Conditional logistic regression

Exercises

In this exercise we look at the data of a matched case-control study. The analysis to be used for this data is conditional logistic regression. We can perform this type of analysis with the `clogit` function which is located in the `survival` package. The data set we use is called endom and is already loaded.

The data are from the Los Angeles study (reported by Mack e.a., 1976) of endometrial cancer. The aim was to study the effect of exegenous oestrogens on the risk of endometrial cancer. Each of the 63 cases was matched to 4 control women, living in the same community, with the same age and marital status. (The difference in age is never more than two years). Variables we will use are: - `STRATUM` stratum number

- `OUTCOME` 0 = control, 1 = case
- `GALL` history of gall-bladder disease
- `HYP` history of hypertension
- `OB` obesitas
- `EST` history of oestrogen use
- `DOS` dose of oestrogens in mg/day
- `DUR` duration of oestrogen use in months
- `NON` use of other drugs:
- `DURATION` duration of oestrogen use (continuous in months)
- `AGE` age of the subject
- `AGEC` age of the matched case
- `AGEGR` age of the matched case in categories:

Question 1

First we recode the outcome as 0/1. We call this variable OUT. Estimate a conditional logistic regression model to investigate the effect of history of oestrogen use on endometrial cancer using `clogit`. Call the fitted model `fit1`. Interpret the results using `summary`.

Question 2

Because the patients are matched on age we cannot put the main effect of this variable in the model. We can however look at interaction effects. Add the interaction with AGEC to the model (and call this fit2) and test if this is statistically significant using `anova`.

Question 3

We think `GALL`, `NON`, `OB` and `HYP` could be important variables. Extend model `fit1` to accommodate this (we leave out the interaction this time). Use again the `summary()` method to get a detailed output and interpret the results. Name the fitted model `fit3`.

Question 4

Perhaps the effect of EST is different if you have gall bladder disease. Add the interaction to the model. Check whether it is significant using an chi-square test.

Question 5

Now we also want to look at the dose of oestrogen. However the variable `DOS` missing for some individuals. When we do embedding tests we need to make sure we are comparing the same individuals. This can be done using the complete.cases() function. Add the dose to the model with EST + GALL + NON + OB + HYP and perform a chi-square test.