第十三章

设置目录

```
setwd("D:\\data\\chapter 13")
```

例 13-1

```
data13.1<-read.csv("13-1.csv")
lm.reg.13.1<-lm(y~X1+X2+X3, data=data13.1)
summary(lm.reg.13.1)
```

```
##
## Call:
## lm(formula = y ~ X1 + X2 + X3, data = data13.1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3749 -0.2747  0.1042  0.1820  0.4277
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4.71489    1.30082  -3.625  0.00228 **
## X1           0.06091    0.02050   2.971  0.00901 **
## X2           0.03563    0.01531   2.327  0.03339 *
## X3           0.04924    0.02866   1.718  0.10507
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2853 on 16 degrees of freedom
## Multiple R-squared:  0.7251, Adjusted R-squared:  0.6736
## F-statistic: 14.07 on 3 and 16 DF,  p-value: 9.464e-05
```

```
lm.X1<-lm(y~X1,data=data13.1)
summary(lm.X1)
```

```
##
## Call:
## lm(formula = y ~ X1, data = data13.1)
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6171 -0.2472  0.0945  0.2436  0.4683
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.60311    1.05873  -1.514  0.14734
## X1           0.08819    0.02132   4.136  0.00062 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3674 on 18 degrees of freedom
## Multiple R-squared:  0.4873, Adjusted R-squared:  0.4588
## F-statistic: 17.11 on 1 and 18 DF,  p-value: 0.00062
```

```
lm.X2<-lm(y~X2,data=data13.1)
summary(lm.X2)
```

```
##
## Call:
## lm(formula = y ~ X2, data = data13.1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.82621 -0.06326  0.02728  0.42039  0.51080
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.54535    1.69061  -0.914    0.373
## X2           0.05468    0.02142   2.553    0.020 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4396 on 18 degrees of freedom
## Multiple R-squared:  0.2658, Adjusted R-squared:  0.225
## F-statistic: 6.516 on 1 and 18 DF,  p-value: 0.01999
```

```
lm.X3<-lm(y~X3,data=data13.1)
summary(lm.X3)
```

```
##
## Call:
## lm(formula = y ~ X3, data = data13.1)
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.66641 -0.10305  0.08157  0.17057  0.51838
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.15531    0.90922  -1.271 0.220029
## X3           0.11656    0.02694   4.326 0.000407 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3592 on 18 degrees of freedom
## Multiple R-squared:  0.5098, Adjusted R-squared:  0.4825
## F-statistic: 18.72 on 1 and 18 DF,  p-value: 0.0004066
```

```
lm.x1.X2<-lm(y~X1+X2,data=data13.1)
summary(lm.x1.X2)
```

```
##
## Call:
## lm(formula = y ~ X1 + X2, data = data13.1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5044 -0.2673  0.1103  0.2665  0.3213
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4.90816    1.36832  -3.587 0.002272 **
## X1           0.08137    0.01762   4.619 0.000245 ***
## X2           0.04623    0.01479   3.125 0.006164 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3013 on 17 degrees of freedom
## Multiple R-squared:  0.6744, Adjusted R-squared:  0.6361
## F-statistic:  17.6 on 2 and 17 DF,  p-value: 7.211e-05
```

```
lm.x1.X3<-lm(y~X1+X3,data=data13.1)
summary(lm.x1.X3)
```

```
##
## Call:
## lm(formula = y ~ X1 + X3, data = data13.1)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.50217 -0.23191  0.01003  0.07585  0.60006
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.47578    0.98274  -2.519   0.0221 *
## X1           0.05413    0.02278   2.376   0.0295 *
## X3           0.07612    0.02944   2.586   0.0192 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3202 on 17 degrees of freedom
## Multiple R-squared:  0.632,  Adjusted R-squared:  0.5887
## F-statistic:  14.6 on 2 and 17 DF,  p-value: 0.0002039
```

```
lm.x2.X3<-lm(y~X2+X3,data=data13.1)
summary(lm.x2.X3)
```

```
##
## Call:
## lm(formula = y ~ X2 + X3, data = data13.1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5806 -0.2004  0.1358  0.1723  0.3689
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.85331    1.37756  -2.071  0.05387 .
## X2           0.02917    0.01831   1.593  0.12957
## X3           0.09870    0.02819   3.502  0.00273 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3448 on 17 degrees of freedom
## Multiple R-squared:  0.5735, Adjusted R-squared:  0.5233
## F-statistic: 11.43 on 2 and 17 DF,  p-value: 0.0007156
```

```
lm.x1.X2.X3<-lm(y~X1+X2+X3,data=data13.1)
summary(lm.x1.X2.X3)
```

```
##
```

```
## Call:
## lm(formula = y ~ X1 + X2 + X3, data = data13.1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3749 -0.2747  0.1042  0.1820  0.4277
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4.71489    1.30082  -3.625  0.00228 **
## X1           0.06091    0.02050   2.971  0.00901 **
## X2           0.03563    0.01531   2.327  0.03339 *
## X3           0.04924    0.02866   1.718  0.10507
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2853 on 16 degrees of freedom
## Multiple R-squared:  0.7251, Adjusted R-squared:  0.6736
## F-statistic: 14.07 on 3 and 16 DF,  p-value: 9.464e-05
```

```
lm.x1.X3<-lm(y~X1+X3,data=data13.1)
summary(lm.x1.X3)
```

```
##
## Call:
## lm(formula = y ~ X1 + X3, data = data13.1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.50217 -0.23191  0.01003  0.07585  0.60006
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.47578    0.98274  -2.519   0.0221 *
## X1           0.05413    0.02278   2.376   0.0295 *
## X3           0.07612    0.02944   2.586   0.0192 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3202 on 17 degrees of freedom
## Multiple R-squared:  0.632,  Adjusted R-squared:  0.5887
## F-statistic:  14.6 on 2 and 17 DF,  p-value: 0.0002039
```

例 13-2

```
data13.2<-read.table("13-2.csv",header=T,sep=",")

model1<-glm(Y~X4+X6,data=data13.2,family = binomial())
summary(model1)
```

```
##
## Call:
## glm(formula = Y ~ X4 + X6, family = binomial(), data = data13.2)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.9114  -0.9114  -0.9114   1.4691   1.9214
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.6487     0.4883  -3.376 0.000735 ***
## X4            0.9849     0.4915   2.004 0.045101 *
## X6           -1.0102     0.4484  -2.253 0.024262 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1896.9  on 1492  degrees of freedom
## Residual deviance: 1886.0  on 1490  degrees of freedom
## AIC: 1892
##
## Number of Fisher Scoring iterations: 4
```

```
x1<-cbind(1,1,0)
pai1<-exp(x1%*%coef(model1))/(1+exp(x1%*%coef(model1)))
x2<-cbind(1,0,0)
pai2<-exp(x2%*%coef(model1))/(1+exp(x2%*%coef(model1)))
odd1<-exp(x1%*%coef(model1))
odd2<-exp(x2%*%coef(model1))
OR<-odd1/odd2
OR
```

```
##          [,1]
## [1,] 2.677447
```

```r
model2<-glm(Y~X1+X2+X3+X4+X5+X6+X7+X8,data=data13.2,family = binomial())
summary(model2)
```

```
##
## Call:
## glm(formula = Y ~ X1 + X2 + X3 + X4 + X5 + X6 + X7 + X8, family = binomial(),
##     data = data13.2)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.5251  -0.8947  -0.6493   1.1814   2.3860
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.03969    0.64921  -4.682 2.84e-06 ***
## X1           0.01149    0.11979   0.096  0.92358
## X2          -0.10126    0.05804  -1.745  0.08105 .
## X3           0.10554    0.05868   1.799  0.07207 .
## X4           1.54410    0.50515   3.057  0.00224 **
## X5           0.48536    0.09411   5.157 2.51e-07 ***
## X6          -1.11460    0.46271  -2.409  0.01600 *
## X7          -0.18445    0.08581  -2.150  0.03159 *
## X8           1.22167    0.14244   8.577  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1896.9  on 1492  degrees of freedom
## Residual deviance: 1743.4  on 1484  degrees of freedom
## AIC: 1761.4
##
## Number of Fisher Scoring iterations: 4
```

```r
OR2<-cbind(exp(coef(model2)))
OR2
```

```
##                   [,1]
## (Intercept) 0.04784972
## X1          1.01155702
## X2          0.90369951
## X3          1.11130728
## X4          4.68375466
```

```
## X5          1.62475839
## X6          0.32804554
## X7          0.83155830
## X8          3.39285126
```

```
OR2confint<-exp(confint(model2))
```

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

```
OR2confint
```

```
##                  2.5 %     97.5 %
## (Intercept) 0.01239935  0.1622169
## X1          0.80008878  1.2798437
## X2          0.80646850  1.0126435
## X3          0.99046169  1.2468607
## X4          1.88478328 14.2171488
## X5          1.35242812  1.9563264
## X6          0.12013513  0.7600184
## X7          0.70331795  0.9850163
## X8          2.57154083  4.4959291
```

例 13-3

```
data13.3<-read.csv("13-3.csv")
```

```
library(survival)
cox.13.3<-coxph(Surv(t,status)~(grade+size+relapse),data13.3)
summary(cox.13.3)
```

```
## Call:
## coxph(formula = Surv(t, status) ~ (grade + size + relapse), data = data13.3)
##
##   n= 30, number of events= 27
##
##           coef exp(coef) se(coef)     z Pr(>|z|)
## grade   1.6804    5.3675   0.3817 4.403 1.07e-05 ***
## size    1.0782    2.9393   0.4600 2.344   0.0191 *
## relapse 0.9790    2.6617   0.4602 2.127   0.0334 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##         exp(coef) exp(-coef) lower .95 upper .95
## grade       5.367     0.1863     2.540    11.341
```

```
## size        2.939     0.3402     1.193     7.242
## relapse     2.662     0.3757     1.080     6.560
##
## Concordance= 0.825  (se = 0.064 )
## Rsquare= 0.683   (max possible= 0.992 )
## Likelihood ratio test= 34.42  on 3 df,   p=1.614e-07
## Wald test            = 23.66  on 3 df,   p=2.944e-05
## Score (logrank) test = 33.98  on 3 df,   p=2e-07
```