

Not Everything is Equally Likely

- Say n balls are placed in m urns
 - Each ball is equally likely to be placed in any urn
- Counts of balls in urns are not equally likely!
 - Example: two balls (A and B) placed with equal likelihood in two urns (Urn 1 and Urn 2)
 - Possibilities:

Urn 1	Urn 2
A, B	-
A	B
B	A
-	A, B

Counts:

Urn 1	Urn 2	Prob
2	0	1/4
1	1	2/4
0	2	1/4

Dice – Our Misunderstood Friends

- Roll two 6-sided dice, yielding values D_1 and D_2
- Let E be event: $D_1 + D_2 = 4$
- What is $P(E)$?
 - $|S| = 36$, $E = \{(1, 3), (2, 2), (3, 1)\}$
 - $P(E) = 3/36 = 1/12$
- Let F be event: $D_1 = 2$
- $P(E, \text{ given } F \text{ already observed})$?
 - $S = \{(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6)\}$
 - $E = \{(2, 2)\}$
 - $P(E, \text{ given } F \text{ already observed}) = 1/6$

Conditional Probability

- **Conditional probability** is probability that E occurs *given* that F has already occurred
 - “Conditioning on F”
- Written as $P(E \mid F)$
 - Means “P(E, given F already observed)”
 - Sample space, S, reduced to those elements consistent with F (i.e. $S \cap F$)
 - Event space, E, reduced to those elements consistent with F (i.e. $E \cap F$)
- With equally likely outcomes:

$$P(E \mid F) = \frac{\# \text{ of outcomes in } E \text{ consistent with } F}{\# \text{ of outcomes in } S \text{ consistent with } F} = \frac{|EF|}{|SF|} = \frac{|EF|}{|F|}$$

Conditional Probability

- General definition:

$$P(E | F) = \frac{P(EF)}{P(F)}$$

where $P(F) > 0$

- Holds even when outcomes are not equally likely
- Implies: $P(EF) = P(E | F) P(F)$ (chain rule)
- What if $P(F) = 0$?
 - $P(E | F)$ undefined
 - *Congratulations! You observed the impossible!*

Generalized Chain Rule

- General definition of Chain Rule:

$$\begin{aligned} P(E_1 E_2 E_3 \dots E_n) \\ = P(E_1) P(E_2 \mid E_1) P(E_3 \mid E_1 E_2) \dots P(E_n \mid E_1 E_2 \dots E_{n-1}) \end{aligned}$$

- Ross calls this the “multiplication rule”
- You can call it either (just be consistent)

Slicing Up the Spam

- 24 emails are sent 6 each to 4 users.
- 10 of the 24 emails are spam.
 - All possible outcomes equally likely
 - E = user 1 receives 3 spam emails

- What is P(E)?

$$\frac{\binom{10}{3}\binom{14}{3}}{\binom{24}{6}} \approx 0.3245$$

Slicing Up the Spam

- 24 emails are sent 6 each to 4 users.
- 10 of the 24 emails are spam.
 - All possible outcomes equally likely
 - E = user 1 receives 3 spam emails
 - F = user 2 receives 6 spam emails
- What is $P(E \mid F)$?

$$\frac{\binom{4}{3}\binom{14}{3}}{\binom{18}{6}} \approx 0.0784$$

Slicing Up the Spam

- 24 emails are sent 6 each to 4 users.
- 10 of the 24 emails are spam.
 - All possible outcomes equally likely
 - E = user 1 receives 3 spam emails
 - F = user 2 receives 6 spam emails
 - G = user 3 receives 5 spam emails
 - What is $P(G | F)$?

$$\frac{\binom{4}{5} \binom{14}{1}}{\binom{18}{6}} = 0$$

*No way to choose 5 spam from
4 remaining spam emails!*

Sending Bit Strings

- Bit string with m 0's and n 1's sent on network
 - All distinct arrangements of bits equally likely
 - E = first bit received is a 1
 - F = k of first r bits received are 1's

- Solution 1:

$$P(E | F) = \frac{P(EF)}{P(F)} = \frac{P(F | E)P(E)}{P(F)} = \frac{k}{r}$$

$$P(E) = \frac{n}{m+n}$$

$$P(F | E) = \frac{\binom{n-1}{k-1} \binom{m}{r-k}}{\binom{m+n-1}{r-1}}$$

$$P(F) = \frac{\binom{n}{k} \binom{m}{r-k}}{\binom{m+n}{r}}$$

Sending Bit Strings

- Bit string with m 0's and n 1's sent on network
 - All distinct arrangements of bits equally likely
 - E = first bit received is a 1
 - $F = k$ of first r bits received are 1's
- Solution 2:
 - Realize $P(E | F) = P(\text{picking one of } k \text{ 1's out of } r \text{ bits})$
 - $P(E | F) = \frac{k}{r}$
 - Rock on!

Card Piles

- Deck of 52 cards randomly divided into 4 piles
 - 13 cards per pile
 - Compute $P(\text{each pile contains exactly one ace})$
- Solution:
 - $E_1 = \{\text{Ace Spades (AS) in any one pile}\}$
 - $E_2 = \{\text{AS and Ace Hearts (AH) in different piles}\}$
 - $E_3 = \{\text{AS, AH, Ace Diamonds (AD) in different piles}\}$
 - $E_4 = \{\text{All 4 aces in different piles}\}$
 - Compute $P(E_1 E_2 E_3 E_4)$
$$= P(E_1) P(E_2 | E_1) P(E_3 | E_1 E_2) P(E_4 | E_1 E_2 E_3)$$

Card Piles

$E_1 = \{\text{Ace Spades (AS) in any one pile}\}$

$E_2 = \{\text{AS and Ace Hearts (AH) in different piles}\}$

$E_3 = \{\text{AS, AH, Ace Diamonds (AD) in different piles}\}$

$E_4 = \{\text{All 4 aces in different piles}\}$

$$P(E_1) = 1$$

$$P(E_2 \mid E_1) = 39/51 \quad (39 \text{ cards not in AS pile})$$

$$P(E_3 \mid E_1 E_2) = 26/50 \quad (26 \text{ cards not in AS or AH piles})$$

$$P(E_4 \mid E_1 E_2 E_3) = 13/49 \quad (13 \text{ cards not in AS, AH, AD piles})$$

$$P(E_1 E_2 E_3 E_4) = \frac{39 \cdot 26 \cdot 13}{51 \cdot 50 \cdot 49} \approx 0.105$$

Thomas Bayes

- Rev. Thomas Bayes (1702 –1761) was a British mathematician and Presbyterian minister

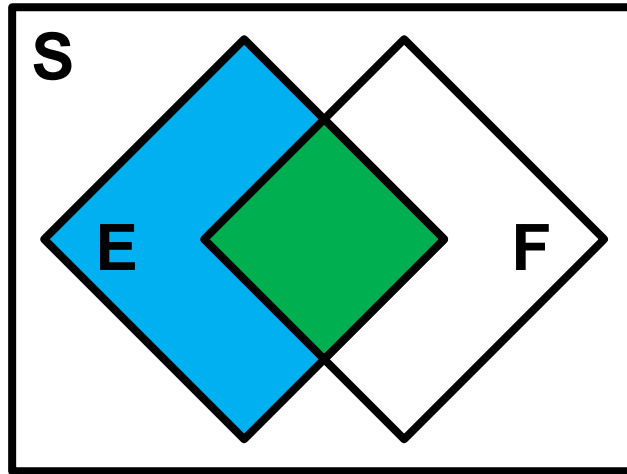


- He looked remarkably similar to Charlie Sheen
 - But that's not important right now...

Background for Bayes' Theorem

- Say E and F are events in S

$$E = EF \cup EF^c$$



Note: $EF \cap EF^c = \emptyset$

$$\text{So, } P(E) = P(EF) + P(EF^c)$$

Bayes' Theorem

- Ross's form:

$$\begin{aligned} P(E) &= P(EF) + P(EF^c) \\ &= P(E | F) P(F) + P(E | F^c) P(F^c) \end{aligned}$$

- Most common form:

$$\begin{aligned} P(F | E) &= P(EF) / P(E) \\ &= [P(E | F) P(F)] / P(E) \end{aligned}$$

- Expanded form:

$$P(F | E) = \frac{P(E | F) P(F)}{P(E | F) P(F) + P(E | F^c) P(F^c)}$$

Bayes' Theorem

- Fully general form:
 - Let F_1, F_2, \dots, F_n be a set of mutually exclusive and exhaustive events
 - “Exhaustive” means one of the events F_1, F_2, \dots, F_n must occur as a result of a particular experiment
 - Event E observed, want to determine which of F_j also occurred, we have:

$$\begin{aligned} P(F_j | E) &= \frac{P(EF_j)}{P(E)} \\ &= \frac{P(E | F_j)P(F_j)}{\sum_{i=1}^n P(E | F_i)P(F_i)} \end{aligned}$$

Bayes is Back!

Improbable Inspiration: The future of software may lie in the obscure theories of an 18th century cleric named Thomas Bayes.

Los Angeles Times (October 28, 1996)

By LESLIE HELM, Times Staff Writer

When Microsoft Senior Vice President Steve Ballmer first heard his company was planning to make a huge investment in an Internet service offering..., he went to Chairman Bill Gates with his concerns.

[...]

Gates began discussing the critical role of "Bayesian" systems.

HIV Testing

- A test is 98% effective at detecting HIV
 - However, test has a “false positive” rate of 1%
 - 0.5% of US population has HIV
 - Let E = you test positive for HIV with this test
 - Let F = you actually have HIV
 - What is $P(F | E)$?

- Solution:

$$P(F | E) = \frac{P(E | F) P(F)}{P(E | F) P(F) + P(E | F^c) P(F^c)}$$

$$P(F | E) = \frac{(0.98)(0.005)}{(0.98)(0.005) + (0.01)(1 - 0.005)} \approx 0.330$$

Why it's Still Good to Get Tested

	HIV +	HIV –
Test +	$0.98 = P(E F)$	$0.01 = P(E F^c)$
Test –	$0.02 = P(E^c F)$	$0.99 = P(E^c F^c)$

- Let E^c = you test negative for HIV with this test
- Let F = you actually have HIV
- What is $P(F | E^c)$?

$$P(F | E^c) = \frac{P(E^c | F) P(F)}{P(E^c | F) P(F) + P(E^c | F^c) P(F^c)}$$

$$P(F | E^c) = \frac{(0.02)(0.005)}{(0.02)(0.005) + (0.99)(1 - 0.005)} \approx 0.0001$$

Simple Spam Detection

- Say 60% of all email is spam
 - 90% of spam has a forged header
 - 20% of non-spam has a forged header
 - Let E = message contains a forged header
 - Let F = message is spam
 - What is $P(F | E)$?

- Solution:

$$P(F | E) = \frac{P(E | F) P(F)}{P(E | F) P(F) + P(E | F^c) P(F^c)}$$

$$P(F | E) = \frac{(0.9)(0.6)}{(0.9)(0.6) + (0.2)(0.4)} \approx 0.871$$

Monty Hall



Let's Make a Deal

- Game show with 3 doors: A, B, and C



- Behind one door is prize (equally likely to be any door)
- Behind other two doors is nothing
- We choose a door
- Then host opens 1 of other 2 doors, revealing nothing
- We are given option to change to other door
- Should we?
 - Note: If we don't switch, $P(\text{win}) = 1/3$ (random)

Let's Make a Deal

- Without loss of generality, say we pick A
 - $P(\text{A is winner}) = 1/3$
 - Host opens either B or C, we always lose by switching
 - $P(\text{win} \mid \text{A is winner, picked A, switched}) = 0$
 - $P(\text{B is winner}) = 1/3$
 - Host must open C (can't open A and can't reveal prize in B)
 - So, by switching, we switch to B and always win
 - $P(\text{win} \mid \text{B is winner, picked A, switched}) = 1$
 - $P(\text{C is winner}) = 1/3$
 - Host must open B (can't open A and can't reveal prize in C)
 - So, by switching, we switch to C and always win
 - $P(\text{win} \mid \text{C is winner, picked A, switched}) = 1$
 - Should always switch!
 - $P(\text{win} \mid \text{picked A, switched}) = (1/3 \cdot 0) + (1/3 \cdot 1) + (1/3 \cdot 1) = 2/3$

A Slight Variant to Clarify Things

- Start with 1,000 envelopes, of which 1 is winner
 - You get to choose 1 envelope
 - Probability of choosing winner = $1/1000$
 - Consider remaining 999 envelopes
 - Probability one of them is the winner = $999/1000$
 - I open 998 of remaining 999 (showing they are empty)
 - Probability the last remaining envelope being winner = $999/1000$
 - Should you switch?
 - Probability winning without switch = $\frac{1}{\text{original \# envelopes}}$
 - Probability winning with switch = $\frac{\text{original \# envelopes} - 1}{\text{original \# envelopes}}$