MAST 6252 - Applied Probability Models
Spring 2021 Mod B
Prof. Michael Braun

Faculty Information
Professor Michael Braun
braunm@smu.edu
Office hours by appointment at
www.calendly.com/braunm

Course Information
Sec. 721: Wednesdays, 6:30pm– 9:45pm
Sec. 021: Thursdays, 8:45am – 12:00pm
Sec. 121: Thursdays, 8:45am – 12:00pm
All class sessions are online.
See Canvas for Zoom links.

Course description

Forecasting customer behavior is a critically important activity for firms. Recent advances in statistics have led to a number of analytical methods that are not only effective in explaining and predicting patterns in customer behavior, but are also easy to implement by a practicing manager. These models use the basic building blocks from probability theory to offer behaviorally plausible perspectives on what people buy, when they buy it, and how much they buy. As more customer-focused data becomes available, the specification and interpretation of these models can become a regular part of the sophisticated manager's toolkit.

The principal focus of this course is on inferring unobserved "types" of customers, and using these inferences to predict what they will do in the future. This is a "hands-on" course, in which students learn to apply probability models to a wide variety of business applications, such as predicting customer retention patterns, computing long-term expected customer value, targeting marketing activities based on latent customer types, and forecasting adoption of new products.

While data-focused students and employees aim to expand their computer programming skills, there is an increasing reliance on methods and algorithms treated like a "black box" without looking inside. But this course will require students to open up the black box. One advantage of studying probability models in pedagogical: they are clear, parsimonious, and easy to estimate, even in a spreadsheet. This allows us to emphasize model building instead of programming, focusing students to think about the relationship between real-world customer behavior patterns in data and model components. The purpose is to understand the behavioral story behind models, how to interpret them, how they work, how to best diagnose their fit of data, and how to know when to use which models. By building these models from scratch, students will not only learn how to use these models effectively, but also why they work so much better than older, more established statistical tools. The emphasis is on

developing good modeling technique, as opposed to manipulating large datasets. This experience creates a more general understanding of how to work with real data, and empowers students to examine alternative interpretations of data with an informed and critical eye.

Learning outcomes and objectives

By the end of the course, students will have learned the following skills (this is not an exhaustive list).

- Students will learn how to think about data generatively, to represent data-generating processes as probability models, and to derive such models that are appropriate to capture customer activity.
- Students will be able to identify a business problem, to choose appropriate data to acquire, and to decide which model(s) to apply to those data.
- Students will practice how to project customer retention / churn and repeat transactions while accounting for differences in customers.
- Students will practice how to compute customer lifetime value (CLV), will how to critique inappropriate calculations of customer value, and will understand the relationship between relationship CLV and several related concepts.
- Students will learn how to choose among model specifications, using both quantitative and qualitative criteria, and will practice defending those choices.
- Students will apply Bayesian methods for inference, classification and forecasting for multiple business applications.

Class structure

Most of the classes will be lecture-based, with a strong emphasis on problem solving, analytical exercises and analysis of datasets. Central to the development of the skills associated with probability modeling is hands-on experience. This is not one of those "30,000-foot-view" survey classes.

We will be using **Piazza** for class discussion and out-of-class learning. Piazza is a free Q&A platform designed to get you great answers from classmates and instructors fast. Think of Piazza as a Q&A wiki for your class. Every question has just a single "students' answer" that students can edit collectively, and a single "instructors' answer" for instructors. The class Piazza page is at:

`https://piazza.com/smu/spring2021/mast6252/home`.

You will receive an email from Piazza instructing you on how to set up your account.

Prerequisites

This course is required of all students who are enrolled in the MS in Business Analytics program. Students are expected to have successfully completed the required statistics and data analytics courses.

Although calculus is not a prerequisite, we will nevertheless refer to integrals and derivatives throughout the course. You will not need to compute derivatives or integrals on your own, but you may find it useful to review basic concepts in univariate calculus before the start of the course. If you want to read a good conceptual review, I recommend *Calculus Made Easy* by Silvanus Thompson, which has survived over 100 years as a masterpiece in mathematical pedagogy. Originally published in 1910, the book is now available for free via Project Gutenberg. The link is: https://www.gutenberg.org/files/33283/33283-pdf.pdf. [1]

Required materials

There is no required textbook for this course, mainly because no suitable book exists. The application of probability models to marketing problems is still a relatively new and evolving field, and much of what you will learn in this class is truly "cutting edge" work. But you will be able to rely on lecture slides, other lecture notes, the required and optional readings, and, of course, the professor.

Many of the concepts, models and techniques that we discuss in this class are explained in published articles in-progress working papers, and purpose-written notes. These documents will be posted on Canvas. I have divided the readings into three groups: Required, Recommended, and Optional. The required readings explain the models that we covered in class. If a reading is flagged as required, interpret that as "the professor thinks that reading this will make it easier for me to learn the material that was covered in class." In fact, some of the required readings give you step-by-step instructions on how to build certain models. The required readings do not introduce new material that is not otherwise covered in class. They are meant to be helpful, not burdensome.

Some of the recommended and optional readings do the same thing, but they are not required because they are written at a technical level that is above that of the class prerequisites. Recommended readings are either somewhat easier, or somewhat more important, and the optional readings. You may find that reading the introductions of the recommended or optional papers, studying the motivating problems, and trying to follow along as much as you can, is a good learning exercise. The other kind of recommended or optional reading is one that illustrates an interesting, practical application of a particular model.

Assignments and grading

The final grade will be computed from the following components, with the associated weighting.

[1] Don't worry about the equations. Focus on the words and the diagrams. The parts that are most important for this class are the Prologue, Chapters, 1-4, Chapters 10-12, and, most importantly, Chapters 17-19. Really, Chapters 17-19 are key because they discuss integration, but you might need to skim some earlier chapters for it all to make sense. And it's the concepts that matter, not the mathematical manipulation. You won't have to solve the integrals directly, but you will need to understand why they are there, and know how to manipulate them a little bit. Ignore anything and everything related to trigonometry, differential equations, integration by parts, or partial fractions.

- Homework assignments: 40%
- Final Exam: 40%
- Participation and attendance: 20%

Homework assignments

There will be three homework assignments, which are to be completed individually. Each will involve analyzing a dataset using the tools and models you will learn in this class. Grading criteria will include not only getting the correct numerical results, but also answering questions about the interpretation and managerial implications of those results. Instructions for submitting assignments will be announced in class and on Canvas.

Assignments are due on Wednesday at 6:00pm for all sections, even those meeting on Thursday.

Assignment	Due date
Assignment 1	March 24, 2021
Assignment 2	April 7, 2021
Assignment 3	April 21, 2021

Attendance and participation

If the class were being taught in person, this section would read:

> Attendance is required. If you must miss class, you should make arrangements to get notes from another student. There will be topics covered in class that are not available in readings or notes, so attendance is strongly recommended. You do not need to tell me in advance if you are going to miss class. You should attend the section in which you are registered. However, if you have an unavoidable conflict, you may attend the other section.

Because the class is being taught online this year, the guidelines are slightly different, but the general principle remains. Although class sessions are recorded and accessible through Panopto, you are better off attending class live. Other than the grading component, there is one major, overriding reason to attend the live class sessions: the ability to ask questions. I am always happy to stop, slow down, and/or repeat anything I say. In fact, I start to worry if there are no questions at all. So please attend the live class sessions.

Note that participation and attendance are not the same thing. I will try to stimulate discussion in class, and I expect everyone to be engaged in the joint learning experience.

This portion of the grade will also take into account completion of pre-class "preparation questions," and/or very short online quizzes. The purpose of these assignments will be to check your progress, and to help us make the best use of our short amount of time together in class.

Final exam

Following SMU policy, the final exam will be taken online, through Canvas. The date and time of the exam will be announced by the Cox Graduate Programs Office. Everyone will take the exam at the same time. Requests to take final exams at different times are rarely granted. The exam will be a combination of multiple choice and short answer questions, as well as some mathematical exercises.

A word of advice for the exam: it is not meant to be "tricky." The objective of the exam is to give students a chance to synthesize all of the course material, and to identify those who know what is going on from those who do not. Very few questions will ask you to regurgitate facts that you have memorized. It is more likely you will be asked to apply concepts of the course to a new task or situation. Students who successfully complete assignments independently, study with classmates throughout the term, review solution sets and learn from mistakes, complete optional or ungraded problem sets, ask for help when needed, and otherwise keep up with the course, tend to do well. Ignoring the professor's advice mand taking shortcuts are examples of reasons students have done poorly.

Other notes regarding grading

Other than specific issues that may necessitate minor changes during the course term, the course requirements are set, and are the same for everyone. There will be no "extra credit" opportunities.

There is no set grading scale for the course (i.e., there is no range of points that corresponds to a particular grade). In a well-taught graduate course, the quantity and difficulty of course material will be a challenge for even the strongest students. You may find that numerical grades on specific assignments may be lower than what you are used to. However, final course grades will reflect this. In the past, the median final grade has been around a B+, and fewer than 10 percent earned grades of C or worse. However, this curve is not guaranteed, and may deviate in either direction. That is, I may decide to issue more A's if students do well, and fewer if not.

Asking questions and getting help

I care a lot about teaching, and I want you to make the most of your Cox educational experience. You should always feel comfortable asking questions in class. If there is something you don't understand, your classmates might be just as confused. So stop me and ask! It is possible (probable?) that I did not explain something as well as I could, so your question gives me an opportunity to try again.

An effective learning strategy in past years has been to review lecture notes, and to replicate the examples from class, that same night. If you can't get the same answers from class, then you will know very quickly what you might not quite understand, and can ask for help.

Outside of class, **all questions about course material must be posted to Piazza**. The sys-

tem is designed for getting you help fast and efficiently from me and your classmates. The sooner you begin asking questions on Piazza (rather than via emails), the sooner you will benefit from the collective knowledge of your classmates and instructor. I encourage you to ask questions when you are struggling to understand a concept—you can even do so anonymously. **I will not respond to emails about course content or logistics. You must post questions to Piazza.** Of course, you may send a private email if your question is of a personal nature (e.g., grades, special circumstances, etc).

I am also available to meet with students in person. If you want to meet with me, you should schedule a time at https://www.calendly.com/braunm. I can always make arrangements to meet with students, although I may not be available the exact time you want.

You should never feel that asking for a meeting is an imposition. I really enjoy helping students. However, given the number of students across all sections of the course, I ask that you try getting help from Piazza first.

Calendar

The list of topics is current at the time of the writing of the syllabus, but may change during the course.

Week 1: Mar. 17/18: Modeling customer retention

- Motivating example: forecasting customer retention for Blue Apron
- Using regression models to model customer behavior (and why you probably shouldn't do it).
- Introduction to discrete timing models and the geometric distribution.
- Understanding unobserved heterogeneity in customer types.
- The "ruse of heterogeneity" and the illusion of increasing customer retention rates.

Week 2: Mar. 24/25: Customer Lifetime Value

Assignment 1 due

- Motivating example: How much are customers worth to me, anyway?
- The idea of lifetime value as a stream of discounted cash flows.
- Incorporating uncertainty about retention rates into ECLV estimates.
- The "perils of ignoring heterogeneity," and how doing just that can cost you a lot of money.

Week 3. Mar 31/Apr 1.: Bayesian Inference and Customer Targeting

- Motivating example: targeting segments in a direct marketing campaign
- The beta-binomial model.
- The importance of Bayesian inference in incorporating information about past customer behavior.

- Estimating segment-specific propensities to purchase.
- Introduction to conditional probability and Bayesian inference.
- Discussion of prior and posterior distributions, and their managerial implications.
- Statistical shrinkage, regression to the mean, and pooling of information.

Week 4. Apr. 7/8: Retention and Valuation for Non-contractual Businesses

Assignment 2 due

- Motivating examples:
  - Forecasting donations to a public radio station
- Using the BG-BB model to estimate the residual lifetime value of existing customers.
- "Contractual" vs "noncontractual" lifetime value settings.

Week 5. Apr. 14/15: Counting processes

- Motivating examples: forecasting media metrics.
- Modeling count data.
- Capturing customer heterogeneity using the gamma distribution.
- Derivation of the negative binomial distribution (NBD), and how to use it as a powerful predictive and descriptive managerial tool.
- Incorporating covariate information into models.

Week 6. Apr. 21/22: Timing processes

Assignment 3 due

- Modeling adoption of new products using data from controlled test markets.

Policies

The listed policies are not exhaustive. Any issues not listed here are covered by standing Cox and SMU policies.

Honor Code

Your work is governed by the Cox School of Business Honor Code. A violation of this honor code is defined as giving or receiving aid on academic work submitted for evaluation without the consent of the professor.

Disability Accommodations

Students who need academic accommodations for a disability must first contact Ms. Rebecca Marin, Coordinator, Services for Students with Disabilities (214-768-4557) to verify the disability and establish eligibility for accommodations. They should then schedule an appointment with the professor to make appropriate arrangements. (See University Policy No. 2.4.)

Accommodations for pregnant and parenting student

Under Title IX students who are pregnant or parenting may request academic adjustments by contacting Elsie Johnson (elsiej@smu.edu) in the Office of the Dean of Students, or by calling 214-768-4564. Students seeking assistance must schedule an appointment with their professors as early as possible, present a letter from the Office of the Dean of Students, and make appropriate arrangements. Please note that academic adjustments are not retroactive and, when feasible, require advance notice to implement.

Covid-19 Attendance Statement

Students who are experiencing COVID-19 symptoms or who have been notified through contact tracing of potential exposure and need to self-quarantine or isolate must follow the protocols laid out in SMU's Contact Tracing Protocol. Because this course is taught entirely online, attendance and participation are expected to continue. Situations for which this is not possible will be handled on a case-by-case basis.

Religious Observance

Religiously observant students who wish to be absent on holidays that require missing class should notify their professors in writing at the beginning of the semester, and should discuss with them, in advance, acceptable ways of making up any work missed because of the absence. (See University Policy No. 1.9.)

Excused Absences for University Extracurricular Activities

Students participating in an officially sanctioned, scheduled University extracurricular activity will be given the opportunity to make up class assignments or other graded assignments missed as a result of their participation. It is the responsibility of the student to make arrangements with the instructor prior to any missed scheduled examination or other missed assignment for making up the work. (University Undergraduate Catalogue)

Sexual Harassment

All forms of sexual harassment including sexual assault, dating violence, domestic violence and stalking are violations of SMU's Title IX Sexual Harassment Policy and may also violate Texas law. Students who wish to file a complaint or receive more information about the grievance process may contact Samantha Thomas, SMU's Title IX Coordinator, at accessequity@smu.edu or 214-768-3601. Please note that faculty are mandatory reporters. If students notify faculty of sexual harassment, they must report it to the Title IX Coordinator. For more information about sexual harassment including resources available to assist students, please visit www.smu.edu/sexualmisconduct.