

Homework9

Yanyu Zheng yz2690

April 13, 2016

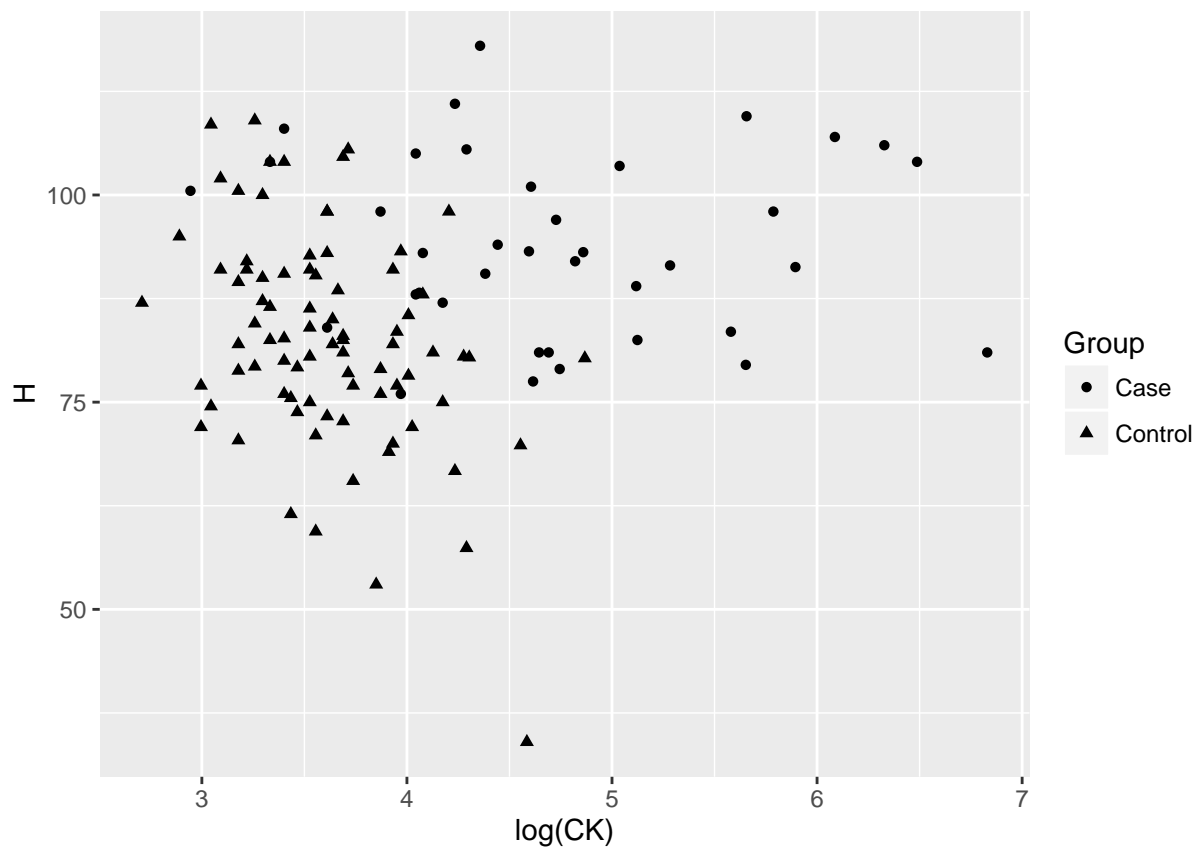
1. Chapter 20, problem 12

a

```
library("Sleuth3")  
library("ggplot2")
```

```
## Warning: package 'ggplot2' was built under R version 3.2.4
```

```
attach(ex2012)  
ggplot(ex2012, aes(x=log(CK), y=H, shape=Group)) + geom_point()
```



Yes, we can see possible connections in the plot.

b

```
myLogit1 = glm(Group ~ CK + I(CK^2), data = ex2012, family = "binomial")
summary(myLogit1)

##
## Call:
## glm(formula = Group ~ CK + I(CK^2), family = "binomial", data = ex2012)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.50536  -0.03915   0.37969   0.51841   2.27337
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  4.177e+00  7.264e-01   5.751 8.87e-09 ***
## CK          -5.798e-02  1.299e-02  -4.463 8.10e-06 ***
## I(CK^2)       5.054e-05  3.268e-05   1.547  0.122
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 149.84  on 119  degrees of freedom
## Residual deviance:  85.47  on 117  degrees of freedom
## AIC: 91.47
##
## Number of Fisher Scoring iterations: 9
```

Squared term is not significant for CK or log(CK).

c

```
myLogit2 = glm(Group ~ log(CK) + I(log(CK)^2), family = "binomial")
summary(myLogit2)

##
## Call:
## glm(formula = Group ~ log(CK) + I(log(CK)^2), family = "binomial")
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.39368  -0.03111   0.38041   0.50222   2.28558
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -9.735     16.298  -0.597   0.550
## log(CK)        8.516      8.358   1.019   0.308
```

```
## I(log(CK)^2)   -1.446      1.063  -1.360    0.174
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 149.840  on 119  degrees of freedom
## Residual deviance:  85.017  on 117  degrees of freedom
## AIC: 91.017
##
## Number of Fisher Scoring iterations: 7
```

I choose the first model with the original scale, since the intercept and CK is significant, while the log scale model has nothing significant.

d

```
myLogit3 = glm(Group ~ log(CK) + H, family = "binomial")
myLogit4 = glm(Group ~ 1, family = "binomial")
anova(myLogit3, myLogit4)
```

```
## Analysis of Deviance Table
##
## Model 1: Group ~ log(CK) + H
## Model 2: Group ~ 1
##   Resid. Df Resid. Dev Df Deviance
## 1         117      61.992
## 2         119     149.840 -2   -87.847
```

e

```
p1 = predict(myLogit3, data.frame(CK = 80, H = 85))
p2 = predict(myLogit3, data.frame(CK = 300, H = 100))
inv1 = exp(p1)/(1+exp(p1))
inv2 = exp(p2)/(1+exp(p2))
inv1/inv2
```

```
##           1
## 908.1816
```

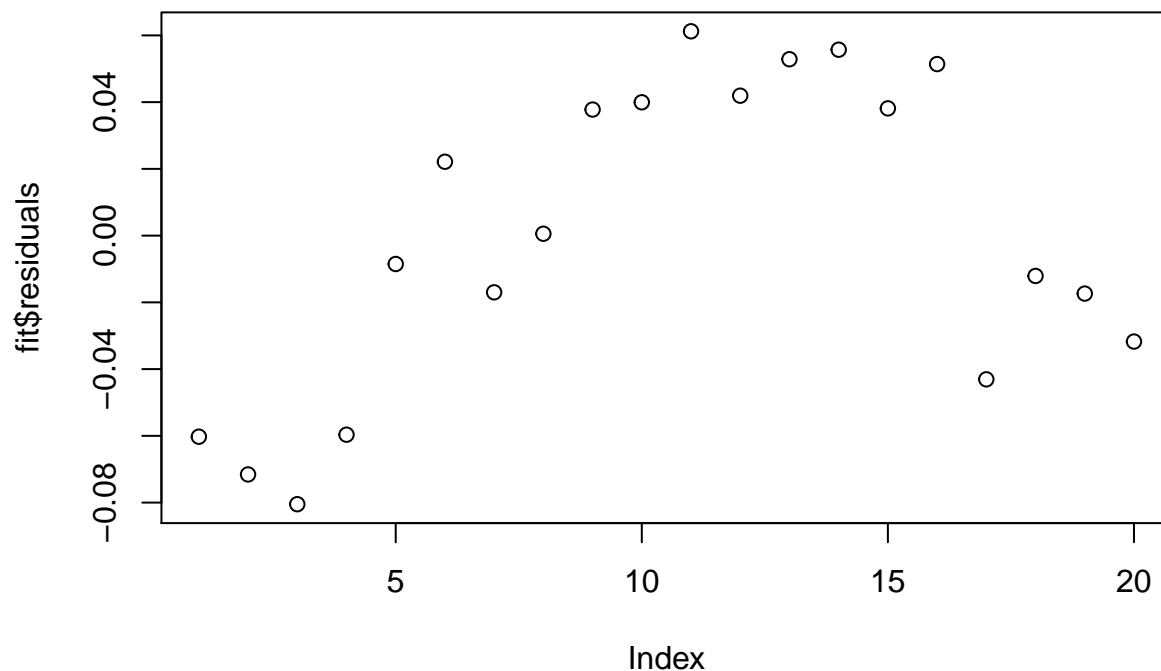
```
detach(ex2012)
```

2. Chapter 21, problem 16

```
attach(ex2116)
odd = Tumor/Total
oddL = exp(odd)/(1+exp(odd))
fit = lm(oddL ~ Dose)
summary(fit)
```

```
##
## Call:
## lm(formula = oddL ~ Dose)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.08048 -0.03458 -0.00397  0.04045  0.06123
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.58026    0.01518  38.236 < 2e-16 ***
## Dose         0.58307    0.12331   4.728 0.000168 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.048 on 18 degrees of freedom
## Multiple R-squared:  0.554, Adjusted R-squared:  0.5292
## F-statistic: 22.36 on 1 and 18 DF, p-value: 0.0001677
```

```
plot(fit$residuals)
```



```
## We can see non-linearity in the plot, try log transfer and square term
fit2 = lm(oddL ~ log(Dose)+log(Dose)^2)
fit2
```

```
##
## Call:
## lm(formula = oddL ~ log(Dose) + log(Dose)^2)
##
## Coefficients:
## (Intercept)    log(Dose)
##    0.80553      0.05826
```

From the model, we can see the dose and the odd is definitely related. And we do a F-test to see if the difference is really significant between group.

```
fitN = lm(oddL~1)
anova(fit,fitN)

## Analysis of Variance Table
##
## Model 1: oddL ~ Dose
## Model 2: oddL ~ 1
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1      18 0.041474
## 2      19 0.092986 -1 -0.051512 22.357 0.0001677 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Again, it is significant.

```
fitNew = lm(Dose ~ odd)
predict(fitNew, data.frame(odd=0.5))
```

```
##           1
## 0.07645605
```

The dose causing 50% tumor rate is about 0.07634605.