

STA 501 Methodologies of Applied Statistics

Spring 2021

CONTACT INFORMATION

Instructor: Cheng Peng
Office: 25 University Ave. 111
Phone: 610-436-2369
Email: cpeng@wcupa.edu

Office Hours: Monday/Tuesday/Wednesday: 1:00 PM-3:00 PM

Live Zoom Lectures: Thursday 6:00 PM – 9:00 PM

[Class Zoom Meeting Link](#)

TEXTBOOKS

Required Textbooks

(Free eBooks via WCU Library),

1. [A Primer in Biological Data Analysis and Visualization Using R](#), by Grepp Hartvigsen, Columbia University Press, 2014. eBook link:
2. [Basic Statistics: An Introduction with R](#), by Tenko Raykov and George Marcoulides, Rowman & Littlefield Publishers.
3. [Introductory Biostatistics](#), by Chap T. Le and Lynn E. Eberly, Wiley, 2016.

Optional Textbooks

(Free eBooks via WCU Library)

1. [Biostatistics for Epidemiology and Public Health Using R](#), by Bertram K. C. Chan.
2. [Statistics Explained: An Introductory Guide for Life Scientists](#), by Steve McKillop, Cambridge University Press, 2005.
3. [Understanding Biostatistics](#) (This is an excellent text but a little technical)

COURSE DESCRIPTION

This course is intended to introduce and teach the commonly used statistical techniques that are most likely to be encountered during graduate-level research. It will be applied in focus, de-emphasizing theory, and formulas to focus on analyzing data, reading output, and interpreting results. We will cover a wide range of topics including, but necessarily limited to, t-tests, multiple regression, ANOVA, chi-square tests, and power/sample size calculations. Most in-class lectures will be supplemented by a lab intended to give hands-on experience to help reinforce the material.

STUDENT LEARNING OUTCOMES (SLO) FOR MS IN APPLIED STATISTICS

1. Demonstrate an understanding of probability and statistical inference, including the fundamental laws of classical probability, discrete and continuous random variables, expectation theory, maximum likelihood methods, hypothesis testing, power, and bivariate and multivariate distribution theory.
2. Demonstrated the ability to apply the elementary methods of statistical analysis, namely those based on classical linear models, categorical methods, and non-parametric ideas to perform data analysis for statistical inference.
3. Demonstrated proficiency in the effective use of computers for research data management and analysis of data with standard statistical software packages.
4. Learned to develop and critically assess the design of experimental studies and the collection of data.
5. Applied one or more methods of statistical inference to a particular area of interest, particularly the program in the elective concentration.
6. Gained practical experience in statistical consulting and communicating with non-statisticians, culminating with interaction with research workers at a local company as part of the internship practicum.

COURSE OBJECTIVES

1. Use R for exploratory data analysis and basic data visualization. [SLO3]
2. Determine the correct statistical analysis for a given set of data. [SLO2]
3. Utilize R to do t-tests, multiple regression, ANOVA, chi-square tests, and power-sample size calculations. [SLO2, SLO3]
4. Correctly interpret the output from all of the above analyses. [SLO2]
5. Communicate the results of these statistical analyses concisely and simply that would be understandable to a non-statistician. [SLO2]
6. Use R Markdown to generate statistical analysis reports. [SLO3]

EVALUATION & GRADING

The course grade will be based on

1. Class attendance and participation (5%)
2. Two Midterm Exams/Projects (20% each, to be held in weeks 6 and 11)
3. Final Exam/Project: 20%
4. Weekly Quiz/Lab Work: 35%

Weekly quiz/Lab work

Weekly quizzes and lab work will be posted on the D2L every Thursday at noon and due at 11:30 PM on Sunday. Solutions to the quizzes will be posted on the course web page on Wednesdays.

Exams

We will have three written examinations which should be written up neatly and organized in R Markdown. The information about the dates and topics to be covered in midterm exams will be posted on the course web page a week before the corresponding scheduled date. A cumulative final exam will be given in the final week according to WCU final exam schedule. Attendance at examinations is crucial. Makeup exams will only be allowed in the event of an excused absence as defined in the WCU Graduate Catalog or prior arrangement with the instructor.

A letter grade will be assigned based on performance in the course, according to the following scale:

Grade	Quality Points	Percentage Equivalents	Interpretation
A	4.00	[93%, 100%]	Superior graduate attainment
A-	3.67	[90%, 93%)	
B+	3.33	[86%, 90%)	Satisfactory graduate attainment
B	3.00	[83%, 86%)	
B-	2.67	[80%, 83%)	
C+	2.33	[76%, 80%)	Attainment below graduate expectations
C	2.00	[73%, 76%)	
C-	1.67	[70%, 73%)	
F	0	< 70%	Failure

Attendance Policy

Attendance is mandatory for this class. I will take attendance during the live Zoom class meeting on Thursdays.

TENTATIVE TOPICS

Following is the list of tentative topics to be covered in the semester. I may modify and list as we move forward during the semester.

Week 1: Setting up computing tools - getting started with R, RStudio, and R Markdown [CLO1, CLO5]

1. Introduction: class structure, topics, assessments, and logistics.
2. Install R, RStudio, and possibly MikTeX
3. Create R Markdown document- Knit HTML, PDF, and WORD file
4. Getting started with R: basic operations, vectors, data frames (R data sets).
5. Install and load R libraries

Week 2: Sampling and Experimental Design [CLO1]

1. Data collection – sampling plans
2. Study / experimental designs
3. Data structures
4. Data granularity levels
5. **R Applications:** (a) loading data to R; (b). sampling population using R

Week 3: Data visualization and Descriptive Statistics [CLO1, CLO4]

1. Descriptive statistics: tables and charts.
2. Numerical measures: centers, spread, and locations
3. Five-number-summary and boxplot
4. **R Applications:** Descriptive statistics with R

Week 4: Standard scores, the Normal, t, Chi-Squared, and F Distributions [CLO1]

1. Introduction to probability distributions
2. Two types of questions: probability and quantile
3. Special distributions: normal, t, chi-square, and F distributions and their applications
4. **R Applications:** solve the two types of questions (focus on normal, t, and chi-squared)

Week 5: Sampling distributions [CLO1]

1. Concepts of the sampling distribution
2. The motivation to study sampling distributions and the two-types of problems
3. Sampling distributions of sample means and proportions
4. Sampling distributions related to sample variances
5. **R Applications:** Two-types of problems associated with sampling distributions

Week 6: Confidence Intervals [CLO1]

1. Motivation and practical formulation
2. Two forms of inference problems: confidence interval and testing hypothesis.
3. Interpretations of confidence intervals: concepts related to confidence intervals
4. Confidence intervals of the population mean and proportion
5. Confidence intervals of the difference between two population means and proportions.
6. **R Applications:** Case study - Some R functions for confidence intervals and testing hypothesis.
7. **Midterm #1**

Week 7: Tests of a single mean/proportion and two means/proportions [CLO2, CLO3]

1. Taxonomy of hypothesis testing
2. Testing a single population mean
3. Testing the difference between two population means
4. Testing population proportion(s)
5. **R Applications:** Case study - Hypothesis testing with R

Week 8: ANOVA [CLO1, CLO2, CLO3, CLO4]

1. The nature of practical questions – Comparing more than t means
2. Assumptions of one-way ANOVA
3. Null and alternative hypotheses
4. Welch's one-way ANOVA
5. **R Applications:** Case study - ANOVA with R and Interpretations

Week 9: Correlation and Simple Linear Regression [CLO1, CLO2, CLO3, CLO4]

1. Relationship between two numerical variables
2. Linear relationship: the strength of linear correlation -coefficient of correlation
3. Least square regression model – structure, assumptions, and interpretation
4. R square and interpretation
5. **R Applications:** Case study - simple linear regression (SLR) with R.

Week 10: Multiple Linear Regression [CLO1, CLO2, CLO3, CLO4]

1. ANOVA as a special MLR – Dummy variables and interpretation
2. ACOVA – MLR with both categorical and numerical predictor variables
3. Assumptions and diagnostics
4. Summarizing output and interpreting coefficients and R square
5. **R Applications:** Case study - MLR with R

Week 11: Binary Categorical Regression [CLO1, CLO2, CLO3, CLO4]

1. Practical question and model formulation
2. Structure of the model
3. Variable selection methods and criteria
4. Interpretation of the regression coefficients
5. **R Applications:** Case study - logistic regression with R
6. **Midterm #2**

Week 12: Frequency Count Regression [CLO1, CLO2, CLO3, CLO4]

1. Practical question and model formulation
2. Model structure and interpretation
3. Regression on rates
4. Variable selection
5. **R Applications:** Case-study – modeling count and rates in biology

Week 13: Procedures Related to Nominal Data [CLO1, CLO2, CLO3, CLO4]

1. Independence of two nominal variables: chi-square-test
2. Chi-square test of goodness-of-fit
3. Conditional independence: CMH test based on 2x2xk tables
4. **R Applications:** Case study – chi-square tests involving nominal variables

Week 14: Power and Sample Size Determination [CLO1, CLO2, CLO3, CLO4]

1. The margin of error and effect size
2. Interpret the statistical power in tests of hypothesis
3. Calculate the sample size to meet the precision of the parameter estimation
4. The sample size required to achieve the power of a statistical test
5. **R Applications:** Case study – power and sample size calculation

ACADEMIC & PERSONAL INTEGRITY

It is the responsibility of each student to adhere to the university's standards for academic integrity. Violations of academic integrity include any act that violates the rights of another student in academic work, that involves misrepresentation of your work, or that disrupts the instruction of the course. Other violations include (but are not limited to): cheating on assignments or examinations; plagiarizing, which means copying any part of another's work and/or using ideas of another and presenting them as one's own without giving proper credit to the source; selling, purchasing, or exchanging of term papers; falsifying of information; and using your work from one class to fulfill the assignment for another class without significant modification. Proof of academic misconduct can result in the automatic failure and removal from this course. For questions regarding Academic Integrity, the No-Grade Policy, Sexual

Harassment, or the Student Code of Conduct, students are encouraged to refer to the Department Graduate Handbook, the Graduate Catalog, the *Ram's Eye View*, and the [University website](#).

STUDENTS WITH DISABILITIES

If you have a disability that requires accommodations under the Americans with Disabilities Act (ADA), please present your letter of accommodations and meet with me as soon as possible so that I can support your success in an informed manner. Accommodations cannot be granted retroactively. If you would like to know more about West Chester University's [Office of Services for Students with Disabilities \(OSSD\)](#), please visit them at 223 Lawrence Center. The OSSD hours of Operation are Monday – Friday, 8:30 a.m. – 4:30 p.m. Their phone number is 610-436-2564, their fax number is 610-436-2600, their email address is ossd@wcupa.edu.

REPORTING INCIDENTS OF SEXUAL VIOLENCE

West Chester University and its faculty are committed to assuring a safe and productive educational environment for all students. To meet this commitment and to comply with Title IX of the Education Amendments of 1972 and guidance from the Office for Civil Rights, the University requires faculty members to report incidents of sexual violence shared by students to the University's Title IX Coordinator, Ms. Lynn Klingensmith. The only exceptions to the faculty member's reporting obligation are when incidents of sexual violence are communicated by a student during a classroom discussion, in a writing assignment for a class, or as part of a University-approved research project. Faculty members are obligated to report sexual violence or any other abuse of a student who was, or is, a child (a person under 18 years of age) when the abuse allegedly occurred to the person designated in the University protection of minor's policy. Information regarding the reporting of sexual violence and the resources that are available to victims of sexual violence is set forth at the webpage for the [Office of Social Equity](#).

EXCUSED ABSENCES POLICY

Students are advised to carefully read and comply with the excused absences policy, including absences for university-sanctioned events, contained in the WCU Graduate Catalog. In particular, please note that the "responsibility for meeting academic requirements rests with the student," that this policy does not excuse students from completing required academic work, and that professors can require a "fair alternative" to attendance on those days that students must be absent from class to participate in a University-Sanctioned Event.

EMERGENCY PREPAREDNESS

All students are encouraged to sign up for the University's free WCU ALERT service, which delivers official WCU emergency text messages directly to your cell phone. For more information, visit [WCU Alert](#). To report an emergency, call the Department of Public Safety at 610-436-3311.

ELECTRONIC MAIL POLICY

It is expected that faculty, staff, and students activate and maintain regular access to University-provided e-mail accounts. Official university communications, including those from your instructor, will be sent through your university e-mail account. You are responsible for accessing that mail to be sure to obtain official University communications. Failure to access will not exempt individuals from the responsibilities associated with this course.