April 12, 2015
Erik Parner

PhD Course in Basic Biostatistics
Exam (J.nr.: 1050/23)

Practical information

Submission options and deadline

Individual solutions should be handed in as a single pdf-file by email to the following email-address:

BBEKSAMEN@BIOSTAT.AU.DK

The pdf-file should be named: fullname.pdf. The file must contain both your written answers and the appendix (see below).

Your solution has to be submitted no later than Friday May 8, 2015, at 9 AM.

Guidelines, requirements, and hints for preparing solutions

- Answer all questions. In particular, be aware that some questions comprise several sub-questions that all must be addressed.
- Plot the data whenever reasonable using scatter plots, histograms, Q-Q plots, etc.
- Always specify the statistical model used in the analysis.
- Any quantification of the findings of a statistical analysis in terms of estimates should be accompanied by confidence intervals. Any comparative statement should be backed up by a test and a p-value.
- Model validation is an integral part of any statistical analysis. It is not necessary to ask for relevant model validation to be performed, this should always be done.
- Include the Stata commands used for the analysis (the do/syntax-files) and log/output-files in appendices.
- No Stata code or output (except graphs) outside the appendix!
- Formulate the conclusions using relevant terms from the context of the study (it is important to be able to translate the findings from the statistical analysis into conclusions regarding the initial scientific question).

Background information on the data

In a Danish cross sectional study among adults 67 years of age and older, living at home, a questionnaire was send to 2000 randomly selected persons. In total 979 returned a full or partly filled questionnaire. The purpose of the study was to determine risk factors for falling accidents. We will here consider the factors shown below on the 909 persons that had a complete questionnaire; data are given in the data file `balance.dta`. The Bergs balance index – the last variable – was however only measured on 102 randomly selected persons among those returning a questionnaire.

Self-evaluated balance is calculated as a sum-score; the participants were asked to evaluate their safety in 12 different basic activities. Each activity was scored from a scale from 1 to 5, where 1 is very unsafe and 5 to very safe. The balance score was calculated as a sum of all 12 answers.

The Berg Balance index is the result of a 45 minutes program based evaluation of balance and mobility. Each of 14 different tasks (sit, stand, lean forward, ect.) was evaluated on a scale from 0 (cannot) to 4 (normal performance) and the total score was hereafter calculated. The Berg Balance index requires an observator and the score is therefore in the study only performed on a subsample of the participants.

The file `balance.dta` contains the data variable:

variable name	storage type	display format	value label	variable label
age	float	%9.0g		Age
sex	float	%9.0g	sexlab	Sex
fall	float	%15.0g	falllab	Indicator of a fall within the last 6 months
balance	float	%9.0g		The balance sum score (scale from 12-60)
berg	float	%9.0g		The Bergs balance index (scale from 0-56)

We will first consider the connection between the person's age and risk of a fall within the last 6 months.

Question 1

a. Divide the participants into the age groups; "under 70 years", "70-74 years", 75-79 years", "80-84 years", "85-89 years" and "90 years or older".
b. Perform a logistic model for women for the odds of fall with the above age groups as explanatory variable. Interpret the parameters.
c. We will now consider if it is reasonable to include age as a continuous variable. Compute the ln(odds) for fall in each age category and make a figure showing ln(odds) against age groups for women only. It is reasonable to assume that ln(odds) is depending linearly on the woman's age?
   Note: The commands

```
logit fall i.agegr
predict logoddsgr, xb
```

will compute the ln(odds) in the groups of the variable `agegr` and store these ln(odds) in the variable `logoddsgr`. See for example Day7.do for further details.

d. Compute the ln(risk) for fall in each age category and make a figure showing ln(risk) against age groups for women. It is reasonable to assume that ln(risk) is depending linearly on the persons age?
   Note: the command
   ```
   binreg fall i.agegr, rr
   ```
   will fit a binary regression model for the relative risk.
e. Repeat the analysis in d. for men.

Question 2

We will here consider a model for the relative risk and include age as a continuous variable.

a. Examine if the association between age and the risk of fall is the same for men and women. Give an interpretation of the estimated parameters in the analysis that is used to answer the question.
b. Compute an age adjusted relative risk of falling comparing men to women.

Question 3

Divide the participates into 6 balance categories: ”below 40”, ”40-41”, ”42-43”, “44-45” ”46-50” and “above 50”.

a. Estimate the risk of falling within each of the balance categories.
b. Estimate the risk difference of each of the balance categories as compared to the “above 50” group.

Question 4

A purpose of the study was also to examine if the self-evaluated balance score could replace the Berg Balance index.

Use the information on the 102 persons who had registered the Berg Balance index to describe the association between the self-evaluated balance score and the Bergs balance index.

Question 5

Among the 102 persons, 37 (36.3%) had a values below 42 for the Berg Balance index. The researchers in the study group suggested to use a cutoff of 48 for the self-evaluated balance score to predict the same group of individuals (Berg Balance index below 42), to

identify a group with bad balance (self-evaluated balance score below 48) and acceptable or god balance (self-evaluated balance score 48 or higher).

a. Among those with a low Berg Balance index estimate the proportion of persons with a self-evaluated balance score below 48 (the sensitivity of using the self-evaluated balance score).
b. Among those with a high Berg Balance index estimate the proportion of persons with a self-evaluated balance score 48 or higher (the specificity of using the self-evaluated balance score).