COURSE
Course: ECON 6374: Probability and Statistics for Economics
CRN: 16410
Semester: Spring 2017
Time: Tuesdays, 6:10 PM – 8:40 PM
Location: Rome 205

INSTRUCTOR
Name: Anthony Kassekert
Campus Address: Monroe 370, Department of Economics
2115 G St., NW, Washington, DC 20052
E-mail: akassekert@gwu.edu
Office hours: By appointment
Cell: 864-918-0112

COURSE DESCRIPTION
Econ 6374 is an introductory graduate econometrics course. Most economists are expected to perform statistical analysis as part of their job, but even those who do not work with data directly are expected to effectively read, critique, and utilize econometric research done by other people. The purpose of this class is to provide students to be both effective producers and smart consumers of statistical knowledge.

The material covered in this class will include basics of research design, probability theory, mathematical expectations, univariate and bivariate statistical comparison tools, and a brief introduction to regression analysis (covered in extensive detail next semester). Some proofs will be walked through during class in order to acquaint students with the notation and logic of modern econometrics, but these will not be used on any exams. To be clear, the level of mathematical rigor will not be sufficient for those wishing to continue on to a Ph.D. in economics.

COURSE PREREQUISITES
Knowledge of calculus and matrix algebra is necessary
Experience using programming tools SPSS, STATA, R, or SAS is advantageous but not required.
Prior exposure to elementary statistics/econometrics as an undergrad is preferred but not required.
TEXTS

2. TI-89 Calculator

For additional reference texts, please consult:

- Econometric Methods with Applications in Business and Economics, 1st Edition, Christiaan Heij, Paul de Boer, Philip Hans Franses, Teun Kloek, Herman K. van Dijk
- Wooldridge, J. Introductory Econometrics: A Modern Approach
- Mittelhammer, Ron C. Mathematical Statistics for Economics and Business

LEARNING OUTCOMES:
By the end of the semester, students will be proficient at performing a variety of univariate statistical tests, linear regression, and the fundamentals of research design. The primary objectives include, but are not limited to:

1. Understanding the fundamental comments of research, including experimental design and causality.
2. Formulating problem statements, research questions, and testable hypotheses.
3. Understanding of the theoretical underpinnings of econometric practice. These include knowing characteristics of frequently used distributions, calculating expectations, the Central Limit Theorem, and the role random sampling plays in modern inference.
4. Perform, interpret, and critique a wide range of univariate and multivariate statistical tests. These include z-test, t-test, Wilcoxon tests, Chi-Square tests, linear regression, and using bootstrapping on tests of medians and means.

In addition to these core technical skills, students will be expected to learn some basic data science issues related to data management and visualization. The fact is that the majority of research time spent by economists is not on the actual analysis, which is often quite simple. Instead, a painstaking amount of time is usually spent on data manipulation and visualizing the results for presentations to policymakers. These secondary goals include:

1. Merging data files from multiple sources (either using SQL or standard merge commands)
2. Recoding observations
3. Converting file formats
4. Statistically sound visual representation of data
AVERAGE MINIMUM AMOUNT OF INDEPENDENT, OUT-OF-CLASS, LEARNING EXPECTED PER WEEK
In a 15-week semester, including exam week, you should expect to spend a minimum of 4 hours a week for each hour of instruction. For a 2 ½ hour course, this means that you should expect to study a minimum of 10 hours outside of class each week. In addition, you should plan to spend at least two hours working on reading each chapter. The majority of time outside of class will be spent on homework assignments. Students should expect to spend 4 to five hours on homework each week.

GRADING
- Problem Sets 25%
- Research Paper 15%
- Data Science Project 10%
- Midterm Exam 25%
- Final Exam 25%

Problem Sets:
- Homework will be provided roughly on a weekly basis.
  - Please note that homework is scored equal to the exams, because repeated experience is key to learning the material.
- Homework will be more difficult than the exams. This is intentional as some material such as data management and proofs will be covered on homework and not in any tests.
  - Some homework will take 4-6 hours (depending on your coding ability).
- The lowest grade homework will be dropped.
- Homework is due at midnight of class day.
  - No late homework will be accepted for full credit.
  - If your name is not on the homework, I reserve the right to grade as zero.
  - No cell phone pictures! Homework needs to be legible and formatted into a single file.
- HW will be in one, single file. (no separate attachments for graphs and code). The instructor will not open more than one file. (learn to merge pdfs)
- Students are responsible for checking that the file uploaded correctly to Blackboard.
- The instructor reserves the right to add quizzes as necessary.

Research Paper and Presentation:
The class project will be a personalized data analysis performed by each student on the topic of their choosing, with instructor approval. The project will require setting a research question, formulating hypotheses, managing data, and presenting results in both written and graphical formats. Part of the grade will include an in-class presentation. Both a detailed document laying out expectations and an example project will be provided after the midterm.

Data Science Project:
The data science project is a large data manipulation and analysis using real data. Students will be given the data set location online and will need to show they can properly merge, format, manipulate, and report on real data in an informative way. Please note, this can be done in SAS, STATA, or R and in all three programs there are numerous methods to get the same result. Therefore, students will also be expected to annotate their code and multiple answers are “correct”.

**Midterm and Final Exams:**
The exams will be focused on understanding econometric concepts and will contain a mix of multiple choice questions and analytic problems that require analysis. A practice exam will be provided before the midterm and the final.

**Class participation:**
Come to class prepared and having reviewed material ahead of time. Please note I will randomly call on students to answer questions.

**Extra Credit and Curves:**
There is no guarantee that any extra credit will be provided. If an extracurricular activity presents itself that warrants additional points, no more than 5% of your grade can be earned. There is an extremely small, though non-zero, chance for a curve. This class is highly objective by design and students should not need a curve if they do all of the assignments and readings.

**CLASS POLICIES**
Attendance is not required or graded but is strongly recommended; there is a strong applied component to this course and students are therefore encouraged to engage closely with the material. Late work is not accepted under normal circumstances; if you have an extraordinary circumstance please contact me in advance for consideration.

**UNIVERSITY POLICY ON RELIGIOUS HOLIDAYS**
1. Students should notify faculty during the first week of the semester of their intention to be absent from class on their day(s) of religious observance;
2. Faculty should extend to these students the courtesy of absence without penalty on such occasions, including permission to make up examinations;
3. Faculty who intend to observe a religious holiday should arrange at the beginning of the semester to reschedule missed classes or to make other provisions for their course-related activities.

For GW’s teaching policies, see [http://www.gwu.edu/~academic/Teaching/main.htm](http://www.gwu.edu/~academic/Teaching/main.htm)

**ACADEMIC INTEGRITY**
I personally support the GW Code of Academic Integrity. It states: “Academic dishonesty is defined as cheating of any kind, including misrepresenting one's own work, taking credit for the work of others without crediting them and without appropriate authorization, and the fabrication of information.” Please note that allowing another student to copy your work is defined as cheating under the Academic Integrity code.
Common examples of academically dishonest behavior include, but are not limited to

1) Cheating
2) Fabrication
3) Plagiarism
4) Falsification and forgery of University academic documents
5) Facilitating academic dishonesty

Sanctions range from failure of the assignment, to failure of the course, to suspension or expulsion from the University. For the remainder of the code, see: http://www.gwu.edu/~ntegrity/code.html

SUPPORT FOR STUDENTS OUTSIDE THE CLASSROOM

DISABILITY SUPPORT SERVICES (DSS)
Any student who may need an accommodation based on the potential impact of a disability should contact the Disability Support Services office at 202-994-8250 in the Marvin Center, Suite 242, to establish eligibility and to coordinate reasonable accommodations. For additional information please refer to: http://gwired.gwu.edu/dss/

Students must make arrangements with the DSS office well in advance of needing to use the service.

UNIVERSITY COUNSELING CENTER (UCC) 202-994-5300
The University Counseling Center (UCC) offers 24/7 assistance and referral to address students' personal, social, career, and study skills problems. Services for students include:

- crisis and emergency mental health consultations
- confidential assessment, counseling services (individual and small group), and referrals

http://gwired.gwu.edu/counsel/CounselingServices/AcademicSupportServices
CLASS SCHEDULE
The class schedule is designed to cover the fundamental topics and techniques in probability and statistics. There will be a large amount of material from outside the require text for the first half of the class while discussing theoretical concepts. This schedule is preliminary and will be adjusted at will by the instructor based on class needs.

<table>
<thead>
<tr>
<th>Meeting</th>
<th>Topic</th>
<th>Moore et. al</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Descriptive Statistics &amp; Data Visualization</td>
<td>Ch. 1 &amp; 2</td>
</tr>
<tr>
<td>2</td>
<td>Research Design</td>
<td>Ch. 3</td>
</tr>
<tr>
<td>3</td>
<td>Probability Theory</td>
<td>Ch. 4 &amp; 5</td>
</tr>
<tr>
<td>4</td>
<td>Probability Distributions</td>
<td>Ch. 4 &amp; 5</td>
</tr>
<tr>
<td>5</td>
<td>Expectations</td>
<td>Ch. 4 &amp; 5</td>
</tr>
<tr>
<td>6</td>
<td>Central Limit Theorem and Law of Large Numbers</td>
<td>Ch. 6</td>
</tr>
<tr>
<td>7</td>
<td>Hypothesis Testing</td>
<td>Ch. 6</td>
</tr>
<tr>
<td>8</td>
<td>One Sample Tests</td>
<td>Ch. 7 &amp; 8</td>
</tr>
<tr>
<td>9</td>
<td>Two Sample Tests</td>
<td>Ch. 7 &amp; 8</td>
</tr>
<tr>
<td>10</td>
<td>Bootstrap &amp; Nonparametric Tests</td>
<td>Ch. 16</td>
</tr>
<tr>
<td>11</td>
<td>Chi Square Tests</td>
<td>Ch. 9</td>
</tr>
<tr>
<td>12</td>
<td>Linear Regression</td>
<td>Ch. 10</td>
</tr>
<tr>
<td>13</td>
<td>Multiple Regression</td>
<td>Ch. 11</td>
</tr>
<tr>
<td>14</td>
<td>Project Presentations</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Final</td>
<td></td>
</tr>
</tbody>
</table>

READING ASSIGNMENTS
Students should read the listed articles and text below before the respective class. These will be provided one week prior to the next class. The first weeks’ reading will be: