STAT2008/4038/6038 Regression Modelling 31/3/2017 (1)

start with a little exploratory data analysis

VI. Summary

So, what have we learned so far? The general procedure for examining the relationship between two variables proceeds as follows:

1. Examine a plot of the data to see if a linear association seems a plausible explanation of the scatterplot;
2. Using scientific or other background information, transform the data appropriately;
3. Fit a linear regression model to the two variables using least-squares estimates;
4. Test the significance of the regression and make required predictions;
5. Use residual plots and normal q-q plots to examine the plausibility of the basic assumptions of the model;
6. If necessary (based on the plots of step 5) transform the data again (or move to a more complicated model structure);
7. Re-fit the regression and again examine the residuals thoroughly;
8. When satisfactory residual analyses have been reached, re-test and re-predict as required (remembering to transform back to appropriate scales if necessary).

A
B
C

move this

We will now take up the point mentioned parenthetically in step 6, and examine more complicated, though still linear, models, where we will allow for the possibility of more than one predictor variable.

we cycle through A, B, C until we are happy with the residual plots ie we have an "appropriate" model & then after 6, 7 we finally do 4 & look at firstly at the ANOVA table, if F test okay "adequate" model then we look at the summary output