Wednesday, Week 6 –

Galapagos Regression Demo

1. Galapagos data downloaded from Aaron Ellison’s Site, imported to Excel.
2. What’s the scientific hypothesis?
3. What’s H0 and Ha?
4. In JMP, Fit Y by X for Island Area and Number of Species. What’s wrong?
5. Transform the data – Log10. Fit Y by X for Log10 Island Area and Log10 Number of Species. Better? Examine results….
6. Confidence Intervals (Fig. 9.4) in JMP
7. Assumptions of Least Squares Regression (p. 257)
   1. Model correctly describes functional relationship between X and Y
   2. X-variable measured without error
   3. For any X, Y value(s) are independent with normally distributed errors.
   4. Variances are constant along the regression line.
8. Diagnostic Tests – Plotting Residuals (differences between Yi’s and Y ̂’s). See Fig 9.5 in book – Residual Patterns.
   1. Expected distribution of residuals for a linear model with a normal distribution of errors. Tells you if the data are well fit by the linear model.
   2. For other patterns, see Figure 9.5.
   3. Q-Q plots <http://www.youtube.com/watch?v=X9_ISJ0YpGw>
9. Monte Carlo Analysis
   1. Demo
   2. To illustrate the p-value you get, use histogram of least squares slopes of randomly shuffled Xs and Ys – see also Fig. 9.8.