Student Information

Full Name : Onat Subaşı
Id Number : 2448835

Answer 1

a)

We have $1 - \alpha = 0.98$, $\alpha = 0.02$. $t_{\alpha/2} = t_{0.01} = 2.602(\text{with} n - 1 = 15 \text{degrees of freedom}$ (retrieved from the textbook). Since the sample size is small and it is not stated as a normal distribution, we used Student's t distribution.
From the given data, we calculate mean $\overline{X} \approx 6.81$ and standard deviation $s \approx 1.02$. Putting the values into the formula for confidence intervals, we have:

$$\begin{aligned}\overline{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}} &= 6.81 \pm 2.326\frac{1.02}{4} \\ &= 6.81 \pm 0.66 \text{ or } [6.14, 7.47]\end{aligned}$$

b)

We test the null hypothesis $H_0 : \mu = 7.5$ against a one-sided left-tail alternative $H_A : \mu < 7.5$.
We use t-test since we do not know the standard deviation of the previous data.
Step 1: Test statistic. We have $s \approx 1.02$, $n = 16$, $\alpha = 0.05$, $\mu_0 = 7.5$, $\overline{X} \approx 6.81$. The test statistic is

$$t = \frac{\overline{X} - \mu_0}{s/\sqrt{n}} = \frac{6.81 - 7.5}{1.02/4} \approx -2.72$$

Step 2: Acceptance and rejection regions. The critical value is

$$-t_\alpha = -t_{0.05} = -1.753 \text{ with } 16 - 1 = 15 \text{ degrees of freedom.}$$

With the left-tail alternative, we
Reject H_0 if $t \leq -1.753$,
Accept H_0 if $t > -1.753$,

Step 3: Result. Our test statistic $t = -2.72$ belongs to the rejection region. Therefore, we reject the null hypothesis. We have sufficient evidence to claim that the improvement is effective.

c)

In this case, since $\overline{X} - \mu_0$ would be guaranteed to be positive since $\overline{X} \approx 6.81$. $s\sqrt{n}$ is also guaranteed to be positive since neither of them can be negative. In this case, t is also guaranteed to be positive.
Since we are comparing t with $-t_\alpha$, which is a negative value, t is guaranteed to be greater than $-t_\alpha$, falling to the acceptance region. Therefore, we can immediately accept H_0.

Answer 2

a)

Null hypothesis $H_0 : \mu = 5000$, Alternative hypothesis $H_A : \mu > 5000$. Ali's claim should be considered as the null hypothesis.

b)

We test the null hypothesis $H_0 : \mu = 5000$ against a one-sided right-tail alternative $H_A : \mu > 5000$.
Step 1: Test statistic. We have $\sigma = 2000$, $n = 100$, $\alpha = 0.05$, $\mu_0 = 5000$, $\overline{X} = 5500$. The test statistic is

$$z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{5500 - 5000}{2000/10} = 2.5$$

Step 2: Acceptance and rejection regions. The critical value is

$$z_\alpha = z_{0.05} = 1.645 \text{ (retrieved from the textbook)}$$

With the right-tail alternative, we
Reject H_0 if $z \geq 1.645$,
Accept H_0 if $z < 1.645$,

Step 3: Result. Our test statistic $z = 2.5$ belongs to the rejection region. Therefore, we reject the null hypothesis. Ahmet has sufficient evidence to claim that there is an increase in the rent prices.

c)

We have the P-value $P = 1 - \Phi(2.5) = 0.0062$. P value is very low, indicating that we would reject the null hypothesis even for a smaller level of significance.

d)

We test the null hypothesis $H_0 : \mu_X - \mu_Y = 0$ against a one-sided right-tail alternative $H_A : \mu_X < \mu_Y$
Step 1: We have $\overline{X} = 5500$, $n = 100$, $\sigma_X = 2000$, $\overline{Y} = 6500$, $m = 60$, $\sigma_Y = 3000$. The test statistic is

$$z = \frac{\overline{Y} - \overline{X} - D}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}}} = \frac{6500 - 5500}{\sqrt{\dfrac{2000^2}{100} + \dfrac{3000^2}{60}}} \approx 2.29$$

Step 2: Acceptance and rejection regions. The critical value is

$$z_\alpha = z_{0.01} = 2.326$$

With the right-tail alternative, we
Reject H_0 if $z \geq 2.326$,
Accept H_0 if $z < 2.326$,

Step 3: Result. Our test statistic $z = 2.29$ belongs to the acceptance region. Therefore, we accept the null hypothesis. We do not have sufficient evidence to claim that prices in Ankara is lower than the prices in Istanbul.

Answer 3

We test H_0 : The number of rainy days in Ankara is independent of the season vs H_A : The number of rainy days in Ankara is dependent on the season.
Observed counts are $Obs = 34, 32, 15, 19$. The corresponding expected counts are

$$Exp(k) = \frac{\text{Total Number of Rainy Days}}{\text{Number of Seasons}} = \frac{34 + 32 + 15 + 19}{4} = 25$$

Observed counts for non-rainy days are $Obs' = 56, 58, 75, 71$. The corresponding expected counts are

$$Exp'(k) = \frac{\text{Total Number of Non-Rainy Days}}{\text{Number of Seasons}} = \frac{56 + 58 + 75 + 71}{4} = 65$$

We compute the chi-square statistic

$$\begin{aligned} x_{obs}^2 &= \sum_{k=1}^{4} \frac{\{Obs(k) - Exp(k)\}^2}{Exp(k)} + \sum_{k=1}^{4} \frac{\{Obs'(k) - Exp'(k)\}^2}{Exp'(k)} \\ &= \frac{(34-25)^2}{25} + \frac{(32-25)^2}{25} + \frac{(15-25)^2}{25} + \frac{(19-25)^2}{25} + \frac{(56-65)^2}{65} + \frac{(58-65)^2}{65} + \frac{(75-65)^2}{65} + \frac{(71-65)^2}{65} = 14.73 \end{aligned}$$

From Table A6 (from the textbook) with $N - 1 = 3$ d.f., the P-value is

$$P = P\{x^2 \geq 10.64\} = \text{between } 0.001 \text{ and } 0.005$$

The P-value is very low, indicating we need a significance level lower than 0.1% to accept the null hypothesis. Therefore, we reject the null hypothesis. We have sufficient evidence to claim that the number of rainy days in Ankara is dependent on the season.

Answer 4

```
pkg load statistics
observed = [34, 32, 15, 19; 56, 58, 75, 71];
first_row = observed(1,:);
second_row = observed(2,:);
cols = size(observed, 2);

expected = sum(first_row/cols);
n_expected = sum(second_row/cols);
x = sum(sum((first_row - expected).^2 ./ expected)) + sum(sum((second_row -
    n_expected).^2 ./ n_expected));

df = cols-1;
p_value = 1 - chi2cdf(x, df);

printf('X_obs^2 = %f\n', x);
printf('P-value: %f\n', p_value);
```

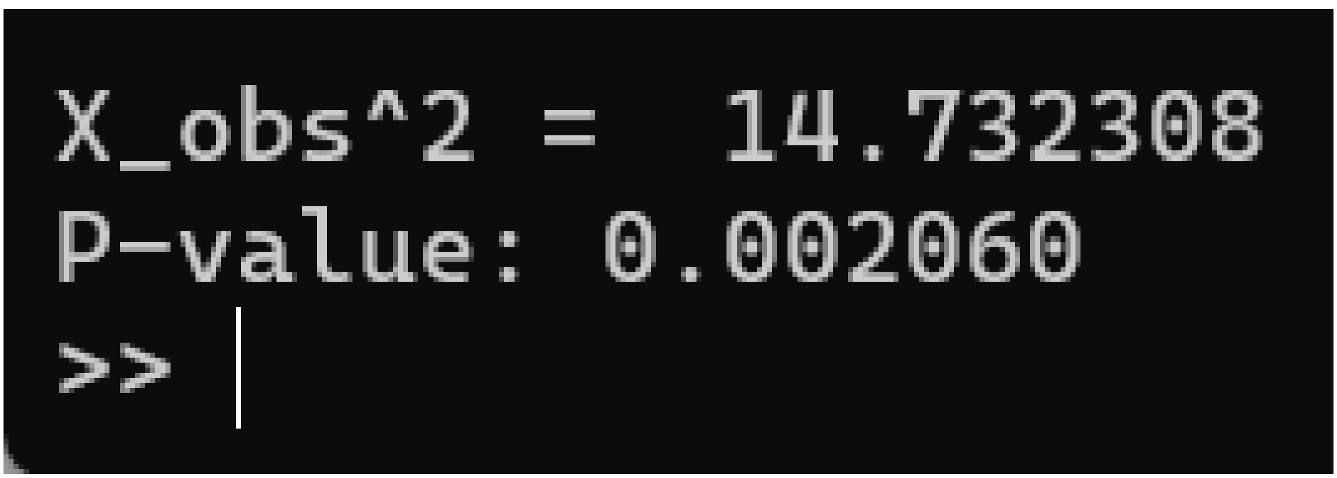

Figure 1: Output of the code above.