

November 2, 2010
Morten Frydenberg

POSTGRADUATE COURSE IN
LINEAR AND LOGISTIC REGRESSION

Day 1

Morning exercises:

You will start by considering the data set used at the first lecture, `lung.dta.` and `lung.sav`

Apart from PEFR in l/min and `height` in cm the data also contains `sex` (1=female; 2=male) and vital capacity, `VC`, in l.

First you **only look at the women**. (in Stata this is done by “`if sex==1`”).

1. Make a scatter plot of PEFR versus `height`.
2. Fit the simple linear regression of PEFR on `height`.
3. What is the estimated slope (with CI) and the interpretation of this estimate.
4. What is the standard deviation around the line (with CI) and the interpretation of this estimate.
5. Estimate (with CI) the mean PEFR for a woman with height 170 cm.

Then you **only look at the men**. (in Stata this is done by “`if sex==2`”).

6. Make a scatter plot of PEFR versus `height`.
7. Fit the simple linear regression of PEFR on `height`.
8. What is the estimated slope (with CI) and the interpretation of this estimate.
9. What is the standard deviation around the line (with CI) and the interpretation of this estimate.
10. Estimate (with CI) the mean PEFR for a man with height 170 cm.

Finally you **compare** what you have found:

11. Make an informal comparison of the variation around the line for the two sexes.
12. Compare the estimated slopes and the confidence intervals.

An approximate comparison of two parameters can be made by using the fact that the standard error of the difference between two independent estimates is given as:

$$se(est_1 - est_2) = \sqrt{se(est_1)^2 + se(est_2)^2}$$

and an approximate 95% confidence interval for the difference by :

$$(est_1 - est_2) \pm 1.96 \cdot se(est_1 - est_2)$$

13. Use this to make a formal inference concerning the difference in the slope for men and women.
14. Make an informal and formal comparison of the mean PEF_R for a man and a woman both of height 170 cm.

We will now turn to the relationship between PEF_R and VC.

15. Make a scatter plot of PEF_R versus VC .
Comment on the plot.
16. Fit the simple linear regression of PEF_R on VC .
Write down the central information found in the output.
17. What is the expected difference (with CI) in PEF_R for two persons who differ 0.5 liter in VC.