

HW9

Hanao Li

November 15, 2019

Problem 1

```
if(!require("pacman")) install.packages("pacman")
```

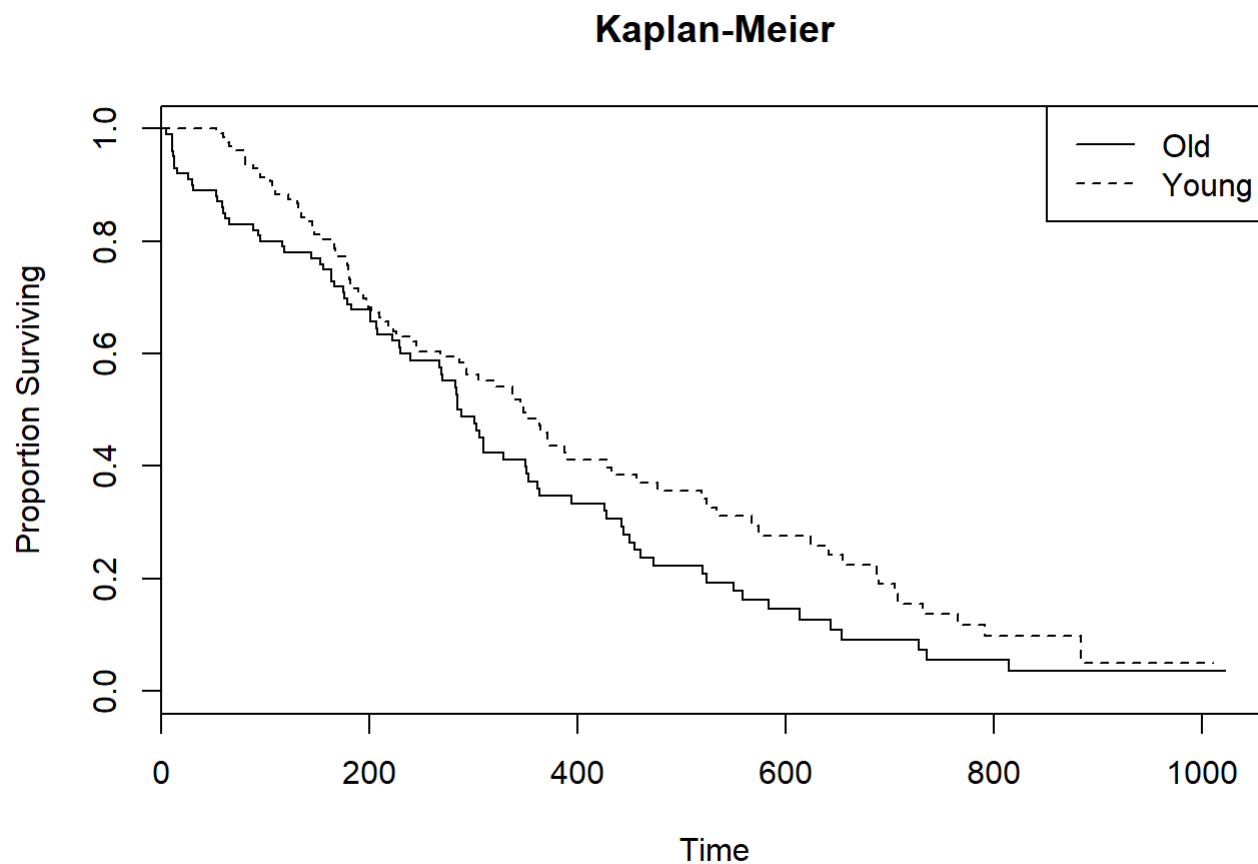
```
## Loading required package: pacman
```

```
p_load(survminer, survival)
```

```
data <- lung  
data$age_group = "OLD"  
data$age_group[which(data$age < 65)] = "YOUNG"
```

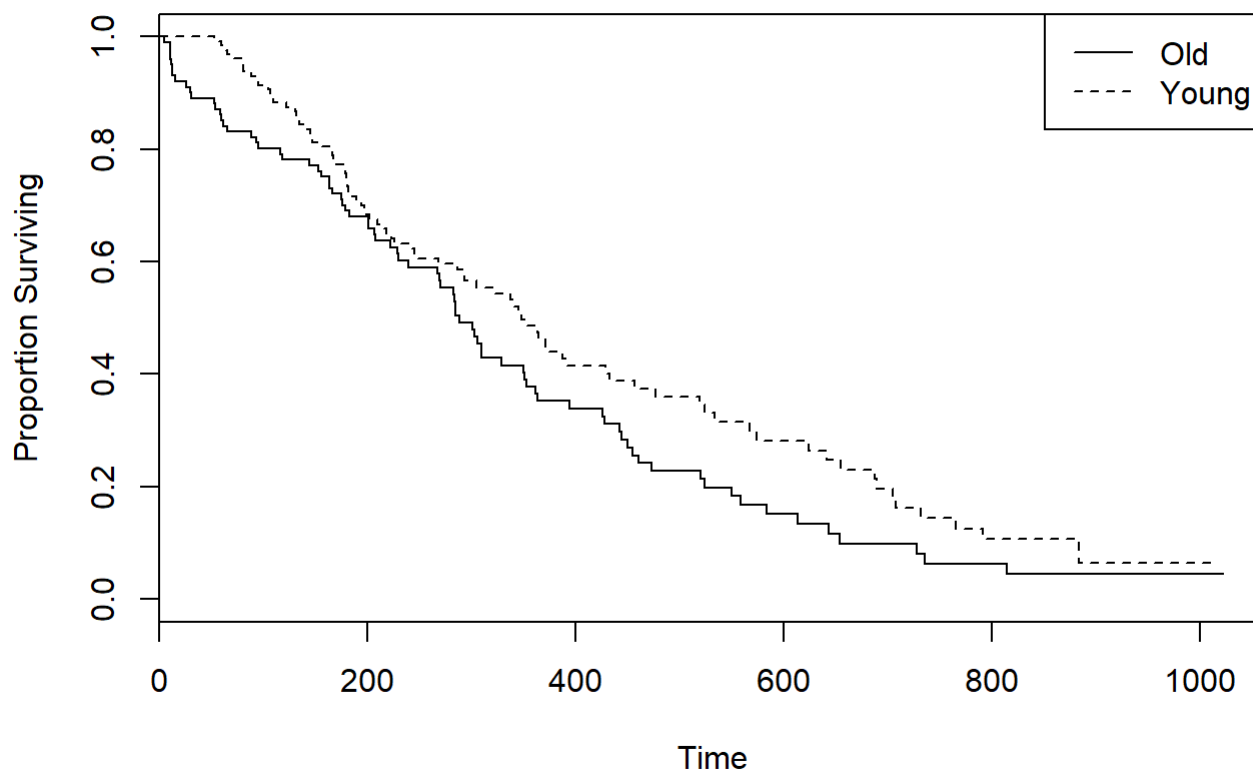
```
# Kaplan-Meier
```

```
km <- survfit(Surv(time,status)~age_group, data=data,type="kaplan-meier")  
plot(km,lty=c(1,2), xlab="Time", ylab = "Proportion Surviving",main="Kaplan-Meier")  
legend("topright", c("Old", "Young"), lty = c(1,2))
```



```
# Fleming-Harrington
fh <- survfit(Surv(time, status)~age_group, data = data,type = "fleming-harrington")
plot(fh,lty = c(1,2), xlab="Time", ylab = "Proportion Surviving", main="Fleming-Harrington")
legend("topright", c("Old", "Young"), lty = c(1,2))
```

Fleming-Harrington



Problem 2

```
# Kaplan-Meier
km
```

```
## Call: survfit(formula = Surv(time, status) ~ age_group, data = data,
##      type = "kaplan-meier")
##
##              n events median 0.95LCL 0.95UCL
## age_group=OLD   100     79   288     239     353
## age_group=YOUNG 128     86   348     291     433
```

```
# Fleming-Harrington
fh
```

```
## Call: survfit(formula = Surv(time, status) ~ age_group, data = data,
##      type = "fleming-harrington")
##
##              n events median 0.95LCL 0.95UCL
## age_group=OLD   100     79   288     239     353
## age_group=YOUNG 128     86   348     291     433
```

From the results, we could see that Kaplan-Meier and Fleming-Harrington leads to the same results. The median survival time is 288 for Old and 348 for Young.

Problem 3

```
data <- data %>% mutate(sex_group = ifelse(sex == 1, "Male", "Female"))
survdifff(Surv(time, status) ~ sex_group, data = data, rho=0)
```

```
## Call:
## survdifff(formula = Surv(time, status) ~ sex_group, data = data,
##      rho = 0)
##
##              N Observed Expected (O-E)^2/E (O-E)^2/V
## sex_group=Female  90      53     73.4      5.68     10.3
## sex_group=Male   138     112     91.6      4.55     10.3
##
## Chisq= 10.3 on 1 degrees of freedom, p= 0.001
```

From the results, we could see that p value is less than 0.05, so we will reject the null hypothesis and conclude that there is a difference in survival between the two groups.