Exercise 1:

Let T denote survival time with survival function $S_T(t)$. Simulate a sample of size $n = 1500$ from a Cox model with hazard rate

$$h(t; x) = t \cdot \exp(0.5x) \ .$$

Use the inverse transform sampling method developed in exercise 5, study sheet 7. Simulate the covariate x_1 from a uniform distribution on the interval $[-3, 3]$ and the censoring times from a uniform distribution on the interval $[0, 6]$. Plot the Cox-Snell residuals against the cumulative hazard rate to check the overall goodness-of-fit of the fitted model. For the derivation of the distribution of the Cox-Snell residuals use the distribution of $Y = -\ln(S_T(T)) \sim \mathcal{E}(\lambda = 1)$.

Exercise 2:

Simulate a sample of size $n = 1500$ from a Cox model with hazard rate

(a) $h(t; x) = t \cdot \exp(\sin(x_1) + 0.5x_2)$.

(b) $h(t; x) = t \cdot \exp(x_1^2 + 0.5x_2)$.

Use the inverse transform sampling method developed in exercise 5, study sheet 7. Simulate the covariate x_1 and x_2 from a uniform distribution on the interval $[-3, 3]$ and the censoring times from a uniform distribution on the interval $[0, 6]$. Obtain the martingale residuals and deviance residuals and check whether one can use them to make conclusions about the functional form of the covariate x_1. The `loess()` function can be used to smooth the residuals.

Exercise 3:

In the lectures, the martingale property has been stated as follows (see slide 7 of the set of slides "Refinements of the semiparametric proportional hazards model"):

$$\mathrm{E}\left[\mathrm{d}M(t)|\mathcal{F}_{t-}\right] = 0 \quad \text{for all } t \ . \tag{1}$$

Show that equation (1) is equivalent to

$$\mathrm{E}\left[M(t)|\mathcal{F}_s\right] = M(s) \quad \text{for all } s < t \ . \tag{2}$$

Exercise 4:

The file `resmelanoma.prn` that is available in the Stud.IP folder "Data" contains survival times from 30 resected melanoma patients (for a description of the data, see the file `resmelanomahelp.txt`).[1] Let `ageg` denote the age group with `ageg = 1` if age < 45 and `ageg = 2` otherwise. Fit the survival times with an `ageg`-stratified Cox proportional hazards model with the covariates sex and treatment received.

Exercise 5:

The file `prison.txt`, which is available in the Stud.IP folder "Data", contains data from an experimental study of recidivism of 432 male prisoners, who were observed for a year after being released from prison.[2] Half of the prisoners were randomly given financial aid when they were released. The following table gives a description of the observed variables:

Variable	Description
`week`	week of first arrest after release, or censoring time
`arrest`	the event indicator, 1 = arrested , 0 = not
`fin`	1 = received financial aid, 0 = not
`age`	in years at the time of release
`race`	1 = black, 0 = others
`wexp`	1 = had full-time work experience, 0 = not
`mar`	1 = married, 0 = not
`paro`	1 = released on parole, 0 = not
`prio`	number of prior convictions
`educ`	codes 2 (grade 6 or less), 3 (grades 6 through 9), 4 (grades 10 and 11), 5 (grade 12), or 6 (some post-secondary)
`emp1 - emp52`	1 = employed in the corresponding week, 0 = not

(a) Fit a Cox model to these data. Use backward selection, which is implemented in the function `stepAIC()` function from the **R** package `MASS`, to find the best model according to the Akaike Information Criterion (AIC).

(b) In the file `prisonlong.txt` each row corresponds to one observation per person per week. Fit a Cox model with the time-dependent variable `employed` to these data.

(c) Create a variable `employed.lag1` which should contain information whether the person was employed in the previous week. Again, fit a Cox model using the variable `employed.lag1` instead to `employed`.

(d) How could you check the assumption of proportional hazards for all the variables of the best model found in (a) using interaction terms with time of observation?

[1] Data of Table 3.1 in Lee, E. T. and Wang, J. W. (2013): *Statistical Methods for Survival Data Analysis*, 4th edition, Wiley.

[2] Rossi, P. H., Berk, R. A. and Lenihan, K. J. (1980): *Money, Work and Crime: Some Experimental Results*, New York: Academic Press.