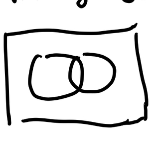
Last time:

Chapter 1

experiment – outcome – event – prob.

3 Axioms of prob.

Assumption of equally likely outcomes

Venn diagram 

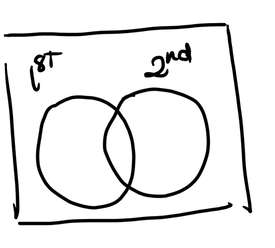
Chapter 2

Inclusion-exclusion : $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A \text{ or } B) = P(A) + P(B) - P(AB)$

Example 1: roll a fair, 6-sided die twice, find

$P(\text{at least one } 6) = \frac{6}{36} + \frac{6}{36} - \frac{1}{36} = \frac{11}{36}$.



Alt. sol. by complement:

$P(\text{at least one } 6) = 1 - P(\text{no } 6)$

$= 1 - P(1^{st} \text{ is not } 6 \text{ and } 2^{nd} \text{ is not } 6)$

$= 1 - \frac{5 \times 5}{6 \times 6} = \frac{11}{36}$

$= 1 - P(1^{st} \text{ is not } 6)\, P(2^{nd} \text{ is not } 6 \mid 1^{st} \text{ is not } 6)$

$= 1 - \frac{5}{6} \times \frac{5}{6} = \frac{11}{36}$.

The multiplication rule. (for computation)

For any events A, B,

$P(AB) = P(A) \cdot P(B \mid A)$

Example 2. 3 cards R G B

draw a card, then another without replacement (w/o repl.)

outcome space $\Omega = \{RG, RB, GR, GB, BR, BG\}$.

$P(R_1 G_2) = \frac{1}{6}$ R_1 : { first draw R }

G_2 : { second draw G }.

2-step:

$P(R_1 G_2) = P(R_1)\, P(G_2 \mid R_1)$ The prob. of drawing G based on the first is R

$= \frac{2}{6} \cdot \frac{1}{2} = \frac{1}{6}$

Example 3. A standard deck of 52 cards

4 suits : H S C D ♡ ♤ ♧ ◇ 13 cards

Deal 2 card w/o repl.

$P(H_1 H_2) = P(H_1)\, P(H_2 \mid H_1)$

$= \frac{13}{52} \times \frac{12}{51}$ → one less H left → one less card left.

Deal 5 cards w/o repl.

$P(H_1 H_2 H_3 H_4 H_5) = P(H_1)\, P(H_2 \mid H_1)\, P(H_3 \mid H_1 H_2) \times P(H_4 \mid H_1 H_2 H_3)\, P(H_5 \mid H_1 H_2 H_3 H_4)$

$= \frac{13}{52} \times \frac{12}{51} \times \frac{11}{50} \times \frac{10}{49} \times \frac{9}{48}$.

(Generalized) multiplication rule

For any events $A_1, A_2, \cdots, A_n$,

$P(A_1 A_2 \cdots A_n) = P(A_1) P(A_2 \mid A_1) P(A_3 \mid A_1 A_2) \cdots P(A_n \mid A_1 A_2 \cdots A_{n-1})$

$P(5 \text{ of the same suit})$

$= P(5H) + P(5S) + P(5C) + P(5D)$.

$= 4 P(5H)$.

Difference between rolling dice ("with repl") & dealing cards (For this course dealing cards implies w/o. repl. unless specified)

Simple Random Sampling (SRS) → Sampling w/o. repl.

Example 4. A deck of cards 26 R 26 B.

- $P(R_1) = \frac{26}{52} = \frac{1}{2}$
- $P(R_2) = \frac{1}{2}$? $P(R_2 \mid R_1)$.

"Law of total prob." = addition + multi.

$P(R_2) = P(R_1 R_2) + P(B_1 R_2)$

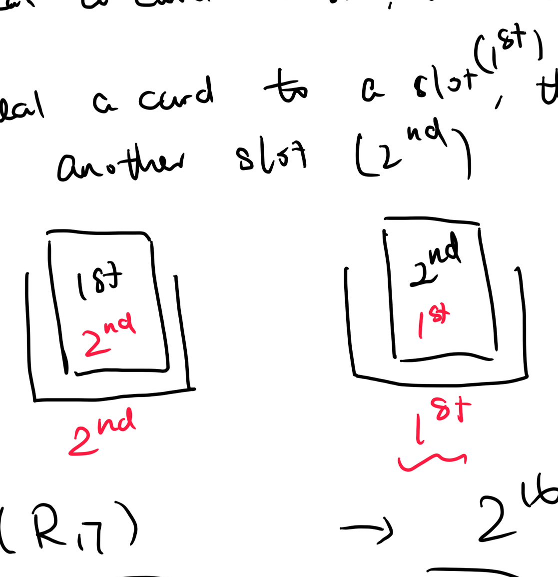
$= P(R_1) P(R_2 \mid R_1) + P(B_1) P(R_2 \mid B_1)$

$= \frac{26}{52} \times \frac{25}{51} + \frac{26}{52} \times \frac{26}{51}$

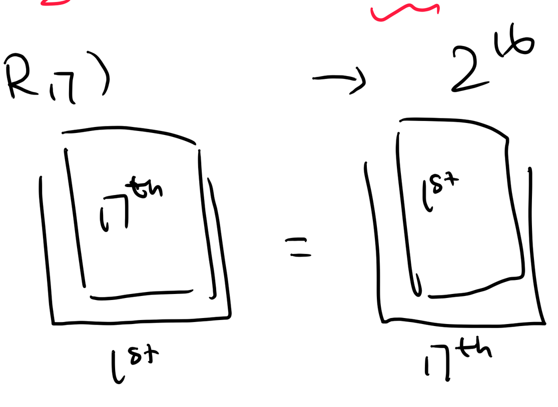
$= \frac{26}{52}$

deal a card first, then deal a second one

↳ deal a card to a slot (1st), then deal a second to another slot (2nd)



- $P(R_{17})$ → 2^{16} possibilities! terms



- $P(R_3 R_7 R_{10} R_{21} R_{40})$

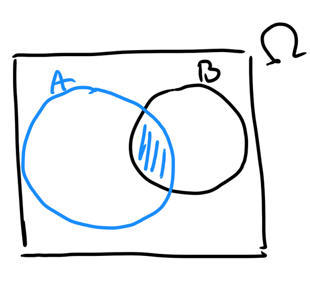
$= P(R_1 R_2 R_3 R_4 R_5)$ ←

by "Symmetry"

Conditional probability.

multi. rule : $P(AB) = P(A)\, P(B \mid A)$

if $P(A) \neq 0$, $P(B \mid A) = \frac{P(AB)}{P(A)}$ (definition!)



Naturally, A happens before B does.

What about $P(A \mid B)$ (backward conditioning).

e.g.1 judge – evidence – guilty or not

you can only infer from evidence

e.g.2 vaccine – effective

"inference"

Example 1. jar 3R 2G

box 1R 4G

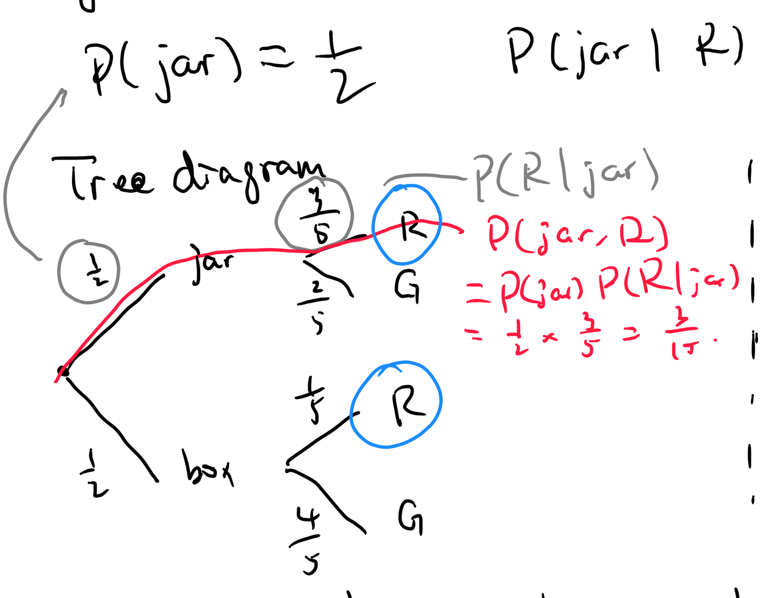
pick a container at random (equally likely)

then pick a ball from it - - - - - - - - - -

you are only shown with the color of the ball (R).

$P(\text{jar}) = \frac{1}{2}$ $P(\text{jar} \mid R) = \frac{P(\text{jar}, R)}{P(R)}$

Tree diagram



$= \frac{3/10}{P(\text{jar}, R) + P(\text{box}, R)}$

$= \frac{3/10}{3/10 + 1/10}$

$= \frac{3}{4}$.

new information has impact on probability.

$P(\text{jar})$ – prior prob.

$P(\text{jar} \mid R)$ – posterior prob. (backward conditioning)

Bayes' rule

$P(R \mid \text{jar})$ – likelihood (forward ..)

{ defn of conditional prob.
addition rule
multi. rule

Example 2: Harvard medical school survey.

disease with prevalence $\frac{1}{1000} = 0.1\%$.
base rate in the population

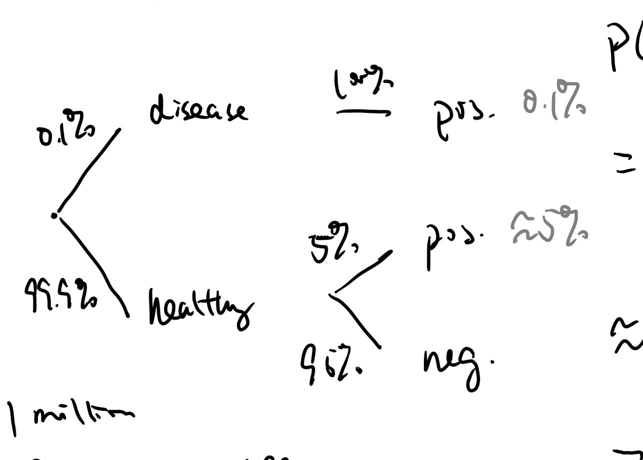
a good test { tells 100% of those who have the disease
has only 5% of reporting false positive

false → actually healthy.
positive → indicates that the person has the disease

It turns out that he results positive in the test, what is the actual chance that he has the disease?

$P(\text{disease} \mid \text{pos.})$

① 95% ② 2%



$P(\text{disease} \mid \text{pos.})$

$= \frac{P(\text{disease}, \text{pos.})}{P(\text{disease}, \text{pos.}) + P(\text{healthy}, \text{pos.})}$

$\approx \frac{0.1\%}{0.1\% + 5\%}$

$= \frac{1}{51} \approx 2\%$

pop. = 1 million

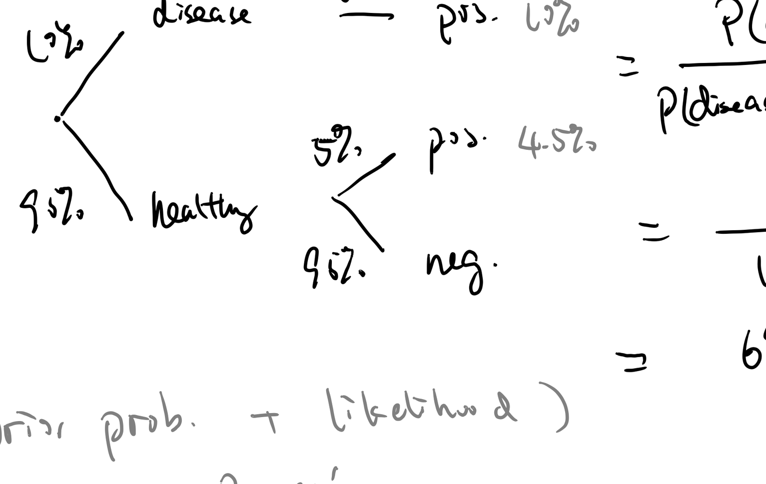
	pos.	neg.
disease	1k	0
healthy	≈50k	≈949k

→ 5% false positive

2% posterior

risk is 20-folded (0.1% → 2%)

Change of base rate. (0.1% → 10%)



$P(\text{disease} \mid \text{pos.})$

$= \frac{P(\text{disease}, \text{pos.})}{P(\text{disease}, \text{pos.}) + P(\text{healthy}, \text{pos.})}$

$= \frac{10\%}{10\% + 4.5\%}$

$= 69\%$.

(prior prob. + likelihood) $\xrightarrow[\text{rule}]{\text{Bayes'}}$ posterior prob.