Student Information

Full Name : Feyza Özlü
Id Number : e2580868

Answer 1

a)

For this question since we are given that measurement errors follow a normal distribution and the measurement device has a guaranteed standard deviation of $\sigma = 2.7$ we can directly use the formula of confidence interval for the mean (σ is known) on textbook page-251, which is:

$$\bar{X} \pm Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

Now, let's first calculate the sample mean:

$$\bar{X} = \frac{6.4 + 9.5 + 8.2 + 10.2 + 7.6 + 11.1 + 8.7 + 7.3 + 9.1}{9} \approx 8.68$$

Also, we are asking 95% confidence interval. Which means our α value is equal to $1 - 95\% = 5\% = 0.05$

From textbook page-250 (9.4) we know that

$$Z_{\alpha/2} = Z_{0.05/2} = Z_{0.025} = 1.960$$

Now, the only thing we need is substituting these values into the formula:

$$\begin{aligned}\bar{X} \pm Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) &= 8.68 \pm 1.960\left(\frac{2.7}{\sqrt{9}}\right) \\ &= 8.68 \pm 1.768 \\ &= [6.916, 10.444]\end{aligned}$$

Hence 95% confidence interval for the population mean is [6.916, 10.444]

b)

For this part we can use the formula of sample size for a given precision in textbook page-254, which is:

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

We are given that margin of maximum 1.25 units with same confidence level which means our $\Delta = 1.25$ and confidence level = 95% again. We already know $Z_{\alpha/2} = 1.960$ from part-a. If we substitute these values into the equation, we get:

$$n \geq \left(\frac{Z_{\alpha/2} \cdot \sigma}{\Delta}\right)^2 = \left(\frac{1.960 \cdot 2.7}{1.25}\right)^2$$
$$= 17.923369$$

Since, sample size has to be an integer we can say that n=18 sample size is required to estimate the population mean with a margin of maximum 1.25 units with same confidence level.

Answer 2

a)

In this question we are given, Leyla claims that the revenue has increased compared to last year, while Mecnun claims that it has remained the same. Since the null hypothesis should be always an equality, the null hypothesis (H_0) should be that the average monthly revenue has remained the same as last year (20,000 TL), meanly Mecnun's claim is the null hypothesis. Moreover, the alternative hypothesis (H_A) should be that the average monthly revenue has increased compared to last year, meanly Leyla's claim is the alternative hypothesis.

$$H_0 : \mu = 20,000 \text{ TL}$$

$$H_A : \mu > 20,000 \text{ TL}$$

b)

For this part we need to make one sided right tail Z-test. First, let's find Z_α value and acceptance region. We are asking 5% significance level. Which means our α value is equal to 5% = 0.05 and our $Z_\alpha = Z_{0.05} = 1.645$ from textbook page-250 (9.4). Since we are conducting one sided right tail Z-test our acceptance region is $(-\infty, 1,645]$. Also, we are given that $\sigma = 3000$, $n = 50$, $\mu_0 = 20000$, $\bar{X} = 22000$.

Now, let's compute the test statistic by using the formula in the textbook page-271 (9.17) which is:

$$Z = \frac{\hat{\theta} - E(\hat{\theta})}{\sqrt{Var(\hat{\theta})}}$$

If we substitute the values inside the formula:

$$Z = \frac{\hat{\theta} - E(\hat{\theta})}{\sqrt{Var(\hat{\theta})}}$$

$$= \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

$$= \frac{22000 - 20000}{3000/\sqrt{50}}$$

$$= \frac{2000 \cdot \sqrt{50}}{3000}$$

$$\approx 4.714$$

Since the calculated z-value is much higher than 1.645, it doesn't belong to the acceptance region. Hence we can reject the null hypothesis. Therefore, Mecnun cannot claim that the average monthly revenue has remained the same compared to last year.

c)

We calculated the Z-value = 4.714 in part-b, which means $Z_{obs} = 4.714$. We can calculate the p-value by using the formula in the table 9.3 in the textbook page-283, since it was a one sided right tail Z-test our formula is:

$$p = P\{Z \geq Z_{obs}\} = 1 - \phi(Z_{obs})$$

If we use octave online to calculate $\phi(Z_{obs}) = \phi(4.714)$ by using *normcdf*(4.714) we see that $\phi(Z_{obs}) = 1.0000$.
So, our p value is

$$p = P\{Z \geq Z_{obs}\} = 1 - \phi(Z_{obs}) = 1 - 1 = 0.$$

which is very low (lower than the 0.01). Therefore, we can reject the null hypothesis at the 5% level of significance. We see now that the average monthly revenue didn't remained the same compared to last year.

d)

We are given that in Leyla and Mecnun's store:

- sample mean(this year) = 22000 TL
- standard deviation = 3000 TL

- sample size = 50

and in competitor's store:

- sample mean(this year) = 24000 TL
- standard deviation = 4000 TL
- sample size = 40

and we are asking 1% significance level. Which means our α value is equal to $1\% = 0.01$ and our $Z_\alpha = Z_{0.01} = 2.326$ from textbook page-250 (9.4).

To compare Leyla and Mecnun's store's current average monthly revenue with that of the competitor's store, we should conduct one sided right tail Z-test for two samples. Since Leyla and Mecnun claim that their store's average monthly revenue is higher than that of the competitor's store:

$$H_0 : \mu_{LeylaMecnun} = \mu_{competitor}$$

$$H_A : \mu_{LeylaMecnun} > \mu_{competitor}$$

First, let's calculate the Var($\hat{\theta}$) by using the formula in the table 9.1 in the textbook page-273 which is:

$$Var(\hat{\theta}) = \frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}$$

If we substitute the values:

$$\begin{aligned} Var(\hat{\theta}) &= \frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y} \\ &= \frac{\sigma_{LeylaMecnun}^2}{n_{LeylaMecnun}} + \frac{\sigma_{competitor}^2}{n_{competitor}} \\ &= \frac{3000^2}{50} + \frac{4000^2}{40} \\ &= 180000 + 400000 \\ &= 580000 \end{aligned}$$

Now, let's calculate the Z-value by using the formula in the table 9.1 in the textbook page-273 which is:

$$Z = \frac{\hat{X} - \hat{Y} - D}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}}$$

Since our null hypothesis is $\mu_{LeylaMecnun} = \mu_{competitor}$

$$D = \mu_{LeylaMecnun} - \mu_{competitor} = 0$$

If we substitute the values:

$$\begin{aligned} Z &= \frac{\hat{X} - \hat{Y} - D}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} \\ &= \frac{22000 - 24000 - 0}{\sqrt{580000}} \\ &= \frac{-2000}{\sqrt{580000}} \\ &\approx -2.626 \end{aligned}$$

Since $Z = -2.626 \leq Z_{0.01} = 2.326$ We can accept the null hypothesis. This means that Leyla and Mecnun can not claim that the average monthly revenue of their store is now higher than that of the competitor's store at a 1% level of significance.

Answer 3

For this question we are going to use Chi-square test for testing independence:

$$H_0 : \text{coffee preferences and gender are independent}$$

$$H_A : \text{coffee preferences and gender are dependent}$$

We are given that:

	Black Coffee	Coffee with Milk	Coffee with Sugar
Male Respondents	52	16	32
Female Respondents	17	63	20

Table 1: Observed preferences for coffee type according to gender

First let's calculate the expected preferences n_{ij} for each entry in the table by using the formula in the slides $n_{ij} = n\frac{n_{i.}}{n}\frac{n_{.j}}{n}$

	Black Coffee	Coffee with Milk	Coffee with Sugar
Male Respondents	$200 \cdot \frac{100}{200} \cdot \frac{69}{200} = 34.5$	$200 \cdot \frac{100}{200} \cdot \frac{79}{200} = 39.5$	$200 \cdot \frac{100}{200} \cdot \frac{52}{200} = 26$
Female Respondents	$200 \cdot \frac{100}{200} \cdot \frac{69}{200} = 34.5$	$200 \cdot \frac{100}{200} \cdot \frac{79}{200} = 39.5$	$200 \cdot \frac{100}{200} \cdot \frac{52}{200} = 26$

Table 2: Expected preferences for coffee type according to gender

Since all values are greater than 5 we don't need to combine the rows we can directly calculate

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

$$\begin{aligned}
X_{obs}^2 &= \frac{(52-34.5)^2}{34.5} + \frac{(16-39.5)^2}{39.5} + \frac{(32-26)^2}{26} + \frac{(17-34.5)^2}{34.5} + \frac{(63-39.5)^2}{39.5} + \frac{(20-26)^2}{26} \\
&= 8.87681159 + 13.9810127 + 1.38461538 + 8.87681159 + 13.9810127 + 1.38461538 \\
&= 48.48487934 \\
&\approx 48.49
\end{aligned}$$

And our degree of freedom is $\nu = (k-1)\cdot(m-1) = (3-1)\cdot(2-1) = 2$ where k is te number of columns and m is the number of rows.
If we look at the Table A6 Table of Chi-Square Distribution in the textbook we saw that our p-value is smaller than 0.001. Hence we have significant evidence to reject null hypothesis. So we can conclude that there is a significant association between gender and coffee preference.