STA2023 – Test 1 Partial List of Topics

Many of these topics will be included on the exam, but not necessarily all of them. Furthermore, other questions may be asked that are not included in this review. In addition to this review, look over the in-class/website handouts and the assigned problems from the textbook.

Most of the questions involving computations or calculator work will be given as applications problems involving real data. In other words, the questions will involve real-world situations similar to the textbook problems. You are encouraged to bring a graphing calculator with you to the exam. In fact, you probably won’t have enough time to complete the exam if you do not use the calculator when prompted.

1. When presented with data…
   A. Know the difference between Categorical and Quantitative variables.
   B. Sketch a relative frequency bar graph, frequency bar graph, pie chart, dot plot, frequency histogram, relative frequency histogram, or stemplot of one variable data.
   C. Be able to create a frequency table or a relative frequency table.
   D. Write a paragraph to interpret the histograms/stemplot: shape, mode, outliers, center, spread/range.
   E. Understand what class width is in both a histogram and stemplot.
   F. Use the mean and the median to help decide if this data is symmetric, left-skewed, or right-skewed.
   G. Be able to change a frequency histogram or frequency table into a relative frequency histogram or relative frequency table.
   H. Find the mean, standard deviation and variance of a set of data.
   I. Understand that a mean can be described as an average and vice versa.
   J. Understand that a proportion can be described as a percentage and vice versa.
   K. Find the 5 number summary for a set of data (find MinX, Q1, Median, Q3, MaxX)
   L. Interpret (write a sentence explaining the word in the context of the problem) each of the numbers in the 5-number summary:
   M. Recognize which is more appropriate to describe a data set: the mean and standard deviation or the five number summary and explain your decision in a sentence.
   N. Using the 1.5*IQR criterion, does this data contain any outliers? (calculate the lower and upper fences and show your work)
   O. Sketch a boxplot of data.
   P. Know which statistics are affected the most by outliers.

**Know how to use your calculator to help you with anything mentioned above.**

Continued on next page…
2. Given the following side-by-side boxplot of …
   A. Write a paragraph (complete sentences) describing each distribution. For example, describe the meaning of the 25\textsuperscript{th}, 50\textsuperscript{th}, and 75\textsuperscript{th} percentiles. Also, discuss the symmetry or skewness of each group (if applicable) and the meaning of any * that may appear on the boxplots.
   B. Compare/contrast the boxplots and state any appropriate conclusions about the similarities and differences between the two groups. Especially comment on the spread and range of the two groups, stating with or without outliers. IQR comparisons can also be mentioned. Be sure to point out any striking comparisons which may reveal information in the overall story problem situation.

3. Survey and Study Design.
   A. Understand that a sample is a subset of a given population.
   B. Know the definitions: population, sample, simple random sample.
   C. Know what SRS stands for.
   D. Know the three basic principles of Experimental Design and what they mean: Control, Randomization and Replication.
   E. Know the difference between an experiment and observational study.
   F. Understand what experimental units are.
   G. Know control groups, placebos, placebo effect, blinding, double-blind, double dummy.
   H. Know sampling methods: Random, Stratified, Cluster, Systematic, Multistage.
   I. Understand what Bias is.
   J. Know what a lurking/confounding variable is.
   K. Know and understand the problems associated with surveys and sampling: nonresponse, response bias, undercoverage, volunteer response, convenience sampling.