

Untitled

Nikhil Gopal

12/4/2020

Code to setup everything:

```
data <- read.csv("C:/Users/d/Google Drive/Notability/Applied Linear Regression Analysis/psets/10/10bdata.csv")
library(survival)
```

```
## Warning: package 'survival' was built under R version 4.0.3
```

```
library(survminer)
```

```
## Warning: package 'survminer' was built under R version 4.0.3
```

```
## Loading required package: ggplot2
```

```
## Loading required package: ggpubr
```

```
## Warning: package 'ggpubr' was built under R version 4.0.3
```

```
#models for transplants existing

#change health and status to factors bc they are categorical data

data$health = as.factor(data$health)
data$transplant_factor = ifelse(!is.na(data$transplant), 1,0)
data$transplant_factor = as.factor(data$transplant_factor)

#1 = failure, 0 = censoring
time.ind.pairs = Surv(time=data$time, event = data$status)

#kaplan meir no transplant
kaplan_meir_fit <- survfit(time.ind.pairs ~ health, data = data)

#estimates of survival function for each level of health
summary(kaplan_meir_fit)
```

```
## Call: survfit(formula = time.ind.pairs ~ health, data = data)
```

```
##
```

```
##           health=0
```

```
##      time n.risk n.event survival std.err lower 95% CI upper 95% CI
```

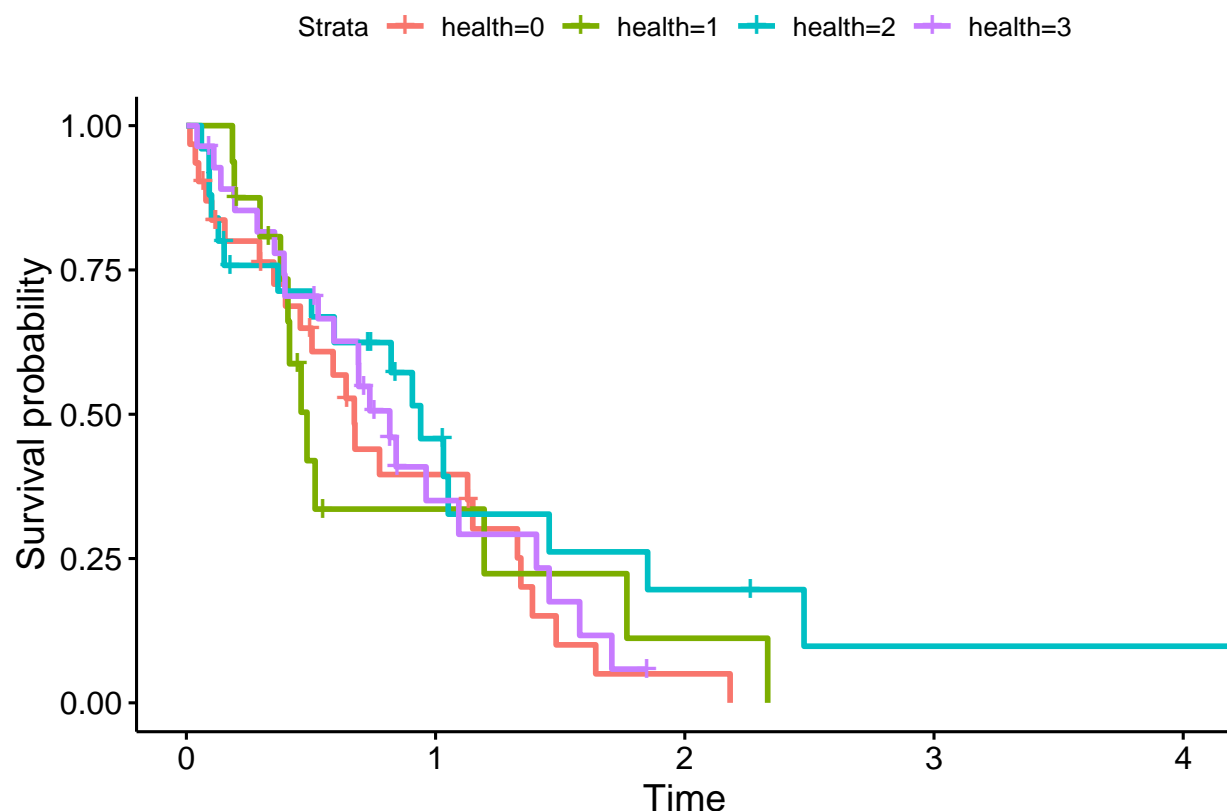
```

## 0.0144      31      1  0.9677  0.0317      0.90750      1.000
## 0.0360      30      1  0.9355  0.0441      0.85288      1.000
## 0.0490      29      1  0.9032  0.0531      0.80492      1.000
## 0.0786      27      1  0.8698  0.0608      0.75847      0.997
## 0.1014      26      1  0.8363  0.0670      0.71478      0.979
## 0.1546      23      1  0.8000  0.0733      0.66846      0.957
## 0.2933      22      1  0.7636  0.0785      0.62430      0.934
## 0.3511      20      1  0.7254  0.0833      0.57919      0.909
## 0.3979      19      1  0.6872  0.0872      0.53586      0.881
## 0.4571      18      1  0.6491  0.0904      0.49405      0.853
## 0.5040      16      1  0.6085  0.0934      0.45043      0.822
## 0.5890      15      1  0.5679  0.0956      0.40838      0.790
## 0.6409      14      1  0.5274  0.0970      0.36779      0.756
## 0.6738      12      1  0.4834  0.0983      0.32446      0.720
## 0.6764      11      1  0.4395  0.0987      0.28294      0.683
## 0.7750      10      1  0.3955  0.0982      0.24318      0.643
## 1.1285       9      1  0.3516  0.0966      0.20520      0.602
## 1.1491       7      1  0.3013  0.0950      0.16250      0.559
## 1.3284       6      1  0.2511  0.0915      0.12300      0.513
## 1.3422       5      1  0.2009  0.0859      0.08694      0.464
## 1.3886       4      1  0.1507  0.0777      0.05484      0.414
## 1.4838       3      1  0.1004  0.0661      0.02767      0.365
## 1.6427       2      1  0.0502  0.0485      0.00757      0.333
## 2.1819       1      1  0.0000      NaN      NA      NA
##
##
##              health=1
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
## 0.185   16      1   0.938  0.0605   0.8261   1.000
## 0.192   15      1   0.875  0.0827   0.7271   1.000
## 0.295   13      1   0.808  0.1000   0.6336   1.000
## 0.377   11      1   0.734  0.1148   0.5405   0.997
## 0.407   10      1   0.661  0.1246   0.4567   0.956
## 0.414    9      1   0.587  0.1306   0.3799   0.908
## 0.460    7      1   0.503  0.1363   0.2962   0.856
## 0.484    6      1   0.420  0.1370   0.2213   0.796
## 0.517    5      1   0.336  0.1328   0.1546   0.729
## 1.195    3      1   0.224  0.1272   0.0734   0.682
## 1.768    2      1   0.112  0.1015   0.0189   0.662
## 2.333    1      1   0.000      NaN      NA      NA
##
##
##              health=2
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
## 0.0601   25      1   0.9600  0.0392   0.8862   1.000
## 0.0912   24      1   0.9200  0.0543   0.8196   1.000
## 0.0919   23      1   0.8800  0.0650   0.7614   1.000
## 0.1010   22      1   0.8400  0.0733   0.7079   0.997
## 0.1284   21      1   0.8000  0.0800   0.6576   0.973
## 0.1515   19      1   0.7579  0.0862   0.6065   0.947
## 0.3676   17      1   0.7133  0.0919   0.5541   0.918
## 0.5024   16      1   0.6687  0.0964   0.5042   0.887
## 0.5932   15      1   0.6241  0.0997   0.4563   0.854
## 0.8216   12      1   0.5721  0.1041   0.4005   0.817
## 0.9068   10      1   0.5149  0.1083   0.3410   0.778
## 0.9400    9      1   0.4577  0.1103   0.2854   0.734

```

```
## 1.0314      7      1  0.3923  0.1123      0.2239      0.687
## 1.0507      6      1  0.3269  0.1110      0.1681      0.636
## 1.4558      5      1  0.2615  0.1063      0.1179      0.580
## 1.8510      4      1  0.1962  0.0978      0.0738      0.521
## 2.4784      2      1  0.0981  0.0849      0.0180      0.535
## 4.4675      1      1  0.0000    NaN          NA          NA
##
##               health=3
##      time n.risk n.event survival std.err lower 95% CI upper 95% CI
## 0.0429    28      1  0.9643  0.0351    0.89794    1.000
## 0.1099    26      1  0.9272  0.0496    0.83491    1.000
## 0.1383    25      1  0.8901  0.0599    0.78013    1.000
## 0.1942    24      1  0.8530  0.0679    0.72977    0.997
## 0.2843    23      1  0.8159  0.0744    0.68239    0.976
## 0.3535    22      1  0.7788  0.0797    0.63725    0.952
## 0.3918    21      1  0.7418  0.0841    0.59392    0.926
## 0.3956    20      1  0.7047  0.0877    0.55212    0.899
## 0.5285    18      1  0.6655  0.0912    0.50883    0.870
## 0.5921    17      1  0.6264  0.0938    0.46702    0.840
## 0.6903    16      1  0.5872  0.0958    0.42655    0.808
## 0.6912    15      1  0.5481  0.0971    0.38734    0.776
## 0.7360    13      1  0.5059  0.0983    0.34565    0.740
## 0.8161    11      1  0.4599  0.0996    0.30089    0.703
## 0.8417     9      1  0.4088  0.1008    0.25219    0.663
## 0.9619     7      1  0.3504  0.1019    0.19818    0.620
## 1.0931     6      1  0.2920  0.1003    0.14898    0.572
## 1.4047     5      1  0.2336  0.0957    0.10464    0.522
## 1.4557     4      1  0.1752  0.0878    0.06560    0.468
## 1.5782     3      1  0.1168  0.0755    0.03290    0.415
## 1.7071     2      1  0.0584  0.0560    0.00893    0.382
```

```
ggsurvplot(kaplan_meir_fit, data = data)
```



```
#Proportional hazards regression, with transplant data included (both transplants and no transplants)
phr.fit.transplants = coxph(time.ind.pairs ~ health + transplant_factor+transplant, data = data)

#models for transplants not existing
data_with_transplant_removed = subset(data, subset = is.na(data$transplant))
time.ind.pairs.nt <- Surv(time = data_with_transplant_removed$time, event = data_with_transplant_removed$event)

#kaplan meir fit no transplant
kmf.nt <- survfit(time.ind.pairs.nt ~ health, data = data_with_transplant_removed)

phr.fit.nt = coxph(time.ind.pairs.nt ~ health, data = data_with_transplant_removed)
```

Just the models:

```
#just the models

#transplants exist
summary(phr.fit.transplants)
```

```
## Call:
## coxph(formula = time.ind.pairs ~ health + transplant_factor +
##       transplant, data = data)
##
## n= 37, number of events= 27
```

```
## (63 observations deleted due to missingness)
##
##               coef exp(coef) se(coef)      z Pr(>|z|)
## health1      0.1703    1.1856   0.8640  0.197  0.84378
## health2     -2.0422    0.1297   0.6652 -3.070  0.00214 **
## health3     -0.1577    0.8541   0.5440 -0.290  0.77196
## transplant_factor1      NA         NA  0.0000      NA      NA
## transplant    -2.7151    0.0662   0.6365 -4.266 1.99e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## health1      1.1856    0.8434   0.21802   6.4475
## health2      0.1297    7.7076   0.03522   0.4779
## health3      0.8541    1.1708   0.29410   2.4807
## transplant_factor1      NA         NA      NA      NA
## transplant    0.0662   15.1059   0.01901   0.2305
##
## Concordance= 0.838 (se = 0.035 )
## Likelihood ratio test= 37.34 on 4 df,  p=2e-07
## Wald test              = 21.63 on 4 df,  p=2e-04
## Score (logrank) test = 24.76 on 4 df,  p=6e-05
```

```
#transplants dont exist
summary(phr.fit.nt)
```

```
## Call:
## coxph(formula = time.ind.pairs.nt ~ health, data = data_with_transplant_removed)
##
## n= 63, number of events= 48
##
##               coef exp(coef) se(coef)      z Pr(>|z|)
## health1  0.22785   1.25590  0.43989  0.518   0.604
## health2  0.67239   1.95891  0.43415  1.549   0.121
## health3 -0.07184   0.93068  0.39932 -0.180   0.857
##
##               exp(coef) exp(-coef) lower .95 upper .95
## health1    1.2559    0.7962    0.5303    2.974
## health2    1.9589    0.5105    0.8365    4.587
## health3    0.9307    1.0745    0.4255    2.036
##
## Concordance= 0.567 (se = 0.046 )
## Likelihood ratio test= 3.49 on 3 df,  p=0.3
## Wald test              = 3.76 on 3 df,  p=0.3
## Score (logrank) test = 3.89 on 3 df,  p=0.3
```

In the model with transplants, it appears that the only statistically significant factors are health status and transplant time.