HW10

Name:- Chandan Suri

1. $\sigma = 9$

Step 1:- Parameter of interest :- mean μ

Data is normally distributed $(N(\mu, \sigma^2))$

$\sigma = 9 \Rightarrow \sigma^2 = 9^2 = 81$

Step 2:-

$\boxed{H_0 : \mu = 75}$ (Null hypothesis)

$\boxed{H_a : \mu < 75}$ (Alternative hypothesis)

$n = 25$ observation

(a)

To find how many standard deviations below the null value we are when $\bar{X} = 72.3$.

Assuming the null hypothesis is true, the test-statistic (z) is given by:-

$$Z = \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} = \frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}}$$ (for normal distribution)

Here, $\bar{X}_n = 72.3$, $\mu = \mu_0 = 75$ (since we are assuming that null hypothesis is true)

$n = 25$, $\sigma = 9$.

$$\therefore \Rightarrow \frac{72.3 - 75}{9/\sqrt{25}} = \frac{-2.7 \times 5}{9} = -\frac{13.5}{9} = \boxed{-1.5}$$

$\therefore$ $\bar{X}_n$ is 1.5 standard deviations below the null value.

Ans.

(b.) $\bar{X}_n = 72.3$, $\alpha = 0.002$ (level of significance).

Step ① & ② are shown in the previous page.

Starting with step 3:-

Step 3:- $\alpha \in (0,1)$ is called the significance level

$$\boxed{\alpha = 0.002}$$

Step 4:- Now, we need to find the test-statistic,

$$T_n(\theta) = \frac{\bar{X}_n - \theta}{\sqrt{\sigma^2/n}}$$ where θ is the parameter of interest

$$\boxed{T_n(\mu_0) = \frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}}}$$ (μ_0 is the null value assuming the null hypothesis is true)

Putting the required values,

$$T_n(\mu_0) = \frac{72.3 - 75}{\sqrt{81/25}} = -1.5$$

Step 5:- Now, we need to calculate p-value,

$$\text{p-value} = P\left(\frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}} \le -1.5 \text{ if } H_0 \text{ is true}\right)$$

$$= \Phi(-1.5)$$

$$= 1 - \Phi(1.5) = 1 - 0.9332 = \underline{0.0668}$$

$\because$ p-value (0.0668) is greater than the level of significance ($\alpha = 0.002$), then we do not reject the null hypothesis. Ans (c.) $\alpha = 0.002$

In this question, we need to find $\beta(70)$, i.e., type-II error where $\theta = 70$.

Step 1:- Determine for which value of the test statistic H_0 is rejected:-

We reject H_0 if,

$$\text{p-value} = P\left(\frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}} \leq t_n(\theta_n) \text{ if } H_0 \text{ is true}\right) \leq 0.002$$

So, we reject H_0 if $t_n(\theta_0) \leq z$ where z is the number at which:-

$$P\left(\frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}} \leq z \text{ if } H_0 \text{ is true}\right) = 0.002$$

The number z is called the critical value of the test statistic.

Because,

$$\frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}} \sim N(0,1) \text{ under } H_0, \text{ we have:-}$$

$$z = \Phi^{-1}(0.002) = -\Phi^{-1}(0.998) = -2.88$$

Step 2:- Determine $\beta(\theta)$

$$\beta(70) = P\left(\frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}} > -2.88 \text{ if } \mu = 70\right)$$

$$= P\left(\frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} + \frac{\mu - \mu_0}{\sqrt{\sigma^2/n}} > -2.88\right)$$

$$= P\left(\frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} > -2.88 - \frac{\mu - \mu_0}{\sqrt{\sigma^2/n}}\right)$$

$$= P\left(\frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} > -2.88 - \left(\frac{70-75}{\sqrt{81/25}}\right)\right)$$

$$= P\left(\frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} > -2.88 + \frac{5 \times 5}{9}\right)$$

$$= P\left(\frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} > -2.88 + \frac{25}{9}\right)$$

$$= P\left(\frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} > -0.10\right) = 1 - \Phi(-0.10)$$

$$= \Phi(0.01) = \underline{\underline{0.5398}}$$

Hence, $\boxed{\beta(70) = 0.5398 = 53.98\%}$

Ans.

(d) Now, we need to find n such that $\underline{\beta(70) = 0.01}$.

We know that:-

$$\beta(70) = P\left(\frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}} > -2.88 \text{ if } \mu = 70\right)$$

$$= P\left(\frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}} > -2.88 - \left(\frac{70-75}{\sqrt{\sigma^2/n}}\right)\right)$$

(from the steps in part (c))

$[\because \mu_0 = 75]$

$$= 1 - \Phi\left(-2.88 - \frac{(70-75)}{\sqrt{81/n}}\right)$$

$[\because \sigma = 9]$

$$\Rightarrow 1 - \Phi\left(-2.88 - \left(\frac{70-75}{\sqrt{81/n}}\right)\right) \overset{!}{=} 0.01$$

$$\Rightarrow \Phi\left(-2.88 - \frac{(70-75)}{\sqrt{81/n}}\right) = 0.99$$

$$\Rightarrow -2.88 - \frac{(70-75)}{\sqrt{81/n}} = \Phi^{-1}(0.99) = 2.326$$

Solving for n,

$$\frac{5}{\sqrt{81/n}} = 2.88 + 2.326$$

$$\frac{5}{9}\sqrt{n} = 2.88 + 2.326$$

$$\sqrt{n} = \frac{(2.88 + 2.326)\,9}{5}$$

$$\Rightarrow n = \left[\frac{(2.88 + 2.326) \times 9}{5}\right]^2$$

$$n = \left(\frac{9}{5} \times 5.206\right)^2 = (9.3708)^2 = 87.81$$

$$\Rightarrow \boxed{n \approx 88} \quad (\text{Rounded up})$$

$\therefore$ Required sample size is 88. Ans

(e.) $\alpha = 0.01$ & $n = 100$.

Type I error is defined as the error when we reject the null hypothesis when it is true,

Thus,

$$P(\text{Type I error}) = \alpha$$

$\because \alpha = 0.01$,

$\therefore$ $\underline{\underline{P(\text{Type I error}) = 0.01}}$ (already provided)

$\therefore$ Probability of Type I error is $\underline{0.01}$.

<u>Ans.</u>

(2.) $n = 58$

Step 1:- Since we are concerned with the average estimated calorie content,

$\therefore$ parameter of interest is the mean μ.

Also, $\because$ σ is unknown (but we have been provided the sample standard deviation S_n),

and our sample-size is large.

$\therefore$ ~~Data is distributed as t_{n-1} distribution~~

Data follows a normal distribution. ~~$[T_n(\mu_0) \sim t_{n-1}]$~~ ~~$(N(0,1))$~~.

& $S_n = 89 \Rightarrow$
$S_n^2 = (89)^2 = \underline{7921}$.

Step 2:-

Here,

$\underline{H_0: \ \mu = 153}$ (Null Hypothesis)

$\underline{H_a: \ \mu > 153}$ (Alternative Hypothesis)

$\boxed{n = 58}$ (given).

Step 3:-

$\alpha \in (0, 1)$ is called the significance level.

$\boxed{\alpha = 0.001}$

Step 4:- Now, we need to find the test statistic,

$$T_n(\theta) = \frac{\bar{X}_n - \theta}{\sqrt{S_n^2/n}}$$

where θ is the parameter of interest. $[\because \sigma \text{ unknown}]$

$$\Rightarrow T_n(\mu_0) = \frac{\bar{X}_n - \mu_0}{\sqrt{S_n^2/n}}$$

(μ_0 is the null value assuming the null hypothesis is true).

Putting in the required values, $\boxed{\bar{X}_n = 191}$ (Sample mean)

$$t_n(\mu_0) = \frac{191 - 153}{\sqrt{7921/58}} = \frac{38 \times \sqrt{58}}{\sqrt{7921}} = 3.2517$$

$$\approx \underline{\underline{3.25}}.$$

Step 5:- Now, we need to calculate the p-value,

$$\text{p-value} = P\left(\frac{\bar{X}_n - \mu_0}{\sqrt{S_n^2/n}} > 3.25 \text{ if } H_0 \text{ is true} \right)$$

$$= 1 - \Phi(3.25) = 1 - 0.9994 = \underline{\underline{0.0006}}$$

$\because$ p-value (0.0006) is less than the level of significance ($\alpha = 0.001$), $\therefore$ we reject H_0 (the null hypothesis).
$[p \leq \alpha]$

Ans.

Thus, we have sufficient evidence to conclude that the true average estimated calorie content is more than 153 calories at 0.1% level of significance.

3. I acknowledge that I have read the examples on pages 16 and 17 of the mentioned pdf.

Ans

4. Observations that are on stopping distance (ft) of a particular truck at 20 mph under specified experimental conditions:-

32.1   30.6   31.4   30.4   31.0   31.9

Maximum allowable stopping distance = 30.

The data is normally distributed.

Also, $\boxed{\alpha = 0.01.}$

Step 1:- Parameter of interest:- mean μ.

It is given that the data is normally distributed $(N(\mu, \sigma^2))$.
But as σ is unknown, we will consider to follow t_{n-1} distribution.
→ we need to determine the values of mean and standard deviation using the data points given.

We know that:-

$$\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{6}[32.1 + 30.6 + 31.4 + 30.4 + 31.0 + 31.9]$$

$$= \frac{1}{6} \times 187.4 = \underline{\underline{31.23}}$$

$$\Rightarrow \boxed{\bar{x}_n = 31.23}$$

Standard deviation ⇒ $(\text{Variance})^2$

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

$$\Rightarrow \frac{1}{6-1}\left[(32.1-31.23)^2 + (30.6-31.23)^2 + (31.4-31.23)^2 + (30.4-31.23)^2 + (31-31.23)^2 + (31.9-31.23)^2\right]$$

$$= \frac{1}{5}[0.7569 + 0.3969 + 0.0289 + 0.6889 + 0.0529 + 0.4489]$$

$$\Rightarrow \frac{1}{5} \times 2.3734 = 0.47468$$

$$S \Rightarrow \sqrt{0.47468} = 0.6889 \simeq 0.69$$

$$\boxed{\therefore S = 0.69}$$

Step 2:-

$H_0: \mu = 30$ (Null hypothesis)

$H_a: \mu > 30$ (Alternative hypothesis)

where the null value is 30.

Step 3:- $\alpha \in (0, 1)$ is the significance level.

It is given as 0.01.

$$\boxed{\therefore \alpha = 0.01}$$

Step 4:- The test statistic is :-

$$T_n(\theta) = \frac{\bar{X}_n - \theta}{\sqrt{s^2/n}}$$

[as we have been provided the sample values & σ is unknown]

$$T_n(\mu_0) = \frac{\bar{X}_n - \mu_0}{s/\sqrt{n}}$$

Putting values,

$$t_n(\mu_0) = \frac{31.23 - 30}{0.69/\sqrt{6}} = \frac{1.23 \times \sqrt{6}}{0.69} = 4.36$$

Step-5 :- Now, we will calculate the p-value,

$$\text{p-value} \Rightarrow P\left(\frac{\bar{X}_n - \mu}{S/\sqrt{n}} > 4.36 \text{ if } H_0 \text{ is true}\right)$$

from the table of the t- distribution with upper-tailed area values, we get that p-value should be less than 0.005 for a test statistic value of 4.36, (with 5 degrees of freedom)

$$\boxed{\text{p-value} < 0.005}$$

$\because \alpha = 0.01$,

$\therefore$ p-value is less than the level of significance, thus, we reject the null hypothesis.

Ans

Therefore, we can conclusively say that the true average stopping distance does not exceed the maximum value of 30 mph.

Ans.

(5.) The calibration of a scale is to be checked by weighing a 10-kg test specimen 25 times.

$n = 25$, $\mu_0 = 10 kg$, $\sigma = 0.2 kg$

Let μ denote the true average weight reading on the scale.

(a)

Null hypothesis:-

$\boxed{H_0: \mu = 10}$

This means that true measured mean in kg is 10.

Alternative hypothesis:-

$\boxed{H_a: \mu \neq 10}$

This means that true mean weight is not 10kg.

Also, level of significance $(\alpha) = 0.01$.

(b)

Step 1:- Parameter of interest:- mean μ

Data is normally distributed $N(\mu, \sigma^2)$.

where $\sigma = 0.2$ kg

Step 2:-

$H_0: \mu = 10$ (Null hypothesis)

$H_a: \mu \neq 10$ (Alternative hypothesis).

where the null value (μ_0) is 10.

Step 3:- The level of significance $(\alpha) = 0.01$ (given)

Step 4:- The test statistic is:-

$$T_n(\theta) = \frac{\bar{X}_n - \theta}{\sqrt{\sigma^2/n}} \quad [\sigma \text{ known}]$$

$$T_n(\mu_0) = \frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}}$$

$$t_n(\mu_0) = \frac{9.85 - 10}{0.2/\sqrt{25}} = \underline{\underline{-3.75}}.$$

<u>Step-5</u>:- Now, p-value can be calculated as:-

$$\text{p-value} = P\left(\left|\frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}}\right| \geq |-3.75| \text{ if } H_0 \text{ is true}\right)$$

$$= P\left(\left|\frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}}\right| \geq 3.75 \text{ if } H_0 \text{ is true}\right)$$

[After solving inequality]

$$= P\left(\frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} \geq 3.75 \text{ and } \frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} \leq -3.75 \text{ if } H_0 \text{ is true}\right)$$

$$= P\left(\frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} \geq 3.75\right) + P\left(\frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}} \leq -3.75\right) \text{ if } H_0 \text{ is true.}$$

$$= 1 - \Phi(3.75) + \Phi(-3.75) \quad \left[\because \Phi(-x) = 1 - \Phi(x)\right]$$

$$= 1 - \Phi(3.75) + (1 - \Phi(3.75))$$

$$= 2\left[1 - \Phi(3.75)\right] = 2\left[1 - 0.9999\right]$$

$$= 2 \times 0.0001 = \boxed{0.0002}$$

$\therefore$ p-value is less than the level of significance (0.01),

Therefore, we reject the null hypothesis.
— Ans.

Thus, we conclude that the ~~true~~ measured mean weight is not 10 kg.
Ans.

(c) To find the probability that recalibration is judged unnecessary, we need to find the probability of the type-II error (when not rejecting the null hypothesis is when it is false) → following from part (b).

Step-1 :- Determine for which values of the test statistic H_0 is rejected:-

we reject H_0 if:-

$$p\text{-value} = P\left(\left| \frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}} \right| \geq |t_n(\theta_0)| \text{ if } H_0 \text{ is true} \right) \leq 0.01$$

(given).

So, we reject H_0 if $t_n(\theta_0) \geq z$ where z is the critical value of the test statistic at which:-

$$\boxed{P\left(\left| \frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}} \right| \geq z \text{ if } H_0 \text{ is true} \right) = 0.01}$$

Because $\dfrac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}} \sim N(0,1)$ under H_0,

we have, $\Phi^{-1}(0.99)$ ~~from the table~~

$$z = \Phi^{-1}(\overset{1-0.01}{\cancel{0.01}}) = \cancel{2.326} = 2.326$$

So, we reject H_0 if $\dfrac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}} \geq 2.326$

Now, we need to determine $\beta(\theta)$ when $\theta = \mu = 10.1$ & $\theta = \mu = 9.8$

So,

to determine $\beta(\mu)$ for any $\mu \neq 10$, we need to find out the threshold between P-value $\leq \alpha$ and P-value $> \alpha$ in terms of mean.

from step ①,

we know,

$$P\left(\left|\frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}}\right| \geq z \text{ if } H_0 \text{ is true}\right) = 0.01$$

$$\Rightarrow 2P\left(\frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}} \leq z \text{ if } H_0 \text{ is true}\right)$$

$$\Rightarrow 2\Phi\left(\frac{\bar{X}_n - 10}{0.2/\sqrt{25}}\right) \Rightarrow 2\Phi\left(\frac{\bar{X}_n - 10}{0.04}\right) = 0.01$$

$$\Phi\left(\frac{\bar{X}_n - 10}{0.04}\right) = 0.005 \Rightarrow \frac{\bar{X}_n - 10}{0.04} = \Phi^{-1}(0.005)$$

$$\because \frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}} \sim N(0,1) \text{ under } H_0$$

$$\therefore \Phi^{-1}(0.005) = -2.58 \quad \text{(from the table)}$$

$$\Rightarrow \frac{\bar{X}_n - 10}{0.04} = -2.58 \Rightarrow \boxed{\bar{X}_n = 9.8968}$$

Thus, we would reject H_0 at $\alpha = 0.01$ iff the observed value of $\bar{X}_n$ is ≤ 9.8968,

thus, by symmetry of the normal distribution,

for $\geq 10 + (10 - 9.8968) = 10.1032$.

So, we would not reject H_0 at $\alpha = 0.01$ if :-

$$\boxed{9.8968 < \bar{X} < 10.1032}$$

Now, to determine the probability of the type-II error,

$$\beta(10.1) = P(9.8968 < \bar{X}_n < 10.1032) \text{ when } \mu = 10.1$$

$$= P\left(\left| \frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}} \right| < 2.58 \right) \text{ if } \mu = 10.1$$

$$= P\left(-2.58 < \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} + \frac{\mu - \mu_0}{\sqrt{\sigma^2/n}} < +2.58 \right)$$

$$= P\left(-2.58 - \frac{\mu - \mu_0}{\sqrt{\sigma^2/n}} < \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} < 2.58 - \frac{\mu - \mu_0}{\sqrt{\sigma^2/n}} \right)$$

$$= P\left(-2.58 - \frac{10 - 10.1}{0.2/\sqrt{25}} < \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} < 2.58 - \frac{10.0 - 10.1}{0.2/\sqrt{25}} \right)$$

$$= P\left(-5.08 < \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} < 0.08 \right) = \Phi(0.08) - \Phi(-5.08)$$

$$= 0.5319 - 0 = \boxed{0.5319} \quad (53.19\%)$$

Ans.

for $\mu = 9.8$,

$$\beta(9.8) = P\left(9.8968 < \bar{X}_n < 10.1032 \text{ if } \mu = 10.1\right)$$

$$= P\left(\left|\frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}}\right| < -2.58\right) \text{ if } \mu = 10.1$$

$$= P\left(-2.58 < \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} + \frac{\mu - \mu_0}{\sqrt{\sigma^2/n}} < 2.58\right)$$

$$= P\left(-2.58 - \frac{\mu - \mu_0}{\sqrt{\sigma^2/n}} < \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} < 2.58 - \frac{\mu - \mu_0}{\sqrt{\sigma^2/n}}\right)$$

$$= P\left(-2.58 - \frac{10 - 9.8}{0.2/\sqrt{25}} < \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} < 2.58 - \frac{10 - 9.8}{0.2/\sqrt{25}}\right)$$

$$= P\left(2.42 < \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} < 7.58\right)$$

$$= \Phi(7.58) - \Phi(2.42) = 1 - 0.9922$$

$$= \underline{\underline{0.0078}} \quad \underline{\underline{(0.78\%)}}$$

<u>Ans.</u>

(6.) $n = 15$, mean ($\bar{x}_n$) = 3.05 mm

$S_n = 0.34$ mm , $\mu = 3.2$ mm

$\alpha = 0.05$

Step 1:- Parameter of interest:- mean μ

Data is normally distributed, $\therefore N(\mu, \sigma^2)$

but because σ is unknown, it will follow t-distribution. with $(n-1) = 14$ degrees of freedom.

Step 2:- $H_0: \mu = 3.2$ (Null hypothesis)

$H_a: \mu \neq 3.2$ (Alternative hypothesis)

where ~~the~~ null value is 3.2.

Step 3:- $\alpha = 0.05$ is the level of significance.

Step 4:- The test statistic is given by:-

$$T_n(\mu_0) = \frac{\bar{X}_n - \mu_0}{S_n/\sqrt{n}} \qquad [\because \sigma \text{ unknown}]$$

$$t_n(\mu_0) = \frac{3.05 - 3.2}{0.34/\sqrt{15}} = \underline{-1.71}$$

Step 5:- p-value $= P\left(\left|\frac{\bar{X}_n - \mu_0}{S_n/\sqrt{n}}\right| \geq |-1.71| \text{ if } H_0 \text{ is true}\right)$

$= P\left(\left|\frac{\bar{X}_n - \mu_0}{S_n/\sqrt{n}}\right| \geq 1.71 \text{ if } H_0 \text{ is true}\right)$

$= P\left(\frac{\bar{X}_n - \mu_0}{S_n/\sqrt{n}} \geq 1.71\right) + P\left(\frac{\bar{X}_n - \mu_0}{S_n/\sqrt{n}} \leq -1.71\right)$ if H_0 is true from the table of t- distribution,

we can get value of these probabilities directly such that

$$P\left(\frac{\bar{X}_n - \mu_0}{S_n/\sqrt{n}} \geq 1.71\right) = P\left(\frac{\bar{X}_n - \mu_0}{S_n/\sqrt{n}} \leq -1.71\right)$$

$$\Rightarrow 2 \times P\left(\frac{\bar{X}_n - \mu_0}{S_n/\sqrt{n}} \geq 1.71\right) = 2 \times 0.055 = \underline{0.11}$$

Step-6:

Since $\alpha = 0.05$ and p-value is greater than the level of significance, we do not reject H_0.

Therefore, the data does not strongly suggest that the true average thickness of lenses is something other than what is desired.

Ans

7. $\sigma = 0.5$

Step-1:- Parameter of interest:- standard deviation σ

Data follows a χ^2 distribution with n-1 degrees of freedom.

$\sigma = 0.5 \Rightarrow \sigma^2 = (0.5)^2 = 0.25$ $(n-1 = 9)$

Step2:-

$H_0: \sigma^2 = \sigma_0^2 = 0.25$ (Null hypothesis)

$H_a: \sigma^2 > \sigma_0^2 \Rightarrow \sigma^2 > 0.25$ (Alternative hypothesis)

$n = 10$ (10 different specimens)

Step3:- $\alpha \in (0,1)$ is called the significance level.

$$\boxed{\alpha = 0.01}$$

Step4:-

Now, we need to find the test statistic,

$$t_n(\sigma_0) = \frac{(n-1)S_n^2}{\sigma_0^2} = \frac{(10-1)\times(0.58)^2}{0.25}$$

$$= \frac{9\times(0.58)^2}{0.25} = 12.1104$$

Step5:- Now, we need to calculate p-value,

$$\text{p-value} = P\left(\frac{(n-1)S_n^2}{\sigma^2} > 12.1104 \text{ if } H_0 \text{ is true}\right)$$

$\Rightarrow$ for χ_9^2 distribution, from the table,

$$\boxed{\text{p-value is } > 0.100}$$

$\therefore$ p-value is greater than the level of significance $(\alpha = 0.01)$, then, we do not reject the null hypothesis.

$\therefore$ The uniformity specification cannot be strongly contradicted.

Ans.

(8)

Ethnicity:	(1) African American	(2) Asian	(3) Caucasian	(4) Hispanic	Total
frequency	57	11	330	6	404
Census Proportions (True)	0.177	0.032	0.734	0.057	1

The null hypothesis is:-

$$H_0: p_1 = 0.177,\ p_2 = 0.032,\ p_3 = 0.734,\ p_4 = 0.057$$

Alternative hypothesis $\Rightarrow$ H_1: At least one p_i (proportions) is not equal to its null value. ($p_{i,0}$)

And level of significance $(\alpha) = 0.01$

Here $n = 404$ people.

Test-statistic can be calculated as:-

$$T(X, p_{1,0}, \dots, p_{k,0}) = \sum_{k=1}^{k} \left(\frac{(N_k - n p_{k,0})^2}{n p_{k,0}} \right)$$

where N_k is the number of X_i's equal to the k th category/value.

$$\therefore\ t(x, p_{1,0}, p_{2,0}, p_{3,0}, p_{4,0}) = \sum_{k=1}^{4} \left(\frac{(N_k - n p_{k,0})^2}{n p_{k,0}} \right)$$

$$= \left(\frac{(57 - 404 \times 0.177)^2}{404 \times 0.177} \right) + \left(\frac{(11 - 404 \times 0.032)^2}{404 \times 0.032} \right) +$$

$$\left(\frac{(330 - 404 \times 0.734)^2}{404 \times 0.734} \right) + \left(\frac{(6 - 404 \times 0.057)^2}{404 \times 0.057} \right)$$

$$= \left(\frac{(57-71.508)^2}{71.508}\right) + \left(\frac{(11-12.928)^2}{12.928}\right) + \left(\frac{(330-296.536)^2}{296.536}\right) + \left(\frac{(6-23.028)^2}{23.028}\right)$$

$$= 2.9435 + 0.2875 + 3.7764 + 12.5913$$

$$= 19.5987 = \underline{19.60} \quad \text{(rounded)}$$

The degrees of freedom can be calculated as :)

$$df = k - 1 = 4 - 1 = 3.$$

As the above statistic follows a χ^2_3 distribution,

Now, we first need to calculate p-value and look into the table for χ^2-distribution with degrees of freedom as 3,

$$\text{p-value} = P\big(T(X, p_{1,0}, p_{2,0}, p_{3,0}, p_{4,0}) \text{ is more extreme than } t(x, p_{1,0}, p_{2,0}, p_{3,0}, p_{4,0}) \text{ if } H_0 \text{ is true}\big)$$

$$\Rightarrow P\big(T(X, p_{1,0}, p_{2,0}, p_{3,0}, p_{4,0}) > \underbrace{t(x, p_{1,0}, p_{2,0}, p_{3,0}, p_{4,0})}_{19.60} \text{ if } H_0 \text{ is true}\big)$$

from the table we can see that to get an upper tailed area 19.60 with degrees of freedom as 3, the p-value has to be $\in (0, 0.001)$.

$\because$ p-value $\in (0, 0.001)$ is less than the level of significance $(\alpha = 0.01)$, we need to reject the null hypothesis and conclude that there is atleast one value of the proportion that is different from the claimed proportion. Ans

(9.)

There are 4 groups in which admission dates of patients are categorized.

$n = 200$ patients.

1st group consists of 11, second group 24, third group 69 and 4th group 96 patients.

To investigate whether there is any relationship between the patient's admission for alcoholism and patient's birthday, we need to perform chi-squared test ~~for independence~~

The null & alternative ~~hypothesis~~ are as follows:-

H_0: There is no relationship between patient's admissin for alcoholism and patient's birthday.

H_a: There is a relationship between patient's admission for alcoholism and patient's birthday.

Next, we need to calculate the test-statistic which is computed as follows:-

$$\cancel{T(x,y) = \sum_{k=1}^{k}\sum_{l=1}^{l}\frac{(N_{kl}-E_{kl})^2}{E_{kl}}} \quad \cancel{\text{where } E_{kl} = \frac{N_{k,\cdot}N_{\cdot,l}}{n}}$$

$$T(X, p_{1,0}, \ldots, p_{k,0}) = \sum_{k=1}^{k}\left(\frac{(N_k - np_{k,0})^2}{np_{k,0}}\right)$$

$$\Rightarrow t(n, p_{1,0}, p_{2,0}, p_{3,0}, p_{4,0}) = \sum_{k=1}^{4}\left(\frac{(N_k - np_{k,0})^2}{np_{k,0}}\right)$$

first, we will need to calculate the expected frequencies according to the birthday groups formed.

for the first group,

$$\text{Expected frequency } (np_{1,0}) \Rightarrow 200 \times \frac{14}{365}$$

$$= 7.6712$$

2nd group,

$$\text{Expected frequency } (np_{2,0}) \Rightarrow 200 \times \frac{46}{365}$$

$$= 25.2055$$

3rd group,

$$\text{Expected frequency } (np_{3,0}) \Rightarrow 200 \times \frac{120}{365}$$

$$\Rightarrow 65.7534$$

4th group,

$$\text{Expected frequency } (np_{4,0}) \Rightarrow 200 \times \frac{185}{365}$$

$$\Rightarrow 101.3699$$

So, to calculate test-statistic,

$$\Rightarrow \left(\frac{(11 - 7.6712)^2}{7.6712}\right) + \left(\frac{(24 - 25.2055)^2}{25.2055}\right) + \left(\frac{(69 - 65.7534)^2}{65.7534}\right) + \left(\frac{(96 - 101.3699)^2}{101.3699}\right)$$

$$\Rightarrow 1.4445 + 0.0577 + 0.1603 + 0.2845$$

$$\Rightarrow \underline{1.947} \Rightarrow \underline{\underline{1.95}}$$

The degrees of freedom can be calculated as $\Rightarrow$

$df = k-1 = 4-1 = \underline{\underline{3}}$.

Now, we need to calculate the p-value and look into the table of χ^2 distribution with degrees of freedom = 3.

$$p\text{-value} = P\left(T\left(X, P_{1,0}, P_{2,0}, P_{3,0}, P_{4,0}\right)\right) \text{ is more extreme}$$
$$\text{than } t\left(x, P_{1,0}, P_{2,0}, P_{3,0}, P_{4,0}\right) \text{ if } H_0 \text{ is true}).$$

$$\Rightarrow P\left(T\left(X, P_{1,0}, P_{2,0}, P_{3,0}, P_{4,0}\right) > t\left(x, P_{1,0}, P_{2,0}, P_{3,0}, P_{4,0}\right) \text{ if } H_0 \text{ is true}\right)$$

from the table, we can see that to get an upper-tailed area of 1.95 with degrees of freedom = 3, the p-value has to be greater than 0.1.

Since, p-value is greater than the level of significance (0.01), we cannot reject the null Hypothesis.

Conclusively, there is no sufficient evidence to conclude that there is a relationship between patient's admission for alcoholism and patient's birthday at 1% significance level.

(10) Level of significance (α) = 0.01.

The claim is to test whether there is an association between extent of binge drinking and age group. If the claim is rejected, then there is no association between extent of binge drinking and age group.

So, the null hypothesis is:-

H_0: Binge drinking and age group are independent.

Alternative hypothesis is:-

H_a: Binge drinking and age group are dependent.

# Episodes (X)	Age Group (Y) (A1) 18-20	(A2) 21-23	(A3) ≥24	Total
(E1) None	357	293	592	1242
(E2) 1-2	218	285	354	857
(E3) 3-4	184	218	185	587
(E4) ≥5	328	331	147	806
Total	1087	1127	1278	3492

Now, to calculate the value of the test-statistic:-

$$E_{E1,A1} = \frac{1087 \times 1242}{3492} = 386.61$$

$$E_{E1,A2} = \frac{1127 \times 1242}{3492} = 400.84$$

$$E_{E1,A3} = \frac{1278 \times 1242}{3492} = 454.55$$

$$E_{E2,A1} = \frac{1087 \times 857}{3492} = 266.77$$

$$E_{E2,A2} = \frac{1127 \times 857}{3492} = 276.59$$

$$E_{E2,A3} = \frac{1278 \times 857}{3492} = 313.64$$

$$E_{E3,A1} = \frac{1087 \times 587}{3492} = 182.72$$

$$E_{E3,A2} = \frac{1127 \times 587}{3492} = 189.45$$

$$E_{E3,A3} = \frac{1278 \times 587}{3492} = 214.83$$

$$E_{E4,A1} = \frac{1087 \times 806}{3492} = 250.89$$

$$E_{E4,A2} = \frac{1127 \times 806}{3492} = 260.13$$

$$E_{E4,A3} = \frac{1278 \times 806}{3492} = 294.98$$

$$T(X,Y) \Rightarrow \sum_{k=1}^{k} \sum_{l=1}^{L} \left(\frac{(N_{kl} - E_{kl})^2}{E_{kl}} \right)$$

$$\Rightarrow \frac{(357 - 386.61)^2}{386.61} + \frac{(293 - 400.84)^2}{400.84} + \frac{(592 - 454.55)^2}{454.55} +$$

$$\frac{(218 - 266.77)^2}{266.77} + \frac{(285 - 276.59)^2}{27659} + \frac{(354 - 313.64)^2}{313.64} +$$

$$\frac{(184-182.72)^2}{182.72}+\frac{(218-189.45)^2}{189.45}+\frac{(185-214.83)^2}{214.83}+$$

$$\frac{(328-250.89)^2}{250.89}+\frac{(331-260.13)^2}{260.13}+\frac{(147-294.98)^2}{294.98}$$

$\Rightarrow 2.267 + 29.013 + 41.563 + 8.916 + 0.256 + 5.914 + 8.967\times10^{-3}$

$+ 4.302 + 4.142 + 23.700 + 19.308 + 74.236$

$\Rightarrow 213.6259 \simeq \boxed{213.63}$

The degrees of freedom can be computed as $\Rightarrow$

$$df = (r-1)(c-1)$$
$$= (4-1)(3-1) = 3\times2 = \underline{\underline{6}}$$

Ans.

Considering p-value for the same $\Rightarrow$

p-value $\Rightarrow P(T(X,Y)$ more extreme than $t(x,y)$ if H_0 is true)

$= P(T(X,Y) > t(x,y)$ if H_0 is true)

$= P(T(X,Y) > 213.63)$ if H_0 is true.

This follows a $\chi^2_{(K-1)(L-1)} = \boxed{\chi^2_6}$ distribution.

From the table of chi-squared curved tail areas, we get the p-value to be $\Rightarrow \boxed{<0.001}$.

Thus, p-value is less than the level of significance (0.01). $\boxed{\text{p-value} \le \alpha}$ $\rightarrow$, $\therefore$ we can reject the null hypothesis.

Therefore, there is sufficient evidence to conclude that there is an association between the extent of binge drinking & age group. Therefore, they are dependent. (associated).

Ans

11. The data on number of failures in each of the three configurations modes is given.

The hypothesis can be given as:-

Null hypothesis → H_0: The configurations does not have an effect on the type of failures

Alternative hypothesis:-

H_a: The configurations has an effect on the type of failures. (are dependent)

Failure Mode

Configurations	1	2	3	4	Total
1	20	44	17	9	90
2	4	17	7	12	40
3	10	31	14	5	60
Total	34	92	38	26	190

Now, to calculating the value of the test-statistic:-

$$E_{1,1} = \frac{34 \times 90}{190} = 16.105$$

$$E_{1,2} = \frac{92 \times 90}{190} = 43.579$$

$$E_{1,3} = \frac{38 \times 90}{190} = 18$$

$$E_{1,4} = \frac{26 \times 90}{190} = 12.316$$

$$E_{2,1} = \frac{34 \times 40}{190} = 7.158$$

$$E_{2,2} = \frac{92 \times 40}{190} = 19.368$$

$$E_{2,3} = \frac{38 \times 40}{190} = 8$$

$$E_{2,4} = \frac{26 \times 40}{190} = 5.474$$

$$E_{3,1} = \frac{34 \times 60}{190} = \cancel{10.7347} \; 10.737$$

$$E_{3,2} = \frac{92 \times 60}{190} = 29.053$$

$$E_{3,3} = \frac{38 \times 60}{190} = 12$$

$$E_{3,4} = \frac{26 \times 60}{190} = 8.211$$

Now, we will calculate the test statistic with the E_{kl} values above,

$$T(X,Y) \Rightarrow \sum_{k=1}^{K} \sum_{l=1}^{L} \left(\frac{(N_{kl} - E_{kl})^2}{E_{kl}} \right)$$

$$\Rightarrow \frac{(20-16.105)^2}{16.105} + \frac{(44-43.579)^2}{43.579} + \frac{(17-18)^2}{18} + \frac{(9-12.316)^2}{12.316}$$

$$+ \frac{(4-7.158)^2}{7.158} + \frac{(17-19.368)^2}{19.368} + \frac{(7-8)^2}{8} + \frac{(12-5.474)^2}{5.474}$$

$$+ \frac{(10-10.737)^2}{10.737} + \frac{(31-29.053)^2}{29.053} + \frac{(14-12)^2}{12} + \frac{(5-8.211)^2}{8.211}$$

$$\Rightarrow 0.942 + 4.067 \times 10^{-3} + 0.056 + 0.893 + 1.393 +$$

$$0.290 + 0.125 + 7.780 + 0.051 + 0.131$$

$$+ 0.333 + 1.256$$

$$= 13.254 \approx \underline{\underline{13.25}}$$

The degrees of freedom can be computed as $\Rightarrow$

$$df = (3-1)(4-1) = \underline{\underline{6}}$$

Considering p-value $\Rightarrow$

p-value $\Rightarrow$ $P(T(X,Y)$ more extreme than $t(x,y)$ if H_0 is ~~true~~ true$)$

$$\Rightarrow P(T(X,Y) > t(x,y) \text{ if } H_0 \text{ is true})$$

$$\Rightarrow P(T(X,Y) > 13.25 \text{ if } H_0 \text{ is true})$$

Since, the test statistic follows a χ^2_6 distribution.

From the table of chi-squared curved tail areas, we get the p-value to be $\Rightarrow$ $\boxed{p \in (0.035, 0.04)}$
(for test statistic computed)

Thus, p-value is greater than the level of significance (0.01),

~~p-value ≥ α~~ $\boxed{p\text{-value} > \alpha}$

$\therefore$ we do not reject the null-hypothesis.

Thus, there is not sufficient evidence to conclude that the configuration is dependent (or associated) with the failure mode/ type of failure.

Ans