Lecture 4: Conditional Probability

$$P\left(\bigcup_{j=1}^{n} A_j\right) = 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n+1} \cdot \frac{1}{n!}$$

$$P(\text{no match}) = P\left(\bigcap_{j=1}^{n} A_j^c\right) = \frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3} + \cdots + (-1)^n \cdot \frac{1}{n!}$$

$$\approx \frac{1}{e}$$

Independence Events (completely different concept from disjointness)

Events A, B are independent if

$$P(A \cap B) = P(A) \cdot P(B) \qquad \text{(true both ways)}$$

→ Independence says if we know that A occurs it tells us nothing whatsoever about whether B occurs or not.

→ Disjointness says if A occurs, B can't occur.

✓ A, B, C are independent if

$$P(A \cap B) = P(A) \cdot P(B)$$
$$P(A \cap C) = P(A) \cdot P(C)$$
$$P(B \cap C) = P(B) \cdot P(C)$$
$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot B(C)$$

$P(A \cap B) = P(A, B)$

→ So, Independence means multiply when we are trying to find the probability of intersection.

slogan, not a definition

~~Therefore,~~ Similarly for events $A_1, A_2, A_3, \ldots, A_n$ to be independent —

(X) Any two of $\{A_1, A_2, A_3, \ldots, A_n\}$ have to be independent,
Any three of $\{A_1, A_2, A_3, \ldots, A_n\}$ have to be independent,
⋮
All of them $\{A_1, A_2, \ldots A_n\}$ has to be independent.

Newton-Pepys Problem (1693)

atleast → Union
complement → intersection

Have fair dice; which is most likely?

(A) atleast one 6 with 6 dice. ← truth

(B) atleast 2 sixes with 12 dice.

(C) atleast 3 sixes with 18 dice. ← Pepys strongly believed

Soln: $P(A) = 1 - \underbrace{\left(\frac{5}{6}\right)^6}_{\text{all six dices are not 6}} \approx 0.665$ ✓

$$P(B) = 1 - \underbrace{\left(\frac{5}{6}\right)^{12}}_{\text{all non-sixes}} - \underbrace{\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^{11}(12)}_{\text{exactly 1 six}} \approx 0.619$$

12 → six could be any of the 12 die (12 posibilities)

$$P(C) = 1 - \underbrace{\sum_{k=0}^{2}\binom{18}{k}\left(\frac{1}{6}\right)^k\left(\frac{5}{6}\right)^{18-k}}_{\text{Binomial probability}} \approx 0.597$$

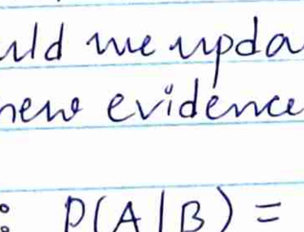
∴ A is most likely, C is less likely.

Conditional Probability (Conditioning is the soul of statistics)

How should we update our uncertainty and update our probabities based on new evidence?
How should we update beliefs/probs./uncertainties based on new evidence?

Definition: $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$, if $P(B) > 0$

Intuition (I) pebble world: 9 pebbles, total mass is 1



$P(A|B)$ = get rid of pebbles that are not in B ~~or B^c~~ (i.e., B^c)
Renormalize to make total mass 1 again.

B is the new world/Universe.

→ Now, part of A that is outside of B is irrelevant.

Renormalize as new world has total mass not equal to 1.

<u>Intuition (II)</u> frequentist world (repeat experiment many times)

⟶ select repeatation where B occurred.

→ Among those, what fraction of time ~~when A occurred~~.
did A also occur?

(among events when B occurred)

<u>Theorem (1)</u> $P(A \cap B) = P(A|B) \cdot P(B)$

$= P(B|A) \cdot P(A)$

<u>Theorem (2)</u> $P(A_1 \cap A_2 \cap \ldots \cap A_n) =$

(n! theorem) $P(A_1) P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdots \cancel{P(A_n|A_1 \cap A_2 \ldots}$

$P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1})$

<u>Theorem (3)</u> $P(A|B) = \dfrac{P(A)\, P(B|A)}{P(B)}$ [Bayes' Rule]