Math 1240: Statistical Analysis I

Credit hours: **4 credits**
Prerequisites: **Placement in ACCUPLACER Grid 5 or MATH 1200 with a grade of C or better or MATH 1179 with a grade of C or better**

Course Description:

An introduction to elementary statistics, this course covers methods used in the collection, presentation, analysis, and interpretation of data. Topics include frequency distributions, measures of central tendency and dispersion and sampling, with emphasis on estimation and hypothesis testing.

Course Objectives:

1. Form a firm foundation in the basics of sampling and data presentation.
2. Develop rigorous descriptive and inferential methods to analyze and interpret data.
3. Employ fundamental concepts of probability within the context of data sets and probability experiments.

Learning Outcomes:

1. Identify the similarities, differences, advantages, and disadvantages between samples and populations, sampling methods, levels of data, measures of center, measures of dispersion, discrete and continuous variables, and measuring scales.

2. Utilize stem and leaf plots, frequency distributions and class-based histograms, and relative and cumulative frequency tables to present data appropriately to infer shape, center, dispersion, skewness, and the presence of outliers by observation.

3. Employ the concepts of probability including basic terms (e.g. outcome, sample space), elementary rules (e.g. conditional probability, complement), empirical and classical probabilities, and counting techniques. Calculate probabilities, expected values, and explain the mean and variance for probability distributions of discrete random variables (e.g. Binomial and Poisson distributions).

4. Translate back and forth from the distribution of continuous random variables to the standard normal distribution to determine percentiles, outliers, and event proportions and probabilities.

5. Investigate the distribution of sample statistics and determine and interpret confidence intervals for population parameters via the standard normal curve, t-distribution and the χ^2 – distribution.

6. Learn and apply the language of hypothesis testing (e.g. type I and type II errors, null and alternative hypotheses). Perform hypothesis tests involving one population parameter, hypothesis tests involving two populations (e.g. mean for independent samples, mean for dependent samples, goodness-of-fit tests, tests for independence) and interpret results.

7. Use simple linear regression analysis to discover trends within scatter plots and predict outcomes with the least-squares regression line subject to a standard error.

Course Topics:

I. GENERAL CONCEPTS
- A. Populations and samples
  1. Census vs. sampling study
  2. Sampling methods
  3. Types of bias
- B. Variables
  1. Qualitative vs. quantitative
  2. Discrete vs. continuous
- C. Levels of data
  1. Nominal
  2. Ordinal
  3. Interval
  4. Ratio

II. PRESENTATION OF DATA
- A. Stem and leaf plot
- B. Boxplot
- C. Other presentations (e.g. pie chart) and misrepresentations
- D. Classes
- E. Frequency distributions and histograms
- F. Relative frequency, cumulative and relative cumulative frequency distributions
- G. Shape, center, dispersion, skewness, kurtosis and the presence of outliers via observation only

III. DESCRIPTIVE STATISTICS
- A. Measures of central tendency
  1. Mean
  2. Median
  3. Mode
  4. Midrange
  5. Advantages and disadvantages of each measure of center
  6. Weighted means
- B. Measures of dispersion
  1. Range
  2. Variance
  3. Standard deviation
- C. Distribution of data
  1. Chebyshev's Inequality
  2. Percentiles
  3. Normal distributions and the Empirical Rule
  4. Z-scores
  5. Identifying outliers numerically
- D. Grouped data
  1. Measures of center
  2. Measures of dispersion

IV. PROBABILITY
- A. Vocabulary
  1. Experiment
  2. Trial
  3. Outcome
  4. Sample space
  5. Event
- B. Types of probability

1. Empirical
2. Classical
3. Subjective

C. Basic and general rules of probability
D. Conditional probability
E. Counting techniques
    1. Fundamental counting principles
    2. Permutations
    3. Combinations
    4. Probabilities involving counting techniques
F. Probability trees
G. Contingency tables

V. DISTRIBUTIONS OF DISCRETE RANDOM VARIABLES
A. Probability distributions
    1. Graphs
    2. Tables
    3. Outliers
B. Discrete random variables
    1. Expected value (mean)
    2. Variance
    3. Standard deviation
    4. Computations including “at most” and “at least” probabilities
C. Binomial experiments
    1. Expected value (mean)
    2. Variance
    3. Standard deviation
    4. Computations including “at most” and “at least” probabilities
D. Poisson experiments
    1. Expected value (mean)
    2. Variance
    3. Standard deviation
    4. Computations including “at most” and “at least” probabilities

VI. DISTRIBUTIONS OF CONTINUOUS RANDOM VARIABLES
A. Introduction to continuous distributions
B. Normal distributions and the bell curve
    1. Properties
    2. Graphs
    3. Standard normal density curve table
    4. Percentiles and z-scores
    5. Transformations to and from z-scores
    6. Computations including “at most” and “at least” probabilities
    7. The normal approximation to the binomial probability distribution
C. Normal probability plots
D. Determination of outliers

VII. DISTRIBUTION OF THE SAMPLE MEAN AND SAMPLE PROPORTION
A. Sampling distributions
B. Distribution of the sample mean
    1. Standard error of the mean
    2. Central Limit Theorem
    3. Conditions for normality
C. Distribution of the sample proportion
    1. Standard deviation
    2. Conditions for normality

VIII. CONFIDENCE INTERVALS

- A. Point estimates
- B. Margin of error factors
- C. Level of significance and level of confidence
- D. Estimating a population proportion
    1. Verification of normality
    2. Constructing a confidence interval
    3. Interpretation
    4. Sample size necessary for a specified error
- E. Estimating a population mean
    1. Student's t-distribution
        a. Properties
        b. Degrees of freedom
        c. Table
    2. Constructing a confidence interval
    3. Interpretation
    4. Sample size necessary for a specified error
- F. Estimating a population variance or standard deviation
    1. Chi-square distribution
        a. Properties
        b. Degrees of freedom
        c. Table
    2. Constructing a confidence interval
    3. Interpretation

IX. HYPOTHESIS TESTING: ONE POPULATION

- A. Null hypothesis
- B. Alternative hypothesis
- C. Type I error
- D. Type II error
- E. P-value
- F. Statistical significance vs. practical significance
- G. Testing a population proportion
    1. Notation
    2. Critical values
    3. Classical approach vs. p-value approach
    4. Interpretation
- H. Testing a population mean
    1. Notation
    2. Critical values
    3. Classical approach vs. p-value approach
    4. Interpretation
- I. Testing a population standard deviation
    1. Notation
    2. Critical values
    3. Classical approach vs. p-value approach
    4. Interpretation

X. HYPOTHESIS TESTING: TWO POPULATIONS

- A. Inference about two means
    1. Dependent samples
    2. Independent samples
        a. Welch's approximate t

XI. GOODNESS-OF-FIT TESTS AND CHI-SQUARE INDEPENDENCE TESTS

A. Expected counts vs. observed counts
B. Chi-Square distribution
C. Goodness-of-fit test
D. Contingency tables
E. Marginal distribution
F. Conditional distribution
G. Chi-Square independence test

XII. SIMPLE LINEAR CORRELATION AND REGRESSION ANALYSIS

A. Correlation
B. Scatter plot
    1. Positive association
    2. Negative association
C. Sample linear correlation coefficient r
    1. Properties of r
    2. Testing for the significance of correlation
D. Least squares regression model
    1. Interpreting slope and intercepts
    2. Graphs
    3. Using the least squares regression model
    4. Standard error