Last lecture

Conditional Probability (Ch 2.3)

- Examples
- Solver
- 3 doors problem revisited

$$P(B|A) = \begin{cases} \dfrac{P(AB)}{P(A)} & \text{if } P(A) > 0 \\ \text{undefined} & P(A) = 0 \end{cases}$$

Law of Total Probability (Ch 2.10)

- Bayes formula

$$P(A) = \sum_i P(AE_i) = \sum_i P(A|E_i)P(E_i)$$

$P(A|E_i)$

$P(AE_i)$

$$P(E_i|A) = \frac{P(A|E_i)P(E_i)}{P(A)} = \frac{P(A|E_i)P(E_i)}{\sum_j P(A|E_j)P(E_j)}$$

Agenda

Bayes formula (Ch 2.10)

- Examples

Independent Events/ RVs (Ch 2.4)

- Definition
- Motivation
- Examples and Facts

- Distributions (Ch 2.4)
  - Bernoulli
  - Binomial

Disease problems

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

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

} Test in lab / test company

⇒ CDC job.

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

Really patient cares.

$P(T) = P(T|A)P(A) + P(T|A^c)P(A^c)$

$$\frac{P(AT)}{P(T)} = \frac{P(T|A)P(A)}{P(T|A)P(A) + P(T|A^c)P(A^c)} = \frac{0.9 \times 0.01}{0.9 \times 0.01 + 0.05 \times 0.99}$$

$$= \frac{0.009}{0.009 + \sim 0.05}$$

$$\approx \frac{0.009}{0.059} \approx \frac{1}{6}$$

Disease problems

According to CDC survey on smoker

- 18% of adults are smokers
- 15% of women are smokers
- Population = 50% men + 50% women

$P(S) = 18\%$

$P(S|W) = 15\%$

$P(W) = 50\%$

- What fraction of adult smokers are women

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

$$= \frac{15\% \times 50\%}{18\%} = \frac{7.5}{18} = \frac{5}{12}$$

Disease problems

100 W.

15 WS × 65%   85% × 5%   65%   5%

$P(S|W) = 15\%$

$P(S|W)$

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

13ℓ ℓ

$P(C|W,S^c) = \ell$

- If I pick a female lung cancer patient, how likely she is a smoker?

$P(C|W,S) = 13\ell$

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

$$\boxed{P(S \mid C, W) = \frac{P(C \mid S, W)\, P(S \mid W)}{P(C \mid W)}}$$

$$\frac{13\ell \times 15\%}{\ell \times 85\% + 13\ell \times 15\%}$$

$P(C \mid W, S^c)$ $P(S^c \mid W)$

Independent Events/ RVs

Definition

A and B are events, they are mutually indenpent if

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

A happens or not does not affect B

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

$P(B)$

$$P(A|B) = \boxed{\frac{P(AB)}{P(B)}} = \frac{P(A)\cancel{P(B)}}{\cancel{P(B)}}$$

Facts

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

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

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

A, B roll a die

$|\Omega_{AB}| = 36.$

$|\Omega_A|$ 6 $|\Omega_B|$ 6

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)

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

$P(B|A) = \frac{1}{6} = P(B)$

Probabilistic independence

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

$A = \{2, 4, 6\}$

$B = \{3, 6\}$

$P(B|A) = \frac{1}{3} = P(B)$

A and B are independent!

Slido

#4285709

Choose “independent” RVs/ Events

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

B. $\{X \text{ is even}\}$ vs. $\{X \text{ is prime}\}$

2 3 5 7 11 13

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

D. $\{\text{Any midterms on } Y\}$ vs. $\{Y \text{ is holiday}\}$