Mathematics for Computer Science

MIT 6.042J/18.062J

Conditional Probability & Independence

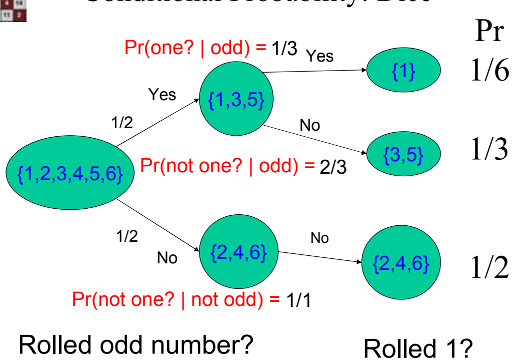
Conditional Probability: Dice

"Knowledge" changes probabilities:

Pr{die rolled 1} = 1/|{1,2,3,4,5,6}|
= 1/6.

Pr{die rolled 1 knowing that die rolled odd number}
= 1/|{1,3,5}|
= 1/3.

Conditional Probability: Dice



Conditional Probability

$\Pr\{A \mid B\} ::=$ probability of event A *given* that event B has occurred. Formally,

$$\Pr\{A \mid B\} ::= \frac{\Pr\{A \cap B\}}{\Pr\{B\}}$$

Product Rule

$$\Pr\{A \cap B\}$$
$$= \Pr\{A \mid B\} \Pr\{B\}$$

Conditional Probability: Monty Hall

Pr{prize at Door 1 | Carol opens 2}
= 1/2.

Really! Outcomes:
(Prize Door, Contestant Door, Carol Door)

$$[\text{Carol opens } 2] = \{\underbrace{(1,1,2)}_{\Pr=\frac{1}{18}}, \underbrace{(1,3,2)}_{\Pr=\frac{1}{9}}, (3,3,2), (3,1,2)\}$$

Conditional Probability: Monty Hall

This suggests the contestant may as well stick, since the probability is 1/2 *given what he knows* at the moment of choosing.

Not so: Contestant knows *more* than door opened by Carol -- also knows: which door he chose himself!

Conditional Probability: Monty Hall

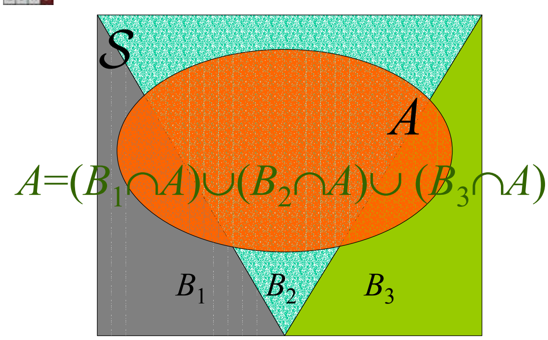
Pr{prize at Door 1|Contestant chose 1 & Carol opens 2}

= 1/3.

[Contestant chose 1 & Carol opens 2]

$$= \{\underbrace{(1,1,2)}_{\Pr=\frac{1}{18}}, \underbrace{(3,1,2)}_{\Pr=\frac{1}{9}}\}$$

Law of Total Probability



Law of Total Probability

$$A=(B_1\cap A)\cup(B_2\cap A)\cup(B_3\cap A)$$

$$\Pr\{A\}= \Pr\{B_1\cap A\}+\Pr\{B_2\cap A\} +\Pr\{B_3\cap A\}$$

$$= \Pr\{A|B_1\}\cdot\Pr\{B_1\} + \Pr\{A|B_2\}\cdot\Pr\{B_2\} + \Pr\{A|B_3\}\cdot\Pr\{B_3\}$$

Law of Total Probability

Let $\mathcal{S}$ be the disjoint union of $B_0, B_1, \ldots$. Then

$$\Pr\{A\} = \sum_{i\in N} \Pr\{A \cap B_i\}$$

$$= \sum_{i\in N} \Pr\{A \mid B_i\} \Pr\{B_i\}.$$

In-class Problem

Problem 1

Introduction to Probability

Independence

Independent Events

A: Baby born at Mass General Hospital between 1:00 am and 1:05 am.

B: Jupiter's moon IO is full.

Independent Events

Does Event *A* (baby is born) have anything to do with Event *B* (IO is full)?

Of course not!

Babies & Full Moons

So the events are *independent:* IO phase has no effect on birth frequency.

Babies & Full Moons

But wait a minute:

My sweet Aunt Daisy believed in Astrology. She thought celestial events could influence babies.

We would say "nonsense," there's no effect.

Babies & Full Moons

Wait another minute! Physics says there IS an effect:

IO full and IO "new" are different distances from Earth.

C:\42\pub\jup-radio_070115.htm

** INFORMATION FOR AMATEUR RADIO ASTRONOMERS ** JUPITER DECAMETRIC EMISSIONS ** JUPITER EPHEMERIS 01 Jul 1994, 0000UTC, Julian Day: 2449534.5, GMT Sidereal Time: 18h35m17s ….

C:\42\pub\jup-radio_070115.htm

SUMMARY: Jupiter's HF emissions are …heard on earth when Jupiter's magnetic field "sweeps" the earth every 9h55m27s and at other times when Io's geometric position influences activity.

Babies & Full Moons

It's True: IO's magnetic field and gravitational pull on the baby are different in different phases!

But their influence on birth times is undetectable.

Babies & Full Moons

If we compared

- All daily birth statistics
- daily birth statistics when IO was full,

we would see no difference!

Babies & Full Moons

Baby frequency betw 1&1:05AM is same when IO is full:

Pr{Baby born 1-1:05AM | IO is full}
= Pr{Baby born 1-1:05AM}

A and *B* are independent!

Definitions of Independence

Definition 1:

Events A and B are independent iff

$$\Pr\{A\} = \Pr\{A \mid B\}.$$

Definition 2:

Events A and B are independent iff

$$\Pr\{A\} \cdot \Pr\{B\} = \Pr\{A \cap B\}.$$

Definitions of Independence

Equivalent:

$\Pr\{A\} = \Pr\{A \mid B\}$ iff

$\Pr\{A\} = \dfrac{\Pr\{A \cap B\}}{\Pr\{B\}}$ iff

$\Pr\{A\} \cdot \Pr\{B\} = \Pr\{A \cap B\}.$

 November 8, 2002

Definitions of Independence

Small bug: need $\Pr\{B\} \neq 0$ for Def. 1.
Def. 2 works even if 0:

$$\Pr\{A\} \cdot \Pr\{B\} = \Pr\{A \cap B\}.$$

 November 8, 2002

Independence

$\Pr\{A\} \cdot \Pr\{B\} = \Pr\{A \cap B\}.$

Symmetric! So,
A independent of B iff
B independent of A.

 November 8, 2002

Independence

Quickies: Reflexive? Transitive?
Intuition for Symmetry?

 November 8, 2002

Independence

A is *independent* of B means it is independent of whether or not B occurs:

A independent of B iff
A independent of $\overline{B}$.

 November 8, 2002

Independence

A independent of B iff
A independent of $\overline{B}$.
Simple proof using:

$\Pr\{A-B\} = \Pr\{A\} - \Pr\{A \cap B\}.$

DO IT NOW!

 November 8, 2002

The Birthday "Paradox"

Puzzle: n students in a room.
What is the probability that two students have the same birthday (month, day)
for n = 2, 10, 23, 30, 107?

 November 8, 2002

The Birthday "Paradox"

So with 10 students have 10/365
≈1/30 chance 2 have same b'day?
Not really, it's more like 1/10.
With 30 students, maybe 3·(30/365)
≈1/3 chance?
No, it's more than 2 to 1!

 November 8, 2002

The Birthday "Paradox"

Let's stop guessing and figure it out.
Choose 2 students at random.
Pr{students have different birthday}?

$$= 1 - \frac{1}{365}$$

 November 8, 2002

The Birthday "Paradox"

We're assuming MIT students are equally likely to have each of 365 possible birthdays.
(Also, probability student has any particular birthday is independent of taking 6.042 this term.)

 November 8, 2002

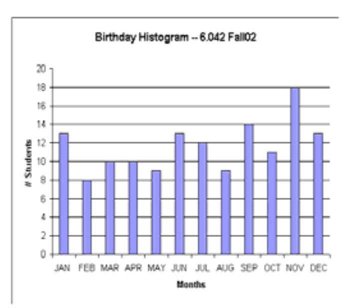
The Birthday "Paradox"

Not really same for each month:
140 students in the 6.042 class reported their birthdays.

 November 8, 2002

Class Birthdays

JAN	13
FEB	8
MAR	10
APR	10
MAY	9
JUN	13
JUL	12
AUG	9
SEP	14
OCT	11
NOV	18
DEC	13
	140



 November 8, 2002

Class Birthdays

November twice as popular as February.

But close enough.

 November 8, 2002

The Birthday "Paradox"

So we'll assume that if we choose 2 students at random:

Pr{students have the same birthday}

$$= \frac{1}{365}$$

 November 8, 2002

The Birthday "Paradox"

Choose another 2 students independently of first two.

Pr{neither pair has same birthday}?

= Pr{1st pair not same birthday and 2nd pair not same birthday}

= Pr{1st pair not same birthday} × Pr{2nd pair not same birthday}

 November 8, 2002

The Birthday "Paradox"

Pr{neither pair has same birthday}

$$= \left(1 - \frac{1}{365}\right)^2$$

 November 8, 2002

The Birthday "Paradox"

Choose another 253 *pairs* of students independently of first pairs.

Pr{no pair has same birthday}?

$$= \left(1 - \frac{1}{365}\right)^{253} \approx \frac{1}{2}$$

 November 8, 2002

The Birthday "Paradox"

But with $n = 23$ students, have

$$\binom{23}{2} = 253 \text{ pairs of students.}$$

 November 8, 2002

The Birthday "Paradox"

So, with 23 students:

$$\Pr\{\text{no pair has same birthday}\} \approx \frac{1}{2}$$

$$\Pr\{\text{some pair has same birthday}\}$$

$$\approx 1 - \frac{1}{2} = \frac{1}{2}$$

 November 8, 2002

The Birthday "Paradox"

With 140 students

$$\Pr\{\text{no pair has same birthday}\}$$

$$= \left(1 - \frac{1}{365}\right)^{\binom{140}{2}} = \left(1 - \frac{1}{365}\right)^{9730}$$

 November 8, 2002

The Birthday "Paradox"

With 140 students

$$\Pr\{\text{no pair has same birthday}\}$$

$$= \left(1 - \frac{1}{365}\right)^{365\left(\frac{9730}{365}\right)} \approx e^{-\left(\frac{9730}{365}\right)}$$

$$\approx \frac{1}{400,000,000,000}$$

 November 8, 2002

The Birthday "Paradox"

In fact have 17 pairs and 2 triples of students in 6.042 with same birthday:

- Jan 1
- Jan 8
- Feb 16
- Feb 23
- Mar 3
- Mar 10
- Apr 16
- Apr 18
- May 17
- Jun 3
- Jun 9
- Jul 25
- Aug 19
- Sep 4
- Sep 22
- Oct 29
- Nov 4
- Nov 14
- Dec 21

 November 8, 2002

Shared Birthdays

TK Focht
Thomas Wilson

Jerry Ing
David Tsai

Jordan Brayanov
Amy Luxenberg

James Tolbert
Alex Vandiver
Gabriel Cunningham

Chris Maes
Casey Dugan

Cecilia Henriquez
Joanna Liang

Sam Kwei
Rocco Pigneri

Robert Kwok
Michael Ogrydziak

Daniel Robey
Isaac Dancy

Aye Moah
Stanley Woods

Miguel Ferreira
Ibrahim Tadros

Alejandro Sedeno
Enrique Zolezzi

Victor Williamson
Samir Meghani

Howard Chou
Gabriel Lopez-Betanzos

Timur Tokmouline
Stephanie Tyll

Edana Gallagher
Brendan Shields
Javier Velez

David Chau
Yi Zhang

Dongsung Huh
Ojonimi Ocholi

Amanda Smith
Brian Williams

 November 8, 2002

The Birthday "Paradox"

Wait! Whether a pair of students in 6.042 has same birthday is not independent of other pairs:

If (Joy, Jen) have same b'day, and (Joy, Mike) do too, then

$\Pr\{(\text{Jen,Mike}) \text{ same b'day}\} = 1.$

 November 8, 2002

The Birthday "Paradox"

Only non-overlapping pairs have independent probabilities of same b'day.

The Birthday "Paradox"

But

as long as #students << # birthdays,

say 23 << 365,

pairs w/same b'day not likely to overlap,

so act like independent.

The Birthday "Paradox"

Accurate formulas for Pr{some pair has same b'day} are in Notes 10.

In-class Problem

Problems 2 & 3

(We didn't get to these in class.)