Instructor: Prof. Spyros Konstantopoulos (spyros@msu.edu)
Office hours: By appointment (please email me to schedule an appointment)
Office Location: 450 Erickson Hall

T.A.s:
Cheng-Hsien Li (lichenh@msu.edu)
Office hours: TBD
Location: Graduate Student Lounge – Erickson Hall

Ting Shen (shentin1@msu.edu)
Office hours: TBD
Location: Graduate Student Lounge – Erickson Hall

Jing-Ru Xu (xujingru@msu.edu)
Office hours: TBD
Location: Graduate Student Lounge – Erickson Hall

Office hours are on Fridays and Mondays (TBD)

Class Hours: Tuesday, 4:10pm to 7:00pm
Classroom: 116 Farrall Agricultural Engineering Hall

Course Content
This course introduces students to techniques of data analysis and statistical inference commonly
used in educational, sociological, economic, and psychological research. Students will conduct
analyses in SPSS using data sets such as the NELS88, the ECLS-K 1998-99, Tennessee Star, and
Add Health. These data bases are among the largest and most important collected by the federal
government, including extensive measurements of students’ beliefs, aspirations, attitudes, health
behaviors, test scores, and background, as well as related information from teachers, parents, and
schools. The major topics are univariate and multiple regression, one- and two-factor analysis of
variance with multiple comparisons and interactions, and logistic regression. Knowledge of
basic algebra is required, as is an understanding of the fundamental principles of descriptive
statistics and hypothesis testing (as taught, for example, in CEP 932 or equivalent). Knowledge
of calculus is not required.

Course Objectives
By the end of the course the student should have demonstrated the ability to:
1. recognize continuous and discrete (or categorical) variables and choose appropriate statistical procedures accordingly

2. describe the relationship between predictor variables and a continuous outcome variable

3. find point estimates and confidence intervals and do hypothesis tests for regression coefficients

4. formulate multiple regression models appropriate for various research problems and interpret computer output relevant to those models

5. delineate assumptions of linear statistical models and examine data to evaluate conformity to those assumptions

6. formulate analysis-of-variance models, estimate their parameters, and test hypotheses about those parameters

7. recognize similarities and differences between regression and analysis-of-variance models

8. identify and control sources of error through experimental design and statistical adjustment

9. identify observations which may be dependent, and explain the limitations of ordinary techniques for these data

10. formulate logistic regression models for binary dependent variables and interpret computer output relevant to those models

11. write coherent summaries and interpretations of data analyzed by the above procedures

**Required Text:**


**Alternative texts and references:**


**Evaluation**

Grades will be based on points accumulated on assignments and examinations. There will be five homework assignments, and one final. There will be 100 total possible points, distributed as follows:

- **Final Exam** (scheduled time only) 25%
- **Homework assignments** 75%

*Homework assignments may be done in a group of two or three students. Each homework assignment counts for 15% of the course grade.*

The overall course grade ranges in terms of total points will be:

- 4.0 > 90%, 3.5 > 85%, 3.0 > 80%, 2.5 > 70%, 2.0 > 60%

**If you would like to appeal any grade on your homework you must make the appeal in writing and wait at least one day after the homework has been returned to you.**

**Late Assignment Policy**

Homework assignments are due at the beginning of class on the day they are due. If you decide to hand in the assignment late, it will be penalized an additional 10% for each day it is late. This means that homework handed in after class starts will be penalized 10%. The homework will be penalized an additional 10% for each subsequent day it is late (e.g., homework that is handed in the day after it is due will be penalized 20%).
How to do well in this course

A) Assignments
   1) Allow at least 10 hours per assignment (more likely 20)
   2) Come to class
   3) Be thorough—respond to all parts of the questions
   4) Be punctual—this class can bury you if you get too far behind
   5) Ask questions in class and contact the instructor and the TAs
   6) Read thoroughly, relevant to lectures

B) Final Exam
   1) Review assignments
   2) Review lectures
   3) Synthesize and get the big picture!

Other Issues

Students with disabilities: Reasonable accommodations for persons with documented disabilities will be made available. Please feel free to speak with us if there are issues of which we should be aware.

Academic Honesty and Integrity: Students are assumed to be honest, and course work is assumed to represent the student’s own work. Violations of the academic integrity policy such as cheating, plagiarism, selling course assignments or academic fraud are grounds for academic action and/or disciplinary sanction as described in the University’s student conduct code.

Incidents of Plagiarism: They will be taken very seriously and will be pursued. Students are strongly cautioned not to copy any text verbatim without appropriate quotations and source citations.

For University regulations on academic dishonesty and plagiarism, please refer to: https://www.msu.edu/unit/ombud/academic-integrity/plagiarism-policy.html

The instructor reserves the right to make any changes he considers academically advisable. Changes will be announced in class, it is your responsibility to keep up with any changed policies, schedules, and assignments.
<table>
<thead>
<tr>
<th>Date</th>
<th>Day</th>
<th>Topic covered</th>
<th>Assignments &amp; readings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-8</td>
<td>1</td>
<td>Introduction to course Review: definitions Review: Distributions and sampling distributions</td>
<td>AF: Chapters 1-8</td>
</tr>
<tr>
<td>1-15</td>
<td>2</td>
<td>Review: Expected Value and Bias Review: Hypothesis testing, T-test</td>
<td>AF: Chapters 1-8</td>
</tr>
<tr>
<td>1-22</td>
<td>3</td>
<td>Introduction to simple linear regression</td>
<td>AF: Chapter 9</td>
</tr>
<tr>
<td>1-29</td>
<td>4</td>
<td>Simple linear regression continued. SPSS Lab 1.</td>
<td></td>
</tr>
<tr>
<td>2-5</td>
<td>5</td>
<td>Multiple linear regression</td>
<td>HW1 Due: T-test and simple Regression</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AF: Chapters 10-11</td>
</tr>
<tr>
<td>2-12</td>
<td>6</td>
<td>Collinearity and inference in multiple regression</td>
<td>AF: Chapter 14</td>
</tr>
<tr>
<td>2-19</td>
<td>7</td>
<td>Interactions, dummy variables, omitted variable bias</td>
<td>HW2 Due: Multiple Regression 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AF: Chapter 13</td>
</tr>
<tr>
<td>2-26</td>
<td>8</td>
<td>SPSS Lab 2: Building models with interactions and dummies</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Introduction to ANOVA</td>
<td></td>
</tr>
<tr>
<td>3-12</td>
<td>9</td>
<td>One- and Two-way ANOVA</td>
<td>AF: Chapter 12</td>
</tr>
<tr>
<td>3-19</td>
<td>10</td>
<td>Regression and ANOVA</td>
<td>HW3 Due: Multiple Regression 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple hypothesis tests</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Change in R²</td>
<td></td>
</tr>
<tr>
<td>3-26</td>
<td>11</td>
<td>Introduction to multilevel models (random effects models)</td>
<td>AF: Chapter 16</td>
</tr>
<tr>
<td>4-2</td>
<td>12</td>
<td>Categorical outcome variables and logistic regression</td>
<td>HW4 Due: ANOVA and F-test</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AF: Chapter 15</td>
</tr>
<tr>
<td>4-9</td>
<td>13</td>
<td>Introduction to advanced methods</td>
<td>AF: Chapter 16</td>
</tr>
<tr>
<td>4-16</td>
<td>14</td>
<td>Experimental and quasi-experimental methods</td>
<td></td>
</tr>
<tr>
<td>4-23</td>
<td>15</td>
<td>Review of course</td>
<td>HW5 Due: Logistic regression</td>
</tr>
<tr>
<td>Week of 4/29</td>
<td></td>
<td>Final exam TBD</td>
<td></td>
</tr>
</tbody>
</table>