HW7

anjal hussan

12/13/2017

7.24 Nutrition at Starbucks, Part I.

The scatterplot below shows the relationship between the number of calories and amount of carbohydrates (in grams) Starbucks food menu items contain.

Since Starbucks only lists the number of calories on the display items, we are interested in predicting the amount of carbs a menu item has based on its calorie content.

a. Describe the relationship between number of calories and amount of carbohydrates (in grams) that Starbucks food menu items contain.

Ans: As number of calories increase the carbohydrates also increases.

b. In this scenario, what are the explanatory and response variables?

Ans: Explanatory - Calories Response - Carbohydrates

In physics, we would do the opposite as energy is the effect (dependent) and carbs are a cause (independent).

c. Why might we want to fit a regression line to these data?

Ans: Customers will want to know how many Calories on average are coming from carbohydrates

7.26 Body measurements, Part III.

Exercise 7.15 introduces data on shoulder girth and height of a group of individuals. The mean shoulder girth is 107.20 cm with a standard deviation of 10.37 cm. The mean height is 171.14 cm with a standard deviation of 9.41 cm. The correlation between height and shoulder girth is 0.67.

a. Write the equation of the regression line for predicting height.

Ans:

```
eq <- (9.41 / 10.37) * .67
qq <- (eq * -107.2) + 171.14

cat("The equation is: ", "y = ", round(qq, digits = 2), " + ", round(eq, digits = 2),
"x")
```

```
## The equation is:  y =  105.97  +  0.61 x
```

b. Interpret the slope and the intercept in this context. Ans: The slope shows that if height increases, the shoulder girth also increases.

The intercept means that if a person were to have 0 cm in shoulder girth they would be 105.9651 cm tall on average. Note that is is not physically possible, so practically it means that people are taller than they are wide at the shoulder.

c. Calculate R22 of the regression line for predicting height from shoulder girth, and interpret it in the context of the application.

Ans:

```
R <- 0.67^2
R
```

```
## [1] 0.4489
```

44.89% of the model is explained by shoulder girth.

d. A randomly selected student from your class has a shoulder girth of 100 cm. Predict the height of this student using the model.

Ans:

```
height <- 105.97  +  0.61 * 100
cat("Height of this student is: ", height)
```

```
## Height of this student is:  166.97
```

e. The student from part (d) is 160 cm tall. Calculate the residual, and explain what this residual means.

Ans:

```
cat("The residual is", 160-height)
```

```
## The residual is -6.97
```

There is a 6.7626cm difference between our prediction and what actually measured. Also negative Reidual means our regression is over estimating.

f. A one year old has a shoulder girth of 56 cm. Would it be appropriate to use this linear model to predict the height of this child?

Ans: No, Data is taken from a sample of adults and can only be used to make predictions on an adult population. Children do not belong to the population.

7.30 7.30 Cats, Part I.

The following regression output is for predicting the heart weight (in g) of cats from their body weight (in kg). The coeffcients are estimated using a data set of 144 domestic cats.

Note that the people who made the measurements here made a physics error. Grams and Kilograms measure mass. Weight is the force of gravity applied to an object by Earth, and is measured in Newtons or dynes. Weight increases as mass increases by Newton’s universal law of gravitation, but they are not equivalent.

a. Write out the linear model.

Ans: y = -0.357 + 4.034 * x

b. Interpret the intercept.

Ans: If the body weight is 0kg, Hearts weight would be negative.This means probably there is an error in the data

Mass as defined by physics measures an objects inertia and how strongly it interacts with gravity. In both cases, material objects like cats and their hearts must have mass greater than 0. A literal interpretation of the slope would be non-physical.

c. Interpret the slope.

Ans: Every 1 kg a cat increases in mass, hearts mass increases by 4.034 grams.

d. Interpret R2. Ans: 64.66% of our linear model is described by the cat's body mass.

e. Calculate the correlation coefficient.

Ans:

```
R <- sqrt(.6466)
cat("The correlation coeffiecient is: ", R)
```

```
## The correlation coeffiecient is:  0.8041144
```

7.40 Rate my professor.

Many college courses conclude by giving students the opportunity to evaluate the course and the instructor anonymously. However, the use of these student evaluations as an indicator of course quality and teaching effectiveness is often criticized because these measures may reflect the influence of non-teaching related characteristics, such as the physical appearance of the instructor. Researchers at University of Texas, Austin collected data on teaching evaluation score (higher score means better) and standardized beauty score (a score of 0 means average, negative score means below average, and a positive score means above average) for a sample of 463 professors.

The scatterplot below shows the relationship between these variables, and also provided is a regression output for predicting teaching evaluation score from beauty score.

a. Given that the average standardized beauty score is -0.0883 and average teaching evaluation score is 3.9983, calculate the slope. Alternatively, the slope may be computed using just the information provided in the model summary table.

Ans:

```
y <- 3.9983
x <- -.0883
int_est <- 4.01

slop <- (y-int_est)/x
cat("The slope is: ", slop)
```

```
## The slope is:  0.1325028
```

b. Do these data provide convincing evidence that the slope of the relationship between teaching evaluation and beauty is positive? Explain your reasoning.

Ans:The data shows that the slope of the relationship between teaching evaluation and beauty is positive. Since the p-value is close to 0 we can conclude the slope to be positive.

c. List the conditions required for linear regression and check if each one is satisfied for this model based on the following diagnostic plots.

Ans: There are some trend on the scatterplot. The residuals plots are indicating that the residuals as normal. There is constant variability. Assuming that these obersvations of professor ratings are independent. Hence, The conditions for linear regression are satisfied.

(End of this file)