Aarhus, December 1st 2017
Morten Frydenberg

Applied Statistical Analysis with Missing Data

Homework to be completed before the course begins

Consider the dataset Homework2017.dta.
We will in the following look at these five persons

A) 50 year old man with BMI = 24 kg/m^2
B) 50 year old woman with BMI = 24 kg/m^2
C) 42 year old man with BMI = 31 kg/m^2
D) 50 year old woman with BMI = 31 kg/m^2
E) 40 year old man with BMI = 31 kg/m^2

Part A

Model 1:
Make a normal linear regression with systolic blood pressure as the dependent variable and sex, age and BMI as the explanatory variables.
Age should be used as is, while BMI should be entered into the model as log-transformed.
The reference person should be a man, age 45 and with BMI 24.

1. Based on **Model 1** find the expected systolic blood pressure for persons A to E.

2. Based on **Model 1** find the expected difference in systolic blood pressure between A and B, D and E, and B and E.

Model 2:
Expand model 1 to allow for an effect modification (interaction) between sex and age.

3. Find, based on **Model 2,** the expected systolic blood pressure for persons A to E.

4. Find, based on **Model 2,** the expected difference in systolic blood pressure between A and B, D and E, and B and E.

Model 3:
Return to model 1 (i.e., drop the interaction term), but now, instead of using BMI as log-transformed, you should enter it as a categorical variable with four levels BMI<20, 20≤BMI<25, 25≤BMI<30, and 30≤BMI, where you use the category 20≤BMI<25 as reference.

5. Based on **Model 3** find the expected systolic blood pressure for persons A to E.

6. Based on **Model 3** find the expected difference in systolic blood pressure between A and B, D and E, and B and E.

Part B

Model 4:
You should now consider the binary outcome, high blood pressure (systolic blood pressure ≥150mmHg).
Make a logistic regression with high blood pressure as the dependent variable and sex, age and BMI as the explanatory variables.
Age should be used as is, while BMI should again be introduced as log-transformed.
The reference person should be a man, age 45 and with BMI 24.

7. Based on **Model 4** find the risk of high blood pressure for persons A to E.

8. Based on **Model 4** estimate the odds ratio for high blood pressure comparing A and B, D and E, and B and E.

Model 5:
Expand model 4 to allow for effect modification (interaction) between sex and age.

9. Based on **Model 5** find the risk of high blood pressure for persons A to E.

10. Based on **Model 5** find the odds ratio for high blood pressure comparing A and B, D and E, and B and E.

Model 6:
Return to model 4, but now, instead of using BMI as log-transformed, you should enter it as a categorical variable with four levels BMI<20, 20≤BMI<25, 25≤BMI<30, and 30≤BMI, where you use the category 20≤BMI<25 as reference.

11. Based on **Model 6** estimate the risk of high blood pressure for persons A relative to E.

12. Based on **Model 6** find the odds ratio for high blood pressure comparing A and B, D and E, and B and E.