ECON 5043: Managerial Economics II

Machine Learning and Causal Inference

University of Oklahoma
Department of Economics

Spring 2022

Instructor:	Professor Le Wang
Email:	le.wang@ou.edu
Office Hours:	13:30-14:30pm (Monday)
Web:	http://www.lewangecon.com
TA:	Cheng Ma (Email: ma0000@ou.edu)
TA Office Hours:	TR 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 Github constantly for updates.** There may be major changes regarding the course such as grading policy and exams if deemed necessary, but they will be explicitly announced in class at least one month ahead.

Email is the fastest way to contact me.

2 Course Description

Managerial economics is primarily concerned with the applications of economic methodologies to key management decisions within organizations such as government and business. Virtually all decisions depend on forecasts. A quantitative approach to forecasting based on modern computer programs, as opposed to informal intuition or a qualitative approach, has proven to deliver more accurate forecasts, thereby leading to more effective plans and decision making. The quantitative approach is indeed what we will focus on in this course.

Our course consists of two parts. The first part is focused on tools for predictive analytics such as regression and logit/probit models, while the second part is focused on tools for program (or policy) evaluation. This distinction reflects different needs of decision making: business decisions are often based on the predictions of certain outcome variables, while government decisions often involve understanding the impacts of certain policy or programs on certain outcomes. For the former, what actually determines the outcome variables to be forecasted and why are not necessarily important; what matters is a precise forecast. For the second, we are actually interested in whether a program has any *causal* impacts on the outcomes of

interest. These differences reflect the subtle differences between *correlation* and *causality* (cause-effect), which is often mentioned in the undergraduate statistics/econometrics courses, but not emphasized or carefully examined. Although these two set of tools are presented separately, they are closely related. Specifically, the first set of methods (which themselves are built on the concepts and tools learned in Econ 5023) are the workhorse models upon which the second set of methods are built.

The world has changed rapidly in the past decades. One significant change is the availability of large or big datasets. These data are big not only because the sample size (N) is big but also because the number of dimensions (p) is large and often even larger than the sample size. The traditional tools, especially the conventional regression and logit/probit models, may not be adequate to handle such data (for example, the regression estimates are not unique when $p > N$). To this end, we will also discuss some extensions of these methods that could handle high-dimensional datasets, if time permits.

Our presentations of these methods are intended to be as general as possible. In-class examples of these techniques are also not necessary only in the business or government sectors; instead, a wide range of situations are examined. It is my hope that this way you can see and understand the generality of these techniques and apply them in the situations where others may not necessarily make the connections.

It is, again worth mentioning that, just as Econ 5023 (Statistics for Data Science), this is an applied course; you will learn not only theoretical concepts and tools, but also how to use R to analyze real-life data. R will again be heavily used. You should develop competency in R in order to apply the techniques learned *Learning by Doing*!

3 Required Readings

Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, An Introduction to Statistical Learning with Applications in R, freely available at `http://www-bcf.usc.edu/~gareth/ISL/ISLR%20Sixth%20Printing.pdf`.

Joshua D. Angrist and Jorn-Steffen Pischke, Mastering Metrics, Princeton University Press. This book provides some intuitive explanations of the techniques on program evaluation. (A slightly more formal and rigorous treatment can be found in their book "Mostly Harmless Econometrics").

4 Statistical Software

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/`.

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.

5 Assignment, Testing, and Grading

5.1 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 some quizzes and homework and move the weight of the midterm to the final. And because of this system, I do NOT give any make-up exams/quiz/homework, as stated above. If you insist on taking a make-up quiz or exam, you would lose the privilege of dropping the lowest one.

5.2 Tests

There will be one midterm (**tentative date: 03/09/2022**) and one final exam (May 11, 2022, W 4:30 - 6:30 pm, Cate Center One 0326). 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 would make the final worth 50% of the course grade in the case of two midterms). I'd like to have the flexibility to speed up or slow down depending on how I think the class understands the material. But should there be any changes to the schedule, it would be announced two weeks in advance.

5.3 Homework Assignment

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. 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.

5.4 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.

5.5 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.

5.6 Grading

The weights in the final grade are assigned as follows:

Homework Assignments:	25
Midterm I:	25
Final	35
Quick Quizzes	15

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.

6 University Policies

6.1 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.

6.2 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`.

6.3 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.

6.4 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`.

6.5 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.

7 Tentative Course Outline

Note that the schedule is subject to change depending on the pace of the course. Not all material would be covered, and extensions of the basic models not listed below may be added to deepen our understanding of these techniques. Moreover, the material will not be necessarily covered in the same order as below.

7.1 First Part: Statistics and Machine Learning

1. Basic Multivariate Statistics: Marginal, Joint, and Conditional Distributions and Independence
2. Measures of Linear Relations
3. Conditional Distribution
4. Conditional Expectation (Mean)
5. Linear Regression
6. Linear Model Selection and Regularization
7. Maximum Likelihood Estimation and Logit and Probit Models (Optional)

7.2 Second Part: Causal Inference

1. Causal Inference, Potential Outcome Frameworks, and Parameters of Interest
2. Randomization
3. Regression and Causal Inference
4. Matching
5. Regression Discontinuity Designs
6. Differences-in-Differences and Synthetic Control Approach (Optional)
7. Quantile Regression (Optional)