

# STAT40810: Stochastic Models : Assignment 3

Ian Towey

20 November 2016

## Question 1

The model chosen that was:

$$\text{Survival} \sim \log\text{BUN} + \text{HGB} * \text{platelet}$$

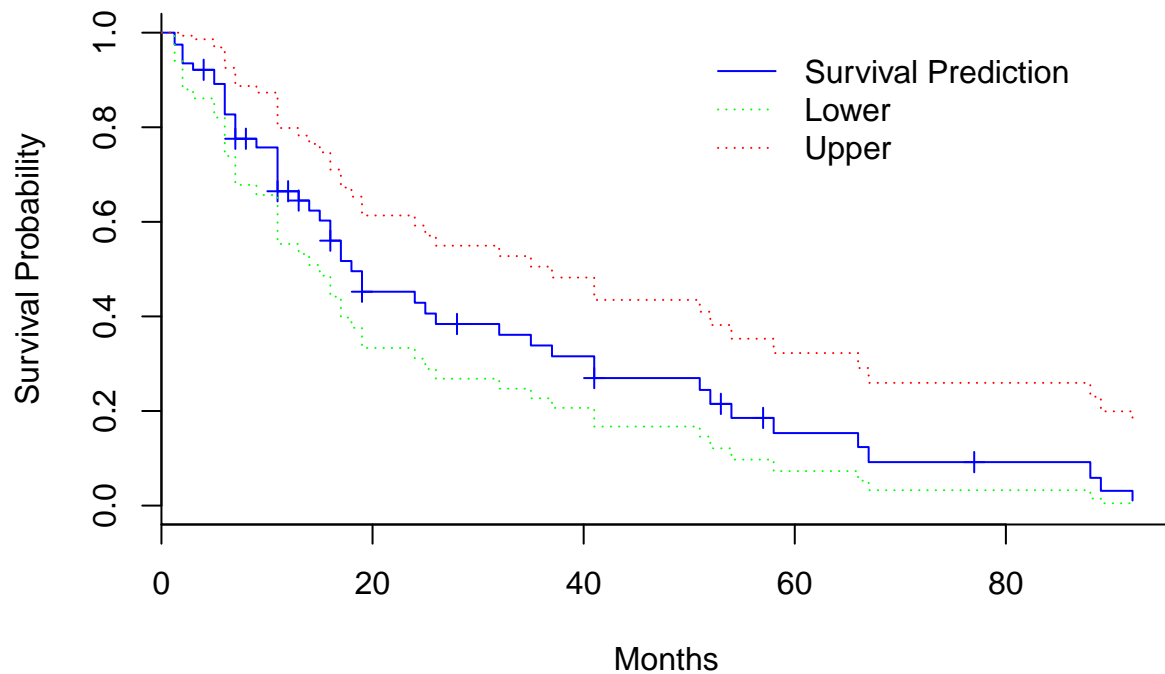
.

The model was discovered by fitting each covariate separately reviewing significance of p-value at  $\alpha = 0.05$ . Then adding interaction pairs and again reviewing significance of p-value at  $\alpha = 0.05$ .

```
## Call:
## coxph(formula = Surv(time, vstatus) ~ logBUN + HGB:platelet,
##       data = myeloma)
##
##      n= 65, number of events= 48
##
##               coef exp(coef) se(coef)      z Pr(>|z|)
## logBUN          1.58934   4.90052  0.61291  2.593  0.00951 **
## HGB:platelet -0.06885   0.93346  0.03242 -2.124  0.03368 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## logBUN            4.9005     0.2041     1.474    16.2913
## HGB:platelet      0.9335     1.0713     0.876     0.9947
##
## Concordance= 0.669  (se = 0.051 )
## Rsquare= 0.171  (max possible= 0.991 )
## Likelihood ratio test= 12.16  on 2 df,   p=0.00229
## Wald test            = 13.49  on 2 df,   p=0.001174
## Score (logrank) test = 13.99  on 2 df,   p=0.0009179
```

The model summary is displayed above. Things of note in this the output include:

- The number of patients at the start of the study was 65.
- At the end of the study there was 48 events, therefore 17 right-censored patients.
- Covariate **logBUN** (log BUN at diagnosis) and interaction covariates **HGB** (hemoglobin at diagnosis) \* **PLATELET** (platelets at diagnosis: 0=abnormal, 1=normal) are significant at  $\alpha = 0.05$  level.
- The level of BUN at diagnosis time increases the hazard by 1.5 per month
- The level of interaction covariate slightly decreases the hazard
- The “Concordance” value, is a measure of the predictive discrimination of a covariate.
- The “Rsquared” value, reflects the improvement in the fit of the model with the covariate compared to the null model.
- The hypothesis tests Likelihood ratio test, Wald test and Score test all test the hypothesis that  $H_0 : \beta = 0$  at the  $\alpha = 0.05$  level. The results of these three tests shows the test statistic for all three tests are much greater than 0 so we reject the  $H_0$



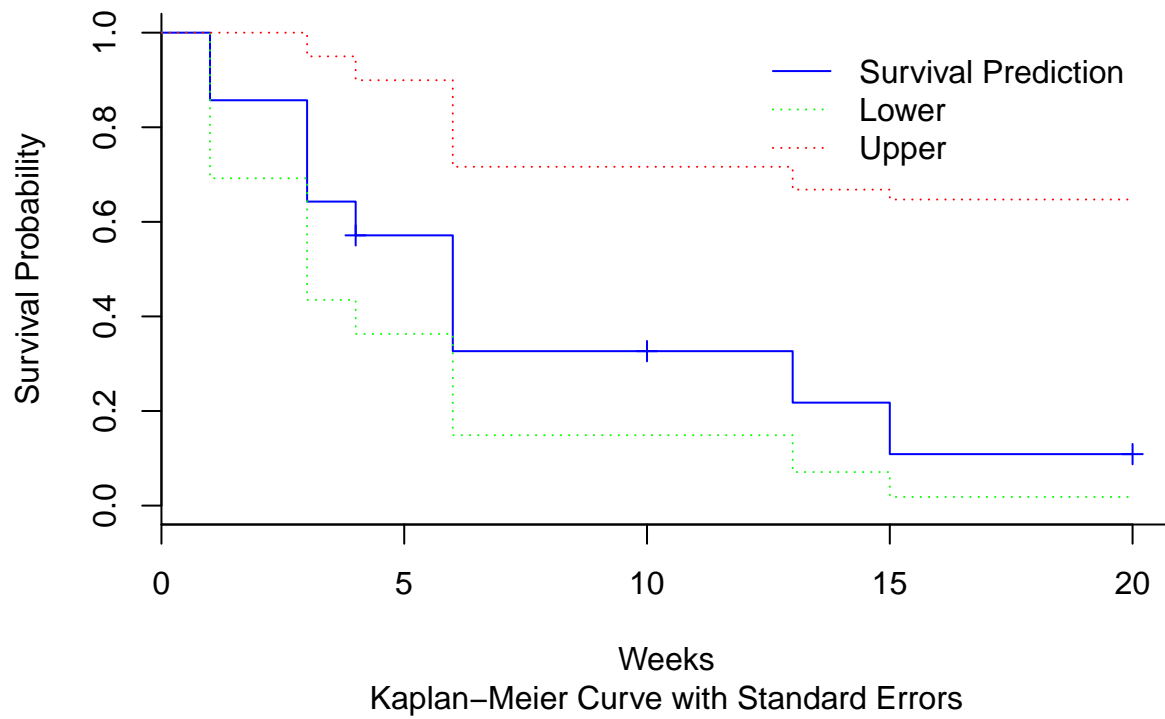
Analysis of the survival function plot above, we see that the media survival time (in months) for a patient is approximately 20 months. The probability of survival for a patient after 80 months is about 10%

## Question 2

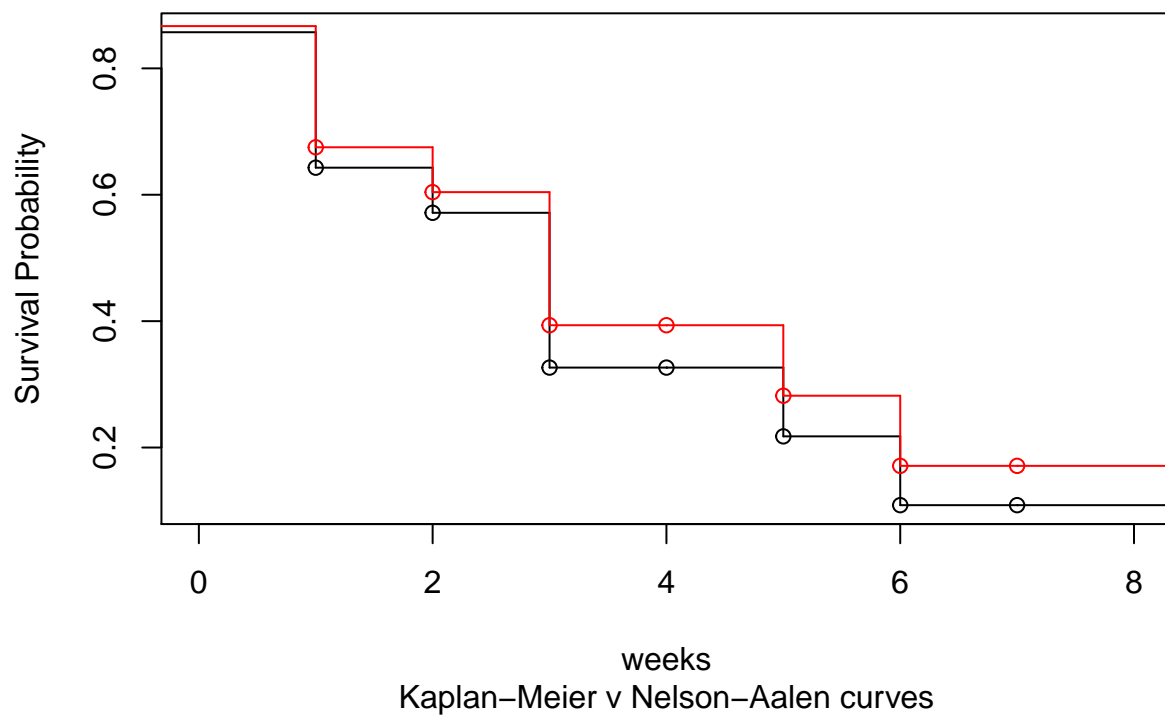
```
## [1] 4 13 3 6 4+ 3 1 1 10+ 6 6 20+ 3 15
```

(a) Kaplan-Meier estimate of the survival function and curve plot of the survival function.

```
## Call: survfit(formula = survdat ~ 1, se = TRUE)
##
##   time  n.risk  n.event  survival std.err lower 95% CI upper 95% CI
##    1      14       2    0.857  0.0935    0.6921    1.000
##    3      12       3    0.643  0.1281    0.4351    0.950
##    4       9       1    0.571  0.1323    0.3630    0.899
##    6       7       3    0.327  0.1309    0.1488    0.716
##   13       3       1    0.218  0.1246    0.0709    0.668
##   15       2       1    0.109  0.0990    0.0183    0.647
```



(b) Comparison of the Kaplan-Meier v Nelson-Aalen estimators



The Kaplan-Meier estimator is unbiased for the survival function. The Nelson-Aalen estimator is unbiased for cumulative hazard. The Nelson-Aalen estimator is always larger than negative log of the Kaplan-Meier estimator.