存活分析 HW3

B082040005 高念慈

2023-05-22

HW from 125~

```
data(larynx)
```

cox model

```
Fit = coxph(formula = Surv(time, delta) ~ factor(stage), data = larynx)
summary(Fit)
```

```
## Call:
## coxph(formula = Surv(time, delta) ~ factor(stage), data = larynx)
##
##   n= 90, number of events= 50
##
##                   coef exp(coef) se(coef)     z Pr(>|z|)
## factor(stage)2 0.06481   1.06696  0.45843 0.141   0.8876
## factor(stage)3 0.61481   1.84930  0.35519 1.731   0.0835 .
## factor(stage)4 1.73490   5.66838  0.41939 4.137 3.52e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                exp(coef) exp(-coef) lower .95 upper .95
## factor(stage)2     1.067     0.9372    0.4344      2.62
## factor(stage)3     1.849     0.5407    0.9219      3.71
## factor(stage)4     5.668     0.1764    2.4916     12.90
##
## Concordance= 0.668  (se = 0.037 )
## Likelihood ratio test= 16.49  on 3 df,   p=9e-04
## Wald test            = 19.24  on 3 df,   p=2e-04
## Score (logrank) test = 22.88  on 3 df,   p=4e-05
```

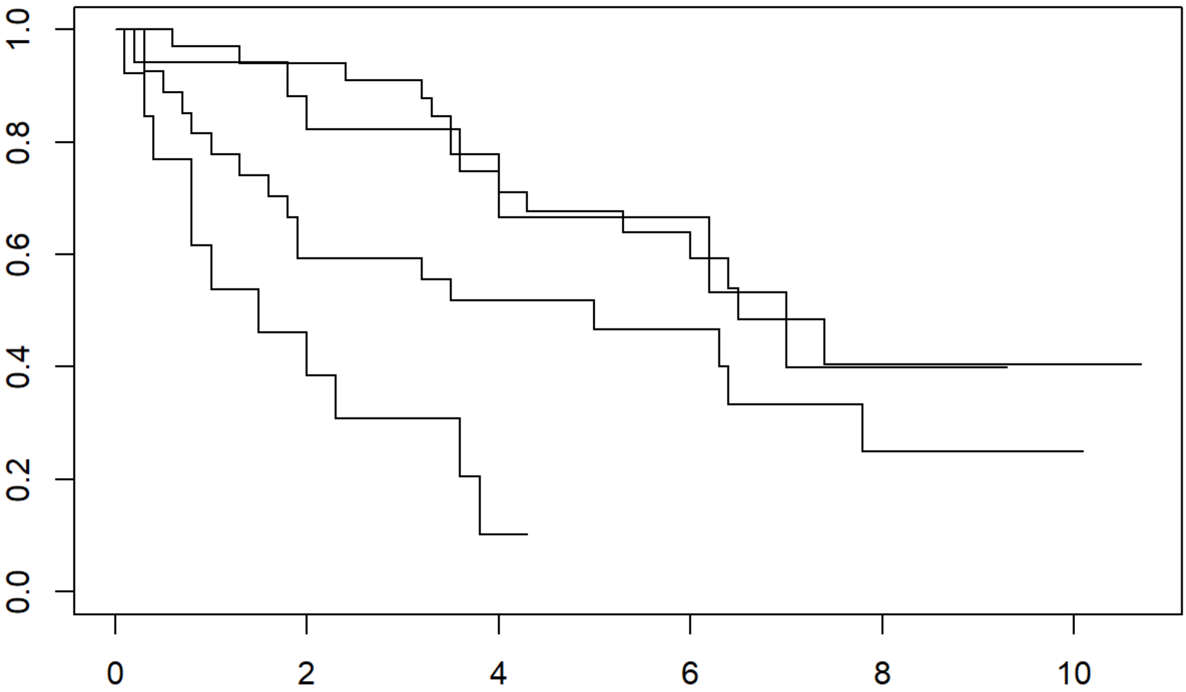
- K-M estimate

```
Fit2 = survfit(formula = Surv(time, delta) ~ factor(stage), data = larynx)
summary(Fit2)
```

```
## Call: survfit(formula = Surv(time, delta) ~ factor(stage), data = larynx)
##
##                 factor(stage)=1
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##   0.6     33       1    0.970  0.0298        0.913        1.000
##   1.3     32       1    0.939  0.0415        0.861        1.000
##   2.4     31       1    0.909  0.0500        0.816        1.000
##   3.2     29       1    0.878  0.0573        0.772        0.998
##   3.3     27       1    0.845  0.0637        0.729        0.980
##   3.5     25       2    0.778  0.0744        0.645        0.938
##   4.0     23       2    0.710  0.0819        0.566        0.890
##   4.3     21       1    0.676  0.0847        0.529        0.864
##   5.3     18       1    0.639  0.0879        0.488        0.836
##   6.0     14       1    0.593  0.0927        0.436        0.806
##   6.4     11       1    0.539  0.0987        0.377        0.772
##   6.5     10       1    0.485  0.1025        0.321        0.734
##   7.4      6       1    0.404  0.1129        0.234        0.699
##
##                 factor(stage)=2
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##   0.2     17       1    0.941  0.0571        0.836        1.000
##   1.8     16       1    0.882  0.0781        0.742        1.000
##   2.0     15       1    0.824  0.0925        0.661        1.000
##   3.6     11       1    0.749  0.1103        0.561        0.999
##   4.0      9       1    0.665  0.1255        0.460        0.963
##   6.2      5       1    0.532  0.1557        0.300        0.945
##   7.0      4       1    0.399  0.1641        0.178        0.894
##
##                 factor(stage)=3
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##   0.3     27       2    0.926  0.0504        0.832        1.000
##   0.5     25       1    0.889  0.0605        0.778        1.000
##   0.7     24       1    0.852  0.0684        0.728        0.997
##   0.8     23       1    0.815  0.0748        0.681        0.975
##   1.0     22       1    0.778  0.0800        0.636        0.952
##   1.3     21       1    0.741  0.0843        0.593        0.926
##   1.6     20       1    0.704  0.0879        0.551        0.899
##   1.8     19       1    0.667  0.0907        0.511        0.870
##   1.9     18       2    0.593  0.0946        0.433        0.810
##   3.2     16       1    0.556  0.0956        0.396        0.778
##   3.5     15       1    0.519  0.0962        0.361        0.746
##   5.0     10       1    0.467  0.0995        0.307        0.709
##   6.3      7       1    0.400  0.1053        0.239        0.670
##   6.4      6       1    0.333  0.1068        0.178        0.625
##   7.8      4       1    0.250  0.1078        0.107        0.582
##
##                 factor(stage)=4
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##   0.1     13       1    0.923  0.0739       0.7890        1.000
##   0.3     12       1    0.846  0.1001       0.6711        1.000
##   0.4     11       1    0.769  0.1169       0.5711        1.000
##   0.8     10       2    0.615  0.1349       0.4004        0.946
##   1.0      8       1    0.538  0.1383       0.3255        0.891
##   1.5      7       1    0.462  0.1383       0.2566        0.830
##   2.0      6       1    0.385  0.1349       0.1934        0.765
```

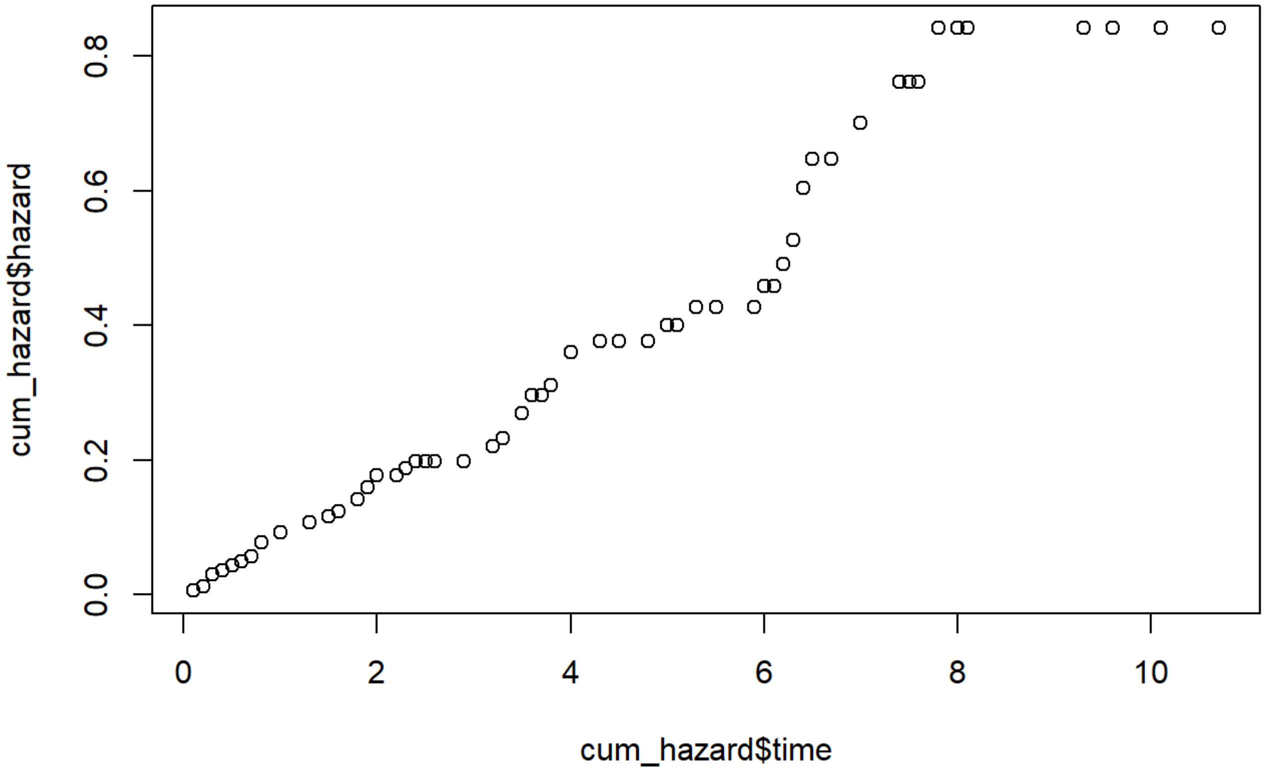
```
##    2.3       5        1     0.308  0.1280       0.1361        0.695
##    3.6       3        1     0.205  0.1196       0.0654        0.643
##    3.8       2        1     0.103  0.0940       0.0170        0.618
```

```
plot(Fit2)
```



Cox baseline cumulative hazard function

```
cum_hazard <- basehaz(Fit)
plot(cum_hazard$time, cum_hazard$hazard)
```



Global

```
Fit2 = coxph(formula = Surv(time, delta) ~ factor(stage) + age, data = larynx)
summary(Fit2)
```

```
## Call:
## coxph(formula = Surv(time, delta) ~ factor(stage) + age, data = larynx)
##
##   n= 90, number of events= 50
##
##                    coef exp(coef) se(coef)     z Pr(>|z|)
## factor(stage)2 0.14004   1.15032  0.46249 0.303   0.7620
## factor(stage)3 0.64238   1.90100  0.35611 1.804   0.0712 .
## factor(stage)4 1.70598   5.50678  0.42191 4.043 5.27e-05 ***
## age            0.01903   1.01921  0.01426 1.335   0.1820
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                exp(coef) exp(-coef) lower .95 upper .95
## factor(stage)2     1.150     0.8693    0.4647     2.848
## factor(stage)3     1.901     0.5260    0.9459     3.820
## factor(stage)4     5.507     0.1816    2.4086    12.590
## age                1.019     0.9811    0.9911     1.048
##
## Concordance= 0.682  (se = 0.039 )
## Likelihood ratio test= 18.31  on 4 df,   p=0.001
## Wald test            = 21.15  on 4 df,   p=3e-04
## Score (logrank) test = 24.78  on 4 df,   p=6e-05
```

Local

Wald test:

```
Fit2 = coxph(formula = Surv(time, delta) ~ factor(stage) + age, data = larynx)

beta1hat = Fit2$coefficients[1:3]
beta10 = rep(0,3)

var11 = Fit2$var[1:3,1:3]

chiWald = t(beta1hat-beta10)%*%solve(var11)%*%(beta1hat-beta10)
chiWald
```

```
##          [,1]
## [1,] 17.92097
```

```
1-pchisq(chiWald,3) # pvalue:0.0004566683，拒絕H0，係數不顯著為0
```

```
##              [,1]
## [1,] 0.0004566683
```

Likelihood ratio test:

- reduced model(stage係數為0)

```
Fit.reduced = coxph(Surv(time,delta) ~ age, data=larynx)
summary(Fit.reduced)
```

```
## Call:
## coxph(formula = Surv(time, delta) ~ age, data = larynx)
##
##   n= 90, number of events= 50
##
##        coef exp(coef) se(coef)     z Pr(>|z|)
## age 0.02328   1.02356  0.01449 1.607    0.108
##
##     exp(coef) exp(-coef) lower .95 upper .95
## age     1.024      0.977    0.9949     1.053
##
## Concordance= 0.555  (se = 0.046 )
## Likelihood ratio test= 2.63  on 1 df,   p=0.1
## Wald test            = 2.58  on 1 df,   p=0.1
## Score (logrank) test = 2.6  on 1 df,   p=0.1
```

```
Fit2$loglik           # Full
```

```
## [1] -196.8635 -187.7074
```

```
Fit.reduced$loglik    # Reduce
```

```
## [1] -196.8635 -195.5478
```

```
logLik(Fit2)          # df=4
```

```
## 'log Lik.' -187.7074 (df=4)
```

```
logLik(Fit.reduced)   # df=1
```

```
## 'log Lik.' -195.5478 (df=1)
```

```
chiLR = 2*(Fit$loglik[2] - Fit.reduced$loglik[2])
1-pchisq(chiLR,3)     # p-value:0.003110812，拒絕H0，係數不顯著為0 # df:4-1
```

```
## [1] 0.003110812
```

Score test:

- under H0 時的係數估計

```
Fit0 = coxph(formula = Surv(time, delta) ~ factor(stage) + age,
             data = larynx,
             init = c(0,0,0,Fit.reduced$coefficients), iter=0)
summary(Fit0)
```

```
## Call:
## coxph(formula = Surv(time, delta) ~ factor(stage) + age, data = larynx,
##     init = c(0, 0, 0, Fit.reduced$coefficients), iter = 0)
##
##   n= 90, number of events= 50
##
##                    coef exp(coef) se(coef)     z Pr(>|z|)
## factor(stage)2 0.00000   1.00000  0.39089 0.000    1.000
## factor(stage)3 0.00000   1.00000  0.34132 0.000    1.000
## factor(stage)4 0.00000   1.00000  0.58598 0.000    1.000
## age            0.02328   1.02356  0.01469 1.585    0.113
##
##                exp(coef) exp(-coef) lower .95 upper .95
## factor(stage)2     1.000      1.000    0.4648     2.151
## factor(stage)3     1.000      1.000    0.5122     1.952
## factor(stage)4     1.000      1.000    0.3171     3.153
## age                1.024      0.977    0.9945     1.053
##
## Concordance= 0.555  (se = 0.046 )
## Likelihood ratio test= 0  on 4 df,   p=1
## Wald test            = 0  on 4 df,   p=1
## Score (logrank) test = 20.98  on 4 df,   p=3e-04
```

- score.vector

```
objects(Fit0)
```

```
##  [1] "assign"            "call"              "coefficients"
##  [4] "concordance"       "contrasts"         "formula"
##  [7] "iter"              "linear.predictors" "loglik"
## [10] "means"             "method"            "n"
## [13] "nevent"            "residuals"         "score"
## [16] "terms"             "timefix"           "var"
## [19] "wald.test"         "xlevels"           "y"
```

```
Fit0$score
```

```
## [1] 20.97831
```

```
score.vector = colSums(coxph.detail(Fit0)$score)
score.vector       # -2.457351   3.089920   7.480648   5.386351e-10
```

```
## factor(stage)2 factor(stage)3 factor(stage)4            age
##  -2.457351e+00   3.089920e+00   7.480648e+00   5.386351e-10
```

```
# Fit0$var

chiSC = t(score.vector[1:3])%*%Fit0$var[1:3,1:3]%*%score.vector[1:3]
1-pchisq(chiSC,3)  # p-value:0.0001063739，拒絕H0，係數不顯著為0
```

```
##              [,1]
## [1,] 0.0001063739
```

HW:

The bfeed data frame has 927 rows and 10 columns.
Format(格式)
This data frame contains the following columns:

- Duration : Duration of breast feeding, weeks
- delta : Indicator of completed breast feeding (1=yes, 0=no)
- race : Race of mother (1=white, 2=black, 3=other)
- poverty(貧窮): Mother in poverty (1=yes, 0=no)
- yschool: Education level of mother (years of school)
- Covariates : yschool, poverty, race
- Time to event random variable : duration of breast feeding
- Right censoring indicator : delta (0: censored)

Note:
Confidence intervals (CI) below are all for 99% CI. Alpha level is 0.05 for all testing.

```
data(bfeed)
head(bfeed)
```

```
##   duration delta race poverty smoke alcohol agemth ybirth yschool pc3mth
## 1       16     1    1       0     0       1     24     82      14      0
## 2        1     1    1       0     1       0     26     85      12      0
## 3        4     0    1       0     0       0     25     85      12      0
## 4        3     1    1       0     1       1     21     85       9      0
## 5       36     1    1       0     1       0     22     82      12      0
## 6       36     1    1       0     0       0     18     82      11      0
```

(1)

We would like to investigate (調查) how the covariates
yschool, poverty and race would affect the risk function.

- cox PH model

```
model = coxph(formula = Surv(duration, delta) ~ yschool + poverty + factor(race), data = bfee
d)
summary(model)
```

```
## Call:
## coxph(formula = Surv(duration, delta) ~ yschool + poverty + factor(race),
##     data = bfeed)
##
##   n= 927, number of events= 892
##
##                    coef exp(coef) se(coef)      z Pr(>|z|)
## yschool       -0.05103   0.95025  0.01868 -2.732   0.0063 **
## poverty       -0.19088   0.82623  0.09222 -2.070   0.0385 *
## factor(race)2  0.13219   1.14133  0.10307  1.283   0.1997
## factor(race)3  0.22862   1.25687  0.09436  2.423   0.0154 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## yschool          0.9502     1.0524    0.9161    0.9857
## poverty          0.8262     1.2103    0.6896    0.9899
## factor(race)2    1.1413     0.8762    0.9326    1.3968
## factor(race)3    1.2569     0.7956    1.0447    1.5122
##
## Concordance= 0.553  (se = 0.011 )
## Likelihood ratio test= 16.77  on 4 df,   p=0.002
## Wald test            = 17.12  on 4 df,   p=0.002
## Score (logrank) test = 17.12  on 4 df,   p=0.002
```

- Write down the model

	white	black	other
$race2$	0	1	0
$race3$	0	0	1

- Model

$$h(t|z) = h_0(t)e^{-0.05103 \times yschool + 0.8262 \times poverty + 1.1413 \times race2 + 1.2569 \times race3}$$

- estimate the coefficients (beta)

$$\hat{\beta} + -Z_{\frac{\alpha}{2}} \times se(\hat{\beta})$$

```
model$coefficients
```

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

```
z = qnorm(0.995, lower.tail = TRUE)
z # 2.575829
```

```
## [1] 2.575829
```

```
sqrt(model$var)
```

```
##             [,1]      [,2]       [,3]       [,4]
## [1,] 0.01868233 0.0236540        NaN 0.01685410
## [2,] 0.02365400 0.0922244        NaN        NaN
## [3,]        NaN       NaN 0.10307376 0.04077863
## [4,] 0.01685410       NaN 0.04077863 0.09435684
```

- 99% confidence intervals

```
# yschool : -0.05103355
c(-0.05103355 - z*0.01868233, -0.05103355 + z*0.01868233)
```

```
## [1] -0.099156043 -0.002911057
```

```
# poverty : -0.19088136
c(-0.19088136 - z*0.0922244, -0.19088136 + z*0.0922244)
```

```
## [1] -0.42843567  0.04667295
```

```
# race2 : 0.13219468
c(0.13219468 - z*0.10307376, 0.13219468 + z*0.10307376)
```

```
## [1] -0.1333057  0.3976951
```

```
# race3 : 0.22862381
c(0.22862381 - z*0.09435684, 0.22862381 + z*0.09435684)
```

```
## [1] -0.0144233  0.4716709
```

```
rbind(lower_CI = model$coef - z*sqrt(diag(model$var)),
      upper_CI = model$coef + z*sqrt(diag(model$var)))
```

```
##               yschool     poverty factor(race)2 factor(race)3
## lower_CI -0.099156056 -0.42843566    -0.1333057    -0.0144233
## upper_CI -0.002911052  0.04667294     0.3976951     0.4716709
```

Please explain your results.

- 有 99% 的信心水準各係數會落在各信賴區間內
- exp(coef)

1. yschool : 0.95025

全母乳喂養的機率，每多讀一年書多0.95025倍(多兩年多0.95025^2倍)

2. poverty : 0.82623

全母乳喂養的機率，貧窮比不貧窮多0.82623倍

3. race

- race2 : 1.14133

全母乳喂養的機率，黑人比白人多1.14133倍

- race3 : 1.25687

全母乳喂養的機率，其他人種比白人多1.25687倍

(2)

If the mother’s years of school increases one year, how would the risk is changed.

- yschool:-0.05103355，exp(-0.05103355)=0.9502468

全母乳喂養的機率，每多讀一年書多 0.95025 倍(多兩年多0.95025^2倍)

(3)

Calculate the RR (relative risk) for mothers in poverty relative to mothers not in poverty. Calculate the CI for the RR.

- poverty:-0.19088136，exp(-0.19088136)=0.82623

全母乳喂養的機率，貧窮比不貧窮多0.82623倍

- relative risk CI

$$P(e^{\hat{\beta}-Z_{\frac{\alpha}{2}}\times se(\hat{\beta})} < e^{\beta} < e^{\hat{\beta}+Z_{\frac{\alpha}{2}}\times se(\hat{\beta})}) = 0.99$$

```
# poverty : -0.19088136
relative_risk = exp(-0.19088136)
relative_risk                          # 0.8262306
```

```
## [1] 0.8262306
```

```
lower_poverty = (model$coef - z*sqrt(diag(model$var)))[2]
upper_poverty = (model$coef + z*sqrt(diag(model$var)))[2]
c(exp(lower_poverty), exp(upper_poverty))
```

```
##   poverty   poverty
## 0.6515275 1.0477793
```

(4)

Calculate the RR for mothers whose race are black relative to mothers **whose race is white**. Calculate the CI for the RR

```
# race2 : 0.13219468
relative_risk_WB = exp(0.13219468)
relative_risk_WB                          # 1.14133
```

```
## [1] 1.14133
```

- relative risk CI

```
lower_race_B = (model$coef - z*sqrt(diag(model$var)))[3]
upper_race_B = (model$coef + z*sqrt(diag(model$var)))[3]
c(exp(lower_race_B), exp(upper_race_B))
```

```
## factor(race)2 factor(race)2
##     0.8751975     1.4883902
```

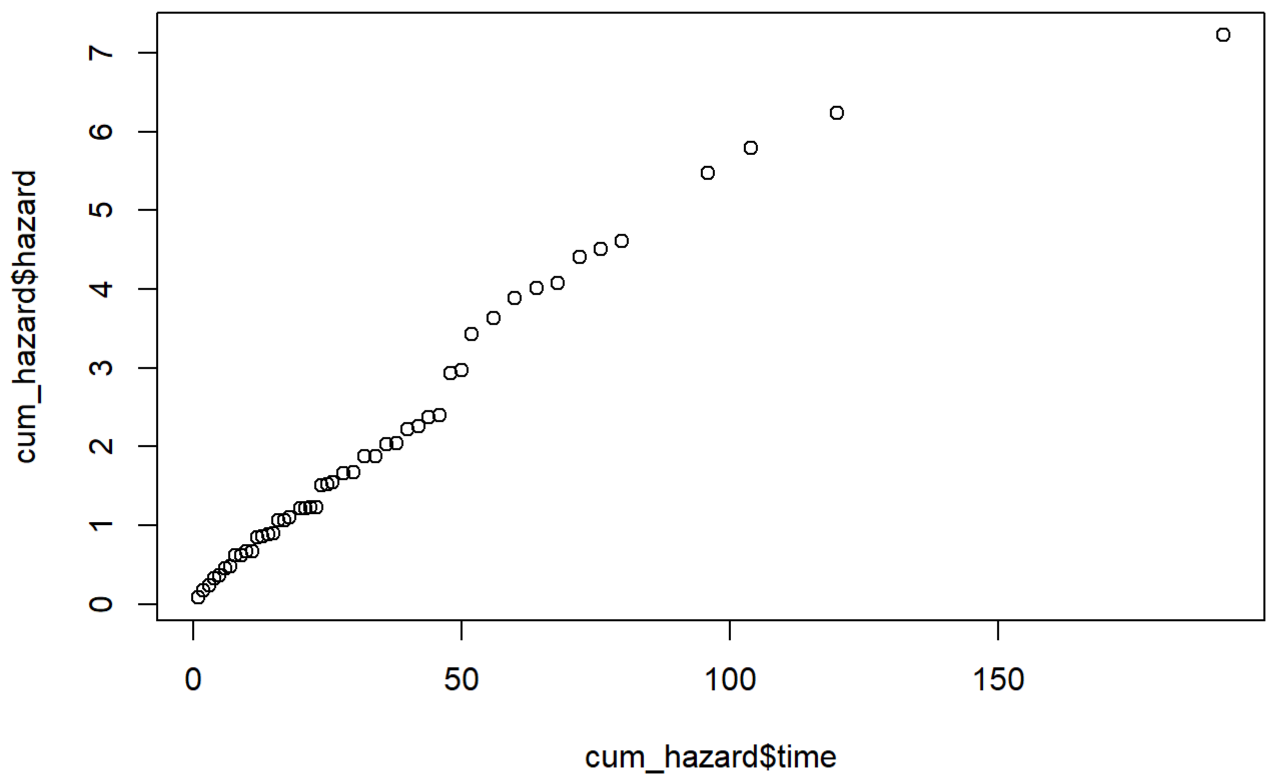
(5)

Estimate the Breslow’s baseline hazard function and plot the estimate of cumulative baseline hazard function.

- 使用 Cox 模型進行預測建模 - 所有關於基線危害 (https://missingdatasolutions.rbind.io/2022/12/cox-baseline-hazard/)

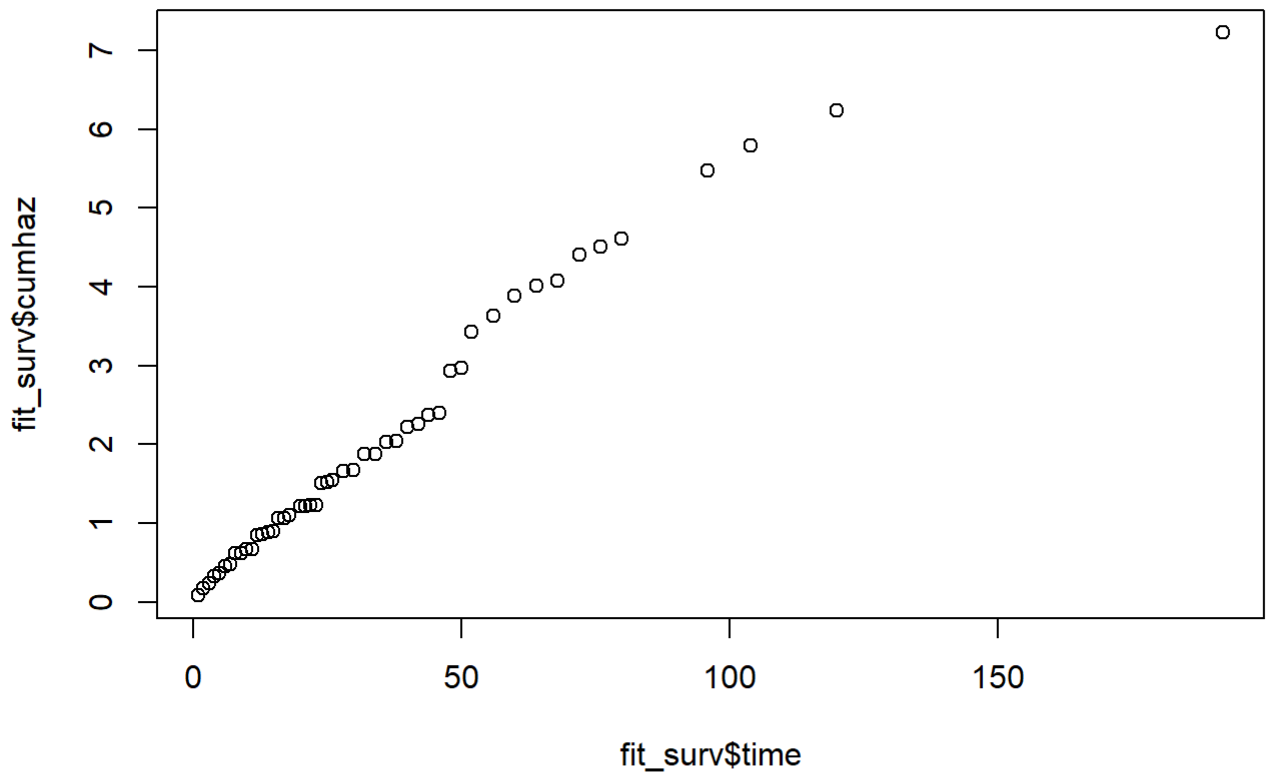
cumulative baseline hazard function

```
cum_hazard <- basehaz(model,centered = T) # model$means
plot(cum_hazard$time, cum_hazard$hazard)
```



cumulative baseline hazard function

```
fit_surv <- survfit(model)             # fit_surv$cumhaz
plot(fit_surv$time, fit_surv$cumhaz)
```




```
cum_hazard$hazard == fit_surv$cumhaz
```

```
##  [1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [16] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [31] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
## [46] TRUE TRUE TRUE TRUE
```

baseline hazard function

```
hazard = data.frame(hazard = c(cum_hazard$hazard[1],diff(cum_hazard$hazard)),time = cum_hazar
d$time)
hazard
```

```
##         hazard time
## 1  0.084348725    1
## 2  0.085230712    2
## 3  0.063957463    3
## 4  0.100024938    4
## 5  0.029206617    5
## 6  0.092033037    6
## 7  0.026467246    7
## 8  0.138717252    8
## 9  0.006254606    9
## 10 0.040645514   10
## 11 0.004385720   11
## 12 0.181613019   12
## 13 0.013476553   13
## 14 0.019259645   14
## 15 0.014023813   15
## 16 0.165509126   16
## 17 0.003358435   17
## 18 0.034268302   18
## 19 0.111315362   20
## 20 0.007875803   21
## 21 0.007934263   22
## 22 0.000000000   23
## 23 0.286494022   24
## 24 0.005395066   25
## 25 0.027388334   26
## 26 0.111761793   28
## 27 0.012568385   30
## 28 0.202287794   32
## 29 0.007747947   34
## 30 0.141596188   36
## 31 0.018024384   38
## 32 0.178601368   40
## 33 0.033065949   42
## 34 0.118215009   44
## 35 0.025290794   46
## 36 0.529583950   48
## 37 0.045019071   50
## 38 0.459261753   52
## 39 0.200406348   56
## 40 0.253770071   60
## 41 0.123998967   64
## 42 0.068680484   68
## 43 0.327240076   72
## 44 0.100387811   76
## 45 0.110511833   80
## 46 0.857698197   96
## 47 0.320938567  104
## 48 0.438432871  120
## 49 0.989485610  192
```

(6)

Test whether the race is a significantly important for risk.
Use Wald test, likelihood ratio test, and the score test.
You have to write

- the null hypothesis/alternative hypothesis
- the form of the test statistic and its distribution with the corresponding degrees of freedom under Ho
- the rejection region
- your conclusions.

Local

Wald test:

```
### Global
model = coxph(formula = Surv(duration, delta) ~ yschool + poverty + factor(race), data = bfee
d)
summary(model)
```

```
## Call:
## coxph(formula = Surv(duration, delta) ~ yschool + poverty + factor(race),
##     data = bfeed)
##
##   n= 927, number of events= 892
##
##                   coef exp(coef) se(coef)      z Pr(>|z|)
## yschool       -0.05103   0.95025  0.01868 -2.732   0.0063 **
## poverty       -0.19088   0.82623  0.09222 -2.070   0.0385 *
## factor(race)2  0.13219   1.14133  0.10307  1.283   0.1997
## factor(race)3  0.22862   1.25687  0.09436  2.423   0.0154 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## yschool          0.9502     1.0524    0.9161    0.9857
## poverty          0.8262     1.2103    0.6896    0.9899
## factor(race)2    1.1413     0.8762    0.9326    1.3968
## factor(race)3    1.2569     0.7956    1.0447    1.5122
##
## Concordance= 0.553  (se = 0.011 )
## Likelihood ratio test= 16.77  on 4 df,   p=0.002
## Wald test            = 17.12  on 4 df,   p=0.002
## Score (logrank) test = 17.12  on 4 df,   p=0.002
```

```
### Local
beta1hat = model$coefficients[3:4]
beta10 = rep(0,2)

var11 = model$var[3:4,3:4]

chiWald = t(beta1hat-beta10)%*%solve(var11)%*%(beta1hat-beta10)
chiWald
```

```
##         [,1]
## [1,] 6.64734
```

```
1-pchisq(chiWald,2)
```

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

- the null hypothesis/alternative hypothesis $H_0 : \beta^1 = \beta_{10} = (0,0) vs \;\; H_a : \beta^1 \neq \beta_{10}$
- the form of the test statistic and its distribution with the corresponding degrees of freedom under Ho $under \; H_0 \; , \chi^2_w = (\hat{\beta^1} - \beta_{10})^T (Var(\hat{\beta^1}))^{-1} (\hat{\beta^1} - \beta_{10})$~$\chi^2_2$
- the rejection region $\chi^2_w > \chi^2_{2,0.05} \; or p - value < 0.05$
- your conclusions.
- pvalue:0.0360204 < 0.05，拒絕H0，係數不顯著為0，種族對全母乳喂養有顯著影響

Likelihood ratio test:

- reduced model(race係數為0)

```
model.reduced = coxph(formula = Surv(duration, delta) ~ yschool + poverty, data = bfeed)
summary(model.reduced)
```

```
## Call:
## coxph(formula = Surv(duration, delta) ~ yschool + poverty, data = bfeed)
##
##   n= 927, number of events= 892
##
##             coef exp(coef) se(coef)      z Pr(>|z|)
## yschool -0.05713   0.94447  0.01855 -3.080  0.00207 **
## poverty -0.17016   0.84353  0.09168 -1.856  0.06344 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##         exp(coef) exp(-coef) lower .95 upper .95
## yschool    0.9445      1.059    0.9108    0.9794
## poverty    0.8435      1.185    0.7048    1.0096
##
## Concordance= 0.544  (se = 0.011 )
## Likelihood ratio test= 10.34  on 2 df,   p=0.006
## Wald test            = 10.23  on 2 df,   p=0.006
## Score (logrank) test = 10.21  on 2 df,   p=0.006
```

```
model$loglik            # Full
```

```
## [1] -5191.115 -5182.730
```

```
model.reduced$loglik    # Reduce
```

```
## [1] -5191.115 -5185.947
```

```
logLik(model)           # df=4
```

```
## 'log Lik.' -5182.73 (df=4)
```

```
logLik(model.reduced)   # df=2
```

```
## 'log Lik.' -5185.947 (df=2)
```

```
chiLR = 2*(model$loglik[2] - model.reduced$loglik[2])
1-pchisq(chiLR,4-2)     # df:4-2
```

```
## [1] 0.04005433
```

- the null hypothesis/alternative hypothesis $H_0 : \beta^1 = \beta_{10} = (0, 0) vs \;\; H_a : \beta^1 \neq \beta_{10}$
- the form of the test statistic and its distribution with the corresponding degrees of freedom under Ho $under \; H_0 \; , \chi^2_{LR} = 2(LL(Full) - LL(Reduced))$~$\chi^2_{4-2}$
- the rejection region $\chi^2_{LR} > \chi^2_{2,0.05} \; orp - value < 0.05$
- your conclusions.

- p-value:0.04005433 < 0.05，拒絕H0，係數不顯著為0，種族對全母乳喂養有顯著影響

Score test:

- under H0 時的係數估計

```
model0 = coxph(formula = Surv(duration, delta) ~ yschool + poverty + factor(race),
               data = bfeed,
               init = c(model.reduced$coefficients,0,0), iter=0)
summary(model0)
```

```
## Call:
## coxph(formula = Surv(duration, delta) ~ yschool + poverty + factor(race),
##     data = bfeed, init = c(model.reduced$coefficients, 0, 0),
##     iter = 0)
##
##   n= 927, number of events= 892
##
##                    coef exp(coef) se(coef)      z Pr(>|z|)
## yschool       -0.05713   0.94447  0.01879 -3.040  0.00237 **
## poverty       -0.17016   0.84353  0.09235 -1.843  0.06540 .
## factor(race)2  0.00000   1.00000  0.10627  0.000  1.00000
## factor(race)3  0.00000   1.00000  0.10059  0.000  1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## yschool          0.9445      1.059    0.9103    0.9799
## poverty          0.8435      1.185    0.7039    1.0109
## factor(race)2    1.0000      1.000    0.8120    1.2316
## factor(race)3    1.0000      1.000    0.8211    1.2179
##
## Concordance= 0.544  (se = 0.011 )
## Likelihood ratio test= 0  on 4 df,   p=1
## Wald test            = 0  on 4 df,   p=1
## Score (logrank) test = 6.67  on 4 df,   p=0.2
```

- score.vector

```
objects(model0)
```

```
##  [1] "assign"            "call"              "coefficients"
##  [4] "concordance"       "contrasts"         "formula"
##  [7] "iter"              "linear.predictors" "loglik"
## [10] "means"             "method"            "n"
## [13] "nevent"            "residuals"         "score"
## [16] "terms"             "timefix"           "var"
## [19] "wald.test"         "xlevels"           "y"
```

```
model0$score
```

```
## [1] 6.669268
```

```
score.vector = colSums(coxph.detail(model0)$score)
score.vector   # 6.004529e-10 -4.278418e-10  8.718373e+00  2.263892e+01
```

```
##       yschool       poverty factor(race)2 factor(race)3
##  6.004529e-10 -4.278418e-10  8.718373e+00  2.263892e+01
```

```
# model0$var

chiSC = t(score.vector[3:4])%*%model0$var[3:4,3:4]%*%score.vector[3:4]
1-pchisq(chiSC,2)
```

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

- the null hypothesis/alternative hypothesis $H_0 : \beta^1 = \beta_{10} = (0,0) vs \;\; H_a : \beta^1 \neq \beta_{10}$
- the form of the test statistic and its distribution with the corresponding degrees of freedom under Ho $under \; H_0 \;, \chi^2_{sc} = (\frac{\partial LL}{\partial \beta^1}(\beta_{10}, \hat{\beta^2}(\beta_{10})))^T (Var((\frac{\partial LL}{\partial \beta^1})(\beta_{10}, \hat{\beta^2}(\beta_{10}))))^{-1} (\frac{\partial LL}{\partial \beta^1}(\beta_{10}, \hat{\beta^2}(\beta_{10})))$~ χ^2_2
- the rejection region $\chi^2_{sc} > \chi^2_{2,0.05} \; or p - value < 0.05$
- your conclusions.
- p-value:0.03562763 < 0.05，拒絕H0，係數不顯著為0，種族對全母乳喂養有顯著影響