MTH 372: Assignment 2

Ashwin Sheoran
2020288

Ans 1)

```
> rm(list=ls())
>
> set.seed(123) # set seed for reproducibility
>
> ## sample size
> n<-10
>
> popn_mean <- 17
>
> popn_standard_deviation <- 0.5
>
> mu <- 15
>
> alpha <- 0.05
>
> ztest_statistic<-(popn_mean - mu)/(popn_standard_deviation/sqrt(n))
> cat("Test statistic: ",ztest_statistic,"\n")
Test statistic:  12.64911
>
> # we calculate the critical value
> # given alpha = 0.05, so for two-tailed we divide by 2
> ztest_critical_value<-qnorm(p=alpha,0,1,lower.tail = FALSE)
> cat("Test Critical Value: ",ztest_critical_value,"\n")
Test Critical Value:  1.644854
> # Here the test_critical_value lies in the rejection region thus, we reject H0
>
> p_value <-  pnorm(ztest_critical_value,0,1,lower.tail = FALSE)
> cat("P value: ",p_value,"\n")
P value:  0.05
>
> if(abs(ztest_statistic) > ztest_critical_value | p_value < alpha){
+   cat("Reject the null hypothesis that bread height is <= 15 cm.\n")
+ } else {
+   cat("Fail to reject the null hypothesis that bread height is <= 15 cm.\n")
+ }
Reject the null hypothesis that bread height is <= 15 cm.
```

<u>Hypothesis:</u>

Given $\mu_0 = 15$

Let H_0 = The mean height of the bread baked by the baker is less than or equal to 15 cm.

H_1 = The mean height of the bread baked by the baker is greater than 15 cm.

Therefore, our hypothesis becomes:
$H_0 : \mu \leq \mu_0$ v/s $H1 : \mu > \mu_0$

Testing:
We will use a one-tailed t-test to test our hypothesis at a 5% significance level. The test statistic is calculated as follows: (here we have used z test)
ztest_statistic<-(popn_mean - mu)/(popn_standard_deviation/sqrt(n))
where n = sample size, and μ0 = 15.

Results:
Using the given values, we find that the sample mean height of the bread is 17 cm, and the standard deviation is 0.5 cm. We calculate the test statistic as:
The test statistic: 12.64911
P value: 0.05

We reject the null hypothesis because the calculated test statistic is greater than the critical value.

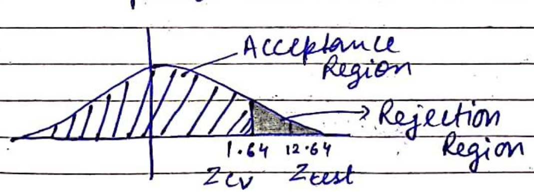
Inference:
Since the calculated test statistic is greater than the critical value, we reject the null hypothesis.

Conclusion:
Based on the results of our hypothesis test, we **can conclude that the data support the baker's claim, and we reject the null hypothesis that the average height of bread baked by the baker is less than or equal to 15 cm.**

Here $Z_{Test} = \frac{\bar{x} - \mu}{\sigma} \sqrt{n} = \frac{17 - 15}{0.5} \sqrt{10}$
$= 12.64$

Now using Z-Table $Z_{0.05} = 1.64 \rightarrow Z_{cv}$

∴ Our Graph for 1-Tailed will be



Ans2)

```
> rm(list=ls())
> # Given values
> n <- 75
> sample_mean <- 17.4
> sample_standard_deviation <- 6.3
> mu <- 15
> alpha <- 0.05
>
> # Calculate the test statistic
> test_stat <- (sample_mean - mu)/(sample_standard_deviation/sqrt(n))
>
> # Calculate the degrees of freedom
> df <- n - 1
>
> # Calculate the critical values from t-distribution
> left_critical_value <- qt(alpha/2, df = df, lower.tail = TRUE)
> right_critical_value <- qt(alpha/2, df = df, lower.tail = FALSE)
>
> # Calculate the p-value
> p_value <- 2 * pt(abs(test_stat), df = df, lower.tail = FALSE)
>
> # Print the results
> cat("Test statistic:", test_stat, "\n")
Test statistic: 3.299144
> cat("Left critical value:", left_critical_value, "\n")
Left critical value: -1.992543
> cat("Right critical value:", right_critical_value, "\n")
Right critical value: 1.992543
> cat("p-value:", p_value, "\n")
p-value: 0.001493164
>
> # Compare the test statistic with critical values and p-value with alpha
> if(test_stat < left_critical_value | test_stat > right_critical_value | p_value < alpha){
+   cat("Reject the null hypothesis that the population mean time on death row is 15 years.\n")
+ } else {
+   cat("Fail to reject the null hypothesis that the population mean time on death row is 15 year
s.\n")
+ }
Reject the null hypothesis that the population mean time on death row is 15 years.
```

Introduction:

In this report, we will conduct a hypothesis test to determine if the meantime on death row for a population of inmates is likely to be 15 years. We have a sample of 75 death row inmates, and we know the mean sample length of time on death row is 17.4 years with a sample standard deviation of 6.3 years. We will use a hypothesis test to determine if there is enough evidence to reject the null hypothesis that the mean time on death row is 15 years.

Hypothesis:

Given $\mu_0 = 15$

Let H_0 = The population's mean time on death row is 15 years.

H_1 = The population's mean time on death row is not 15 years.

Therefore, our hypothesis becomes:
$H_0 : \mu = \mu_0$ v/s $H_1 : \mu \neq \mu_0$

<u>Testing:</u>
We will use a two-tailed t-test to test our hypothesis at a 5% significance level. The test statistic is calculated as follows: (here we have used t test)
Test = (sample_mean - μ0)/(sample_standard_deviation/sqrt(n))
where n = sample size, and μ0 = 15.

<u>Results:</u>
Using the given values, we find that the sample mean length of time on death row is 17.4 years, and the sample standard deviation is 6.3 years. We calculate the test statistic as:
Test statistic: 3.299144
Left critical value: -1.992543
Right critical value: 1.992543
p-value: 0.001493164

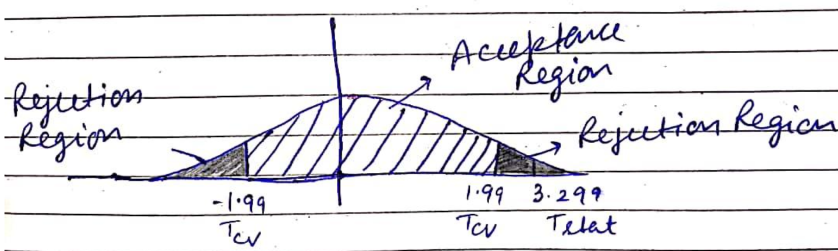
<u>Inference:</u>
Since the calculated test statistic is greater than the positive critical value, we reject the null hypothesis.

<u>Conclusion</u>:
Based on the results of our hypothesis test, we can conclude **Reject the null hypothesis that the population mean time on death row is 15 years.**

Here Our T stat = 3.299

Using the T-table $t_{0.05} = 1.99$ (T_{cv})

So our 2 Tailed graph



Ans 3)

```
> rm(list=ls())
>
> ## Sample size
>
> n<-10
>
> ## Data
> x <- c(14.3, 12.6, 13.7, 10.9, 13.7, 12.0, 11.4, 12.0, 12.6, 13.1)
> alpha <- 0.05
>
> sample_mean<-mean(x)
> sample_standard_deviation <- sd(x)
> mu<-12
> t_stat <- (sample_mean - mu)/(sample_standard_deviation/sqrt(n))
>
> # Calculate the degrees of freedom
> degree <- n - 1
>
> p_value <- 2 * pt(-abs(t_stat), df = degree)
>
>
> # Calculate the critical values from t-distribution
> left_critical_value <- qt(alpha/2, df = degree, lower.tail = TRUE)
> right_critical_value <- qt(alpha/2, df = degree, lower.tail = FALSE)
>
> # Print the results
> cat("Test statistic:", t_stat, "\n")
Test statistic: 1.835644
> cat("P value:", p_value, "\n")
P value: 0.09959876
>
> cat("Left critical value:", left_critical_value, "\n")
Left critical value: -2.262157
> cat("Right critical value:", right_critical_value, "\n")
Right critical value: 2.262157
>
> # Compare the test statistic with critical values and p-value with alpha
> if ((t_stat < left_critical_value | t_stat > right_critical_value) | (p_value < alpha)) {
+   cat("Reject the null hypothesis that the mean yield is 12.0 quintals per hectare.\n")
+ } else {
+   cat("Fail to reject the null hypothesis that the mean yield is 12.0 quintals per hectare.\n")
+ }
Fail to reject the null hypothesis that the mean yield is 12.0 quintals per hectare.
```

<u>Hypothesis:</u>

Given μ_0 = 12.0 (expected yield)

Let H0 = The new variety of green gram does not give a significantly different yield compared to the expected yield of 12.0 quintals per hectare.

H1 = The new variety of green gram gives a significantly different yield compared to the expected yield of 12.0 quintals per hectare.

Therefore, our hypothesis becomes:

H0: $\mu = \mu_0$ v/s H1: $\mu \neq \mu_0$

Testing:

We used a two-tailed t-test to test our hypothesis.
We will make our inferences using 2 ways, by test statistic and p-value
The test statistic is calculated as below:
t_stat <- (sample_mean - mu)/(sample_standard_deviation/sqrt(n))

From this, we get our p-value as:
p_value <- 2 * pt(-abs(test_stat), df = degree)

Results:

We conducted an experiment on 10 randomly selected farmers' fields to test the hypothesis. The recorded yields (quintals/hectare) were as follows: 14.3, 12.6, 13.7, 10.9, 13.7, 12.0, 11.4, 12.0, 12.6, 13.1.
Test statistic: 1.835644
P value: 0.09959876
Left critical value: -2.262157
Right critical value: 2.262157

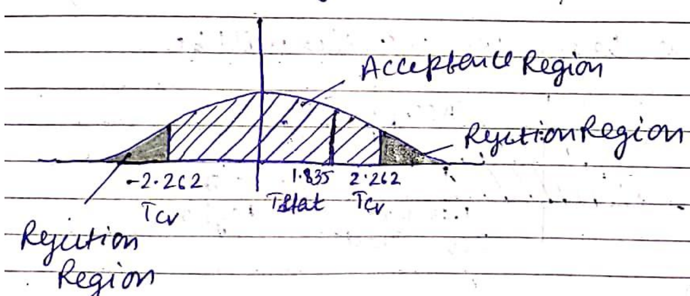
Inference:

As we can see, test_stat > left_critical_value and test_stat < right_critical_value at the 5% significance level, and the p-value is greater than 0.05. Therefore, we fail to reject the null hypothesis.

Here Tstat = 1.835

$t_{0.05}$ = 2.262 (Tcv)

Our 2-Tailed graph.



Ans 4)

```
> rm(list = ls())
>
> ## Here H0 => mean time boys  spend  = mean time girls  spend on playing sports
> ## vs
> ## H1 => mean time boys  spend != mean time girls  spend on playing sports
>
> n_boys <- 16
> n_girls <- 9
> sample_mean_boys <- 3.2
> sample_mean_girls <- 2
> sample_var_boys <- 1
> sample_var_girls <- 0.75
> alpha <- 0.05
> df <- min(n_boys -1 ,  n_girls - 1)
>
> # Calculate the test statistic and critical values
> test_stat <- (sample_mean_boys - sample_mean_girls) / sqrt((sample_var_boys/n_boys) + (sample_var_girls/n_girls))
> left_critical_value <- qt(p = alpha/2, df = df, lower.tail = TRUE)
> right_critical_value <- qt(p = alpha/2, df = df, lower.tail = FALSE)
>
> # Calculate the p-value
> p_value <- 2 * pt(q = abs(test_stat), df = df, lower.tail = FALSE)
>
> # Print the results
> cat("Test Statistics: ", test_stat, "\n")
Test Statistics:  3.142338
> cat("Critical Values: Left: ", left_critical_value, ", Right: ", right_critical_value, "\n")
Critical Values: Left:  -2.306004 , Right:  2.306004
> cat("P-Value: ", p_value, "\n")
P-Value:  0.01375661
>
> # Test the hypothesis
> if ((test_stat < left_critical_value | test_stat > right_critical_value) | (p_value < alpha)) {
+   cat("Reject the null hypothesis. The mean amount of time spent playing sports per day is different between boys and girls.\n")
+ } else {
+   cat("Fail to reject the null hypothesis. \n")
+ }
Reject the null hypothesis. The mean amount of time spent playing sports per day is different between boys and girls.
```

<u>Hypothesis:</u>

Let H0 be the null hypothesis that the mean amount of time spent playing sports per day is the same for boys and girls, and H1 be the alternative hypothesis that the mean amount of time spent playing sports per day is different for boys and girls. We can express the hypothesis as:

H0: $\mu\_boys = \mu\_girls$

H1: $\mu\_boys \neq \mu\_girls$

<u>Testing:</u>

We will use a two-sample t-test assuming equal variances to test our hypothesis. We will make our inferences using both the test statistic and the p-value.

The test statistic is calculated as follows:

test_stat <- (sample_mean_boys - sample_mean_girls) / sqrt((sample_var_boys/n_boys) + (sample_var_girls/n_girls))

```
left_critical_value <- qt(p = alpha/2, df = df, lower.tail = TRUE)
right_critical_value <- qt(p = alpha/2, df = df, lower.tail = FALSE)
p_value <- 2 * pt(q = abs(test_stat), df = df, lower.tail = FALSE)
```

<u>Results:</u>
Using the sample data provided, we can calculate the test statistic, critical value, and p-value as follows:

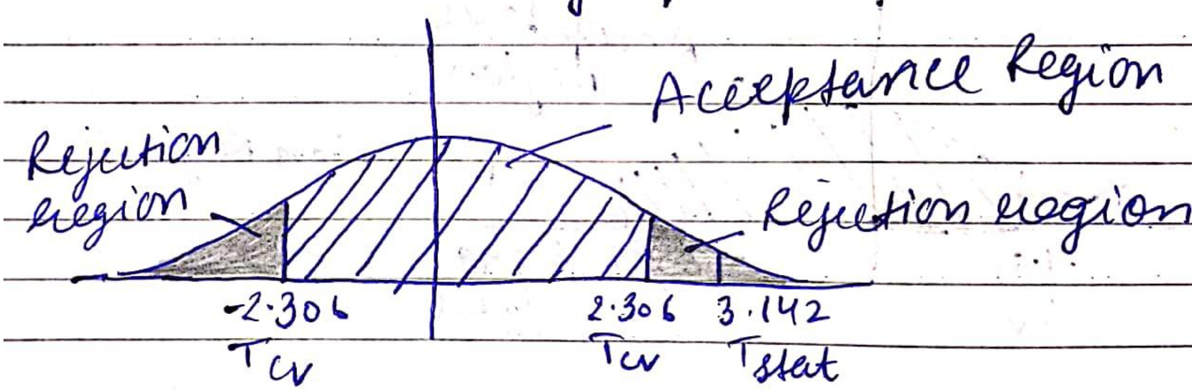
Test Statistics: 3.142338
Critical Values: Left: -2.306004 , Right: 2.306004
P-Value: 0.01375661

<u>Inference:</u>
As test_stat > right_critical_value and p_value < alpha. Therefore Reject the null hypothesis. The mean time spent playing sports per day differs between boys and girls.

$T_{stat} = 3.142$

$t_{0.05} = 2.306 \ (T_{cv})$

Our 2-Tailed graph



Ans 5)

```
> rm(list=ls())
> # H0 : mu of difference of pairs = 0 vs H1 : mu of difference of pairs != 0
> # Data
> food_A <- c(49, 53, 51, 52, 47, 50, 52, 53)
> food_B <- c(52, 55, 52, 53, 50, 54, 54, 53)
> n <- 8
>
> # Calculate the difference in weight for each child
> diff_weight <- food_B - food_A
>
> # Calculate the sample mean and standard deviation of the difference
> sample_mean <- mean(diff_weight)
> sample_sd <- sd(diff_weight)
>
> # Set the significance level
> alpha <- 0.05
>
> # Calculate the test statistic and p-value
> test_stat <- (sample_mean - 0) / (sample_sd / sqrt(length(diff_weight)))
> p_value <- 2 * pt(abs(test_stat), df = length(diff_weight) - 1, lower.tail = FALSE)
> left_critical_value <- qt(p = alpha/2,df=n-1,lower.tail = TRUE)
> right_critical_value <- qt(p = alpha/2,df=n-1,lower.tail = FALSE)
>
> # Print the results
> cat ("Sample Mean: ",sample_mean,"\n" )
Sample Mean:  2
> cat ("Sample Sd: ",sample_sd,"\n" )
Sample Sd:  1.309307
> cat("Test Statistics: ", test_stat, "\n")
Test Statistics:  4.320494
> cat("Critical Values: Left: ", left_critical_value, ", Right: ", right_critical_value, "\n")
Critical Values: Left:  -2.364624 , Right:  2.364624
> cat("P-Value: ", p_value, "\n")
P-Value:  0.003478084
>
> # Test the hypothesis
> if ((test_stat < left_critical_value) | (test_stat > right_critical_value) | (p_value < alpha)) {
+   cat("Reject the null hypothesis. There is a significant difference in the average change in weight of children due to Food B.\n")
+ } else {
+   cat("Fail to reject the null hypothesis. There is no significant difference in the average change in weight of children due to Food B.\n")
+ }
Reject the null hypothesis. There is a significant difference in the average change in weight of children due to Food B.
```

<u>Hypothesis:</u>

Let H0 be the null hypothesis that there is no difference in the average change in weight of children due to Food B, and H1 be the alternative hypothesis that there is a difference in the average change in weight of children due to Food B.

<u>Testing:</u>

We will use a paired two-sample t-test to test our hypothesis. We will make our inferences using both the test statistic and the p-value.

The test statistic is calculated as follows:

test_stat <- (sample_mean - 0) / (sample_sd / sqrt(length(diff_weight)))

p_value <- 2 * pt(abs(test_stat), df = length(diff_weight) - 1, lower.tail = FALSE)

<u>Results:</u>

Using the sample data provided, we can calculate the test statistic, critical value, and p-value as follows:

Test Statistics: 4.320494

Critical Values: Left: -2.364624 , Right: 2.364624
P-Value: 0.003478084

<u>Inference:</u>
Here as test_stat > right_critical_value, Therefore we Reject the null hypothesis. There is a significant difference in children's average weight change due to Food B.

$$T_{stat} = \frac{\bar{x} - \mu_x}{S_x/\sqrt{n}}$$

$$S_x = 1.309$$

$$T_{stat} = \frac{2}{1.309}\sqrt{8} = 4.32$$

$$T_{cv} = 2.36$$

Our 2 Tailed graph