Applied Statistics I

FINAL EXAM
Topics to Review

Basic Computing Skills

- Getting data into SPSS
  - Entering data manually
  - Loading an SPSS data file
  - Identifying types of variables (e.g., scale, nominal, categorical)
- Obtaining descriptive statistics and plots using SPSS (including histograms and boxplots).
- Using statistical procedures for inference (test and confidence intervals) and correlation and regression.

Data Production

- Data basics:
  - Individuals (subjects, cases)
  - Variables: categorical versus quantitative variables
- Methods:
  - Observational studies and experiments.
  - Confounding variables
  - Sampling
    - Terminology: sample, population, sampling frame
    - Difference between probability and non-probability samples
    - Methods
      - Voluntary samples
      - Convenience samples
      - Simple random samples
      - Stratified samples
    - Types of Bias
    - Drawing random samples with the Randomizer
  - Experiments
    - Terminology:
      - Factors (and factor levels), treatments, and response variables.
    - Designs
      - Completely randomized designs
      - Block designs
      - Matched pair designs

Data Analysis

- Basic approach
  - Start with graphs, look for:
    - Pattern: center, spread, shape
    - Deviations: outliers
  - Add numerical descriptions based on what you see.
- Graphs: stemplot, histogram, boxplot, and modified boxplot.
- Comparative plots: side-by-side boxplots and back-to-back stemplots.
- Carry out and present a data analysis using the 4-step method: state, formulate, solve, and conclude.
- Numerical descriptions:
  - Resistant: median, quartiles, IQR
  - Non-resistant: $\bar{x}$ and s
  - Five-number summary
- Identification of outliers using the $1.5 \times$ IQR rule.
- Application and interpretation of the 68-95-99.7 rule.
- Characteristics of normal distributions.
- Standardizing values from normal distributions (e.g., IQ scores).
- Finding areas under the normal curve using the Rossman-Chance applet.
- Carrying out backwards normal calculations using the Rossman-Chance applet.

Probability

- Density curves
- Relative frequency definition of probability

Sampling Distributions

- Parameters and statistics
- Sampling distribution of $\bar{X}$ under various assumptions
- The Central Limit Theorem

Theory of Confidence Intervals and Significance Tests

- Conditions for inference
- Properties of confidence intervals
  - Role of standard deviation, level of confidence and sample size
  - Interpretation of *confidence* and a confidence interval
- Significance test concepts:
  - Null and alternative hypotheses
  - *p*-value and level of significance (α)
  - Interpretation of the *p*-value in context
  - One- and two-sided tests
  - Practical vs statistical significance
  - Duality of significance tests and confidence intervals
- Sample size for desired margin of error *m*

Confidence Intervals and Significance Tests for Means and Proportions

- Four-step method approach to inference.
- Identifying study design: one sample, paired, or two-sample design
- Recognition of parameter or parameters about which inference is required (e.g., means, proportions, or differences in means or proportion)
- Developing appropriate hypotheses to address a research question
- Discerning between one- and two-sided alternative hypotheses
- Computation of one and two-sample *t*-tests and confidence intervals for means
- Computation of one- and two-sample *z*-tests for proportions.
- Computation of one- and two-sample plus-four confidence interval for proportions.
- Know and be able to check conditions for use of procedures.
- Interpretation of test results and confidence intervals
- Sample size for desired margin of error *m* (one-sample proportions only)

Scatterplots and Correlation

- Making and interpreting scatterplots (manually and using the calculator)
- Computing and interpreting the correlation coefficient *r*.
- Properties of the correlation coefficient

Regression

- Finding the equation for the least squares regression line.
- Carefully interpreting the regression slope in the context of the data.
- Proportion of variability explained in a regression (r^2).
- Finding fitted values and residuals.

STUDY HELPS

- End of section and end of chapter summaries.
- "Statistical Thinking Revisited" on pp. 657-659.

SCOPE

- Chapters 1-5, 7 (review of Chapters 1-5), 8-11, 14-22 (omitting most optional sections consistent with coverage in class)