1. Let X be distributed as the binomial distribution with sample size n and proportion parameter p. Show that, for large n, the sample proportion of X/n satisfies

$$\frac{(X/n-p)}{\sqrt{p(1-p)/n}} \sim N(0,1)$$

Ans:- let X be a random variable that follows a binomial distribution, with parameters

n – number of trials

p – probability of success in each trial

$\therefore \quad X \sim \text{Binomial}(n,p)$

→ Mean and Variance of X

$$\mu_X = E[X] = n.p$$

$$\sigma_X^2 = Var(X) = n.p.(1-p)$$

From, Central limit theorem,

$$Z = \frac{X-\mu_X}{\sigma_X}$$

converges in distribution to a standard normal distribution;

$$Z \sim N(0,1)$$

substituting the mean and variance of X:

$$Z = \frac{X-n.p}{\sqrt{n.p.(1-p)}} \sim N(0,1)$$

$$\Rightarrow \quad \boxed{Z = \frac{X/n-p}{\sqrt{p(1-p)/n}} \sim N(0,1)}$$

2. A special purpose coating must have the proper abrasion. The standard deviation is known to be 21. Consider a random sample of 49 abrasion measurements.

   (a) Find the probability that the sample mean lies within 2 units of population mean.

   (b) Find the number k such that $P(-k \leq \bar{X} - \mu \leq k) = 0.9$

Ans: given,

$\sigma = 21$

$n = 49$

(a) We need to find $P(\mu - 2 < \bar{X} < \mu + 2)$

$$\text{standard error} = \frac{\sigma}{\sqrt{n}} = \frac{21}{\sqrt{49}} = \frac{21}{7} = 3$$

$$Z_1 = \frac{(\mu - 2) - \mu}{S.E} = \frac{-2}{3}$$

$$Z_2 = \frac{(\mu + 2) - \mu}{SE} = \frac{2}{3}$$

$$\therefore \quad P\left(-\frac{2}{3} \leq Z \leq \frac{2}{3}\right)$$

3. Find the probability that a random sample of 25 observations from a normal population with variance $\sigma^2 = 6$ will have a sample variance S^2 (i) greater than 9.1 (ii) between 3.462 and 10.745.

Ans:: given,

- population variance $\sigma^2 = 6$
- Sample size, $n = 25$

$$\therefore \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$$

(i) Probability that the sample variance is greater than 9.1

$\rightarrow P(s^2 > 9.1)$

$$\Rightarrow \frac{(n-1)s^2}{\sigma^2} = \frac{(25-1)\times 9.1}{6} = \frac{24 \times 9.1}{6} = 36.4$$

$\therefore$ we need to find,

$$P(\chi^2_{(24)} > 36.4)$$

in chi-square distribution table, $\chi^2_{24} = 36.4$ corresponds to 0.05

$$\therefore \boxed{P(s^2 > 9.1) \simeq 0.05}$$ A

(ii) Probability that the sample variance is between 3.462 and 10.745

here, we need to find :- $P(3.462 \le s^2 \le 10.745)$

So, for $s^2 = 3.462 \rightarrow \frac{(n-1)s^2}{\sigma^2} = \frac{(25-1)\,3.462}{6} = 13.848$

& $s^2 = 10.745 \rightarrow \frac{(n-1)s^2}{\sigma^2} = \frac{(25-1)\,10.745}{6} = 42.98$

$\therefore P(13.848 \le \chi^2_{24} \le 42.98)$, here, $\chi^2_{24} = 13.848 \rightarrow 0.95$

$\chi^2_{24} = 42.98 \rightarrow 0.01$ $= \boxed{0.94}$ A

4. A student at a large midwestern university questioned $n = 40$ students concerning the amount of time they spent doing community service during the past month. The data on times, in hours, are presented in the following table -

0	0	0	0	0	0	0	1	1	1
2	2	2	2	2	3	3	3	3	4
4	4	4	5	5	5	5	5	5	5
5	5	6	6	6	8	10	15	20	25

Give a point estimate of the population mean and state a 95% error margin.

Ans:- given,

$n = 40$

$(1-\alpha) = 0.95 \Rightarrow \alpha = 0.05$

$\therefore \alpha/2 = 0.025$

$t_{\alpha/2} \Rightarrow t_{0.025}(df) \Rightarrow t_{(0.025)(39)} \Rightarrow t_{(0.975)} = 2.021$

$$\bar{X} = \frac{0+0+\dots\dots+25}{40} = 4.55$$

Now,

Confidence Interval $= \bar{X} \pm$ Table.value. (Estimator)

$$= 4.55 \pm t_{\alpha/2} \times \left(\frac{S}{\sqrt{n}}\right)$$

$$= 4.55 \pm (2.021)\left(\frac{5.168}{\sqrt{40}}\right)$$

$$= 4.55 \pm 1.651$$

$$= (2.899, 6.201)$$

$$\therefore \boxed{(2.89 < \mu < 6.20)}$$

5. A limnologist wishes to estimate the mean phosphate content per unit volume of lake water. It is known from studies in previous years that the standard deviation has a fairly stable value $\sigma = 4$. How many water samples must be analyzed to be 90% certain that the error of estimation does not exceed 0.8 mg?

Ans:- given,

$\sigma = 4$

$(1-\alpha) = 0.90 \Rightarrow \alpha = 0.10 \Rightarrow Z_{0.05} = 1.645$

$n = ?$

$E = 0.8\text{ mg}$

here, we can use the formula for sample size estimation for population mean,

$$n = \left(\frac{Z_{\alpha/2} \cdot \sigma}{E}\right)^2$$

$$\Rightarrow n = \left(\frac{1.645 \times 4}{0.8}\right)^2$$

$$\therefore \boxed{n = 68}$$

$\therefore$ The limnologist needs to analyze 68 water samples to be 90% certain that the error of estimation does not exceed 0.8 mg.

6. A city health department wishes to determine if the mean bacteria count per unit volume of water a a lake beach is within the safety level of 200. A researcher collected 10 samples of unit volume and found the bacteria counts to be

175	190	205	193	184
207	204	193	196	180

Do the data strongly indicate that there is no cause for cencern. Test with $\alpha = 0.05$.

Ans: given,

$n = 10$

$\alpha = 0.05 \Rightarrow \alpha/2 = 0.025$

- Null Hypothesis (H_0): The mean bacteria count is equal to 200

$H_0 : \mu = 200$

- Alternate Hypothesis (H_1): The mean bacteria count != 200

$H_1 : \mu \neq 200$

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Now,

$$\bar{x} = \frac{175+190+205+193+184+207+204+193+196+180}{10}$$

$= 224.7$

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} \Rightarrow s^2 = 1623.41 \therefore s = 40.31$$

Now,

$$t = \frac{224.7 - 200}{40.31/\sqrt{10}} \Rightarrow t = 1.93$$

$df = n-1 = 9 \quad \therefore t_{(\alpha/2)(df)} = t_{(0.025)(9)} \Rightarrow 2.262$

Since, t value = 1.93 computed is between $-2.262 < t < 2.262$

$\therefore$ we fail to reject the null hypothesis.

7. Consider the following sample for fat content of $n = 10$ randomly selected hot dogs:

$$25.2 \quad 21.3 \quad 22.8 \quad 17.0 \quad 29.8 \quad 21.0 \quad 25.5 \quad 16.0 \quad 20.9 \quad 19.5$$

Assuming that these were selected from normal population, obtain point estimators for population mean and variance. Also, obtain a 95% CI for population mean and variance. Also, test $H_0 : \mu = 8.5$ versus $H_1 : \mu \neq 8.5$ with $\alpha = 0.05$.

Ans:- given, $n = 10$, $1-\alpha = 0.95 \Rightarrow \alpha = 0.05$

$$\overline{X} = \frac{25.2 + 21.3 + 22.8 + 17.0 + 29.8 + 21.0 + 25.5 + 16.0 + 20.9 + 19.5}{10}$$

$$\Rightarrow \overline{X} = 21.89$$

and, sample variance $(s^2) = (25.2-21.89)^2, (21.3-21.89)^2, \ldots\ldots (19.5-21.89)^2$

$$= 17.06$$

$\therefore \quad s = \sqrt{17.06} = 4.13$

here. $n = 10 \therefore df = 9$

$$\Rightarrow t_{\alpha/2}(df) = t_{0.025}(9) = 2.262$$

Now, $\quad \text{Error} = t_{\alpha/2}(df) \cdot \frac{s}{\sqrt{n}} = 2.262 \cdot \frac{4.13}{\sqrt{10}} = 2.96$

$\therefore$ confidence interval $= \overline{X} \pm$ table.value. (E)

$$= 21.89 \pm (2.262)(2.96)$$

$$= (18.93, 24.85)$$

→ Hypothesis test, $\quad H_0 :- \mu_0 = 8.5$ → Null Hypothesis

$\quad H_1 :- \mu_0 \neq 8.5$ → Alternate Hypothesis.

Now, $\quad t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{21.89 - 8.5}{4.13/\sqrt{10}} = 10.22$

Since, $t = 10.22$ is much greater than 2.262, $\boxed{\text{we reject null hyp.}}$

8. Assume that the lifespan of Indian males is normally distributed with unknown mean and standard deviation. A sample of 30 mortality histories of Indian males show that $\bar{x} = 71.3$ years and $s^2 = 128$ square years. Determine the observed values of 95 % confidence intervals for population mean and variance.

Ans:- given,

$n = 30$ $\quad\quad 1-\alpha = 0.95$

$\bar{x} = 71.3$ $\quad\quad \Rightarrow \alpha = 0.05$

$s^2 = 128$ $\quad\quad \therefore \alpha/2 = 0.025$

$\therefore s =$ $\quad\quad$ here, t-table as $n \leq 30$

Now,

(i) Confidence interval for mean (μ) is $= \bar{X} \pm t_{\alpha/2(n-1)} \frac{s}{\sqrt{n}}$

$$\Rightarrow 71.3 \pm t_{0.025(29)} \cdot \frac{11.313}{\sqrt{30}}$$

$$= 71.3 \pm 2.045 \cdot \frac{11.313}{\sqrt{30}}$$

$$= \boxed{(67.08, 75.52)}$$ Ans

(ii) Confidence interval for population variance

$$= \left(\frac{(n-1) \cdot s^2}{\chi^2_{\alpha/2\,(n-1)}}, \frac{(n-1) \cdot s^2}{\chi^2_{1-\alpha/2\,(n-1)}} \right)$$

$$= \left(\frac{(29) \cdot (128)}{45.72}, \frac{(29) \cdot (128)}{16.05} \right)$$

$$= \boxed{(81.16, 231.19)}$$ Ans

9. According to Nielsen Media Research, the average number of hours of TV viewers per household per week in the United States is 50.4 hours. Suppose the standard deviation is 11.2 hours and a random sample of 49 U.S. households is taken.

(a) What is the probability that the sample average is more than 52 hours?

(b) Suppose the population standard deviation is unknown. If 71% of all sample means are greater than 49 hours and the population mean is still 50.4 hours, what is the population S.D.?

Ans:- given,

$\mu = 50.4$

$\sigma = 11.2$

$n = 49$

σ_x = standard error of the mean,

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{11.2}{\sqrt{49}} = \frac{11.2}{7} = 1.6$$

Now,

(a) $$Z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} \Rightarrow Z = \frac{52 - 50.4}{1.6} = 1.00$$

cumulative probability for $z = 1.00$ is 0.8413

$\therefore$ $P(\bar{x} > 52) = 1 - 0.8413 = 0.1587$

$\therefore$ The probability that the sample average is more than 52 hours is 0.1587 or 15.87%.

(b) $\bar{x} = 49$,

as 71% of sample means are greater than 49 hours,

Z-score $= 1 - 0.71 = 0.29 \rightarrow -0.55$

Now,

$$Z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} \Rightarrow -0.55 = \frac{49 - 50.4}{\sigma_{\bar{x}}}$$

$\therefore$ $\sigma_{\bar{x}} = 2.55$

as,

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \Rightarrow 2.55 = \frac{\sigma}{\sqrt{49}}$$

$\therefore$ $\boxed{\sigma = 17.85}$ Ans

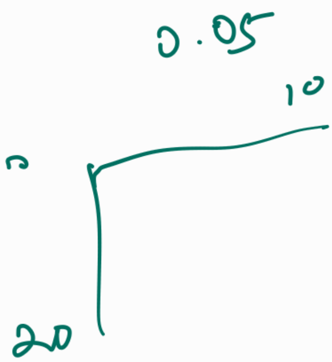
10. The strength of steel wire made by an existing process is normally distributed with a mean of 1250 and a standard deviation of 150. A batch of wire is made by a new process, and a random sample consisting of 25 measurements gives an average strength of 1312. Assume that the standard deviation does not change. Is there evidence at the 99% level of significance that the new process gives a larger mean strength than the old?

Ans:-

11. Two additives to Portland cement are being tested for their effect on the strength of concrete. 21 batches were made with Additive A, and their strengths showed standard deviation $s_A = 41.3$. 16 batches were made with the same percentage of Additive B, and their strengths showed standard deviation $s_B = 26.2$. Assume that the strengths of concrete follow a normal distribution. Is there evidence at the 90% level of significance that concrete made with Additive A and concrete made with Additive B have different variabilities?

$A \to 21 \to S_A = 41.3$

$B \to 16 \to S_B = 26.2$



$H_0 \to S_A^2 = S_B^2$

$H_1 \to S_A^2 \neq S_B^2$

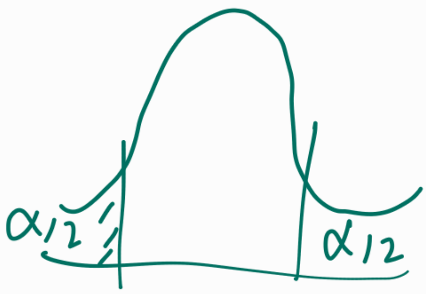
$90 = \alpha$

CI

$90 \to$ CI

0.10

$0.45 = \alpha$



$\boxed{0.05}$

$$F = \frac{S_A^2}{S_B^2} = \frac{S^2_{(21-1)}}{S^2_{(16-1)}} : \frac{S^2_{20}}{S^2_{10}}$$