Student Information

Full Name : Aytaç SEKMEN
Id Number : 2575983

Answer 1

a)

Step-1) Let's first calculate the mean for this sample: $\mu \approx 6.81$
Note: In this question it is not stated that the given sample is from Normal distribution. Also the population deviation is not given. In addition to these, the sample size $n = 16 < 30$ so I should use t-distribution instead of Standard Normal Distribution.

Step-2) Now let's find the sample standard deviation by using the formula of (8.4) which is given in the page 219 in book: $s = 1.06$.
Step-3) Degree of freedom: 16-1=15.
Step-4) Now, lets calculate the α by using formula of $(1-\alpha) = 0.98$ then we can find that $\alpha = 0.02$.
Step-5) When I look at the t-distribution table with these values I can clearly see that: $t_{0.01} \approx 2.60$.
Then by using the formula 9.9 given in the page 259, I can calculate the confidence interval as follows:
$6.81 \pm (2.60)\dfrac{1.06}{\sqrt{16}} \approx 6.81 \pm 0.69 =$ [6.12;7.5]

b)

Let me first call that C0 is the initial gasoline consumption such that C0=7.5 and C1 is the gasoline consumption with improved version of engine. So lets define the null and alternative hypothesis now.
Null Hypothesis H_0: C0=C1
Alternative Hypothesis: C0>C1

So as we can clearly see from above, this is one-sided left tailed. Since we dont know the population standard deviation, and since sample size $n = 16 < 30$, we should use T-test here.
Step-1) Let's first find the t-statistic like this:

$$Z = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} = \frac{6.81 - 7.5}{\frac{1.06}{\sqrt{16}}} \approx -2.60$$

Step-2) Now let's determine the degree of freedom and α. Firstly, d.f.= 16-1=15 and $\alpha = 0.05$ because alternative hypothesis is single-sided. When I look at from T distribution table, I see the value of 1.753. Since my $t = -2.6291 \leq -1.753$ I can reject my null hypothesis and claim that the improvement is effective, at a 5% level of significance.

c)

First of all there is two way of thinking in this question, since I couldn't decide what is meant with "without any calculations". I'll share both way of my thinking:

1^{st} way: Let's now determine new hypothesis such that:
Null Hypothesis : $\mu = 6.5$
Alternative Hypothesis: $\mu \neq 6.5$
Now by using the theorem 9.21 in the book, we can clearly say that since "6.5" is in the confidence interval we found in part-a. So it is actually clear that null hypothesis that I derived is correct. Since it is correct, we can immediately accept H_0 without any calculations, with 2% significance level.

2^{nd} way: Actually by just "looking" at the calculation steps we can see that we substract μ from $\bar{X}$. Since in case of our $\mu = 6.5$, this substract will have positive sign. So it is not necessary to know the exact value of new t-statistic, we can immediately say it will be greater than -1.753 (because it is positive) so that we can immediately accept H_0 without any calculations.

Answer 2

a)

Since null hypotheses should be in the form of equality. Ali's claim is the null hypothesis and Ahmet's claim is the alternative hypothesis. And we can clearly see from Ahmet's claim that our situation is right-tailed. To sum up:
Null Hypothesis: $\mu = 5000$
Alternative Hypothesis: $\mu > 5000$

b)

Step-1) Let's first find the test-statistic:

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{5500 - 5000}{\frac{2000}{\sqrt{100}}} = 2.5$$

Step-2) Now we should check whether it is on acceptance or reject region.
$z_\alpha = z_{0.05} = 1.645$

Since our $Z = 2.5 \geq 1.645 = z_a$ we should rejectt H_0. So we can actually claim that alternative hypothesis is correct, in other words Ahmet can he claim that there is an increase in the rent prices compared to the last year.

c)

Since our situation is right tailed we should do the corresponding calculation. According to the page 283 in the book, we can say that our P value should be calculated using the following computation: $1-\phi(Z_obs) = 1-\phi(2.5) = 1-0.9938 = 0.0062$ which is clearly less than 0.01. This actually states that Ahmet can reject the null hypothesis not only at the 5%, but also at the 1% and even 0.05% level of significance.

d)

Step-1) For this question lets first say that μ_A represent the average rent price in Ankara and μ_I represents the average rent price in Istanbul. And σ_A represent the population standard deviation in Ankara and σ_I represents the population standard deviation in Istanbul. And, n represents the sample size of rent prices in Ankara, m represents the sample size of rent prices in Istanbul. Then our null hypothesis becomes $H_0 : \mu_A = \mu_I$ or $\mu_A - \mu_I = D = 0$ and alternative hypothesis becomes: $H_A : \mu_A < \mu_I$ or $\mu_A - \mu_I = D < 0$. So it is single sided left tail hypothesis.
Step-2) Now lets find the z-statistic by using this formula:

$$Z = \frac{\mu_A - \mu_I - D}{\sqrt{\frac{\sigma_A^2}{n} + \frac{\sigma_I^2}{m}}} = \frac{5500 - 6500}{\sqrt{\frac{2000^2}{100} + \frac{3000^2}{60}}} \approx -2.29$$

Step-3)α=0.01 and when I look from the Standard Normal Distribution table I can clearly see that $-z_a = -z_{0.01} = -2.33$
And since Z= $-2.29 > -2.33 = -z_a$ I should accept the Null Hypothesis, H_0. Which means that we can't claim that prices in Ankara is cheaper than Istanbul.

Answer 3

First we should say that Null Hypothesis represents that they are independent. (which actually represents that product of marginal probabilities equals to the joint probabilites. This is shown in10.4 at page 311 in the book, so I thought it is enought to verbally write it.) And alternative hypothesis represents that they are dependent. Also I should say that there are total of 360 days and 90 days for each season. And there are 100 rainy days and 260 non-rainy days.

Now I have to calculate the expected number of sampling units in each category, by using the formula at page 311. Let me identify the columns with "j" and the rows with "i". So my Exp(i,j) such that j=1,2,3,4 and i=1,2 should be like these:

$$E(1,1) = \frac{90 \times 100}{360} = 25 \; E(1,2) = \frac{90 \times 100}{360} = 25 \; E(1,3) = \frac{90 \times 100}{360} = 25$$

$$E(1,4) = \frac{90 \times 100}{360} = 25$$

$$E(2,1) = \frac{90 \times 260}{360} = 65 \; E(2,2) = \frac{90 \times 100}{360} = 65 \; E(2,3) = \frac{90 \times 100}{360} = 65$$

$E(2,4) = \frac{90 \times 100}{360} = 65$

Now it is time to calculate the X^2_{obs}.

$X^2_{obs} = \frac{(34-25)^2}{25} + \frac{(32-25)^2}{25} + \frac{(15-25)^2}{25} + \frac{(19-25)^2}{25} + \frac{(56-65)^2}{65} + \frac{(58-65)^2}{65} + \frac{(75-65)^2}{65} + \frac{(71-65)^2}{65} \approx 14.73$

Now lets find the degree of fredom $d.f. = (4-1)(2-1) = 3$. Now when we look at Chi-Square Distribution table, we can clearly see that our P-value stands in the interval of $0.005 \geq P \geq 0.001$ So it is obvious that our P-value< 0.01. So this means that we should reject the null hypothesis. So as stated in alternative hypothesis it is clear that the number of rainy days in Ankara is dependent on the season.

Answer 4

Codes:

```
% Chi square calculator

final_x2=0;
total=0;
input_array=[34,32,15,19;56,58,75,71];
number_of_rows=size(input_array,1);
number_of_columns=size(input_array,2);
numbers_in_rows=zeros(number_of_rows,1);
numbers_in_columns=zeros(number_of_columns,1);
degree_of_freedom=(number_of_rows-1)*(number_of_columns-1);
for i=1:number_of_rows
for j=1:number_of_columns
total=total+input_array(i,j);
numbers_in_rows(i,1)=numbers_in_rows(i,1)+input_array(i,j);
numbers_in_columns(j,1)=numbers_in_columns(j,1)+input_array(i,j);
end
end
for i=1:number_of_rows
for j=1:number_of_columns
average=numbers_in_rows(i,1)*numbers_in_columns(j,1)/total;
final_x2=final_x2+((input_array(i,j)-average)^2)/average;

end
end
p_value=1-chi2cdf(final_x2,degree_of_freedom);
```

```
fprintf('X^2-value: %.4f\n', final_x2);
fprintf('P-value: %.4f\n', p_value);
```

Screenshot: