```
> library(survival)
> ###
> #Q1
> ###
> time = c(18,9,28,31,39,19,45,6,8,15,23,28,7,12,9,8,2,26,10,4,3,4,18,8,3,14,3,13,13,35)
> status = c(1,1,0,1,0,0,0,1,1,1,1,0,1,1,1,1,1,0,1,1,1,1,1,1,1,1,1,1,1,0)
> x1 = c(0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1)
> x2 = c(0,1,0,1,1,1,1,1,1,1,1,0,0,1,0,0,0,1,0,1,0,0,0,1,1,1,1,0,1,1,0)
> survival_data = Surv(time, event = status)
>
> #Non-parametric approach
> survdiff(formula = survival_data ~ x1+x2)
Call:
survdiff(formula = survival_data ~ x1 + x2)

           N Observed Expected (O-E)^2/E (O-E)^2/V
x1=0, x2=0 4        2     4.92     1.736     2.433
x1=0, x2=1 9        6     8.16     0.571     0.961
x1=1, x2=0 9        7     5.11     0.697     0.954
x1=1, x2=1 8        8     4.80     2.126     2.965

 Chisq= 5.7  on 3 degrees of freedom, p= 0.1
>
> #Semi-parametric approach
> coxph(formula = survival_data ~ x1+x2)
Call:
coxph(formula = survival_data ~ x1 + x2)

     coef exp(coef) se(coef)     z      p
x1 1.0463    2.8472   0.4581 2.284 0.0224
x2 0.3586    1.4313   0.4401 0.815 0.4152

Likelihood ratio test=5.76  on 2 df, p=0.05607
n= 30, number of events= 23
>
```

```
> ###
> #2a
> ###
> set.seed(123457)
> s<-sample(1:1601, size=1000)
> Q4 = read.csv("ass4.csv")[s,]
> colnames(Q4) = c("lstay", "age", "trt", "gender", "marstat", "hlstat", "cens")
> t = Surv(Q4$lstay, 1-Q4$cens)
>   age = Q4$age
>   trt = Q4$trt
>   gender = Q4$gender
>   marstat = Q4$marstat
>   hlstat2 = (Q4$hlstat==2)+0
>   hlstat3 = (Q4$hlstat==3)+0
>   hlstat4 = (Q4$hlstat==4)+0
>
>  ### Step 1: Full Model
>    m0 = coxph( t~age + trt + gender + marstat + hlstat2 + hlstat3 + hlstat4 )
>    summary(m0)
Call:
coxph(formula = t ~ age + trt + gender + marstat + hlstat2 +
    hlstat3 + hlstat4)

  n= 1000, number of events= 801

              coef exp(coef)  se(coef)      z Pr(>|z|)
age       0.001168  1.001169  0.004817  0.242 0.808439
trt      -0.043763  0.957181  0.073271 -0.597 0.550324
gender    0.350924  1.420380  0.086931  4.037 5.42e-05 ***
marstat   0.207470  1.230561  0.102052  2.033 0.042055 *
hlstat2  -0.326060  0.721762  0.132839 -2.455 0.014106 *
hlstat3  -0.421340  0.656167  0.123798 -3.403 0.000665 ***
hlstat4  -0.123711  0.883635  0.123936 -0.998 0.318190
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

         exp(coef) exp(-coef) lower .95 upper .95
age         1.0012     0.9988    0.9918    1.0107
trt         0.9572     1.0447    0.8291    1.1050
gender      1.4204     0.7040    1.1979    1.6842
marstat     1.2306     0.8126    1.0075    1.5030
hlstat2     0.7218     1.3855    0.5563    0.9364
hlstat3     0.6562     1.5240    0.5148    0.8364
hlstat4     0.8836     1.1317    0.6931    1.1266

Concordance= 0.574  (se = 0.011 )
Likelihood ratio test= 47.41  on 7 df,   p=5e-08
Wald test            = 49.75  on 7 df,   p=2e-08
Score (logrank) test = 50.38  on 7 df,   p=1e-08
```

```
>
>  ### Step 2: Reduced Model
>    m1 = coxph( t~            gender + marstat + hlstat2 + hlstat3)
>    summary(m1)
Call:
coxph(formula = t ~ gender + marstat + hlstat2 + hlstat3)

  n= 1000, number of events= 801

            coef exp(coef) se(coef)      z Pr(>|z|)
gender   0.34594   1.41332  0.08537  4.052 5.07e-05 ***
marstat  0.20775   1.23090  0.10067  2.064   0.0390 *
hlstat2 -0.22374   0.79952  0.09244 -2.420   0.0155 *
hlstat3 -0.32308   0.72392  0.08088 -3.994 6.49e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

        exp(coef) exp(-coef) lower .95 upper .95
gender     1.4133     0.7076    1.1956    1.6707
marstat    1.2309     0.8124    1.0105    1.4994
hlstat2    0.7995     1.2507    0.6670    0.9583
hlstat3    0.7239     1.3814    0.6178    0.8483

Concordance= 0.573  (se = 0.011 )
Likelihood ratio test= 46.16  on 4 df,   p=2e-09
Wald test            = 48.41  on 4 df,   p=8e-10
Score (logrank) test = 48.99  on 4 df,   p=6e-10
```

```
> ###
> #2b
> ###
>
> #Wald test
>    x<-c(1,2,7)
>    subpar<-m0$coefficient[x]
>    subvar<-m0$var[x,x]
>    subvarin<-solve(subvar)
>    subwaldtest<-t(subpar)%*%subvarin%*%subpar
>    subdf<-length(m0$coefficient)-length(m1$coefficient)
>    rbind(c("sub-wald","df","pval"),c(subwaldtest,subdf,1-pchisq(subwaldtest,subdf)))
     [,1]              [,2] [,3]
[1,] "sub-wald"        "df" "pval"
[2,] "1.26832234350982" "3"  "0.736668636389971"
>
> #Likelihood Ratio test
>    LRtest<-2*(m0$loglik[2]-m1$loglik[2])
>    df=length(m0$coefficient)-length(m1$coefficient)
>    rbind(c("LR","df","pval"),c(LRtest,df,1-pchisq(LRtest,df)))
     [,1]              [,2] [,3]
[1,] "LR"              "df" "pval"
[2,] "1.25324025656118" "3"  "0.740265425327963"
>
> #Score test
>    initial<-rep(0,7)
>    xx<-c(3,4,5,6)
>    initial[xx]<-m1$coefficient
>    scoretest<-coxph( t~age + trt + gender + marstat + hlstat2 + hlstat3 + hlstat4,init=initial )
>    df<-length(m0$coefficient)-length(m1$coefficient)
>    rbind(c("score","df","pval"),c(scoretest$score,df,1-pchisq(scoretest$score,df)))
     [,1]              [,2] [,3]
[1,] "score"           "df" "pval"
[2,] "1.26923367940956" "3"  "0.736451481448187"
```