ECON 5023: Statistics for Decision Making

UNIVERSITY OF OKLAHOMA

DEPARTMENT OF ECONOMICS

FALL 2020

INSTRUCTOR:	PROFESSOR LE WANG
OFFICE:	VIRTUAL
EMAIL:	LE.WANG@OU.EDU
OFFICE HOURS:	M 2:45-3:45PM OR BY APPOINTMENT
WEB:	HTTP://WWW.LEWANGECON.COM
TA:	CHENG MA (MA0000@OU.EDU)
TA OFFICE HOURS:	TUESDAY AND THURSDAY OR BY APPOINTMENT

1 REMARK

This syllabus is *only* tentative and subject to change. I will update it according to the progress of this course. **It is your responsibility to check Canvas and our course page on Github** `https://github.com/lewangecon/stat20` **constantly for updates.**

It is your responsibility to set up your Canvas correctly to receive notifications and announcements, and you would be the person who bear the consequences of missing any important announcements such as homework deadlines and exam times and locations.

Email is the fastest way to contact me. Please include **ECON 5023** in the subject line.

2 COURSE DESCRIPTION

This course focuses on the *applications* of statistics or econometrics in the decision-making process. A statistical approach to decision making, unlike informal intuition or a qualitative approach, is based on two things: modern computer programs and analysis of large-scale data. The traditional "paper-and-pencil" type of introductory statistics courses usually covers some theoretical concepts and techniques, but fails to include much programming and data analysis, which is nowadays at the heart of "data science". Because of this, emphasis of this course will be placed on combining programming techniques and statistical concepts simultaneously through the analysis of real-life data sets taken from various sources. Our goal is to demonstrate how economists use data analysis to aid decision making and to answer important questions about complex social and economic issues. Throughout the process, we will learn how to programming with R and to manipulate and explore data in R, although this is not an R course.

At the end of the semester, you should gain a basic understanding of how one could apply statistical techniques to analyze micro, macroeconomic, financial events and issues, or even

other social issues. You will learn not only theoretical concepts and tools, but also how to use statistical softwares to analyze real-life data. The 'application' aspect entails becoming familiar with statistical software (R will be heavily used). You will develop competency in R in order to apply the techniques learned. *Learning by Doing*!
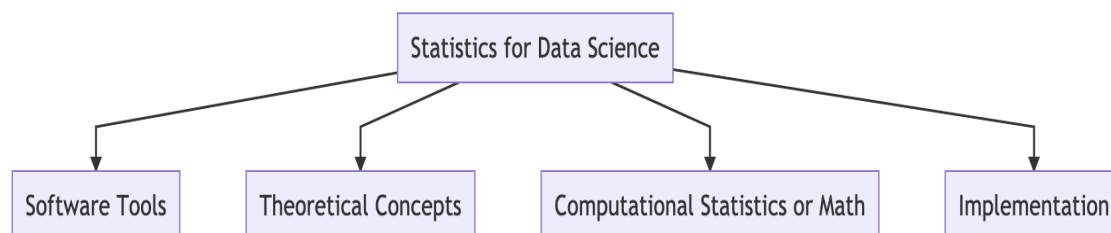
3 Our goals: Three Layers

Our ultimate goal is to provide statistical foundations for further learning of data science and machine learning.

1. We will review **basic concepts**, which have been covered in your previous statistics courses. This is to make sure that everyone is on the same page. But more importantly, we want to take this opportunity to introduce how you can implement these concepts using a programming language and apply them them to some real-life data.

2. We will focus on the "Why" behind the basic concepts to gain **better and deeper** understanding of them. We will provide a unified framework to organize and connect all these concepts and tools, as opposed to "pieces". Throughout the process, we will offer a more rigorous **re-introduction** of some basic concepts, as well as intuition and historical background behind their development. Hopefully such process will help you think about what may be a more appropriate method to apply in a real-life setting, and what you can do differently (with some creativity).

3. We will introduce some **new concepts** or methods. Some can only be introduced when computers are available to us (such as Monte Carlo simulation and Bootstrap), some are more complicated extensions of the basic concepts (e.g., exponential smoothing methods if time permits), and some are concepts more specific to decision making (e.g., criteria for model selections).

A theme that runs through the entire learning process is "hands-on real-life applications' where we will apply what we will learn to simplified yet real-life topics. The examples are taken from different fields to underscore the variety of possible applications of our materials. For example,

1. **Discrete Variables:** Classification and Text Analysis in Machine Learning

2. **Continuous Variables**: Risk Management in Finance

3. **Computational Statistics**: Numerical Integration and Differentiation.

4 Possible Issues in the process

4.1 Conceptual Understanding vs. *Real* Understanding: Importance of Practice

A common scenario (From Barbara Oakley's WSJ Opinion): I just don't see how I could have done so poorly. I understood it when you taught it in class.

That one think he/she understood an idea in a classroom "doesn't necessarily mean that this person *truly* understood the idea. It certainly doesn't mean that the student with retain that idea. And it absolutely doesn't mean that the student has mastered the idea."

"Understanding is key. But not superficial, light-bulb moments of understanding or conceptual understanding. In [a field like this], true and deep understanding comes with the mastery gained through practice."

4.2 Asking and Answering Vague Questions

Sometimes, students are afraid to ask a question especially when he or she cannot formulate it clearly yet. Sometimes, the students complain that the professor sometimes asks vague questions that they do not know how to answer.

4.3 Frustration

"Frustration is natural when you start programming in R, because it is such a stickler for punctuation, and even one character out of place will cause it to complain. But while you should expect to be a little frustrated, take comfort in that it?s both typical and temporary: it happens to everyone, and the only way to get over it is to keep trying." R for Data Science by Garrett Grolemund and Hadley Wickham

4.4 General Solutions: Communication

Lack of (positive and constructive) communication is the biggest obstacle to learning. Lets ensure that this does not happen.

5 Online Learning

My goal for this course is to replicate in-person classes as much as I can in a virtual environment. Our lectures will be live and won't be recorded. Recordings are also not allowed. Just as with in-person classes, missing one lecture will entail some cost, but the impact is minimum. It won't be detrimental to your learning, and one can easily catch up by going through the assigned material, including slides, and the notes that you can borrow from your classes.

No matter what we do, however, online learning will surely be different from in-person learning. Throughout the semester, we may encounter many challenges pertaining to online learning. One of such challenges is that when in a virtual environment where interactions may

not be the same as in-person classes, it may take a different approach to stay engaged. So, I would ask you try your best to avoid distractions and stay engaged.

1. Try to give the Zoom meeting your full attention. Turn off email, text, and other instant notifications to avoid distraction.

2. If you are moving around the room, wrestling with your cat, or need to step away for a second, turn off your video to avoid distracting others.

3. Avoid multi-tasking. You'll retain the discussion better if you refrain from replying to emails or text messages during the meeting and wait to work on that PowerPoint presentation until after the meeting ends.

Throughout, the most important thing is again **communication**. Show understanding and patience to each other, and collectively come up creative solutions to address any challenges that may get in the way of our learning. Hopefully, by addressing them, we may be able to fully exploit some of the advantages of online learning.

6   Required Reading

Note: We do not follow the books closely. Rather, I will use my own structure to organize select topics. As a graduate student, it is your responsibility to find out what would be the corresponding material in the book and independently look for additional readings if needed. Two approaches: look for the index at the back of the book, or Google material on the topic.

Statistical Techniques in Business & Economics by Lind, Marchal, and Wathen (hereafter LMW).

Forecasting: Principles and Practice by Rob J. Hyndman and George Athanasopoulos (hereafter FPP). The latest version of the book is available for free at https://otexts.org/fpp2/. All the examples are written in R.

Introduction to Probability and Statistics Using R by G. Jay Kerns freely available at `http://www.atmos.albany.edu/facstaff/timm/ATM315spring14/R/IPSUR.pdf`

R for Data Science freely available at `https://r4ds.had.co.nz`

The first book provides detailed material on introductory statistics and the applications of these elementary concepts and methods in business and economics, while the second book provides a more specific purpose for statistics and econometrics, namely forecasting. I will also try to supply my slides or provide pointers (additional readings) for the topics that are not covered or in less detail by these two books.

7 Statistical Software and Datacamp

1. R can be downloaded at `http://cran.r-project.org/bin/windows/base/`.

2. RStudio, a powerful IDE for R, can be downloaded at `http://www.rstudio.com/`.

3. R Markdown, `https://help.github.com/articles/about-remote-repositories/`

Bring a laptop to classroom (please email me if that would be a problem). However, you should use it only when we need to use R for practice or a quiz.

1. Datacamp: `https://www.datacamp.com/home`. I also requested free access to many different courses on R, Statistics, and Machine Learning. It is integrated into our course to both enhance and expand your in-class learning. **Send me your OU email account to request access to these materials immediately after the first class**.

Question-Asking Etiquette

1. You need to show your independent work on Google or somewhere else before asking your TA or me. This is also a way to help you develop skills to solve problems and issues independently.

2. When asking about coding, you need to include the exact error message and the code to replicate such message. You cannot simply say that my code does not work.

8 Useful Data Sources

1. Analyze Data for Free `http://www.asdfree.com/`

9 RECOMMENDED READINGS

There is no single textbook that covers every topic well. Here I also list some books that may be useful for certain topics under "Introductory Level", and some books for more advanced and technical treatments under "Advanced Level".

9.1 Machine Learning

1. Sugiyama, Masashi. *Introduction to statistical machine learning*. Morgan Kaufmann, 2015.

2. Robert, C. (2014). *Machine learning, a probabilistic perspective.*

9.2 Introductory Level

1. OpenIntro Statistics 3rd Edition. Available for free at `https://www.openintro.org/stat/textbook.php`

2. *Business Forecasting: Text Alone* by J.Holton Wilson, Barry P. Keating, and Solutions Inc. Publisher: McGraw-Hill Publishing Co.; 5th edition (August 5, 2011) ISBN-10: 0072979674; ISBN-13: 978-0072979671. This textbook contains more detailed discussions of the contents and many useful examples. The software is however dated.

3. Elements of Forecasting by Francis X. Diebold, Cengage Learning; 4 edition (December 8, 2006), ISBN-10: 0324359047

4. *An Introduction to Statistical Learning* by Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, available for free at `http://www-bcf.usc.edu/~gareth/ISL/`

9.3 Advanced Level

1. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* by Trevor Hastie, Robert Tibshirani, Jerome Friedman. available for free at `https://web.stanford.edu/~hastie/ElemStatLearn/`.

2. Introduction to Econometrics, Fourth Edition, Maddala and Lahiri, Wiley: 2009, ISBN: 9780470015124

3. Introductory Econometrics, Fourth Edition: Wooldridge, South-Western: 2009, ISBN: 9780324660545

4. Symposium on Econometric Tools, Journal of Economic Perspectives, Vol. 15, No. 4, Fall 2001

5. Time Series Data Analysis Using EViews, First Edition, Gusti Ngurah Agung, Wiley: 2008, ISBN: 9780470823675

6. Applied Time Series Econometrics, First Edition, by Lukepohl and Krazig, Cambridge: 2004, ISBN: 0521547873

7. *Applied Econometric Time Series, 2e or 3e* by Walter Enders

10 Assignment, Testing, and Grading

10.1 Policy regarding illness and special circumstances

It is expected that there will be students who become ill or face some other special circumstances that prevent them from attending a class, taking a test, or completing a problem set. To accommodate students in such situations, I allow you drop the lowest grades in homework assignments and quizzes as explained below. So, do not worry about re-taking the test or homework assignment you miss due to unexpected events.

Remark: This grading system already takes into account the possibility of missing a quiz or exam *for any reasons*! I do not want such a rare event to adversely impact your grades. That's why I allow you drop a midterm and some quizzes. And because of this system, I do NOT give any make-up exams, as stated above. If you insist on taking a make-up quiz or exam, you would lose the privilege of dropping the lowest one.

10.2   Tests

There will be one midterm (**tentative date: 10/21/2020**) and one final exam. The format of these exams will be announced later. **No make-up exams will be given.** If a student misses a midterm examination for **any reason**, the weight of that examination will be added to that of the final examination (e.g., missing Midterm I would make the final worth 50% of the course grade). I'd like to have the flexibility to speed up or slow down depending on how I think the class understands the material. So, midterms will be announced at least one week in advance.

10.3   Homework Assignment and R Markdown

There will be homework assignments following each topic. The total number of homework assignments will depend on the progress of the class. **Homework assignments will be due at the beginning of class and late homework assignments will not be graded for credit.** It is important to know that the homework assignments are very important in that the basic ideas covered by them invariably show up on the mid-term and final exams. If you know you are going to be missing a class on the day a homework exercise is due, hand in your homework **in advance** to receive full credit for your work.

- You should try to type all the homework using R Markdown. This is also courteous to your TA who would be grading the homework.

- Group discussion is encouraged when working on the problem sets. Your answers should show individual understanding of the materials and be written in your own words. If you collaborate with other student(s), everyone in your group should be acknowledged at the beginning of the homework. Identical homework assignments would receive a zero for both students.

10.4   Quick Quizzes

There will also be many short quizzes (short answer and/or multiple-choice) given throughout the semester. Such quizzes are not announced beforehand and will usually be administered in the first five to ten minutes of the class. These quizzes are designed to help you understand better the materials covered and find out whether you have any problems with the materials and have done any assigned readings that I may have asked you to do. **No make-up quizzes will be given**. If you must miss a class due to legitimate circumstances beyond your control, be sure and contact me **beforehand** so that I will know of your circumstances. If excused, I will correspondingly excuse you from any QQ that is given that day. The lowest quiz grade will also be dropped.

10.5   Mistakes in Grading

Your TA and myself are only human. Sometimes it is possible that your TA makes some mistakes in grading or misplaces your homework, quiz, or exam papers. If that occurs, please report the incidence within **three** days after they are handed back and recorded on Canvas; after that, I would not be able to make any changes except in special circumstances such as illness that prevent you from doing so. Discuss first with your TA about your concerns or objections, and if it is not resolved, then present a written document to me for final decision on the matter.

10.6   Grading

The weights in the final grade are assigned as follows:

Homework Assignments (including Datacamp exercises):	20
Midterm I:	25
Final	35
Quick Quizzes	15
Participation Credits: Daily Responses	5

The total points are capped at 100. The grading scale is:

90-100	A
80-90)	B
70-80)	C
60-70)	D
0-60)	F

A number with a parenthesis means any number less than it. For example 90) means any number less than 90, but not equal to 90. It could be 89.9999999. I reserve the right to change the grading scale. The required score for the final grade could be lower but not higher. For example, the current requirement for an A is 90 - 100, but later I may change the cut-off point to 85-100, but won't change it to 97-100.

11   University Policies

11.1   Accommodations for Students with Disabilities

If you are a student with a documented disability who will require accommodations in this course, please register with the Disability Resource Center (Goddard Health Center, Room 166, 325-3852). Students who are already registered with the Office of Disability Services and wish to receive accommodations in this course are strongly encouraged to share their Accommodation Letter with me in a timely manner so I can provide an appropriate contact to discuss accommodations necessary to ensure full participation and facilitate your educational opportunities. Students with disabilities must be registered with the Disability Resource Center before receiving academic adjustments.

11.2 Academic Honesty

Cheating is strictly prohibited at the University of Oklahoma, because it devalues the degree you are working hard to get. As a member of the OU community it is your responsibility to protect your educational investment by knowing and following the rules. For specific definitions on what constitutes cheating, review the Student's Guide to Academic Integrity at `http://integrity.ou.edu/students.html`.

11.3 Religious Observance

It is the policy of the University to excuse the absences of students that result from religious observances and to reschedule examinations and additional required classwork that may fall on religious holidays, without penalty.

11.4 Title IX Resources and Reporting Requirement

For any concerns regarding gender-based discrimination, sexual harassment, sexual assault, dating/domestic violence, or stalking, the University offers a variety of resources. To learn more or to report an incident, please contact the Sexual Misconduct Office at 405/325-2215 (8 to 5, M-F) or smo@ou.edu. Incidents can also be reported confidentially to OU Advocates at 405/615-0013 (phones are answered 24 hours a day, 7 days a week). Also, please be advised that a professor/GA/TA is required to report instances of sexual harassment, sexual assault, or discrimination to the Sexual Misconduct Office. Inquiries regarding non-discrimination policies may be directed to: Bobby J. Mason, University Equal Opportunity Officer and Title IX Coordinator at 405/325-3546 or bjm@ou.edu. For more information, visit `http://www.ou.edu/eoo.html`.

11.5 Adjustments for Pregnancy/Childbirth Related Issues

Should you need modifications or adjustments to your course requirements because of documented pregnancy-related or childbirth-related issues, please contact your professor or the Disability Resource Center at 405/325-3852 as soon as possible. Also, see `http://www.ou.edu/eoo/faqs/pregnancy-faqs.html` for answers to commonly asked questions.