

Homework 6

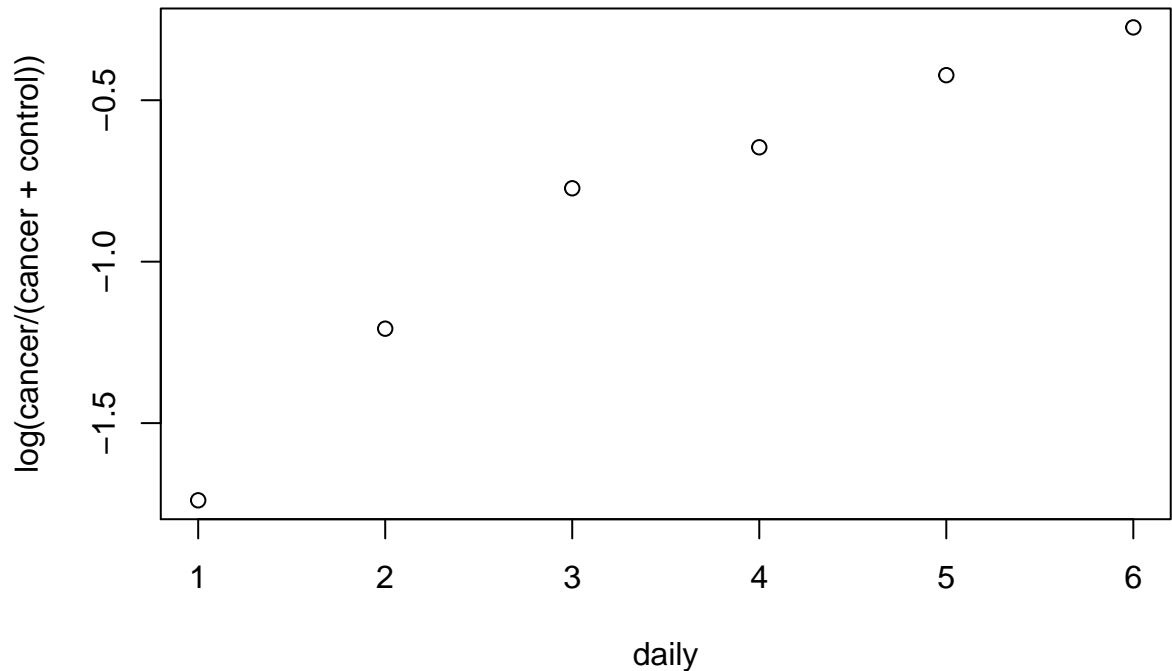
Jing Leng

November 6, 2014

1

a

```
cancer <- c(13, 55, 489, 475, 293, 38)
control <- c(61, 129, 570, 431, 154, 12)
y <- cbind(cancer, control)
daily = c(1,2,3,4,5,6)
lung <- data.frame(daily = c(1,2,3,4,5,6), rate = cancer / (cancer + control) )
m1 <- glm(y ~ daily, family = binomial)
plot(daily, log(cancer/(cancer + control)))
```



7.4797 degrees of freedom = 4

G =

b

```
phat <- fitted.values(m1)
ni = rowSums(y)
nphat = ni * phat
chis <- sum((y[,1] - nphat) ^ 2/ (nphat * (1-phat)))
chis
```

```
## [1] 7.362
```

```
1-pchisq(chis,1)
```

```
## [1] 0.006661
```

We reject the null hypothesis.

c

```
m0 <- glm(y ~ 1, family = binomial)
anova(m0, m1)
```

```
## Analysis of Deviance Table
##
## Model 1: y ~ 1
## Model 2: y ~ daily
##   Resid. Df Resid. Dev Df Deviance
## 1         5      130.9
## 2         4       7.5  1      123
```

```
1-pchisq(123.42,1)
```

```
## [1] 0
```

```
library(aod)
wald.test(b = coef(m1), Sigma = vcov(m1), Terms = 1)
```

```
## Wald test:
## -----
##
## Chi-squared test:
## X2 = 107.1, df = 1, P(> X2) = 0.0
```

LRT: 123.42, p=0 Wald: 107.1, p = 0 Both tests reject the null hypothesis.

d

```
summary(m1)$coef
```

```
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.6003    0.15466  -10.35 4.299e-25
## daily        0.4463    0.04159   10.73 7.286e-27
```

```
confint(m1)
```

```
## Waiting for profiling to be done...
```

```
##           2.5 % 97.5 %
## (Intercept) -1.9059 -1.2994
## daily        0.3654  0.5285
```

e

```
y2 <- matrix(c(68, 190, 964, 1001, 331, 166), ncol = 2)
m2 <- glm(y2 ~ c(1,2,3), family = binomial)
summary(m2)$coef
```

```
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -4.980      0.17525  -28.41 1.365e-177
## c(1, 2, 3)     2.239      0.07379   30.34 3.403e-202
```

```
confint(m2)
```

```
## Waiting for profiling to be done...
```

```
##              2.5 % 97.5 %
## (Intercept) -5.331 -4.644
## c(1, 2, 3)   2.097  2.386
```

```
estimate: 2.238741, confidence interval: (2.096861 2.386239)
```

2

a

```
deaths <- c(2,32,12,104,28,206,28,186,31,102)
py <- c(18793,52407,10673,43248,5710,28612,2585,12663,1462,5317)
logpy <- log(py)
age <- gl(5,2)
smoke <- gl(2,1,length=10)
dataf <- data.frame(age=age,smoke=smoke,deaths=deaths,
                    py=py,logpy=logpy)
m1 <- glm(deaths ~ age+smoke,offset=logpy,family=poisson,
          data=dataf)
summary(m1)
```

```
##
## Call:
## glm(formula = deaths ~ age + smoke, family = poisson, data = dataf,
##      offset = logpy)
##
## Deviance Residuals:
##      1      2      3      4      5      6      7      8
## -2.1800  0.9018 -1.3080  0.5104 -0.1379  0.0513  0.2289 -0.0873
##      9     10
##  1.9191 -0.9124
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -7.919      0.192  -41.30 < 2e-16 ***
```

```

## age2          1.484      0.195    7.61  2.8e-14 ***
## age3          2.628      0.184   14.30 < 2e-16 ***
## age4          3.351      0.185   18.13 < 2e-16 ***
## age5          3.700      0.192   19.25 < 2e-16 ***
## smoke2        0.355      0.107    3.30  0.00096 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
## Null deviance: 935.091  on 9  degrees of freedom
## Residual deviance: 12.134  on 4  degrees of freedom
## AIC: 79.2
##
## Number of Fisher Scoring iterations: 4

dataf$inter <- c(0,1,0,2,0,3,0,4,0,5)
m2 <- glm(deaths~ age+smoke+inter,offset=logpy,family=poisson, data=dataf)
summary(m2)

##
## Call:
## glm(formula = deaths ~ age + smoke + inter, family = poisson,
## data = dataf, offset = logpy)
##
## Deviance Residuals:
##      1      2      3      4      5      6      7      8
## -0.8775  0.2712  0.2075 -0.0689  0.6201 -0.2186 -0.4381  0.1754
##      9     10
##  0.0060 -0.0033
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -8.5856     0.3050  -28.15 < 2e-16 ***
## age2           1.7346     0.2133   8.13  4.2e-16 ***
## age3           3.1484     0.2533  12.43 < 2e-16 ***
## age4           4.1420     0.3193  12.97 < 2e-16 ***
## age5           4.7310     0.3840  12.32 < 2e-16 ***
## smoke2        1.4450     0.3729   3.88  0.00011 ***
## inter        -0.3087     0.0973  -3.17  0.00150 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
## Null deviance: 935.0909  on 9  degrees of freedom
## Residual deviance:  1.5464  on 3  degrees of freedom
## AIC: 70.61
##
## Number of Fisher Scoring iterations: 4

```

b

```
obsrate <- deaths/py
lam <- obsrate
eta <- log(lam)
z <- eta
w <- deaths
# do 5 Fisher scoring iterations
for (i in 1:4) {
m3 <- lm(z ~ age + smoke, weights=w) # regress z on X weights w
eta <- fitted(m3)
lam <- exp(eta)
z <- eta + (obsrate - lam)/lam
w <- py * lam }
coef(m3) # beta hat
```

```
## (Intercept)      age2      age3      age4      age5      smoke2
##      -7.9194      1.4840      2.6275      3.3505      3.7001      0.3545
```

```
varbeta <- summary(m3)$cov.unscaled # (X'WX)^-1
sqrt(diag(varbeta)) # standard errors
```

```
## (Intercept)      age2      age3      age4      age5      smoke2
##      0.1918      0.1951      0.1837      0.1848      0.1922      0.1074
```

```
varbeta
```

```
##      (Intercept)      age2      age3      age4      age5
## (Intercept)      0.036772 -0.0289194 -0.0286912 -0.0287139 -0.029050
## age2      -0.028919  0.0380651  0.0294597  0.0294582  0.029436
## age3      -0.028691  0.0294597  0.0337555  0.0294798  0.029447
## age4      -0.028714  0.0294582  0.0294798  0.0341505  0.029446
## age5      -0.029050  0.0294357  0.0294469  0.0294458  0.036948
## smoke2     -0.009212 -0.0006159 -0.0009015 -0.0008731 -0.000453
##      smoke2
## (Intercept) -0.0092119
## age2      -0.0006159
## age3      -0.0009015
## age4      -0.0008731
## age5      -0.0004530
## smoke2      0.0115292
```

```
vcov(m1)
```

```
##      (Intercept)      age2      age3      age4      age5
## (Intercept)      0.036772 -0.0289194 -0.0286913 -0.0287139 -0.029050
## age2      -0.028919  0.0380651  0.0294597  0.0294582  0.029436
## age3      -0.028691  0.0294597  0.0337555  0.0294798  0.029447
## age4      -0.028714  0.0294582  0.0294798  0.0341506  0.029446
## age5      -0.029050  0.0294358  0.0294470  0.0294459  0.036948
```

```

## smoke2      -0.009212 -0.0006159 -0.0009015 -0.0008731 -0.000453
##              smoke2
## (Intercept) -0.0092119
## age2        -0.0006159
## age3        -0.0009015
## age4        -0.0008731
## age5        -0.0004530
## smoke2      0.0115292

```