Student Information

Name : Yusuf Sami Lök

ID : 2521748

Answer 1

a)

We have $n = 16$,
$\bar{X} = 6.806$ from $\dfrac{(8.4 + 7.8 + 6.4 + ... + 8.5)}{16}$,
Standard deviation, $s = 1.055$ from the formula $s^2 = \dfrac{\Sigma(X_i - \bar{X})^2}{n-1}$
We have to use t-table because the number of elements in the sample is less than 30, and we don't know the standard deviation of the population.
The critical value of t distribution with $n - 1 = 15$ degrees of freedom is $t_{\alpha/2} = t_{0.01} = 2.602$
Margin $= t_{\alpha/2} \cdot \dfrac{s}{\sqrt{n}}$, $\alpha = 1 - 0.98 = 0.02$, and $t_{\alpha/2} = t_{0.01} = 2.602$ from the t-table with 15 degrees of freedom.
Margin $= 2.602 \cdot \dfrac{1.055}{4} = 2.596$
Therefore, the μ interval is $[\bar{X} - Margin, \bar{X} + Margin] = [6.806 - 0.686, 6.806 + 0.686] = [6.12, 7.50]$.

b)

Let us test

$$H_0 : \mu = 7.5 \quad vs \quad H_A : \mu < 7.5$$

From a, we have sample statistics $n = 16, \bar{X} = 6.81$, and $s = 1.06$. Compute the T-statistic,
$t = \dfrac{\bar{X} - 7.5}{\frac{s}{\sqrt{n}}} = \dfrac{6.81 - 7.5}{\frac{1.06}{4}} = -2.60$
The rejection region is $(-\infty, -1.75)$ where we used T-distribution with $16 - 1 = 15$ degrees of freedom and $\alpha = 0.05$ because of the left tail alternative.
Since t is the element of the rejection region, we reject the null hypothesis and conclude that there is a significant evidence of reduction in the gasoline consumption.
Therefore, we can say that the improvement is effective at a 95% level of confidence, $\mu < 7.5$.

c)

When the mean is 6.5, the test value is always positive. Since we have a one sided left tail alternative, and we need a negative test value to reject the null hypothesis, we can not reject H_0 immediately.

Answer 2

a)

The null hypothesis $H_0 : \mu = 5000$ and the alternative hypothesis is $H_A : \mu > 5000$. Therefore, Ali's claim should be considered as the null hypothesis.

b)

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$Z = \frac{5500 - 5000}{\frac{2000}{10}} = 2.5$$

α is $1 - 0.95 = 0.05$. Since we use the right tail Z test, we should use z_α.
$z_{0.05} = 2.5$. Rejection region is $[1.645, +\infty]$.
Since the calculated Z is in the rejection region, we reject the null hypothesis and conclude that there is an increase in the rent prices compared to the last year at a 95% confidence level.

c)

We calculated Z in part a). $Z = 2.5$
$P = P\{Z \geq 2.5\} = 1 - P\{Z < 2.5\} = 1 - 0.9938 = 0.0062$ from the z-table.
Since $0.0062 < \alpha$, we can reject the null hypothesis by using p-value. The meaning of P value about Ahmet's claim is that as the p-value decreases, we reject the null hypothesis more strongly and accept the alternative hypothesis with the same level of strength.

d)

Let X is the sample from Ankara, and Y is the sample from Istanbul.
Let us test

$$H_0 : \bar{X} = \bar{Y} vs H_A : \bar{X} < \bar{Y}$$

$Z = \dfrac{\bar{X} - \bar{Y}}{\sqrt{\dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m}}}$ where n is the number of elements in the sample X, and m is the number of

elements in the sample Y.
$Z = \dfrac{5500 - 6500}{\sqrt{\dfrac{2000^2}{100} + \dfrac{3000^2}{60}}} = -2.294$
$\alpha = 1 - 0.99 = 0.01$
$z_{0.01} = 2.33$ from the z-table with $\alpha = 0.01$. We should use negative of this value in the comparison test because we construct a test with a left tail alternative.

Since -2.294(Z value) is not less than or equal to -2.33($-z_\alpha$)($-2.294 > -2.33$), we can not reject the null hypothesis. Therefore, at a 1% level of significance, they can not claim that the prices in Ankara is lower than the prices in Istanbul.

Answer 3

H_0 : "season and the number of rainy days are independent factors" vs H_A : "these factors are dependent" Here is the table that includes the expected values of the variables:

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

We have to find the χ^2_{obs} value from the formula:

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

where $Obs(i,j) = n_{ij}$ are observed counts, and
$Exp(i,j) = \dfrac{(n_{i.})(n_{.j})}{n}$ are estimated expected counts, and
χ^2_{obs} has $(k-1)(m-1)$ degree of freedom.
$\chi^2_{obs} = \dfrac{\{34-25\}^2}{25} + \dfrac{\{32-25\}^2}{25} + \dfrac{\{15-25\}^2}{25} + \dfrac{\{19-25\}^2}{25} + \dfrac{\{56-65\}^2}{65} + \dfrac{\{58-65\}^2}{65} + \dfrac{\{75-65\}^2}{65} + \dfrac{\{71-65\}^2}{65} = 14.732$
$\chi^2_{obs} = 14.732$ and degree of freedom is $(2-1)\cdot(4-1) = 1\cdot 3 = 3$.

We find an interval: $0.001 < P < 0.005$, from the chi-square table from the book.
Since the evidence against the null hypothesis is strong, they are dependent from the chi_square test. When the p-value is so small, we reject the null hypothesis more strongly. The less value of p-value, the more strongly we reject the null hypothesis basically.

Answer 4

```
X = [34 32 15 19;56 58 75 71];
Row = sum(X')'; Col = sum(X); Tot = sum(Row);
k = length(Col); m = length(Row);
e = zeros(size(X));
for i=1:m;
    for j=1:k;
        e(i,j) = Row(i)*Col(j)/Tot;
    end;
end;
chisq = (X-e).^2./e;
chistat = sum(sum(chisq));
Pvalue = 1-chi2cdf(chistat,(k-1)*(m-1));

fprintf ("chi^2 = %.8f\n " , chistat ) ;
fprintf ("P_val = %.8f\n " , Pvalue ) ;
```

```
octave:13> source("my_script.m")
chi^2 = 14.73230769
 P_val = 0.00206031
```