MKTG/STAT 476/776:
Applied Probability Models in Marketing

Spring 2018 (Monday/Tuesday/Wednesday 3-6PM)

Professor Peter Fader and TAs: Elliot Oblander (head TA), Danielle Clanaman, Jordan Farrer, Gautam Kamath, Ajjit Narayanan (group email: mktg476776ta@wharton.upenn.edu)

Motivations and Objectives

Over the past five decades, statisticians have developed a number of models that have proven to be highly effective in their ability to explain and predict empirical patterns within many areas in business and the social sciences. These models use some basic "building blocks" from probability theory to offer behaviorally plausible perspectives on different types of timing, counting, and choice processes. Researchers in marketing have actively contributed to (and benefited from) these models for a wide variety of applications, such as new product sales forecasting, analyses of media usage, targeted marketing programs, estimation of customer lifetime value, and even overall corporate valuation. Other disciplines have seen equally broad utilization of these techniques.

As new forms of information technology provide increasingly rich descriptions of individual-level shopping/purchasing behavior, these models offer great value to practicing managers, particularly those interested in pursuing CRM ("customer relationship management") activities. Furthermore, as more managers become comfortable with non-linear optimization techniques (using, for example, the "Solver" feature within Microsoft Excel), the specification and interpretation of these models can become a regular part of the sophisticated manager's toolkit. Taken as a whole, the methodological approaches covered in this course are well-suited to address the types of questions that are being asked with increasing frequency and interest by investors and managers of today's data-intensive businesses.

The principal objectives of this course are:

- To familiarize students with probability models and their role in marketing, information systems, supply chain management, actuarial science, operations research, public policy, and other related areas.

- To provide students with the analytical and empirical skills required to develop probability models and apply them to problems of genuine managerial interest.

- To have students develop good instincts to judge the appropriateness, performance, and value of different kinds of models in a variety of managerial settings.

- To encourage students to think critically about statistical methods and managerial perspectives that are common in certain domains but not always the best ways to approach all data-oriented decision problems.

Prerequisites

This course is open to students at any level (undergraduate, MBA, PhD) who have sufficient mathematical skills to handle the sophisticated methods that will be introduced and featured here. It is essential that students have some familiarity with basic integral calculus. Furthermore, a mid-level probability/statistics course would be helpful, but aptitude to learn and fully understand the methods covered in such a course is far more important than mere exposure to them. Finally, there is no need to have taken any marketing (or business) courses before this one.

Smart and highly motivated students are encouraged to take the course sooner (e.g., sophomore year or first-year MBA) rather than later. This course provides a nice foundation for other more advanced modeling courses that can and should be taken after this one.

Course Organization and Materials

Every session will be lecture-based, with a strong emphasis on real-time problem solving, including mathematical derivations and numerical investigations using Microsoft Excel. Central to the development of the skills associated with probability modeling is hands-on experience. To this end, a set of homework exercises will be assigned for most sessions.

There is no formal textbook for the course (since no suitable book exists), but lecture notes covering most of the material presented in class will be made available immediately after each session on Canvas (https://canvas.upenn.edu/courses/1379877). Most of the Excel spreadsheets used in class will be made available to the students, and some journal articles will be suggested as illustrations/applications of the techniques discussed. But most of these readings are just recommended – there will be no formal pre- or post-class reading assignments for any session.

Teaching Approach

The methods covered in this course will be quite unfamiliar to most students at the start of the semester. As such, it is important to ensure that the first exposure is impactful and that there are opportunities to work with the materials multiple times and through multiple formats. To make this possible we will utilize a fairly unique "heads up" learning system in the classroom. The basic elements include:

- Strongly recommended (but never mandatory) classroom attendance
- The use of laptops in the classroom is strongly discouraged
- Presentation decks, spreadsheets, etc., will not be provided until after class
- Each session will be recorded and made available to students (in a multimedia format) soon after each class session

These steps are intended to help students keep their "heads up" to focus on the main points in each session. Students are encouraged to ask questions about key conceptual issues, managerial applications, and the overall modeling philosophy; however, questions about more minor technical issues should be addressed by reviewing the presentation decks and recordings after class (and interacting with the TAs and other students).

Students are expected to create their own complete set of class notes after attending each session and working through the decks/recordings. It is fine for students to collaborate on this task, but it's best for each student to create his/her own notes. Any kind of "divide and conquer" approach will be counterproductive for the student (particularly with regard to the final exam).

Evaluation and Grading

Homework (5%): These exercises will be both analytical and numerical in nature. It is OK for students to communicate about specific problems, but each student must write up each problem independently and hand in their own work. Completed assignments must be uploaded to Canvas; hard copy will not be accepted.

Class Participation (15% of final grade): Although there are no formal case discussions, students are expected to be actively engaged in the lectures, which will include frequent "cold calls" to ensure that everyone is following (and participating in) the conversation.

Project 1 (25%): For the first paper, students will be asked to find a specific type of dataset and analyze it carefully. Papers will be graded using an innovative collaborative platform, the Wharton Online Ordinal Peer Performance Evaluation Engine (WHOOPPEE). Details about the assignment and grading process will be discussed in class.

Project 2 (25%): The second paper will be more standardized – all students will be given a common dataset to analyze (and WHOOPPEE will be used again for grading).

Final Exam (30%): The final exam (date TBA) will be a structured set of questions to assess students' conceptual understanding of the course material. It will not require any detailed mathematical derivations or extensive numerical calculations. It will be a closed-book exam but students can bring a one-page (two-sided) "cheat sheet" as a reference.

All relevant University of Pennsylvania policies regarding academic integrity must be followed. Students may not submit work that has been prepared by (or in conjunction with) someone else, without explicit instructor permission. Any student who in any way misrepresents somebody else's work as their own will face severe disciplinary consequences.

Attending Different Sections

The three sections of the course are basically identical and interchangeable: the same material will be covered in each one. It is fine if students want to switch sections from week to week, and there is no need to ask (or notify) anyone in advance. I encourage students to attend any of the sections but watch the videos from one (or both) of the sessions that they didn't attend – slight differences from one section to another can be a helpful way to learn the material better.

Course Schedule

Note that the university's academic calendar treats Wednesday 1/10 as a Monday, so Week #1 consists of 1/10, 1/16, and 1/17. Then the regular schedule will begin on January 22.

Week 1 ("M" 1/10, T 1/16, W 1/17): Introduction to probability models

Motivating problem: Forecasting customer retention. Comparisons to traditional regression-based models: "curve-fitting" vs. "model-building." Careful derivation of a parametric mixture model (the beta-geometric). Coverage of maximum likelihood estimation and the Microsoft Excel Solver tool. Discussion about the philosophy and objectives of probability modeling.

Week "1A" (Date/time TBA): Math/stat review

Optional Q&A session on the basic calculus, probability, and statistics issues covered in Week 1.

Week 2 (M 1/22, T 1/23, W 1/24): Models for count data

Introduction to the Poisson process and its extension to the negative binomial distribution. Evaluating goodness-of-fit. Generalizing the model to allow for "spikes" at 0 or 1.

Week 3 (M 1/29, T 1/30, W 1/31): More on count models

Alternative estimation approaches for count models ("Means and zeroes" and "method of moments"). Dealing with problems of limited/missing data: truncated and shifted NBD models.

Week 4 (M 2/5, T 2/6, W 2/7): Repeated Choice Processes and Empirical Bayes methods

Choice vs. counting. The binomial distribution. The beta distribution as a mixture model. Parameter estimation and inference. Conditional distributions and expectations. Combining population information ("priors") with observed data for individuals. Regression-to-the-mean.

Week 5 (M 2/12, T 2/13, W 2/14): Timing models

Motivating problem: forecasting new product adoption. Implementing and evaluating different timing models, particularly the exponential-gamma. Dealing with grouped data and right censoring. Introducing hazard functions. Derivation and discussion of other timing models (e.g., Weibull), and the linkages among them. Exploring the interplay between timing and counting processes.

Week 6 (M 2/19, T 2/20, W 2/21): Customer-base analysis
Project 1 due

Combining the basic building blocks to create integrated models to estimate customer lifetime value and related concepts.

Week 7 (M 2/26, T 2/27, W 2/28): Customer-base analysis (cont.)

More CLV-oriented applications.

Spring Break – no class (M 3/5, T 3/6, W 3/7)

Week 8 (M 3/12, T 3/13, W 3/14): Introducing covariates

Poisson regression and NBD regression for counting models. Proportional hazard methods and covariate effects for timing models. General discussion about the different role of covariates from the perspective of an econometrician vis-à-vis a probability modeler. Applications.

Week 9 (M 3/19, T 3/20, W 3/21): Finite mixture and latent class methods

Looking at non-parametric (discrete) approaches to capturing heterogeneity. Interpreting support points versus cluster characteristics. Estimation issues. Overview of selection criteria for non-nested models.

Week 10 (M 3/26. T 3/27, W 3/28): Multi-item choice models

The multinomial choice process and the Dirichlet mixing distribution. Interplay between the beta and Dirichlet distributions.

Week 11 (M 4/2, T 4/3, W 4/4): Fun with Dirichlet!
Project 2 due

Further examination of the Dirichlet-multinomial choice model and its astonishing patterns. Discussion of Ehrenberg's "empirical laws."

Week 12 (M 4/9, T 4/10, W 4/11): Integrated models

Combined models of counting, timing, and/or choice. Particular focus on the BB/NBD as a working example.

Week 13 (M 4/16, T 4/17, W 4/18): Nonstationary processes

Overview and comparison of techniques such as renewal processes, learning models, hidden Markov methods, and other approaches to capture dynamics over time

Week 14 (M 4/23, T 4/24, W 4/25): Applications and wrap-up

More applications/extensions, preparation for final exam, and final comments, questions, and discussion about probability models overall