Ćwiczenia 6

2023-12-06

Zadanie 1

```
df <- data.frame(
  Time  = c(143, 165, 188, 188, 190, 192, 206, 208, 212, 216, 220, 227, 230, 235, 246, 265, 303, 216, 24
           142, 157, 163, 198, 205, 232, 232, 232, 233, 233, 233, 233, 239, 240, 261,
           280, 280, 295, 295, 323, 204, 344),
  group = rep(1:2, c(19, 22)) |> factor(),
  cens  = c(rep(1:0, c(17, 2)),
           rep(1:0, c(20, 2)))
)

df <- df |>
  mutate(Time_c = Surv(Time, cens))
```

```
model_surv <- survfit(Time_c ~ group, data = df)
summary(model_surv)
```

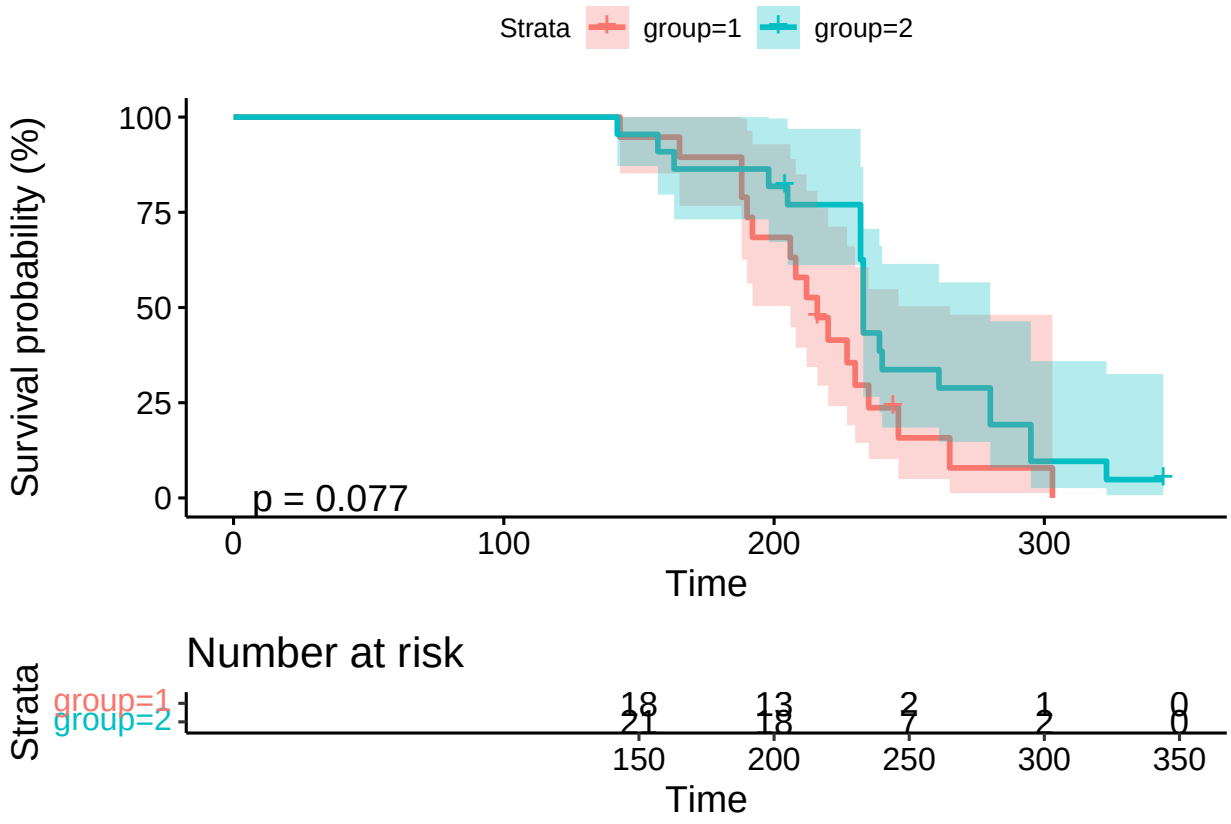
```
## Call: survfit(formula = Time_c ~ group, data = df)
##
##                 group=1
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##   143     19       1   0.9474  0.0512       0.8521        1.000
##   165     18       1   0.8947  0.0704       0.7669        1.000
##   188     17       2   0.7895  0.0935       0.6259        0.996
##   190     15       1   0.7368  0.1010       0.5632        0.964
##   192     14       1   0.6842  0.1066       0.5041        0.929
##   206     13       1   0.6316  0.1107       0.4480        0.890
##   208     12       1   0.5789  0.1133       0.3946        0.850
##   212     11       1   0.5263  0.1145       0.3435        0.806
##   216     10       1   0.4737  0.1145       0.2949        0.761
##   220      8       1   0.4145  0.1145       0.2412        0.712
##   227      7       1   0.3553  0.1124       0.1911        0.661
##   230      6       1   0.2961  0.1082       0.1447        0.606
##   235      5       1   0.2368  0.1015       0.1023        0.548
##   246      3       1   0.1579  0.0934       0.0495        0.504
##   265      2       1   0.0789  0.0728       0.0130        0.481
##   303      1       1   0.0000     NaN           NA           NA
##
##                 group=2
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##   142     22       1   0.9545  0.0444      0.87136        1.000
##   157     21       1   0.9091  0.0613      0.79656        1.000
##   163     20       1   0.8636  0.0732      0.73151        1.000
##   198     19       1   0.8182  0.0822      0.67189        0.996
##   205     17       1   0.7701  0.0904      0.61180        0.969
```

```
##    232     16       3   0.6257  0.1051       0.45020         0.870
##    233     13       4   0.4332  0.1082       0.26548         0.707
##    239      9       1   0.3850  0.1063       0.22408         0.662
##    240      8       1   0.3369  0.1034       0.18465         0.615
##    261      7       1   0.2888  0.0992       0.14731         0.566
##    280      6       2   0.1925  0.0864       0.07991         0.464
##    295      4       2   0.0963  0.0647       0.02580         0.359
##    323      2       1   0.0481  0.0469       0.00712         0.326
```

```
print(model_surv, rmean = "common")
```

```
## Call: survfit(formula = Time_c ~ group, data = df)
##
##          n events rmean* se(rmean) median 0.95LCL 0.95UCL
## group=1 19     17    219      9.07    216     206     265
## group=2 22     20    241     10.79    233     232     280
##     * restricted mean with upper limit =  344
```

```
ggsurvplot(model_surv,
           data = df,
           risk.table = TRUE,
           pval = TRUE,
           conf.int = TRUE,
           fun = 'pct')
```



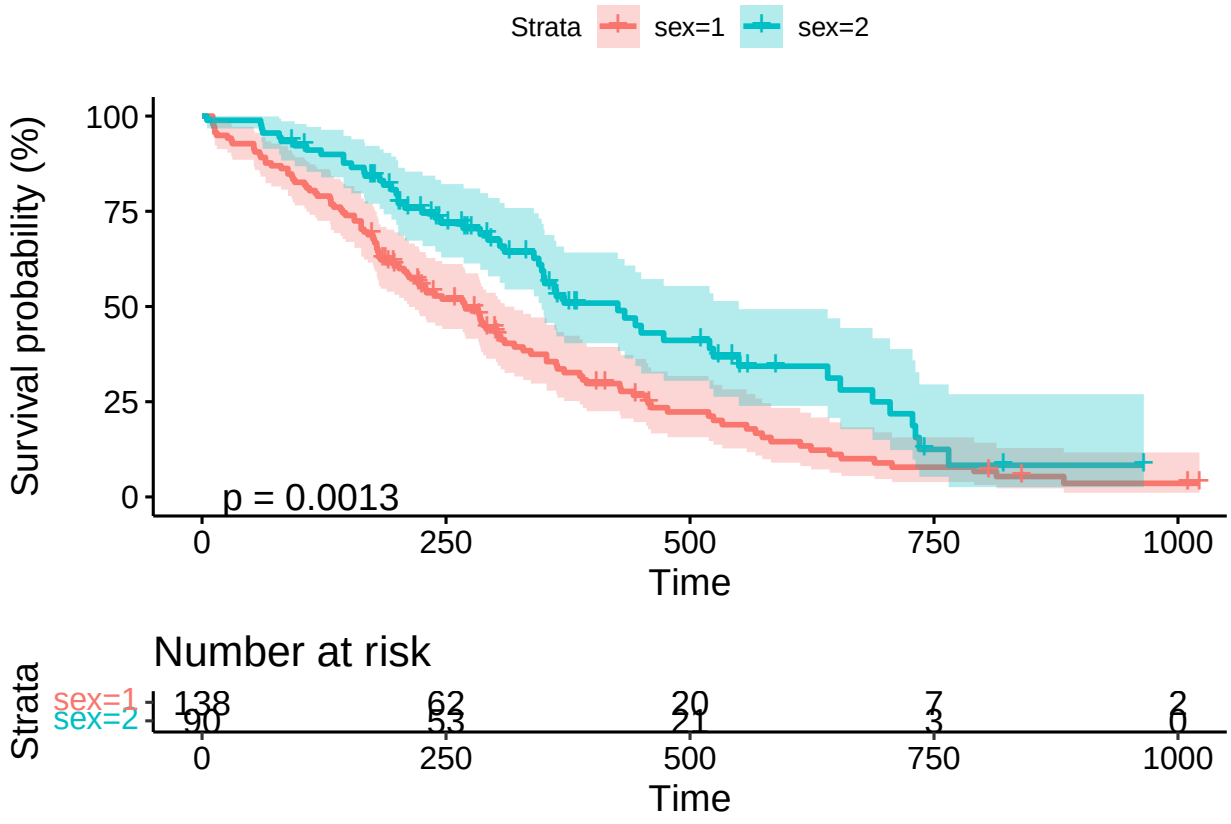
Zadanie 2

```
(df <- cancer |>
  select(time, status, sex, meal.cal)|>
```

```
  mutate(status = status - 1) |>
  mutate(time_c = Surv(time, status)) |>
  select(-status, - time) |>
  tibble())
```

```
## # A tibble: 228 x 3
##      sex meal.cal time_c
##    <dbl>    <dbl> <Surv>
##  1     1     1175   306
##  2     1     1225   455
##  3     1       NA  1010+
##  4     1     1150   210
##  5     1       NA   883
##  6     1      513  1022+
##  7     2      384   310
##  8     2      538   361
##  9     1      825   218
## 10     1      271   166
## # i 218 more rows
```

```
model_surv <- survfit(time_c ~ sex, data = df)
ggsurvplot(model_surv,
           data = df,
           risk.table = TRUE,
           pval = TRUE,
           conf.int = TRUE,
           fun = 'pct')
```



```r
print(model_surv, rmean = "common")
```

```
## Call: survfit(formula = time_c ~ sex, data = df)
##
##         n events rmean* se(rmean) median 0.95LCL 0.95UCL
## sex=1 138    112    326      22.9    270     212     310
## sex=2  90     53    461      34.7    426     348     550
##     * restricted mean with upper limit =  1022
```

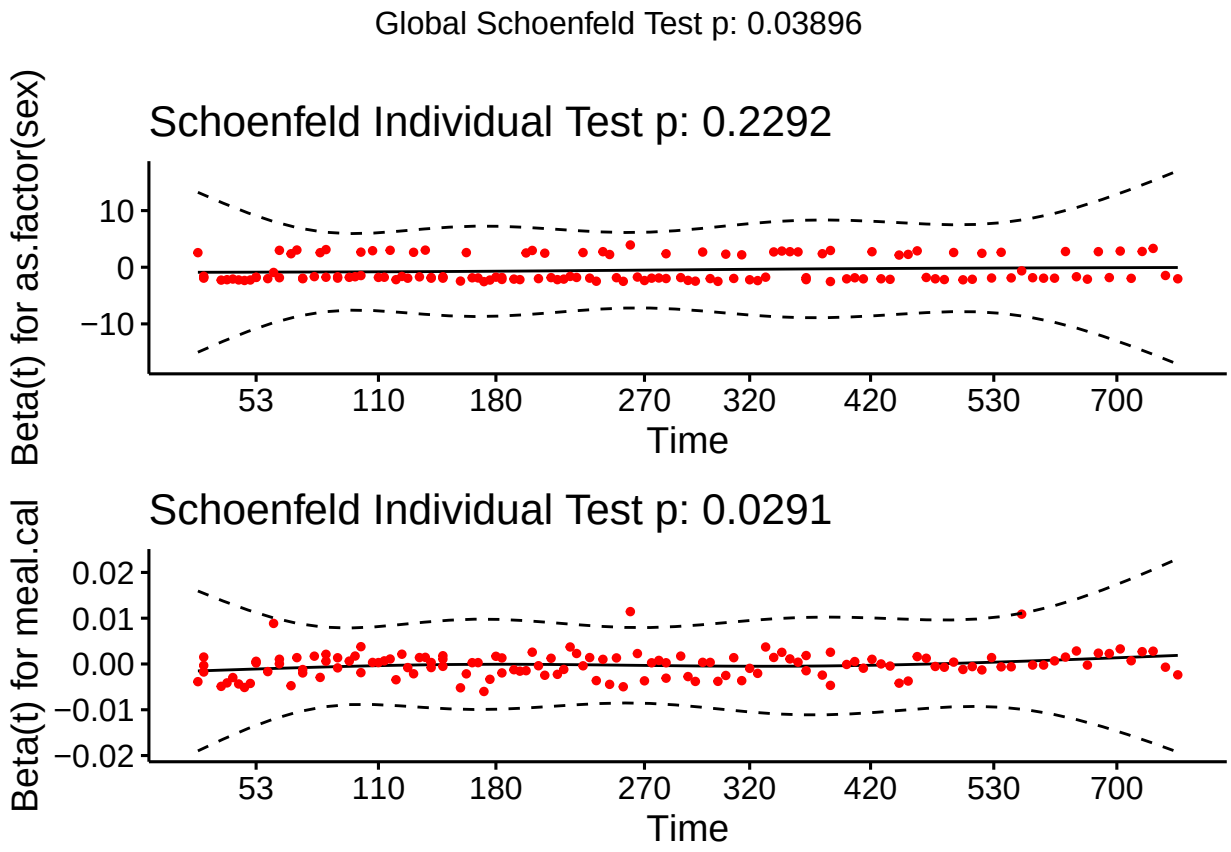
```r
model_surv <- coxph(time_c ~ as.factor(sex) + meal.cal, data = df)
summary(model_surv)
```

```
## Call:
## coxph(formula = time_c ~ as.factor(sex) + meal.cal, data = df)
##
##   n= 181, number of events= 134
##    (47 observations deleted due to missingness)
##
##                        coef  exp(coef)   se(coef)      z Pr(>|z|)
## as.factor(sex)2 -0.5276151  0.5900104  0.1893980 -2.786  0.00534 **
## meal.cal        -0.0002164  0.9997836  0.0002343 -0.924  0.35570
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                 exp(coef) exp(-coef) lower .95 upper .95
## as.factor(sex)2    0.5900      1.695    0.4070    0.8552
## meal.cal           0.9998      1.000    0.9993    1.0002
##
## Concordance= 0.598  (se = 0.028 )
## Likelihood ratio test= 8.5  on 2 df,   p=0.01
## Wald test            = 8.08  on 2 df,   p=0.02
## Score (logrank) test = 8.24  on 2 df,   p=0.02
```

```r
cox.zph(model_surv)
```

```
##                chisq df     p
## as.factor(sex)  1.45  1 0.229
## meal.cal        4.76  1 0.029
## GLOBAL          6.49  2 0.039
```

```r
ggcoxzph(cox.zph(model_surv))
```



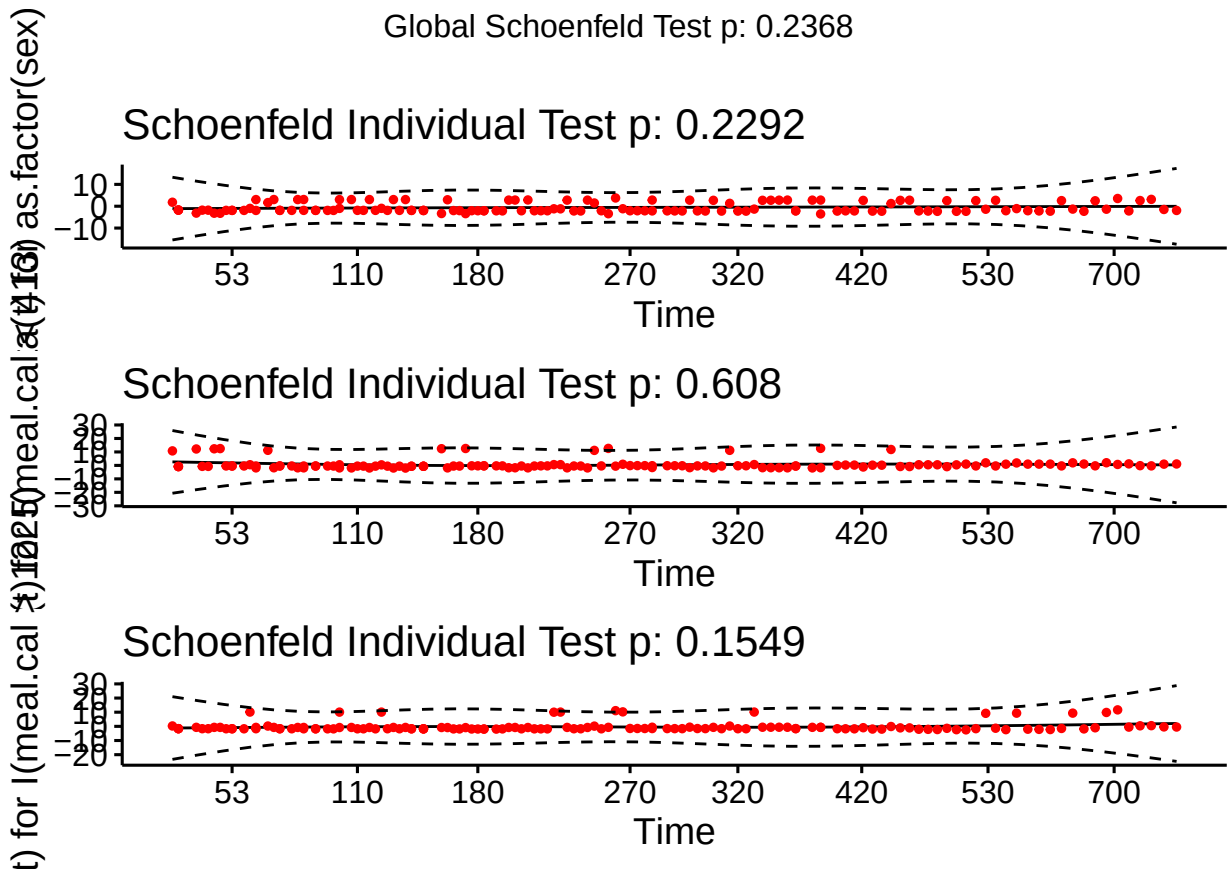

```r
model_surv <- coxph(time_c ~ as.factor(sex) + I(meal.cal < 413) + I(meal.cal > 1225), data = df)
print(summary(model_surv))
```

```
## Call:
## coxph(formula = time_c ~ as.factor(sex) + I(meal.cal < 413) +
##     I(meal.cal > 1225), data = df)
##
##   n= 181, number of events= 134
##    (47 observations deleted due to missingness)
##
##                            coef exp(coef) se(coef)      z Pr(>|z|)
## as.factor(sex)2         -0.5844    0.5575   0.1925 -3.035   0.0024 **
## I(meal.cal < 413)TRUE    0.6877    1.9892   0.3121  2.203   0.0276 *
## I(meal.cal > 1225)TRUE -0.2132    0.8080   0.2970 -0.718   0.4730
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                        exp(coef) exp(-coef) lower .95 upper .95
## as.factor(sex)2           0.5575     1.7939    0.3822     0.813
## I(meal.cal < 413)TRUE     1.9892     0.5027    1.0789     3.668
## I(meal.cal > 1225)TRUE    0.8080     1.2376    0.4514     1.446
##
## Concordance= 0.594  (se = 0.025 )
## Likelihood ratio test= 12.5  on 3 df,   p=0.006
## Wald test            = 12.4  on 3 df,   p=0.006
## Score (logrank) test = 12.63  on 3 df,   p=0.006
```

```
print(cox.zph(model_surv))
```

```
##                     chisq df    p
## as.factor(sex)      1.446  1 0.23
## I(meal.cal < 413)   0.263  1 0.61
## I(meal.cal > 1225) 2.023  1 0.15
## GLOBAL              4.238  3 0.24
```

```
ggcoxzph(cox.zph(model_surv))
```




```
model_surv1 <- coxph(time_c ~ as.factor(sex) + I(meal.cal < 413), data = df)
model_surv <- coxph(time_c ~ as.factor(sex), data = df |> na.omit())
anova(model_surv, model_surv1)
```

```
## Analysis of Deviance Table
##  Cox model: response is  time_c
##  Model 1: ~ as.factor(sex)
##  Model 2: ~ as.factor(sex) + I(meal.cal < 413)
##    loglik  Chisq Df Pr(>|Chi|)
## 1 -575.08
## 2 -572.91 4.3409  1    0.03721 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Zadanie 3

Dane:

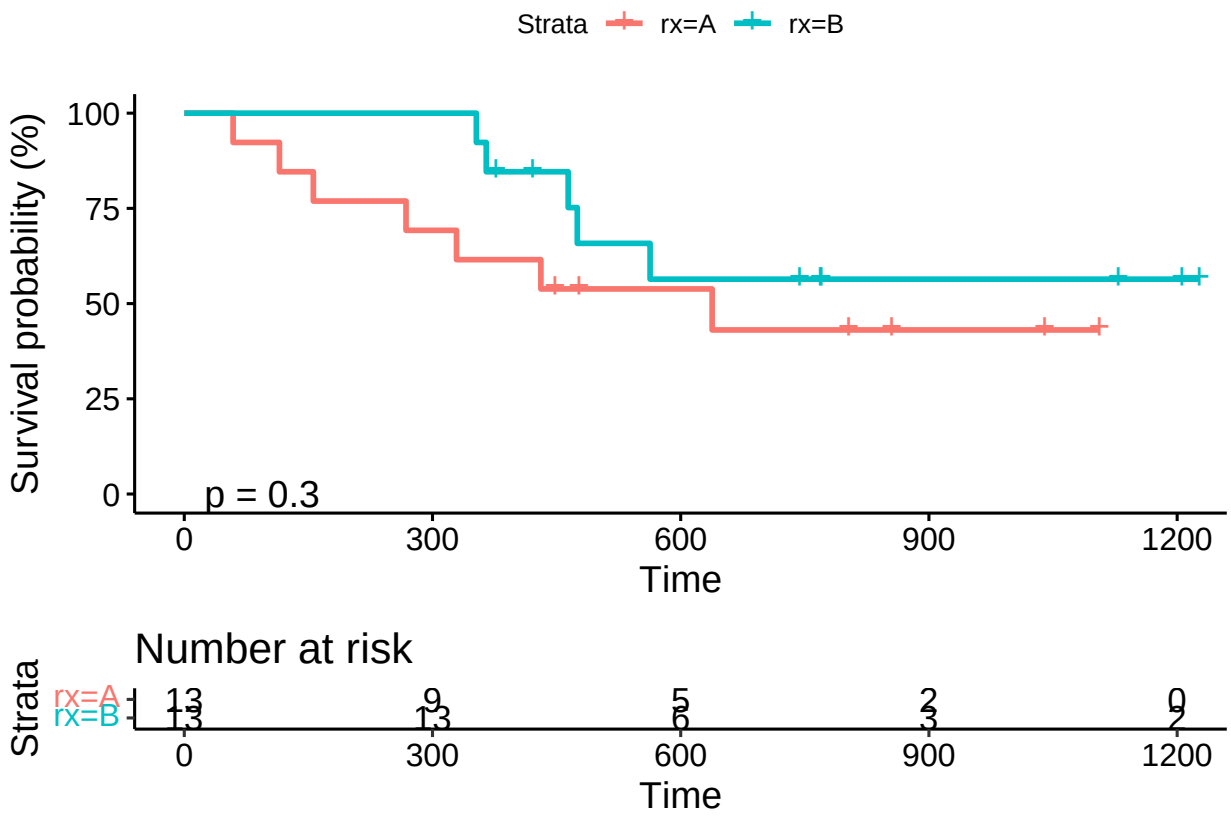
```
ovarian %>%
  mutate(rx = factor(
    rx,
    levels = c(1, 2),
    labels = c("A", "B")),
    resid.ds = factor(
      resid.ds,
      levels = c(1, 2),
      labels = c("no", "yes")),
    ecog.ps = factor(
      ecog.ps,
      levels = c(1, 2),
      labels = c("good", "bad")),
    age = factor(ifelse(age >= 50, "old", "young")),
    futime = Surv(futime, fustat)) -> df
```

```
model_surv <- survfit(futime ~ rx, data = df)
print(model_surv, rmean = "common")
```

a, b, c)

```
## Call: survfit(formula = futime ~ rx, data = df)
##
##       n events rmean* se(rmean) median 0.95LCL 0.95UCL
## rx=A 13      7    702       137    638     268      NA
## rx=B 13      5    889       115     NA     475      NA
##     * restricted mean with upper limit =  1227
```

```
ggsurvplot(model_surv,
           data = df,
           risk.table = TRUE,
           pval = TRUE,
           fun = 'pct')
```



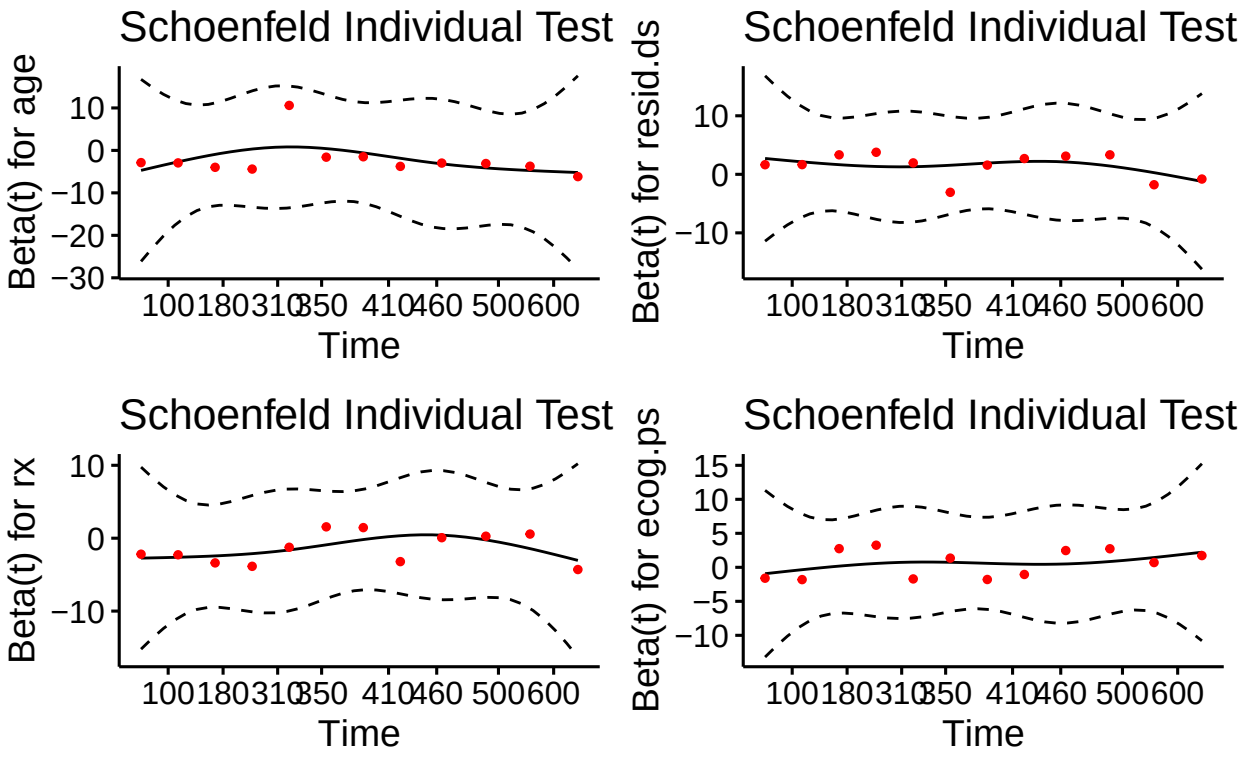

```
model_surv <- coxph(futime ~ . - fustat, data = df)
print(cox.zph(model_surv))
```

d, e)

```
##           chisq df    p
## age        1.06  1 0.30
## resid.ds   1.51  1 0.22
## rx         1.11  1 0.29
## ecog.ps    2.15  1 0.14
## GLOBAL     5.31  4 0.26
```

```
ggcoxzph(cox.zph(model_surv))
```

Global Schoenfeld Test p: 0.2571




```
summary(model_surv)
```

```
## Call:
## coxph(formula = futime ~ . - fustat, data = df)
##
##   n= 26, number of events= 12
##
##                coef exp(coef) se(coef)      z Pr(>|z|)
## ageyoung    -2.2013    0.1107   1.1069 -1.989   0.0467 *
## resid.dsyes  1.4470    4.2503   0.7292  1.984   0.0472 *
## rxB         -1.3814    0.2512   0.6448 -2.142   0.0322 *
## ecog.psbad   0.5859    1.7966   0.6329  0.926   0.3546
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##             exp(coef) exp(-coef) lower .95 upper .95
## ageyoung       0.1107     9.0368   0.01264    0.9686
## resid.dsyes    4.2503     0.2353   1.01801   17.7453
## rxB            0.2512     3.9803   0.07100    0.8891
## ecog.psbad     1.7966     0.5566   0.51966    6.2113
##
## Concordance= 0.782  (se = 0.065 )
## Likelihood ratio test= 12.19  on 4 df,   p=0.02
## Wald test            = 9.02  on 4 df,   p=0.06
## Score (logrank) test = 11.97  on 4 df,   p=0.02
```

```r
model_surv_reduced <- step(model_surv)
```

f)

```
## Start:  AIC=65.78
## futime ~ (fustat + age + resid.ds + rx + ecog.ps) - fustat
##
##            Df    AIC
## - ecog.ps   1 64.658
## <none>        65.775
## - resid.ds  1 68.377
## - rx        1 68.489
## - age       1 69.939
##
## Step:  AIC=64.66
## futime ~ age + resid.ds + rx
##
##            Df    AIC
## <none>        64.658
## - resid.ds  1 66.452
## - rx        1 66.945
## - age       1 68.537
```

```r
anova(model_surv, model_surv_reduced)
```

```
## Analysis of Deviance Table
##  Cox model: response is  futime
##  Model 1: ~ (fustat + age + resid.ds + rx + ecog.ps) - fustat
##  Model 2: ~ age + resid.ds + rx
##    loglik  Chisq Df Pr(>|Chi|)
## 1 -28.888
## 2 -29.329 0.8828  1     0.3474
```