IS606 - Homework 8

Daniel Dittenhafer

December 3, 2015

8.2 Baby Weights, Part II (p395)

a) Write the equation of the regression line.

$$y = 120.07 - 1.93x_{parity}$$

b) Interpret the slope in the context, and calculate the predicted birth weight of first borns and others. The slope 120.07 indicates the first born (parity = 0) would be predicted to weigh 120.07 ounces. The others born, based on the slope of -1.93 would be 118.14 oz.

c) Is there a statistically significant relationship between the average birth weight and parity? Given the p-value of 0.1052 for the parity parameter, I conclude there is not a statistically significant relationship between average birth weight and parity.

8.4 Absenteeism (p397)

a) Write the equation of the regression line.

$$y = 18.93 - 9.11x_{eth} + 3.10x_{sex} + 2.15x_{lrn}$$

b) Interpret each one of the slopes in this context.

- The slope of `eth` indicates that, all else being equal, there is a 9.11 day reduction in the predicted absenteeism when the subject is no aboriginal.

- The slope of `sex` indicates that, all else being equal, there is a 3.10 day increase in the predicted absenteeism when the subject is male.

- The slope of `lrn` indicates that, all else being equal, there is a 2.15 day increase in the predicted absenteeism when the subject is a slow learner.

c) Calculate the residual for the first observation in the data set: a student who is aboriginal, male, a slow learner, and missed 2 days of school. Using the R code below, we compute the predicted absentee days and further compute the residual against the actual missed days of school.

```
eth <- 0
sex <- 1
lrn <- 1
actualDaysMissed <- 2

absDaysPred <- 18.93 - 9.11 * eth + 3.1 * sex + 2.15 * lrn
absDaysPred
```

```
## [1] 24.18
```

```
residual <- actualDaysMissed - absDaysPred
residual
```

[1] -22.18

The residual is -22.18 days.

d) The variance of the residuals is 240.57 and the variance of the number of absent days for all students in the data set is 264.17. Calculate the R^2 and adjusted R^2. Note that there are 146 observations in the data set. The following R code computes the R^2 and adjusted R^2:

```
n <- 146
k <- 3
varRes <- 240.57
varOut <- 264.17

R2 <- 1 - (240.57 / 264.17)
R2
```

[1] 0.08933641

```
adjustedR2 <- R2 * ( (n-1) / (n-k-1) )
adjustedR2
```

[1] 0.0912238

8.8 Absenteeism, Part II (p399)

8.16

8.18