Time-to-Event Analysis: Practical 2 Solutions

Instructor: Graeme L. Hickey

27th October 2014

1 Problem 1

We begin by inspecting the data.

```
library("survival")
```

```
## Loading required package: splines
```

```
dim(veteran)
```

```
## [1] 137   8
```

```
head(veteran)
```

```
##   trt celltype time status karno diagtime age prior
## 1   1 squamous   72      1    60        7  69     0
## 2   1 squamous  411      1    70        5  64    10
## 3   1 squamous  228      1    60        3  38     0
## 4   1 squamous  126      1    60        9  63    10
## 5   1 squamous  118      1    70       11  65    10
## 6   1 squamous   10      1    20        5  49     0
```

```
summary(veteran)
```

```
##       trt            celltype       time           status      
##  Min.   :1.000   squamous :35   Min.   :  1.0   Min.   :0.0000  
##  1st Qu.:1.000   smallcell:48   1st Qu.: 25.0   1st Qu.:1.0000  
##  Median :1.000   adeno    :27   Median : 80.0   Median :1.0000  
##  Mean   :1.496   large    :27   Mean   :121.6   Mean   :0.9343  
##  3rd Qu.:2.000                  3rd Qu.:144.0   3rd Qu.:1.0000  
##  Max.   :2.000                  Max.   :999.0   Max.   :1.0000  
##      karno          diagtime          age            prior       
##  Min.   :10.00   Min.   : 1.000   Min.   :34.00   Min.   : 0.00  
##  1st Qu.:40.00   1st Qu.: 3.000   1st Qu.:51.00   1st Qu.: 0.00  
##  Median :60.00   Median : 5.000   Median :62.00   Median : 0.00  
##  Mean   :58.57   Mean   : 8.774   Mean   :58.31   Mean   : 2.92  
##  3rd Qu.:75.00   3rd Qu.:11.000   3rd Qu.:66.00   3rd Qu.:10.00  
##  Max.   :99.00   Max.   :87.000   Max.   :81.00   Max.   :10.00
```

We create a Cox proportional hazards regression model including just one variable, `trt`, and summarize it.

```
fit0 <- coxph(Surv(time, status) ~ trt, data = veteran)
summary(fit0)
```

```
## Call:
## coxph(formula = Surv(time, status) ~ trt, data = veteran)
##
##   n= 137, number of events= 128
##
##        coef exp(coef) se(coef)     z Pr(>|z|)
## trt 0.01774   1.01790  0.18066 0.098    0.922
##
##     exp(coef) exp(-coef) lower .95 upper .95
## trt     1.018     0.9824    0.7144      1.45
##
## Concordance= 0.525  (se = 0.026 )
## Rsquare= 0   (max possible= 0.999 )
## Likelihood ratio test= 0.01  on 1 df,   p=0.9218
## Wald test            = 0.01  on 1 df,   p=0.9218
## Score (logrank) test = 0.01  on 1 df,   p=0.9218
```

The hazard ratio is 1.02 (95% CI: 0.71-1.45; $P = 0.92$), which does not reject the null hypothesis of $\beta_{\texttt{trt}} = 0$ (equivalently $\text{HR}_{\texttt{trt}} = 1$), therefore we conclude that the new treatment has no benefit on survival.

We can add in other variables one-by-one. You might end up with a different final model from me if you make modelling decisions along the way. I opted to fit the model below.[1]

```
fit1 <- coxph(Surv(time, status) ~ trt + celltype + prior + karno,
              data = veteran)
summary(fit1)
```

```
## Call:
## coxph(formula = Surv(time, status) ~ trt + celltype + prior +
##     karno, data = veteran)
##
##   n= 137, number of events= 128
##
##                       coef exp(coef)  se(coef)      z Pr(>|z|)
```

[1] Regression model building and selection is a broad topic that crosscuts through most of statistical modelling, therefore I won't focus on the specifics of why I chose this model.

```
## trt                 0.255036  1.290508  0.201489  1.266    0.206
## celltypesmallcell   0.833094  2.300425  0.269632  3.090    0.002 **
## celltypeadeno       1.174203  3.235562  0.299650  3.919 8.91e-05 ***
## celltypelarge       0.393718  1.482483  0.282233  1.395    0.163
## prior               0.008354  1.008389  0.020675  0.404    0.686
## karno              -0.031456  0.969033  0.005201 -6.048 1.47e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                   exp(coef) exp(-coef) lower .95 upper .95
## trt                   1.291     0.7749    0.8695     1.915
## celltypesmallcell     2.300     0.4347    1.3561     3.902
## celltypeadeno         3.236     0.3091    1.7984     5.821
## celltypelarge         1.482     0.6745    0.8526     2.578
## prior                 1.008     0.9917    0.9683     1.050
## karno                 0.969     1.0320    0.9592     0.979
##
## Concordance= 0.737  (se = 0.03 )
## Rsquare= 0.36   (max possible= 0.999 )
## Likelihood ratio test= 61.23  on 6 df,   p=2.53e-11
## Wald test            = 63.37  on 6 df,   p=9.284e-12
## Score (logrank) test = 66.56  on 6 df,   p=2.067e-12
```

The coefficients all seem sensible: all are positive ($\Rightarrow$ HR > 1) with the exception of the Karnofsky performance score, but in this case the greater the score the better, thus it is intuitively correct. The coefficient for treatment effect has increased from 0.018 to 0.255, equivalent to the hazard ratio increasing from to 1.02 to 1.29. However, this remains non-significant and therefore we would still not conclude that there is a differential effect between treatments.

2 Problem 2

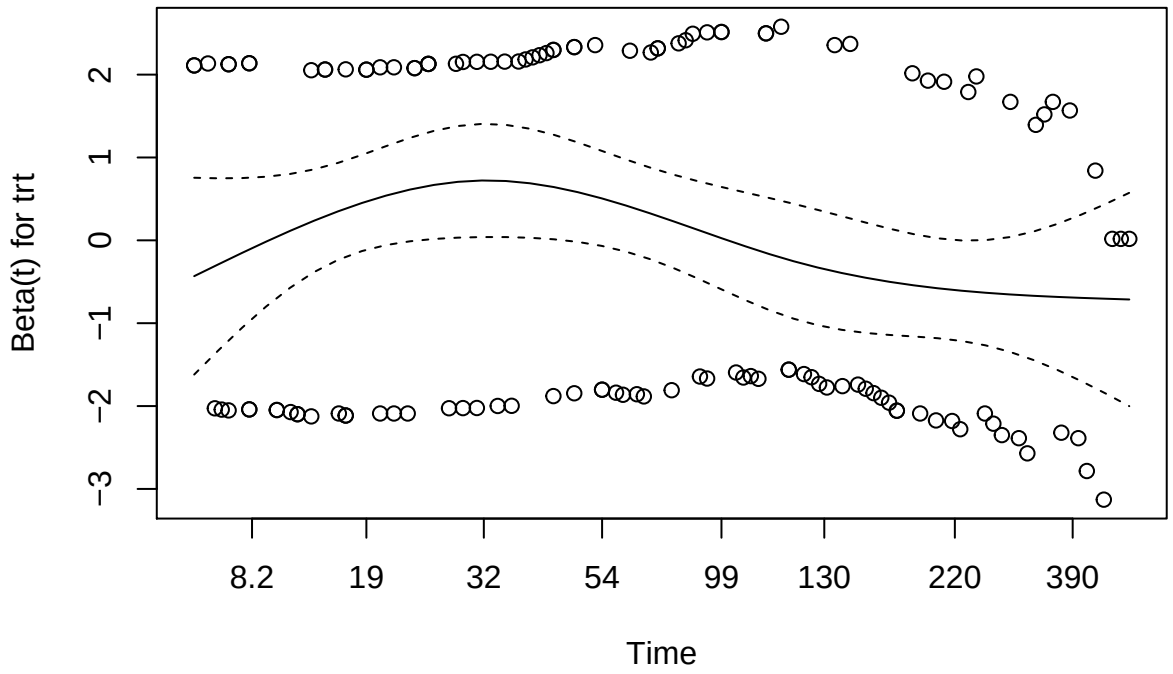
To assess the PH assumption for `trt` we can apply a number of options. The most straightforward from a syntax viewpoint is to use the `cox.zph()` function.

For the model with treatment assignnment only, we use the following code.

```
vet.ph0 <- cox.zph(fit0)
print(vet.ph0)
```

```
##       rho chisq      p
## trt -0.16   3.3 0.0691
```

```
plot(vet.ph0)
```



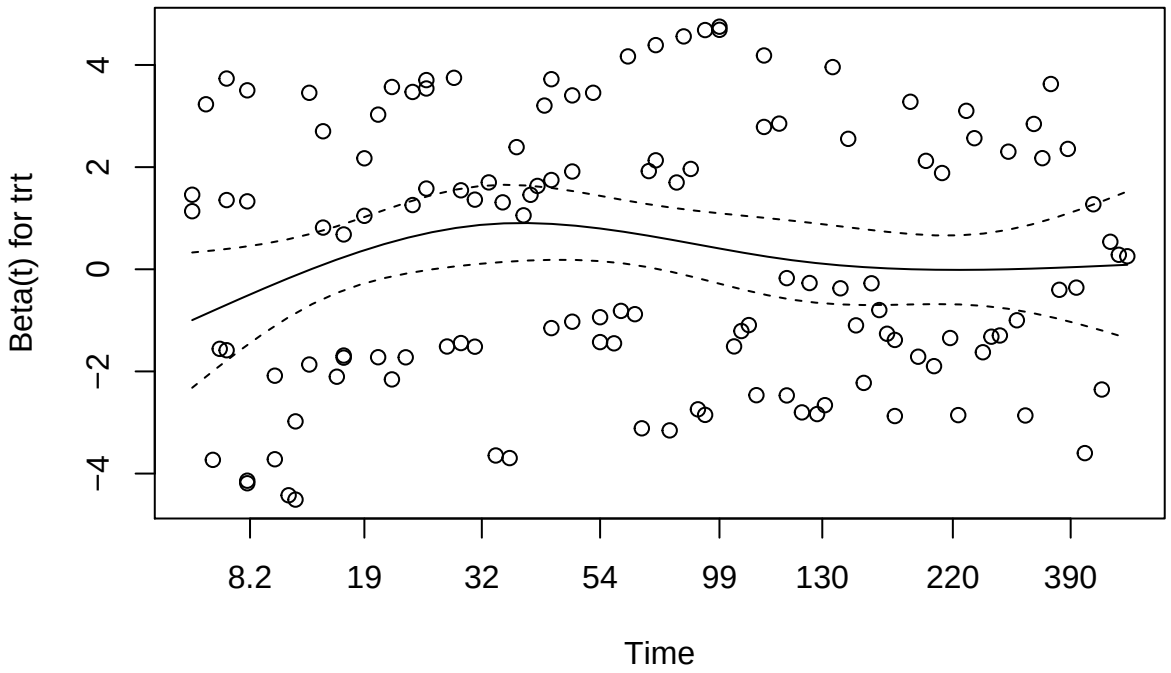
This indicates that there is slight evidence towards the PH assumption not being supported by the data, although it is not statistically significant.

For the model with treatment assignnment and other relevant variables, we use the following code.

```
vet.ph1 <- cox.zph(fit1)
print(vet.ph1)
```

```
##                       rho    chisq       p
## trt                0.00718  0.00809 0.92831
## celltypesmallcell  0.04839  0.34497 0.55698
## celltypeadeno      0.15517  3.50367 0.06123
## celltypelarge      0.17173  4.03494 0.04457
## prior             -0.14746  2.95550 0.08559
## karno              0.26377  8.64441 0.00328
## GLOBAL                 NA 20.55242 0.00221
```

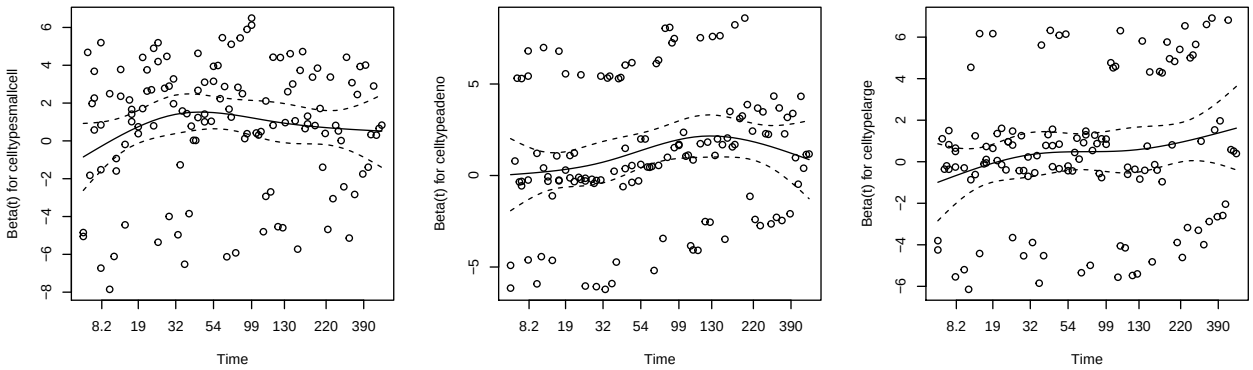
```
plot(vet.ph1, var = 1)
```



Now we find that the Schoenfeld residuals are consistent with the assumption of proportional hazards.

As part of the output from `vet.ph1` we also find that the Grambsch and Therneau global test gives evidence that the proportional hazards assumption is violated, and this might be partially driven by `celltype`. Inspection of the residual plots for `celltype` indicates that adeno and large cell types are marginally significant (one just above, and one just below $P = 0.05$).

```
par(mfrow = c(1, 3)) # We will print 3 for each dummy variable
plot(vet.ph1, var = 2:4)
```

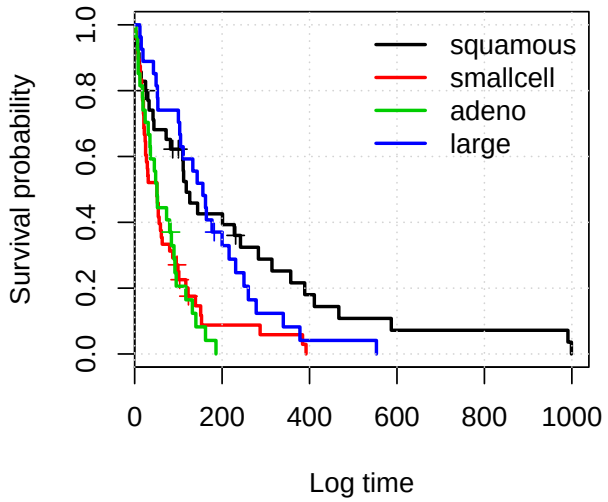


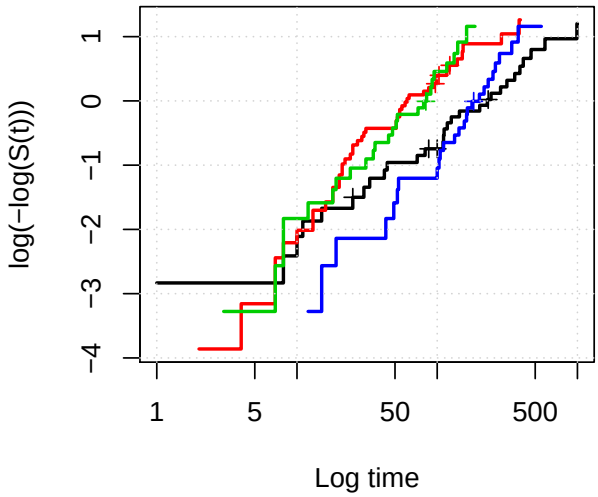
You might be wondering why there are 3 smoothed Schoenfeld residual plots. Recall that `celltype` can take any of four values: squamous, small cell, adeno, and large. R absorbs `squamous` into the intercept (or rather baseline hazard function) — called the *reference level* — and all hazard ratios for the other 3 cell types are ratios with respect to this reference. R does this by creating 3 binary 'dummy' variables: $x_{\texttt{smallcell}}$, $x_{\texttt{adeno}}$ and $x_{\texttt{large}}$, with each x_i taking value 1 if the cell type is i, and 0 otherwise.

In actual fact the Karnofsky performance score strongly rejects the PH hypothesis; however we might be able to consider including transformations of this variable to alleviate this.

We could of also looked at plots of (1) the Kaplan-Meier curves; (2) $\log\left[\hat{H}(t)\right]$ against log time, $\log(t)$, equivalent to plots of $\log\left[-\log(\hat{S}(t))\right]$ against log time, $\log(t)$.

```
fit.km <- survfit(Surv(time, status) ~ celltype, data = veteran)

par(mfrow = c(1, 2)) # Two plots side-by-side

# Kaplan-Meier curves
plot(fit.km,
     col = 1:4,
     lwd = 2,
     xlab = "Log time",
     ylab = "Survival probability")

grid()

# Add a legend
legend("topright",
       levels(veteran$celltype),
       col = 1:4,
       lwd = 2,
       lty = rep(1, 4),
       bty = "n")

# Complementary log-log plot
plot(fit.km, fun = "cloglog",
     col = 1:4,
     lwd = 2,
     xlab = "Log time",
     ylab = expression(paste("log(-log(", hat(S), "(t)))")))

grid()
```





What we find is that some of the Kaplan-Meier curves cross. Also, there is some suggestion on non-proportionality in the plots of log cumulative hazard against log time.

We can refit the model using `celltype` as a stratification variable.

```
fit2 <- coxph(Surv(time, status) ~ trt + strata(celltype) + prior +
                karno,
              data = veteran)
summary(fit2)
```

```
## Call:
## coxph(formula = Surv(time, status) ~ trt + strata(celltype) +
##     prior + karno, data = veteran)
##
##   n= 137, number of events= 128
##
##             coef exp(coef)  se(coef)      z Pr(>|z|)
## trt    0.220628  1.246860  0.201543  1.095    0.274
## prior  0.014911  1.015023  0.021111  0.706    0.480
## karno -0.036181  0.964465  0.005585 -6.478 9.29e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##       exp(coef) exp(-coef) lower .95 upper .95
## trt      1.2469     0.8020    0.8400    1.8509
## prior    1.0150     0.9852    0.9739    1.0579
## karno    0.9645     1.0368    0.9540    0.9751
##
## Concordance= 0.697  (se = 0.06 )
## Rsquare= 0.268   (max possible= 0.993 )
## Likelihood ratio test= 42.8  on 3 df,   p=2.711e-09
## Wald test            = 42.88  on 3 df,   p=2.609e-09
```

```
## Score (logrank) test = 45.95  on 3 df,   p=5.826e-10
```

The coefficient for treatment is almost unchanged at 0.221, giving a hazards ratio of 1.25, which remains non-significant ($P = 0.27$). The obvious drawback is that we can no longer quantify the effect of cell type on the hazard of death.

3 Problem 3

To fit a Weibull regression model to the `veteran` data, I continue to include the same variables as before in my model.

```
fit.w <- survreg(Surv(time, status) ~ trt + celltype + prior + karno,
                 dist = "weibull",
                 data = veteran)
summary(fit.w)
```

```
##
## Call:
## survreg(formula = Surv(time, status) ~ trt + celltype + prior +
##     karno, data = veteran, dist = "weibull")
##                       Value Std. Error      z        p
## (Intercept)         3.84814    0.46315  8.309 9.69e-17
## trt                -0.20313    0.18271 -1.112 2.66e-01
## celltypesmallcell  -0.80763    0.24048 -3.358 7.84e-04
## celltypeadeno      -1.11617    0.25771 -4.331 1.48e-05
## celltypelarge      -0.38993    0.25426 -1.534 1.25e-01
## prior              -0.00587    0.01898 -0.309 7.57e-01
## karno               0.02922    0.00459  6.368 1.92e-10
## Log(scale)         -0.07236    0.06625 -1.092 2.75e-01
##
## Scale= 0.93
##
## Weibull distribution
## Loglik(model)= -715.8   Loglik(intercept only)= -748.1
##  Chisq= 64.56 on 6 degrees of freedom, p= 5.3e-12
## Number of Newton-Raphson Iterations: 5
## n= 137
```

I won't fully interpret the model at this stage. However, there are two key features we should note:

1. The coefficient from treatment is -0.203. Ignoring the standard error (and hence P-value) momentarily, we would interpret this as saying a patient with the new test

treatment takes $\exp(-0.203) = 0.82$ times longer to reach death. In other words, we would infer a beneficial effect. However, the coefficient is not significantly different from zero, so our conclusion is that there is no evidence to suggest a differential effect, as per the Cox regression model findings above.

2. The scale parameter is < 1, which corresponds with $\hat{p} > 1$, or rather an increasing hazard function. However, the P-value is 0.28, which suggests that a simpler exponential model would be reasonable. However, it should be noted that either model might not be appropriate given the lack of proportionality in the hazards.

4 Problem 4

At this stage the course attendee should have sufficient practical knowledge to answer Problem 4 independently.