hw9.R

jiayuan

Sun Nov 15 21:20:04 2015

```r
##HW9 Appendix
#2A
hyperstudy <- read.csv("hyperstudy.csv",header=T)
attach(hyperstudy)
table(hypertension)
```

```
## hypertension
##   0   1
## 533 147
```

```r
147/(147+533)
```

```
## [1] 0.2161765
```

```r
#2B
log.out <- glm(hypertension ~ age + smoke + relevel(factor(race),ref='1'),
               family=binomial(link=logit))
summary(log.out)
```

```
##
## Call:
## glm(formula = hypertension ~ age + smoke + relevel(factor(race),
##     ref = "1"), family = binomial(link = logit))
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.1704  -0.7539  -0.5374  -0.4513   2.1851
##
## Coefficients:
##                                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)                    -1.578289   0.487966  -3.234  0.00122
## age                            -0.005791   0.009726  -0.595  0.55155
## smoke                           1.057311   0.193989   5.450 5.03e-08
## relevel(factor(race), ref = "1")2  0.689845   0.241574   2.856  0.00430
## relevel(factor(race), ref = "1")3 -0.370916   0.344169  -1.078  0.28116
## relevel(factor(race), ref = "1")4 -0.164278   0.347479  -0.473  0.63638
##
## (Intercept)                    **
## age
## smoke                          ***
## relevel(factor(race), ref = "1")2 **
## relevel(factor(race), ref = "1")3
## relevel(factor(race), ref = "1")4
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## 
## (Dispersion parameter for binomial family taken to be 1)
## 
##     Null deviance: 709.96  on 679  degrees of freedom
## Residual deviance: 669.63  on 674  degrees of freedom
## AIC: 681.63
## 
## Number of Fisher Scoring iterations: 4
```

```r
exp(coef(log.out))
```

```
##                     (Intercept)                             age
##                       0.2063278                       0.9942257
##                           smoke relevel(factor(race), ref = "1")2
##                       2.8786197                       1.9934060
## relevel(factor(race), ref = "1")3 relevel(factor(race), ref = "1")4
##                       0.6901021                       0.8485065
```

```r
exp(confint(log.out))
```

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

```
##                                      2.5 %     97.5 %
## (Intercept)                     0.07830377 0.5317231
## age                             0.97540226 1.0133630
## smoke                           1.97321281 4.2256692
## relevel(factor(race), ref = "1")2 1.23531107 3.1915894
## relevel(factor(race), ref = "1")3 0.33722581 1.3135277
## relevel(factor(race), ref = "1")4 0.41248113 1.6279665
```

```r
#2C
library("epicalc")
```

```
## Loading required package: foreign
## Loading required package: survival
## Loading required package: MASS
## Loading required package: nnet
```

```
lroc(log.out)$auc
```
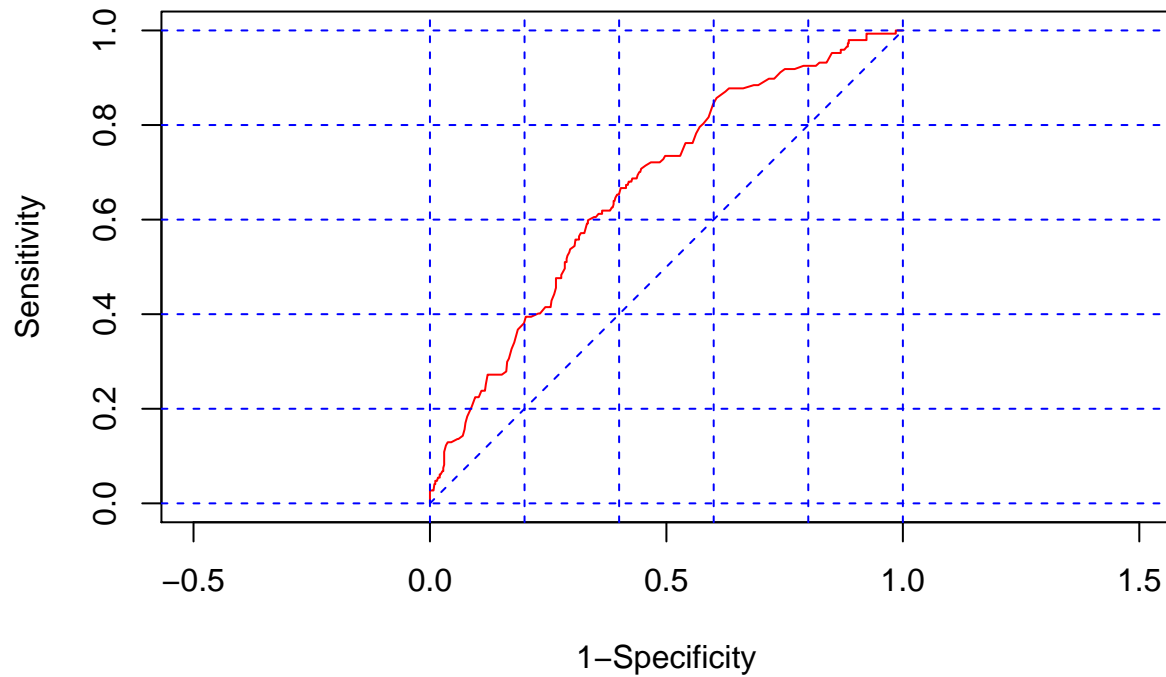


```
## [1] 0.6656775
```

```
#2E
log.reduced <- glm(hypertension ~ age + smoke,
                   family=binomial(link=logit))
summary(log.reduced)
```

```
##
## Call:
## glm(formula = hypertension ~ age + smoke, family = binomial(link = logit))
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.9359  -0.8366  -0.5680  -0.5364   2.0120
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.356461   0.470966  -2.880  0.00397 **
## age         -0.008217   0.009581  -0.858  0.39111
## smoke        1.004348   0.190711   5.266 1.39e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 709.96  on 679  degrees of freedom
## Residual deviance: 680.88  on 677  degrees of freedom
## AIC: 686.88
```

```
##
## Number of Fisher Scoring iterations: 4
```

```
#3A
log.int <- glm(hypertension ~ age + factor(race) + smoke + age*smoke,
                family = binomial(link=logit))
summary(log.int)
```

```
##
## Call:
## glm(formula = hypertension ~ age + factor(race) + smoke + age *
##     smoke, family = binomial(link = logit))
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.2251  -0.7856  -0.5376  -0.4154   2.3070
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.54740    0.68507  -0.799  0.42426
## age           -0.02796    0.01452  -1.926  0.05415 .
## factor(race)2  0.68207    0.24218   2.816  0.00486 **
## factor(race)3 -0.39184    0.34639  -1.131  0.25796
## factor(race)4 -0.17709    0.34813  -0.509  0.61096
## smoke         -0.86467    0.93459  -0.925  0.35487
## age:smoke      0.04111    0.01965   2.092  0.03647 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 709.96  on 679  degrees of freedom
## Residual deviance: 665.21  on 673  degrees of freedom
## AIC: 679.21
##
## Number of Fisher Scoring iterations: 4
```

```
#3C
1-pchisq(7.30 ,8)
```

```
## [1] 0.5046378
```