

Study Guide for Exam 3

Quantitative Data Analysis (Chapter 8)

- Descriptive statistics: statistics used to describe the distribution of and relationships among variables
 - o Central tendency: the most common value (for variables measured at the nominal level) or the value around which cases tend to center (for a quantitative variable)
 - o Variability/variation: the extent to which cases are spread out through the distribution or clustered around just one value
 - o Skewness: the extent to which cases are clustered more at one or the other end of the distribution of a quantitative variable rather than in a symmetric pattern around its center. Skew can be positive (to the right) with numbers tapering in a positive direction or negative with number tapering off in negative direction
 - o Know how to interpret graphs (i.e., histogram, bar graph, and line graph)
 - o Measures for central tendency: Central tendency is summarized with one of three statistics:
 - Mode: most frequent value in a distribution; also termed probability average
 - Median: position average that divides a distribution in half (the 50th percentile)
 - Mean: The arithmetic or weighted average computed by adding up the value of all the cases and dividing by the total number of cases
 - o Measures of variation: Captures how widely and densely spread the numbers are
 - Range: the true upper limit in a distribution minus the true lower limit (or the highest rounded value minus the lowest rounded value, plus 1)
 - Interquartile range: The range in a distribution between the end of the 1st quartile and the beginning of the 3rd quartile
 - Variance: A statistic that measures the variability of a distribution as the average squared deviation of each case from the mean.
 - Standard deviation: distance from the mean that covers a clear majority of cases (about 2/3). It is the square root of the variance
 - o Which central tendency statistics and which variability statistics are appropriate under what conditions?
 - Mode: Most appropriate for nominal data. Can easily give a misleading impression of a distribution's central tendency. Problems occur with bimodal distributions. (2 categories with roughly equal numbers of cases and clearly more cases than the other categories) However, it can be used to characterize central tendency of variables at the nominal level. Because it is the most probable value, it can be used to answer questions such as which ethnic group is most common in a given school.
 - Median: For ordered data. Not appropriate for variables that are measured at the nominal level; their values cannot be put in order, so there is no meaningful middle position.
 - Mean: For interval and ratio data. Only makes sense if the values of the cases can be treated as actual quantities. That is if they reflect an interval or ratio level of measurement or if we assume that an ordinal measure can be treated as an interval. Should not measure qualitative variables such as religion.
 - Range: Not a good summary measure since outliers can alter it drastically. It is to identify the whole range of possible values that might be encountered.
 - Interquartile range: avoids problem outliers and shows the range where most cases lie.
 - Variance: mainly used to compute standard deviation. Conventionally used to measure variability with the closely related SD than with the variance.

- Standard deviation: preferred measure of variability particularly when a variable is normally distributed. It can tell you quickly about how wide the variation is of any set of cases or the range in which most cases will fall.
- The normal curve: symmetric distribution shaped like a bell and centered around the population mean.
- Inferential statistics and hypothesis testing
 - o Inferential stats: Estimate the degree of confidence that can be placed in generalizations from a sample to the population
 - o Null hypothesis: Statement of no association or no difference
 - Ex: There is no difference between male and female students on number of Facebook friends.
 - o Alternative hypothesis: Statement that there is an association between variables or there is a difference between groups
 - Ex1: Male and female students differ on the number of Facebook friends they have. (nondirectional)
 - Ex2: Male students have more Facebook friends than female students. (directional)
 - o Statistical significance—what does the p-value mean? The mathematical likelihood that an association is not due to chance, judged by a criterion the analyst sets. Basically concludes a relationship exists. (Often the probability is less than 5 out of 100 or $p < .05$)
 - If the p-value is less than .05 ($p < .05$) → the possibility of the null hypothesis being true is very small
 - We **reject the null hypothesis** (and accept the alternative hypothesis)
 - We conclude that the association or group difference found is **statistically significant**
- Association between variables
 - o Chi-square: an inferential statistic used to test hypotheses about relationships between two or more variables in cross tabulation. Customarily reported in a summary form such as $p < .05$, which can be translated as probability that the association was due to chance is less than 5 out of 100. Association between variables measured at *nominal* or *ordinal* level
 - Example: Is gender associated with frequency of Facebook profiles update?
 - Null hypothesis: no association
 - Alternative hypothesis: there is an association; or females updates their Facebook profiles more frequently than males
 - o Correlation:
 - Association between variables measured at *interval* or *ratio* level
 - Correlation between two variables ranges from -1 to +1
 - Close to +1 or -1 → strong correlation
 - Close to 0 → no correlation
 - Is extraversion related to number of Facebook photos?
 - The correlation analysis shows: $r = .37, p = .04$
 - We reject the null hypothesis and conclude that students who are more extraverted have more Facebook friends
- Comparing groups (t test)
 - o Difference in means between **two** groups
 - o t: the ratio of difference between the group means and the variability within groups
 - o Do males and females differ in number of facebook photos?
 - o Null hypothesis: Males and females do not differ
 - o Alternative hypothesis: Males and females differ in number of facebook photos

- o Females have more facebook photos than males
- o Results of t test indicate that $t = 3.15, p = .004$
- o Thus, we **reject the null hypothesis** and conclude that difference between males and females are **statistically significant** - females have more facebook photos than males

Qualitative Data Analysis (Chapters 9 & 10; class notes)

- Common features of qualitative methods:
 - o Exploratory research question about what people think and how they act and why in some social setting.
 - o Designs focus on previously unstudied processes and unanticipated phenomena
 - o Have an orientation to social context to the interconnections between social phenomena rather than to their discrete features.
 - o Focus on human subjectivity on the meanings that participants attach to events and that people give to their lives.
 - o Sensitivity to the subjective role of the researcher. Keep track of their own actions in and to that social process.
- Analyzing qualitative data
 - o What's distinctive about qualitative data analysis? Focus on text rather than numbers
 - o Methods:
 - Participant observation
 - Intensive interviewing
 - Focus groups
 - Written text
- Approaches in qualitative research
 - o Phenomenology: To describe or understand experiences or phenomenon within a given context
 - Rich thematic description of the lived experience
 - o Grounded theory: systematic theory developed inductively based on observations that are summarized into conceptual categories, reevaluated in the research setting and gradually refined and linked to other conceptual categories
 - o Ethnography: study and systematic recording of human cultures
 - Seeks to describe and understand the richness and detail of natural social world
 - Data are usually collected through participant observation and intensive interview
 - Avoid **ethnocentrism**
 - What is emic perspective, and what is etic perspective?
 - Emic: representing a setting in terms of the participants
 - Etic: representing a setting in terms that the researcher brings to the study
 - What do educational ethnographers study? Studying school/classroom culture
- Differences between qualitative and quantitative methods
 - o Qualitative focuses on meanings rather than quantifiable phenomena
 - o Collection of data on a few cases rather than little data on many cases
 - o Study in depth detail without predetermined categories or directions rather than emphasis on analyses and categories determined in advance
 - o Conception of the researcher as an instrument rather than as the designer of objective instruments to measure particular variables
 - o Sensitivity to context rather than seeking universal generalizations