

Multivariate Analysis for the Behavioral Sciences,
Second Edition (Chapman and Hall/CRC, 2019)

Examples of Chapter 7:
Survival Analysis

Kimmo Vehkalahti and Brian S. Everitt

14 November 2018

Contents

Examples	2
Table 7.1: Age at First Sexual Intercourse for Women in Two Diagnostic Groups	2
Table 7.2	3
Figure 7.1	4
Figure 7.2	5
Table 7.3: Data for Heroin Addicts Being Treated with Methadone	7
Figure 7.4	9
Table 7.4	10

Table 7.1: Age at First Sexual Intercourse for Women in Two Diagnostic Groups

##	Diagnosis	Agefs	Age	Status
## 1	RAN	30	30	0
## 2	RAN	24	24	0
## 3	RAN	12	18	1
## 4	RAN	21	42	1
## 5	RAN	19	19	0
## 6	RAN	18	39	1

##	Diagnosis	Agefs	Age	Status
## 79	NC	19	26	1
## 80	NC	20	23	1
## 81	NC	16	20	1
## 82	NC	15	21	1
## 83	NC	17	21	1
## 84	NC	18	21	1

Table 7.2

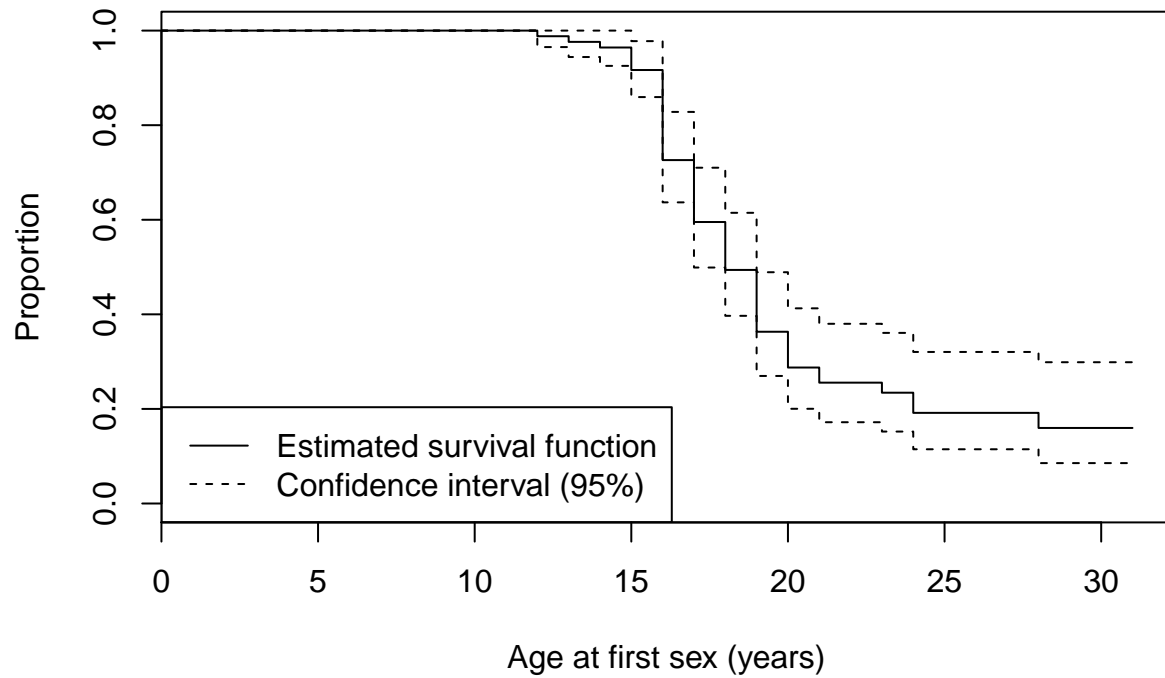
```
attach(first_sex)
library("survival")
summary(survfit(Surv(Agefs, Status) ~ 1))

## Call: survfit(formula = Surv(Agefs, Status) ~ 1)
##
##      time  n.risk  n.event  survival  std.err  lower 95% CI  upper 95% CI
##      12      84       1    0.988  0.0118    0.9652    1.000
##      13      83       1    0.976  0.0166    0.9441    1.000
##      14      82       1    0.964  0.0202    0.9254    1.000
##      15      81       4    0.917  0.0302    0.8594    0.978
##      16      77      16    0.726  0.0487    0.6368    0.828
##      17      61      11    0.595  0.0536    0.4990    0.710
##      18      47       8    0.494  0.0551    0.3969    0.615
##      19      34       9    0.363  0.0551    0.2697    0.489
##      20      24       5    0.288  0.0530    0.2003    0.413
##      21      18       2    0.256  0.0517    0.1719    0.380
##      23      12       1    0.234  0.0516    0.1521    0.361
##      24      11       2    0.192  0.0503    0.1147    0.320
##      28       6       1    0.160  0.0510    0.0854    0.299

detach(first_sex)
```

Figure 7.1

```
attach(first_sex)
plot(survfit(Surv(Agefs, Status) ~ 1), xlab = "Age at first sex (years)", ylab = "Proportion")
legend("bottomleft", c("Estimated survival function", "Confidence interval (95%)"), lty = 1:2)
```

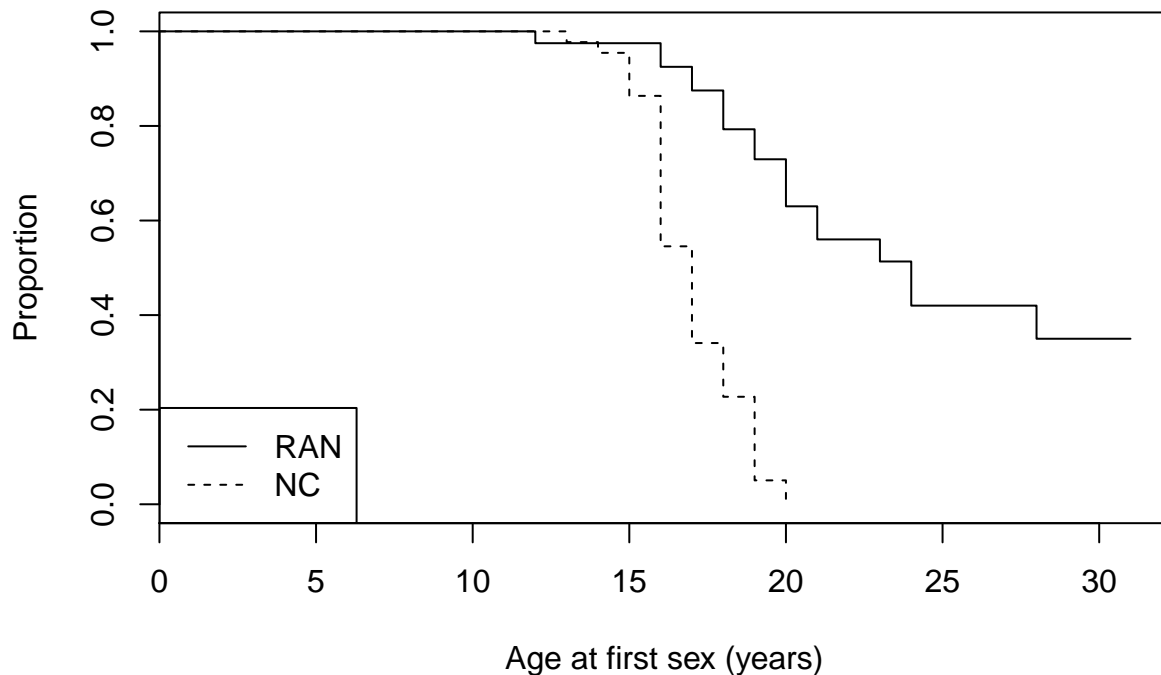


```
detach(first_sex)
```

Figure 7.2

The following gives also the log-rank test and Cox regression results that are referred to in the text.

```
attach(first_sex)
plot(survfit(Surv(Agefs,Status) ~ Diagnosis), xlab = "Age at first sex (years)",
     ylab = "Proportion", lty = 1:2)
legend("bottomleft", c("RAN", "NC"), lty = 1:2)
```



```
# logrank test
survdif(Surv(Agefs, Status) ~ Diagnosis)

## Call:
## survdif(formula = Surv(Agefs, Status) ~ Diagnosis)
##
##           N Observed Expected (O-E)^2/E (O-E)^2/V
## Diagnosis=RAN 40      19    40.8      11.6      46.4
## Diagnosis=NC  44      43    21.2      22.3      46.4
##
##  Chisq= 46.4  on 1 degrees of freedom, p= 9.81e-12

#cox regression
sex_cox <- coxph(Surv(Agefs, Status) ~ Diagnosis)
summary(sex_cox)

## Call:
## coxph(formula = Surv(Agefs, Status) ~ Diagnosis)
##
```

```

##    n= 84, number of events= 62
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## DiagnosisNC 2.1371    8.4746   0.3457 6.183 6.3e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## DiagnosisNC      8.475      0.118      4.304      16.69
##
## Concordance= 0.726 (se = 0.041 )
## Rsquare= 0.438 (max possible= 0.996 )
## Likelihood ratio test= 48.46 on 1 df,  p=3.364e-12
## Wald test               = 38.23 on 1 df,  p=6.301e-10
## Score (logrank) test = 49.92 on 1 df,  p=1.602e-12
contrasts(Diagnosis)

##      NC
## RAN  0
## NC   1
detach(first_sex)

```

Table 7.3: Data for Heroin Addicts Being Treated with Methadone

```
heroin_dat <- structure(list(
  Clinic = structure(c(1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L,
    1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L,
    1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L,
    1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L,
    1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L,
    1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L,
    1L, 1L, 1L, 1L, 1L, 1L, 1L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L,
    2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L,
    2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L,
    2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 1L, 1L, 1L,
    2L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L,
    1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L,
    2L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L,
    2L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 2L,
    1L, 1L, 1L, 2L, 1L, 2L, 2L, 1L, 1L, 2L, 2L, 1L, 1L),
    .Label = c("Clinic 1", "Clinic 2"), class = "factor"),
  Status = c(1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
    1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1,
    1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 1,
    1, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, 1,
    1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0,
    1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0,
    0, 1, 0, 0, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0,
    1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1,
    0, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0,
    0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1),
  TimeDays = c(428, 275, 262, 183, 259, 714, 438, 796, 892, 393, 161, 836, 523, 612,
    212, 399, 771, 514, 512, 624, 209, 341, 299, 826, 262, 566, 368, 302, 602, 652,
    293, 564, 394, 755, 591, 787, 739, 550, 837, 612, 581, 523, 504, 785, 774, 560,
    160, 482, 518, 683, 147, 563, 646, 899, 857, 180, 452, 760, 496, 258, 181, 386,
    439, 563, 337, 613, 192, 405, 667, 905, 247, 821, 821, 517, 346, 294, 244, 95, 376,
    212, 96, 532, 522, 679, 408, 840, 148, 168, 489, 541, 205, 475, 237, 517, 749, 150,
    465, 708, 713, 146, 450, 555, 460, 53, 122, 35, 532, 684, 769, 591, 769, 609, 932,
    932, 587, 26, 72, 641, 367, 633, 661, 232, 13, 563, 969, 1052, 944, 881, 190, 79, 884,
    170, 286, 358, 326, 769, 161, 564, 268, 611, 322, 1076, 2, 788, 575, 109, 730, 790,
    456, 231, 143, 86, 1021, 684, 878, 216, 808, 268, 222, 683, 496, 389, 126, 17, 350,
    531, 317, 461, 37, 167, 358, 49, 457, 127, 7, 29, 62, 150, 223, 129, 204, 129, 581, 176,
    30, 41, 543, 210, 193, 434, 367, 348, 28, 337, 175, 149, 546, 84, 283, 533, 207, 216,
    28, 67, 62, 111, 257, 136, 342, 41, 531, 98, 145, 50, 53, 103, 2, 157, 75, 19, 35, 394,
    117, 175, 180, 314, 480, 325, 280, 204, 366, 531, 59, 33, 540, 551, 90, 47),
  Prison = structure(c(1L, 2L, 1L, 1L, 2L, 1L, 2L, 2L, 1L,
    2L, 2L, 2L, 1L, 1L, 2L, 2L, 2L, 2L, 1L, 2L, 2L, 2L, 1L, 1L,
    2L, 2L, 2L, 2L, 1L, 1L, 1L, 1L, 2L, 2L, 1L, 1L, 1L, 1L, 2L, 2L, 1L,
    1L, 2L, 1L, 1L, 1L, 2L, 2L, 1L, 1L, 1L, 1L, 2L, 2L, 1L, 1L,
    1L, 1L, 1L, 2L, 1L, 2L, 1L, 2L, 2L, 2L, 1L, 1L, 1L, 2L, 1L,
```

```

1L, 1L, 2L, 1L, 1L, 1L, 1L, 2L, 1L, 1L, 1L, 2L, 1L, 2L, 1L,
1L, 1L, 1L, 1L, 2L, 2L, 1L, 1L, 2L, 1L, 2L, 2L, 2L, 2L,
1L, 1L, 2L, 1L, 1L, 1L, 2L, 2L, 1L, 1L, 1L, 2L, 1L, 2L,
1L, 2L, 1L, 1L, 2L, 2L, 1L, 2L, 2L, 2L, 1L, 2L, 2L, 1L,
1L, 2L, 2L, 1L, 2L, 2L, 2L, 1L, 2L, 2L, 1L, 1L, 2L, 1L,
1L, 1L, 1L, 2L, 2L, 1L, 2L, 2L, 2L, 1L, 2L, 1L, 1L, 2L, 1L,
2L, 2L, 1L, 2L, 2L, 2L, 2L, 2L, 1L, 1L, 1L, 1L, 1L, 2L, 2L,
1L, 1L, 2L, 1L, 2L, 2L, 2L, 1L, 2L, 1L, 1L, 2L, 1L, 1L, 2L,
2L, 1L, 2L, 2L, 1L, 1L, 2L, 1L, 2L, 1L, 1L, 2L, 2L, 2L,
2L, 1L, 2L, 1L, 2L, 2L, 1L, 1L, 2L, 1L, 1L, 2L, 2L, 2L,
1L, 1L, 1L, 1L),
.Label = c("No prison record", "Prison record"), class = "factor"),

Dose = c(50, 55, 55, 30, 65, 55, 65,
60, 50, 65, 80, 60, 55, 70, 60, 60, 75, 80, 80, 80, 60, 60,
55, 80, 65, 45, 55, 50, 60, 80, 65, 60, 55, 65, 55, 80, 60,
60, 60, 65, 70, 60, 60, 80, 65, 65, 35, 30, 65, 50, 65, 70,
60, 60, 60, 70, 60, 60, 65, 40, 60, 60, 80, 75, 65, 60, 80,
80, 50, 80, 70, 80, 75, 45, 60, 65, 60, 60, 55, 40, 70, 80,
70, 35, 50, 80, 65, 65, 80, 80, 50, 75, 45, 70, 70, 80, 65,
60, 50, 50, 55, 80, 50, 60, 60, 40, 70, 65, 70, 70, 40, 100,
80, 80, 110, 40, 40, 70, 70, 70, 40, 70, 60, 70, 80, 80,
80, 80, 50, 40, 50, 40, 45, 60, 60, 40, 40, 80, 70, 40, 55,
80, 40, 70, 80, 70, 80, 90, 70, 60, 70, 40, 80, 80, 60, 100,
60, 40, 40, 100, 40, 55, 75, 40, 60, 65, 50, 75, 60, 55,
45, 60, 40, 20, 40, 60, 40, 60, 40, 40, 65, 50, 65, 55, 60,
60, 40, 50, 70, 55, 45, 60, 50, 40, 60, 80, 50, 45, 80, 55,
50, 50, 50, 50, 60, 55, 60, 55, 60, 40, 45, 40, 55, 50, 50,
50, 60, 60, 55, 40, 60, 80, 40, 60, 60, 70, 50, 60, 90, 50,
55, 50, 45, 60, 80, 65, 40, 45)),

.Names = c("Clinic", "Status", "TimeDays", "Prison", "Dose"),
row.names = c(NA, -238L), class = "data.frame")

head(heroin_dat, n = 10)

```

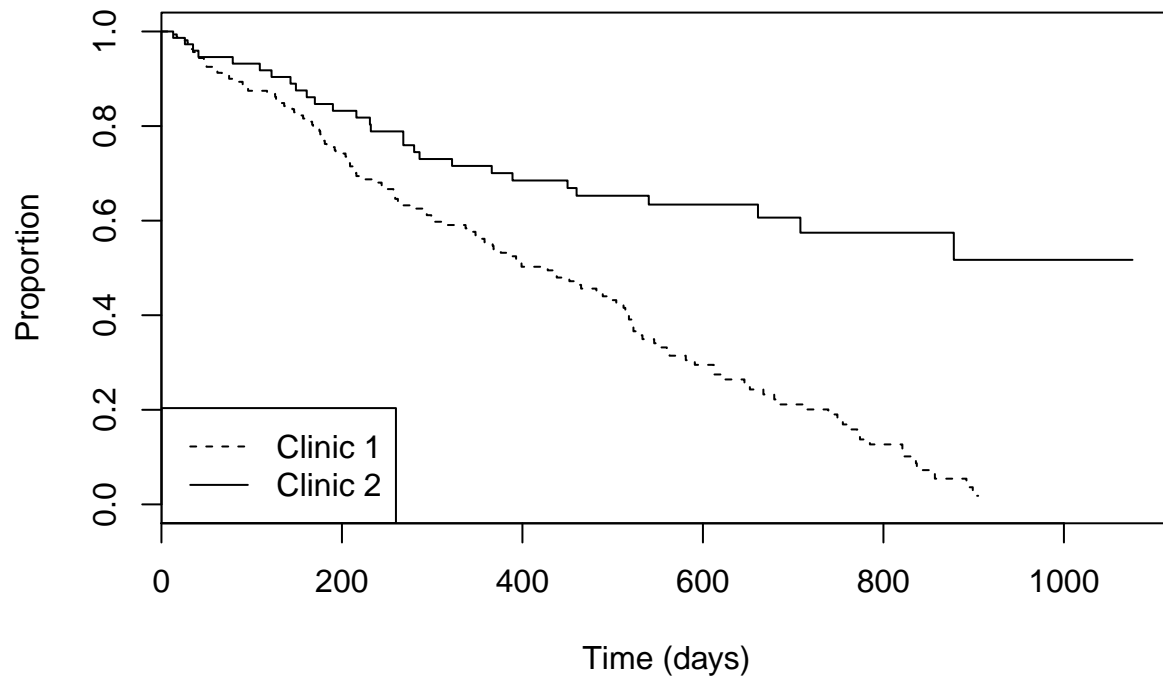
```

##      Clinic Status TimeDays      Prison Dose
## 1 Clinic 1      1      428 No prison record 50
## 2 Clinic 1      1      275   Prison record 55
## 3 Clinic 1      1      262 No prison record 55
## 4 Clinic 1      1      183 No prison record 30
## 5 Clinic 1      1      259   Prison record 65
## 6 Clinic 1      1      714 No prison record 55
## 7 Clinic 1      1      438   Prison record 65
## 8 Clinic 1      0      796   Prison record 60
## 9 Clinic 1      1      892 No prison record 50
## 10 Clinic 1     1      393   Prison record 65

```


Figure 7.4

```
attach(heroine_dat)
plot(survfit(Surv(TimeDays, Status) ~ Clinic), xlab = "Time (days)", ylab = "Proportion",
     lty = c(2,1)) # 1:2
legend("bottomleft", c("Clinic 1", "Clinic 2"), lty = c(2,1)) # 1:2
```



```
contrasts(Clinic)
```

```
##          Clinic 2
## Clinic 1         0
## Clinic 2         1
```

Table 7.4

```
heroin_cox <- coxph(Surv(TimeDays, Status) ~ Prison + Dose + Clinic)
summary(heroin_cox)
```

```
## Call:
## coxph(formula = Surv(TimeDays, Status) ~ Prison + Dose + Clinic)
##
##      n= 238, number of events= 150
##
##              coef exp(coef)    se(coef)      z Pr(>|z|)
## PrisonPrison record  0.326555  1.386184  0.167225  1.953  0.0508 .
## Dose                -0.035369  0.965249  0.006379 -5.545 2.94e-08 ***
## ClinicClinic 2      -1.009896  0.364257  0.214889 -4.700 2.61e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## PrisonPrison record    1.3862    0.7214    0.9988    1.9238
## Dose                    0.9652    1.0360    0.9533    0.9774
## ClinicClinic 2         0.3643    2.7453    0.2391    0.5550
##
## Concordance= 0.665  (se = 0.026 )
## Rsquare= 0.238  (max possible= 0.997 )
## Likelihood ratio test= 64.56  on 3 df,   p=6.228e-14
## Wald test               = 54.12  on 3 df,   p=1.056e-11
## Score (logrank) test = 56.32  on 3 df,   p=3.598e-12
```

Other results mentioned in the text following Table 7.4:

```
# Backward elimination etc.
```

```
step(heroin_cox, direction="backward")
```

```
## Start: AIC=1352.52
```

```
## Surv(TimeDays, Status) ~ Prison + Dose + Clinic
```

```
##
```

```
##           Df      AIC
```

```
## <none>      1352.5
```

```
## - Prison    1 1354.3
```

```
## - Clinic    1 1376.9
```

```
## - Dose      1 1381.3
```

```
## Call:
```

```
## coxph(formula = Surv(TimeDays, Status) ~ Prison + Dose + Clinic)
```

```
##
```

```
##           coef exp(coef) se(coef)      z      p
```

```
## PrisonPrison record  0.32655   1.38618  0.16722  1.95  0.051
```

```
## Dose                -0.03537   0.96525  0.00638 -5.54 2.9e-08
```

```
## ClinicClinic 2      -1.00990   0.36426  0.21489 -4.70 2.6e-06
```

```
##
```

```
## Likelihood ratio test=64.6 on 3 df, p=6.23e-14
```

```
## n= 238, number of events= 150
```

```
heroin_zph <- cox.zph(heroin_cox)
```

```
heroin_zph
```

```
##           rho chisq      p
```

```
## PrisonPrison record -0.0382  0.22 0.639369
```

```
## Dose                0.0724  0.70 0.402749
```

```
## ClinicClinic 2      -0.2578 11.19 0.000824
```

```
## GLOBAL              NA 12.62 0.005546
```

```
# For Exercise 7.1, use strata as follows:
```

```
#heroin_cox <- coxph(Surv(TimeDays, Status) ~ Prison + Dose + strata(Clinic))
```