**DATA 1204 Statistical and Predictive Modeling for Analytics DATA 1204**

**Assignment 2-Hypothesis Testing**

**Srilasya Garigipaty Student #100822953**

**1.State the Hypothesis Statement that helps solve the research question:**

**For two -tail test, hypothesis statement is:**

After exercise, the blood pressure of the population will change from the average of 138.28 before exercise.

Hypothesis Testing

Ho: μ = 138.28 (null hypothesis)

Ha: μ ≠ 138.28 (alternative hypothesis)

**For one-tail test, the hypothesis statement is:**

After exercise, the blood pressure of the population will be greater than the average of 138.28 before exercise. Hypothesis Testing Ho: μ = 138.28 (null hypothesis) Ha: μ > 138.28(alternative hypothesis)

**2.Provide a written *step-by-step outline* on how you would prove (or dis-prove) your hypothesis statements that you developed above.**

**Types of Hypothesis Tests**:

**One-Tail Test**- Test to see if the mean or some other population parameter is greater than or smaller than the null value.

**Two-Tail Test-** Test to see if there is any difference from the null value.

Both tests are done by evaluating the evidence against the null hypothesis and checking the p-values against significance level.

**P-value-** measure that quantifies the strength of the evidence made about a population parameter. It is the strength of the evidence against null hypothesis and in favor of the alternative hypothesis. A p-value is compared relative to the value of significance level. If p-value is greater than or less than significance level the null hypothesis is rejected.

**Steps in Conducting a Hypothesis Test:**

The following are the steps in conducting a Hypothesis test using a z-test and p-value approach:

1. **State the null and alternative hypothesis Ha**



1. **Specify a significance level α**

**3. Compute the test statistic *z***

****

population has a **mean μ, standard deviation σ**.

Sample size is **n**. Sample mean 

**4. Assuming the null is true and decide by:**

o Comparing the test statistic *z* with the critical value or values.

o Comparing the p-value with the significance value α

**5. State conclusion**

**3.Conduct the analysis you outlined in #2 in R**

**Two Tail Test:**

#intall library

#Install openintro

#load library

library(openintro)

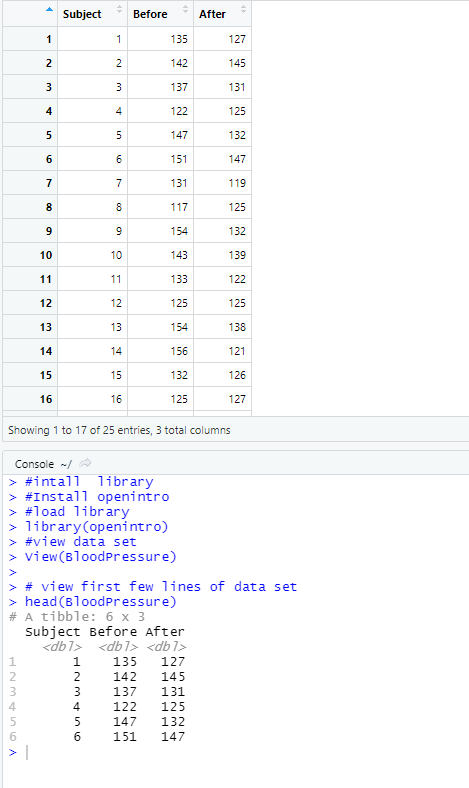
#view data set

View(BloodPressure)

# view first few lines of data set

head(BloodPressure)

**Results:**



x<- BloodPressure$After

#mean of sample

mean.x<-mean(x)

mean.x

#standard Deviation of sample

sd.x<-sd(x)

sd.x

#Standard Error of sample

SE.x <-sd(x)/sqrt(length(x))

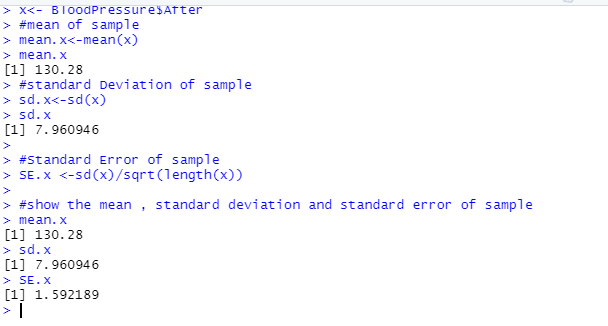
#show the mean , standard deviation and standard error of sample

mean.x

sd.x

SE.x

**Results:**



#The Before exercise blood pressure readings were taken as the population readings

p<- BloodPressure$Before

#mean of population

mean.p<-mean(p)

mean.p

#standard Deviation of population

sd.p<-sd(p)

sd.p

#Standard Error of population

SE.p <-sd(p)/sqrt(length(p))

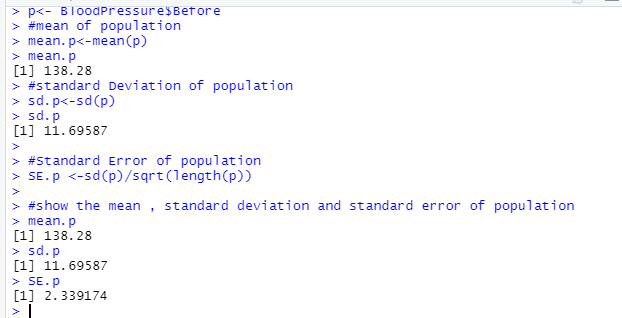
#show the mean , standard deviation and standard error of population

mean.p

sd.p

SE.p

**Results:**



#Histogram of Blood Pressure after Exercise

x=BloodPressure$After

h<-hist(x, breaks=5, col="red", xlab="After",

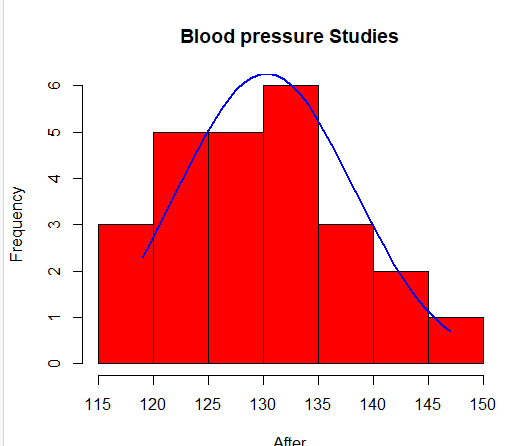
main="Blood pressure Studies")

xfit<-seq(min(x),max(x),length=40)

yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))

yfit <- yfit\*diff(h$mids[1:2])\*length(x)

lines(xfit, yfit, col="blue", lwd=2)

 **normal distribution**

#Hypothesis Testing

#Step 1: Null and alternative hypotheses as follows:

#Ho: μ = μo

#Ha: μ ≠ μo

#Furthermore, since we are testing if the blood pressure average of population has changed from 138.28, we have the following setup:

#Ho: μ = 138.28

#Ha: μ ≠ 138.28

# We want to test hypothesis that the average After result is 138.28

#Step 2 Calculate z parameters

mean.p<-138.28 # Specify the mean #assuming mean of the population dataset is 138.28

alpha <- 0.05 # Specify the significance level

sigma <- 11.69587 # population standard deviation assumed same as sample standard deviation

n <- nrow(BloodPressure) # get the sample size

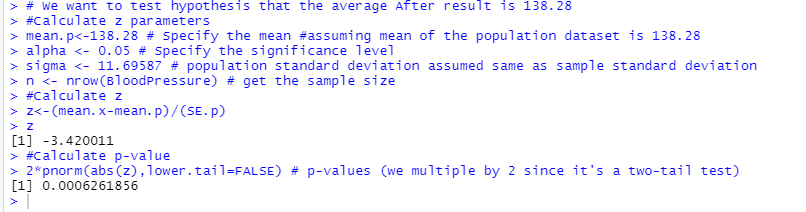
#Step 3 Calculate z

z<-(mean.x-mean.p)/(SE.p)

z

#Step 4 Calculate p-value

2\*pnorm(abs(z),lower.tail=FALSE) # p-values (we multiple by 2 since it's a two-tail test)

**Results:** 

Step 5: Since p-value of **0.000626** is much lower than 0.05 confidence interval, therefore we

reject the null hypothesis that μ = 138.28.

Conclusion: We reject the null hypothesis that the population mean (or blood pressure after exercise) is equal to 138.28 at the 0.05 level and accept the alternative that blood pressure has changed after exercise from before exercise.

**One Tail Test:**

#Hypothesis Testing

#Step 1: Null and alternative hypotheses as follows:

#Ho: μ = μo

#Ha: μ ≠ μo

#Furthermore, since we are testing if the blood pressure average of population is greater than 138.28, we have the following setup:

#Ho: μ = 138.28

#Ha: μ > 138.28

# We want to test hypothesis that the average After result is greater than 138.28

#Step 2 Calculate z parameters

mean.p<-138.28 # Specify the mean #assuming mean of the population dataset is 138.28

alpha <- 0.05 # Specify the significance level

sigma <- 11.69587 # population standard deviation assumed same as sample standard deviation

n <- nrow(BloodPressure) # get the sample size

#Step 3 Calculate z

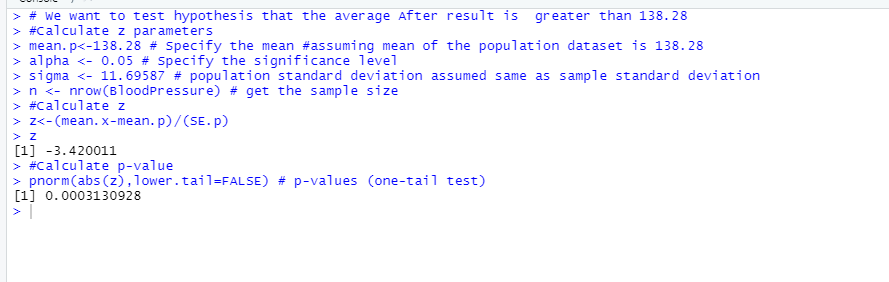
z<-(mean.x-mean.p)/(SE.p)

z

# Step 4 Calculate p-value

pnorm(abs(z),lower.tail=FALSE) # p-values (one-tail test)

**Results:**



Step 5: Since p-value of 0.0003130928 is much lower than 0.05 confidence interval, therefore we reject the null hypothesis that μ = 138.28. Conclusion: We reject the null hypothesis that the population mean (or blood pressure after exercise) is equal to 138.28 at the 0.05 level. In other words, the mean blood pressure after exercise is actually greater than 138.28.

**4.Summary of findings and conclusion**

Sample x <- Blood Pressure after exercise values

|  |  |
| --- | --- |
| mean | 130.28 |
| Standard deviation | 7.96 |
| Standard error | 1.592189 |

Population <- Blood Pressure before exercise values

|  |  |
| --- | --- |
| mean | 138.28 |
| Standard deviation | 11.69587 |
| Standard error | 2.339174 |

**Two-Tail Test:**

Z = -3.420011

p-value= 0.0006261856

Since p-value of **0.000626** is much lower than 0.05 confidence interval and lies in rejection region therefore we reject the null hypothesis that μ = 138.28.

Conclusion: We reject the null hypothesis that the population mean (or blood pressure after exercise) is equal to 138.28 at the 0.05 level and accept the alternative that blood pressure has changed after exercise from before exercise. There is enough evidence to reject null hypothesis.

**One Tail Test:**

Z=--3.420011

p-value=0.0003130928

Since p-value of 0.0003130928 is much lower than 0.05 confidence interval and lies in rejection region therefore we reject the null hypothesis that μ = 138.28. Conclusion: We reject the null hypothesis that the population mean (or blood pressure after exercise) is equal to 138.28 at the 0.05 level. In other words, the mean blood pressure after exercise is actually greater than 138.28. There is enough evidence to reject null hypothesis.