

# Final thoughts

Lecture 22

December 2



# Looking ahead to statistics (wikipedia)

## Probability

The word ***probability*** derives from the [Latin](#) *probare* (to prove, or to test). Informally, *probable* is one of several words applied to uncertain events or knowledge, being more or less interchangeable with *likely*, *risky*, *hazardous*, *uncertain*, and *doubtful*, depending on the context. *Chance*, *odds*, and *bet* are other words expressing similar notions. As with the [theory of mechanics](#) which assigns precise definitions to such everyday terms as *work* and *force*, so the [theory of probability](#) attempts to quantify the notion of *probable*.

**Statistics** is a broad [mathematical](#) discipline which studies ways to collect, summarize and draw conclusions from [data](#). It is applicable to a wide variety of [academic disciplines](#) from the physical and social [sciences](#) to the [humanities](#), as well as to [business](#), [government](#), and [industry](#).

# Basic Idea of Class

Consider problem with uncertainty.

Come up with appropriate model – realizing the simplifying assumptions you make.

Use this model to answer questions at hand about likelihood of events or expected behavior.

Understand what these results mean.

# Concepts/Tools

- Probability very natural in game a chance and problems with finite outcomes

Combinatorics

- Conditional probability allows you to factor what you know into the problem.

Surprising results – Bayes Rule

# Key Concepts

- Event model

Break down problems into discrete events that you can figure out the probabilities of. Combine to get the probability of interest.

- Mutual Exclusive
- Independence

# Random Variables

Three kinds

discrete, continuous, mixed

Defined by pdf, cdf:

What are properties of pdf/cdf?

How do you use them to calculate probabilities?

# Discrete Random Variables

## Our Favorites

- Uniform
- Bernoulli
- Binomial
- Negative Binomial
- Geometric
- Hypergeometric
- Poisson (approximation to binomial)

# Continuous Random Variables

## Our Favorites

- Normal
- Uniform
- Exponential
- Gamma
- Beta

# Expectation

Average case behavior of a random variable or function of random variables for a given distribution.

How do you calculate  $E(g(X))$  or  $E(g(X, Y))$ ?

Using pdf

Using mgf

# Joint or Multivariate Probability Distributions

Can be discrete, continuous, mixed

Joint distribution

Marginal distribution

Conditional distribution

Derive them.

Apply them.

Use them to calculate probabilities.

# Expectation and Multivariate Distributions

Covariance and Correlation

Conditional Expectation

# Functions of RV

Practical uses

Methods

CDF

mgf

transforms

Interesting case:

Order statistics

# Central Limit Theorem (wiki)

- **Central limit theorems** are a set of weak-convergence results in [probability theory](#). Intuitively, they all express the fact that any sum of many [independent](#) identically distributed [random variables](#) will tend to be distributed according to a particular "attractor distribution". The most important and famous result is simply called *The Central Limit Theorem* which states that if the summed variables have a finite variance then they will be approximately [normally distributed](#). Since many real processes yield distributions with finite [variance](#), this explains the ubiquity of the normal [distribution](#).