STAT3032_Homework 5

Guizhen Yu

November 7, 2017

Answer for 6.10

Answer for 6.10.1

```
m6101 <- lm(quality ~ gender + numYears + pepper + discipline + easiness + ra
terInterest, data = Rateprof)
summary(m6101)
```

```
##
## Call:
## lm(formula = quality ~ gender + numYears + pepper + discipline +
##     easiness + raterInterest, data = Rateprof)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.63978 -0.42534  0.03105  0.41535  1.26088
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -0.18066    0.24240  -0.745  0.45658
## gendermale          0.04678    0.06492   0.721  0.47162
## numYears            0.01760    0.01005   1.751  0.08085 .
## pepperyes           0.56166    0.09934   5.654 3.22e-08 ***
## disciplineSocSci    0.01865    0.08889   0.210  0.83393
## disciplineSTEM      0.29475    0.08148   3.618  0.00034 ***
## disciplinePre-prof  0.09656    0.09139   1.057  0.29144
## easiness            0.51288    0.04245  12.082  < 2e-16 ***
## raterInterest       0.54413    0.05937   9.165  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5892 on 357 degrees of freedom
## Multiple R-squared:  0.5158, Adjusted R-squared:  0.505
## F-statistic: 47.54 on 8 and 357 DF,  p-value: < 2.2e-16
```

```
print("t-value of b2 is equal to 1.751")
```

```
## [1] "t-value of b2 is equal to 1.751"
```

```
print("b2 != 0 is: ")
```

```
## [1] "b2 != 0 is: "
```

```
2*pt(1.751,357, lower.tail = F)
```

```
## [1] 0.08080476
```

```
print("br <= 0 is: ")
```

```
## [1] "br <= 0 is: "
```

```
pt(1.751, 357, lower.tail = T)
```

```
## [1] 0.9595976
```

```
print("br >= 0 is: ")
```

```
## [1] "br >= 0 is: "
```

```
pt(1.751, 357, lower.tail = F)
```

```
## [1] 0.04040238
```

Answer for 6.10.2

```
m6101Fvalue <- Anova(m6101, type = 2)
sqrt(m6101Fvalue$`F value`)
```

```
## [1]  0.7206099  1.7507479  5.6537293  2.1925254 12.0817399  9.1646509
## [7]         NA
```

```
summary(m6101)
```

```
##
## Call:
## lm(formula = quality ~ gender + numYears + pepper + discipline +
##     easiness + raterInterest, data = Rateprof)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.63978 -0.42534  0.03105  0.41535  1.26088
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -0.18066    0.24240  -0.745  0.45658
## gendermale          0.04678    0.06492   0.721  0.47162
## numYears            0.01760    0.01005   1.751  0.08085 .
## pepperyes           0.56166    0.09934   5.654 3.22e-08 ***
## disciplineSocSci    0.01865    0.08889   0.210  0.83393
## disciplineSTEM      0.29475    0.08148   3.618  0.00034 ***
## disciplinePre-prof  0.09656    0.09139   1.057  0.29144
## easiness            0.51288    0.04245  12.082  < 2e-16 ***
## raterInterest       0.54413    0.05937   9.165  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5892 on 357 degrees of freedom
## Multiple R-squared:  0.5158, Adjusted R-squared:  0.505
## F-statistic: 47.54 on 8 and 357 DF,  p-value: < 2.2e-16
```

If you compare the sqrt of model 6.10.1's F-value with the t-value in summary, you will find they are the same. This means F-test and t-test produce the same conclusions.

Answer for 6.10.3

```
m6101Fvalue
```

```
## Anova Table (Type II tests)
##
## Response: quality
##                Sum Sq  Df  F value    Pr(>F)
## gender          0.180   1   0.5193  0.471621
## numYears        1.064   1   3.0651  0.080848 .
## pepper         11.098   1  31.9647 3.218e-08 ***
## discipline      5.007   3   4.8072  0.002698 **
## easiness       50.680   1 145.9684 < 2.2e-16 ***
## raterInterest  29.161   1  83.9908 < 2.2e-16 ***
## Residuals     123.949 357
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Anova(m6101, type = 3)
```

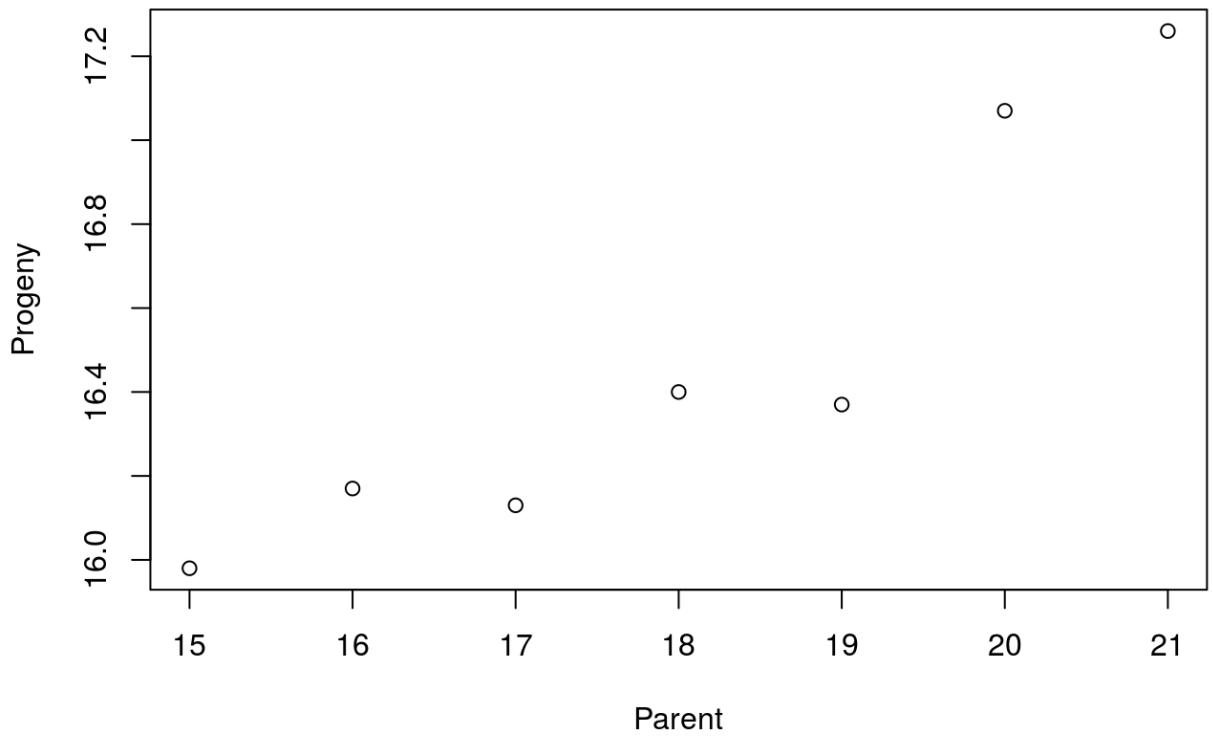
```
## Anova Table (Type III tests)
##
## Response: quality
##                Sum Sq  Df  F value    Pr(>F)
## (Intercept)     0.193   1   0.5555  0.456584
## gender          0.180   1   0.5193  0.471621
## numYears        1.064   1   3.0651  0.080848 .
## pepper         11.098   1  31.9647 3.218e-08 ***
## discipline      5.007   3   4.8072  0.002698 **
## easiness       50.680   1 145.9684 < 2.2e-16 ***
## raterInterest  29.161   1  83.9908 < 2.2e-16 ***
## Residuals     123.949 357
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

If you compare the table from type II and type III, you will find that all regressors have the same values of mean square, F statistics, and p values. This is because the model is an main effect model, and type II and type III model use same NH and AH to test each regressor. (except intercept, because there is no intercept in type II)

Answer for 7.7

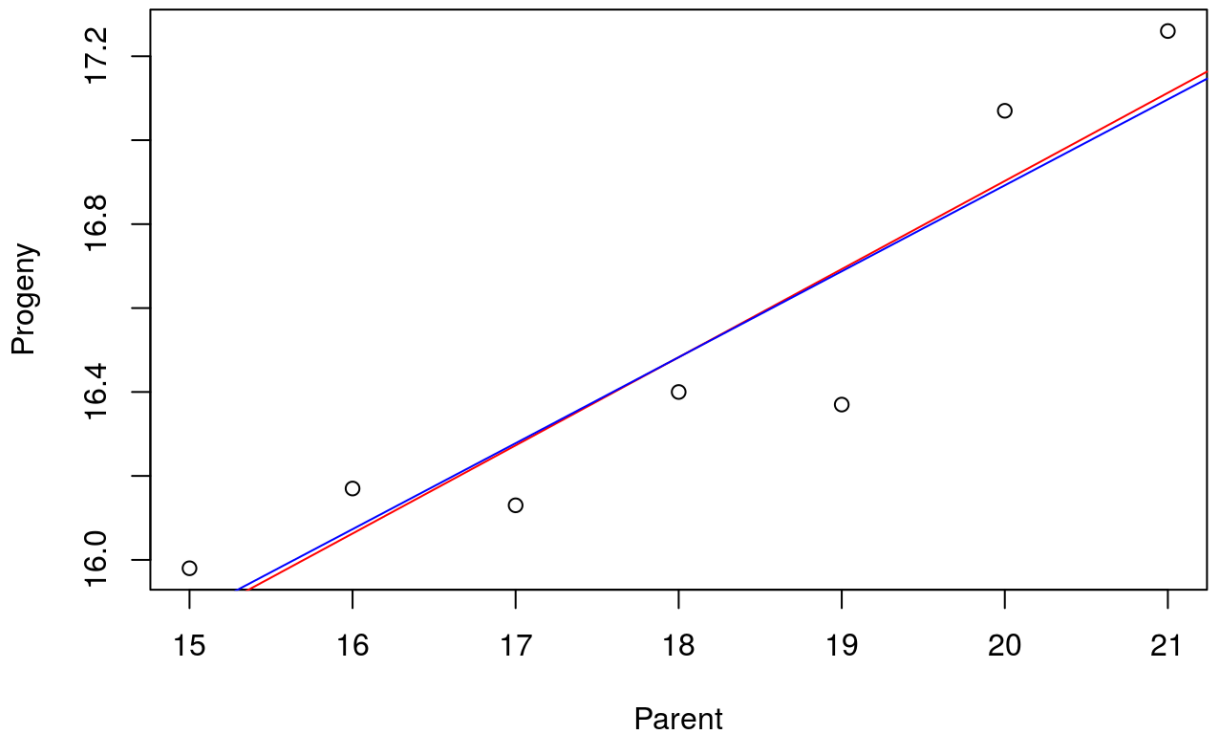
Answer for 7.7.1

```
plot(Progeny ~ Parent, data = galtonpeas)
```



Answer for 7.7.2

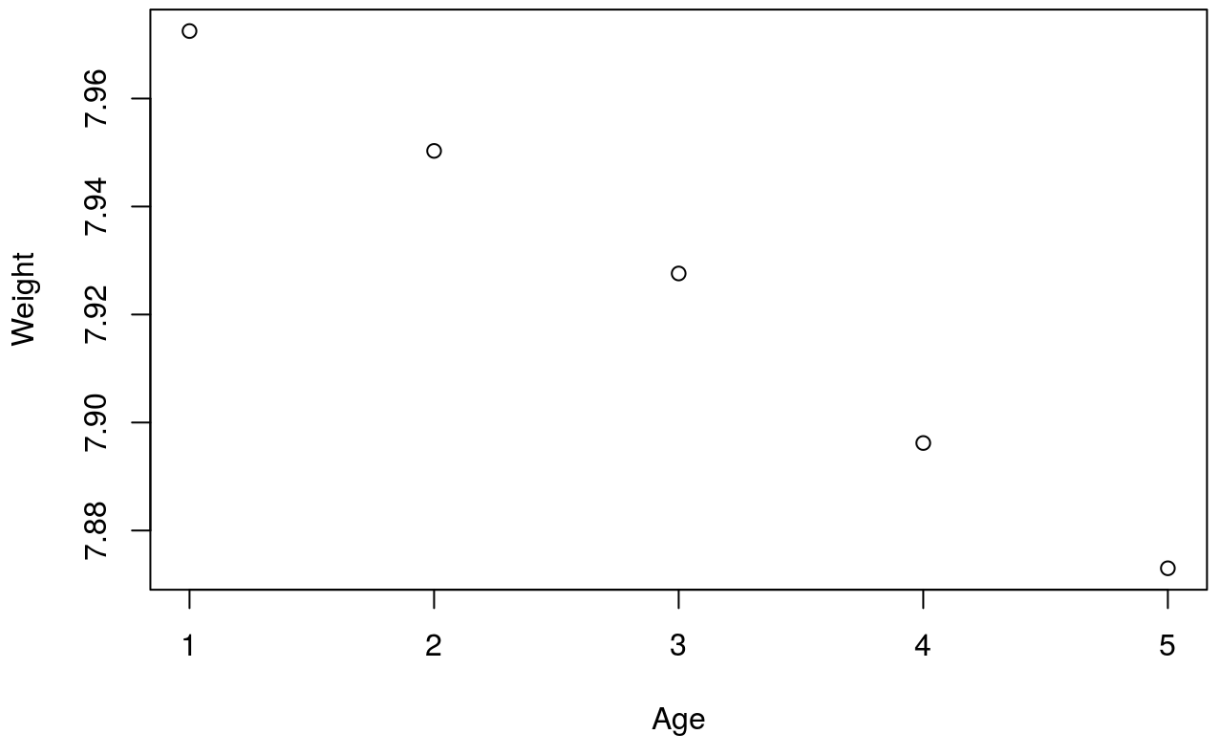
```
m77.weighted <- lm(Progeny ~ Parent, data= galtonpeas, weights = 1/SD^2)
m77.unweighted <- lm(Progeny ~ Parent, data = galtonpeas)
plot(Progeny ~ Parent, data = galtonpeas)
abline(m77.unweighted, col = "red")
abline(m77.weighted, col = "blue")
```



Answer for 7.8

Answer for 7.8.1

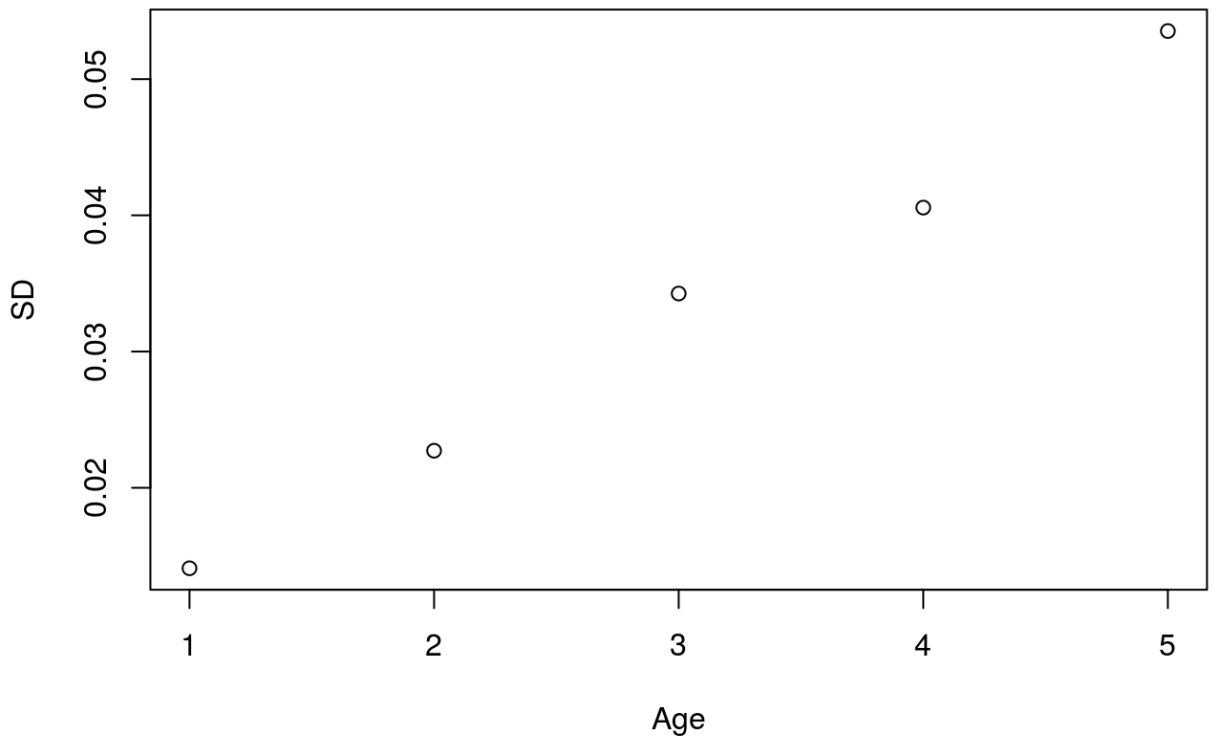
```
plot(Weight ~ Age, data = jevons)
```




```
print("This plot means weight will change when age changes in sample. Therefo
re, the usual assumptions that values in vector W are all the same are nolong
er applictable.")
```

```
## [1] "This plot means weight will change when age changes in sample. Theref
ore, the usual assumptions that values in vector W are all the same are nolon
ger applictable."
```

```
plot(SD ~ Age, data = jevons)
```




```
print("This plot means SD will change as Age changes, therefore, the variance are not constant for E(Y|X) therefore, we have to use WLS to make variance constant.")
```

```
## [1] "This plot means SD will change as Age changes, therefore, the variance are not constant for E(Y|X) therefore, we have to use WLS to make variance constant."
```

Answer for 7.8.2

```
m78.weighted <- lm(Weight ~ Age, data= jevons, weights = SD^2/Age)
summary(m78.weighted)
```

```
##
## Call:
## lm(formula = Weight ~ Age, data = jevons, weights = SD^2/Age)
##
## Weighted Residuals:
##          1          2          3          4          5
## -3.915e-05  9.568e-06  6.863e-05 -4.780e-05  4.205e-07
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  8.0008525  0.0034857 2295.34 1.82e-10 ***
## Age         -0.0255740  0.0009314  -27.46 0.000106 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.36e-05 on 3 degrees of freedom
## Multiple R-squared:  0.996,  Adjusted R-squared:  0.9947
## F-statistic: 753.9 on 1 and 3 DF,  p-value: 0.000106
```

Answer for 7.8.3

```
print("The fitted regression consistent with the known standard weight for a
new coin. based on the model, for a coin that is age 0(new), the predicted we
ight is 8.0008, while the standard weight is 7.9876. therefore, they are cons
istent. ")
```

```
## [1] "The fitted regression consistent with the known standard weight for a
new coin. based on the model, for a coin that is age 0(new), the predicted we
ight is 8.0008, while the standard weight is 7.9876. therefore, they are cons
istent. "
```