Introduction to Biostatistics

Distance Education Online Course

http://ccghe.org

Course Syllabus

Welcome

Welcome. This course has been made available through joint efforts of the Johns Hopkins Center for Clinical Global Health Education and the Johns Hopkins Bloomberg School of Public Health. Funding has been provided by NIH Office of AIDS Research and Fogarty International.

Course Description

Introduction to Biostatistics provides an introduction to selected important topics in biostatistical concepts and reasoning through lectures, exercises, and bulletin board discussions. This course represents an introduction to the field and provides a survey of data and data types. Specific topics include tools for describing central tendency and variability in data; methods for performing inference on population means and proportions via sample data; statistical hypothesis testing and its application to group comparisons; issues of power and sample size in study designs; and random sample and other study types. While there are some formulae and computational elements to the course, the emphasis is on interpretation and concepts.

Objectives

Upon completion of the course you will be able to:

- Recognize and give examples of different types of data arising in public health and clinical studies
- Interpret differences in data distributions via visual displays
- Calculate standard normal scores and resulting probabilities
- Calculate and interpret confidence intervals for population means and proportions
- Interpret and explain a p-value
- Perform a two-sample t-test and interpret the results; calculate a 95% confidence interval for the difference in population means
- Select an appropriate test for comparing two populations on a continuous measure, when the two sample t-test is not appropriate
- Understand and interpret results from Analysis of Variance (ANOVA), a technique used to compare means amongst more than two independent populations
- Choose an appropriate method for comparing proportions between two groups; construct a 95% confidence interval for the difference in population proportions
- Understand and interpret relative risks and odds ratios when comparing two populations
- Understand why survival (timed to event) data requires its own type of analysis techniques
- Construct a Kaplan-Meier estimate of the survival function that describes the "survival experience" of a cohort of subjects
- Interpret the result of a log-rank test in the context of comparing the "survival experience" of multiple cohorts

- Describe different kinds of studies
- Understand confounding and interaction in studies
- Use SPSS/STATA package to:
    - Perform two sample comparisons of means and create confidence intervals for the population mean differences
    - Compare proportions amongst two independent populations
    - Interpret output from the statistical software package STATA related to the various estimation and hypothesis testing procedures covered in the course

Course Format
The content of this course is divided into 10 Lecture Modules, each corresponding to a grouping of lectures. All the required course work can be accessed from the Course Content (http://moodle.ccghe.net/course/view.php?id=41) page. The lecture sections are presented sequentially and should be completed in that order. Each of these sections combines audio presentation and slides—just like attending lectures in class – as well as the option to download the audio and/or slides separately. You may return to any previous section at any point and review its contents at your convenience. Below is a listing of each Lecture Module with the sections for each listed along with the time length of each presentation.

Pre and Post-Course Knowledge Assessments
There is a Pre- and a Post-Course Knowledge Assessment. The Pre-Assessment must be done prior to accessing any of the lecture material and the Post-Assessment should be taken after you have viewed all lectures to optimize you ability to do well. You may take each only 1 time.

Course requirements and Criteria for Earning a Certificate
Participants who complete the course according to the following guidelines will be able to print a Certificate of Achievement.

In order to receive a Certificate, the student must achieve the following:
1. Register to the website and enroll for the course online at www.ccghe.org
2. Achieve a Post-Course Knowledge Assessment score of 65% or better.
3. Complete the final course evaluation

Course Evaluation
At the end of this course, we urge you to complete the final course evaluation. The evaluation only takes a few minutes and it is valuable to us as a way to make program improvements for future courses. The current course is improved over past courses, because participants took the time to tell us how we are doing. We are grateful for their comments and suggestions.

**Funding for the development of this course is from NIH Fogarty International Center.*

Lecture Modules

Lecture 1 -- Describing Data: Part I
Section A : What Role Does Biostatistics Play in Public Health (Why Do I Need this Stuff?) 14:25
Section B : Types of Data 7:31
Section C : Continuous Data: Numerical Summary Measures; Sample Estimates versus Population Measures 24:20
Section D : Visually Displaying Continuous Data: Histograms 8:41
Section E : Stem and Leaf Plots, Box Plots 12:13
Section F : Samples versus Populations, Part 2: Sample Distribution versus Underlying "Population Distribution" 10:01

Lecture 2 -- Describing Data Part II
Section A : The Normal Distribution 7:44
Section B : Variability in the Normal Distribution: Calculating Normal Scores 14:24
Section C : Normal Scores and Variability in Non-Normal Data 8:10

Lecture 3 -- Sampling Variability and Confidence Intervals
Section A : The Random Sampling Behavior of a Sample Mean Across Multiple Random Samples 20:04
Section B : The Theoretical Sampling Distribution of the Sample Mean and Its Estimate Based on a Single Sample 13:07
Section C : Estimating Confidence Intervals for the Mean of a Population Based on a Single Sample of Size n: Some Examples 10:19
Section D : True Confessions Biostat Style: What We Mean by Approximately Normal and What Happens to the Sampling Distribution of the Sample Mean with Small n 9:42
Section E : The Sample Proportion as a Summary Measure for Binary Outcomes and the CLT 10:05
Section F : The Theoretical Sampling Distribution of the Sample Proportion and Its Estimate Based on a Single Sample 7:27
Section G : Estimating Confidence Intervals for the Proportion of a Population Based on a Single Sample of Size n: Some Examples 6:46
Section H : Small Sample Considerations for Confidence Intervals for Population Proportions 9:51

Lecture 4 -- An Introduction to Hypothesis Testing: The Paired t-Test
Section A : The Paired t-Test; the Confidence Interval Component 12:57
Section B : The Paired t-Test; the Hypothesis Testing Component 16:11
Section C : The Paired t-Test; Two More Examples 14:32
Section D : The p-Value in Even More Detail! 19:40

Lecture 5 -- Comparing Means among Two (or More) Independent Populations
Section A : Two Sample t-test: The Resulting Confidence Interval 17:42
Section B : Two Sample t-test: Getting a p-value 7:32
Section C : Two Sample t-test, Approach with Smaller Samples 10:07
Section D : Two Sample t-test, Two Choices 4:30
Section E : The Unpaired t-test: More Examples 9:26
Section F : Non-Parametric Analogue to the Two Sample t-test 11:14
Section G : Comparing Means between More than Two Independent Populations 13:09

Lecture 6 -- Simple Linear Regression
Section A : Review: The Equation of a Line 9:47
Section B : Linear Regression: Motivating Example 18:22
Section C : Simple Linear Regression: More Examples 15:21
Section D : Simple Linear Regression Model: Estimating the Regression Equation—Accounting for Uncertainty in the Estimates 19:16

Lecture 7 -- Comparing Proportions Between Two Independent Populations

Section A : The Two Sample z-Test for Comparing Proportions between Two Independent Populations: The Confidence Interval Approach 12:47

Section B : Two Sample z-test: Getting a p-value 9:39

Section C : The Chi-Square Test: Mechanics 11:48

Section D : More Examples of Comparing Two Proportions 5:13

Section E : Fisher's Exact Test 10:22

Section F : Measures of Association: Risk Difference, Relative Risk, and the Odds Ratio 20:12

Section G : FYI: Sampling Behavior of Relative Risks/Odds Ratios 10:30

Lecture 8 -- When Time is of Interest: The Case for Survival Analysis

Section A : Motivating the Need 15:01

Section B : Estimating the Survival Curve: The Kaplan Meier Approach 18:17

Section C : Statistical Inference on Survival Curves 16:44

Lecture 9 -- Study Design

Section A : Making the Case for Randomized Controlled Studies 12:46

Section B : Methods of Randomization 7:07

Section C : But John, What If Randomization Is Not Possible? 15:44

Section D : Another Non-Randomized Study Design: The Case-Control Design 12:14

Lecture 10 -- Confounding and Effect Modification

Section A : Confounding: An Introduction 14:06

Section B : Confounding, More Examples 19:36

Section C : Statistical Interaction/Effect Modification 12:03