存活分析 作業2

孫浩哲 M072040002

2019年5月23日

## 1.

# (1)

require(KMsurv)

## Loading required package: KMsurv

require(survival)

## Loading required package: survival

require(dplyr)

## Loading required package: dplyr

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

require(survminer)

## Loading required package: survminer

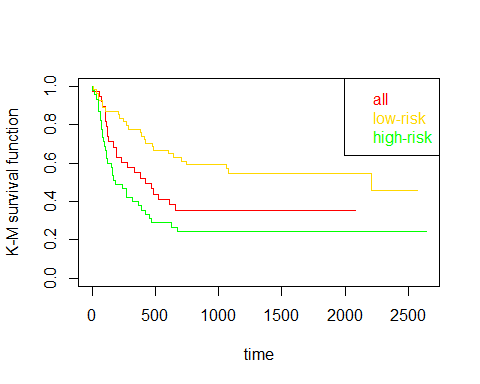
## Loading required package: ggplot2

## Registered S3 methods overwritten by 'ggplot2':  
## method from   
## [.quosures rlang  
## c.quosures rlang  
## print.quosures rlang

## Loading required package: ggpubr

## Loading required package: magrittr

data(bmt)  
fitmo=survfit(Surv(t2,d3)~factor(group),data=bmt)  
plot(fitmo,col=c("red","gold","green"),xlab="time",ylab="K-M survival function")  
legend("topright", legend = c("all", "low-risk", "high-risk"),col=c("red","gold","green"),text.col=c("red","gold","green"))



The disease-free survival functions for the three populations are not same.

# (2)

：All hazard functions are same. ：At least one hazard is different.

fitt=survdiff(Surv(t2,d3)~factor(group),data=bmt)  
fitt

## Call:  
## survdiff(formula = Surv(t2, d3) ~ factor(group), data = bmt)  
##   
## N Observed Expected (O-E)^2/E (O-E)^2/V  
## factor(group)=1 38 24 21.9 0.211 0.289  
## factor(group)=2 54 25 40.0 5.604 11.012  
## factor(group)=3 45 34 21.2 7.756 10.529  
##   
## Chisq= 13.8 on 2 degrees of freedom, p= 0.001

We reject the , that is, at least one hazard is different.

# (3)

fit3=survdiff(Surv(t2,d3)~factor(group),data=bmt,rho=1)  
fit3

## Call:  
## survdiff(formula = Surv(t2, d3) ~ factor(group), data = bmt,   
## rho = 1)  
##   
## N Observed Expected (O-E)^2/E (O-E)^2/V  
## factor(group)=1 38 16.6 15.7 0.0558 0.102  
## factor(group)=2 54 15.8 27.1 4.6533 11.920  
## factor(group)=3 45 25.7 15.4 6.8769 12.420  
##   
## Chisq= 15.7 on 2 degrees of freedom, p= 4e-04

Under , the test statistic is and the rejection region is .¶

Since , we reject the null hypothesis.

## 2.

# (1)

data(bfeed)  
fit\_model = coxph(Surv(duration,delta)~yschool+factor(poverty)+factor(race),data=bfeed)  
fit\_model %>% coef

## yschool factor(poverty)1 factor(race)2 factor(race)3   
## -0.05103355 -0.19088136 0.13219468 0.22862381

fit\_model %>% confint(level=0.99)

## 0.5 % 99.5 %  
## yschool -0.09915606 -0.002911052  
## factor(poverty)1 -0.42843566 0.046672942  
## factor(race)2 -0.13330574 0.397695102  
## factor(race)3 -0.01442330 0.471670916

# (2)

The risk will decrease years if the mother’s years of school increases one year.

# (3)

summary(fit\_model)->summ  
summ$coefficients

## coef exp(coef) se(coef) z Pr(>|z|)  
## yschool -0.05103355 0.9502468 0.01868233 -2.731648 0.006301848  
## factor(poverty)1 -0.19088136 0.8262306 0.09222440 -2.069749 0.038475847  
## factor(race)2 0.13219468 1.1413305 0.10307376 1.282525 0.199658543  
## factor(race)3 0.22862381 1.2568691 0.09435684 2.422970 0.015394188

exp(confint(fit\_model,level=0.99))

## 0.5 % 99.5 %  
## yschool 0.9056014 0.9970932  
## factor(poverty)1 0.6515275 1.0477793  
## factor(race)2 0.8751975 1.4883902  
## factor(race)3 0.9856802 1.6026699

The RR is and the 99%CI is .

# (4)

fit\_model$var

## [,1] [,2] [,3] [,4]  
## [1,] 3.490296e-04 0.0005595115 -4.438761e-05 0.0002840607  
## [2,] 5.595115e-04 0.0085053392 -1.113393e-03 -0.0004845856  
## [3,] -4.438761e-05 -0.0011133929 1.062420e-02 0.0016628964  
## [4,] 2.840607e-04 -0.0004845856 1.662896e-03 0.0089032128

sd = sqrt(0.0106+0.0089-2\*0.0017) # Var(X-Y) = Var(X)+Var(Y)-2Cov(X,Y)  
ans = 0.1322-0.2286+c(-qnorm(0.995),qnorm(0.995))\*sd  
ans %>% exp

## [1] 0.654924 1.259149

By the result we get in (3), the RR .

The CI of RR is

# (5)

：all coefficients of race are 0 ：at least one of coefficient of race are not 0.

rejection region:

##wald test  
beta1hat=fit\_model$coefficients[3:4]  
var11=fit\_model$var[3:4,3:4]  
chi=(beta1hat %>% t)%\*%solve(var11)%\*% beta1hat # test-statistic  
1-pchisq(chi,2) #chi-square distribution with df 2

## [,1]  
## [1,] 0.0360204

##likelihood ratio test  
fit\_model.reduced=coxph(Surv(duration,delta)~yschool+factor(poverty),data=bfeed)  
LR=2\*(fit\_model$loglik[2]-fit\_model.reduced$loglik[2]) # test-statistic  
1-pchisq(LR,2) #chi-square distribution with df 2

## [1] 0.04005433

##score test  
fit0=coxph(Surv(duration,delta)~yschool+factor(poverty)+factor(race),data=bfeed,init=c(fit\_model.reduced$coefficients,0,0),iter=0)  
score.vector=colSums(coxph.detail(fit0)$score)  
chiSC=t(score.vector[3:4]%\*%fit0$var[3:4,3:4]%\*%score.vector[3:4])   
# test-statistic  
1-pchisq(chiSC,2) #chi-square distribution with df 2

## [,1]  
## [1,] 0.03562763

qchisq(0.95,2)

## [1] 5.991465

No matter Wald test, likelihood ratio test and score test, we can reject the , that is, we accept at least one of coefficient of race are not 0.