Student Information

Full Name : BAŞAR YILMAZ

Id Number : 2644409

Answer 1

a)

We have $n = 16 \leq 30$, so it is a small sample size. Furthermore, the standard deviation of the population is not given. That means we need to use $t - test$ in this question. First, we derive the statistical data from the given sample.

$$\bar{x} \approx 6.81$$

$$s \approx 1.06$$

$$n = 16$$

We need to construct a 98% confidence interval so $\alpha = 0.02$ and $t_{\alpha/2} = t_{0.01} \approx 2.60$ (from t table).

We need to find:

$$\bar{x} \pm t_{0.01} * \frac{s}{\sqrt{n}} = 6.81 \pm 2.6 * \frac{1.06}{\sqrt{16}} = 6.81 \pm 0.69$$

Our confidence interval:

$$[6.12, 7.50]$$

b)

Null hypothesis and alternative hypothesis:

$$H_0 : \mu = 7.5$$

$$H_A : \mu < 7.5$$

Since our significance level is 5%:

$$\alpha = 0.05$$

We need to conduct a left-tail t-test with $\alpha = 0.05$ and degree of freedom $= n - 1 = 15$.

$$t_{0.05} \approx -1.75 \text{ (from t table)}$$

Now we need to derive t_{test} from our sample values.

$$t_{test} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{6.81 - 7.5}{\frac{1.06}{\sqrt{16}}} = -2.6$$

We found that:

$$t_{test} < t_{0.05}$$

So, we can reject the null hypothesis, and it can be claimed that improvement is effective with 5% significance level.

c)

We cannot reject H_0 without any calculations as 6.5 liters is in the interval we found in part a) with 2% confidence.

Answer 2

a)

$$H_0 : \mu = 5000$$

$$H_A : \mu > 5000$$

The null hypothesis is Ali's claim.

b)

We can conduct a right-tail $z - test$ with $\alpha = 0.05$.

It can be found on the z-table:

$$z_{0.05} = 1.645$$

Now we need to derive z_{test} from our sample values.

$$z_{test} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{5500 - 5000}{\frac{2000}{\sqrt{100}}} = 2.5$$

We found that:

$$z_{test} > z_{0.05}$$

So, we can reject the null hypothesis and claim that, with a 5% significance level, there is an increase in the rent prices compared to last year.

c)

In order the find $p - value$, we need to find:

$$p = P(z > 2.5) = 1 - P(z < 2.5) = 0.0062 \approx 0.01$$

This means the probability of observing a sample mean of 5500 or higher is very low with 1% probability. Therefore, we have sufficient evidence to support Ahmet's claim.

d)

	$\bar{x}$ (Mean)	s (Standard Deviation)	n (Sample Size)
İstanbul	6500 TL	3000 TL	60
Ankara	5500 TL	2000 TL	100

To determine if the prices in Ankara are lower than those in İstanbul, we can perform a two-sample z-test.

We have the following information:

Ankara sample: $n_1 = 100$, $\bar{x}_1 = 5500\text{TL}$, $s_1 = 2000\text{TL}$

İstanbul sample: $n_2 = 60$, $\bar{x}_2 = 6500\text{TL}$, $s_2 = 3000\text{TL}$

The null hypothesis and alternative hypothesis:

$$H_0 : \bar{x}_1 = \bar{x}_2$$

$$H_A : \bar{x}_1 < \bar{x}_2$$

The formula for the two-sample z-test is:

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Calculating the z-value:

$$z = \frac{5500 - 6500}{\sqrt{\frac{2000^2}{100} + \frac{3000^2}{60}}} = -\frac{1000}{\sqrt{190000}} \approx -\frac{1000}{435.89} \approx -2.29$$

Since the question asks for 1% significance level $\alpha = 0.01$. We should check z-table to find $z_{0.01} \approx -2.33$.

We fail to reject the null hypothesis since it is left-tail $z\!-\!test$ and the $z = -2.29 > z_{0.01} = -2.33$. This means that at a 1% level of significance, there is insufficient evidence to claim that the prices in Ankara are significantly lower than in İstanbul.

Answer 3

First, our null and alternative hypothesis are:

$$H_o: \ rainy\ days\ and\ seasons\ are\ independent$$

$$H_A: \ rainy\ days\ and\ seasons\ are\ dependent$$

We can calculate the Expected Table for each data.

$$E_{1,1} = \frac{90*100}{360} = 25 \ E_{1,2} = \frac{90*100}{360} = 25... \ E_{2,1} = \frac{90*260}{360} = 65 \ E_{2,2} = \frac{90*260}{360} = 65...$$

Finally, our expected table will be like that:

	Winter	Spring	Summer	Autumn	Total
Rainy	25	25	25	25	100
Non-Rainy	65	65	65	65	260
Total	90	90	90	90	360

Now from the data in the expected table, we can calculate our chi-square test statistic:

$$\begin{aligned} X_c^2 = & \frac{(34-25)^2}{25} + \frac{(32-25)^2}{25} + \frac{(15-25)^2}{25} + \frac{(19-25)^2}{25} \\ & + \frac{(56-65)^2}{65} + \frac{(58-65)^2}{65} + \frac{(75-65)^2}{65} + \frac{(71-65)^2}{65} \\ & = 3.24 + 1.96 + 4 + 1.44 + 1.25 + 0.75 + 1.54 + 0.55 \\ & = 14.73 \end{aligned}$$

Chi-square Test Statistic: The calculated chi-square test statistic is 14.73.

Degrees of Freedom: Degrees of freedom (df) = (rows - 1) * (columns - 1) = 3.

P-value: The p-value for the chi-square test statistic of 14.73 with 3 degrees of freedom is approximately 0.002. I used the function $1 - chi2cdf(14.73, 3)$ in Octave.

Based on the p-value of 0.002, which is smaller than the significance level of 0.05, we can reject the null hypothesis. Thus, sufficient evidence suggests a statistically significant association between the variables "rainy days" and "season" in Ankara. Therefore, the number of rainy days in Ankara depends on the season.

Answer 4

```
observed = [34 32 15 19; 56 58 75 71];
expected = [25 25 25 25; 65 65 65 65];

chi2obs = sum(sum( (observed - expected).^2 ./ expected) );

df = (size(observed, 1) - 1) * (size(observed, 2) - 1);
p_value = 1 - chi2cdf(chi2obs, df);

printf("Chi-square Test Statistic: %f\n", chi2obs);
printf("P-value: %f\n", p_value);
```

```
Chi-square Test Statistic: 14.732308
P-value: 0.002060
```

Figure 1: X^2_{obs} and $p - value$