Assignment3-StochasticModels

Claijon da Silva

8/5/2017

Introduction

This exercice will make the use of Kaplan-Meier estimate and Cox Regression to perform a Survival analysis on a data set of 1445 criminals. R Programming language was used to assist in the analysis.

Data

The frist few rows show the structure of the table.

```
##   durat fail drugs black married educ
## 1    72    0     0     0       1    7
## 2    75    0     0     1       0   12
## 3     9    1     0     0       0    9
## 4    25    1     1     0       0    9
## 5    81    0     1     0       0    9
## 6    79    0     0     1       0   12
```

- durat Time observed for
- fail 1=Reoffended; 0=Censored
- drugs 1=Yes; 0=No
- black 1=Yes; 0=No
- married 1=Yes; 0=No
- educ Years of schooling

Time To Reoffend

Showing that meddian 71 and mean 55.37 are the average times for reoffending.

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.00   27.00   71.00   55.37   76.00   81.00
```

Showing that 38.2% of this population experienced the failure event (did not reoffend)

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   0.000   0.000   0.382   1.000   1.000
```

Kaplan-Meier Estimate of time to reoffending for this population

```
## [1] 72+ 75+  9  25  81+ 79+
```

Fitting the survival

```
## Call: survfit(formula = Surv(time, event) ~ 1, data = offend)
##
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##     1   1445       8    0.994 0.00195        0.991        0.998
```

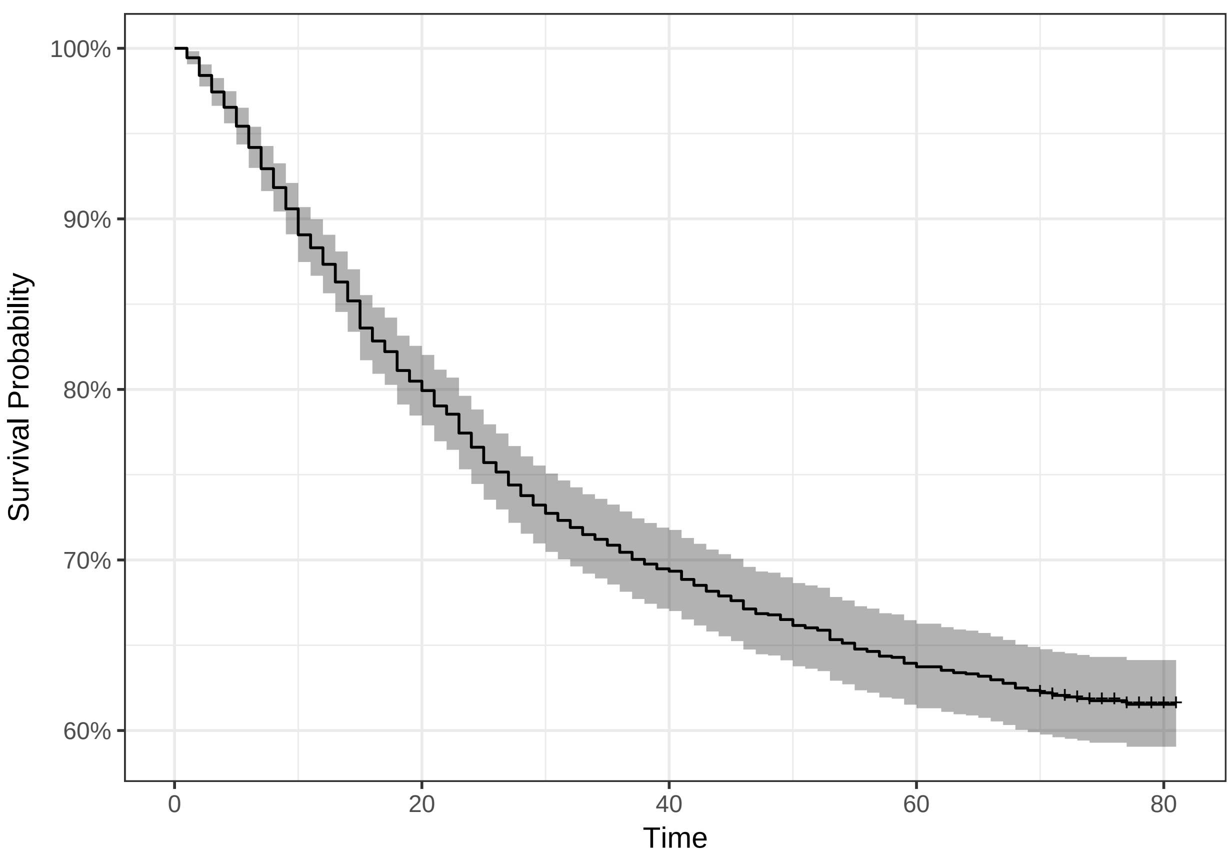
```
##     2   1437       15    0.984 0.00329         0.978         0.991
##     3   1422       14    0.974 0.00416         0.966         0.983
##     4   1408       13    0.965 0.00481         0.956         0.975
##     5   1395       16    0.954 0.00549         0.944         0.965
##     6   1379       18    0.942 0.00616         0.930         0.954
##     7   1361       18    0.929 0.00674         0.916         0.943
##     8   1343       16    0.918 0.00720         0.904         0.933
##     9   1327       18    0.906 0.00768         0.891         0.921
##    10   1309       22    0.891 0.00821         0.875         0.907
##    11   1287       11    0.883 0.00845         0.867         0.900
##    12   1276       14    0.873 0.00875         0.856         0.891
##    13   1262       15    0.863 0.00905         0.845         0.881
##    14   1247       16    0.852 0.00934         0.834         0.870
##    15   1231       23    0.836 0.00974         0.817         0.855
##    16   1208       11    0.828 0.00992         0.809         0.848
##    17   1197        9    0.822 0.01006         0.803         0.842
##    18   1188       16    0.811 0.01030         0.791         0.832
##    19   1172        9    0.805 0.01043         0.785         0.826
##    20   1163        8    0.799 0.01054         0.779         0.820
##    21   1155       13    0.790 0.01071         0.770         0.812
##    22   1142        7    0.785 0.01080         0.765         0.807
##    23   1135       16    0.774 0.01100         0.753         0.796
##    24   1119       12    0.766 0.01114         0.745         0.788
##    25   1107       13    0.757 0.01128         0.735         0.780
##    26   1094        8    0.752 0.01137         0.730         0.774
##    27   1086       11    0.744 0.01148         0.722         0.767
##    28   1075        9    0.738 0.01157         0.715         0.761
##    29   1066        8    0.732 0.01165         0.710         0.755
##    30   1058        7    0.727 0.01172         0.705         0.751
##    31   1051        6    0.723 0.01177         0.700         0.747
##    32   1045        6    0.719 0.01182         0.696         0.743
##    33   1039        6    0.715 0.01188         0.692         0.739
##    34   1033        4    0.712 0.01191         0.689         0.736
##    35   1029        5    0.709 0.01195         0.686         0.732
##    36   1024        6    0.704 0.01200         0.681         0.728
##    37   1018        6    0.700 0.01205         0.677         0.724
##    38   1012        4    0.698 0.01208         0.674         0.722
##    39   1008        4    0.695 0.01211         0.671         0.719
##    40   1004        2    0.693 0.01213         0.670         0.718
##    41   1002        7    0.689 0.01218         0.665         0.713
##    42    995        5    0.685 0.01222         0.662         0.709
##    43    990        5    0.682 0.01225         0.658         0.706
##    44    985        4    0.679 0.01228         0.655         0.703
##    45    981        4    0.676 0.01231         0.652         0.701
##    46    977        7    0.671 0.01236         0.647         0.696
##    47    970        4    0.669 0.01238         0.645         0.693
##    48    966        1    0.668 0.01239         0.644         0.693
##    49    965        4    0.665 0.01242         0.641         0.690
##    50    961        5    0.662 0.01245         0.638         0.686
##    51    956        2    0.660 0.01246         0.636         0.685
##    52    954        2    0.659 0.01247         0.635         0.684
##    53    952        8    0.653 0.01252         0.629         0.678
##    54    944        3    0.651 0.01254         0.627         0.676
##    55    941        5    0.648 0.01257         0.624         0.673
```

```
##    56    936         2    0.646 0.01258         0.622         0.671
##    57    934         4    0.644 0.01260         0.619         0.669
##    58    930         1    0.643 0.01260         0.619         0.668
##    59    929         5    0.639 0.01263         0.615         0.665
##    60    924         3    0.637 0.01265         0.613         0.663
##    62    921         3    0.635 0.01266         0.611         0.661
##    63    918         2    0.634 0.01267         0.610         0.659
##    64    916         1    0.633 0.01268         0.609         0.659
##    65    915         2    0.632 0.01269         0.607         0.657
##    66    913         3    0.630 0.01270         0.605         0.655
##    67    910         3    0.628 0.01272         0.603         0.653
##    68    907         4    0.625 0.01274         0.600         0.650
##    69    903         2    0.624 0.01275         0.599         0.649
##    70    901         2    0.622 0.01275         0.598         0.648
##    71    796         2    0.621 0.01277         0.596         0.646
##    72    706         1    0.620 0.01278         0.595         0.645
##    73    621         1    0.619 0.01280         0.594         0.644
##    74    513         1    0.617 0.01283         0.593         0.643
##    77    292         1    0.615 0.01296         0.590         0.641
```

The survival table shows time from 1 to 77. We see that the subjects starting are 1445 that at time 1, 8 had the event happening (reoffended). The survival hate was 0.994. The time 2 shows that 1437 (that is 1445 - 8 - the censored ones from time 1) is the number of criminals. From these 15 had the event (reoffended). The hazard rate for time 2

$$\frac{15}{1437} = 0.0104384$$

The survival rate goes down from 0.994 to 0.615



Cox Proportional Harzard Model

A Cox regression model has been proposed to compare the time to reoffending, and particular interest is in the effect of the education duration and its effect on survival. Using the coxph() function, develop an appropriate model for predicting the time to reoffending.

```
## Call:
## coxph(formula = Surv(time, event) ~ X, data = offend)
##
##   n= 1445, number of events= 552
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## Xdrugs    0.22951   1.25798  0.09796  2.343  0.01913 *
## Xblack    0.42125   1.52386  0.08727  4.827 1.39e-06 ***
## Xmarried -0.32605   0.72177  0.10503 -3.104  0.00191 **
## Xeduc    -0.01975   0.98045  0.01707 -1.157  0.24736
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##          exp(coef) exp(-coef) lower .95 upper .95
## Xdrugs      1.2580     0.7949    1.0382    1.5242
## Xblack      1.5239     0.6562    1.2843    1.8081
## Xmarried    0.7218     1.3855    0.5875    0.8867
## Xeduc       0.9804     1.0199    0.9482    1.0138
##
## Concordance= 0.578  (se = 0.013 )
## Rsquare= 0.026   (max possible= 0.995 )
## Likelihood ratio test= 38.27  on 4 df,   p=9.858e-08
## Wald test            = 37.29  on 4 df,   p=1.569e-07
## Score (logrank) test = 37.64  on 4 df,   p=1.328e-07
```

Interpreting the coefficients:

- criminals that use drugs will reoffend faster
- back criminals will reoffend faster
- criminals that are married will take longer to reoffend
- the more years in school the righer will talke for reoffences

Hazard Interpretation:

- criminals that use drugs will have 25.8% increase in the hazard rates
- black criminals will have 52.4% increase in the hazard rates
- criminals that are married will have 27.8% lower hazard rates
- criminals with more years in school will have 1.9% lower hazard rates

Conclusion

The best model will not need the education covariate included as the analysis shows that it's insignificant its addition. We then have that we do not have enought evidence to suggest that years of education decreases the hazard rates.