Overview and Objectives:
This course is designed for researchers in the clinical and translational sciences who want to understand and use linear regression techniques. Topics covered include multiple linear regression, regression diagnostics, analysis of variance, analysis of covariance, confounding, mediation, moderation, interactions, and model selection. At the completion of the course, trainees should be able to demonstrate an understanding of the basic principles of ANOVA and linear regressions, to assess their adequacy and assumptions, to analyze simple data sets taken from the fields of medicine and public health, to summarize results from regression models via written communication, and to recognize situations that require more advanced methodologies.

Responsibilities:

- There will be reading assignments in the textbook. All readings are expected to be completed before class. You should anticipate class discussions about the material you are expected to read.
- Students will be assigned four homework assignments that will be graded. These assignments will be used to reinforce material reviewed in class. All homework assignments will be assigned a due date. You are encouraged to discuss and work together on problems, but you must write up your results individually, i.e. very similar papers will not be accepted. Homework assignments are to be turned in at the beginning of class on the due date. No assignment will be accepted via email. Late homework assignments will be penalized 10% per day past the due date (unless prior arrangements have been made with the Dr. Tudorascu) until the following lecture, at which no late homeworks will be accepted.
- There will be a quiz on TBD consisting of short answer and multiple choice questions. The focus of this quiz will be on understanding 1) what to do and when to do it, 2) how to interpret what you did, and 3) more focused questions on the important issues. We will discuss the results on the last day of class.
- There will be a final project due one week from the last day of class. This project will be a compilation of the homework assignments but with synthesized text that pulls the analyses together in a coherent, scientific manner. This project must be turned in on time to receive full credit, and will be subject to the same late penalties as the homework (outlined above). Students must work on the final project independently. Discussion, consulting, or working with other people is not permitted.

Course Requirements:

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class participation and attendance</td>
<td>0%</td>
</tr>
<tr>
<td>Homework assignments</td>
<td>55%</td>
</tr>
</tbody>
</table>
Final project 30%
Quiz 15%

Attendance Policy:

- Students are expected to sign-in to each class (computer provided in suite lobby). If a problem is encountered with the sign-in system, please contact the course instructor(s) as well as Allie Giel (alg190@pitt.edu) immediately.

Course Grading Scale:

For the computation of the final course grade as well as for the course assignments, the following grading scale will be used:

93 - 100 = A  86 – 89 = B+  76 – 79 = C+  66 – 69 = D+
90-92 = A-  80 – 85 = B  70 – 75 = C  60 – 65 = D  < 60 = F

Required Textbook(s):


Supplemental Textbook(s):


Books are available at the ICRE library in Parkvale 306.

Website resources:

All homework assignments, course information, and communication will be available at [http://courseweb.pitt.edu](http://courseweb.pitt.edu).

Academic Integrity:

Students in this course will be expected to comply with the [University of Pittsburgh’s Policy on Academic Integrity](http://www.provost.pitt.edu/info/ai1.html). Any student suspected of violating this obligation for any reason during the semester will be required to participate in the procedural process, initiated at the instructor level, as outlined in the University Guidelines on Academic Integrity. This may include, but is not limited to, the confiscation of the examination of any individual suspected of violating University Policy. Furthermore, no student may bring any unauthorized materials to an exam, including dictionaries and programmable calculators.
Course Schedule

Session 1: Correlation and Simple Linear Regression

At the conclusion of this lecture, the student will be able to:
Describe correlation and the linear regression model and explain its appropriate use answering biomedical research questions.

Topics:
1. Course description
2. Review of correlation
3. Review of simple linear regression (model, assumptions, estimation, inference, prediction)
4. Prediction and confidence bands
5. Introduction to multiple linear regression

Competencies:
Data Management and Biomedical Informatics: Organize datasets (variable display and structure) appropriately for given statistical techniques.
Applied Analytic Techniques: Determine and apply a range of appropriate statistical techniques to answer research questions and explain the implications of missing data on conclusions drawn from statistical results.

Required Reading(s): Vittinghoff Ch 1 (Introduction), Ch 2 (Explor & Descr) and Ch 3 pgs 33-42 (Correlation & SLR)

Supplemental reading: Kleinbaum Ch 4 -7 (Intro Regression, SLR, Correlation, ANOVA table)

Homework assignment(s):
Homework 1: Correlation, simple linear regression, and multiple linear regression
Due: TBD

Session 2: Multiple Linear Regression and Dummy Variables

At the conclusion of this lecture, the student will be able to:
Confidently use the linear regression model with multiple explanatory variables to discover important relationships between variables.

Topics:
1. Multiple (multivariable) regression model
2. Estimation
3. Evaluating model and predictor significance
4. Dummy (indicator) variables
5. Introduction to confounding, mediation, and interaction
**Competencies:**
*Data Management and Biomedical Informatics:* Organize datasets (variable display and structure) appropriately for given statistical techniques.
*Applied Analytic Techniques:* Determine and apply a range of appropriate statistical techniques to answer research questions and explain the implications of missing data on conclusions drawn from statistical results

**Required Reading(s):** Vittinghoff Chapter 4 69-81 (Linear regression and categorical predictors)

Supplemental reading: Kleinbaum Ch 8-10 & Sections 12.1-12.3 (MLR, Correlations, Intro to dummy vars)

---

**Session 3: Confounding and Interaction**

***Instructor: Tudorascu***

**At the conclusion of this lecture, the student will be able to:**
Identify and appropriately address confounding and interaction.

**Topics:**
1. Covariates
2. Confounding
3. Interaction and testing for equal slopes

---

**Competencies:**
*Applied Analytic Techniques:* Describe appropriate data analysis plans for addressing specific research questions.

**Required Reading(s):** Vittinghoff Chapter 4 89-108 (Confounding, Mediation, and Interaction)
Supplemental reading: Kleinbaum Chapters 11 and 12 (Confounding and Interaction)

**Homework assignment(s):**
**Homework 2: Confounding, interaction, and regression diagnostics**
Due

**Due Today:**
**Homework 1: Correlation, simple linear regression, and multiple linear regression**

---

**Session 4: Mediation, Moderation and Diagnostics**

***Instructor: Rothenberger***

**At the conclusion of this lecture, the student will be able to:**
Understand mediator and moderation and be able to perform diagnostic tests on a regression model.

**Topics:**
1. Mediation and moderation
2. Regression diagnostics
3. Model assumptions
**Competencies:**
*Data Management and Biomedical Informatics:* Organize datasets (variable display and structure) appropriately for given statistical techniques.

*Applied Analytic Techniques:* Determine and apply a range of appropriate statistical techniques to answer research questions and explain the implications of missing data on conclusions drawn from statistical results.

**Required Reading(s):**
Vittinghoff Chapter 4 109-130 (Model Assumptions & Fit) and Ch 10 421-423 (Collinearity & number of predictors)

---

**Session 5: ANOVA, ANCOVA and Two-Way ANOVA**

**Instructor:** Rothenberger

**At the conclusion of this lecture, the student will be able to:**

Describe the ANOVA model and explain when it should be used compared to the linear regression. Describe the concept of adjusted means from ANCOVA and appropriate application of two-way ANOVA.

**Topics:**

1. ANOVA vs. dummy coding
2. Reference cell coding
3. ANCOVA and linear regression
4. Adjusted means example
5. Two-way ANOVA

**Homework assignment:**
*Homework 3: ANOVA, ANCOVA, two-way ANOVA (Due September 25th)*

**Required Reading(s):** Vittinghoff Chapter 3 pgs 28-32 (t-test and ANOVA) and pages 81-84 (Multiple comparisons)

Supplemental reading: Kleinbaum Ch 17 (One-way ANOVA)

**Due Today:**
*Homework 2: Confounding, interaction, and regression diagnostics*
At the conclusion of this lecture, the student will be able to:

Correctly select the best linear regression model to fit exploratory data.

Topics:

1. Model selection
2. Best prediction of Y
3. Steps in selecting the best regression equation
4. Criteria for selecting a model
5. Backward, forward, and stepwise procedures
6. Evaluating a primary predictor and multiple important predictors
7. Introduction to hierarchical regression
8. Review of class topics

Competencies:

Applied Analytic Techniques: Determine and apply a range of appropriate statistical techniques to answer research questions and explain the implications of missing data on conclusions drawn from statistical results.

Required Reading(s): Vittinghoff Chapter 10 (Predictor Selection sections 10.2, 10.3, and 10.4) and Ch 8 257 & 267 (Hierarchical Data)
Supplemental reading: Kleinbaum Ch 14 and 16 (Diagnostics and Selecting the Best Regression Equation)

Homework assignment(s): Homework 4: Model selection

At the conclusion of this lecture, the student will be able to:

Determine the meaningful sample size and power for a given biomedical research question.

Topics:

1. Sample size and power calculations

Competencies:

Sampling: Identify appropriate study populations and sample size, control and comparison groups, and possible sources of bias for research problems.

Applied Analytic Techniques: Determine and apply a range of appropriate statistical techniques to answer research questions and explain the implications of missing data on conclusions drawn from statistical results.
**Required Reading(s):** Vittinghoff Chapter 4, section 4.8 (Sample Size)

**Due Today:** Homework 3: ANOVA, ANCOVA, two-way ANOVA

**Session 8: Review of quiz, Correlated Data, Missing data, Assignment of project**

**Instructor:** Tudorascu

---

**At the conclusion of this lecture, the student will be able to:**

- Explain how and when to use the linear regression and ANOVA models.
- Recognize studies that have or require correlated data.
- Explain different types of missing data mechanisms.

**Topics:**

1. Review of regression and ANOVA
2. Correlated data
3. Missing data

**Competencies:**

*Applied Analytic Techniques:* Determine and apply a range of appropriate statistical techniques to answer research questions and explain the implications of missing data on conclusions drawn from statistical results.

*Written Communication:* Organize and report statistical results.

**Required Reading(s):**

- Required Reading: Vittinghoff Chapter 13 Summary
- Supplemental Reading: Vittinghoff Chapter 7 Repeated Measures Analysis pgs 261-271
  Vittinghoff Chapter 11 Missing data (skim this if you are interested)

**Due Today:** Homework 4: Model Selection