Nithin S
221IT085

# IT302 Lab Assignment 4

Q1. A machine puts an average of 5 grams of jelly beans in bags, with a standard deviation of 0.25 grams. If 45 bags are randomly chosen, write a R/Python code to find the probability that the mean amount per bag in the sampled bags is less than 3.5 grams?

Code

```
mean <- 5
sd <- 0.25
n <- 45
sample_mean <- 3.5
SE <- sd / sqrt(n)
z <- (sample_mean - mean) / SE
probability <- pnorm(z)
probability
```

Output

```
> mean <- 5
> sd <- 0.25
> n <- 45
> sample_mean <- 3.5
> SE <- sd / sqrt(n)
> z <- (sample_mean - mean) / SE
> probability <- pnorm(z)
> probability
[1] 0
>
```

Probability is 0

Q2. A consumer-reports group is testing whether a gasoline additive changes a car’s gas mileage. A test of seven cars finds an average improvement of 0.5 miles per gallon with a standard deviation of 3.77. Is this difference significantly greater than 0.7 miles per gallon? Assume the values are normally distributed. Test at 1% level of significance.

Code

$$H_0 : \mu = 0.7$$
$$H_a : \mu > 0.7$$

```
mu <- 0.7
x <- 0.5
s <- 3.77
n <- 7
t <- (x-mu)/(s/sqrt(n))
df<-n-1
p_value <- pt(t, df, lower.tail = FALSE)
cat("T value: ",t,"\n")
cat("P Value: ",p_value)
alpha<-0.1
if (p_value < alpha) {
  cat("Reject the null hypothesis\n")
} else {
  cat("Fail to reject the null hypothesis\n")
}
```

Output

```
> mu <- 0.7
> x <- 0.5
> s <- 3.77
> n <- 7
> t <- (x-mu)/(s/sqrt(n))
> df<-n-1
> p_value <- pt(t, df, lower.tail = FALSE)
> cat("T value: ",t,"\n")
T value:  -0.1403582
> cat("P Value: ",p_value)
P Value:  0.5535148> alpha<-0.1
> if (p_value < alpha) {
+   cat("Reject the null hypothesis\n")
+ } else {
+   cat("Fail to reject the null hypothesis\n")
+ }
Fail to reject the null hypothesis
> 
```

Since the p-value (0.5535) is significantly higher than the significance level of 0.01 (1%), we **do not reject the null hypothesis**. This indicates that there is insufficient evidence to conclude that the gasoline additive produces a statistically significant improvement in gas mileage greater than 0.7 miles per gallon.

Q3. A car manufacturing company wants to test if its car meets the air pollution standard that the mean emission is less than 20 ppm of carbon. A random sample of 10 cars of this company is selected and the cars are tested. The emission level (ppm) of each of 10 cars as follows
20.5, 15.6, 16.2, 22.5, 17.9, 16.4, 12.7, 19.4, 13.9, 16.6
Is there enough evidence to conclude that the cars meet the air pollution standard? Test at 1% level of significance.

Code

$$H0 : \mu \geq 20$$
$$H1 : \mu < 20$$

```
emissions <- c(20.5, 15.6, 16.2, 22.5, 17.9, 16.4, 12.7, 19.4, 13.9, 16.6)
n<-10
mu<-20
alpha<-0.01
x_bar=sum(emissions)/n
cat("X_bar is:", x_bar, "\n")
sd <- sqrt(sum((emissions - x_bar)^2) / (n - 1))
cat("SD is: ",sd,"\n")
t=(x_bar-mu)/(sd/(sqrt(n)))
cat("t = ",t,"\n")
df<-n-1
critical_value <- qt(alpha, df, lower.tail = TRUE)
cat("Critical Value: ",critical_value,"\n")


if (t < critical_value) {
  cat("Reject the null hypothesis\n")
} else {
  cat("Fail to reject the null hypothesis\n")
}
```

Output

```
> emissions <- c(20.5, 15.6, 16.2, 22.5, 17.9, 16.4, 12.7, 19.4, 13.9, 16.6)
> n<-10
> mu<-20
> alpha<-0.01
> x_bar=sum(emissions)/n
> cat("X_bar is:", x_bar, "\n")
X_bar is: 17.17
> sd <- sqrt(sum((emissions - x_bar)^2) / (n - 1))
> cat("SD is: ",sd,"\n")
SD is:  2.981443
> t=(x_bar-mu)/(sd/(sqrt(n)))
> cat("t = ",t,"\n")
t =  -3.00165
> df<-n-1
> critical_value <- qt(alpha, df, lower.tail = TRUE)
> cat("Critical Value: ",critical_value,"\n")
Critical Value:  -2.821438
>
>
> if (t < critical_value) {
+   cat("Reject the null hypothesis\n")
+ } else {
+   cat("Fail to reject the null hypothesis\n")
+ }
Reject the null hypothesis
> |
```

Since the calculated t-statistic (-3.00) is less than the critical t-value (-2.82), we **reject the null hypothesis**. This conclusion indicates that the data provide sufficient evidence to conclude that the engine's true mean emission level is less than 20 ppm, thereby satisfying the pollution standard at the 1% significance level.

Q4. A machine fills 12-ounce of soda cans. Due to government regulation, the standard deviation of weights can not exceed 0.03 ounce. A sample of 8 cans is selected and the number of ounces of soda in each can is as given below. Test at 5% significance level if the machine is functioning properly.
12.3, 12.1, 12.02, 11.98, 12.00, 12.05, 11.97, 11.99

Code

- Null Hypothesis ( $H_0$ ): The machine is working properly/correctly, i.e., ( $\sigma \leq 0.03$ ).
- Alternate Hypothesis ( $H_1$ ): The machine is not working properly/correctly, i.e., ( $\sigma > 0.03$ ).

```
machines<-c(12.3, 12.1, 12.02, 11.98, 12.00, 12.05, 11.97, 11.99)
n <- length(machines)
alpha<-0.05
mu<-0.03
sample_sd <- sd(machines)
print(sample_sd)
df<-n-1
print(df)
chi_square_statistic <- (df * (sample_sd^2))/mu^2
cat("Chi Square: ",chi_square_statistic,"\n")
critical_value <- qchisq(1 - alpha, df)
cat("Critical value: ",critical_value)
if (critical_value < chi_square_statistic) {
  cat("Reject the null hypothesis\n")
} else {
  cat("Fail to reject the null hypothesis\n")
}
```

Output

```
> machines<-c(12.3, 12.1, 12.02, 11.98, 12.00, 12.05, 11.97, 11.99)
> n <- length(machines)
> alpha<-0.05
> mu<-0.03
> sample_sd <- sd(machines)
> print(sample_sd)
[1] 0.1090789
> df<-n-1
> print(df)
[1] 7
> chi_square_statistic <- (df * (sample_sd^2))/mu^2
> cat("Chi Square: ",chi_square_statistic,"\n")
Chi Square:  92.54167
> critical_value <- qchisq(1 - alpha, df)
> cat("Critical value: ",critical_value)
Critical value:  14.06714> if (critical_value < chi_square_statistic) {
+   cat("Reject the null hypothesis\n")
+ } else {
+   cat("Fail to reject the null hypothesis\n")
+ }
Reject the null hypothesis
>
```

Since our calculated Chi-square value/result (92.63) is more than the critical value (14.067), **we reject/dismiss the null hypothesis.** This shows that the machine's standard deviation of pop filling excide/surpasses the permitted/controlled limit/restrain of 0.03 ounces, meaning the machine is not working properly/correctly as per Government/administrative standards.

Q5. A college bookstore claims that, on average, a college student will pay Rs. 101.75 per class for textbooks. A student group investigates this claim by randomly selecting ten courses from the course catalog and finding the textbook costs for each.
The data collected is 140, 125, 150, 124, 143, 170, 125, 94, 127, 53
Do a test at 5% level of significance of H0: μ=101.75 against the alternative hypothesis H1: μ>101.75? Assume the values are normally distributed

Code

$$H_0 : \mu = 101.75$$

$$H_1 : \mu > 101.75$$

```r
data <- c(140, 125, 150, 124, 143, 170, 125, 94, 127, 53)
mu <- 101.75
alpha <- 0.05
x_bar <- mean(data)
n <- length(data)
sd <- sqrt(sum((data - x_bar)^2) / (n - 1))

t <- (x_bar - mu) / (sd / sqrt(n))
cat("T value: ",t)

df <- n - 1

p_value <- 1 - pt(t, df)
cat("P-value:", p_value, "\n")


if (p_value < alpha) {
  cat("Reject the null hypothesis\n")
} else {
  cat("Fail to reject the null hypothesis\n")
}
```

Output

```
> data <- c(140, 125, 150, 124, 143, 170, 125, 94, 127, 53)
> mu <- 101.75
> alpha <- 0.05
> x_bar <- mean(data)
> n <- length(data)
> sd <- sqrt(sum((data - x_bar)^2) / (n - 1))
>
> t <- (x_bar - mu) / (sd / sqrt(n))
> cat("T value: ",t)
T value:  2.291015>
> df <- n - 1
>
> p_value <- 1 - pt(t, df)
> cat("P-value:", p_value, "\n")
P-value: 0.02384771
>
>
> if (p_value < alpha) {
+   cat("Reject the null hypothesis\n")
+ } else {
+   cat("Fail to reject the null hypothesis\n")
+ }
Reject the null hypothesis
> |
```

After conducting the hypothesis test, we find that the p-value is less than 0.05 . Thus, we **reject the null hypothesis** and conclude that there is significant evidence to support the claim that the average textbook cost per course is greater than Rs 101.75

Q6. The data set normtemp (package "UsingR") contains measurements of 130 healthy, randomly selected individuals. The variable temperature contains normal body temperature. Perform a test to see if the commonly assumed value of 98.6 °F is correct. Test at 5% level of significance.

Code

H0: μ = 98.6
H1: μ != 98.6

```r
library(UsingR)
data(normtemp)
temperatures <- normtemp$temperature

n <- length(temperatures)
mean_temp <- mean(temperatures)
sd_temp <- sd(temperatures)
mu <- 98.6
t_stat <- (mean_temp - mu) / (sd_temp / sqrt(n))
df <- n - 1
alpha <- 0.05
critical_t <- qt(1 - alpha/2, df)
p_value <- 2 * (1 - pt(abs(t_stat), df))

cat("Sample Mean:", mean_temp, "\n")
cat("Sample Standard Deviation:", sd_temp, "\n")
cat("T-Statistic:", t_stat, "\n")
cat("Critical Value:", critical_t, "\n")
cat("P-Value:", p_value, "\n")

if (p_value < alpha) {
  cat("Reject the null hypothesis\n")
} else {
  cat("Fail to reject the null hypothesis\n")
}
```

Output

```
> data(normtemp)
> temperatures <- normtemp$temperature
>
> n <- length(temperatures)
> mean_temp <- mean(temperatures)
> sd_temp <- sd(temperatures)
> mu <- 98.6
> t_stat <- (mean_temp - mu) / (sd_temp / sqrt(n))
> df <- n - 1
> alpha <- 0.05
> critical_t <- qt(1 - alpha/2, df)
> p_value <- 2 * (1 - pt(abs(t_stat), df))
>
> cat("Sample Mean:", mean_temp, "\n")
Sample Mean: 98.24923
> cat("Sample Standard Deviation:", sd_temp, "\n")
Sample Standard Deviation: 0.7331832
> cat("T-Statistic:", t_stat, "\n")
T-Statistic: -5.454823
> cat("Critical Value:", critical_t, "\n")
Critical Value: 1.978524
> cat("P-Value:", p_value, "\n")
P-Value: 2.410632e-07
>
> if (p_value < alpha) {
+   cat("Reject the null hypothesis\n")
+ } else {
+   cat("Fail to reject the null hypothesis\n")
+ }
Reject the null hypothesis
>
```

P-value  less than alpha.
**Reject the null hypothesis.** There is sufficient evidence to suggest that the mean body temperature is not 98.6 °F.

Q7. In the babies (package “UsingR”) data set, the variable dht contains the father’s height. Do a significance test of the null hypothesis that the mean height is 68 inches against an alternative that it is taller. Remove the values of 99 from the data, as these indicate missing data. Test at 10% level of significance.

Code

H0: $\mu=68$
H1: $\mu>68$

```r
data("babies")

dht_clean <- babies$dht[babies$dht != 99]

mu <- 68
alpha <- 0.10

sample_mean <- mean(dht_clean)
sample_sd <- sd(dht_clean)
n <- length(dht_clean)

t_stat <- (sample_mean - mu) / (sample_sd / sqrt(n))
cat("Test Statistic (t):", t_stat, "\n")

df <- n - 1
p_value <- pt(t_stat, df, lower.tail = FALSE)
cat("p-value:", p_value, "\n")

if (p_value < alpha) {
  cat("Reject the null hypothesis\n")
} else {
  cat("Fail to reject the null hypothesis\n")
}
```

Output

```
> data("babies")
>
> dht_clean <- babies$dht[babies$dht != 99]
>
> mu <- 68
> alpha <- 0.10
>
> sample_mean <- mean(dht_clean)
> sample_sd <- sd(dht_clean)
> n <- length(dht_clean)
>
> t_stat <- (sample_mean - mu) / (sample_sd / sqrt(n))
> cat("Test Statistic (t):", t_stat, "\n")
Test Statistic (t): 20.79567
>
> df <- n - 1
> p_value <- pt(t_stat, df, lower.tail = FALSE)
> cat("p-value:", p_value, "\n")
p-value: 2.35261e-76
>
> if (p_value < alpha) {
+   cat("Reject the null hypothesis\n")
+ } else {
+   cat("Fail to reject the null hypothesis\n")
+ }
Reject the null hypothesis
>
```

**P-value is less than alpha. Reject the null hypothesis**. There is evidence that the mean height is greater than 68 inches.