Lecture notes 1
Introduction

- About EE178
- Prerequisites (calculus & linear algebra)
- Course Outline
- History
- Mathematical Models
- Role of the Math
- Why Probabilistic Analysis?

- Two Famous Laws

Ackknowledgement These lecture notes draw heavily on notes by Ron Howard, Al Drake, John Tsitsiklis, and Dimitris Bertsekas.

About EE178

Introduction to probabilistic systems analysis.

Basic probability: no prior probability experience is assumed.

Short introduction to statistics and some statistical signal processing.

- *Probability theory:* Mathematical models of random phenomena and their use to predict quantifiable future behavior, e.g., would expect percentage of heads in a sequence of flips of fair coin to approach 50% for a large number of flips.
- *Statistics:* How do you get the models? How do you make inferences from data?

- *Statistical signal processing:* Applications of probability and statistics to specific random signals such as speech, audio, images, video, EKGs, etc. in order to predict, smooth, extract, compress, infer, find, classify, estimate, interpolate, modify, . . .

  Optimize "average" value of some performance measure.

Course Logistics: See course information sheet at
`http://www.stanford.edu/class/ee178/`.
Course also has a secure Web site which will be used for grade and other information at
`http://eeclass.stanford.edu/ee178/`

Prerequisites

- You should be familiar with basic calculus, e.g., you should be able to sum geometric progressions and related sequences and integrate exponentials and related functions. You should be able to do basic integral manipulation like making linear variable substitutions.
- You should be familiar with basic linear algebra, e.g., understand vector and matrix notation, be able to multiply vectors with matrices and matrices with each other, you should be familiar with the ideas of an inverse of a matrix and the determinant of a matrix.

Course Outline

1. Introduction
   - About EE178
   - Prerequisites (calculus & linear algebra)
   - Course Outline
   - History
   - Mathematical Models
   - Role of the Math
   - Why Probabilistic Analysis?
   - Two Famous Laws
2. Probabilistic Models
   - Basic probabilistic model: an experiment
     - Sample space
     - Events (algebra of events, set theory)
     - Probability measure (law, distribution)

3. Conditional probability
   - Conditional probability
   - Chain rule (multiplicative rule)
   - Independence
   - Total Probability
   - Bayes Rule
4. Counting
   - Principles of Counting
   - Examples
   - Binomial Probabilites
5. Random variables
   - Random Variables
   - Examples
   - Probability mass functions (pmfs) and How to Find Them
   - Famous random variables: Binomial, Bernoulli, Geometric, and Poisson
   - Functions of Random Variables

   - Derived Distributions

6. Expectation
   - Review of RVs, pmfs
   - Expectation
   - Examples
7. Multiple random variables
   - Marginal and Joint PMF's
   - Linearity of Expectation
   - Conditional PMF's
   - Total Probability for PMF's
   - Conditional Expectation
   - Independent Random Variables
   - Sums of Independent Random Variables
   - Sample Means
8. Continuous random variables

- Continuous Random Variables
- Probability Density Functions (PDF's)
- Examples: Uniform, Exponential, Laplacian, Gaussian
- Expectation
- Cumulative Distribution Functions (CDF's)

9. Multiple continuous random variables

- PDF Review
- Multiple random variables
  - Joint PDF's
  - Independence
  - Conditioning
- Example: stick breaking
- A few more properties
- Examples: Buffon's needle and darts
- Multivariate Gaussian random variables

10. Derived densities

- Derived pdfs

- Examples

11. Transforms

- Transforms
- Moment generating properties
- Inversion of transforms
- Examples
- Transforms of sums of independent RVs

12. Iterated expectation

- Conditional expectation and iterated expectation revisited
- Conditional variances
- Sums of a random number of random variables

13. Least squares estimation

- Review of Iterated Expectations
- Least Squares Estimation

14. Stochastic Processes I: Bernoulli and Poisson processes

- Stochastic Processes
- Bernoulli Process
- Binomial process
- Interarrival times
- Poisson process
- Random telegraph wave
- Arrivals
- Interarrival times
- Properties

15. Stochastic Processes II: Markov Chains

- Discrete State Processes
- Markov Chains
- State Classification
- Steady State Probabilities
- Birth and Death Processes

16. The law of large numbers

- Sums of Random Variables Reviewed
- Convergence of Random Variables
- The Chebychef and Markov Inequalities
- The Weak Law of Large Numbers

17. The central limit theorem

- Gaussian Density Revisited
- Standardized Sums
- The CLT

18. An introduction to statistics

- Probability vs Statistics
- Significance Testing
- Hypothesis Testing
- Estimation
- Baysesian Statistics

19. Review

History

- Ancient times
  - Egypt, Greece, Rome, India
  - predict weather and river flooding based on past observations
  - census for taxes (e.g., Augustus)
  - gaming: dice (e.g., Augustus again)
  - divination (e.g., I Ching)
- 11th century
  - Domesday book (William the Conquerer's statistical study of population and occupations for taxation)
  - Agricultural inventories
  - Marine shipping/insurance (14th-15th centuries)
- 16th century
  - mortality statistics (epidemics)

  - Beginnings of theory, counting to determine chance, simple combinatorics, beginnings of "calculus of probability," additive and multiplicative rules
  - Cardano: "stable" behavior of dice
  - Galileo: "clustering" of measurements of dice around "true" values
- 17th century
  - Pascal, Fermat, Huygens: Independent trials, combinatorics, expectation, more calculus
- 18th- early 19th centuries
  - Bernoullis, (*The art of conjecturing* by James Bernoulli), first law of large numbers
  - Bayes: conditional probability
  - Euler: Lottery for Frederick the Great
  - LaPlace: central limit theorem (Gaussian law: also found earlier by Robert Adrian and later by Gauss), Demoivre-Laplace theorem of approximation (certain discrete probabilities approximated by Gaussian distribution)

  - Gauss: distribution of random errors, FFT
  - Poisson
- 19th century
  - Russian school (St. Petersburg): Lobachevsky, Chebychev, Markov
  - Physics: Boltzman, Brown, Gibbs, Maxwell
- Modern
  - Axiomatic probability: Kolmogorov (Khinchine, Borel, Cantelli, von Mises)
  - Computers, simulations, Monte Carlo methods

Mathematical Models

- Used as substitute for reality, lead to quantitative results and simulation methods.
- Main components:
  - Sample space — a description of all possible outcomes of an experiment
  - Algebra of events (or sets) — A common nonambiguous language for describing what can happen in an experiment
  - Probability law or measure — the relevant measure in sample space, based on algebra of events plus three simple axioms. Can prove theorems relating probability measure to intuitive notions of probability, likelihood, relative frequency.
  - Calculus, to manipulate probability measures and compute expectations.

Role of the Math

- Probabilistic analysis is mathematical, but intuition often dominates and guides the math.
- BUT beware, sometimes results are counterintuitive at first glance, probability theory has many apparent paradoxes.
- Usually the hardest part of probabilistic analyis is correct formulation of the problem. After that it is mostly calculus.
- Probabilistic analysis is easy in hindsight. Many students have looked at the answers of old homeworks and exams without paying their dues by first trying the problems and have found the answers to be easy and obvious, only to have a rude awakening on unfamiliar problems in an exam. Like carpentry and motorcycle maintenance, you have to work the problems to understand the methods and build the correct intuition.

Why Probabilistic Analysis?

Applications

- Engineering
  - Statistical signal processing
  - Comminications
  - Systems and control
  - Decision and resource allocation under uncertainty (e.g., communications networks)
  - Reliability (dealing with noise, error control, failures)
- Economics and finance
- Physics, statistical mechanics, thermodynamics
- Medicine, FDA, drugs and procedures
- Linguistics, automatic speech recognition and translation

Two Famous Laws

- *Laws of large numbers* (also called ergodic theorems)

If a sequence of random numbers X_n, $n = 1, 2, 3, \ldots$ are described by the same probability measure and are mutually independent, then

$$\frac{1}{N}\sum_{n=1}^{N} X_n = \frac{X_1 + X_2 + \cdots + X_N}{N}$$

convergences to a constant as $N \to \infty$.

- *Central limit theorem* If a sequence random numbers X_n, $n = 1, 2, 3, \ldots$ are described by the same probability measure and are mutually independent and if the constant of the previous law is m, then

$$\frac{1}{\sqrt{N}}\sum_{n=1}^{N} X_n = \frac{X_1 + X_2 + \cdots + X_N}{\sqrt{N}}$$

converges to a random quantity with a distribution of a very specific form, called a *Normal* or *Gaussian* or *bell-shaped* distribution.

There are many such results, but these are the simplest and most famous. They justify the widespread use of means, averages, and moments in probabilistic analysis and the widespread use of Gaussian models.