Recommended Courses for the Analytics-Focused Student

BSTA 526/527
Biostatistics for Epidemiologic Methods

The first half of this will cover graphical methods, probability, discrete and continuous distributions, estimation, confidence intervals, and one sample hypothesis testing. Emphasis is placed on understanding the proper application and interpretation of the methods. The second half of this course will cover two sample hypothesis testing, nonparametric techniques, sample size determination, correlation, regression, analysis of variance, and analysis of covariance. Emphasis is placed on understanding the proper application and underlying assumptions of the methods presented. Laboratory sessions focus on the use of the STATA statistical package and applications to clinical data. (527) The first half of this covers concepts in biostatistics as applied to epidemiology, primarily categorical data analysis, analysis of case-control, cross-sectional, cohort studies, and clinical trials. Topics include simple analysis of epidemiologic measures of effect; stratified analysis; confounding; interaction, the use of matching, and sample size determination. The second half of this course covers concepts in biostatistics as applied to epidemiology, primarily multivariable models in epidemiology for analyzing case-control, cross-sectional, cohort studies, and clinical trials. Topics include logistic, conditional logistics, and Poisson regression methods; simple survival analyses including Cox regression. Emphasis is placed on understanding the proper application and underlying assumptions of the methods presented. Laboratory sessions focus on the use of the STATA statistical package and applications to clinical data.

BSTA 630/631
Statistical Methods for Data Analysis I/II
Shults and Putt, Gimmony

This first course in statistical methods for data analysis is aimed at first year Biostatistics degree candidates. It focuses on the analysis of continuous data, and includes descriptive statistics, such as central tendencies, dispersion measures, shapes of a distribution, graphical representations of distributions, transformations, and testing for goodness of fit for a distribution. Populations, samples, hypotheses of differences and equivalence, and errors will be defined. One and two sample t-tests, analysis of variance, correlation, as well as non-parametric tests and correlations will be covered. Estimation, including confidence intervals, and robust methods will be discussed. The relationship between outcome variables and explanatory variables will be examined via regression analysis, including single linear regression, multiple regression, model fitting and testing, partial correlation, residuals, multicollinearity. Examples of medical and biologic data will be used throughout the course, and use of computer software demonstrated. This is the second half of the methods sequence and focuses on categorical data and survival data. Topics in categorical data to be covered include defining rates, incidence and prevalence, the chi-squared test, Fisher's exact test and its extension, relative risk and odds-ratio, sensitivity, specificity, predictive values, logistic regression with goodness of fit tests, ROC curves, Mantel-Haenszel test, McNemar's test, the Poisson model, and the Kappa statistic. Survival analysis will include defining the survival curve, censoring, and the hazard function, the Kaplan-Meier estimate, Greenwood's formula and confidence bands, the log rank test, and Cox's proportional hazards regression models. Examples of medical and biologic data will be used throughout the course, and use of computer software demonstrated.

BSTA670
Statistical Computing

This is a course in computing algorithms useful in statistical research and advanced statistical applications. Topics include computer arithmetic; matrix algebra; numerical optimization methods with application to maximum-likelihood estimation and GEEs; spline smoothing and penalized likelihood; numerical integration, random number generation and simulation; Gibbs sampling; bootstrap methods; missing data problems and EM; imputation; data augmentation algorithms; and Fourier transforms. (1.0 course unit, fall.) Prerequisites: BSTA 621, BSTA 651; permission of instructor.

CIS Programming Courses :
CIS 190 C++ Programming
CIS 191 Linux / Unix skills
CIS 192 Python Programming

This are .5 credit courses on programming languages offered by the Engineering school. More information here.

CIS 261
Discrete Probability, Stochastic Processes, and Statistical Inference

This course tightly integrates the theory and applications of discrete probability, discrete stochastic processes, and discrete statistical inference in the study of computer science. The course will introduce the Minimum Description Length Paradigm to unite basic ideas about randomness, inference and computation. Students will be expected to use the Maple programming environment in homework exercises which will include numerical and symbolic computations, simulations, and graphical displays.

CIS 419/520
Machine Learning

This course covers the foundations of statistical machine learning. The focus is on probabilistic and statistical methods for prediction and clustering in high dimensions. Topics covered include SVMs and logistic regression, PCA and dimensionality reduction, and EM and Hidden Markov Models.

CIS 450/550
Database & Information Systems

Structured information is the lifeblood of commerce, government, and science today. This course provides an introduction to the broad field, covering a range of topics relating to structured data, ranging from data modeling to logical foundations and popular languages, to system implementations. We will study the theory of relational and XML data design, the basics of query languages, databases' role in the Web, information integration, object-to-relational mappings, the connections between databases, key-value stores and "big data", database tuning and internals, and recent ""cloud"" data processing techniques. The course project will exercise students' knowledge of database design, AJAX Web programming, and information integration, to build a social network photo sharing site. Database design, relational algebra and calculus, query languages (SQL, Datalog, XQuery), views, object and XML mappings, indexing, consistency, database tuning, servlet programming with AJAX, Map/Reduce and its relationship to SQL, basics of information retrieval and recommendation systems.

CIS 521
Fundamentals of AI

Modern AI uses a collection of techniques from a number of fields in the design of intelligent systems: probability, statistics, logic, operations research, optimal control and economics, to name a few. This course covers basic modeling and algorithmic tools from these fields underlying current research and highlights their applications in computer vision, robotics, and natural language processing.

ECON 103
Statistics for Economics

The course focuses on elementary probability and inferential statistical techniques. The course begins with a survey of basic descriptive statistics and data sources and then covers elementary probability theory, sampling, estimation, hypothesis testing, correlation, and regression. The course focuses on practical issues involved in the substantive interpretation of economic data using the techniques of statistical inference. For this reason empirical case studies that apply the techniques to real-life data are stressed and discussed throughout the course, and students are required to perform several statistical analyses of their own.

ECON 104
Econometrics

This course is designed to introduce students to econometric techniques and their applications in economic analysis and decision-making. The main objective of the course is to train the student in (i) handling economic data; (ii) quantitative analyses of economic models with probabilistic tools; (iii) econometric techniques, their application as well as their statistical and practical interpretation; (iv) implementing these techniques on a computer. Estimation and inference procedures are formally analyzed for simple econometric models and illustrated by empirical case studies using real-life data. The course covers linear regression models, simultaneous-equations models, discrete choice models and univariate time

series models. Estimation and Inference is conducted using least squares and likelihood based techniques. Students are required to perform several econometric analyses of their own.

ECON 222
Advanced Econometric Techniques and Applications

This course introduces students to advanced study in econometrics, with an emphasis on methods used in microeconomic applications and in evaluating the effects of social interventions. The methods covered include methods for handling limited dependent variables (useful, for example, in forecasting the demand for a new good), maximum likelihood estimators, and flexible semiparametric and non parametric estimation methods, and randomized and nonexperimental methods of estimating treatment effects. Applications of econometrics to the field of program evaluation will also be studied.

ECON 224
Macro-Modeling

This is an advanced undergraduate course in models of economic growth. Students will be introduced to the workhorse theoretical models that are used to understand growth by modern macroeconomic researchers and policy makers. The types of questions that we will address include: Why are some countries richer than others? Why do some countries grow quickly while others stagnate? Why did modern economic growth start in Western Europe? What can governments do to accelerate economic growth? How does economic growth interact with demographic and geographic factors? We will build theoretical models that can be used to answer these questions. There will be a strong focus on emphasizing the microeconomic foundations of models, and using the language of mathematics to express the underlying assumptions and assess their implications for policy. Hence, there are strict mathematical prerequisites. We will also compare the predictions of our models with the data. Thus, a fair amount of econometrics will be required. A class in statistics and econometrics is highly recommended.

ENM 321
Engineering Statistics

This course covers the topics in probability and statistics with emphasize on the application of probability theories and statistical techniques to practical engineering problems. Mathematical derivations of theorems will be presented whenever it is necessary to illustrate the concepts involved, however.

ENM 503
Introduction to Probability and Statistics

Combinatorics, introduction to probability, conditional probability and independence, random variables, special discrete and continuous distributions, descriptive statistics, expectation, variance, covariance, joint probability, moment generating functions, stochastic models and applications, Markov chains, queuing models, order statistics.

ESE 301/302
Engineering Probability

Basic ideas of probability theory. Combinatorics. Random variables and functions of random variables. Means, moments and generating functions. Order statistics and special distributions. Inequalities and the central limit theorem. Principles and engineering applications of statistical inference. The basic topics covered are parameter estimation, confidence intervals, and hypothesis testing. Additional topics may include analysis of variance (ANOVA) and/or linear regression. Each method is treated both from theoretical and applied viewpoints, including software analysis of selected data sets.

ESE303
Stochastic Systems Analysis and Simulation

Stochastic systems analysis and simulation (ESE 303) is a class that explores stochastic systems which we could loosely define as anything random that changes in time. Stochastic systems are at the core of a number of disciplines in engineering, for example communication systems and machine learning. They also find application elsewhere, including social systems, markets, molecular biology and epidemiology. The goal of the class is to learn how to model, analyze and simulate stochastic systems. With respect to analysis we distinguish between what we could call theoretical and experimental analysis. By theoretical analysis we refer to a set of tools which let us discover and understand properties of the system. These analyses can only take us so far and is usually complemented with numerical analysis of experimental outcomes. Although we use the word experiment more often than not we simulate the stochastic system in a computer and analyze the outcomes of these virtual experiments. The class's material is divided in four blocks respectively dealing with Markov chains, continuous time Markov chains, Gaussian processes and stationary processes. Emphasis is placed in the development of toolboxes to analyze these different classes of processes and on describing their applications to complex stochastic systems in different disciplines. Particular examples include: (i) the problem of ranking web pages by a search engine; (ii) the study of reputation and trust in social networks; (iii) modeling and analysis of communication networks; (iv) the use of queues in the modeling of transportation networks; (v) stochastic modeling and simulation of biochemical reactions and gene networks; (vi) arbitrages, pricing of stocks, and pricing of options through Black-Scholes formula; and (vii) linear filtering of stochastic processes to separate signals of interest from background noise. For more information visit the class's web page at http://alliance.seas.upenn.edu/ ~ese303/wiki/.

FNCE392
Financial Engineering
Krisha Ramaswamy

This class offers an analysis of advanced derivative pricing models. It aims at reviewing the main models and modeling techniques used in practical applications, understanding their applicability and limitations, and at building an integrated framework allowing students to: 1) decide what stochastic factors (e.g., volatility, jumps, one or more interest rate factors, default intensities) should be incorporated in a reasonable pricing model for a given derivative; 2) formulate a consistent model incorporating the chosen factors; 3) calibrate the model using market data; 4) price the derivative and identify a hedging strategy. To allow sufficient flexibility in the choice of the pricing model, the class will not place any special emphasis on closed-form valuation formulas, relying instead on the full generality afforded by the martingale approach to asset pricing as numerically implemented by Monte-Carlo simulation. Students will be asked to implement and calibrate the models introduced in the class using software of their choice: Mathematica or Matlab, in conjunction with Crystal Ball (an Excel add-in for Monte Carlo simulation), are strongly recommended. Although every effort will be made to introduce the models as intuitively as possible, the class will be by its nature very quantitative and will require a significant amount of work116 and Math 260 (the analogues of Math 114 and Math 240, respectively). These courses will cover essentially the same material as 114 and 240, but more in depth and involve discussion of the underlying theory as well as computations.

MATH 114
Calculus II

Functions of several variables, vector-valued functions, partial derivatives and applications, double and triple integrals, conic sections, polar coordinates, vectors and vector calculus, first order ordinary differential equations. Applications to physical sciences. Use of symbolic manipulation and graphics software in calculus.

MATH 116
Honors Calculus

Students who are interested in math or science might also want to consider a more challenging Honors version of Calculus II and III, Math 116 and Math 260 (the analogues of Math 114 and Math 240, respectively). These courses will cover essentially the same material as 114 and 240, but more in depth and involve discussion of the underlying theory as well as computations.

MATH 240/241
Calculus III / Calculus IV

Linear algebra: vectors, matrices, systems of linear equations, eigenvalues and eigenvectors. Series solutions of ordinary differential equations, Laplace transforms and systems of ordinary differential equations. Use of symbolic manipulation and graphics software. IV: Sturm-Liouville

problems, orthogonal functions, Fourier series, and partial differential equations including solutions of the wave, heat and Laplace equations, Fourier transforms. Use of symbolic manipulation and graphics software.

MATH 312/313/412
Linear Algebra

Various Linear transformations, Gauss Jordan elimination, eigenvalues and eigenvectors, theory and applications. Mathematics majors are advised that MATH 312 cannot be taken to satisfy the major requirements.

MATH 530
Mathematics of Finance

This course presents the basic mathematical tools to model financial markets and to make calculations about financial products, especially financial derivatives. Mathematical topics covered: stochastic processes, partial differential equations and their relationship. No background in finance is assumed.

MKTG 101
Introduction to Marketing

The objective of this course is to introduce students to the concepts, analyses, and activities that comprise marketing management, and to provide practice in assessing and solving marketing problems. The course is also a foundation for advanced electives in Marketing as well as other business/social disciplines. Topics include marketing strategy, customer behavior, segmentation, market research, product management, pricing, promotion, sales force management and competitive analysis.

MKTG 212/712
Data and Analysis for Marketing Decisions

Prerequisite(s): MKTG 101, STAT 101. Students are highly encouraged to take statistics in the semester immediately preceding this course. (Former course title Marketing Research.) Firms have access to detailed data of customers and past marketing actions. Such data may include in-store and online customer transactions, customer surveys as well as prices and advertising. Using real-world applications from various industries, the goal of the course is to familiarize students with several types of managerial problems as well as data sources and techniques, commonly employed in making effective marketing decisions. The course would involve formulating critical managerial problems, developing relevant hypotheses, analyzing data and, most importantly, drawing inferences and telling convincing narratives, with a view of yielding actionable results.

MKTG 271/771
Models for Marketing Strategy
Eliashberg

In today's business environment, marketing executives are involved in complex decision-making and they become responsible for return on their marketing investments. The first objective of this course is to help participants become better executives. By exposing students to various analytical and computer-based tools, developed for solving marketing problems, it will help to prepare them for careers in industries such as consumer packaged goods, hi-tech, financial services, media and entertainment, pharmaceutical, consulting, and venture capital. The course's main focus is on various existing models, such as models that predict the consumer's dynamic adoption of an innovative product. However, at some point in their career, students may find themselves facing business problems for which a model can assist in making decisions, but no existing model is available. Hence, the second objective of the course is to provide participants with critical skills necessary to evaluate new models to which they may be exposed by attending presentations or reading the literature. The models to be discussed in the class have been implemented and proven useful in a wide range of industries (e.g., business-to-consumers and business-to-business). The course is not only about models, however. It also covers modeling needs. Some industries such as the media and entertainment or the pharmaceutical industries present unique problems and modeling needs. The third objective of the course is to expose participants to the nature and essence of such idiosyncratic problems as well as modeling needs in such industries. Overall, the course will make participants understand better critical marketing problems by analyzing them rigorously and will enhance their skills in either designing or evaluating models-based strategies.

MKTG 476/STAT 776
Applied Probability Models
Pete Fader

This course will expose students to the theoretical and empirical “building blocks” that will allow them to construct, estimate, and interpret powerful models of customer behavior. Over the years, researchers and practitioners have used these models for a wide variety of applications, such as new product sales, forecasting, analyses of media usage, and targeted marketing programs. Other disciplines have seen equally broad utilization of these techniques. The course will be entirely lecture-based with a strong emphasis on real-time problem solving. Most sessions will feature sophisticated numerical investigations using Microsoft Excel. Much of the material is highly technical.

MKTG 309/809
Experiments for Business Decision Making

In the past decade, massive shifts in how companies interact with their customers have suddenly made field experiments an economically feasible way to learn about a variety of business questions such as what types of promotions are most effective, what products should be stocked at a store, how e-mail promotions should be designed, how sales staff should be compensated, etc. Many marketers engaged in online retailing, direct-marketing, online advertising, media management, etc. are rapidly embracing a “test and learn” philosophy and a number of platforms such as Google Website Optimizer, have been developed to facilitate rigorous field experiments in the online environment. Just as with the quality revolution in manufacturing during the 1980s and 1990s, the rapid rise of the “test and learn” philosophy in marketing has created a huge demand for those who can design, field, and analyze marketing experiments. Through this course, you will learn and practice a wide range of critical skills, from the statistical methods used to design and analyze experiments to the management and strategy required to execute an experiment and act on the results. Although the cases and examples will focus on marketing problems, the material covered can be applied in a number of other domains particularly operations management and product design.

MKTG 940/941
Measurement and Data Analysis in Marketing
Raghu Iyengar

In this course we consider models for binary, count, and continuous data including contingency table models, logistic and profit regression, ANOVA, ANCOVA, conjoint analysis, and OLS. In addition we cover multidimensional techniques such as MDS, cluster analysis, principal components analysis, factor analysis, and discriminant analysis. We utilize the statistics package SPlus 2000, and also BUGS for implementing many of the techniques described in a Bayesian manner.

OIDD 101
An Introduction to Operations, Information, and Decisions

OIDD 101 explores a variety of common quantitative modeling problems that arise frequently in business settings, and discusses how they can be formally modeled and solved with a combination of business insight and computer-based tools. The key topics covered include capacity management, service operations, inventory control, structured decision making, constrained optimization and simulation. This course teaches how to model complex business situations and how to master tools to improve business performance. The goal is to provide a set of foundational skills useful for future coursework at Wharton as well as providing an overview of problems and techniques that characterize disciplines that comprise Operations and Information Management.

OIDD 105
Developing Tools for Data Access and Analysis

This course provides an introduction to the construction of data analysis tools that are commonly used for business applications, especially in consulting and finance. The course builds on the spreadsheet and analytical skills developed in OPIM101, providing a much more extensive treatment of spreadsheet application development and database management. The first portion of the course will focus on programming in VBA, the embedded programming language in the Microsoft Office suite of applications. This will be supplemented with discussion of industry best practice in software development, such as specification development, interface design, documentation, and testing. The second portion of the class will emphasize data access and analysis utilizing SQL, the industry standard language for interacting with database software.

OIDD 290
Decision Processes
Joseph Simmons

This course is an intensive introduction to various scientific perspectives on the processes through which people make decisions. Perspectives covered include cognitive psychology of human problem-solving, judgment and choice, theories of rational judgment and decision, and the mathematical theory of games. Much of the material is technically rigorous. Prior or current enrollment in STAT 101 or the equivalent, although not required, is strongly recommended.

OIDD 311
Business Computer Languages
Steve Kimbrough

Best way to learn how to code (Matlab / Python) with no previous experience in real computer science. Spring semester. This course is taught with the more descriptive title of "Scripting for Business Analytics." "Business Analytics" refers to modeling and analysis undertaken for purposes of management and supporting decision making. The varieties of techniques and methods are numerous and growing, including simple equational models, constrained optimization models, probabilistic models, visualization, data analysis, and much more. Elementary modeling of this sort can be undertaken in Excel and other spreadsheet programs, but "industrial strength" applications typically use more sophisticated tools, based on scripting languages. Scripting languages are programming languages that are designed to be learned easily and to be used for special purposes, rather than for large-scale application programming. This course focuses on the special purposes associated with business analytics and teaches MATLAB and Python in this context. MATLAB and Python are widely used in practice (both in management and in engineering), as are the business analytic methods covered in the course. Prior programming experience is useful, but not required or presumed for this course.

OIDD 314/662
Enabling Technologies

Conducting business in a networked economy invariably involves interplay with technology. The purpose of this course is to improve understanding of technology (what it can or cannot enable) and the business drivers of technology-related decisions in firms. We will be discussing some of the new and most disruptive technologies right now to stimulate thought on new applications for commerce and new ventures, as well as their implications to the tech industry as a whole. Topics include social media, online advertising, big data, and cloud computing. *IN SPRING 2015 THIS COURSE WILL BE OFFERED AS A FULL-CU COURSE THAT ONLY MEETS DURING Q3. WE WILL MEET TWICE A WEEK FOR 3 HOURS PER DAY TO ENSURE THAT THE TOTAL CLASSROOM TIME IS THE SAME AS OTHER 1-CU COURSES. THIS FORMAT WILL FREE UP TIME FOR STUDENTS TO WORK ON COURSE PROJECTS DURING Q4. WE WILL NOT HAVE REGULAR CLASS TIME IN Q4 AS A RESULT.* The course will take a layered approach (from network infrastructure) to data infrastructure to applications infrastructure, or direct enablers of commerce) to first, understanding and then, thinking about technology enablers. Network infrastructure layers include fundamentals of wired and wireless infrastructure technologies such as protocols for networking, broadband technologies - for last (DSL, Cable etc) and other miles (advances in optical networking) and digital cellular communications. Data infrastructure layers include usage tracking technologies, search technologies and data mining. Direct application layers include personalization technologies (CRM), design technologies for content and exchanges, software renting enablers, application service provision, agents and security mechanisms. Finally some emerging technology enablers (such as Bluetooth, biometrics and virtual reality) are identified and discussed.

OIDD 321
Management Science

Understanding how to use data and business analytics can be the key differential for a company's success or failure. This course is designed to introduce fundamental quantitative decision making tools for a broad range of managerial decision problems. Topics covered include linear, nonlinear and discrete optimization, dynamic programming, and simulation. Students will apply these quantitative models in applications of portfolio management, electricity auctions, revenue management for airlines, manufacturing, advertising budget allocation, and healthcare scheduling operations. Emphasis in this course is placed on mathematical modeling of real world problems and implementation of decision making tools.

OIDD 353 - Mathematical Modeling and its Application in Finance

Quantitative methods have become fundamental tools in the analysis and planning of financial operations. There are many reasons for this development: the emergence of a whole range of new complex financial instruments, innovations in securitization, the increased globalization of the financial markets, the proliferation of information technology and the rise of high-frequency traders, etc. In this course, models for hedging, asset allocation, and multi-period portfolio planning are developed, implemented, and tested. In addition, pricing models for options, bonds, mortgage-backed securities, and other derivatives are studied. The models typically require the tools of statistics, optimization, and/ or simulation, and they are implemented in spreadsheets or a high-level modeling environment, MATLAB. This course is quantitative and will require extensive computer use. The course is intended for students who have strong interest in finance. The objective is to provide students the necessary practical tools they will require should they choose to join the financial services industry, particularly in roles such as: derivatives, quantitative trading, portfolio management, structuring, financial engineering, risk management, etc. Prospective students should be comfortable with quantitative methods such as basic statistics and the methodologies (mathematical programming and simulation) taught in OIDD 612 Business Analytics and OIDD 321 Management Science (or equivalent). Students should seek permission from the instructor if the background requirements are not met.

OIDD 410/672
Decision Support Systems
Hill

The past few years have seen an explosion in the amount of data collected by businesses and have witnessed enabling technologies such as database systems, client-server computing and artificial intelligence reach industrial strength. These trends have spawned a new breed of systems that can support the extraction of useful information from large quantities of data. Understanding the power and limitations of these emerging technologies can provide managers and information systems professionals new approaches to support the task of solving hard business problems. This course will provide an overview of these techniques (such as genetic algorithms, neural networks, and decision trees) and discuss applications such as fraud detection, customer segmentation, trading, marketing strategies and customer support via cases and real datasets.

OIDD 661
Systems Analysis, Design, and Implementation

At its surface this course introduces students to the management and technical issues associated with planning and designing large-scale computer systems. It does so in part as an elaboration of Fred Brooks's observation that "The technology, the surrounding organization, and the traditions of the craft conspire to define certain items of paperwork." But if that were our only goal, we would soon find ourselves mired in (and probably arguing about) the minutiae of how such paper items ought to be constructed - not a very helpful pedagogical exercise. So then, at a deeper level we seek to understand why the conspiracy endures, and why in spite of it, systems still take too long and cost too much to build as a systems project's team members struggle to understand one another across disparate discourse communities and world views, differences in experience and training, and over long periods of time. More than anything else, within the context of working with the main tools and techniques of systems analysis and design, this course treats communication, corroboration, and thinking within the boundaries of a technology-oriented project as its primary subjects.

OIDD 664
Data Based and Information Management Systems

Data and information are critical to the modern organization. Whether used in knowledge management, business intelligence, enterprise resource planning (ERP), product design, marketing, personalization and other aspects of managing customer relationships (CRM), the underlying principles of data management are the same. This course aims to provide a practical introduction to the fundamental principles. Examples and exercises will cover the relational database tools at the core of ERP, CRM, and on-line exchanges and portals. However, the course will also use the same basic foundations to consider emerging technologies and standards such as XML, ebXML, UDDI, etc.

OIDD 910/ESE504
Optimization Theory

Various Introduction to mathematical programming for PhD students who would like to be intelligent and sophisticated consumers of mathematical programming theory but do not plan to specialize in this area. Integer and nonlinear programming are covered, including the fundamentals of each area together with a sense of the state-of-the-art and expected directions of future progress.

OIDD 930/931
Stochastic Models
Rieders

This course introduces mathematical models describing and analyzing the behavior of processes that exhibit random components. The theory of stochastic processes will be developed based on elementary probability theory and calculus. Topics include random walks, Poisson processes, Markov chains in discrete and continuous time, renewal theory, and martingales. Applications from the areas of inventory, production, finance, queueing and communication systems will be presented throughout the course. This course is really only appropriate for students with a very strong math background (i.e. a semester of linear algebra, a semester of real analysis, a semester of advanced probability, and exposure to differential equations) and an interest in going into academia.

OIDD 934
Dynamic Programming and Stochastic Models

Prerequisite(s): OPIM930. The course goal is to provide a brief but fairly rigorous introduction to the formulation and solution of dynamic programs. Its focus is primarily methodological. We will cover discrete state space problems, over finite or infinite time horizon, with and without discounting. Structured policies and their theoretical foundation will be of particular interest. Computational methods and approximation methods will be addressed. Applications are presented throughout the course, such as inventory policies, production control, financial decisions, and scheduling.

PSYC 739
Probabilistic Models of Perception and Cognition
Stocker

Probability theory has become an increasingly popular and successful framework for modeling human perceptual and cognitive behavior. This course will provide a careful introduction to probability theory and the various ways it has been applied in psychology and neuroscience. Goal is to make students understand the most important state-of-the-art probabilistic models in perception and cognition, what they reveal about the brain's underlying computations and strategies in dealing with uncertainty, and how such computations can potentially be performed by populations of neurons.

STAT 101/102
Introductory Business Statistics

Various Data summaries and descriptive statistics; introduction to a statistical computer package; Probability: distributions, expectation, variance, covariance, portfolios, central limit theorem; statistical inference of univariate data; Statistical inference for bivariate data: inference for intrinsically linear simple regression models. This course will have a business focus, but is not inappropriate for students in the college.

STAT 111/112
Introductory Statistics
Wyner

Introduction to concepts in probability. Basic statistical inference procedures of estimation, confidence intervals and hypothesis testing directed towards applications in science and medicine. The use of the JMP statistical package.

STAT 430
Probability

Various Discrete and continuous sample spaces and probability; random variables, distributions, independence; expectation and generating functions; Markov chains and recurrence theory.

STAT 431
Statistical Inference

An introduction to the mathematical theory of statistics. Estimation, with a focus on properties of sufficient statistics and maximum likelihood estimators. Hypothesis testing, with a focus on likelihood ratio tests and the consequent development of "t" tests and hypothesis tests in regression and ANOVA. Nonparametric procedures.

STAT 433/533
Stochastic Processes
Steele, Mossel

Prerequisite(s): STAT 430, or permission of instructor. An introduction to Stochastic Processes. The primary focus is on Markov Chains, Martingales and Gaussian Processes. We will discuss many interesting applications from physics to economics. Topics may include: simulations of path functions, game theory and linear programming, stochastic optimization, Brownian Motion and Black-Scholes.

STAT 453/ BEPP 853
Actuarial Statistics
Jean Lemaire

This course covers models for insurer's losses, and applications of Markov chains. Poisson processes, including extensions such as non-homogeneous, compound, and mixed Poisson processes are studied in detail. The compound model is then used to establish the distribution of losses. An extensive section on Markov chains provides the theory to forecast future states of the process, as well as numerous applications of Markov chains to insurance, finance, and genetics. The course is abundantly illustrated by examples from the insurance and finance literature. While most of the students taking the course are future actuaries, other students interested in applications of statistics may discover in class many fascinating applications of stochastic processes and Markov chains.

STAT 470
Data Analytics and Statistical Computing

This course will introduce a high-level programming language, called R, that is widely used for statistical data analysis. Using R, we will study and practice the following methodologies: data cleaning, feature extraction; web scrubbing, text analysis; data visualization; fiing statistical models; simulation of probability distributions and statistical models; statistical inference methods that use simulations (bootstrap, permutation tests).

STAT471/701
Intermediate Statistics/Advanced Statistics for Management
Linda Zhao

Modern Data Mining: Statistics or Data Science has been evolving rapidly to keep up with the modern world. While classical multiple regression and logistic regression technique continue to be the major tools we go beyond to include methods built on top of linear models such as LASSO and Ridge regression. Contemporary methods such as KNN (K nearest neighbor), Random Forest, Support Vector Machines, Principal Component Analyses (PCA), the bootstrap and others are also covered. Text mining especially through PCA is another topic of the course. While learning all the techniques, we keep in mind that our goal is to tackle real problems. Not only do we go through a large collection of interesting, challenging real-life data sets but we also learn how to use the free, powerful software "R" in connection with each of the methods exposed in the class.

STAT 472/712
Decision Making Under Uncertainty
Stine

Fundamentals of modern decision analysis with emphasis on managerial decision making under uncertainty and risk. The basic topics of decision analysis are examined. These include payoffs and losses, utility and subjective probability, the value of information, Bayesian analysis, inference and decision making. Examples are presented to illustrate the ideas and methods. Some of these involve: choices among investment alternatives; marketing a new product; health care decisions; and costs, benefits, and sample size in surveys.

STAT 473
Bioinformatics

An introduction to the use of statistical methods in the increasingly important scientific areas of genomics and bioinformatics. The topics to be covered will be decided in detail after the initial class meeting, but will be taken from the following: - background probability theory of one and many random variables and of events; background statistical inference theory, classical and Bayesian; Poisson processes and Markov chain; the analysis of one and many DNA sequences, in particular shotgun sequencing, pattern analysis and motifs; substitution matrices, general random walk theory, advanced statistical inference, the theory of BLAST, hidden Markov models, microarray analysis, evolutionary models.

STAT 474
Modern Regression

Function estimation and data exploration using extensions of regression analysis: smoothers, semiparametric and nonparametric regression, and supervised machine learning. Conceptual foundations are addressed as well as hands-on use for data analysis.

STAT 520/521
Applied Economics
Dylan Small

This is a course in econometrics for graduate students. The goal is to prepare students for empirical research by studying econometric methodology and its theoretical foundations. Students taking the course should be familiar with elementary statistical methodology and basic linear algebra, and should have some programming experience. Topics include conditional expectation and linear projection, asymptotic statistical theory, ordinary least squares estimation, the bootstrap and jackknife, instrumental variables and two-stage least squares, specification tests, systems of equations, generalized least squares, and introduction to use of linear panel data models. Topics include system estimation with instrumental variables, fixed effects and random effects estimation, M-estimation, nonlinear regression, quantile regression, maximum likelihood estimation, generalized method of moments estimation, minimum distance estimation, and binary and multinomial response models. Both theory and applications will be stressed.

STAT 531
Advanced Stochastic Processes

Markov chains, Markov processes, and their limit theory. Renewal theory. Martingales and optimal stopping. Stable laws and processes with independent increments. Brownian motion and the theory of weak convergence. Point processes. Requires a strong background in measure theory and real analysis.

STAT 541
Statistical Methodology
Andreas Buja

This is a course that prepares 1st year PhD students in statistics for a research career. This is not an applied statistics course. Topics covered include: linear models and their high-dimensional geometry, statistical inference illustrated with linear models, diagnostics for linear models, bootstrap and permutation inference, principal component analysis, smoothing and cross-validation.

STAT 542
Bayesian Methods and Computation
Shane Jensen

Sophisticated tools for probability modeling and data analysis from the Bayesian perspective. Hierarchical models, mixture models and Monte Carlo simulation techniques.

STAT 550
Mathematical Statistics
Dylan Small

Decision theory and statistical optimality criteria, sufficiency, point estimation and hypothesis testing methods and theory.

STAT622
Statistical Modeling
Justin Bleich

This six-week, elective MBA course continues the required MBA statistics course, Stat 613. It expands the material covered in Stat 613 in several ways, adding both breadth (e.g., logistic regression) and depth to the coverage of regression (e.g., more diagnostics, model selection). The course emphasizes the models for decision making from large data sets, as common in data-mining. Lectures feature extensive analysis of large data sets from marketing, personal finance, and management. The course presumes that students are familiar with the inferential methods covered in STAT 613 (including hypothesis tests, confidence intervals, p-values) as well as the use and interpretation of least squares regression models. The course also uses JMP as in STAT 613. Beginning with a review of these concepts, STAT 622 covers related methodologies that produce fits that resemble and extend regression models. The methodologies include those that expand the nature of the predictors (as in the use of special transformations in time series and the construction of regression trees) and allow the use of categorical reponses (logistic regression). The course concludes by exploring the relationship between a regression fit to observational data and an analysis of variance estimated from experimental data, as in a conjoint analysis.

STAT 901
Stochastic Processes
Dean Foster

Martingales, optimal stopping, Wald's lemma, age-dependent branching processes, stochastic integration, Ito's lemma.

STAT 955
Stochastic Calculus and Financial Applications Steele

Selected topics in the theory of probability and stochastic processes. Requires a strong background in measure theory and real analysis.