FDMAT 222: Biostatistics
Course Goals and Learning Outcomes

Department of Mathematics

February 13, 2009

Introduction

Fundamental expectations for FDMAT 222 classes:

1. Faculty will apply the BYU-Idaho Learning Model throughout the course
2. Whenever possible, examples, exercises, and cases will be based on real and relevant data
3. Students will complete a group project in which they collect real data and prepare a formal (written or oral) presentation of the results
4. Some common departmental assessments will be administered; these may be formative, and are not necessarily exams

Course Goals

- Students will be able to make informed decisions based on data
- Students will be able to correctly apply a variety of statistical procedures and tests
- Students will know the uses, capabilities and limitations of various statistical procedures
- Students will be able to interpret the results of statistical procedures and tests

Learning Outcomes

- Students will develop skills in SPSS
- Students will be able to determine the correct procedures to use in a given situation
- Students will be able to explain how the Central Limit Theorem applies in inference
- Students will be able to interpret the meaning of confidence intervals in context
- Students will be able to interpret the results of hypothesis tests
- Students will be able to make an informed decision, based on the results of inferential procedures
- Students will become a statisticians with a small "s"

Topics Lists for FDMAT 221, 222, and 223

Topics	Approx Class Hrs.	FDMAT 221 Business Statistics	FDMAT 222 Biostatistics	FDMAT 223 Social Science Statistics
Quantitative Data				
Descriptive Statistics				
Sampling and Data Collection	1	√	√	√
Measures of Center; Histograms	1	√	√	√
Measures of Spread	1.5	√	√	√
Bridge Topics				
Probability	0.5	√	√	√
Mean and Variance of a Discrete RV	1	√		
Density Curves; Normal Distribtuion; QQ plots	1	√	√	√
Values and Areas Under a Normal Curve	0.5	√	√	√
Law of Large Numbers; Central Limit Theorem	1	√	√	√
Sampling Distribution of a Sample Mean	1	√	√	√
Inferential Statistics				
CI for μ with σ Known	1	√	√	√
CI for μ with σ Unknown	0.5	√	√	√
CI for Two Means, Dependent (Paired) Samples	0.5	√	√	√
CI for Two Means, Independent Samples	2	√	√	√
Introduction to Hypothesis Testing	1	√	√	√
Hypothesis Test for One Mean with σ Known	1	√	√	√
Hypothesis Test for One Mean with σ Unknown	0.5	√	√	√
Sample Size & Power	≈2		Module 1†	
Hypothesis Test for Two Means, (Paired)	0.5	√	√	√
Hypothesis Test for Two Means, Indep. Samples	1	√	√	√
Hypothesis Test for Several Means (ANOVA)	1		√	√
Post-hoc testing (Bonferroni, Tukey)	1		√	√
ANCOVA	–		Module 2‡	
Introduction to Nonparametric Tests*	0.1			√*
Bivariate Data				
Descriptive Statistics				
Scatterplots, Correlation	1	√	√	√
Covariance	0.5	√		
Bridge Topics				
Least-squares Regression	1	√	√	√
Introduction to Multiple Regression*	0.1		√*	
Inferential Statistics				
Regression Diagnostics	2	√	√	√
Hypothesis Test for Regression Coefficients	1	√	√	√
Categorical Data				
Descriptive Statistics				
Graphical and Numerical Summaries	1	√	√	√
Bridge Topics				
Sampling Distribution of a Sample Proportion	1	√	√	
Inferential Statistics				
Confidence Interval for One Proportion	1	√	√	
Confidence Interval for Two Proportions	1	√	√	
Hypothesis Test for One Proportion	1	√	√	
Hypothesis Test for Two Proportions	1	√	√	
Hypothesis Test for Independence/Several Prop.	1	√		√
Hypothesis Test for Goodness-of-Fit	1	√		
Odds Ratio & Relative Risk	–		Module 3§	
Statistical Software Skills				
Using Excel	0.5	√		
Using SPSS	0.5		√	√
Project				
Analyzing research results	0.5	√	√	√
Group research project	4	√	√	√

*This represents a very brief exposure to the topic. Students are informed that the methods exist, but do not implement them.

†Exercise Science students

‡Nursing students

§Biology and Health Science students

Specific Learning Outcomes

1 General Knowledge

Students will be able to

1. create the appropriate graph (histogram, boxplot, bar chart, pie chart, or Pareto chart) to represent values in a given data set
2. identify the appropriate confidence interval to apply in a given situation
3. identify the appropriate hypothesis test to apply in a given situation
4. analyze a case study using the appropriate descriptive and inferential procedures
5. critically analyze the basic statistical arguments in research articles from their field
6. identify critical issues in conducting research with human subjects

2 Lesson Topics

Overview

2.1 Course Overview

Students will be able to

1. discuss ways in which they will seek learning by faith
2. explain the course policies
3. access course resources (course outline, lesson schedule, preparation activities, homework assignments, assessments, communication tools, etc.)
4. communicate with their instructor

Quantitative Data

Descriptive Statistics

2.2 Sampling and Data Collection

Students will be able to

1. identify the difference between a categorical and a quantitative variable
2. classify quantitative random variables as either discrete or continuous
3. distinguish between an observational study and an experiment
4. distinguish between a population and a sample
5. distinguish between a parameter and a statistic
6. design a simple observational study or experiment
7. discuss the characteristics of the following sampling schemes:
   (a) simple random sample (SRS)

   (b) systematic sample
   (c) cluster sample
   (d) stratified sample
   (e) convenience sample
8. identify specific instances in which each type of sampling scheme should be used
9. explain the importance of using a random sample
10. discuss important considerations in designing a sample survey including:
   (a) bias
   (b) undercoverage
   (c) nonresponse
   (d) voluntary response samples
   (e) convenience samples
   (f) poorly worded questions
11. identify the following when presented with the description of an experiment:
   (a) subjects (or experimental units)
   (b) factors
   (c) treatment groups
12. explain the three principles of experimental design: control, randomization, and replication
13. correctly utilize the basic vocabulary associated with data collection

2.3 Getting Started with Statistical Software; Histograms; Measures of Center

Students will be able to

1. enter and edit data in SPSS
2. change variable attributes in SPSS
3. obtain summary statistics in SPSS
4. graphically summarize quantitative data by creating a histogram in SPSS
5. illustrate quantitative data using a histogram
6. interpret data presented graphically in a histogram
7. discuss the shape, center, and spread of data illustrated on a histogram
8. explain the mean, median, and standard deviation in a distribution of values
9. calculate the following summaries of quantitative data in SPSS :
   (a) mean
   (b) median
   (c) standard deviation

2.4 Measures of Spread: Percentiles and Boxplots

Students will be able to

1. Use SPSS to compute percentiles (including the quartiles)
2. interpret percentiles (including the quartiles)
3. obtain the 5-number summary for a given data set
4. create a boxplot using SPSS
5. interpret data presented graphically in a boxplot

2.5 Measures of Spread: Standard Deviation and Variance

Students will be able to

1. calculate the standard deviation by hand for a small data set (Optional)
2. use SPSS to compute the standard deviation of a data set
3. explain in their own words what the standard deviation is
4. find the variance given the standard deviation

Bridge Topics

2.6 Probability

Students will be able to

1. calculate simple probabilities
2. apply the following probability rules:
   (a) a probability is a number between 0 and 1
   (b) the sum of the probabilities of all possible outcomes is 1
   (c) the probability that an event does not occur is 1 minus the probability that the event does occur

2.7 Density Curves; Normal Distribtuion; QQ plots

Students will be able to

1. explain the two properties of a density curve:
   (a) density curves lie on or above the X-axis
   (b) the total area under a density curve equals 1
2. interpret the area under a density curve as a probability
3. identify the shape of density curves (right-skewed, left-skewed, symmetric, and/or bell-shaped)
4. identify the normal density curve
5. state the properties of the normal density curve
6. locate the mean, median, and standard deviation, given a graph of a normal density curve
7. calculate a z-score for an individual observation, given the population mean and standard deviation
8. interpret a z-score

9. estimate probabilities using the Empirical (68-95-99.7) Rule for symmetric, bell-shaped distributions
10. identify unusual events
11. create a Q-Q plot in SPSS
12. visually assess normality (using a Q-Q plot and/or a histogram)

2.8 Converting Between Values and Areas Under a Normal Curve

Students will be able to

1. obtain the probability of events involving a normal probability density curve
2. determine percentiles for a normally distributed random variable

2.9 Law of Large Numbers; Central Limit Theorem

Students will be able to

1. explain the Law of Large Numbers in their own words
2. explain the Central Limit Theorem (CLT) in their own words
3. determine if a sample mean will be normally distributed based on the shape of the parent population and the sample size

2.10 Sampling Distribution of a Sample Mean

Students will be able to

1. explain what a sampling distribution is
2. determine the sampling distribution of a sample mean when the parent population is normally distributed
3. determine the sampling distribution of a sample mean when the parent population is not normally distributed, but the sample size is large

Inferential Statistics

2.11 CI for μ with σ Known

Students will be able to

1. state what is meant by a 95% (or any other percentage) confidence interval
2. describe the role of the margin of error in a confidence interval
3. explain how the margin of error changes with the sample size and the level of confidence
4. distinguish between the level of confidence and the probability of capturing the true mean (0 or 1)
5. obtain a confidence interval for a single mean with σ known
6. check if the assumptions for creating the confidence interval are satisfied
7. interpret a confidence interval for a single mean with σ known in context

2.12 CI for μ with σ Unknown

Students will be able to

1. obtain a confidence interval for a single Mean with σ unknown
2. check if the assumptions for creating the confidence interval are satisfied
3. interpret a confidence interval for a single mean with σ unknown in context

2.13 CI for Two Means, Dependent Samples (Paired Data)

Students will be able to

1. obtain a confidence interval for the difference of two means with dependent samples
2. check if the assumptions for creating the confidence interval are satisfied
3. interpret a confidence interval for a difference of two means (with dependent samples) in context

2.14 CI for Two Means, Independent Samples

Students will be able to

1. explain the difference between dependent and independent samples
2. obtain a confidence interval for the difference of two means with independent samples
3. check if the assumptions for creating the confidence interval are satisfied
4. interpret a confidence interval for a difference of two paired means (with independent samples) in context

2.15 Introduction to Hypothesis Testing

Students will be able to

1. properly state the null and alternative hypotheses in a testing situation
2. explain the meaning of a P-value without using technical jargon
3. define the α-level
4. determine whether a one-sided or two-sided test is appropriate for a given data set
5. demonstrate how the alternative hypothesis (one- or two-sided) affects the computation of the P-value
6. assess statistical significance by comparing the P-value to the α-level
7. write a conclusion from a statistical test that explains the result in the context of the original problem
8. make an informed decision based on the results of a hypothesis tests

2.16 Hypothesis Test for One Mean with σ Known

Students will be able to

1. conduct a hypothesis test for a single Mean with σ known
2. verify if the assumptions of the hypothesis test are satisfied
3. make an informed decision based on the results of a hypothesis test for one mean with σ known

2.17 Hypothesis Test for One Mean with σ Unknown

Students will be able to

1. conduct a hypothesis test for a single mean with σ unknown
2. verify if the assumptions of the hypothesis test are satisfied
3. make an informed decision based on the results of a hypothesis test for one mean with σ unknown

2.18 Hypothesis Test for Two Means, Dependent Samples (Paired Data)

Students will be able to

1. conduct a hypothesis test for a comparison of two means (dependent samples)
2. verify if the assumptions of the hypothesis test are satisfied
3. make an informed decision based on the results of a hypothesis test for a difference of two means with dependent samples

2.19 Hypothesis Test for Two Means, Independent Samples

Students will be able to

1. conduct a hypothesis test for a comparison of two means (independent samples)
2. verify if the assumptions of the hypothesis test are satisfied
3. make an informed decision based on the results of a hypothesis test for a difference of two means with independent samples

2.20 Hypothesis Test for Several Means (ANOVA)

Students will be able to

1. conduct a hypothesis test for equality of several means (one-way ANOVA)
2. conduct post-hoc tests using the Bonferroni adjustment in an ANOVA setting
3. conduct post-hoc Tukey tests for pairwise comparisons of means in an ANOVA setting
4. verify if the assumptions of the hypothesis test are satisfied
5. make an informed decision based on the results of an ANOVA test

2.21 Nonparametric Tests

1. (This topic is not necessarily covered in FDMAT 222.)

Bivariate Data

Descriptive Statistics

2.22 Scatterplots, Correlation

Students will be able to

1. create a scatterplot of bivariate data using SPSS

2. analyze the overall pattern in a scatter plot to assess the form (linear, nonlinear), direction (positive or negative association), and strength (weak, moderate, strong) of the relationship between two quantitative variables
3. identify potential outliers in a scatter plot
4. use SPSS to calculate the correlation coefficient
5. interpret the correlation coefficient, r, as a measure of the strength of the linear relationship between two variables
6. explain the properties of the correlation coefficient
   (a) the correlation of X and Y is the same as the correlation between Y and X (i.e. there is no distinction between explanatory and response variables.)
   (b) r is a number between -1 and 1
   (c) positive values of r imply a positive linear relationship between the two variables
   (d) negative values of r imply a negative linear relationship between the two variables
   (e) values of r close to zero suggest there is a weak correlation between the two variables
   (f) if r is close to 1, it is evidence of a strong positive linear relationship between the two variables
   (g) if r is close to -1, there is evident of a strong negative linear relationship between the two variables
   (h) if r equals 1 or -1, then there is a perfect linear relationship between the two variables–i.e. the points are all in a line
   (i) the correlation coefficient measures the strength of the linear relationship between two variables; it does not give the strength of a nonlinear relationship, no matter how strong
   (j) the correlation coefficient is affected by outliers
7. distinguish between correlation and causation

Bridge Topics

2.23 Least-squares Regression

Students will be able to

1. identify the explanatory and response variables in a study
2. write the simple linear regression model for a given situation
3. display the simple linear regression line on a scatterplot in SPSS
4. obtain the estimated simple linear regression equation using SPSS
5. interpret, when appropriate, the Y-intercept of the regression line in the context of the original problem
6. explain the meaning of the slope of the regression line in the context of the original problem
7. predict values of the dependent variable given a particular value for the independent variable using the regression line
8. recognize settings in which multiple regression is appropriate

Inferential Statistics

2.24 Regression Diagnostics

Students will be able to

1. identify situations where simple linear regression is inappropriate: (nonlinear relationships, non-normal error terms, outliers)
2. use a residual plot and a Q-Q plot of the residuals to assess if simple linear regression is appropriate

2.25 Hypothesis Test for Regression Coefficients

Students will be able to

1. carry out a hypothesis test to determine if the true slope, β_1, is zero
2. explain the conclusions of the test for the slope in the context of the original problem
3. conduct a test of significance for the coefficients of a multiple regression equation
4. make an informed decision based on the results of a test for the significance of regression coefficients

Categorical Data

Descriptive Statistics

2.26 Graphical and Numerical Summaries

Students will be able to

1. graphically summarize categorical data in SPSS
   (a) create a pie chart using SPSS
   (b) create a bar chart using SPSS
   (c) create a Pareto chart by modifying a bar chart in SPSS
2. create plots involving grouped (summarized) data in SPSS
3. interpret data presented graphically using a bar chart, pie chart, and Pareto chart

Bridge Topics

2.27 Sampling Distribution of a Sample Proportion

Students will be able to

1. state the sampling distribution of a sample proportion
2. apply the sampling distribution of a sample proportion to calculate probabilities of events

Inferential Statistics

2.28 Confidence Interval for One Proportion

Students will be able to

1. calculate the appropriate confidence interval for a single proportion
2. check the relevant assumptions for creating confidence intervals for one proportion
3. interpret a confidence interval for one proportion in context

2.29 Confidence Interval for Two Proportions

Students will be able to

1. calculate the appropriate confidence interval for a comparison of two proportions (Optional)
2. check the relevant assumptions for creating confidence intervals for two proportions
3. interpret a confidence interval for a difference of two proportions in context

2.30 Hypothesis Test for One Proportion

Students will be able to

1. conduct a hypothesis test for a single proportion
2. check the relevant assumptions for conducting hypothesis tests with categorical data
3. make an informed decision based on the results of a test for one proportion

2.31 Hypothesis Test for Two Proportions

Students will be able to

1. conduct a hypothesis test for the comparison of two proportions
2. check the relevant assumptions for conducting hypothesis tests with categorical data
3. make an informed decision based on the results of a test for the difference of two proportions

2.32 Hypothesis Test for Several Proportions or χ^2 Test for Independence

Students will be able to

1. conduct a hypothesis test for equality of several proportions
2. make an informed decision based on the results of a test for the equality of several proportions or a test for independence
3. (This topic is not necessarily covered in FDMAT 222.)

2.33 Hypothesis Test for Goodness-of-Fit

Students will be able to

1. (This topic is not necessarily covered in FDMAT 222.)

2.34 Modules

Students will be able to

1. conduct basic sample size calculations (Module 1)
2. conduct basic power calculations (Module 1)
3. explain when an ANCOVA analysis is appropriate (Module 2)
4. conduct an ANCOVA analysis (Module 2)
5. determine which terms are significant in an ANCOVA analysis (Module 2)
6. calculate an odds ratio (Module 3)

7. calculate a relative risk (Module 3)
8. interpret an odds ratio (Module 3)
9. interpret a relative risk (Module 3)
10. conduct a hypothesis test for the significance of an odds ratio (Module 3)
11. conduct a hypothesis test for the significance of a relative risk (Module 3)

2.35 Project

Students will

1. work in groups on a term project
2. read a peer-reviewed journal article in their field
3. report on the use of statistics in a peer-reviewed journal article
4. collect data to answer a specific research question
5. analyze their data using both descriptive and inferential statistics
6. present their results in a formal (written or oral) report