Catalog description: An introduction to the mathematics of statistics. Topics include exploratory data analysis, estimators, estimation, and hypothesis testing. Conceptual statistics and the mathematical development of statistical tests are studied.

Prerequisites: Successful completion of MTH 241 and MTH 250

Required text, software and other material:


- R - free software environment for statistical computing and graphics. Available for download here [http://www.r-project.org/](http://www.r-project.org/)


Overview: Statistics is the science of reasoning from data. It is both an exciting intellectual discipline and a powerful scientific tool. Statistics is a mathematical science, in the sense that it makes use of mathematics extensively, but it is not a branch of mathematics. This course will introduce you to fundamental concepts and methods of Statistics. Some of the key ideas to be studied include data collection strategies and the scope of conclusions based on the collection strategy, the role of randomness in collecting data and drawing conclusions, graphical and numerical summaries of data, simulation based inference, assessing statistical significance, and estimating with confidence.

Goals:

- Communicate your knowledge of statistical ideas effectively

- Apply and interpret the results of a variety of statistical techniques, including both descriptive and inferential methods

- Understand the fundamental ideas of statistics, including variability, distribution, sampling, experimentation, confidence, and significance

- Understand and apply simulation based inference

- Analyze and assess statistical arguments, such as those found in the popular press as well as in scholarly publications
Objectives:

- Conduct and interpret descriptive analyses of data, including graphical and numerical summaries, for categorical and quantitative data
- Analyze studies with the understanding of the differences between observational studies and controlled experiments and the scope of conclusions that each permits
- Interpret and explain population distributions, sample distributions, sampling distributions, bootstrap and randomization distributions
- Apply fundamental concepts of statistical inference, such as confidence and significance, explaining the limitations of these procedures

These objectives will be assessed through homework, exams and projects.

Computer Use: We will make extensive use of computers in this course. They will prove useful in at least three ways:

- for performing calculations and creating graphics necessary for analyzing data
- for conducting simulations to approximate long-run behavior of random phenomena
- for addressing “what if” questions that allow you to explore statistical concepts

The computer package R is a command-line driven environment with many packages and libraries available. Syntax is VERY important. If a command does not work, you should first check carefully that your syntax is 100% correct. You will be given a complete reference guide for all of the commands we use in this course. Your job is to create the ANALYSIS of the computer output. You should be able to debug your syntax errors since I am giving you a model file. Once you have correctly substituted data files, your errors are, most likely, due to misunderstandings of the statistical input and analysis.

Expectations:

In class: Most class meetings will consist of both lecture and activities and examples exploring statistical ideas and techniques. Please come to class prepared and willing to work during class time and to collaborate with your peers and to ask questions of me. This preparation will not only help you to learn the material and perform well in the course, but it will also produce a much more enjoyable learning environment for all of us. Class attendance is very strongly encouraged, as the in-class activities should prove to be valuable learning experiences. Needless to say, you are responsible for everything presented in class, the sections assigned for reading, and assignments used for formative evaluations.

Outside-of-class: Extensive work for this course, typically involving 8-12 hours per week will be divided among:

- reading and working through sections in the text and other books
- reviewing, completing and correcting your class notes
- working on assignments and writing papers for your projects
• struggling with the computer software  
• preparing for exams

**Integrity:** Central to the purpose and pursuit of any academic community is academic integrity. All members of the Westminster community, including students, faculty, staff, and administrators, are expected to maintain the highest standards of honesty and integrity, in keeping with the philosophy and mission of the College. Academic dishonesty is a profound violation of this code of behavior.

The paragraph above is taken from the Westminster College 2016-2017 Undergraduate Catalog, page 54. It is imperative that you never submit the work of others as though it is your own work nor should you ever allow anyone else to use your work without giving credit to you. The penalty for academic dishonesty in this class is minimally the grade of 0 on the assignment and, except for unusual circumstances, a grade of F for the course. Any event of academic dishonesty is reported to the Dean of the College. Other details of violations and consequences are given in the Catalog.

**Group work is expected.** Within the group work context, it is possible to misunderstand exactly what it means to be responsible for “doing your own work.” As such, I wish to define specifically what I expect. Group study is a proper and effective way to study if all of the participants have done their full share of the work. You may and should discuss problems together and reach conclusions together. But it is a form of dishonesty for a student who has not attended class, read the assignment, or thought about the problem on his or her own to try to use the ideas developed by a group or claim credit for work to which he or she has not contributed. It is also a form of dishonesty to encourage or allow such practices on the part of others.

Aside from integrity issues, writing solutions on your own allows you to determine how much of the material you understand versus how much you can “do.” At some point you are responsible for understanding each and every piece of the problems and will need to be able to write the complete solutions in your own words.

**Assessment:**

**Graded homework:** There will be homework problems from the text assigned daily, some of which will be graded by me and some of which will not be graded. You should expect to hand in carefully written and word-processed solutions (most will be using pieces of the reference guide in R that I give you) to all assigned problems. Each problem will be recorded individually and you should turn in both the Rmarkdown and html output for EACH problem until told otherwise. Your lowest 3 grades will be dropped. **Homework is due before 11:05 each class day. Late assignments receive a 0 (zero) for the grade.** Discussion of the assignments with anyone and everyone is encouraged, but all submitted work must be written independently. In particular, you are not permitted to see the code or the written analysis another student is submitting. Statistics requires substantial practice. I will be assigning a minimal amount of homework. Students working toward a final grade of above a C in this class should expect to do approximately 2.5 times the amount of assigned problems
from the text. Some of this work should be the practice problems as suggested on WileyPLUS course web site.

**Exams:**

- Two unit exams - one at the completion of Unit B approximately Oct. 10, one at the completion of Unit D Nov. 21 (NOTE: Nov. 21st is the day before Thanksgiving break begins. You should inform me by the end of the day on August 29th if you will not be here for this exam.)
- Comprehensive final exam on Monday, Dec. 11 at 3:30 pm.

Makeup examinations procedures as explained on page 60 of the 2016-2017 Undergraduate Catalog will be followed. Students requesting makeup examinations should expect a more difficult examination than that given during the regularly scheduled exam time.

**Projects:** These projects are described in your text at the end of the corresponding sections. More details and due dates will be posted on D2L at least one week before the Project is due.

1. Unit A Project 2 due the following Tuesday after we complete Unit A.
2. Unit A Project 3 due the following Thursday after we complete Unit A.
3. Unit B your choice of one of the following - Project 3 or 4. Due the class period after we complete Unit B.
4. Unit B Project 5. Due the following Tuesday after we complete Unit B.
5. Due the following Tuesday after we complete Unit C.
   (a) (1%) Return to your second project. Complete the inference using (and justifying) the inferential techniques of Unit C.
   (b) (4%) If you did Project 3 in unit B, now do Project 4 from Unit C. If you did Project 4 in unit B now do Project 3 in Unit C.
6. Unit D Project 1 due Nov. 16th (which is before I expect to finish Unit D.) The scholarly article should come from a peer reviewed journal such as *Science* or *Nature.*
Grade Calculations:

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework</td>
<td>10%</td>
</tr>
<tr>
<td>Unit A and B exam</td>
<td>20%</td>
</tr>
<tr>
<td>Unit C and D exam</td>
<td>20%</td>
</tr>
<tr>
<td>Comprehensive final exam</td>
<td>20%</td>
</tr>
<tr>
<td>Projects</td>
<td>5% each</td>
</tr>
</tbody>
</table>

- For every unexcused absence beyond the first one, you will lose 5% from your final percentage. For example, if you have a 82% with 3 unexcused absences your final grade will be recorded as $82 - (3 - 1) \times 5 = 72\%$.

- If you do not earn higher than a 55 on the Final, you will not receive higher than a C- for your final grade regardless of your other grades in the course.

Grade cutoffs will be no higher than A: 93, A-: 90, B+: 87, B: 83, B-: 80, C+: 77, C: 73, C-: 70, D: 60.

See you in class!