As 2 Q1

Sumanth Donthula

2022-09-19

Problem 1

Loading the file and assigning the Muscle Mass as Y and Age as X

```
Data=read.table("AS_2_Q_1_Data.txt", header = FALSE, sep = "")
Y = Data[1]$V1
X = Data[2]$V2
```

Creating a linear model and assigning the values of estimators.

```
Model=lm(Y~X)
Model
```

```
##
## Call:
## lm(formula = Y ~ X)
##
## Coefficients:
## (Intercept)            X
##      156.35        -1.19
```

```
Beta0=as.numeric(Model$coefficients["(Intercept)"])
Beta1=as.numeric(Model$coefficients["X"])
```

1.a.

Hypothesis Testing

Let's Start a two sided Null Hypothesis Testing

Assumption: H0: Beta1 is a Positive Integer Beta1>=0

Ha: Beta1 is a Negative Integer Beta1<0

Decision: As we can see the P value is 4.12*10^-19 which is far less than alpha=0.05. So we reject the null hypothesis.

Conclusion: Beta1 is a Negative Integer Beta1<0

The P value of the two sided test is 4.12*10^-19

```
SXY = sum(X *Y ) - length(X) * mean(X) * mean(Y)
SYY = sum( Y * Y) - length(Y) * (mean(Y))^2
SXX = sum( X * X) - length(X) * (mean(X))^2
```

```
RSS = SYY - Beta1 * SXY
SIGMA=(RSS/(length(X)-2))^0.5

Tobs=Beta1*(SXX)^0.5/SIGMA
Tobs
```

```
## [1] -13.19326
```

```
2*pt(q=Tobs, df=58,lower.tail=TRUE)
```

```
## [1] 4.123987e-19
```

1.b.

No even the test of Non Zero Beta0 is significant. We can not provide information on amount of muscle mass at birth of a female child because we do not have data collected in that range.

1.c.

The 95 % confidence interval of muscle mass for the woman whose ages differ by one year is Beta1(The Slope of the Linear Model). Calculating the confidence interval gives a range (-1.370545,-1.009446)

It is not necessary to know the specific ages because the difference in muscle mass of women whose age is differed by one year is Beta1 which is the slope of the regression model.

```
n=0.05
t=qt(n/2, length(X)-2, lower.tail=TRUE)

lb=Beta1+t*(SIGMA/SXX^0.5)
ub=Beta1-t*(SIGMA/SXX^0.5)
lb
```

```
## [1] -1.370545
```

```
ub
```

```
## [1] -1.009446
```