

**Although you are responsible for all material taught in Statistics 141/151 here is checklist of the 141/151 material that I assume you already know going into Stat 252.**

Introduction

- a) Population vs. Sample
- b) Parameter vs. Statistic
- c) Descriptive vs. Inferential statistics

Sample Data

- a) Categorical vs. Numerical data
- b) Discrete vs. Continuous numerical data
- c) Observational studies vs. Randomized experiments
- d) Confounding variables
- e) Sampling techniques (simple random sample)
- f) Types of sampling bias (response, selection, non-response) and their effects.

Graphical Methods

- a) Frequency, Relative frequency (distribution)
- b) Histogram, Boxplot, Scatterplot
- c) Describing these plots (shape, skew, center, spread, outliers, etc.)

Numerical Measures

- a) Mean, Median
- b) Proportion
- c) Standard deviation, Variance
- d) Quartiles, Interquartile range, outliers
- e) Z-scores
- f) Percentiles
- g) The Empirical Rule (or 68-95-99.7 Rule)

Probability and Sampling Distributions

- a) What is a probability?
- b) Population Distribution
- c) Discrete vs. Continuous variables
- d) Density curve
- e) Normal distribution
- f) Standardized variable, z-score
- g) Standard normal distribution
- h) Sampling distributions
- i) Sampling variability
- j) Central Limit Theorem
- k) Sampling distribution of the sample mean
- l) Sampling distribution of the sample proportion

One Sample Inference for the Population Mean

- a) The t-distribution
- b) Confidence Intervals for the Population Mean
  - Point estimate
  - Confidence level and critical values
  - Standard error
  - Sample size calculation

c) Hypothesis Testing for the Population Mean

- Null and alternative hypothesis
- 1-sided and 2-sided tests
- Type I error, Type II error
- Level of significance
- Test statistics
- P-value

Two Sample Inference for the Difference in Population Means

- a) Confidence Intervals for the difference in population means: paired samples, independent samples.
- b) Hypothesis testing for the difference in population means: paired samples, independent samples.

Simple Linear Regression

- a) Scatterplots
- b) Pearson's correlation
- c) Fitting a line: Least squares method
- d) Coefficient of determination.
- e) Assumptions of simple linear regression (known to some extent)

ANOVA

- a) The F-test
- b) Assumptions
- c) Sums of squares (also known as sum-of-squares residuals)
- d) Degrees of freedom for different sources of variation
- e) F distribution
- f) ANOVA tables