ASDS 5301: Project 2

Administration of a drug was suspected to cause reduced alkaline phosphatase levels in adult males, a population known to have an alkaline phosphatase mean (µ) of 60 with a standard deviation (σ) of 15. To test this, a sample of 100 men were treated and the data on their alkaline phosphatase levels are presented in the file “alkaline.txt”. Write a report by answering the following questions:

1. Import the alkaline data as a data frame by using a suitable R/SAS function

• Converted the given data set from Text(.txt) to excel(.xlsx) format to make the data clear.

•

• To read excel data in R, “readxl” package need to be installed

• Reading the installed packages using library() function.

#converted the given csv file in to excel file  
# Import alkaline data as a data frame  
#install readxl package from packages  
library(readxl)  
data <- read\_excel("alkaline.xlsx")  
head(data)

#Using “head()” function in R we can see the first 6 rows of the data

#To view entire data we can use “View()” function

## # A tibble: 6 × 2  
## patient alkaline  
## <dbl> <dbl>  
## 1 1 64.0  
## 2 2 62.9  
## 3 3 55.5  
## 4 4 79.7  
## 5 5 58.8  
## 6 6 78.1

1. State the null and alternate hypotheses for the testing under consideration

The null and alternate hypothesis are :

• Null Hypothesis: The null hypothesis assumes that there is no effect of the drug, and the population mean alkaline phosphatase level remains at the known population mean of 60.

• Alternate Hypothesis: The alternative hypothesis suggests that the drug causes a reduction in alkaline phosphatase levels in adult males, leading to a population mean that is less than 60.

This is a one-sided test, as we are only interested in a decrease in alkaline phosphatase levels.

1. Test the hypothesis at 5% level of significance using a suitable R/SAS function. Use both p-value method and critical value method. Draw appropriate conclusions.

**P-Value Method**:

# Given mean is 60  
mean <- 60  
# Given standard deviation is 15  
std <- 15  
# Sample statistics  
sp\_mean <- mean(data$alkaline)  
cat("sample mean of data is ",sp\_mean,"\n")

## sample mean of data is 62

sp\_sd <- sd(data$alkaline)  
cat(" sample standad deviation of data is ",sp\_sd,"\n")

## sample standad deviation of data is 12.58917

# For null hypothesis  
cat("According to Null hypothes,the true alkaline phosphate mean is less than or equal to ", mean,"\n")

## According to Null hypothes,the true alkaline phosphate mean is less than or equal to 60

# For alternate hypothesis  
cat("According to Alternate hypothes,the true alkaline phosphate mean is greater than ", mean,"\n")

## According to Alternate hypothes,the true alkaline phosphate mean is greater than 60

#Perform Z Test  
#we need to install asbio package  
library(asbio)

## Loading required package: tcltk

library(BSDA)

## Loading required package: lattice

##   
## Attaching package: 'BSDA'

## The following object is masked from 'package:datasets':  
##   
## Orange

z\_test <- z.test(data$alkaline , alternative = "less", mu = mean, sigma.x = std, conf.level = 0.95)  
print(z\_test)

##   
## One-sample z-Test  
##   
## data: data$alkaline  
## z = 1.3333, p-value = 0.9088  
## alternative hypothesis: true mean is less than 60  
## 95 percent confidence interval:  
## NA 64.46728  
## sample estimates:  
## mean of x   
## 62

# Declare alpha  
alpha = 0.05  
# Extract p value  
p\_value = z\_test$p.value  
# Compare p-value with critical value  
if(p\_value < alpha) {  
 cat("Reject Null Hypothesis: There is significant evidence that the drug affects alkaline phosphatase levels","\n")  
} else {  
 cat("Fail to reject Null Hypothesis: There is no significant evidence that the drug affects alkaline phosphatase levels","\n")  
}

## Fail to reject Null Hypothesis: There is no significant evidence that the drug affects alkaline phosphatase levels

# Calculate critical value  
critical\_value <- qt(alpha, length(data$alkaline) - 1, lower.tail = TRUE)  
cat("Critical value is :",critical\_value,"\n")

## Critical value is : -1.660391

# Display p-value  
cat("P-value:", p\_value,"\n")

## P-value: 0.9087884

Description for P value: As the p-value is greater than the significance level of 0.05. Therefore, we fail to reject the null hypothesis. This implies that there is insufficient evidence to conclude that the drug causes a reduction in alkaline phosphatase levels in adult males.

Description for critical value: Additionally, since the critical value is -1.660391 and the test statistic from the data is not less than this critical value, we again fail to reject the null hypothesis. Hence, we can conclude that we do not have enough evidence to suggest that the drug causes a reduction in alkaline phosphatase levels in adult males.

In summary, based on the p-value method and a significance level of 0.05, there is no significant evidence to support the alternate hypothesis that the drug affects alkaline phosphatase levels to a degree where the mean becomes greater than 60. Therefore, the null hypothesis is not rejected.

(iv ) Calculate the power of the test for the alternative µ = 55.

# Power using Z test  
# To perform power.z.test() we need to install asbio package  
  
power <- power.z.test(sigma = std, n = length(data$alkaline), alpha = alpha, effect = mean - 55,   
 test = "one.tail", strict = FALSE)$power  
cat("Power of the test for µ = 55:", power)

## Power of the test for µ = 55: 0.9543404

**Description**:

In summary, the calculated power of the one-tailed Z-test is quite high, indicating that the test is likely to correctly identify a true difference in means if the true mean alkaline phosphate level is 55, with a 95.4% probability. This high power value is a positive aspect of the statistical test, as it suggests that the test is sensitive enough to detect meaningful effects.

1. Plot the power curve of the test for a suitable set of alternatives. Display the power curve and comment.

**Create Sequence of Alternative Means:**

alternative\_means <- seq(45,75, by = 0.1): This line generates a sequence of alternative means ranging from 45 to 75 in increments of 0.1. These alternative means represent the values that you want to test against the null hypothesis to observe how statistical power varies as the true mean changes.

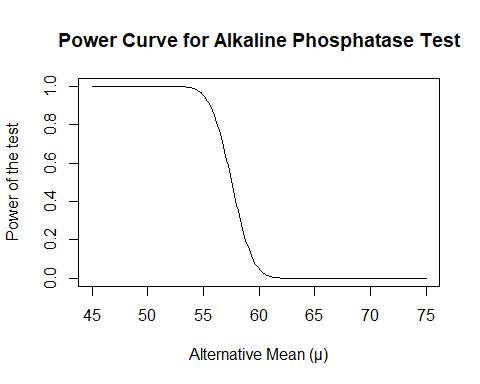
**Calculate Power for Each Alternative Mean:**

power\_values <- sapply(alternative\_means, function(mu) {...}): This code uses sapply to iterate through the sequence of alternative means and calculates the statistical power for each alternative mean using the power.z.test() function. It stores the power values in the power\_values vector.

**Plot the Power Curve:**

plot(alternative\_means, power\_values, type = "l", ...): This line creates a line plot of the alternative means on the x-axis and their corresponding power values on the y-axis. It labels the x and y axes appropriately and provides a title for the plot.

# Create sequence of alternative means  
alternative\_means <- seq(45,75, by = 0.1)  
# Calculate power for each alternative mean  
power\_values <- sapply(alternative\_means, function(mu) {  
 power.z.test(sigma = std, n = length(data$alkaline), alpha = alpha, effect = mean - mu,   
 test = "one.tail", strict = FALSE)$power  
})  
# Plot the power curve  
plot(alternative\_means, power\_values, type = "l",   
 xlab = "Alternative Mean (μ)", ylab = "Power of the test",   
 main = "Power Curve for Alkaline Phosphatase Test")



**Description**: The code generates a power curve that illustrates how the statistical power of a one sample z test changes as the population mean varies over a specifies range of alternative means. It helps assess the test’s sensitivity to different effect sizes. The power of the test reduces continuously after 55 till 60, it becomes 0 at 60 and then becomes uniform after 60 till 70. In summary, this code segment is a good practice for exploring the statistical power of your test across a range of alternative scenarios, providing insights into the test's performance under different conditions.

1. Calculate the sample size required to achieve a power of 80%, 85% and 90%. What can you conclude from your findings.

# Calculate sample sizes for desired powers  
samplesizes = sapply(c(0.80, 0.85, 0.90), function(power) {  
 power.z.test(sigma = std, power = power, alpha = alpha, effect = mean - 55,   
 test = "one.tail", strict = FALSE)$n  
})  
cat("Sample sizes required for 80% is ", round(samplesizes[1]),"\n")

## Sample sizes required for 80% is 56

cat("Sample sizes required for 85% is ", round(samplesizes[2]),"\n")

## Sample sizes required for 85% is 65

cat("Sample sizes required for 90% is ", round(samplesizes[3]),"\n")

## Sample sizes required for 90% is 77

**Description**: To achieve a power of 80% for detecting a difference in mean alkaline phosphatase levels with the drug (assuming a reduction), a sample size of 56 is needed. This indicates that if there is a true difference in means, there is an 80% chance of detecting it. If we want a higher power of 85%, a larger sample size of 65 is required. A larger sample size increases the likelihood of detecting a true difference if it exists. Finally, to achieve a very high power of 90%, an even larger sample size of 77 is needed. This indicates a strong ability to detect a true difference in means if it exists.

1. Now, assume that administration of the drug is suspected to alter the alkaline phosphatase level
2. Repeat parts (ii) and (iii) and draw proper conclusions in each case

# Repeat parts (ii) and (iii)  
# For null hypothesis  
cat("According to Null hypothes the true alkaline phosphate mean is ", mean,"\n")

## According to Null hypothes the true alkaline phosphate mean is 60

# For alternate hypothesis  
cat("According to Alternate hypothes the true alkaline phosphate mean is not equal to ", mean,"\n")

## According to Alternate hypothes the true alkaline phosphate mean is not equal to 60

Null & Alternate Hypothesis: The null and alternate hypothesis are:

• Null Hypothesis: The mean alkaline phosphatase level after drug administration is 60.

• Alternate Hypothesis: The mean alkaline phosphatase level after drug administration is not equal to 60.

# Perform Z Test  
# To perform power.z.test() we need asbio package installed  
ztest2 <- z.test(data$alkaline, alternative = "two.sided", mu = mean, sigma.x = std, conf.level = 0.95)  
print(ztest2)

##   
## One-sample z-Test  
##   
## data: data$alkaline  
## z = 1.3333, p-value = 0.1824  
## alternative hypothesis: true mean is not equal to 60  
## 95 percent confidence interval:  
## 59.06005 64.93994  
## sample estimates:  
## mean of x   
## 62

# Extract p value  
p\_value2 = ztest2$p.value  
# Compare p-value with critical value  
if(p\_value2 < alpha) {  
 cat("Reject Null Hypothesis: There is significant evidence that the drug affects alkaline phosphatase levels.","\n")  
} else {  
 cat("Fail to reject Null Hypothesis: There is no significant evidence that the drug affects alkaline phosphatase levels.","\n")  
}

## Fail to reject Null Hypothesis: There is no significant evidence that the drug affects alkaline phosphatase levels.

# Calculate critical value  
critical\_value2 <- qt(alpha, length(data$alkaline) - 1, lower.tail = FALSE)  
  
# Display p-value and confidence interval  
cat("P-value:", p\_value2,"\n")

## P-value: 0.1824232

cat("Critical value:", critical\_value2,"\n")

## Critical value: 1.660391

Hypothesis Testing: As the p-value is greater than the significance level of 0.05 therefore, we fail to reject the null hypothesis. This suggests that there is insufficient evidence to conclude that the mean alkaline phosphatase level after drug administration is different from 60.

The test statistic is between the critical lower and upper t-values for a two-tailed test at a significance level of 0.05. This confirms the result from the p-value that we fail to reject the null hypothesis.

Hence, we do not have enough evidence to suggest that the mean alkaline phosphatase level after drug administration is different from 60 at a 5% significance level.

1. Calculate the power of the test for the alternatives µ = 65 and µ = 55, separately.

Power of Test For Alternative 65 and 55

# Calculate power for μ = 65  
power\_65 <- power.z.test(sigma = std, n = length(data$alkaline), alpha = alpha, effect = mean - 65,   
 test = "two.tail", strict = TRUE)$power  
  
# Calculate power for μ = 55  
power\_55 <- power.z.test(sigma = std, n = length(data$alkaline), alpha = alpha, effect = mean - 55,   
 test = "two.tail", strict = TRUE)$power  
  
cat("Power of the test for µ = 65:", power\_65,"\n")

## Power of the test for µ = 65: 0.9151813

cat("Power of the test for μ = 55:", power\_55,"\n")

## Power of the test for μ = 55: 0.9151813

**Description**:

The code calculates and displays the statistical power for two different values of the true mean, 65 and 55.

It is essential to calculate the power of a test for different true mean values to understand how well the test can detect deviations from the null hypothesis.

The high power values for both μ = 65 and μ = 55 (approximately 0.915) suggest that the test has a good chance of correctly identifying a difference from the null hypothesis, whether the true mean is 65 or 55.

In summary, this code segment is valuable for assessing the test's ability to detect differences in means when considering two different alternative mean values and provides insight into the test's sensitivity to variations in the true mean.

1. Repeat part (vi) and draw proper conclusion.

# Repeat part (vi)  
# Calculate sample sizes for desired powers  
samplesizes2 = sapply(c(0.80, 0.85, 0.90), function(power) {  
 power.z.test(sigma = std, power = power, alpha = alpha, effect = mean - 55,   
 test = "two.tail", strict = TRUE)$n  
})  
cat("Sample sizes required for 80% is ", round(samplesizes2[1]),"\n")

## Sample sizes required for 80% is 71

cat("Sample sizes required for 85% is ", round(samplesizes2[2]),"\n")

## Sample sizes required for 85% is 81

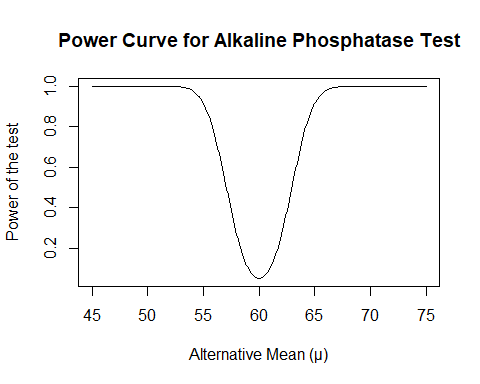
cat("Sample sizes required for 90% is ", round(samplesizes2[3]),"\n")

## Sample sizes required for 90% is 95

**Description**: The provided sample sizes demonstrate a progression in statistical power: To achieve a power of 80% for detecting a difference in mean alkaline phosphatase levels with the drug (assuming a reduction), a sample size of 56 is needed. This indicates that if there is a true difference in means, there is an 80% chance of detecting it. If we want a higher power of 85%, a larger sample size of 65 is required.from a good power of 80% to a higher power of 85%, and further to a very high power of 90%. A larger sample size generally enhances the ability to detect effects, making the study more robust and reliable in drawing conclusions about the population.

1. Draw the power curve of the test for the alternativesµ = (45, 45.1, · · · , 59.9, 60.1, 60.2, · · · , 75) and comment

# Create sequence of alternative means  
alternative\_means2 <- seq(45,75, by = 0.1)  
  
# Calculate power for each alternative mean  
power\_values2 <- sapply(alternative\_means2, function(mu) {  
 power.z.test(sigma = std, n = length(data$alkaline), alpha = alpha, effect = mean - mu,   
 test = "two.tail", strict = TRUE)$power  
})  
  
# Plot the power curve  
plot(alternative\_means2, power\_values2, type = "l",   
 xlab = "Alternative Mean (μ)", ylab = "Power of the test",   
 main = "Power Curve for Alkaline Phosphatase Test")



**Description**:The code generates a power curve that illustrates how the statistical power of a one sample z test changes as the population mean(alternative mean) 45 to 75 in incriments of 0.1 .This curve helps assess the test’s ability to detect differnces from a null hypothesis mean of 60,showing that power decreases as the alternative mean approches 60 and then rises as it goes beyond 60.

**Source** **Code**: “Using R software”

# Import alkaline data as a data frame

#install readxl package from packages

library(readxl) data <- read\_excel(“alkaline.xlsx”)

head(data)

# Given mean is 60

mean <- 60

# Given standard deviation is 15

std <- 15

# Sample statistics

sp\_mean <- mean(data$alkaline)

cat("sample mean of data is ",sp\_mean,"\n")

# For null hypothesis

cat("According to Null hypothes,the true alkaline phosphate mean is less than or equal to ", mean,"\n")

# For alternate hypothesis

cat("According to Alternate hypothes,the true alkaline phosphate mean is greater than ", mean,"\n")

#Perform Z Test

#we need to install asbio package

library(asbio)

library(BSDA)

z\_test <- z.test(data$alkaline , alternative = “less”, mu = mean, sigma.x = std, conf.level = 0.95) print(z\_test)

# Declare alpha alpha = 0.05

# Extract p value p\_value = z\_test$p.value

# Compare p-value with critical value

if(p\_value < alpha)

{

cat("Reject Null Hypothesis: There is significant evidence that the drug affects alkaline phosphatase levels","\n") }

else {

cat("Fail to reject Null Hypothesis: There is no significant evidence that the drug affects alkaline phosphatase levels","\n")

}

# Calculate critical value

critical\_value <- qt(alpha, length(data$alkaline) - 1, lower.tail = TRUE) cat(“Critical value is :”,critical\_value,“”) # Display p-value cat(“P-value:”, p\_value,“”) # Power using Z test # To perform power.z.test() we need to install asbio package

power <- power.z.test(sigma = std, n = length(data$alkaline), alpha = alpha, effect = mean - 55, test = "one.tail", strict = FALSE)$power cat(“Power of the test for µ = 55:”, power) # Create sequence of alternative means alternative\_means <- seq(45,75, by = 0.1) # Calculate power for each alternative mean power\_values <- sapply(alternative\_means, function(mu) { power.z.test(sigma = std, n = length(data$alkaline), alpha = alpha, effect = mean - mu, test = "one.tail", strict = FALSE)$power }) # Plot the power curve plot(alternative\_means, power\_values, type = “l”, xlab = “Alternative Mean (μ)”, ylab = “Power of the test”, main = “Power Curve for Alkaline Phosphatase Test”) # Calculate sample sizes for desired powers samplesizes = sapply(c(0.80, 0.85, 0.90), function(power) { power.z.test(sigma = std, power = power, alpha = alpha, effect = mean - 55, test = “one.tail”, strict = FALSE)$n }) cat("Sample sizes required for 80% is ", round(samplesizes[1]),"\n") cat("Sample sizes required for 85% is ", round(samplesizes[2]),"\n") cat("Sample sizes required for 90% is ", round(samplesizes[3]),"\n") # Repeat parts (ii) and (iii) # For null hypothesis cat("According to Null hypothes the true alkaline phosphate mean is ", mean,"\n") # For alternate hypothesis cat("According to Alternate hypothes the true alkaline phosphate mean is not equal to ", mean,"\n") # Perform Z Test # To perform power.z.test() we need asbio package installed ztest2 <- z.test(data$alkaline, alternative = “two.sided”, mu = mean, sigma.x = std, conf.level = 0.95) print(ztest2) # Extract p value p\_value2 = ztest2$p.value # Compare p-value with critical value if(p\_value2 < alpha) { cat("Reject Null Hypothesis: There is significant evidence that the drug affects alkaline phosphatase levels.","\n") } else { cat("Fail to reject Null Hypothesis: There is no significant evidence that the drug affects alkaline phosphatase levels.","\n") } # Calculate critical value critical\_value2 <- qt(alpha, length(data$alkaline) - 1, lower.tail = FALSE)

# Display p-value and confidence interval

cat(“P-value:”, p\_value2,“”) cat(“Critical value:”, critical\_value2,“”) # Calculate power for μ = 65 power\_65 <- power.z.test(sigma = std, n = length(data$alkaline), alpha = alpha, effect = mean - 65, test = "two.tail", strict = TRUE)$power

# Calculate power for μ = 55

power\_55 <- power.z.test(sigma = std, n = length(data$alkaline), alpha = alpha, effect = mean - 55, test = "two.tail", strict = TRUE)$power

cat(“Power of the test for µ = 65:”, power\_65,“”) cat(“Power of the test for μ = 55:”, power\_55,“”) # Repeat part (vi) # Calculate sample sizes for desired powers samplesizes2 = sapply(c(0.80, 0.85, 0.90), function(power) { power.z.test(sigma = std, power = power, alpha = alpha, effect = mean - 55, test = “two.tail”, strict = TRUE)$n }) cat("Sample sizes required for 80% is ", round(samplesizes2[1]),"\n") cat("Sample sizes required for 85% is ", round(samplesizes2[2]),"\n") cat("Sample sizes required for 90% is ", round(samplesizes2[3]),"\n") # Create sequence of alternative means alternative\_means2 <- seq(45,75, by = 0.1) # Calculate power for each alternative mean power\_values2 <- sapply(alternative\_means2, function(mu) { power.z.test(sigma = std, n = length(data$alkaline), alpha = alpha, effect = mean - mu, test = “two.tail”, strict = TRUE)$power }) # Plot the power curve plot(alternative\_means2, power\_values2, type = “l”, xlab = “Alternative Mean (μ)”, ylab = “Power of the test”, main = “Power Curve for Alkaline Phosphatase Test”)