Laboratory C: Regression – Minitab

NAME:
Student Number:

This sheet should be completed and handed in at the end of the tutorial class.

Brain and body weights of mammals
Is there a relationship between the brain weights and the body weights of mammals, and, if so, what type of relationship? Data that can help answer this question were collected as part of a large study into sleep in mammals, and are found in Weisberg, S. 1985, *Applied Linear Regression*, 2nd edn, John Wiley & Sons. They show the average body weights (kg) and brain weights (g) for 62 mammals.

C1. Retrieve the data from the worksheet **mammals.mtw**. Explain why it would be reasonable to use brain weight as the response and body weight as the predictor variable. Obtain a graph of brain weight versus body weight. What problem is apparent?

C2. For data, such as these, that are spread over several orders of magnitude, a useful transformation is the logarithm. Use **Calc > Calculator** to get the natural logarithms of body weights, *Logbody*, with the expression **loge(c2)**. Likewise, calculate *Logbrain*, the natural logarithms of the brain weights. Obtain a graph of *Logbrain* versus *Logbody*. Does a linear relationship seem reasonable?

C3. Use **Stat > Regression > Regression** to obtain the least-squares regression line of *Logbrain* on *Logbody*. Select **Storage** and store fitted values (or fits) and standardised residuals. Write down the equation of the regression line, the standard errors of the coefficients and the values of s and R^2.

C4. Use **Stat > Regression > Fitted Line Plot** to see the results of the regression graphically. Do you think that log brain weights can be successfully predicted from log body weights? Why or why not? Can you use the regression equation to write down the equation linking brain weight and body weight, rather than their logs?

C5. Minitab has created two new columns, *SRES1* and *FITS1*. Standardised residuals are residuals which have been scaled to have a standard deviation of 1. (They already have a mean of 0.) Obtain residual plots with **Stat > Regression > Residual Plots**, entering *SRES1* and *FITS1* in the boxes. The residuals can be checked for normality using the normal probability plot (and the histogram). Do they seem to be normally distributed? The I-chart shows the residuals in order, and can sometimes help to pick up cyclic or other patterns. The residual plot shows (standardised) residuals versus fitted values. Which animals have the highest and lowest standardised residual? Use brushing to find out. Are you surprised by the answers?

C6. Obtain 95% confidence intervals for the intercept and the slope of the regression line. Minitab doesn't do this automatically; you have to use the formula *estimate* ? t^* *times standard error*. You can find the appropriate value of t_8 from **Calc > Probability Distributions > t**, selecting **Inverse cumulative probability**, degrees of freedom from the error line of the ANOVA table, and 0.975 as input constant.

C7. A specialist in the biology of mammals claims that brain weight is proportional to body weight raised to the power ¾. Does this seem reasonable on the basis of this set of data? Carry out a hypothesis test to answer the question. What will be your null and alternative hypotheses?

C8. Primates are the highest order of mammals and include man, apes, monkeys, tarsiers, lemurs and tree shrews. To investigate the brain-body relationship for primates, we must first create a variable that shows which animals they are. Label a column *Primate*, and enter the value 0 if the animal is not a primate and 1 if it is. There are eight primates in the data. (Check with **Stat > Tables > Tally** that you have the right number.) A variable such as *Primate* is called an *indicator variable*. Can you explain why?

C9. Is the relationship between brain weight and body weight significantly different for primates than for other mammals? Answer this question by fitting a multiple linear regression model with *Logbrain* as response and *Logbody* and *Primate* as predictors. Minitab will do the calculations with two predictors as easily as with one, and the overall form of the output is very similar. Carry out a hypothesis test to see whether the coefficient of *Primate* could be taken to be zero. Why are we interested in this particular value? What is the result of the test, and what can we conclude from this?

C10. The smallest primate is the pen-tailed tree shrew from Malaysia, which has an average body weight of only 45g (that is 0.045kg). We want to find an estimate and a 95% prediction interval for its brain weight. Run the multiple regression again, but this time select **Options** and enter the values of the predictors into the **Prediction intervals** box. You will need to enter –3.101, the natural log of the body weight, and then a 1, since the animal is a primate. Minitab will give you an estimate (labelled 'FIT') and a 95% prediction interval (labelled '95%PI') for the natural log of the brain weight. How will you get these back to the original scale? What are your final values?

New Minitab commands and subcommands used in this laboratory class:
Stat > Regression > Regression
Stat > Regression > Fitted Line Plot
Stat > Regression > Residual Plots