Last lecture

Random Variables (RV) (Ch 2.1)

- Mean and Variance recap (Ch 2.2)

Conditional Probability (Ch 2.3)

- Motivation
- Examples
- Solver

$$\left\{ P(A|B) = \frac{P(AB)}{P(B)} \text{ if } P(B)>0 \right.$$

Agenda

Law of Total Probability (Ch 2.10)

- Bayes formula

Independent Events/ RVs (Ch 2.4)

- Definition
- Motivation
- Examples and Facts

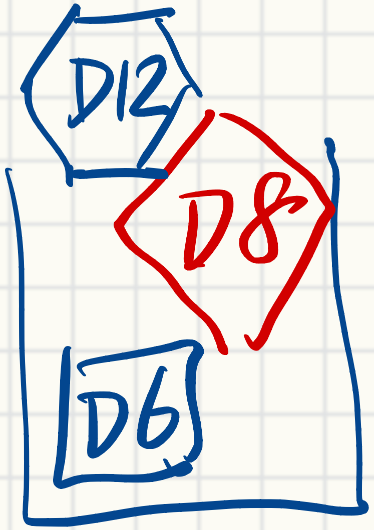
Distributions

- Bernoulli (Ch 2.4.3)
- Binomial (Ch 2.4.4)

Law of total probability

Law of total probability

	E_1	E_2	E_3	E_4
A	$A \cap E_1$	$A \cap E_2$	...	...
A^c				



$A = \{6\}$

- Case-by-case discussion law…
- $P(A)$ is the summed of “Partitioned conditional probability”
- $P(A) = \sum_i P(A|E_i)P(E_i)$

$= \sum_i P(A E_i)$

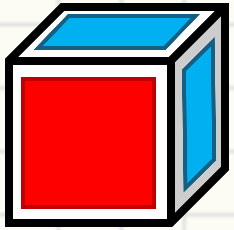
Law of total probability

$P(A) = P(B) = P(C)$

$= \frac{1}{3}$

There are 3 dice A, B, C in the bag

- $A = [R \times 1; B \times 5]$
- $B = [R \times 2; B \times 4]$
- $C = [R \times 3; B \times 3]$



A, B, C are partition & mutually exclusive

Draw one die and roll many times

- $P(R_1) = P(R_1 A) + P(R_1 B) + P(R_1 C)$

  $\hookrightarrow \underbrace{P(R_1|A)}_{\frac{1}{6}} \; \underbrace{P(A)}_{\frac{1}{3}}$

  $= \frac{1}{6} \times \frac{1}{3} + \frac{2}{6} \times \frac{1}{3} + \frac{3}{6} \times \frac{1}{3}$

  $= \frac{1}{3}$

- $P(R_2|R_1)$

$$P(R_2|R_1) = \frac{P(R_2R_1)}{P(R_1)} \qquad \frac{1}{3}$$

$$= P(R_2|R_1,A)P(A) + P(R_2|R_1,B)P(B) + P(R_2|R_1,C)P(C)$$

$$= 3 \times P(R_2R_1)$$

$$= 3 \times [P(R_2R_1|A)P(A) + P(R_2R_1|B)P(B) + P(R_2R_1|C)P(C)]$$

$$= 3\left[\frac{1}{6} \times \frac{1}{6} \times \frac{1}{3} + \frac{2}{6} \times \frac{2}{6} \times \frac{1}{3} + \frac{3}{6} \times \frac{3}{6} \times \frac{1}{3}\right]$$

Bayes Formula

Conditional probability + Law of total probability

- How do we get $P(B|A)$ from $P(A|B)$?

- $P(B|A) = \dfrac{P(AB)}{P(A)} = \dfrac{P(A|B)P(B)}{P(A)}$

- $P(E_i|A) = \dfrac{P(E_iA)}{P(A)} = \dfrac{P(A|E_i)P(E_i)}{\sum_j P(A|E_j)P(E_j)}$

Law of total prob

Disease problems

Assume there is a disease A, and the corresponding test T

$T \triangleq$ test positive

$A \triangleq$ Have disease

- What do the followings mean?
- $P(T|A) = 0.9$
- $P(T|A^c) = 0.05$
- $P(A) = 0.01$

healthy

factory test.

- $P(A|T) =$

$$\frac{P(T|A)P(A)}{P(T)} = \frac{P(T|A)P(A)}{P(T|A)P(A)+P(T|A^c)P(A^c)}$$

$$= \frac{0.9\times 0.01}{0.9\times 0.01 + 0.05\times 0.99} \approx \frac{1}{6}$$

Disease problems

According to CDC survey on smoker

- 18% of adults are smokers $\rightarrow P(S) = 0.18$
- 15% of women are smokers $\rightarrow P(S|W) = 0.15$
- Population = 50% men + 50% women

$P(W^c) = 0.5 \qquad P(W) = 0.5$

- What fraction of adult smokers are women

$$P(W|S) = \frac{P(S|W)P(W)}{P(S)}$$

Disease problems

According to CDC survey on smoker vs. lung cancer

- 15% of women are smokers
- Compared to nonsmokers, women who smoke are 13 times likely to get lung cancer $P(C \mid \cancel{W}, S) = 13 P(C \mid \cancel{W}, S^c)$
- If I pick a female lung cancer patient, how likely she is a smoker?

$$\underbrace{P(C \mid S)}_{13\ell} = 13 P(C \mid S^c)\ \ell$$

$$P(S \mid \cancel{W}, C) = \frac{P(C \mid S) P(S)}{P(C)} = \frac{13\ell \times 0.15}{\underbrace{\ell}_{P(C \mid S^c)} \cdot \underbrace{0.85}_{P(S^c)} + \underbrace{13\ell \cdot 0.15}_{P(C \mid S)}}$$

Independent Events/ RVs

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

Definition

A and B are events, they are mutually independent if

- $P(B|A) = P(B)$ or
- $P(AB) = P(A)P(B)$

Facts

- $P(B|A) = P(B)$ implies $P(A|B) = P(A)$
- If $P(A) = 0$, B is independent of A

Definition

A and B are events, they are mutually independent if

- $P(B|A) = P(B)$ or
- $P(AB) = P(A)P(B)$

RVs X and Y are independent if A and B are independent for any $X \in A$ and $Y \in B$

Motivation

Independent is a common but strong property

- $P(AB) = P(A)P(B)$ factorize the pmf
  - Compute the pmf easily
  - Will skip Ch 2.3 affect only my HW2 and Midterm 1?
  - If I join this club, will it affect my GPA?

- Decide the model complexity
  - What really affects the results?
  - What do I need to ask when reviewing a loan request?
  - What input data do I need to predict the defect?

Examples

Physically independence – Toss a coin and roll a die (N, X)

- $A \triangleq \{N = H\}$
- $B \triangleq \{X = 6\}$

Probabilistic independence

- $A \triangleq X$ is even
- $B \triangleq \{X \equiv 0\ (mod\ 3)\}$

$P(A) = \frac{1}{2}$

$P(A|B) = \frac{P(AB)}{P(B)} = \frac{\|\{6\}\|}{\|\{3,6\}\|} = \frac{1}{2} = P(A)$

$\Rightarrow$ X True

for D8

Slido

#3409457

Choose "independent" RVs/ Events

A. Pick X from 52 playing card,
{X is RED} vs. {X is prime}

$P\{\heartsuit \diamondsuit \mid \{2, 3, 5, 7, 11, 13\}\} = P\{\heartsuit \diamondsuit\}$

B. {X is even} vs. {X is prime}

$\frac{6}{13} \neq P(\text{Even} \mid \text{Prim}) = \frac{1}{6}$

C. Pick Y from 365 days,
{Y is rainy at Champaign} vs. {Y is a holiday}

?

D. {Any midterms on Y} vs. {Y is a holiday}

$P(M) \gg P(M \mid H)$

Terms and Facts

- A, B, C are pairwise independent if $(A, B), (B, C), (A, C)$ are mutually independent
  - Toss a fair coin twice
  - $A \triangleq$ {First coin is Head}
  - $B \triangleq$ {second coin is Head}
  - $C \triangleq$ {toss results are the same}

$P(ABC) = \frac{1}{4} = \frac{\|\{H, H\}\|}{\|\Omega\|}$

$\neq P(A)P(B)P(C)$

$= \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$

$P(C|A) = P(C)$

$P(C|B) = P(C)$ $\Rightarrow$ A, B, C are NOT independent

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

Terms and Facts

- A, B, C are independent they are pairwise independent and $P(ABC) = P(A)P(B)P(C)$

- $A_1, A_2, \ldots A_i$ are independent if
$$P\left(A_{i_1} A_{i_2} \ldots A_{i_k}\right) = P\left(A_{i_1}\right) P\left(A_{i_2}\right) \ldots P\left(A_{i_k}\right)$$