

Name: _____

MAT 222

Spring 2019

Minitab Lab 1

“When life gives you lemons, make lemonade. I read that on a can of lemonade but I like to think it applies to life.”

–Andy Dwyer, Parks and Recreation

The following exercises make use of the `abalone.mpj` file. This data is from a study by Dr. Sam Waugh’s Ph.D. thesis at the University of Tasmania. The goal of the study was to study whether one could predict the age of abalone using physical measurements. The variables in the study were sex (male, female, and infant), length (mm), diameter (mm), height (mm), whole weight (g), shucked weight (g), viscera weight (g), shell weight (g), and rings (each ring is 1.5 years). Use Minitab to complete the following exercises.

Problem 1: Give a boxplot of the variable ‘whole weight’. Does the distribution seem normal? Are there any outliers?

Problem 2: Find the 5-number summary for the variable ‘whole weight’. Find the mean and standard deviation for this variable.

Problem 3: Give a histogram of the variable ‘diameter’. Does is the distribution normal? If not, is the distribution left skewed, right skewed, or neither?

Problem 4: Give a scatterplot of the variable ‘rings’ (on the x -axis) and ‘whole weight’ (y -axis). Describe any trends you see. Does the distribution seem linear?

Problem 5: Find the equation of the linear regression for the variable ‘rings’ (on the x -variable) and ‘whole weight’ (y -variable). What is the R^2 value? Using the R^2 value, discuss the linearity of the distribution.

Problem 6: Find *exactly* how many total female, male, and infant (I) abalones were examined without using a plot. Give a barchart for the variable ‘sex’. Did the study examine approximately equal numbers of each type? Give a pie chart for the variable ‘sex’. Using the pie chart, discuss whether equal numbers of females, males, and infant abalones were examined.

Problem 7: Assume that the true standard deviation for the variable ‘height’ is $\sigma = 0.04183$. Find a 99% confidence interval for the mean of the variable ‘height’. Perform a hypothesis test, using $\alpha = 0.01$, for the variable ‘height’ with $H_0 : \mu = 0.1385$ and $H_a : \mu > 0.1385$. State your conclusions verbally in a complete sentence. Be sure to give the test statistic and p -value for your test.

Problem 8: Find the tally of the ‘variable rings’. How many abalone had 12 rings? Using the summarized data in a one-sample proportion test, find a 99% confidence interval for the proportion of abalone with 12 rings. Using this confidence interval, does it seem likely that 4% of abalone have 12 rings?

Problem 9: Find a histogram for the variable ‘shell weight’ with a normal fit. Does the distribution seem normal from this plot? Run a normality test on the variable ‘shell weight’, using this normality plot, is the variable ‘shell weight’ normally distributed?

C L Blake, C J Merz. UCI repository of machine learning databases University of California, Irvine, Department of Information and Computer Sciences. 1998, Sam Waugh (1995) "Extending and benchmarking Cascade-Correlation", PhD thesis, Computer Science Department, University of Tasmania.