|  |
| --- |
| Photo displaying partial image of two pie charts on a canvas-textured page |
| Distribution Analysis  Assignment-2 |
| |  |  |  | | --- | --- | --- | | Jay Panchal – 100960958 | 5/27/24 | Statistical and Predictive Modeling | |

Contents

[**Question 1:** 2](#_Toc168861803)

[**Hypothesis Statement:** 2](#_Toc168861804)

[**Question 2:** 2](#_Toc168861805)

[**Normality of dataset based on the QQ plot and Histogram:** 2](#_Toc168861806)

[**Question 3:** 4](#_Toc168861807)

[**Step-by-step Outline:** 4](#_Toc168861808)

[**Question 4: Conducting Analysis in R** 5](#_Toc168861809)

[**Question 5: Summary and findings and conclusion.** 6](#_Toc168861810)

# **Question 1:**

## **Hypothesis Statement:**

As a researcher we want to test**, if** an individual has followed the diet program properly for the past 6 months, **then** their cholesterol level has changed significantly compared to before the diet program.

# **Question 2:**

## **Normality of dataset based on the QQ plot and Histogram:**

**QQ-Plot:**

A graph of a graph

Description automatically generated with medium confidence

* Based on the QQ plot, we can clearly say that relation between data is normally distributed as all the data points are near to the reference line.

**Histogram:**

A graph of a normality

Description automatically generated

* The data is normally distributed. The histogram shows a bell curve, which is one of the characteristics of the normal distribution. The data is centered around the 5 and distribution is symmetric, with the tails of the distribution decreasing gradually.

# **Question 3:**

## **Step-by-step Outline:**

**Step 1: Null hypothesis (H0) and Alternative hypothesis (Ha)**

**Ho: μ = μo:** if an individual has followed the diet program properly for the past 6 months, then their cholesterol level has **not changed** significantly.

**Ha: μ ≠ μo:** if an individual has followed the diet program properly for the past 6 months, then their cholesterol level has **changed** significantly.

Now, we are testing if the cholesterol level is 5.95, we have the following setup:

Ho: μ = 5.95

Ha: μ ≠ 5.95

**Step 2: The significance level (α) = 0.05**

**Step 3: Choose which test we need to perform (One sided or two sided) and calculate p-value.**

* Two- tail Test as we are calculating difference from null hypothesis.
* mu - sample mean
* mu0 - Population mean
* alpha (α) – significance level (usually 0.05)
* sigma - population standard deviation
* n - sample size
* Calculate z

z = (mu - mu0) / (sigma / sqrt(n))

* Calculate p-value

p <- 2\*pnorm(abs(z),lower.tail=FALSE)

**Step 4: Conclusion: Either support or reject the null hypothesis.**

By comparing p-value to significance level (α) we will decide whether null hypothesis is rejected or not.

# **Question 4: Conducting Analysis in R**

**A screenshot of a computer

Description automatically generated**

**Result:**

**A screenshot of a computer program

Description automatically generated**

# **Question 5: Summary and findings and conclusion.**

* The objective of this analysis was to determine whether an individual’s cholesterol level has changed after following the specific diet program for the past 6 months, for which we will conduct the hypothesis testing.
* For the same, we stated the hypothesis statement and we then we will plot the histogram to check the normality of the data.
* As histogram shows, cholesterol data is approximately normally distributed.
* Based on the hypothesis test, z score is -11.13135 and p- value is 8.828976e-29.
* Since p-value (8.828976e-29) is much lower than significance level (0.05), we reject the null hypothesis.
* **Conclusion:** We reject the null hypothesis that the population mean cholesterol level is equal to 5.95 at the 0.05 significance level. In other words, the analysis provides strong evidence that the mean cholesterol level after the diet program is significantly different from the population mean 5.95.