Student Information

Full Name : Musa Alper Yaman
Id Number : 2581155

Answer 1

a)

We can calculate confidence interval for population mean by using the formula:

$$\text{Confidence interval} = \bar{x} \pm z_{\frac{\alpha}{2}}(\frac{\sigma}{\sqrt{n}})$$

Where, $\bar{x}$ = Sample mean, n = Sample size, σ = Standard derivation, z = Critical value.
To calculate sample mean we can use:

$$\bar{x} = \frac{(x_1 + x_2 + \cdots + x_n)}{n}$$
$$\bar{x} = \frac{6.4 + 9.5 + 8.2 + 10.2 + 7.6 + 11.1 + 8.7 + 7.3 + 9.1}{9}$$
$$\bar{x} \cong 8.678$$

To calculate α and Critical value:

$$(1-\alpha) = 0.95 \implies \alpha = 0.05 \implies \frac{\alpha}{2} = 0.025$$
$$q_{0.025} = -z_{0.025} \cong -1.960$$
$$q_{0.975} = z_{0.025} \cong 1.960$$

It is given that standard derivation $\sigma = 2.7$. Calculate the interval by inserting the values:

$$\text{Confidence interval} = \left(8.678 - 1.960\left(\frac{2.7}{9}\right), 8.678 + 1.960\left(\frac{2.7}{9}\right)\right)$$
$$\cong (6.914, 10.442)$$

b)

We can use the following formula to find minimum value of sample size:

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

From the first part we know, $z = 1.96$, in the question given $\sigma = 2.7, \Delta = 1.25$, insert the values to find n:

$$n \geq \left(\frac{1.96 \cdot 2.7}{1.25}\right)^2$$
$$n = \left(\frac{5.292}{1.25}\right)^2$$
$$n = (4.233)^2$$
$$n \geq 17.923$$

We can say that n is minimum **18**.

Answer 2

a)

Null hypothesis (H_0) = The average monthly revenue of the store is the same as last year ($\mu = 20000$)
Alternative hypothesis (H_A) = The average monthly revenue of the store have changed from the last year ($\mu \neq 20000$)
As a result Mecnun's claim is null hypothesis.

b)

Population standard derivation is given and sample size is large (n = 50), so we can use z-test:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Insert the given values, sample mean $\bar{x} = 22000$, last years average $\mu = 20000$, population standard derivation $\sigma = 3000$, sample size n = 50:

$$z = \frac{22000 - 20000}{\frac{3000}{\sqrt{50}}} = \frac{2000}{\frac{600}{\sqrt{2}}} = \frac{2000}{300\sqrt{2}} \cong 4.714$$

By using Standard normal distribution table, we found that critical z-value for %5 level of significance is ± 1.96 since calculated z value (4.713) is greater than critical z-value (1.96), we can reject null hypothesis and claim that the average monthly revenue has changed compared to the last year.

c)

From the previous parts, we found that $z \cong 4.714$. We can calculate the p value by using the formula:

$$P = P\{Z \geq Z_{obs}\}$$
$$P = P\{Z \geq Z_{4.713}\}$$
$$P = 1 - \phi(4.713)$$

We can use Octave online to calculate $1 - \phi(4.713)$:

```
octave:2> 1-normcdf(4.731)
ans = 1.1171e-06
```

As we calculated, P is almost 0. So there is almost no probability that observing a z is greater than 4.71. We can say that the average monthly revenue has increased compared to the last year.

d)

We can conduct a hypothesis test:

- Null Hypothesis (H_0) = The average revenue of Leyla and Mecnun's store is less than or equal to the competitor's store.
- Alternative Hypothesis (H_0) = The average revenue of Leyla and Mecnun's store is higher than the competitor store.

Parameters are:

- Leyla and Mecnun's store's this year average revenue: $\bar{x_1} = 22000$
- Leyla and Mecnun's store's this year standard derivation: $\sigma_1 = 3000$
- Competitor's store's average revenue: $\bar{x_2} = 24000$
- Competitor's store's standard derivation: $\sigma_2 = 4000$
- Leyla and Mecnun's store sample size: $n_1 = 50$
- Competitor's store sample size: $n_2 = 40$
- level of significance: $\alpha = 0.01$

We will use z-test to compare means of them:

$$z = \frac{(\bar{x_1} - \bar{x_2})}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$
$$z = \frac{(22000 - 24000)}{\sqrt{\frac{3000^2}{50} + \frac{4000^2}{40}}}$$
$$z = \frac{-2000}{\sqrt{\frac{9000000}{50} + \frac{16000000^2}{40}}}$$
$$z = \frac{-2000}{\sqrt{180000 + 400000}}$$
$$z = \frac{-2000}{761.57}$$
$$z \cong -2.626$$

For $\alpha = 0.01$ the z value is equals to 2.33.

$$z = -2.63 < 2.33$$

Since calculated z value is smaller than the z value for $\alpha = 0.01$, we accept the null hypothesis. As a result, we cannot claim that average monthly revenue of Leyla and Mecnun's store is higher than competitor's store's average revenue.

Answer 3

To determine if there is a significant association between gender and coffee preference, we can use Chi-square test of independence.
The hypothesis for the chi-square test is:

- Null Hypothesis (H_0) = There is no association between gender and coffee preference
- Alternative Hypothesis (H_0) = There is an association between gender and coffee preference

Then, we should calculate Expected Value:

$$\text{Expected Value} = \frac{\text{Row Total} \cdot \text{Column Total}}{\text{Overall Total}}$$

By using Expected Value formula we cant construct Expected frequency table:

	Black Coffee	**Coffee with Milk**	**Coffee with Sugar**	**Total**
Male	52	16	32	100
Female	17	63	20	100
Total	69	79	52	200

Table 1: Contingency table showing the relationship between gender and coffee preference.

	Black Coffee	**Coffee with Milk**	**Coffee with Sugar**	**Total**
Male	34.5	39.5	26	100
Female	34.5	39.5	26	100
Total	69	79	52	200

Table 2: Expected Frequency Table

Now, we will use Chi-square test, where O = Observed Frequency, E = Expected Frequency:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

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

$$\chi^2 = \frac{306.25}{34.5} + \frac{552.25}{39.5} + \frac{36}{26} + \frac{306.25}{34.5} + \frac{552.25}{39.5} + \frac{36}{26}$$

$$\chi \cong 6.96$$

Calculate df = (number of rows - 1) $\cdot$ (number of columns - 1):

$$df = (3-1) \cdot (2-1) = 2$$

For $df = 2$ and $\alpha = 2$, the Chi-square value is approximately 5.991 (By chi-square distribution table).
As a result, $\chi > 6.991$, so we can reject null hypothesis. Thus, there is a significant association between gender and coffee preference.