

STAT 113: EXAM 1 STUDY GUIDE

COLIN REIMER DAWSON, FALL 2019

(1) Chapter 1

1.1 Identifying Cases and Variables. You should be able to

- Recognize what the cases/observational units are in a dataset.
- Recognize what the variables are.
- Distinguish variables based on whether they are categorical vs. quantitative, and
- explanatory vs. response.

1.2 Sampling. You should be able to

- Recognize potential source of bias in sampling procedures.
- Recognize that completely random sampling is the gold standard, but not usually achievable in practice.
- Identify the appropriate population that sample information can be generalized to.

1.3 Study Design. You should be able to

- Distinguish between observational and experimental designs.
- Identify possible confounding variables in observational studies, and explain why they meet the definition of confounding variables.
- Distinguish between mere association and association due to causation.
- Come up with multiple possible explanations for associations.

(2) Chapter 2

2.1 Categorical Data. You should be able to

Date: October 7, 2019.

- Create (on paper) and interpret one and two-way frequency and relative frequency tables
- Recognize when the principles in what I called “the cardinal rule of bar charts” are violated and why it is misleading
- Calculate and interpret conditional (within-context) proportions, recognize when it makes sense to condition, and make comparisons between conditional proportions to examine differences between contexts.

2.2 Shape and Center. You should be able to

- Calculate and interpret means and medians, recognize their properties (e.g., robustness/resistance), and decide when it is appropriate/useful to describe a dataset with each.
- Draw and interpret dot plots and histograms.
- Recognize right- and left-skewed distributions, identify the impact of skew on the mean and median, and identify possible outliers.

2.3-2.4 Variability.

A. Quantile-based measures. You should be able to

- Recognize problems with using range to measure variability.
- Interpret range and interquartile range.
- Understand what quartiles and percentiles mean, and recognize correspondence between quartiles and median/min/max.
- Draw a box plot from the five number summary
- Give the five number summary, range/interquartile range, from a box plot
- Recognize skewness/outliers from a box plot

B. Metric-based measures. You should be able to

- Understand standard deviation as “typical distance from the mean”.
- Use the 68/95 rule with bell-shaped distributions
- Estimate the standard deviation from a density plot or histogram if the distribution is bell-shaped.

- C. Know when to choose what measure of variability (e.g., which are resistant/robust, and when is that important)
- D. Use histograms/boxplots/summary statistics to compare distributions of a quantitative variable for different levels of a categorical variable.

2.5 Association Between Numeric Variables. You should be able to

- Interpret scatterplots.
- Recognize linear association from a scatterplot.
- Estimate correlation from a scatterplot (to a first approximation).
- Distinguish between slope of line and degree of fit to the line.

2.6 Simple Linear Regression. You should be able to

- Interpret the intercept and slope of a regression line.
- Use the equation of a regression line to make predictions about y given a particular x .
- Use a graph of a regression line to estimate the predicted value of y given a particular x .
- Recognize the meaning of a residual.
- Distinguish between good and bad linear fits by looking at a residual plots.

(3) Chapter 3

3.1 Sampling Distributions. You should be able to

- Recognize and explain the difference between a population and a sample, and between a parameter and a statistic.
- Compute point estimates of parameters using a sample.
- Understand what a sampling distribution is, what the cases are in a sampling distribution, and what the individual values in a sampling distribution represent.
- Understand what a standard error is, how it relates to standard deviation, and how it contrasts with variability of individual cases in a sample.
- Recognize where a sampling distribution of means is centered.

- Visually estimate standard error from a dot plot or histogram of a sampling distribution.