BIOSTATYSTYKA - LABORATORIUM 1

```r
library("foreign")
library("survival")
```

Zadanie 1.

```r
seasick <- read.dta("C:\\Users\\Marta\\Desktop\\Marta\\studia\\rok4\\Biostatystyka\\1\\seasick_eng_data.dta")
head(seasick)

##   intens time vomit
## 1      1   30     1
## 2      1   50     1
## 3      1   50     0
## 4      1   51     1
## 5      1   66     0
## 6      1   82     1

# dla pierwszego eksperymentu:

sea1.KM <- survfit(Surv(time, vomit) ~ intens, data = seasick, conf.type = "none",
    subset = intens == 1)
summary(sea1.KM)

## Call: survfit(formula = Surv(time, vomit) ~ intens, data = seasick,
##     subset = intens == 1, conf.type = "none")
##
##   time n.risk n.event survival std.err
##     30     21       1    0.952  0.0465
##     50     20       1    0.905  0.0641
##     51     18       1    0.854  0.0778
##     82     16       1    0.801  0.0894
##     98     15       1    0.748  0.0981
```
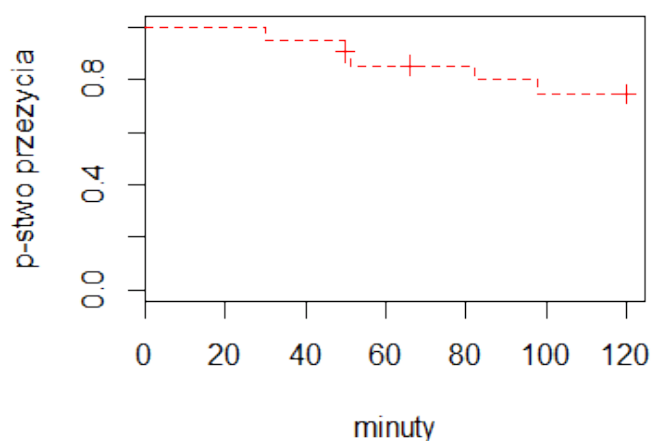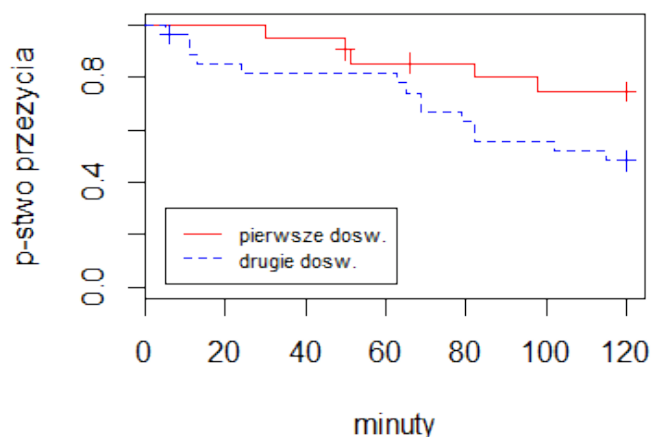
```r
plot(sea1.KM, col = "red", lty = 2, xlab = "minuty", ylab = "p-stwo przezycia")
```

```r
# dla obu eksperymentow:

sea.KM <- survfit(Surv(time, vomit) ~ intens, data = seasick, conf.type = "none")
summary(sea.KM)

plot(sea.KM, col = c("red", "blue"), lty = c(1, 2), xlab = "minuty", ylab = "p-stwo przezycia")
legend(5, 0.3, c("pierwsze dosw.", "drugie dosw."), col = c("red", "blue"),
    lty = c(1, 2), cex = 0.7)
```



```r
# porowynywanie krzywych przezycia:

sea.test <- survdiff(Surv(time, vomit) ~ intens, data = seasick, rho = 0)
sea.test

## Call:
## survdiff(formula = Surv(time, vomit) ~ intens, data = seasick,
##     rho = 0)
##
##            N Observed Expected (O-E)^2/E (O-E)^2/V
## intens=1 21        5     8.86      1.68      3.21
## intens=2 28       14    10.14      1.47      3.21
##
##  Chisq= 3.2  on 1 degrees of freedom, p= 0.0733

# p-value duze, zatem przyjmujemy hipoteze, czyli krzywe przezycia nie
# roznia sie istotnie
```

Zadanie 2.

```r
nsclc <- read.dta("C:\\Users\\Marta\\Desktop\\Marta\\studia\\rok4\\Biostatystyka\\1\\nsclc_eng.dta")
head(nsclc)

##   patient mutation survtime survind tnm expression
## 1       1        1    24.51       1   2          1
## 2       2        1    27.13       1   2          1
## 3       3        0    34.84       0   2          1
```
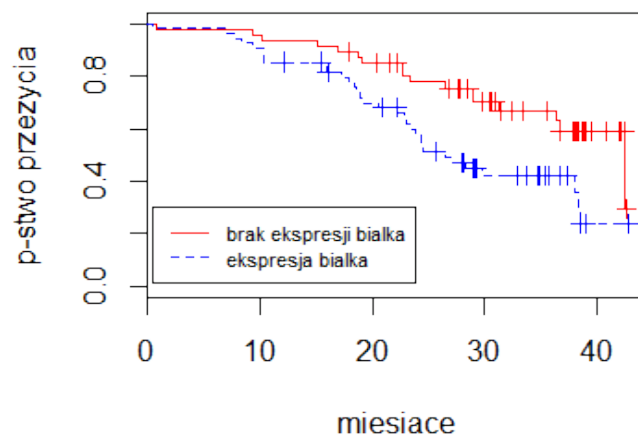
```
## 4          4       1     35.70       0    2           1
## 5          5       1     27.90       0    3           1
## 6          6       1     25.67       0    3           1
```

```r
nsclc.KM <- survfit(Surv(survtime, survind) ~ expression, data = nsclc, conf.type = "none")

plot(nsclc.KM, col = c("red", "blue"), lty = c(1, 2), xlab = "miesiace", ylab = "p-stwo przezycia")
legend(0.4, 0.3, c("brak ekspresji bialka", "ekspresja bialka"), col = c("red",
    "blue"), lty = c(1, 2), cex = 0.7)
```



```r
nsclc.test <- survdiff(Surv(survtime, survind) ~ expression, data = nsclc)
nsclc.test  # krzywe roznia sie istotnie

## Call:
## survdiff(formula = Surv(survtime, survind) ~ expression, data = nsclc)
##
##                N Observed Expected (O-E)^2/E (O-E)^2/V
## expression=0 47       17     26.2      3.25      7.11
## expression=1 55       32     22.8      3.75      7.11
##
##  Chisq= 7.1  on 1 degrees of freedom, p= 0.00769

# test warstwowy:

nsclc.test.w <- survdiff(Surv(survtime, survind) ~ expression + strata(tnm),
    data = nsclc)
nsclc.test.w

## Call:
## survdiff(formula = Surv(survtime, survind) ~ expression + strata(tnm),
##     data = nsclc)
##
##                N Observed Expected (O-E)^2/E (O-E)^2/V
## expression=0 47       17     25.5      2.82      6.05
## expression=1 55       32     23.5      3.06      6.05
##
##  Chisq= 6.1  on 1 degrees of freedom, p= 0.0139
```
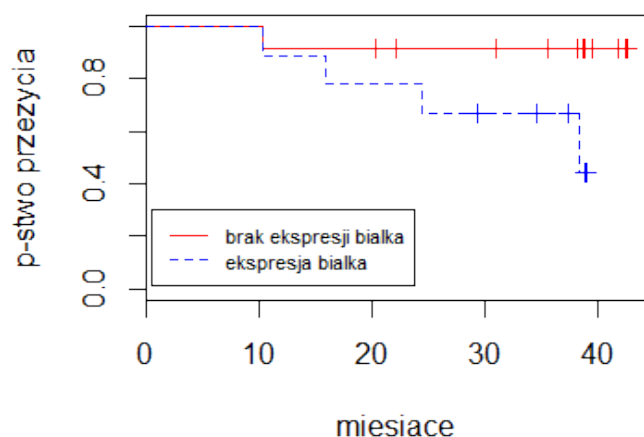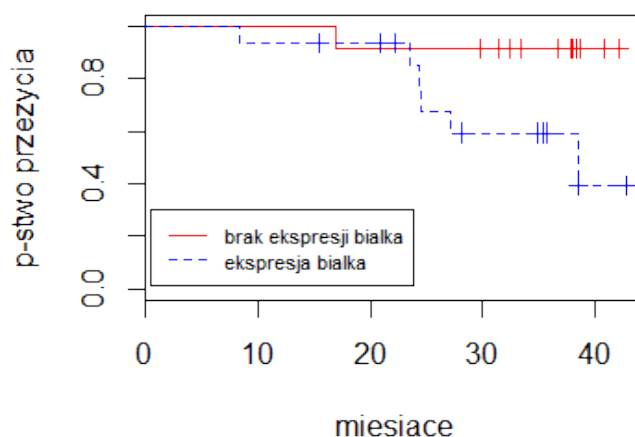
```
# wykresy dla warstw oddzielnie:

nsclc.KM.1 <- survfit(Surv(survtime, survind) ~ expression, data = nsclc, conf.type = "none",
    subset = tnm == 1)
plot(nsclc.KM.1, col = c("red", "blue"), lty = c(1, 2), xlab = "miesiace", ylab = "p-stwo przezycia",
    main = "TNM = 1")
legend(0.4, 0.3, c("brak ekspresji bialka", "ekspresja bialka"), col = c("red",
    "blue"), lty = c(1, 2), cex = 0.7)
nsclc.KM.2 <- survfit(Surv(survtime, survind) ~ expression, data = nsclc, conf.type = "none",
    subset = tnm == 2)
plot(nsclc.KM.2, col = c("red", "blue"), lty = c(1, 2), xlab = "miesiace", ylab = "p-stwo przezycia",
    main = "TNM = 2")
legend(0.4, 0.3, c("brak ekspresji bialka", "ekspresja bialka"), col = c("red",
    "blue"), lty = c(1, 2), cex = 0.7)
nsclc.KM.3 <- survfit(Surv(survtime, survind) ~ expression, data = nsclc, conf.type = "none",
    subset = tnm == 3)
plot(nsclc.KM.3, col = c("red", "blue"), lty = c(1, 2), xlab = "miesiace", ylab = "p-stwo przezycia",
    main = "TNM = 3")
legend(0.4, 0.3, c("brak ekspresji białka", "ekspresja bialka"), col = c("red",
    "blue"), lty = c(1, 2), cex = 0.7)
```