 \simeq \underline{19.48}$

upper bound $\Rightarrow 2669 \times 0.0927 \simeq \underline{247.58}$

21) A research worker wishes to estimate the mean of population by using sufficiently large sample. The probability is 0.95 that the sample mean will not differ from the true mean by more than 25% of the std deviation. How large a sample should be taken?

A) $P(|\bar{x} - \mu| \leq 0.25\sigma) = 0.95$

$\Rightarrow 0.25\sigma = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

$\Rightarrow 0.25\cancel{\sigma} = \frac{1.96\cancel{\sigma}}{\sqrt{n}} \Rightarrow 0.25 = \frac{1.96}{\sqrt{n}} \Rightarrow \sqrt{n} = 7.84$

$\Rightarrow n = (7.84)^2 \Rightarrow n = 61.4656$

$\Rightarrow$ Rounding off $n = \underline{62}$

The sample of 62 should be taken.