NAME: PRAGYA AGARWAL

STUDENT ID: 1001861779

ASSIGNMENT NO: 05

ASSIGNMENT-05

Ques.1 we are interested to study a sensor system and we know that the sensor noise is distributed normally and has a mean of 4.65 and a standard deviation of 0.39. There is a new sensor which we want to see how it works so we take 10 measurements and the results are: {4.55, 5.35, 5.3, 4.6, 4.9, 5.5, 4.2, 5.1, 4.1, 4.0}. Can you show that the sensors generate different data?

Ans.1 For the given question following is given →

mean = $\mu = 4.65$

SD = $\sigma = 0.39$

As it is said that the data is distributed normally, we use z test for hypothesis testing.

Step1:

H_0 = Null Hypothesis = $\mu_0 = \mu_1$ (Generate same data)

H_1 = Alternative Hypothesis = $\mu_0 \neq \mu_1$ (two-tail test)

Step2: State the significance level. As value is not given, let us assume the significance level as 5% = 0.05

$\Rightarrow \frac{0.05}{2} = 0.025$

$\phi^{-1}(z_{0.025}) = \phi^{-1}(1-0.025)$

$= \phi^{-1}(0.975)$

$\phi^{-1}(z_{0.025}) = 1.96$

Step 3:

Given $\mu = 4.65$

$\sigma = 0.39$

Mean of sample (x) =

$$\frac{4.55 + 5.35 + 5.3 + 4.6 + 4.9 + 5.5 + 4.2 + 5.1 + 4.1 + 4}{10}$$

$$\text{mean} = \frac{47.6}{10}$$

mean = 4.76

Given $n = 10$

$$Z \text{ score} = \frac{x - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$= \frac{4.76 - 4.65}{\frac{0.39}{\sqrt{10}}} = \frac{0.11}{0.123}$$

$$\boxed{Z \text{ score} = 0.8919}$$

Step 4: $1.96 > 0.8919$ i.e. $-1.96 < 0.8919 < 1.96$

As the Zscore does not fall under the rejection region, we accept the null hypothesis and reject the alternative hypothesis.
Therefore, the sensors do not generate different data.

Ques. 2 In order to evaluate the runtime performance of an algorithm, we measured it's run time performance on 15 samples and results are:
{17, 16.3, 18.2, 17.9, 16, 16.5, 18.9, 19.9, 18, 18.3, 18.1, 15.7, 17.6, 17.7, 17.1}. The state-of-the art algorithm that has an average runtime performance of 17 and a standard deviation of 5.82. Is there any difference between the first algorithms and state-of-the-art algorithm?

Ans. 2 The given distribution is not normally distributed and the size of the sample is small i.e. less than 30. Therefore, we will use t-test.

Step 1: $H_0 = \mu_1 = \mu_2$ = Null Hypothesis = Both Algorithms have the same performance

$H_1 = \mu_1 \neq \mu_2$ = Alternative Hypothesis = Both Algorithms have different performance

Step 2: For finding critical value of t we will assume that $\alpha = 0.05$. Hence, we will be conducting non-directional test (two-sided)

$\therefore t_{critical}$ for degree of freedom i.e. $(n-1)$
$= (15-1)$
$= 14$

$\therefore t_{critical}$ for degree of freedom is 14 & significance % 95
$= 2.145$

Step 3 → $$t_{absolute} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

We know that :

$$\mu = 17$$
$$s = 5.82$$
$$n = 15$$

$$\bar{x} = \frac{17+16.3+18.2+17.9+16+16.5+18.9+19.9+18+18.3+18.1+15.7+17.6+17.7+17.1}{15}$$

$$= \frac{263.2}{15}$$

$$\bar{x} = 17.54$$

$$t_{absolute} = \frac{17.54-17}{\frac{5.82}{\sqrt{15}}}$$

$$= \frac{0.54}{1.5027}$$

$$t_{absolute} = \cancel{0.35}\ 0.3593$$

$$\boxed{t_{absolute} = 0.36}$$

Step 4: As the threshold is 2.145 & $-2.145 < 0.36 < 2.145$, we accept the null hypothesis and reject the alternate hypothesis.

∴ Both the algorithms have same performance.

Ques.3 In order to develop a new approach to solve a problem, a student develops a new approach (method 1) and the result she gets are: {12, 4, 6, 7, 12, 4, 7, 4, 15, 10, 8, 10, 9, 5, 9}. Method 2 which is the current state-of-the-art approach, has average performance of 7 and a standard deviation of 3. The student compares the results and discovers that it seems that method 1 outperforms method 2 and she wants to prove it using significance testing with a two-tailed 1% significance threshold.

Ans.3 The sample given is not normally distributed and the size of the sample is small as it is less than 30. Hence, we will use t-test.

Step1: $H_0 = \mu_1 = \mu_2$ = Null Hypothesis i.e. method 1 & method 2 performs the same

$H_1 = \mu_1 \neq \mu_2$ = Alternative hypothesis i.e. method 1 & method 2 performs different

Step2: Given,

$\alpha = 0.01$

$\Rightarrow$ Degree of freedom $= n-1 = 15-1 = 14$

we need to perform two-tailed test

$\therefore t_{critical} = 2.977$

Step3: For finding $t_{absolute} \rightarrow$

$$t_{absolute} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

Given,

$\mu = 7$

$s = 3$

$n = 15$

$$\bar{x} = \frac{12+4+6+7+12+4+7+4+15+10+8+10+9+5+9}{15}$$

$\bar{x} = 8.13$

$$\Rightarrow t_{absolute} = \frac{8.13-7}{\frac{3}{\sqrt{15}}}$$

$$= \frac{1.13}{0.7746}$$

$$= 1.45818$$

$$\boxed{t_{absolute} = 1.46}$$

Step 4: As the threshold is 2.977 & $1.46 < 2.977$, we accept the null hypothesis and reject the alternative hypothesis.

$\therefore$ Both methods performs same.

Ques. 4 A researcher wants to do a research on people's body height (in meter) in a specific city so he measures the height of 36 random persons and the results are:
{1.8, 1.72, 1.85, 2.48, 1.58, 1.65, 1.9, 1.83, 1.8, 1.7, 1.75, 1.83, 1.6, 1.65, 1.5, 1.42, 1.82, 1.87, 1.53, 1.65, 1.99, 1.78, 1.69, 1.52, 1.62, 1.63, 1.75, 1.8, 1.72, 1.6, 1.82, 1.62, 1.87, 1.75, 1.82, 1.52}. Suppose that from previous studies which has been done 50 years ago, the average height of persons was determined to be 1.73 and standard deviation is 0.2 in this city.

Ans. 4 It is given that sample size is greater than 30. Therefore, we will use Z-test.

Step 1: $H_0 = \mu = 1.73$

$H_1 = \mu \neq 1.73$

Step 2: Let us assume $\alpha = 0.05$

It is two-tail test

$\Rightarrow \frac{\alpha}{2} = \frac{0.05}{2} = 0.025$

$\phi^{-1}(Z_{0.025}) = \phi^{-1}(1 - 0.025)$

$= \phi^{-1}(0.0975)$

$\phi^{-1}(Z_{0.025}) = 1.96$

Step 3: Given,

$\mu = 1.73$

$\sigma = 0.2$

To calculate z,

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$\bar{x} = \frac{1.8+1.72+1.85+1.48+1.58+1.65+1.9+1.83+1.8+1.7+1.75+1.83+1.6+1.65+1.5+1.42+1.82+1.87+1.53+1.65+1.99+1.78+1.67+1.52+1.62+1.63+1.95+1.8+1.72+1.6+1.82+1.62+1.87+1.75+1.82+1.52}{36}$$

$\bar{x} = 1.706$

$$\Rightarrow z = \frac{1.706 - 1.73}{\frac{0.2}{\sqrt{36}}}$$

$$= \frac{-0.024}{0.033}$$

$z = -0.72$

Absolute value is 0.72

Step 4 : $-1.96 < 0.72 < 1.96$

Therefore we accept the null hypothesis & reject the alternative hypothesis.

The average height has remained same for last 50 years. (No change in data) So we cannot use this data to show average height of city is changed in past 50 years.