MRes in Medical Statistics
MMB8028

School of Mathematics and Statistics

Practical session on regression and correlation

1. Open the worksheet HBHT.MTW containing the data on haemoglobin and height. Plot Hb against height. The regression of Hb on height can be found using commands under the **Stat** menu.

Does the result of the regression make sense?

Perform the regression separately within each sex. What are the general lessons here for regression analysis?

[Note: see below for ways of separating male and female data]

2. Open the worksheet GRAVES.MTW. This contains data on the ophthalmic index OI and age in years (the sex of the patient can be ignored for this exercise) of patients with Graves' Disease. Calculate the $\log_{10}$ OI and then regress it on age. What is the predicted $\log_{10}$ OI at age 50? Within what limits does $\log_{10}$ OI lie for a 50-year-old? What is wrong with regressing OI on age?

3. The worksheet ANSC.MTW comprises four sets of pairs of point (x,y). Find the correlation between y and x for each of the four sets of data; use the **Basic Statistics** part of the **Stat** menu. Now plot the data. Comment.

Separating males and females.

In HBHT.MTW the data for males and females are distinguished by the value in the column labelled Sex (1=female, 2=male). If you want to perform separate regressions for the different sexes then it might be convenient to construct columns for Hb and height that comprise only males or only females. To do this you can cut and paste in the Data Window in the usual Windows manner. Alternatively, you can use the Columns to Columns… item on the Copy part of the **Data** menu. If yo do this, you will need to enter the columns appropriately and also click on the Subset the Data… button in the Copy Columns to Columns dialogue box.