Student Information

Name : Unal Hasanli

ID : 2589877

Answer 1

a)

$n = 16$(**sample size**)
$\bar{X} = \frac{(8.4+7.8+6.4+6.7+6.6+6.6+7.2+4.1+5.4+6.9+7.0+6.9+7.4+6.5+6.5+8.5)}{16} \approx 6.81$ **(sample mean)**
$s \approx 1.06$(**sample standard deviation**) from the formula $\sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$.
$v = n - 1 = 15$ (**degrees of freedom**)
$t_{\alpha/2} = t_{0.01} = 2.602$(**critical value**)
Hence, 98% confidence interval for the mean time is $= 6.81 \pm 2.602 \times \frac{1.06}{\sqrt{16}} = [6.12, 7.5]$.

b)

Null hypothesis: The mean gasoline consumption remained the same after the improvement ($\mu = 7.5$).

Alternative hypothesis: The mean gasoline consumption has reduced after the improvement ($\mu < 7.5$).

One can use a t-test to compare the mean consumption before and after the improvement, since sample size is small and population standard deviation is unknown. Hence, we use the formula $t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$ to find test statistic. From part a, $\bar{=} \approx 6.81$, $s \approx 1.06$ and $v = 15$. Now, we substitute values and calculate the t: $\frac{6.81-7.5}{1.06/\sqrt{16}} \approx -2.6$.

Since we have a one-tailed test (H1 is $\mu < 7.5$), we will compare the t-value with the critical value for the left tail. From Table A5, the critical t-value at a significance level of 0.05 and 15 degrees of freedom is 1.713. Given that, we reject Null Hypothesis as $-2.6 < -1.713$.

Since we rejected Null Hypothesis, at a 5% level of significance, we can claim that there is a significant decline in mean gasoline consumption after the improvement.

c)

In that scenario, we would be able to immediately accept the Null hypothesis. The reasoning is that sample's average fuel consumption is greater than the assumed average($6.81 > 6.5$). In fact, if we try to get the t value, the calculation will give a positive value, which can never be less than any negative value(in particular, -1.713).

Answer 2

a)

Null hypothesis(Ali's statement): The rent prices have not changed last year.
Alternative hypothesis: There is an increase in the rent prices.

b)

Considering $\bar{X} = 5500$ and $\sigma = 2000$:

$$Z = \frac{5500 - 5000}{2000/\sqrt{100}} = 2.5$$

From Table A4, critical value for $\alpha = 0.05 \approx 1.645$. As z-score $2.5 > 1.645$, we have sufficient evidence to reject the Null hypothesis in favor of Alternative hypothesis. Therefore, Ahmet can claim that rent prices have increased at a 5% level of significance.

c)

$1 - \phi(Z_{obs}) = 1 - \phi(2.5) = 1 - 0.9938 = 0.0062$. Since $0.0062 < 0.05$, we can reject the null hypothesis. Hence, Ahmet can claim that rent prices have increased with 95% confidence.

d)

Null hypothesis: The rent prices are equal.
Alternative hypothesis: The rent prices in Ankara are lower.
$n_A = 100$, $\bar{X}_A = 5500$ and $\sigma_A = 2000$
$n_I = 60$, $\bar{X}_I = 6500$ and $\sigma_I = 3000$

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

From Table A4, $z_{0.01} \approx 2.325$. Since -2.29 > -2.326, we don't have enough evidence to reject the null hypothesis. Hence, one cannot claim that the rent prices in Ankara are lower than in Istanbul with 1% level of significance.

Answer 3

Null hypothesis: The number of rainy days in Ankara is independent on the season.
Alternative hypothesis: The number of rainy days in Ankara is dependent on the season.
First, we assume independency between seasons and the number of rainy days and calculate expected value for each case using the formula $Exp(i,j) = \frac{n_{i.} \times n_{.j}}{n}$:
Exp(Winter, Rainy)= $\frac{90 \times 100}{360} = 25$ Exp(Spring, Rainy)= $\frac{90 \times 100}{360} = 25$
Exp(Summer, Rainy)= $\frac{90 \times 100}{360} = 25$ Exp(Autumn, Rainy)= $\frac{90 \times 100}{360} = 25$

Exp(Winter, Non-Rainy)= $\frac{90\times 260}{360} = 65$ Exp(Spring, Non-Rainy)= $\frac{90\times 260}{360} = 65$
Exp(Summer, Non-Rainy)= $\frac{90\times 260}{360} = 65$ Exp(Autumn, Non-Rainy)= $\frac{90\times 260}{360} = 65$

$Exp(i,j) = \frac{n_{i.}\times n_{.j}}{n}$	Winter	Spring	Summer	Autumn	$n_{i.}$
Rainy	25	25	25	25	100
Non-Rainy	65	65	65	65	260
$n_{.j}$	90	90	90	90	360

$$X^2_{obs} = \frac{(34-25)^2}{25} + \frac{(32-25)^2}{25} + \frac{(15-25)^2}{25} + \frac{(19-25)^2}{25} + \frac{(56-65)^2}{65} + \frac{(58-65)^2}{65} + $$
$+ \frac{(75-65)^2}{65} + \frac{(71-65)^2}{65} \approx 14.73$. $v = (k-1)\times(m-1) = (2-1)\times(4-1) = 3$. With degree of freedom 3, critical value of chi distribution is less than 14.73 when $\alpha = 0.005$ and greater than 14.73 when $\alpha = 0.01$. Hence, we can reject Null hypothesis with 5% of significance. Precise p-value can calculated to be 0.00206 with special software.

Answer 4

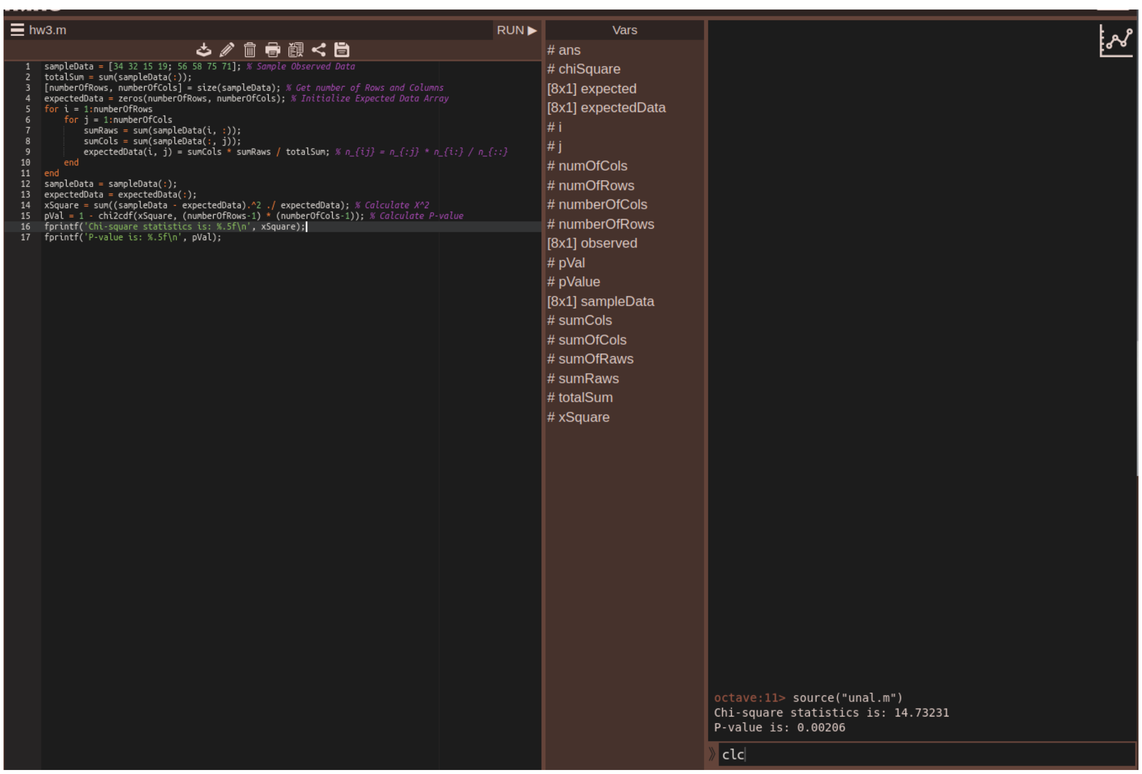

Figure 1: Octave code and corresponding output for part a

```
sampleData = [34 32 15 19; 56 58 75 71]; % Sample Observed Data
totalSum = sum(sampleData(:));
```

```
[numberOfRows, numberOfCols] = size(sampleData); % Get number of Rows
and Columns
expectedData = zeros(numberOfRows, numberOfCols); % Initialize
Expected Data Array
for i = 1:numberOfRows
    for j = 1:numberOfCols
        sumRaws = sum(sampleData(i, :));
        sumCols = sum(sampleData(:, j));
        expectedData(i, j) = sumCols * sumRaws / totalSum; % n_{ij} =
n_{:j} * n_{i:} / n_{::}
    end
end
sampleData = sampleData(:);
expectedData = expectedData(:);
xSquare = sum((sampleData - expectedData).^2 ./ expectedData);
pVal = 1 - chi2cdf(xSquare, (numberOfRows-1) * (numberOfCols-1));
fprintf('Chi-square statistics is: %.5f\n', xSquare);
fprintf('P-value is: %.5f\n', pVal);
```