MD2201: Data Science

Name of the student: Rajeev Tapadia   **Roll No.: 67**

Div: C   **Batch:3**

Date of performance:

Experiment No.3

Title: Hypothesis Testing

Aim: i. To perform hypothesis testing for single sided tail.
   ii. To perform hypothesis testing for Two-sided tail.
Software used: Programming language R.
Code Statement:

1. Suppose the manufacturer claims that the mean lifetime of a ball bearing is 10000hours. The auditing team stated that the mean lifetime is less than what is claimed. On the basis of a randomly chosen sample of 50 ball bearings as given in the dataset, at 0.05 significance level, can we reject the claim of the manufacturer? What will be your interpretation if the significance level is made as 0.01? Consider the data set titled "Hypothesis_csv1.csv".

```r
2.  tValue = (mean - 10000) / (sd / sqrt(n))
3.  criticalValue = qnorm(0.05)
4.
5.  if(tValue < criticalValue) {
6.    cat("reject null hypothesis, lifetime is less")
7.  } else {
8.    cat("fail to reject null hypothesis, lifetime is 10,000")
9.  }
10.
11. criticalValue = qnorm(0.01)
12.
13. if(tValue < criticalValue) {
14.   cat("reject null hypothesis, lifetime is less")
15. } else {
16.   cat("fail to reject null hypothesis, lifetime is 10,000")
17. }
18.
```

```
> source("~/Data-Science/Data-Science-Lab/
reject null hypothesis, lifetime is less
reject null hypothesis, lifetime is less
>
```

19. The nutrition label on a bag of potato chips says that a one ounce (28 gram) serving of potato chips has 130 calories and contains ten grams of fat, with three grams of saturated fat. A random sample of 35 bags yielded a sample mean of 134 calories with a standard deviation of 17 calories. Is there evidence that the nutrition label does not provide an accurate measure of calories in the bags of potato chips? We have verified the independence, sample size, and skew conditions are satisfied. Take alpha as 5% and 1%.

```r
20. tValue = (mean - 130) / (sd / sqrt(n))
21.
22. criticalValue = qnorm(0.025)
23.
24. # using abs for two sided tail
25. if(abs(tValue) < abs(criticalValue)) {
26.    cat("reject null hypothesis, measure is inaccurate")
27. } else {
28.    cat("fail to reject null hypothesis, measure is accurate")
29. }
30.
31. criticalValue = qnorm(0.01)
32.
33. if(abs(tValue) < abs(criticalValue)) {
34.    cat("reject null hypothesis, measure is inaccurate")
35. } else {
36.    cat("fail to reject null hypothesis, measure is accurate")
37. }
38.
```

```
> source("~/Data-Science/Data-Science-Lab/ass
reject null hypothesis, measure is inaccurate
reject null hypothesis, measure is inaccurate
>
```

Conclusion:

This assignment investigated two hypothesis testing scenarios: ball bearing lifetime and potato chip calorie content. By analyzing sample data, we aimed to determine if the actual values differ from manufacturer claims (10,000 hours lifetime and 130 calories per serving). The analysis considers significance levels of 0.05 and 0.01, providing results for both. A full conclusion would depend on the calculated t-values and critical values, but the overall purpose was to assess if there's evidence to reject the claims made about the products.