

Joint Distributions

Will Song

CS109, Stanford University

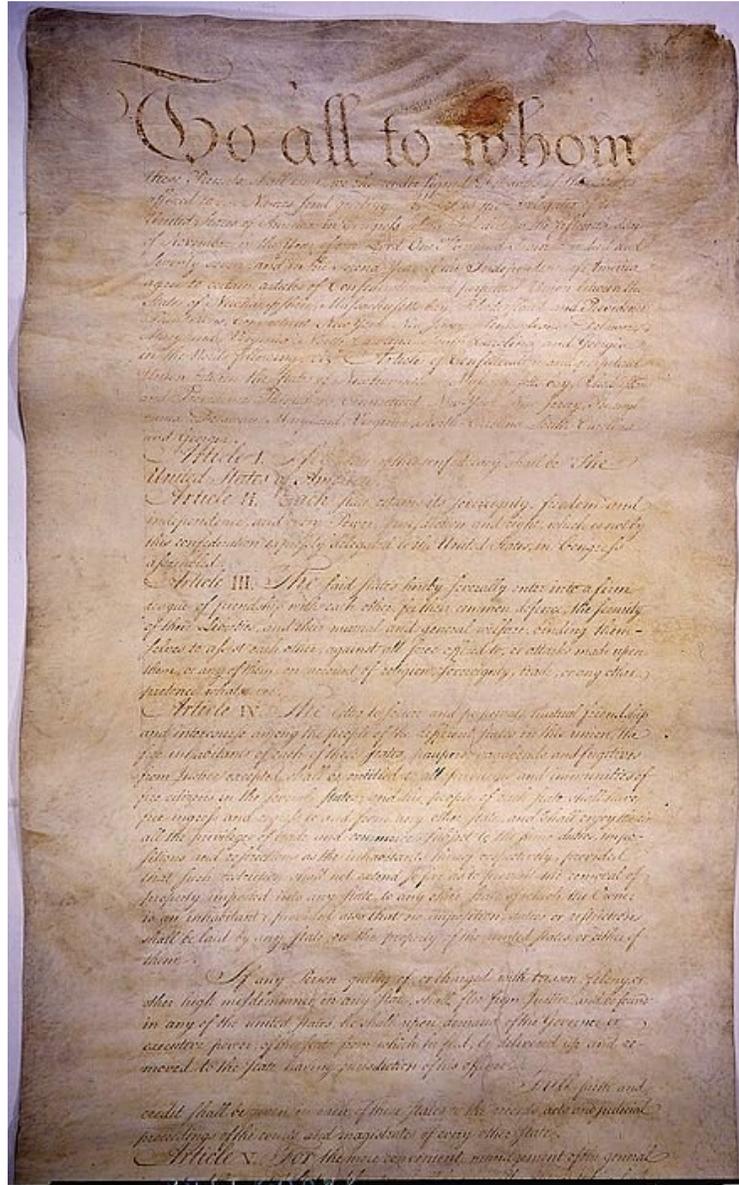
Announcements!

Midterm Next Tuesday 7p—9pm

- Location: 370-370
- Exam preparation going live soon!



Exciting Day!



"The best commentary on the principles of government ever written."
—PRESIDENT AND FOUNDING FATHER THOMAS JEFFERSON

"Read it, underline it, and dog-ear it." —SUPREME COURT JUSTICE ANTONIN SCALIA

THE FEDERALIST PAPERS

John Jay, James Madison,
& Alexander Hamilton

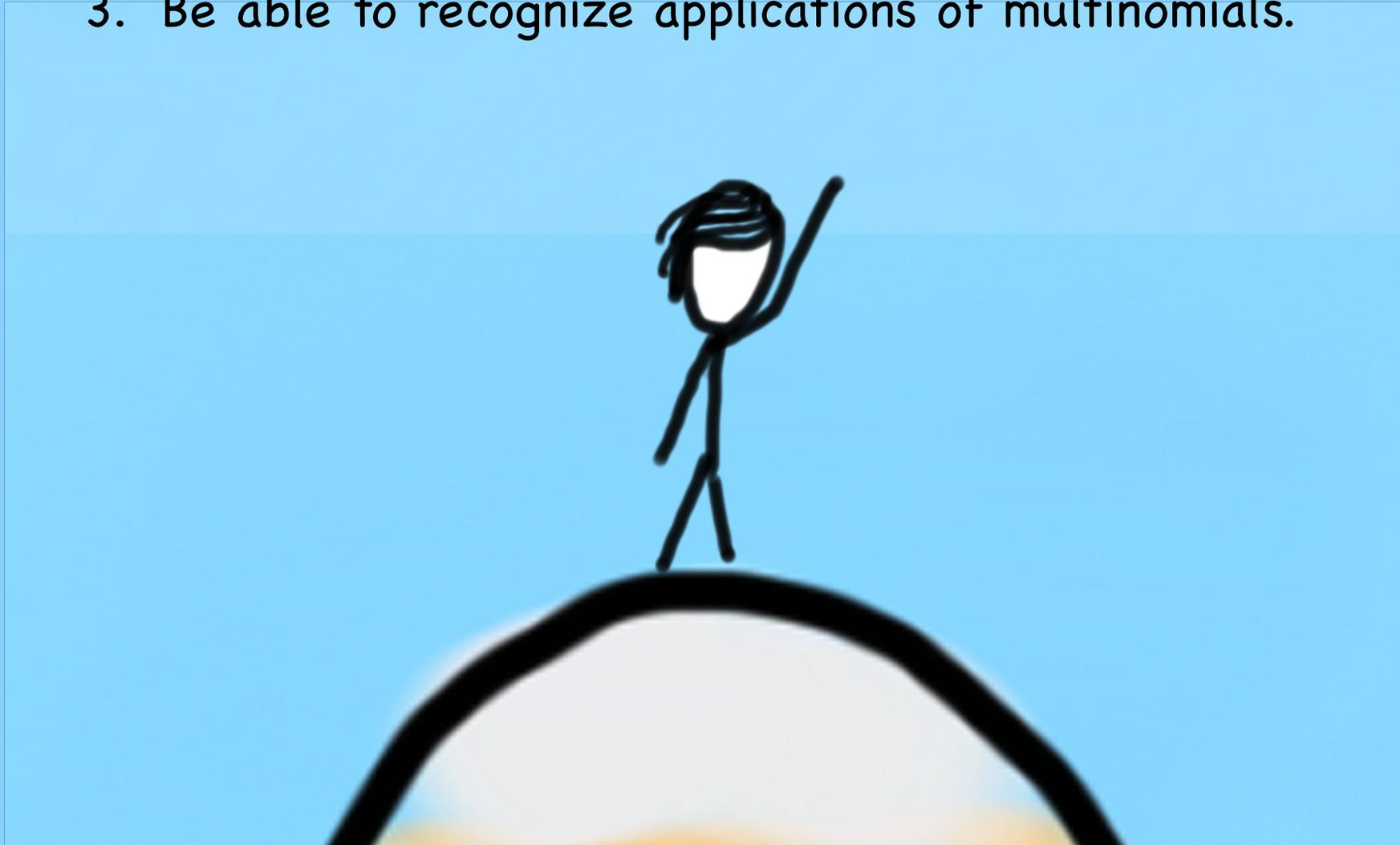


Foreword by
ALAN DERSHOWITZ

The image shows the front cover of the book 'The Federalist Papers'. At the top, there are two quotes: one from President and Founding Father Thomas Jefferson, and another from Supreme Court Justice Antonin Scalia. The title 'THE FEDERALIST PAPERS' is written in large, red, serif capital letters. Below the title is a decorative flourish. Underneath the flourish is a red banner with the authors' names 'John Jay, James Madison, & Alexander Hamilton' in white, italicized serif font. Below the banner are three individual portraits of the authors: John Jay on the left, James Madison in the center, and Alexander Hamilton on the right. At the bottom of the cover, it says 'Foreword by ALAN DERSHOWITZ' in red, serif font.

Learning Goals

1. Be able to Approximate Binomials with a Normal!
2. Be able to reason about joint distributions.
3. Be able to recognize applications of multinomials.



First, some review

Where are we in CS109?

Overview of Topics



Counting
Theory



Core
Probability



Random
Variables



Probabilistic
Models

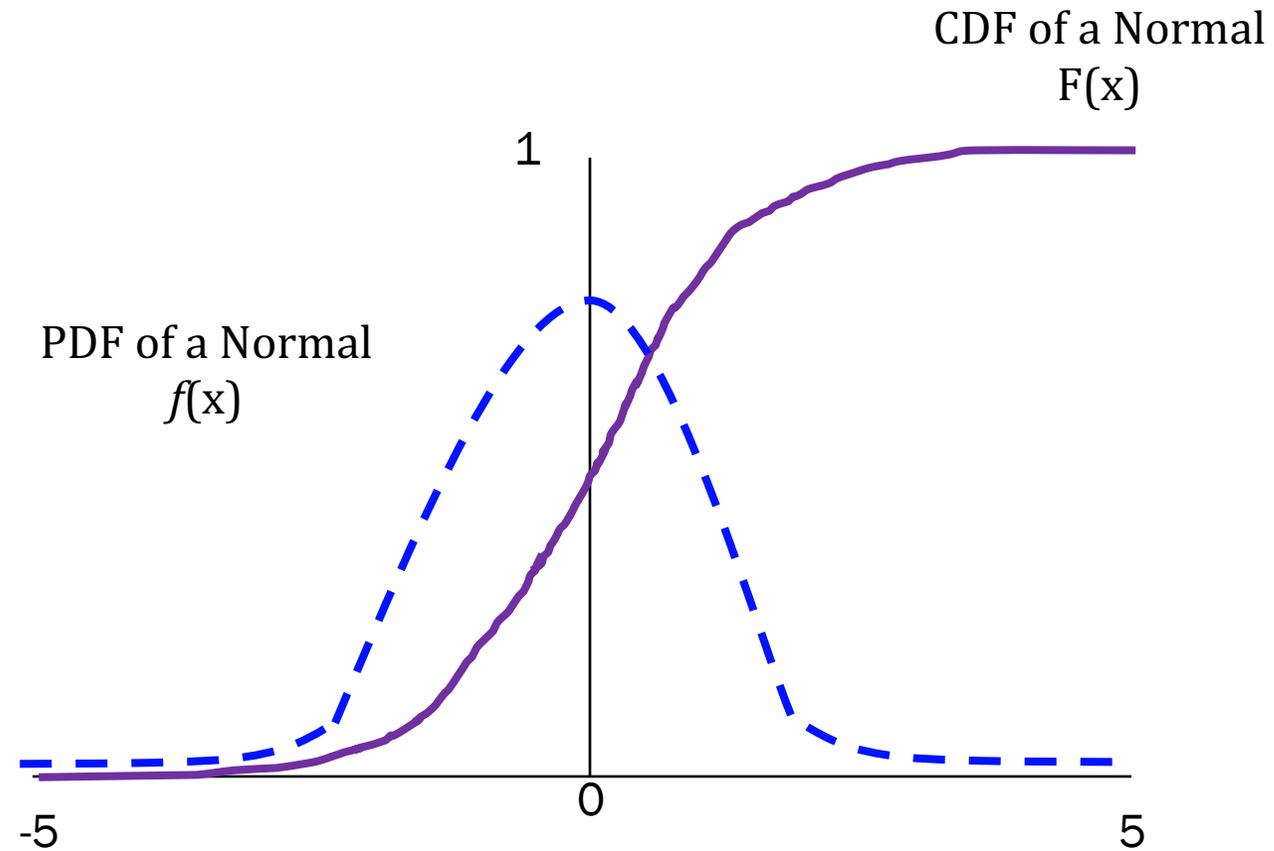


Uncertainty
Theory



Machine
Learning

Normal Distribution



$f(x)$ = derivative of probability

$F(x) = P(X < x)$

Probability Density Function

$$\mathcal{N}(\mu, \sigma^2)$$

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x - \mu)^2}{2\sigma^2}}$$

“exponential”

the distance to the mean

probability density at x

a constant

sigma shows up twice

Does it look less scary like this?

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

This means "e to the power of" and
is common function in code math
libraries

$$f(x) \propto \frac{1}{\sigma} \cdot \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right]$$

This means "proportional to". There is a
constant but there are many cases where we
don't care what it is!

What if you had to take the log of this function?

Cumulative Density Function

$$\mathcal{N}(\mu, \sigma^2)$$

CDF of Standard Normal: A function that has been solved for numerically

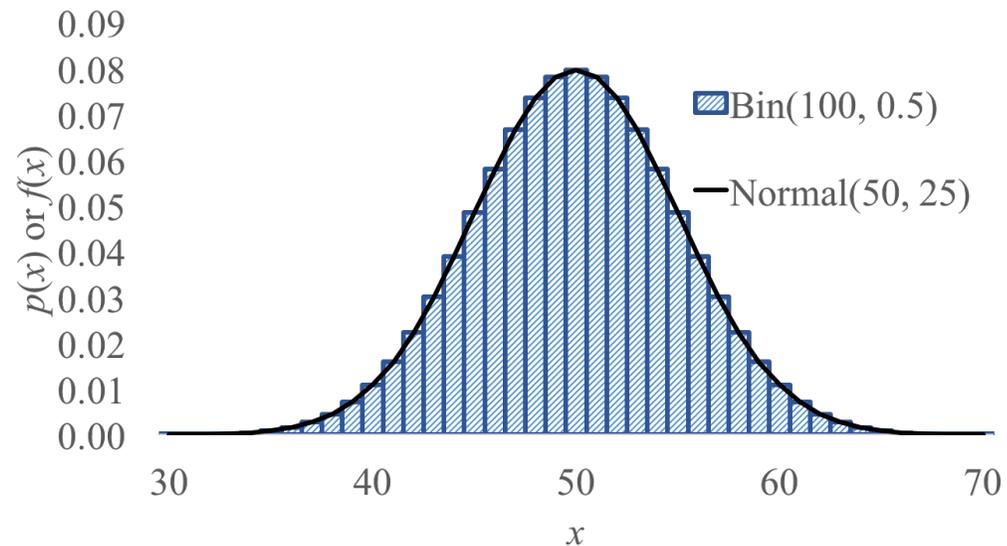
$$F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

The cumulative density function (CDF) of any normal

Don't worry, Normal approximates Binomial



Galton Board



Website testing

- 100 people are given a new website design.
- $X = \#$ people whose time on site increases
- The design actually has no effect, so $P(\text{time on site increases}) = 0.5$ independently.
- CEO will endorse the new design if $X \geq 65$.

What is $P(\text{CEO endorses change})$? Give a numerical approximation.

Approach 1: Binomial

Define

$$X \sim \text{Bin}(n = 100, p = 0.5)$$

Want: $P(X \geq 65)$

Solve

$$P(X \geq 65) \approx 0.0018$$



(this approach is missing something important)

Approach 2: approximate with Normal

Define

$$Y \sim \mathcal{N}(\mu, \sigma^2)$$

$$\mu = np = 50$$

$$\sigma^2 = np(1 - p) = 25$$

$$\sigma = \sqrt{25} = 5$$

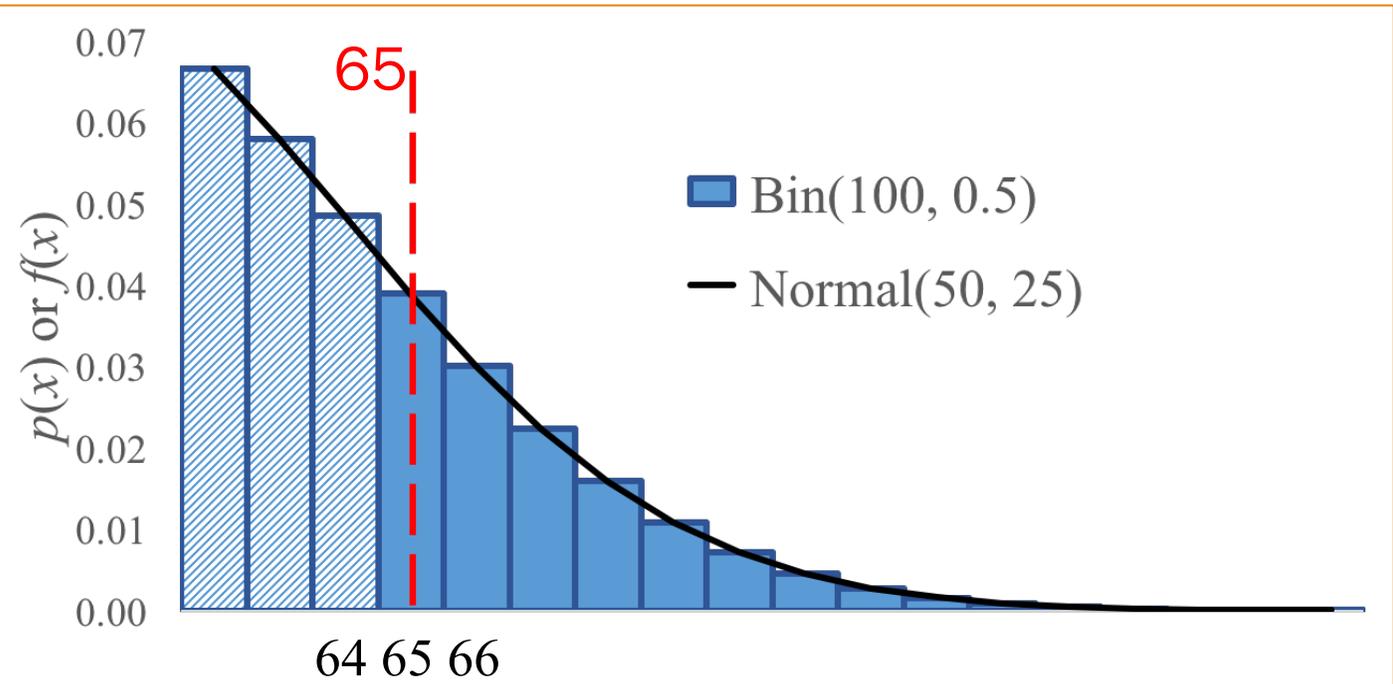
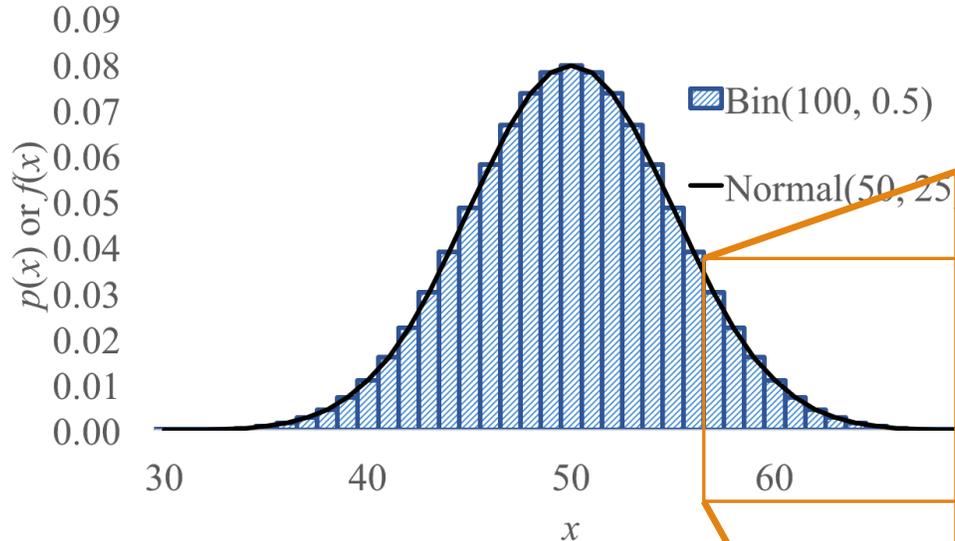
Solve

$$\begin{aligned} P(X \geq 65) &\approx P(Y \geq 65) = 1 - F_Y(65) \\ &= 1 - \Phi\left(\frac{65-50}{5}\right) = 1 - \Phi(3) \approx 0.0013? \end{aligned}$$



Website testing (with continuity correction)

In our website testing, $Y \sim \mathcal{N}(50, 25)$ approximates $X \sim \text{Bin}(100, 0.5)$.



$$P(X \geq 65) \text{ Binomial}$$

$$\approx P(Y \geq 64.5) \text{ Normal}$$

$$\approx 0.0018 \quad \checkmark \text{ the better Approach 2}$$

You must perform a **continuity correction** when approximating a Binomial RV with a Normal RV.

Continuity correction

If $Y \sim \mathcal{N}(np, np(1 - p))$ approximates $X \sim \text{Bin}(n, p)$, how do we approximate the following probabilities?

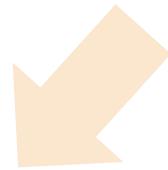
Discrete (e.g., Binomial) probability question	Continuous (Normal) probability question	
$P(X = 6)$	$P(5.5 \leq Y \leq 6.5)$	$P(5.5 \leq Y \leq 6.5)$
$P(X \geq 6)$	$P(Y \geq 5.5)$	
$P(X > 6)$	$P(Y \geq 6.5)$	
$P(X < 6)$	$P(Y \leq 5.5)$... 5 6 7 ...
$P(X \leq 6)$	$P(Y \leq 6.5)$	

Who gets to approximate?

$$X \sim \text{Bin}(n, p)$$

$$E[X] = np$$

$$\text{Var}(X) = np(1 - p)$$



$$Y \sim \text{Poi}(\lambda)$$

$$\lambda = np$$

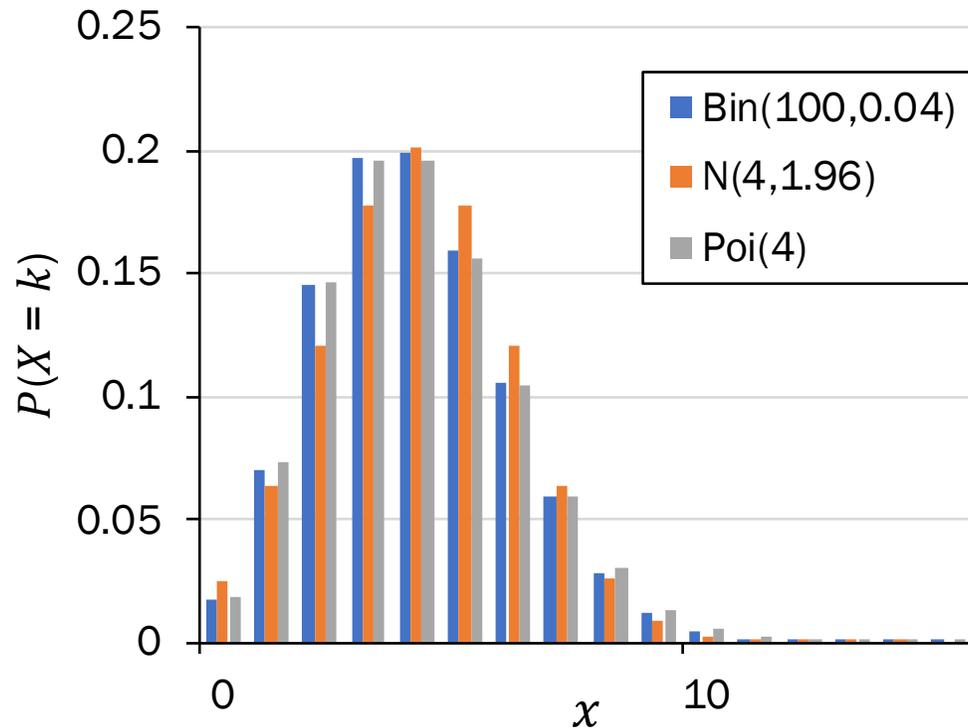
?

$$Y \sim \mathcal{N}(\mu, \sigma^2)$$

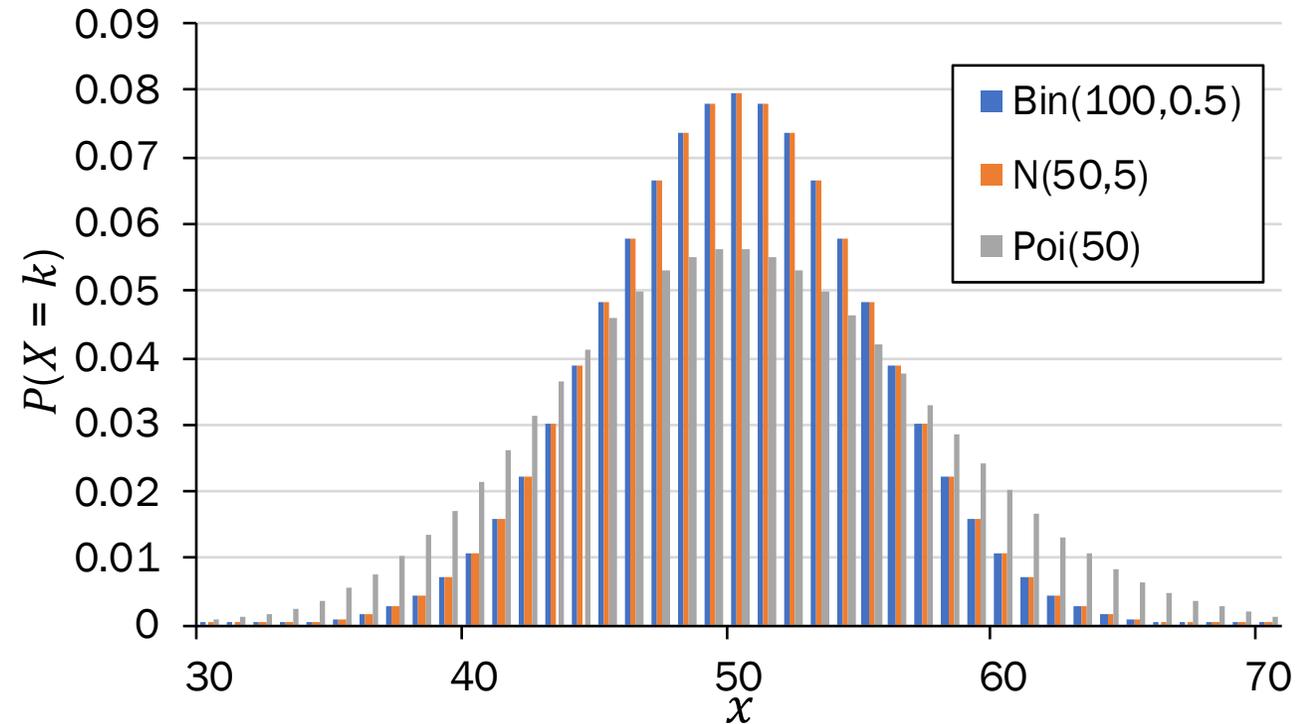
$$\mu = np$$

$$\sigma^2 = np(1 - p)$$

Who gets to approximate?



Poisson approximation
 n large (> 20), p small (< 0.05)
slight dependence okay



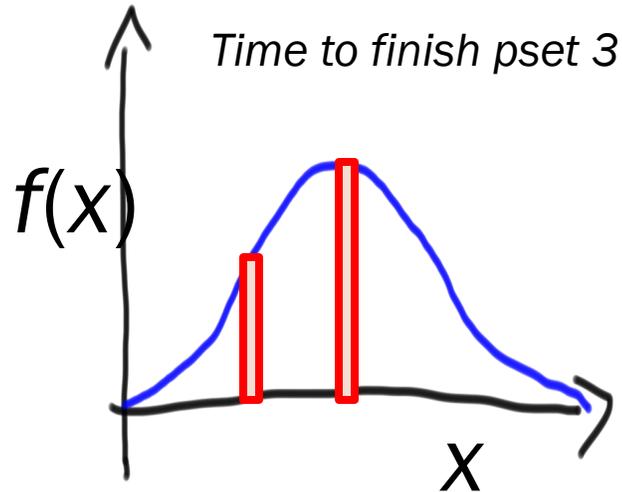
Normal approximation
 n large (> 20), variance large ($np(1-p) > 10$)
independence

1. If there is a choice, either is fine.
2. When using Normal to approximate a discrete RV, use a continuity correction.

Relative Probability of Continuous Variables

$X =$ time to finish pset 3

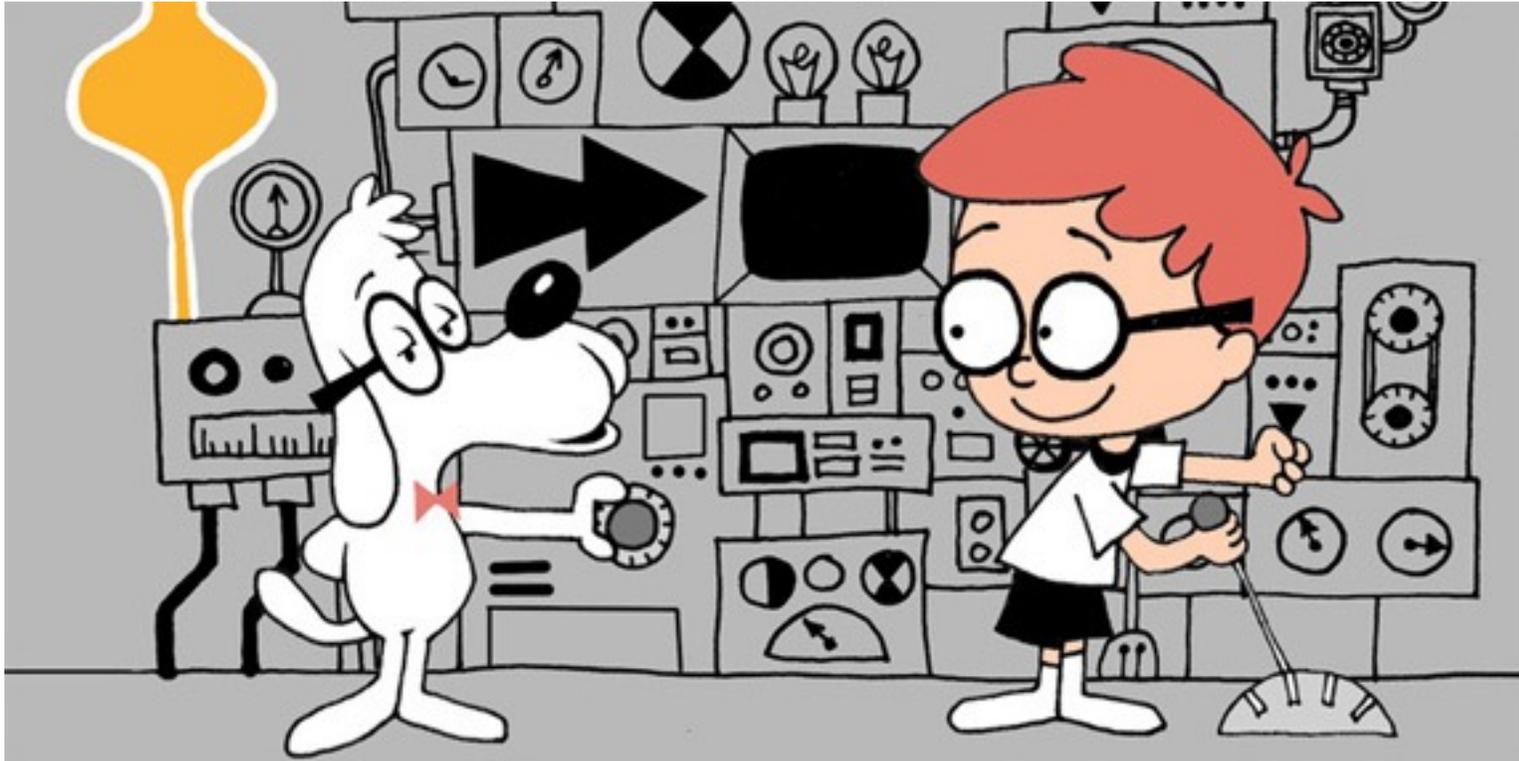
$X \sim N(\mu = 10, \sigma^2 = 2)$



How much more likely are you to complete in 10 hours than in 5?

$$\begin{aligned}\frac{P(X = 10)}{P(X = 5)} &= \frac{\varepsilon f(X = 10)}{\varepsilon f(X = 5)} \\ &= \frac{f(X = 10)}{f(X = 5)} \\ &= \frac{\frac{1}{\sqrt{2\sigma^2\pi}} e^{-\frac{(10-\mu)^2}{2\sigma^2}}}{\frac{1}{\sqrt{2\sigma^2\pi}} e^{-\frac{(5-\mu)^2}{2\sigma^2}}} \\ &= \frac{\frac{1}{\sqrt{4\pi}} e^{-\frac{(10-10)^2}{4}}}{\frac{1}{\sqrt{4\pi}} e^{-\frac{(5-10)^2}{4}}} \\ &= \frac{e^0}{e^{-\frac{25}{4}}} = 518\end{aligned}$$

Recall the good times



Permutations

$n!$

How many ways are
there to order n
objects?

Ways to put elements into fixed size containers

How many ways are there to put n objects into r buckets such that:

n_1 go into bucket 1

n_2 go into bucket 2

...

n_r go into bucket r ?

$$\frac{n!}{n_1!n_2!\dots n_r!} = \binom{n}{n_1, n_2, \dots, n_r}$$

Note: Multinomial > Binomial

Counting unordered objects

Binomial coefficient

How many ways are there to order n objects such that k are indistinct and $(n-k)$ are indistinct

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Called the binomial coefficient because of something from Algebra

Multinomial coefficient

How many ways are there to order n objects such that n_1 are indistinct, n_2 are indistinct etc.

$$\binom{n}{n_1, n_2, \dots, n_r} = \frac{n!}{n_1! n_2! \cdots n_r!}$$

Multinomials generalize Binomials for counting.

End of review

Machine Learning

Uncertainty Theory

Single Random
Variables

Probabilistic Models

Counting

Probability Fundamentals

Discrete Probabilistic Models

The world is full of interesting probability problems



Have multiple random variables interacting with one another

Multiple Random Variables. Start of Digital Revolution



Multiple Random Variables. Start of Digital Revolution

Conditions that match your symptoms

UNDERSTANDING YOUR RESULTS [i](#)

Migraine headache (adult)



Moderate match



Acute Sinusitis



Fair match



Stroke



Fair match



Gender **Male**

Age **30**

[Edit](#)

My Symptoms

[Edit](#)

dizziness, one sided headache

Joint probability mass functions

Roll two 6-sided dice, yielding values X and Y .



X

random variable

$$P(X = 1)$$

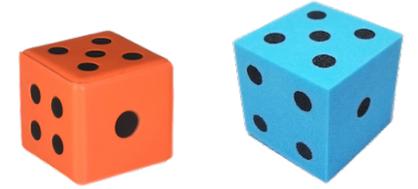
probability of
an event

$$P(X = k)$$

probability mass function

Joint probability mass functions

Roll two 6-sided dice, yielding values X and Y .



X

random variable

$$P(X = 1)$$

probability of
an event

$$P(X = k)$$

probability mass function

X, Y

random variables

$$P(X = 1 \text{ and } Y = 6)$$

$$P(X = 1, Y = 6)$$

recall: the comma

probability of the intersection
of two events

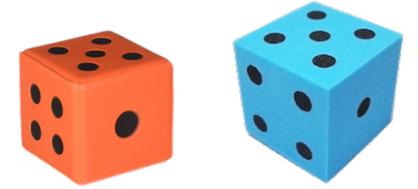
$$P(X = a, Y = b)$$

joint probability mass function

Two dice

Roll two 6-sided dice, yielding values X and Y .

1. What is the joint PMF of X and Y ?



$$P(X = a, Y = b) = 1/36 \quad (a, b) \in \{(1,1), \dots, (6,6)\}$$

		X					
		1	2	3	4	5	6
Y	1	1/36	1/36
	2
	3
	4
	5
	6	1/36	1/36

$P(X = 4, Y = 3)$

Probability table

- All possible outcomes for several discrete RVs
- Not parametric (e.g., parameter p in $\text{Ber}(p)$)

Dating at Stanford. Data from a few years ago

	Single	In a relationship	It's complicated
Freshman	0.13	0.08	0.02
Sophomore	0.17	0.11	0.02
Junior	0.09	0.10	0.02
Senior	0.02	0.07	0.01
5+	0.06	0.09	0.04

Joint is Complete Information!

	Single	Relationship	Complicated
Frosh	0.13	0.08	0.02
Soph	0.17	0.11	0.02
Junior	0.09	0.10	0.02
Senior	0.02	0.07	0.01
5+	0.06	0.09	0.04



A joint distribution is complete information. It can be used to answer any probability question.

Joint table: mutually exclusive and covers sample space.

	Single	Relationship	Complicated
Frosh	0.13	0.08	0.02
Soph	0.17	0.11	0.02
Junior	0.09	0.10	0.02
Senior	0.02	0.07	0.01
5+	0.06	0.09	0.04

Each combination is mutually exclusive, and they span the sample space

$$\sum_{x \in X} \sum_{y \in Y} P(x, y) = 1$$

X is dating status.
Y is year.

Joint table: mutually exclusive and covers sample space.

	Single	Relationship	Complicated
Frosh	0.13	0.08	0.02
Soph	0.17	0.11	0.02
Junior	0.09	?	0.02
Senior	0.02	0.07	0.01
5+	0.06	0.09	0.04

Each combination is mutually exclusive, and they span the sample space

$$\sum_{x \in X} \sum_{y \in Y} P(x, y) = 1$$

X is dating status.
Y is year.

Joint table: mutually exclusive and covers sample space.

	Single	Relationship	Complicated
Frosh	0.13	0.08	0.02
Soph	0.17	0.11	0.02
Junior	0.09	0.10	0.02
Senior	0.02	0.07	0.01
5+	0.06	0.09	0.04

Each combination is mutually exclusive, and they span the sample space

$$\sum_{x \in X} \sum_{y \in Y} P(x, y) = 1$$

X is dating status.
Y is year.

What is the probability someone is in a relationship?

	Single	Relationship	Complicated
Frosh	0.13	0.08	0.02
Soph	0.17	0.11	0.02
Junior	0.09	0.10	0.02
Senior	0.02	0.07	0.01
5+	0.06	0.09	0.04

We can use the law of total probability!
X is dating status. Y is year.

$$P(X = \text{relation}) = \sum_{y \in Y} P(X = \text{relation}, Y = y)$$

$$P(X = \text{single}) = \sum_{y \in Y} P(X = \text{single}, Y = y)$$

$$P(Y = \text{frosh}) = \sum_{x \in X} P(X = x, Y = \text{frosh}) \quad P(Y = \text{soph}) = \sum_{x \in X} P(X = x, Y = \text{soph})$$

Welcome the marginal

Marginal Distribution

For two discrete joint random variables X and Y , the **joint probability mass function** is defined as:

$$P(X = a, Y = b)$$

The **marginal distributions** of the joint PMF are defined as:

$$P(X = a) = \sum_y P(X = a, Y = y)$$

$$P(Y = b) = \sum_x P(X = x, Y = b)$$

Use marginal distributions to get a 1-D RV from a joint PMF.

What about 3 Random Variables?

D is disease, S is can smell, F is fever status

$D = 0$

	$S = 0$	$S = 1$
$F = \text{none}$	0.024	0.783
$F = \text{low}$	0.003	0.092
$F = \text{high}$	0.001	0.046

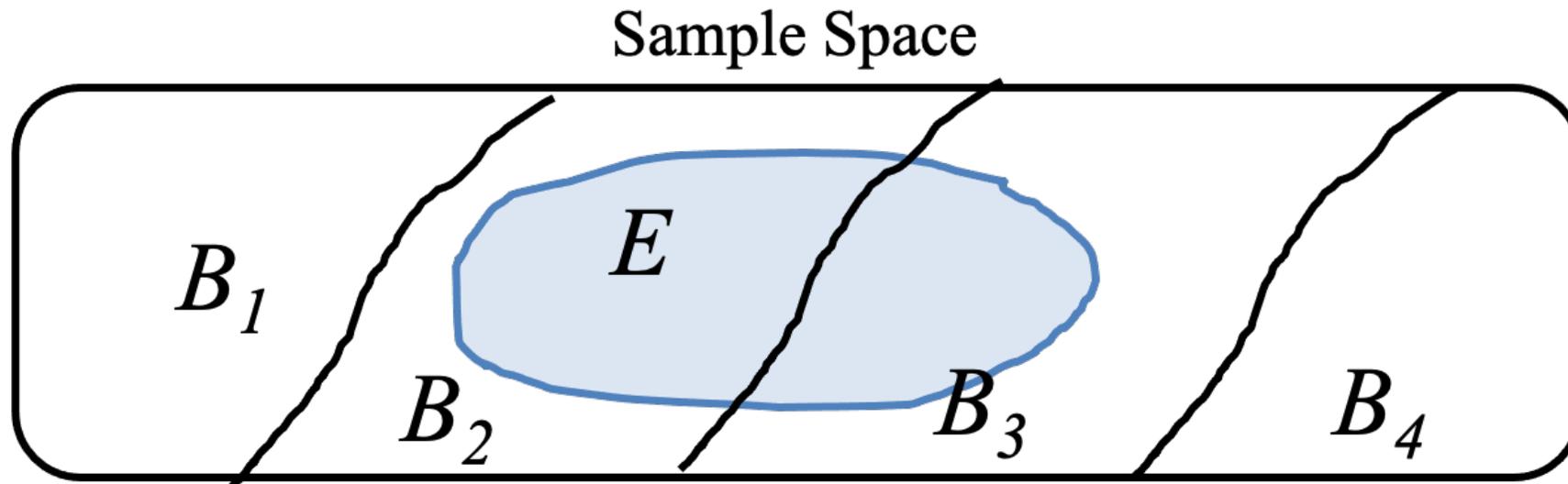
$D = 1$

	$S = 0$	$S = 1$
$F = \text{none}$	0.006	0.014
$F = \text{low}$	0.005	0.011
$F = \text{high}$	0.004	0.011

$$P(D = 1) = \sum_f \sum_s P(D = 1, F = f, S = s)$$

Marginal Distribution. Law of Total Probability for RVs

$$P(X = a) = \sum_y P(X = a, Y = y)$$



Why is that called the marginal?

Key limitation of the joint: it is too big

What about 10 Random Variables?

Imagine you have **10 discrete** RVs which can each take on **5 values**

$$\# \text{ unique assignments} = 5^{10}$$

10 million entries in your joint table.

So, we are going to need models ...

... **probabilistic models** ...

Sometimes the structure of the variables suggests a more efficient representation

Roll 100 dice.

X_1 = How many 1s?

X_2 = How many 2s?

X_3 = How many 3s?

X_4 = How many 4s?

X_5 = How many 5s?

X_6 = How many 6s?

How big is the joint table?

Multinomial RV

Making Binomial Bigger.

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Binomial # of ways of ordering the successes Probability of each ordering of k successes is equal + mutually exclusive

$$P(X = k) = \binom{n}{k} \binom{n-k}{n-k} p^k (1 - p)^{n-k}$$

Binomial # of ways of ordering the successes # of ways of ordering the fails Probability of each ordering of k successes is equal + mutually exclusive

$$P(X = k) = \binom{n}{k} p^k \binom{n-k}{n-k} (1 - p)^{n-k}$$

Case 1 # of ways and probabilities. Case 2 # of ways and probabilities

Generalized Binomial

$$\binom{n}{k} p^k \binom{n-k}{n-k} (1-p)^{n-k}$$

Case 1 # of ways and probabilities.

Case 2 # of ways and probabilities

$$\binom{n}{k_1} p_1^{k_1} \binom{n-k_1}{k_2} p_2^{k_2} \binom{n-k_2-k_1}{k_3} p_3^{k_3}$$

Case 1 # of ways and probabilities.

Case 2 # of ways and probabilities

Case 3 # of ways and probabilities

Notation Detail:

$$\binom{n}{k_1} \binom{n-k_1}{k_2} \binom{n-k_2-k_1}{k_3} = \binom{n}{k_1, k_2, k_3}$$

Fine Print

$$\begin{aligned} k_1 + k_2 + k_3 &= n \\ p_1 + p_2 + p_3 &= 1 \end{aligned}$$

Probability

Binomial RV

What is the probability of getting k successes and $n - k$ failures in n trials?

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Binomial # of ways of ordering the successes

Probability of each ordering of k successes is equal + mutually exclusive

Multinomial RV

What is the probability of getting c_1 of outcome 1, c_2 of outcome 2, ..., and c_m of outcome m in n trials?

$$\binom{n}{c_1, c_2, \dots, c_m} p_1^{c_1} p_2^{c_2} \dots p_m^{c_m}$$

Multinomial RVs also generalize Binomial RVs for probability!

Multinomial Random Variable?

Consider an experiment of n independent trials:

- Each trial results in one of m outcomes. $P(\text{outcome } i) = p_i$, $\sum_{i=1}^m p_i = 1$
- Let $X_i = \#$ trials with outcome i

Joint PMF

$$P(X_1 = c_1, X_2 = c_2, \dots, X_m = c_m) =$$

where $\sum_{i=1}^m c_i = n$ and $\sum_{i=1}^m p_i = 1$

Multinomial Random Variable?

Consider an experiment of n independent trials:

- Each trial results in one of m outcomes. $P(\text{outcome } i) = p_i$, $\sum_{i=1}^m p_i = 1$
- Let $X_i = \#$ trials with outcome i

Joint PMF

$$P(X_1 = c_1, X_2 = c_2, \dots, X_m = c_m) =$$

$$p_1^{c_1} p_2^{c_2} \cdots p_m^{c_m}$$

where $\sum_{i=1}^m c_i = n$ and $\sum_{i=1}^m p_i = 1$

↑
Probability of each ordering is equal + mutually exclusive

Multinomial Random Variable

Consider an experiment of n independent trials:

- Each trial results in one of m outcomes. $P(\text{outcome } i) = p_i$, $\sum_{i=1}^m p_i = 1$
- Let $X_i = \#$ trials with outcome i

Joint PMF

$$P(X_1 = c_1, X_2 = c_2, \dots, X_m = c_m) = \binom{n}{c_1, c_2, \dots, c_m} p_1^{c_1} p_2^{c_2} \dots p_m^{c_m}$$

where $\sum_{i=1}^m c_i = n$ and $\sum_{i=1}^m p_i = 1$

Multinomial # of ways of ordering the outcomes

Probability of each ordering is equal + mutually exclusive

Hello dice rolls, my old friends

$$P(X_1 = c_1, X_2 = c_2, \dots, X_m = c_m) = \binom{n}{c_1, c_2, \dots, c_m} p_1^{c_1} p_2^{c_2} \dots p_m^{c_m}$$

A 6-sided die is rolled 7 times.

What is the probability of getting:

- 1 one
- 0 threes
- 0 fives
- 1 two
- 2 fours
- 3 sixes



Hello dice rolls, my old friends

$$P(X_1 = c_1, X_2 = c_2, \dots, X_m = c_m) = \binom{n}{c_1, c_2, \dots, c_m} p_1^{c_1} p_2^{c_2} \dots p_m^{c_m}$$

A 6-sided die is rolled 7 times.

What is the probability of getting:

- 1 one
- 1 two
- 0 threes
- 2 fours
- 0 fives
- 3 sixes

$$P(X_1 = 1, X_2 = 1, X_3 = 0, X_4 = 2, X_5 = 0, X_6 = 3)$$

$$= \binom{7}{1, 1, 0, 2, 0, 3} \left(\frac{1}{6}\right)^1 \left(\frac{1}{6}\right)^1 \left(\frac{1}{6}\right)^0 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^0 \left(\frac{1}{6}\right)^3 = 420 \left(\frac{1}{6}\right)^7$$

Hello dice rolls, my old friends

A 6-sided die is rolled 7 times.

What is the probability of getting:

- 1 one • 0 threes • 0 fives
- 1 two • 2 fours • 3 sixes

of times
a six appears

$$P(X_1 = 1, X_2 = 1, X_3 = 0, X_4 = 2, X_5 = 0, X_6 = 3)$$

$$= \binom{7}{1, 1, 0, 2, 0, 3} \left(\frac{1}{6}\right)^1 \left(\frac{1}{6}\right)^1 \left(\frac{1}{6}\right)^0 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^0 \left(\frac{1}{6}\right)^3 = 420 \left(\frac{1}{6}\right)^7$$

choose where
the sixes appear

probability
of rolling a six

this many times

Parameters of a Multinomial RV?

$X \sim \text{Bin}(n, p)$ has parameters $n, p \dots$

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

p : probability of success outcome on a single trial

A Multinomial RV has parameters n, p_1, p_2, \dots, p_m (Note $p_m = 1 - \sum_{i=1}^{m-1} p_i$)

$$P(X_1 = c_1, X_2 = c_2, \dots, X_m = c_m) = \binom{n}{c_1, c_2, \dots, c_m} p_1^{c_1} p_2^{c_2} \dots p_m^{c_m}$$

p_i : probability of outcome i on a single trial

The Federalist Papers

Intro to Natural Language Processing

Probabilistic text analysis

Ignoring the order of words...

What is the probability of any given word that you write in English?

- $P(\text{word} = \text{"the"}) > P(\text{word} = \text{"pokemon"})$
- $P(\text{word} = \text{"Stanford"}) > P(\text{word} = \text{"Cal"})$

Probabilities of *counts* of words = Multinomial distribution



A document is a large multinomial.

(according to the Global Language Monitor, there are 988,968 words in the English language used on the internet.)

Model text as a multinomial

Example document:

“Pay for Viagra with a credit-card. Viagra is great. So are credit-cards. Risk free Viagra. Click for free.”

$n = 18$

$$P \left(\begin{array}{l} \text{Viagra} = 2 \\ \text{Free} = 2 \\ \text{Risk} = 1 \\ \text{Credit-card: } 2 \\ \dots \\ \text{For} = 2 \end{array} \middle| \text{spam} \right) = \frac{n!}{2!2! \dots 2!} p_{\text{viagra}}^2 p_{\text{free}}^2 \dots p_{\text{for}}^2$$

It's a Multinomial!

Probability of seeing
this document | spam

The probability of a word in
spam email being viagra

Who wrote the federalist papers?



Old and New Analysis

Authorship of the Federalist Papers

- 85 essays advocating ratification of the US constitution
- Written under the pseudonym “Publius” (really, Alexander **Hamilton**, James **Madison**, John **Jay**)



Who wrote which essays?

- Analyze probability of words in each essay and compare against word distributions from known writings of three authors

Who wrote Federalist Paper 53?

madison.txt

```
1 To the People of the State of New York:
2
3 AMONG the numerous advantages promised by a
wellconstructed Union, none deserves to be more
accurately developed than its tendency to break
and control the violence of faction. The friend
of popular governments never finds himself so
much alarmed for their character and fate, as
when he contemplates their propensity to this
dangerous vice. He will not fail, therefore, to
set a due value on any plan which, without
violating the principles to which he is attached,
provides a proper cure for it. The instability,
injustice, and confusion introduced into the
public councils, have, in truth, been the mortal
diseases under which popular governments have
everywhere perished; as they continue to be the
favorite and fruitful topics from which the
adversaries to liberty derive their most specious
declamations. The valuable improvements made by
the American constitutions on the popular models,
both ancient and modern, cannot certainly be too
much admired; but it would be an unwarrantable
partiality, to contend that they have as
effectually obviated the danger on this side, as
was wished and expected. Complaints are
everywhere heard from our most considerate and
virtuous citizens, equally the friends of public
and private faith, and of public and personal
liberty, that our governments are too unstable,
that the public good is disregarded in the
conflicts of rival parties, and that measures are
too often decided, not according to the rules of
justice and the rights of the minor party, but by
the superior force of an interested and
overbearing majority. However anxiously we may
wish that these complaints had no foundation, the
evidence, of known facts will not permit us to
deny that they are in some degree true. It will
be found, indeed, on a candid review of our
situation, that some of the distresses under
which we labor have been erroneously charged on
the operation of our governments; but it will be
found, at the same time, that other causes will
not alone account for many of our heaviest
misfortunes; and, particularly, for that
prevailing and increasing distrust of public
```

hamilton.txt

```
1 The Utility of the Union in Respect to Commercial
Relations and a Navy
2 Hamilton for the Independent Journal.
3
4 To the People of the State of New York:
5 THE importance of the Union, in a commercial
light, is one of those points about which there
is least room to entertain a difference of
opinion, and which has, in fact, commanded the
most general assent of men who have any
acquaintance with the subject. This applies as
well to our intercourse with foreign countries as
with each other.
6
7 There are appearances to authorize a supposition
that the adventurous spirit, which distinguishes
the commercial character of America, has already
excited uneasy sensations in several of the
maritime powers of Europe. They seem to be
apprehensive of our too great interference in
that carrying trade, which is the support of
their navigation and the foundation of their
naval strength. Those of them which have colonies
in America look forward to what this country is
capable of becoming, with painful solicitude.
They foresee the dangers that may threaten their
American dominions from the neighborhood of
States, which have all the dispositions, and
would possess all the means, requisite to the
creation of a powerful marine. Impressions of
this kind will naturally indicate the policy of
fostering divisions among us, and of depriving
us, as far as possible, of an active commerce in
our own bottoms. This would answer the threefold
purpose of preventing our interference in their
navigation, of monopolizing the profits of our
trade, and of clipping the wings by which we
might soar to a dangerous greatness. Did not
prudence forbid the detail, it would not be
difficult to trace, by facts, the workings of
this policy to the cabinets of ministers.
8
9 If we continue united, we may counteract a policy
so unfriendly to our prosperity in a variety of
ways. By prohibitory regulations, extending, at
the same time, throughout the States, we may
oblige foreign countries to bid against each
```

unknown.txt

```
1 To the People of the State of New York:
2 I SHALL here, perhaps, be reminded of a current
observation, that where annual elections end,
tyranny begins. If it be true, as has often
been remarked, that sayings which become
proverbial are generally founded in reason, it
is not less true, that when once established,
they are often applied to cases to which the
reason of them does not extend. I need not look
for a proof beyond the case before us. What is
the reason on which this proverbial observation
is founded? No man will subject himself to the
ridicule of pretending that any natural
connection subsists between the sun or the
seasons, and the period within which human
virtue can bear the temptations of power.
Happily for mankind, liberty is not, in this
respect, confined to any single point of time;
but lies within extremes, which afford
sufficient latitude for all the variations which
may be required by the various situations and
circumstances of civil society. The election of
magistrates might be, if it were found
expedient, as in some instances it actually has
been, daily, weekly, or monthly, as well as
annual; and if circumstances may require a
deviation from the rule on one side, why not
also on the other side? Turning our attention
to the periods established among ourselves, for
the election of the most numerous branches of
the State legislatures, we find them by no
means coinciding any more in this instance,
than in the elections of other civil
magistrates. In Connecticut and Rhode Island,
the periods are half-yearly. In the other States,
South Carolina excepted, they are annual. In
South Carolina they are biennial as is proposed
in the federal government. Here is a difference,
as four to one, between the longest and
shortest periods; and yet it would be not easy
to show, that Connecticut or Rhode Island is
better governed, or enjoys a greater share of
rational liberty, than South Carolina; or that
either the one or the other of these States is
distinguished in these respects, and by these
causes, from the States whose elections are
different from both. In searching for the
grounds of this doctrine, I can discover but
one, and that is wholly inapplicable to our
case. The important distinction so well
```

Where to start?

We have words, we want to know probability of authorship. We also know probability of words given author...



Well hello again...

Who wrote Federalist Paper 53?

Prob Document given Hamilton

Prior belief it was Hamilton

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}$$

Prob Hamilton given Document

Prob of the document???

Who wrote Federalist Paper 53?

Model document as a multinomial where we care about count of words

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}$$

Who wrote Federalist Paper 53?

Loop over unique words

Prob hamilton would write word i

Number of times word i is in the doc

Prior belief it was Hamilton

Prob Hamilton given Document

Prob of the document???

$$P(H|D) = \frac{\binom{n}{c_1 \dots c_k} \cdot \prod_i h_i^{c_i} \cdot P(H)}{P(D)}$$

Who wrote Federalist Paper 53?

Prob that Hamilton wrote it

$$\begin{aligned} P(H|D) &= \frac{P(D|H)P(H)}{P(D)} \\ &= \frac{P(H) \cdot \binom{n}{c_1 \dots c_m} \cdot \prod_i h_i^{c_i}}{P(D)} \end{aligned}$$

Prob that Madison wrote it

$$\begin{aligned} P(M|D) &= \frac{P(D|M)P(M)}{P(D)} \\ &= \frac{P(M) \cdot \binom{n}{c_1 \dots c_m} \cdot \prod_i m_i^{c_i}}{P(D)} \end{aligned}$$

$$\begin{aligned} \frac{P(H|D)}{P(M|D)} &= \frac{P(H) \cdot \binom{n}{c_1 \dots c_m} \cdot \prod_i h_i^{c_i}}{P(M) \cdot \binom{n}{c_1 \dots c_m} \cdot \prod_i m_i^{c_i}} \\ &= \frac{\prod_i h_i^{c_i}}{\prod_i m_i^{c_i}} \end{aligned}$$

(Detail) The power of Log

Prob that Hamilton wrote it

$$\begin{aligned} P(H|D) &= \frac{P(D|H)P(H)}{P(D)} \\ &= \frac{P(H) \cdot \binom{n}{c_1 \dots c_m} \cdot \prod_i h_i^{c_i}}{P(D)} \end{aligned}$$

Prob that Madison wrote it

$$\begin{aligned} P(M|D) &= \frac{P(D|M)P(M)}{P(D)} \\ &= \frac{P(M) \cdot \binom{n}{c_1 \dots c_m} \cdot \prod_i m_i^{c_i}}{P(D)} \end{aligned}$$

Numerically unstable.

$$\frac{\prod_i \boxed{h_i^{c_i}}}{\prod_i \boxed{m_i^{c_i}}} \begin{array}{l} \leftarrow h_i \leq 1 \\ \leftarrow m_i \leq 1 \end{array}$$

The numerator and denominator will almost vanish!

Instead, take a log!

$$\sum_i c_i \log(h_i) - c_i \log(m_i) = -1353.6. \text{ Madison?}$$

If positive, then ratio on left is greater than 1.

Fin!