

May 11, 2009
Morten Frydenberg

POSTGRADUATE COURSE IN
LINEAR AND LOGISTIC REGRESSION

Day 2

Afternoon exercises:

Part A

Back to the lung function data (`lung.dta.` and `lung.sav`).

We will consider the multiple regression of `PEFR` on `height` and `sex`.

We will consider `sex` as a **categorical variable**.

Fit model with an interaction between `sex` and `height`

1. What is the estimated difference in the slopes (with CI) for men and women?
Are the slopes statistical significant?
You also made such a comparison yesterday morning.
Compare this with what you just found.
2. What is the estimated difference (with CI) in expected `PEFR` for a man and a woman both 170 cm high?
Are this difference statistical significant?
You also made such a comparison yesterday morning.
Compare this with what you just found.

Part B

Here we look at data from the lecture today (`fram200.dta.` and `fram200.sav`).

$$\ln(sbp) = \beta_0 + \beta_1 \cdot age + \beta_2 \cdot woman + \beta_3 \cdot \ln(bmi) + E$$

3. Fit a model with `sex` as a categorical variable and an interaction between `sex` and `age`, using **men** as reference for `sex`.
Spend some time trying to understand the estimates.
4. Fit the same model but now with **women** as reference.
Spend some time trying to understand the estimates.
5. Combine your findings into a conclusion on whether or not there is interaction (effect modification) between `sex` and `age`.