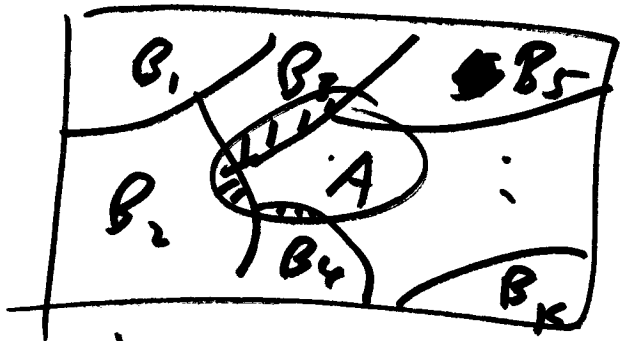
this time: p&c; Bayes's theorem; next time: random variables

read: DS chapters 1 & 2

STAT 131 14 Apr 20 lecture

$\cancel{\sum_{j=1}^{k} P(B_j) = 1}$ $\quad \boxed{P(B_j) > 0}$



$$P(A) = P[(A \text{ and } B_1) \text{ or } (A \text{ and } B_2) \text{ or } \dots \text{ or } (A \text{ and } B_k)]$$

no overlaps

$$= P(A \text{ and } B_1) + P(A \text{ and } B_2) + \dots + P(A \text{ and } B_k)$$

$$= P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \dots + P(B_k)P(A|B_k)$$

$$= \boxed{\sum_{j=1}^{k} P(B_j)P(A|B_j) = P(A)}$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}^n \longleftarrow \mathcal{Y}$$

$$0 \le w_i \le 1, \quad \sum_{i=1}^{n} w_i = 1$$

weighted mean of $\mathcal{Y}$

$$\text{mean } \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{\sum_{i=1}^{n}\left(\frac{1}{n}\right) y_i}{\sum_{i=1}^{n}\frac{1}{n}} = \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i}$$

$\left(w_i = \frac{1}{n}\right)$

$$P(A) = \sum_{j=1}^{k} \underbrace{P(B_j)}_{w_j} \underbrace{P(A|B_j)}_{y_j}$$

mixture (weighted average)

$$\sum_{j=1}^{k} P(B_j) = 1 \qquad 0 < P(B_j) < 1$$

A, B conditionally independent given C iff

$$P(A \text{ and } B|C) = P(A|C)\cdot P(B|C)$$

Bayes's Theorem for events (T/F propositions)

1750 village in rural England

data: unusually large # of people in village died in last month (effect)

Q: why? what's the cause of this data?

(C) cause → effect (E)

(D) ← data

Causal inference

possible causes:

- bad food
- bad water
- bad air
- disease

deterministic causation: hammer always falls down

probabilistic causation: some but not all people die from smoking

$p(\underline{\text{cause}} \mid \underline{\underline{\text{effect}}})$ harder ← we want

$p(\text{effect} \mid \text{cause})$ easier

$$p(C \mid E) \overset{?}{=} p(E \mid C)$$

Bayes's theorem answers this
