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, given F already observed)?
  - S = {(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6)}
  - E = {(2, 2)}
  - P(E, given F 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 | 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:

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

Conditional Probability

- General definition:

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

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:

$$P(E_1E_2E_3 \ldots E_n)$$
$$= P(E_1)P(E_2 \mid E_1)P(E_3 \mid E_1E_2) \ldots P(E_n \mid E_1E_2 \ldots E_{n-1})$$

- 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 | 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)?

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

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:

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

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

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

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

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(picking one of *k* 1's out of *r* 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(each pile contains exactly one ace)
- Solution:
  - E_1 = {Ace Spades (AS) in any one pile}
  - E_2 = {AS and Ace Hearts (AH) in different piles}
  - E_3 = {AS, AH, Ace Diamonds (AD) in different piles}
  - E_4 = {All 4 aces in different piles}
  - Compute $P(E_1 E_2 E_3 E_4)$

    $= P(E_1) P(E_2 \mid E_1) P(E_3 \mid E_1 E_2) P(E_4 \mid E_1 E_2 E_3)$

Card Piles

E_1 = {Ace Spades (AS) in any one pile}
E_2 = {AS and Ace Hearts (AH) in different piles}
E_3 = {AS, AH, Ace Diamonds (AD) in different piles}
E_4 = {All 4 aces in different piles}

$P(E_1) = 1$
$P(E_2 \mid E_1) = 39/51$ (39 cards not in AS pile)
$P(E_3 \mid E_1 E_2) = 26/50$ (26 cards not in AS or AH piles)
$P(E_4 \mid E_1 E_2 E_3) = 13/49$ (13 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

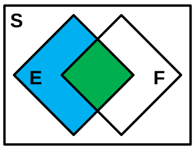
- Rev. Thomas Bayes (1702 –1761) was a British mathematician and Presbyterian minister

- 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 ∪ EF^c



Note: EF ∩ EF^c = ∅

So, P(E) = P(EF) + P(EF^c)

Bayes' Theorem

- Ross's form:

P(E) = P(EF) + P(EF^c)
= P(E | F) P(F) + P(E | F^c) P(F^c)

- Most common form:

P(F | E) = P(EF) / P(E)
= [P(E | F) P(F)] / P(E)

- Expanded form:

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

Bayes' Theorem

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

$$P(F_j \mid E) = \frac{P(EF_j)}{P(E)}$$

$$= \frac{P(E \mid F_j)P(F_j)}{\sum_{i=1}^{n} P(E \mid F_i)P(F_i)}$$

Bayes' is Back, Baby!

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:

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

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

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

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

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

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:

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

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

Odds

- Odds of an event defined as:

$$\frac{P(A)}{P(A^c)} = \frac{P(A)}{1 - P(A)}$$

- Odds of H given observed evidence E:

$$\frac{P(H \mid E)}{P(H^c \mid E)} = \frac{P(H)\,P(E \mid H) / P(E)}{P(H^c)\,P(E \mid H^c) / P(E)} = \frac{P(H)\,P(E \mid H)}{P(H^c)\,P(E \mid H^c)}$$

- After observing E, just update odds by: $\frac{P(E \mid H)}{P(E \mid H^c)}$

Coins and Urns?!

- An urn contains 2 coins: A and B
  - A comes up heads with probability ¼
  - B comes up heads with probability ¾
  - Pick coin (equally likely), flip it, and it comes up heads
  - What are odds that A was picked (note: $A^c = B$)?

$$\frac{P(A \mid \text{heads})}{P(A^c \mid \text{heads})} = \frac{P(A)\ P(\text{heads} \mid A)}{P(A^c)\ P(\text{heads} \mid A^c)} = \frac{\frac{1}{2}\ \frac{1}{4}}{\frac{1}{2}\ \frac{3}{4}} = 1/3$$

  - Odds are 1/3:1 (or probability ¼) that A was picked
  - Note: before observing heads $P(A) / P(A^c) = 1:1$
    - Equally likely to pick A vs. not picking A (1 out of 2 chance)

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(win) = 1/3 (random)

Let's Make a Deal

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