

Review Sheet for the Midterm

1. Be able to identify units of statistical analysis and variables and to determine a variable's level of measurement (nominal, ordinal, interval/ratio) based on variable description or frequency distribution.
2. Be able to construct frequency distributions (counts and percentages) and cumulative frequency distributions, both by hand and using Stata, and interpret them.
3. Be able to calculate the central tendency measures (mean, median, mode) and variability measures (variation ratio, range, interquartile range, variance, standard deviation), both by hand and using Stata. Understand the meaning of each and when it can/should be used. Know which measures of central tendency and variation can be used for what type of variable (nominal, ordinal, interval/ratio). Know how skewness and extreme values (outliers) affect different measures of central tendency.
4. Know how to construct bar graphs, pie charts, stem-and-leaf plots, histograms, and boxplots using Stata and interpret them. Be able to interpret line plots, frequency polygons, and ogives. Know which graphs can be used for each level of measurement.
5. Be able to identify the type of plot based on its appearance and to distinguish univariate plots from bivariate plots. Know what kind of problematic practices are common in graphs that appear in the media and what can make a bar graph, a pie chart, or a line plot look misleading.
6. Be able to create a histogram with a normal curve overlay and determine whether a distribution looks approximately normal or not. Understand what skewness and kurtosis are, be able to visually detect these characteristics for a given distribution graph as well as identify these based on the skewness and kurtosis statistics. Be able to obtain skewness and kurtosis statistics using Stata. Be able to determine the direction of skewness, if any, and distinguish platykurtic and leptokurtic distributions.
7. Know the key properties of the normal distribution. Understand what z-scores are and why they are called standard scores. Understand how z-scores are related to raw scores and to the normal curve. Know how to convert a value into a z-score and a z-score into a value. Be able to find proportions of the distribution above, below, and in between given z-scores. Be able to find the scores (in original metric) that cut off a certain percentage of the distribution.
8. Understand what the sampling error is; know the difference between random sampling error and systematic sampling error and know the sources of each of these two types of error.
9. Understand what a sampling distribution of the mean is. Know its mean, standard deviation, and shape (Central Limit Theorem). Understand how the shape of sampling distribution is related to the size of the sample (t-distribution and normal distribution).
10. Understand what a confidence interval is, and how our ability to construct confidence intervals is related to our knowledge of the sampling distribution of the mean. Know how to construct confidence intervals for the mean, both by hand (for large samples and small samples) and using Stata, and how to interpret them.
11. Know the difference between point estimate and interval estimate, understand the concept of confidence level (CL) and know which CL are typically used. Understand how the level of confidence is related to the probability of error (α). Know how the size of interval is related to the sample size, the level of confidence, and the degree of heterogeneity in the population.
12. Know what error bars are and how to interpret them to compare means across groups. Know the main types of error bars (confidence interval, standard error, standard deviation) and understand the differences in interpretation (and size) depending on the type. Know that error bars should not be used to compare means if the data are paired (come from the same individual or from related individuals, e.g. relatives).
13. Know what the margin of error in opinion polls is and how it is related to confidence intervals. Be able to calculate and interpret the confidence interval (and the margin of error) for a given proportion. Be able to calculate and interpret the maximum margin of error for a specified sample size.
14. Understand how margin of error differs from bias, and know the sources of each for public opinion polls. Know what pollsters do to minimize both the margin of error and bias in polls and be able to evaluate the quality of a public opinion poll. Understand the difference between probability and nonprobability samples in public opinion polls and distinguish between margins of error and credibility intervals.