# p8131_hw9_xy2395

Jack Yan

4/23/2019

```r
library(tidyverse)
library(KMsurv) # data source
library(survival) # problem 3
```

Problem 3

```r
data(tongue)
tongue_df =
  tongue %>% as.tibble() %>%
  mutate(type = if_else(type == 1, "Aneuploid", "Diploid"))
```

```r
# Aneuploid K-M curve
KM1 = survfit(Surv(time,delta)~1, data = subset(tongue_df, type == "Aneuploid"), conf.type = 'log')
plot(KM1,
     conf.int = TRUE, mark.time = TRUE,
     xlab = "Weeks", ylab = "Survival Probability", main="Aneuploid K-M curve",
     cex.lab = 1.5, cex.main = 1.5)
```
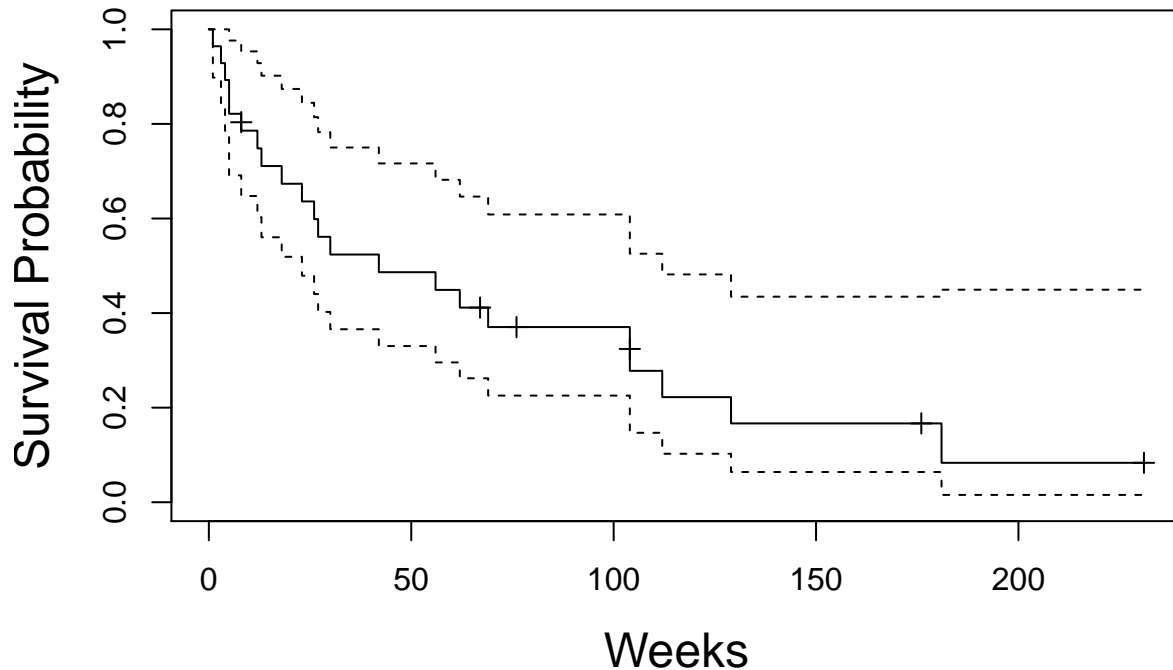
Aneuploid K–M curve
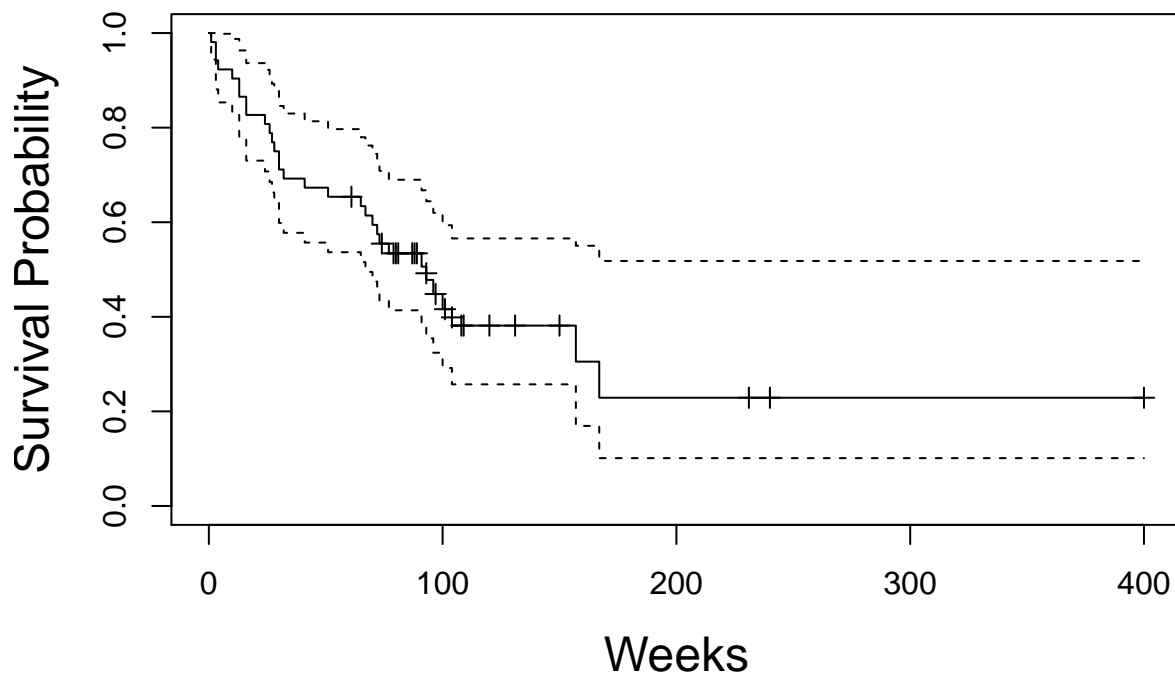

```r
# Diploid K-M curve
KM2 = survfit(Surv(time,delta)~1, data = subset(tongue_df, type == "Diploid"), conf.type = 'log')
plot(KM2,
     conf.int = TRUE, mark.time = TRUE,
```

```
      xlab = "Weeks", ylab = "Survival Probability", main="Diploid K-M curve",
      cex.lab = 1.5, cex.main = 1.5)
```

Diploid K–M curve


```
# one-year survival rate
# Aneuploid
summary(KM1, time = 365/7)
```

```
## Call: survfit(formula = Surv(time, delta) ~ 1, data = subset(tongue_df,
##     type == "Aneuploid"), conf.type = "log")
##
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
## 52.1     34      18    0.654   0.066        0.537        0.797
```

```
# Diploid
summary(KM2, time = 365/7)
```

```
## Call: survfit(formula = Surv(time, delta) ~ 1, data = subset(tongue_df,
##     type == "Diploid"), conf.type = "log")
##
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
## 52.1     13      14    0.486  0.0961         0.33        0.716
```

The estimated 1-year survival rate (95% CI) is 0.654 (0.537, 0.797) for Aneuploid and 0.486 (0.33, 0.716) for Diploid.