620-371: Linear Models

Practice Class 7

21st April, 2009

In this practice class, we shall analyse a large(ish) dataset. Go to the 'Datasets' section of the 620-371 website and download the 'sleep' dataset. This dataset contains (among other things) data on the body weight (kg) and brain weight (g) of 62 mammals. Use the following function to read the data:

```
mammals <- read.table("sleep.txt", header=TRUE)
```

This creates a data frame, `mammals`, with components (among others) named `BodyWt` and `BrainWt`.

1. This data needs a logarithmic transformation. Apply `log` to both `BodyWt` and `BrainWt`. You will have to logarithmise each component individually as the entire data frame contains some non-numeric entries, which cannot be log'ed.
2. We want to fit a linear model explaining brain weight from body weight. Fit a `lm` model to the data. Display the summary of it.
3. Create a scatter plot of the data and superimpose the fitted regression line on it.
4. Create diagnostic plots and look at them. Are the model assumptions satisfied?
5. Calculate:
   (a) The least squares estimator of the parameters, $\mathbf{b}$;
   (b) The vector of residuals, $\mathbf{e}$;
   (c) The residual sum of squares, SS_{Res};
   (d) The regression sum of squares, SS_{Reg}; and
   (e) The estimator for the variance of the errors, s^2.
6. Find a 95% prediction interval for a particular mammal weighing 50 kg.
7. Test for model adequacy — $H_0 : \boldsymbol{\beta} = 0$.
8. Test for model adequacy with a possibly nonzero intercept, $H_0 : \beta_1 = 0$, using the `anova` function.
9. Test the hypothesis $H_0 : \beta_0 = 0$, using the `anova` function.