Student Information

Full Name : Mithat Can Timurcan
Id Number : 2581064

Answer 1

a)

- In order to obtain a 95% confidence interval for our normally distributed sample, we need to use the according z value $z_{\alpha/2}$. Since $1-\alpha$ is given as 0.95, α becomes 0.05 and our according z value becomes $z_{0.025}$ which is 1.96.

- While obtaining a confidence interval we use the following formula:

$$\overline{X} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

- We can find our mean value $\overline{X}$ by doing the following:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{6.4 + 9.5 + 8.2 + 10.2 + 7.6 + 11.1 + 8.7 + 7.3 + 9.1}{9} \approx 8.68$$

- Let's put our values, $\overline{X} \approx 8.68$, $n = 9$, $\sigma = 2.7$, $z_{\alpha/2} = z_{0.025} = 1.96$:

$$\overline{X} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = (8.68) \pm (1.96) \cdot \frac{(2.7)}{\sqrt{9}} = [6.91, 10.44]$$

b)

- In order to find the size that we need to obtain a specific margin, we use the following:

$$\Delta = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \ \rightarrow \ n \geq \left(\frac{z_{\alpha/2} \cdot \sigma}{\Delta}\right)^2$$

- Again, putting our values $\sigma = 2.7$, $z_{\alpha/2} = z_{0.025} = 1.96$ and $\Delta = 1.25$ we get:

$$n \geq \left(\frac{z_{\alpha/2} \cdot \sigma}{\Delta}\right)^2 = \left(\frac{1.96 \cdot 2.7}{1.25}\right)^2 = 17.92 \approx 18$$

- Therefore, our sample size should be more than 18 so that we can get a margin maximum of 1.25.

Answer 2

a)

- Since we want to test if their average revenue has increased or not, we can set up the hypotheses as follows:

$$H_0 : \mu = 20000$$
$$H_A : \mu > 20000$$

- In this case, Mecnun's claim will be the null hypothesis H_0 and Leyla's claim will be the alternative hypothesis H_A.

b)

- In order to test our statistics with a 5% level of significance, we'll use a right-tail z-test with the according z value:

$$\alpha = 0.05 \;\rightarrow\; z_\alpha = z_{0.05} = 1.645$$
$$\begin{cases} \text{accept } H_0 \quad \text{if} \quad Z < 1.645 \\ \text{reject } H_0 \quad \text{if} \quad Z \geq 1.645 \end{cases}$$

- Now we need to compute our Z value and check if it's in the accepted region:

$$Z = \frac{\overline{X} - \mu_0}{\sigma \,/\sqrt{n}} = \frac{22000 - 20000}{3000 \,/\sqrt{50}} \approx 4.71$$

- Since $Z = 4.71$ is in the rejected region, we will reject the null hypothesis. The data provided supports the conclusion that there is a significant increase in the average amount of revenue.

c)

- Since we've found our observed Z value, we can compute the p-value for Mecnun's test by using `normcdf(Z)`:

$$P = \mathbf{P}\{Z \geq Z_{obs}\} = \mathbf{P}\{Z \geq 4.71\} = 1 - \Phi(4.71) \approx 0.000$$

- Since we got an extremely low p-value, we have grounds to reject the null hypothesis not just at the 5% significance level asked in the question, but also at the 1% and even the 0.05% levels.

d)

- In order to test if their average revenue is higher than the competitor's value, we should set up the hypotheses as follows:

$$H_0 : \mu_X - \mu_Y = 0$$
$$H_A : \mu_X - \mu_Y > 0$$

- Then, we should find the acceptance region for the right-tail z-test with 1% significance:

$$\alpha = 0.01 \;\rightarrow\; z_\alpha = z_{0.01} = 2.32$$
$$\begin{cases} \text{accept } H_0 & \text{if} \quad Z < 2.32 \\ \text{reject } H_0 & \text{if} \quad Z \geq 2.32 \end{cases}$$

- Let's compute our Z value and check if it's in the accepted region:

$$Z = \frac{\overline{X} - \overline{Y}}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}}} = \frac{22000 - 24000}{\sqrt{\dfrac{(3000)^2}{50} + \dfrac{(4000)^2}{40}}} \approx -2.62$$

- Since $Z = -2.62$ is not in the rejected region, we can't reject the null hypothesis. Therefore with %1 significance, we can't claim that Leyla and Mecnun's average revenue is higher than the competitor's average revenue.

Answer 3

- We are given the observed table, let us put the values:

$Obs(i,j) = n_{ij}$	Black Coffee	Coffee with Milk	Coffee with Sugar	$n_{i.}$
Male	52	16	32	100
Female	17	63	20	100
$n_{.j}$	69	79	52	200

- We are going to test the following hypotheses:

H_0 : Gender and choice of coffee are independent factors.

H_A : Gender and choice of coffee are dependent factors.

- Now let us compute the estimated expected values:

$$\widehat{Exp}(i,j) = \frac{n_{i.} \cdot n_{.j}}{n}$$

$$\widehat{Exp}(1,1) = \frac{69 \cdot 100}{200} = 34.5 \qquad \widehat{Exp}(2,1) = \frac{69 \cdot 100}{200} = 34.5$$

$$\widehat{Exp}(1,2) = \frac{79 \cdot 100}{200} = 39.5 \qquad \widehat{Exp}(2,2) = \frac{79 \cdot 100}{200} = 39.5$$

$$\widehat{Exp}(1,3) = \frac{52 \cdot 100}{200} = 26 \qquad \widehat{Exp}(2,3) = \frac{52 \cdot 100}{200} = 26$$

$\widehat{Exp}(i,j) = \frac{n_{i.} \cdot n_{.j}}{n}$	Black Coffee	Coffee with Milk	Coffee with Sugar	$n_{i.}$
Male	34.5	39.5	26	100
Female	34.5	39.5	26	100
$n_{.j}$	69	79	52	200

- We can compute our chi-square value:

$$\chi_{obs}{}^2 = \sum_{i=1}^{k} \sum_{j=1}^{m} \frac{\{Obs(i,j) - \widehat{Exp}(i,j)\}^2}{\widehat{Exp}(i,j)}$$

$$\chi_{obs}{}^2 = \frac{(52-34.5)^2}{34.5} + \frac{(17-34.5)^2}{34.5} + \frac{(16-39.5)^2}{39.5} + \frac{(63-39.5)^2}{39.5} + \frac{(32-26)^2}{26} + \frac{(20-26)^2}{26} \approx 48.5$$

- From Table A6, with the degrees of freedom $(3-1)(2-1) = 2$, we find that the p-value $P < 0.001$ and therefore, we reject the null hypothesis. We have a significant evidence that gender is related to the coffee choice.