**Name: Shiva Kumar Peddapuram KSU ID: 811235874**

**Subject :- Applied Statistics student mail id:- speddapu@kent.edu**

3.3.3 Fit the multiple linear regression model with mean function E | ( ) BMI18X = + β β 0 1HT2+ + β β 2WT2 3HT9+ + β β 4WT9 5ST9 (3.27) Find σˆ and R2 . Compute the t-statistics to be used to test each of the βj to be 0 against two-sided alternatives. Explicitly state the hypotheses tested and the conclusions

**Code:-**

**##3.3.3**

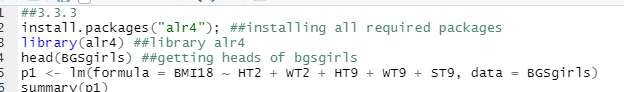
**install.packages("alr4"); ##installing all required packages**

**library(alr4) ##library alr4**

**head(BGSgirls) ##getting heads of bgsgirls**

**p1 <- lm(formula = BMI18 ~ HT2 + WT2 + HT9 + WT9 + ST9, data = BGSgirls)**

**summary(p1)**



**Output: -**

A picture containing scatter chart

Description automatically generated

**While we enter summary function, we get the following output**

Table

Description automatically generated

**Head’s function:**

WT2 HT2 WT9 HT9 LG9 ST9 WT18 HT18 LG18 ST18 BMI18 Soma

67 13.6 87.7 32.5 133.4 28.4 74 56.9 158.9 34.6 143 22.5 5.0

68 11.3 90.0 27.8 134.8 26.9 65 49.9 166.0 33.8 117 18.1 4.0

69 17.0 89.6 44.4 141.5 31.9 104 55.3 162.2 35.1 143 21.0 5.5

70 13.2 90.3 40.5 137.1 31.8 79 65.9 167.8 39.3 148 23.4 5.5

71 13.3 89.4 29.9 136.1 27.7 83 62.3 170.9 36.3 152 21.3 4.5

72 11.3 85.5 22.8 130.6 23.4 60 47.4 164.9 31.8 126 17.4 3.0

Explanation:-

The fitted mean function from this model is

**E(BMI18|X) = 30.8553+(-0.1939) \*HT2+(-0.3177) \*WT2+(0.0080) \*HT9+(0.4197) \*WT9+(-0.044) \*ST9**

**From the above output we can see that residual standard error is 2.14 on 64 degrees of freedom**

**Multiplied r-squared values is 0.4431**

***Adjusted R-squared value is 0.3996*** , **f-statistics is 10.19 on 5 and 64 DF, P-value is 3.2943-07**.

The hypothesis tested by the t-values are that each of the slope values (HT2, WT2, WT9, HT9, ST9) = 0. with other slope values arbitrary vs. slope value! = 0 with other slope values arbitrary.

Furthermore, only WT9 and ST9 have a p-value below 0.05, indicating that these are the only variables capable of explaining BMI variation.

So, from the above code we can note that we got **residuals output of min as -5.0948** **max as 10.4951 And median as -0.2533**

**Question 3.6.2: - Get the regression summary for the regression of BSAAM on these three regressors. Explain what the “t-values” column of your output means**

**Code: -**

library(alr4)

data(water)

p2<- lm(BSAAM~OPBPC+OPRC+OPSLAKE,water)

summary(p2)

Table

Description automatically generated

**output: -**

The hypothesis tested by the t-values are that each of the slope values (OPBPC, OPRC, OPSLAKE) = 0 with other slope values arbitrary vs. slope value! = 0 with other slope values arbitrary.

Here we see that all the variables except OPBPC are significant as the p-value is less than 0.05, hence we reject null hypothesis and conclude that the regression equation is significant, and the regression says there is 83% variation in BSAAM. We can also better explain the variation in response with the BSAAM