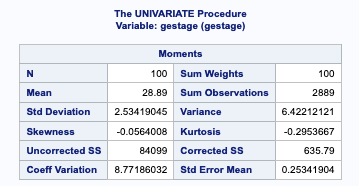
Patrick Fitzgerald

ADS 534 Statistical Modeling

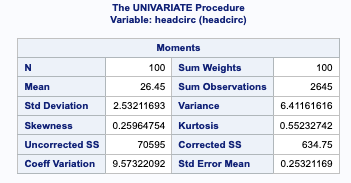
Lab # 1

Numerical Summary of the Data

**Figure 1:** Summary of Gestational Age (weeks)



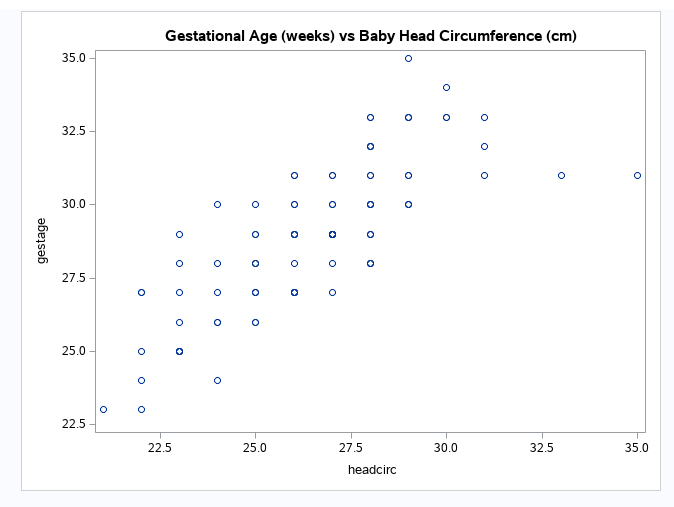
**Figure 2:** Summary of Head Circumference (cm)



Analyze the effect of **headcirc** on **gestage**

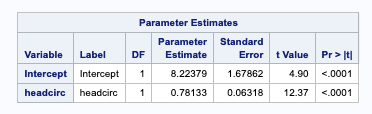
1. Scatterplot of **gestage** versus **headcirc**

**Figure 3:** Scatterplot of Gestational Age vs Head Circumference

****

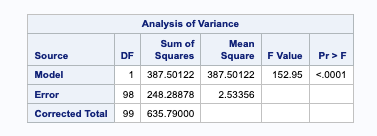
**(b)** What is the model we will fit?

**Figure 4:** Estimates of Fit-Model



1. What is the estimate for the effect of **gestage** on **headcirc**? How do you interpret this?

**Figure 5:** ANOVA table for gestage vs headcirc



Due to a very small p-value we find that it is unlikely we would have obtained such a large F statistic (Figure 5) if β1 were in fact equal to zero. Therefore, we conclude β1 ≠ 0 and there is significant evidence at the α = 0.05 level to conclude that there is a linear relationship between gestational age and head circumference (see next section question **(a)**).

Inferences based on the model

**(a)** Perform the appropriate t-test to determine if there is a significant relationship between **gestage** and **headcirc**?

*H0* : β1 = 0 and *HA* : β1 ≠ 0

Referring to **Figure 5:** ANOVA table for gestage vs headcirc we can see that we have a small p-value (< 0.0001) which indicates that we reject the null hypothesis *H0* : β1 = 0 in favor of the alternative hypothesis *HA* : β1 ≠ 0. There is significant evidence at the α = 0.05 level to conclude that there is a linear relationship between gestational age and head circumference

**(b)** How does this compare to the F-test result given in the output?

The p-values are the same because of the relationship between a t random variable and an F random variable that has 1 numerator degrees of freedom, that is to say for all simple linear regression: (t(n-2))2 = F(1, n-2).

For this example, we can see that the t-value produced in Figure 4 is equal to 12.37 and (12.37)2 = ~152.95.

**(c)** What if we were only interested in testing if increased **gestage** lead to an increase in **headcirc**? Perform this test.

**(d)** Find a two-sided 95% confidence interval for β1, the regression coefficient of **gestage**.

α = 1 – (95/100) = 0.05

Degrees of Freedom (df) = n-2 = 98

Critical Value is t-statistic having (n-2) df and a cumulative probability (p\*) = 0.975

p\* = 1 – α/2 = 0.05/2 =0.975

t = 1.66 (from Appendix B in text)

α = 0.05, p\* = 0.975, df = 98, t = 1.66

Margin of Error = critical value \* standard error

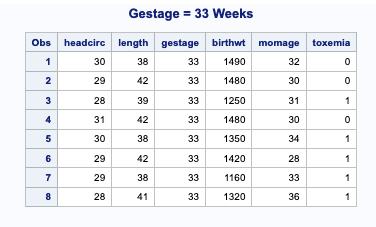
ME = 1.66 \* 0.06318 = 0.105

The 95% CI for β1 (gestage) is 0.781 ± 0.105 which is

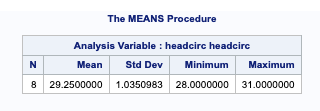
(0.676, 0.886)

**(e)** Find a two-sided 95% confidence interval of the mean value of **headcirc** for those with a **gestage** of 33 weeks.

**Figure 6:** Occurrences of gestage = 33



**Figure 7:** Mean for headcirc where gestage = 33



From the data obtained in Figures 6 & 7, I created a table and using R was able to obtain a 95% CI based on the mean head circumference (cm) for a gestational age of 33 weeks.

R-code:

> headcirc <- c(30, 29, 28, 31, 30, 29, 29, 28 )

> a <- mean(headcirc)

> b <- sd(headcirc)

> n <- 8

> error <- qt(0.975, df=n-1)\*b/sqrt(n)

> left <- a-error

> right <- a+error

> left

[1] 28.38464

> right

[1] 30.11536

So, the 95% CI for head circumference is 28.384 cm to 30.115 cm based on the acquired dataset.

**(f)** How do you interpret this interval?

We can say that we are 95% confident that the mean head circumference in centimeters will fall in the interval 28.384 cm and 30.115 cm for a gestational age of 33 weeks. There is 5% chance that the head circumference at 33-week gestational age will not fall into this interval.

**(g)** Calculate the prediction interval of **headcirc** for a future observation with **gestage** of 33 weeks.

Sample Estimate ± (t-multiplier \* SE)

From our model (Figure X.X) we can see that the predicted value for a gestational age of 33 weeks corresponds to a head circumference of 34 cm.

We can see from the same table (Figure X.X) that the standard error (SE) is 0.06318. Putting all of this together we see:

34 ± (1.66 \* 0.06318) = (33.895, 34.105)

**(h)** How do you interpret this prediction interval?

Our 95% prediction interval for a future observation with gestage of 33 weeks tells us that future head circumferences based on a 33 week gestational age will fall between 33.895 cm and 34.105 cm 95% of the time. There is 5% chance that the head circumference at 33-week gestational age will not fall into this interval.