Session 7 PBC lab

Levi Waldron

Example: Primary Biliary Cirrhosis (PBC)

- Mayo Clinic trial in primary biliary cirrhosis (PBC) of the liver conducted between 1974 and 1984, n=424 patients.
- randomized placebo controlled trial of the drug D-penicillamine.
  - 312 cases from RCT, plus additional 112 not from RCT.
- Primary outcome is (censored) time to death

```
library(survival)
data(pbc)
summary(pbc)
```

```
##        id              time          status            trt
##  Min.   :  1.0   Min.   :  41   Min.   :0.0000   Min.   :1.000
##  1st Qu.:105.2   1st Qu.:1093   1st Qu.:0.0000   1st Qu.:1.000
##  Median :209.5   Median :1730   Median :0.0000   Median :1.000
##  Mean   :209.5   Mean   :1918   Mean   :0.8301   Mean   :1.494
##  3rd Qu.:313.8   3rd Qu.:2614   3rd Qu.:2.0000   3rd Qu.:2.000
##  Max.   :418.0   Max.   :4795   Max.   :2.0000   Max.   :2.000
##                                                  NA's   :106
##       age        sex        ascites            hepato
##  Min.   :26.28   m: 44   Min.   :0.00000   Min.   :0.0000
##  1st Qu.:42.83   f:374   1st Qu.:0.00000   1st Qu.:0.0000
##  Median :51.00           Median :0.00000   Median :1.0000
##  Mean   :50.74           Mean   :0.07692   Mean   :0.5128
##  3rd Qu.:58.24           3rd Qu.:0.00000   3rd Qu.:1.0000
##  Max.   :78.44           Max.   :1.00000   Max.   :1.0000
##                          NA's   :106       NA's   :106
##     spiders           edema             bili             chol
##  Min.   :0.0000   Min.   :0.0000   Min.   : 0.300   Min.   : 120.0
##  1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.: 0.800   1st Qu.: 249.5
##  Median :0.0000   Median :0.0000   Median : 1.400   Median : 309.5
##  Mean   :0.2885   Mean   :0.1005   Mean   : 3.221   Mean   : 369.5
##  3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.: 3.400   3rd Qu.: 400.0
##  Max.   :1.0000   Max.   :1.0000   Max.   :28.000   Max.   :1775.0
##  NA's   :106                                        NA's   :134
##     albumin          copper          alk.phos            ast
##  Min.   :1.960   Min.   :  4.00   Min.   :  289.0   Min.   : 26.35
##  1st Qu.:3.243   1st Qu.: 41.25   1st Qu.:  871.5   1st Qu.: 80.60
##  Median :3.530   Median : 73.00   Median : 1259.0   Median :114.70
##  Mean   :3.497   Mean   : 97.65   Mean   : 1982.7   Mean   :122.56
##  3rd Qu.:3.770   3rd Qu.:123.00   3rd Qu.: 1980.0   3rd Qu.:151.90
##  Max.   :4.640   Max.   :588.00   Max.   :13862.4   Max.   :457.25
##                  NA's   :108      NA's   :106       NA's   :106
##       trig           platelet        protime          stage
##  Min.   : 33.00   Min.   : 62.0   Min.   : 9.00   Min.   :1.000
##  1st Qu.: 84.25   1st Qu.:188.5   1st Qu.:10.00   1st Qu.:2.000
##  Median :108.00   Median :251.0   Median :10.60   Median :3.000
```

```
##  Mean   :124.70   Mean   :257.0   Mean   :10.73   Mean   :3.024
##  3rd Qu.:151.00   3rd Qu.:318.0   3rd Qu.:11.10   3rd Qu.:4.000
##  Max.   :598.00   Max.   :721.0   Max.   :18.00   Max.   :4.000
##  NA's   :136      NA's   :11      NA's   :2       NA's   :6
```

Problems

1. Create a `Surv` object using variables "time" and "status", add this to the pbc dataframe
2. Plot a KM curve for all participants using `library(survminer)` function `ggsurvplot()`.
3. Stratify by treatment and add a p-value to this plot (see `?ggsurvplot`).
4. Check on the command line whether these p-values correspond to results from a log-rank test.
5. Perform a Cox proportional hazards regression, using the "trt" variable as a predictor.
6. Create a log-minus-log plot to test the proportional hazards assumption.