

Analiza Przeżycia Raport 4

Paweł Matłowski
album 249732

25 lutego 2021

Spis treści

1	Lista nr 1	2
1.1	Zadanie nr 1	2
1.2	Zadanie nr 2	3
2	Lista nr 2	6
2.1	Zadanie nr 1	6
2.2	Zadanie nr 2	7
2.3	Zadanie nr 3	8
2.4	Zadanie nr 4	12
2.5	Zadanie nr 5	13

1 Lista nr 1

```
library(survival)
library(ggplot2)
library(survminer)

## Loading required package: ggpubr
## Loading required package: magrittr

df <- lung
df$status <- as.factor(df$status)
df$sex <- as.factor(df$sex)
df$ph.ecog <- as.factor(df$ph.ecog)
df$ph.karno <- as.factor(df$ph.karno)
df$pat.karno <- as.factor(df$pat.karno)
attach(df)
```

1.1 Zadanie nr 1

Przyjmując za zmienną zależną *time* i za indyktor cenzurowania zmienną *status*, dopasujmy model proporcjonalnych hazardów Coxa, biorąc zmienne objaśniające odpowiednio: *age*, *sex*, *ph.ecog*, *ph.karno*, *pat.karno*, *meal.cal*, *wt.loss*.

```
modelCox <- coxph(Surv(time, status==2)~age+sex+ph.ecog+ph.karno+pat.karno+meal.cal+wt.loss)
summary(modelCox)
```

```
## Call:
## coxph(formula = Surv(time, status == 2) ~ age + sex + ph.ecog +
##       ph.karno + pat.karno + meal.cal + wt.loss)
##
##      n= 168, number of events= 121
##      (60 observations deleted due to missingness)
##
##              coef  exp(coef)    se(coef)      z  Pr(>|z|)
## age           5.871e-03  1.006e+00  1.206e-02  0.487    0.6263
## sex2          -6.082e-01  5.443e-01  2.130e-01 -2.855    0.0043 **
## ph.ecog1       6.397e-01  1.896e+00  3.523e-01  1.816    0.0694 .
## ph.ecog2       1.320e+00  3.744e+00  5.445e-01  2.425    0.0153 *
## ph.ecog3       2.554e+00  1.285e+01  1.175e+00  2.173    0.0298 *
## ph.karno60     1.028e+00  2.795e+00  6.900e-01  1.490    0.1363
## ph.karno70     1.003e+00  2.727e+00  6.507e-01  1.542    0.1231
## ph.karno80     1.172e+00  3.229e+00  6.664e-01  1.759    0.0786 .
## ph.karno90     1.314e+00  3.721e+00  6.850e-01  1.918    0.0551 .
## ph.karno100    1.458e+00  4.298e+00  7.743e-01  1.883    0.0597 .
## pat.karno40    -3.519e-01  7.033e-01  1.522e+00 -0.231    0.8171
## pat.karno50     7.532e-01  2.124e+00  1.218e+00  0.619    0.5362
## pat.karno60     1.228e-01  1.131e+00  1.043e+00  0.118    0.9062
```

```
## pat.karno70 -1.740e-01 8.403e-01 1.077e+00 -0.162 0.8717
## pat.karno80 -2.811e-01 7.550e-01 1.076e+00 -0.261 0.7939
## pat.karno90 -6.893e-02 9.334e-01 1.078e+00 -0.064 0.9490
## pat.karno100 -5.681e-01 5.666e-01 1.100e+00 -0.517 0.6054
## meal.cal -4.431e-05 1.000e+00 2.866e-04 -0.155 0.8771
## wt.loss -1.394e-02 9.862e-01 8.358e-03 -1.668 0.0953 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## exp(coef) exp(-coef) lower .95 upper .95
## age 1.0059 0.9941 0.98240 1.0299
## sex2 0.5443 1.8372 0.35852 0.8264
## ph.ecog1 1.8959 0.5275 0.95041 3.7819
## ph.ecog2 3.7443 0.2671 1.28799 10.8852
## ph.ecog3 12.8534 0.0778 1.28434 128.6341
## ph.karno60 2.7954 0.3577 0.72293 10.8094
## ph.karno70 2.7273 0.3667 0.76185 9.7637
## ph.karno80 3.2291 0.3097 0.87474 11.9202
## ph.karno90 3.7213 0.2687 0.97181 14.2496
## ph.karno100 4.2982 0.2327 0.94231 19.6056
## pat.karno40 0.7033 1.4218 0.03563 13.8827
## pat.karno50 2.1237 0.4709 0.19528 23.0960
## pat.karno60 1.1307 0.8844 0.14651 8.7260
## pat.karno70 0.8403 1.1900 0.10179 6.9370
## pat.karno80 0.7550 1.3246 0.09161 6.2218
## pat.karno90 0.9334 1.0714 0.11287 7.7187
## pat.karno100 0.5666 1.7648 0.06566 4.8901
## meal.cal 1.0000 1.0000 0.99939 1.0005
## wt.loss 0.9862 1.0140 0.97013 1.0024
##
## Concordance= 0.663 (se = 0.029 )
## Likelihood ratio test= 34.71 on 19 df, p=0.02
## Wald test = 35.22 on 19 df, p=0.01
## Score (logrank) test = 38.96 on 19 df, p=0.004
```

1.2 Zadanie nr 2

Powtarzamy zadanie nr 1, ale nie uwzględniając zmiennych *age* oraz *meal.cal*. Przyjrzyjmy się także, jak wyglądają współczynniki przy zmiennych *sex* i *ph.ecog*

```
modelCox2 <- coxph(Surv(time, status==2)~sex+ph.ecog+ph.karno+pat.karno+wt.loss)
summary(modelCox2)

## Call:
## coxph(formula = Surv(time, status == 2) ~ sex + ph.ecog + ph.karno +
## pat.karno + wt.loss)
```

```

##      n= 210, number of events= 148
##      (18 observations deleted due to missingness)
##
##              coef exp(coef)  se(coef)      z Pr(>|z|)
## sex2          -0.666328   0.513591  0.184476 -3.612 0.000304 ***
## ph.ecog1       0.519718   1.681553  0.301850  1.722 0.085110 .
## ph.ecog2       1.116277   3.053463  0.468803  2.381 0.017260 *
## ph.ecog3       2.632819  13.912939  1.155307  2.279 0.022673 *
## ph.karno60     0.864977   2.374951  0.670182  1.291 0.196822
## ph.karno70     1.047251   2.849807  0.633549  1.653 0.098332 .
## ph.karno80     1.365861   3.919095  0.647187  2.110 0.034819 *
## ph.karno90     1.206609   3.342133  0.664854  1.815 0.069547 .
## ph.karno100    1.260148   3.525945  0.728086  1.731 0.083493 .
## pat.karno40    0.030517   1.030987  1.496812  0.020 0.983734
## pat.karno50    1.058281   2.881415  1.201632  0.881 0.378478
## pat.karno60    0.130915   1.139871  1.035814  0.126 0.899424
## pat.karno70   -0.144684   0.865296  1.056975 -0.137 0.891122
## pat.karno80   -0.325051   0.722491  1.057123 -0.307 0.758473
## pat.karno90   -0.334883   0.715422  1.062647 -0.315 0.752655
## pat.karno100  -0.507518   0.601988  1.076386 -0.472 0.637283
## wt.loss       -0.014248   0.985853  0.007329 -1.944 0.051896 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## sex2              0.5136    1.94707    0.35776    0.7373
## ph.ecog1          1.6816    0.59469    0.93063    3.0384
## ph.ecog2          3.0535    0.32750    1.21828    7.6531
## ph.ecog3         13.9129    0.07188    1.44551   133.9115
## ph.karno60        2.3750    0.42106    0.63855    8.8331
## ph.karno70        2.8498    0.35090    0.82326    9.8649
## ph.karno80        3.9191    0.25516    1.10230   13.9339
## ph.karno90        3.3421    0.29921    0.90803   12.3012
## ph.karno100       3.5259    0.28361    0.84631   14.6901
## pat.karno40       1.0310    0.96994    0.05485   19.3795
## pat.karno50       2.8814    0.34705    0.27338   30.3695
## pat.karno60       1.1399    0.87729    0.14968    8.6804
## pat.karno70       0.8653    1.15567    0.10901    6.8685
## pat.karno80       0.7225    1.38410    0.09099    5.7366
## pat.karno90       0.7154    1.39778    0.08913    5.7423
## pat.karno100      0.6020    1.66116    0.07301    4.9637
## wt.loss           0.9859    1.01435    0.97179    1.0001
##
## Concordance= 0.668 (se = 0.026 )
## Likelihood ratio test= 42.64  on 17 df,   p=5e-04
## Wald test              = 44.89  on 17 df,   p=3e-04
## Score (logrank) test = 49.06  on 17 df,   p=6e-05

```

```

modelCox2$coefficients['ph.ecog1']

## ph.ecog1
## 0.5197179

modelCox2$coefficients['ph.ecog2']

## ph.ecog2
## 1.116277

modelCox2$coefficients['ph.ecog3']

## ph.ecog3
## 2.632819

modelCox2$coefficients['sex2']

## sex2
## -0.6663276

```

Patrząc na współczynnik $\exp(\text{sex2})$ równy około 0.5, możemy zauważyć że "hazard" dla kobiet (indentyfikator 2) jest dwa razy mniejszy od "hazardu" dla mężczyzn. Innymi słowy, statystyczna szansa na śmierć mężczyzny jest dwa razy większy niż na śmierć kobiety. Analogicznie interpretujemy współczynniki ph.ecog względem ph.ecog0 (całkowita sprawność). Zatem przykładowo przy sprawności według lekarza ph.ecog równej 3, prawdopodobieństwo śmierci pacjenta jest około 13.9 razy od pacjenta całkowicie sprawnego.

2 Lista nr 2

2.1 Zadanie nr 1

Przyjmując za zmienną zależną *time* i za indyktor cenzurowania zmienną *status*, dopasujemy model proporcjonalnych hazardów Coxa, biorąc zmienne objaśniające odpowiednio: *age*, *sex*, *ph.ecog*, *ph.karno*, *pat.karno*, *meal.cal*, *wt.loss*. Następnie zweryfikujemy hipotezę o nieistotności zmiennej *meal.cal* w przyjętym modelu. Przyjmijmy poziom istotności $\alpha = 0.05$.

- $H_0 : \beta_m = 0$
- $H_1 : \beta_m \neq 0$

Jeśli p-value będzie większe od poziomu istotności, nie mamy podstawy do odrzucenia hipotezy, zatem zmienna *meal.cal* jest nieistotna. W przeciwnym razie, zmienna jest istotna w przyjętym modelu.

```
modelCox_mealcal <- coxph(Surv(time, status==2)~age+sex+ph.ecog+ph.karno+pat.karno+meal.
summary(modelCox_mealcal)

## Call:
## coxph(formula = Surv(time, status == 2) ~ age + sex + ph.ecog +
##       ph.karno + pat.karno + meal.cal + wt.loss)
##
##    n= 168, number of events= 121
##    (60 observations deleted due to missingness)
##
##              coef    exp(coef)    se(coef)      z  Pr(>|z|)
## age              5.871e-03  1.006e+00  1.206e-02  0.487   0.6263
## sex2            -6.082e-01  5.443e-01  2.130e-01 -2.855   0.0043 **
## ph.ecog1         6.397e-01  1.896e+00  3.523e-01  1.816   0.0694 .
## ph.ecog2         1.320e+00  3.744e+00  5.445e-01  2.425   0.0153 *
## ph.ecog3         2.554e+00  1.285e+01  1.175e+00  2.173   0.0298 *
## ph.karno60        1.028e+00  2.795e+00  6.900e-01  1.490   0.1363
## ph.karno70        1.003e+00  2.727e+00  6.507e-01  1.542   0.1231
## ph.karno80        1.172e+00  3.229e+00  6.664e-01  1.759   0.0786 .
## ph.karno90        1.314e+00  3.721e+00  6.850e-01  1.918   0.0551 .
## ph.karno100       1.458e+00  4.298e+00  7.743e-01  1.883   0.0597 .
## pat.karno40      -3.519e-01  7.033e-01  1.522e+00 -0.231   0.8171
## pat.karno50       7.532e-01  2.124e+00  1.218e+00  0.619   0.5362
## pat.karno60       1.228e-01  1.131e+00  1.043e+00  0.118   0.9062
## pat.karno70      -1.740e-01  8.403e-01  1.077e+00 -0.162   0.8717
## pat.karno80      -2.811e-01  7.550e-01  1.076e+00 -0.261   0.7939
## pat.karno90      -6.893e-02  9.334e-01  1.078e+00 -0.064   0.9490
## pat.karno100     -5.681e-01  5.666e-01  1.100e+00 -0.517   0.6054
## meal.cal         -4.431e-05  1.000e+00  2.866e-04 -0.155   0.8771
## wt.loss          -1.394e-02  9.862e-01  8.358e-03 -1.668   0.0953 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
##          exp(coef) exp(-coef) lower .95 upper .95
## age          1.0059    0.9941   0.98240    1.0299
## sex2          0.5443    1.8372   0.35852    0.8264
## ph.ecog1       1.8959    0.5275   0.95041    3.7819
## ph.ecog2       3.7443    0.2671   1.28799   10.8852
## ph.ecog3      12.8534    0.0778   1.28434  128.6341
## ph.karno60     2.7954    0.3577   0.72293   10.8094
## ph.karno70     2.7273    0.3667   0.76185    9.7637
## ph.karno80     3.2291    0.3097   0.87474   11.9202
## ph.karno90     3.7213    0.2687   0.97181   14.2496
## ph.karno100    4.2982    0.2327   0.94231   19.6056
## pat.karno40     0.7033    1.4218   0.03563   13.8827
## pat.karno50     2.1237    0.4709   0.19528   23.0960
## pat.karno60     1.1307    0.8844   0.14651    8.7260
## pat.karno70     0.8403    1.1900   0.10179    6.9370
## pat.karno80     0.7550    1.3246   0.09161    6.2218
## pat.karno90     0.9334    1.0714   0.11287    7.7187
## pat.karno100    0.5666    1.7648   0.06566    4.8901
## meal.cal       1.0000    1.0000   0.99939    1.0005
## wt.loss        0.9862    1.0140   0.97013    1.0024
##
## Concordance= 0.663 (se = 0.029 )
## Likelihood ratio test= 34.71 on 19 df,  p=0.02
## Wald test              = 35.22 on 19 df,  p=0.01
## Score (logrank) test = 38.96 on 19 df,  p=0.004
```

Zauważmy, że wartość p w teście Walda dla zmiennej *meal.cal* przyjmuje $p = 0.8771$, zatem ta zmienna jest statystycznie nieistotna w przyjętym modelu.

2.2 Zadanie nr 2

Na podstawie modelu dopasowanego w zadaniu nr 1 zweryfikujemy hipotezę o nieistotności zmiennej *pat.karno* w przyjętym modelu. Przyjmijmy poziom istotności $\alpha = 0.05$.

- $H_0 : \beta_p = 0$
- $H_1 : \beta_p \neq 0$

Jeśli p -value będzie większe od poziomu istotności, nie mamy podstawy do odrzucenia hipotezy, zatem zmienna *pat.karno* jest nieistotna. W przeciwnym razie, zmienna jest istotna w przyjętym modelu. Tym razem badamy zmienną dyskretną, a więc zastosujemy funkcję *Anova*.

```
modelCox_patkarno <- coxph(Surv(time, status==2)~age+sex+ph.ecog+ph.karno+pat.karno+meal
anova(modelCox_patkarno)

## Analysis of Deviance Table
## Cox model: response is Surv(time, status == 2)
## Terms added sequentially (first to last)
##
```

```
##          loglik    Chisq Df Pr(>|Chi|)
## NULL          -512.92
## age           -511.13   3.5613  1   0.059143 .
## sex           -508.42   5.4390  1   0.019692 *
## ph.ecog       -502.51  11.8091  3   0.008067 **
## ph.karno      -499.71   5.6049  5   0.346575
## pat.karno     -497.03   5.3602  7   0.616093
## meal.cal      -497.01   0.0317  1   0.858604
## wt.loss       -495.56   2.9040  1   0.088358 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Zauważmy, że wartość p w teście Walda dla zmiennej *meal.cal* przyjmuje $p = 0.616093$, zatem zakładamy że jest ona statystycznie nieistotna.

2.3 Zadanie nr 3

Korzystając z funkcji *step* dobierzemy odpowiedni model Coxa korzystając odpowiednio z:

- kryterium informacyjnego Akaike’a (*AIC*)
- kryterium *BIC*

```
df2 <- na.omit(df)
modelCoxSTEP <- coxph(Surv(df2$time, df2$status==2)~df2$age+df2$sex+df2$ph.ecog+df2$ph.karno+df2$pat.karno+df2$meal.cal+df2$wt.loss, data=df2, stepAIC=step)
step(modelCoxSTEP)

## Start:  AIC=1019.75
## Surv(df2$time, df2$status == 2) ~ df2$age + df2$sex + df2$ph.ecog +
##      df2$ph.karno + df2$pat.karno + df2$meal.cal + df2$wt.loss
##
##              Df      AIC
## - df2$pat.karno  7 1010.5
## - df2$ph.karno   5 1014.7
## - df2$meal.cal    1 1017.8
## - df2$age         1 1018.0
## <none>            1019.8
## - df2$wt.loss     1 1020.6
## - df2$ph.ecog     3 1021.1
## - df2$sex         1 1026.3
##
## Step:  AIC=1010.47
## Surv(df2$time, df2$status == 2) ~ df2$age + df2$sex + df2$ph.ecog +
##      df2$ph.karno + df2$meal.cal + df2$wt.loss
##
##              Df      AIC
## - df2$ph.karno   5 1006.8
## - df2$meal.cal    1 1008.5
```



```

## - df2$age      1 1009.0
## <none>         1010.5
## - df2$wt.loss  1 1011.7
## - df2$ph.ecog  3 1013.7
## - df2$sex      1 1018.3
##
## Step: AIC=1006.8
## Surv(df2$time, df2$status == 2) ~ df2$age + df2$sex + df2$ph.ecog +
##      df2$meal.cal + df2$wt.loss
##
##              Df      AIC
## - df2$meal.cal  1 1004.8
## - df2$age      1 1005.0
## <none>          1006.8
## - df2$wt.loss  1 1007.4
## - df2$sex      1 1012.4
## - df2$ph.ecog  3 1015.1
##
## Step: AIC=1004.81
## Surv(df2$time, df2$status == 2) ~ df2$age + df2$sex + df2$ph.ecog +
##      df2$wt.loss
##
##              Df      AIC
## - df2$age      1 1003.1
## <none>          1004.8
## - df2$wt.loss  1 1005.5
## - df2$sex      1 1010.5
## - df2$ph.ecog  3 1013.3
##
## Step: AIC=1003.07
## Surv(df2$time, df2$status == 2) ~ df2$sex + df2$ph.ecog + df2$wt.loss
##
##              Df      AIC
## <none>          1003.1
## - df2$wt.loss  1 1003.8
## - df2$sex      1 1009.0
## - df2$ph.ecog  3 1014.0
## Call:
## coxph(formula = Surv(df2$time, df2$status == 2) ~ df2$sex + df2$ph.ecog +
##      df2$wt.loss)
##
##              coef exp(coef) se(coef)      z      p
## df2$sex2      -0.546262  0.579110  0.199796 -2.734 0.006255
## df2$ph.ecog1  0.385355  1.470136  0.235724  1.635 0.102098
## df2$ph.ecog2  1.106652  3.024217  0.285246  3.880 0.000105
## df2$ph.ecog3  2.202191  9.044808  1.044267  2.109 0.034959
## df2$wt.loss   -0.012451  0.987626  0.007694 -1.618 0.105588
##

```

```
## Likelihood ratio test=23.16 on 5 df, p=0.000314
## n= 167, number of events= 120
```

Na podstawie kryterium *AIC* powinniśmy przyjąć model zależny od zmiennych: *sex*, *ph.ecog* oraz *wt.loss*.

```
n = length(df$status==2)
step(modelCoxSTEP, k = log(n))

## Start: AIC=1084.91
## Surv(df2$time, df2$status == 2) ~ df2$age + df2$sex + df2$ph.ecog +
##      df2$ph.karno + df2$pat.karno + df2$meal.cal + df2$wt.loss
##
##              Df      AIC
## - df2$pat.karno  7 1051.6
## - df2$ph.karno   5 1062.7
## - df2$ph.ecog    3 1076.0
## - df2$meal.cal   1 1079.5
## - df2$age        1 1079.7
## - df2$wt.loss    1 1082.4
## <none>           1084.9
## - df2$sex        1 1088.0
##
## Step: AIC=1051.63
## Surv(df2$time, df2$status == 2) ~ df2$age + df2$sex + df2$ph.ecog +
##      df2$ph.karno + df2$meal.cal + df2$wt.loss
##
##              Df      AIC
## - df2$ph.karno   5 1030.8
## - df2$ph.ecog    3 1044.5
## - df2$meal.cal   1 1046.2
## - df2$age        1 1046.8
## - df2$wt.loss    1 1049.4
## <none>           1051.6
## - df2$sex        1 1056.0
##
## Step: AIC=1030.81
## Surv(df2$time, df2$status == 2) ~ df2$age + df2$sex + df2$ph.ecog +
##      df2$meal.cal + df2$wt.loss
##
##              Df      AIC
## - df2$meal.cal   1 1025.4
## - df2$age        1 1025.6
## - df2$wt.loss    1 1028.0
## - df2$ph.ecog    3 1028.8
## <none>           1030.8
## - df2$sex        1 1033.0
##
```

```

## Step:  AIC=1025.39
## Surv(df2$time, df2$status == 2) ~ df2$age + df2$sex + df2$ph.ecog +
##      df2$wt.loss
##
##           Df    AIC
## - df2$age    1 1020.2
## - df2$wt.loss 1 1022.6
## - df2$ph.ecog 3 1023.6
## <none>        1025.4
## - df2$sex     1 1027.7
##
## Step:  AIC=1020.22
## Surv(df2$time, df2$status == 2) ~ df2$sex + df2$ph.ecog + df2$wt.loss
##
##           Df    AIC
## - df2$wt.loss 1 1017.6
## <none>        1020.2
## - df2$ph.ecog 3 1020.8
## - df2$sex     1 1022.7
##
## Step:  AIC=1017.56
## Surv(df2$time, df2$status == 2) ~ df2$sex + df2$ph.ecog
##
##           Df    AIC
## - df2$ph.ecog 3 1015.4
## <none>        1017.6
## - df2$sex     1 1018.9
##
## Step:  AIC=1015.42
## Surv(df2$time, df2$status == 2) ~ df2$sex
##
##           Df    AIC
## <none>        1015.4
## - df2$sex    1 1016.2
## Call:
## coxph(formula = Surv(df2$time, df2$status == 2) ~ df2$sex)
##
##           coef exp(coef) se(coef)      z      p
## df2$sex2 -0.4792    0.6193   0.1966 -2.437 0.0148
##
## Likelihood ratio test=6.25  on 1 df, p=0.01244
## n= 167, number of events= 120

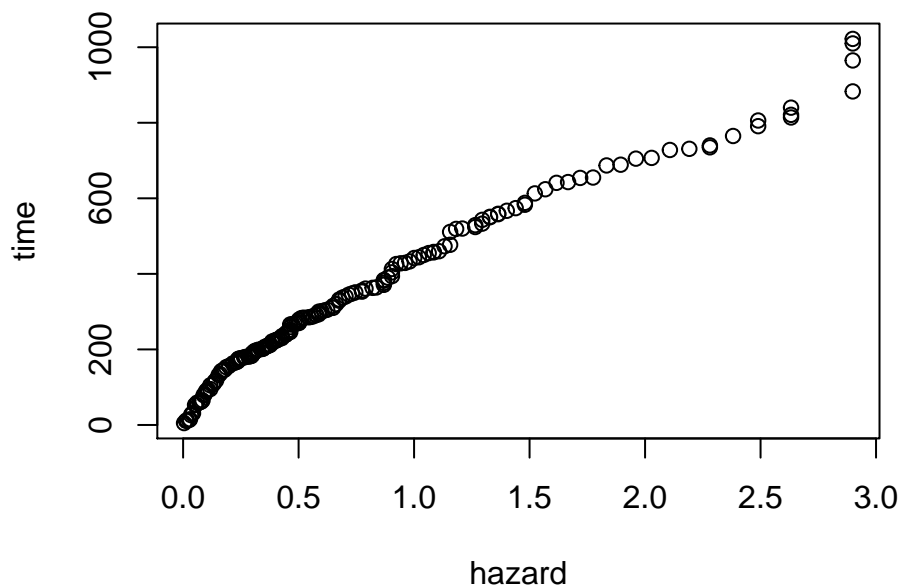
```

Na podstawie kryterium *BIC* powinniśmy przyjąć model zależny tylko od zmiennej *sex*.

2.4 Zadanie nr 4

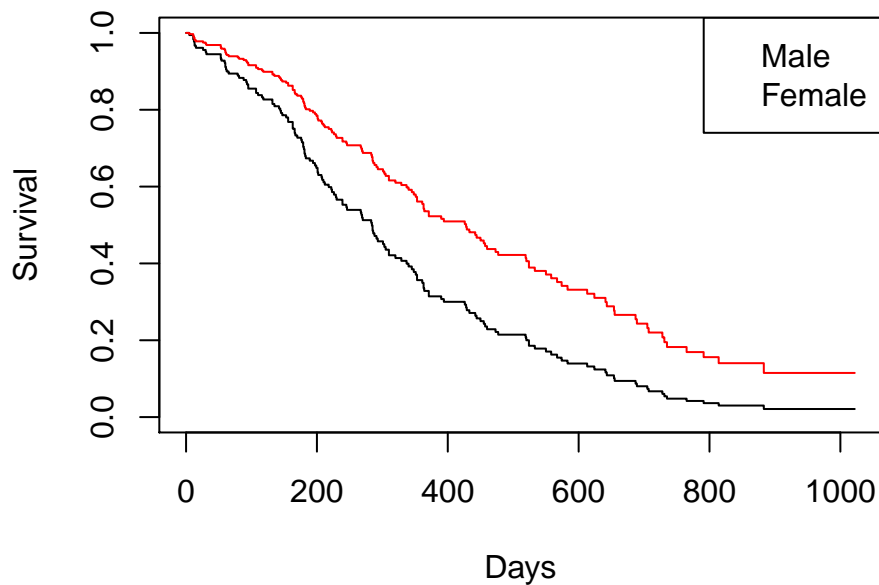
Na podstawie modelu wybranego w zadaniu nr 3 (za pomocą kryterium AIC) zwizualizujemy funkcję hazardu i funkcję przeżycia dla mężczyzny o parametrach ph.ecog=1 i wt.loss=4.

```
Coxseq <- seq(0.01, 0.99, by = 0.01)
ModelCoxAIC <- coxph(Surv(time, status==2)~as.factor(sex)+as.factor(ph.ecog)+wt.loss)
CoxAIC <- basehaz(ModelCoxAIC, centered=TRUE)
plot(CoxAIC)
```



Rysunek 1: Funkcja hazardu

```
v = expand.grid(sex = c(1,2), ph.ecog = 1, wt.loss = 4)
csurv <- survfit(ModelCoxAIC, newdata = v, conf.int=FALSE)
plot(csurv, col=1:2, xlab = "Days", ylab = "Survival")
legend("topright", legend=c("Male","Female"), col=c("red", "black"))
```



Rysunek 2: Funkcja przeżycia z podziałem ze względu na płeć o wartościach `ph.ecog=1`, `wt.loss=4`

2.5 Zadanie nr 5

Zweryfikujemy hipotezę o proporcjonalności hazardu. Wykorzystamy do tego test Grambscha i Therneau, który został zaimplementowany w pakiecie R i można go wykonać przy użyciu funkcji `cox.zhp`.

```
modelCox <- coxph(Surv(time, status==2)~age+sex+ph.ecog+ph.karno+pat.karno+meal.cal+wt.loss)
prop = cox.zph(modelCox)
prop
```

##		chisq	df	p
##	age	0.0378	1	0.8458
##	sex	1.5108	1	0.2190
##	ph.ecog	6.1399	3	0.1050
##	ph.karno	6.8290	5	0.2337
##	pat.karno	6.8012	7	0.4499
##	meal.cal	6.7177	1	0.0095
##	wt.loss	0.0623	1	0.8029
##	GLOBAL	21.3461	19	0.3180

```
propAIC = cox.zph(ModelCoxAIC)
propAIC
```

##		chisq	df	p
----	--	-------	----	---

```
## as.factor(sex)      2.9699  1 0.085
## as.factor(ph.ecog)  5.2090  3 0.157
## wt.loss             0.0493  1 0.824
## GLOBAL              7.9660  5 0.158
```

Zarówno w modelu ze wszystkimi charakterystykami, jak i w modelu wyznaczonym z kryterium AIC, p-value wynosi powyżej 0.05, zatem nie ma podstaw do odrzucenia hipotezy proporcjonalności hazardów. "Moment", w którym powinniśmy zająć się badaniem hipotezy proporcjonalnych hazardów jest wtedy, gdy podejrzewamy, że zmienne objaśniające mogą być zależne od czasu. Założeniem proporcjonalności hazardów jest to, że jest stały w czasie. Możemy to sprawdzić za pomocą p-value zmiennych objaśniających. Jeśli p-value zmiennej jest poniżej poziomu krytycznego, to nasza hipoteza o zależności od czasu jest uzasadniona. Lecz w naszym modelu przyjętym według kryterium AIC nie mamy żadnej takiej zmiennej objaśniającej.