

chapter 5 : Q5.4

Kholoud Basalim

2024-11-19

5.4 For the leukemia survival data in Exercise 4.2:

(a) Use the Wald statistic to obtain an approximate 95% confidence interval for the parameter β_1 .

(b) By comparing the deviances for two appropriate models, test the null hypothesis $\beta_2 = 0$ against the alternative hypothesis $\beta_2 \neq 0$. What can you conclude about the use of the initial white blood cell count as a predictor of survival time?

```
# y: times to death .
y<-c(65,156,100,134,16,108,121,4,39,143,56,26,22,1,1,5,65)
y
## [1] 65 156 100 134 16 108 121 4 39 143 56 26 22 1 1 5 65

#x: log10(initial white blood cell count).
x<-c(3.36,2.88,3.63,3.41,3.78,4.02,4,4.23,3.73,3.85,3.97,4.51,4.54,5,5,4.72,5)
x
## [1] 3.36 2.88 3.63 3.41 3.78 4.02 4.00 4.23 3.73 3.85 3.97 4.51 4.54 5.00 5.00
## [16] 4.72 5.00
```

1) Hypothesis: H_0 vs H_1

$$H_0: \beta_2 = 0 \quad \gg \quad H_0: \mu_i = e^{\beta_1}$$

Vs

$$H_1: \beta_2 \neq 0 \quad \gg \quad H_1: \mu_i = e^{\beta_1 + \beta_2 * x_i}$$

a-Model under H_1 :

```
#To Test H0: Beta2 equal 0 vs H1: Beta2 not equal 0 :
# 1) find model under H0 and H1 :
#model under H1 (M1) :
m1<-glm(y~x, family = Gamma(link = "log"))
summary(m1, dispersion = 1)

##
## Call:
## glm(formula = y ~ x, family = Gamma(link = "log"))
##
## Coefficients:
##           Estimate Std. Error z value Pr(>|z|)
```

```
## (Intercept)    8.4775    1.6548    5.123 3.01e-07 ***
## x             -1.1093    0.3997   -2.776 0.00551 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 1)
##
## Null deviance: 26.282 on 16 degrees of freedom
## Residual deviance: 19.457 on 15 degrees of freedom
## AIC: 173.97
##
## Number of Fisher Scoring iterations: 8
```

b-Model under H_0 :

```
#model under H0 (M0) :
m0<-glm(y~1,family =Gamma(link = "log"))
summary(m0,dispersion =1)

##
## Call:
## glm(formula = y ~ 1, family = Gamma(link = "log"))
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  4.1347      0.2425   17.05  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 1)
##
## Null deviance: 26.282 on 16 degrees of freedom
## Residual deviance: 26.282 on 16 degrees of freedom
## AIC: 178.09
##
## Number of Fisher Scoring iterations: 6
```

2)Find test statistics: $D_0 \geq D_1$

$$\Delta D = D_0 - D_1 \sim \chi^2_{p-q}$$

D0=Deviance of model M0

D1= Deviance of model M1

p= number of parameters in model $M_1=2$

q= number of parameters in model $M_0=1$

```
# 2) Find test statistics
# we can find test statistics by useing Deviance (delta(D)= D0-D1)
D0<- deviance(m0)
D0
```

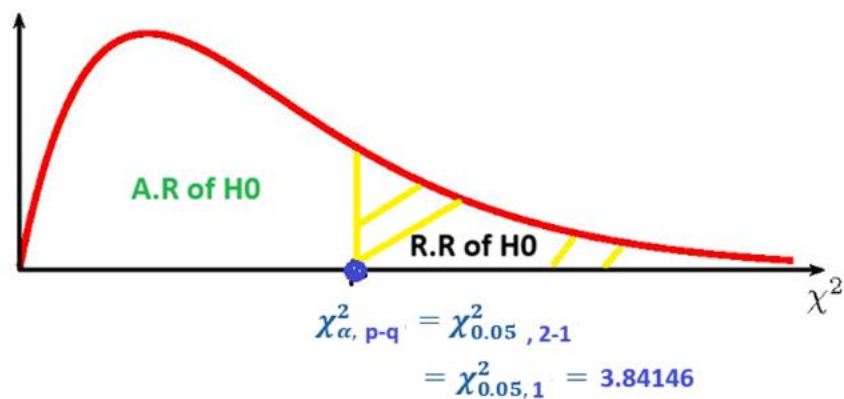
```
## [1] 26.2821
D1<- deviance(m1)
D1
## [1] 19.45653
# or by code : anova(m0 ,m1 , test = "Chisq")
delta_D <- D0-D1
delta_D
## [1] 6.825567
```

$$\Delta D = D_0 - D_1 = 26.2821 - 19.4563 = 6.825567$$

3)To find table value:

use the chi-square table and the degrees of freedom (df) are calculated as p-q.

```
# critical value "from chi table " df=p-q =2-1=1
# p= number of parameter in model M1=2 ,, q= number of parameter in model M0=1
df_diff<- m0$df.residual - m1$df.residual
df_diff
## [1] 1
critical_value<- qchisq(1-0.05 , df_diff)
critical_value
## [1] 3.841459
```



4)Decision:

Since $\Delta D > \chi^2_{0.05, 1}$, We Reject $H_0: \beta_2 = 0$., and we conclude that x has a significant importance in predicting

```

#Decision:
if (delta_D >critical_value) {
  print("Reject H0")
}else{
  print("Do not Reject H0")
}
## [1] "Reject H0"

```

#####

Hypothesis Testing Steps for Model Parameters:

1- Hypothesis:

$$H_0: \beta_i = 0 \quad \text{Vs} \quad H_1: \beta_i \neq 0$$

2- Test statistics:

$$\begin{aligned} \Delta D &= D_0 - D_1 \sim \chi^2_{p-q} \\ &= 2 [\ell(\mathbf{b}_1; \mathbf{y}) - \ell(\mathbf{b}_0; \mathbf{y})] \end{aligned}$$

D_0 = Deviance of model M_0 **and** D_1 = Deviance of model M_1

($D_0 \geq D_1$)

b_0 is the MLE of β_0 for model M_0 **and** b_1 is the MLE of β_1 for model M_1

$$\ell(\mathbf{b}_1; \mathbf{y}) \geq \ell(\mathbf{b}_0; \mathbf{y})$$

3- Critical value

$$\chi^2_{p-q}$$

p= number of parameters in model M_1 **and** q= number of parameters in model M_0

4- Decision:

If $\Delta D > \chi^2_{\alpha, p-q}$, We Reject H_0

If $\Delta D \leq \chi^2_{\alpha, p-q}$, We do not Reject H_0 , and choose model M_0