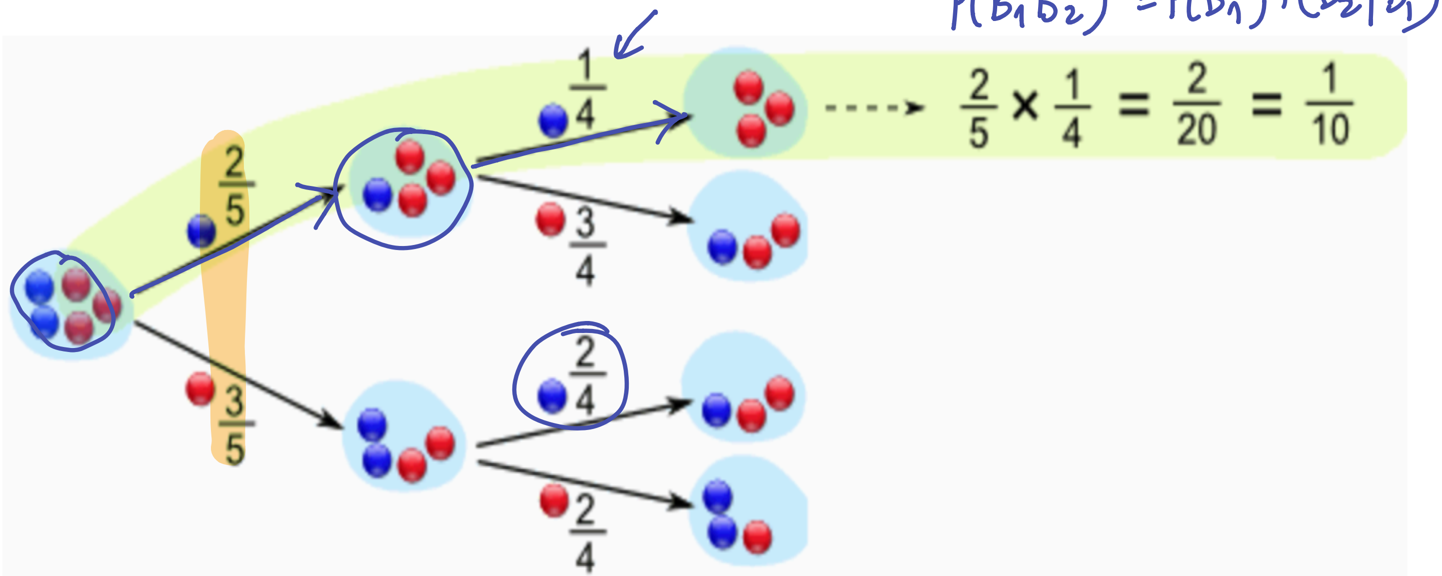
Stat 88: Probability and Mathematical Statistics in Data Science

$P(B_1 B_2) = P(B_1)P(B_2 | B_1)$



Lecture 4: 1/27/2021

Symmetry in Sampling, Bayes' Rule

Sections 2.2, 2.3

Shobhana M. Stoyanov

Agenda

- Kahoot!
- Review the multiplication rule
- Addition rule
- Inclusion Exclusion
- Symmetries in simple random sampling
- Bayes' rule

Multiplication rule

$$P(AB) = P(A|B) \times P(B)$$, $P(B) \neq 0$

- Ex.: Draw a card at random, from a standard deck of 52
  - P(King of hearts) =? $1/52$
- Draw **2** cards one by one, **without** replacement.
  - P(1st card is K of hearts)= $1/52$
  - P(2nd card is Q of hearts| 1st is K of hearts) = $1/51$
  - P(1st card is K of hearts AND 2nd is Q of hearts) = $\underbrace{\frac{1}{52}}_{P(1^{st} K♡)} \cdot \underbrace{\frac{1}{51}}_{P(2^{nd} Q♡ | 1^{st} K♡)}$ ~~P(2nd Q♡)~~

=

- We can also write the "Division Rule" for conditional probability:

$$\boxed{P(A|B) = \frac{P(AB)}{P(B)}, P(B) \neq 0}$$

Cond'l prob

Mistake $\qquad A \cap B = 0$ ϕ

$AB \leftrightarrow A \cap B$, A and B

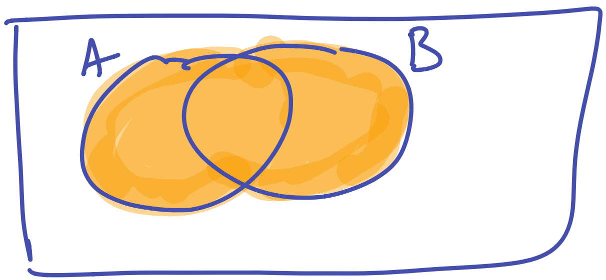
$A \cup B$ for at least one of A or B is true.

Addition rule:

- Addition rule: If A and B are ***mutually exclusive*** events, then the probability that **at least one** of the events will occur is the sum of their probabilities:

$$P(A \cup B) = P(A) + P(B)$$

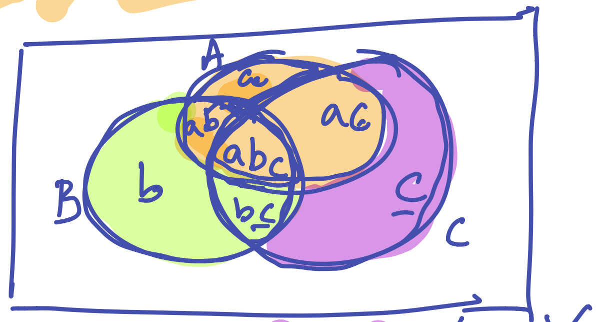
- If they are not mutually exclusive, does this still hold?
- How do we write the event that "at least one of the events A or B will occur? How do we draw it?



Inclusion-Exclusion Formula (general addition rule)

- $P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(AB)$

- $P(A \cup B \cup C) = P(A) + P(B) + P(C)$
  $- P(AB) - P(AC) - P(BC)$
  $+ P(ABC)$

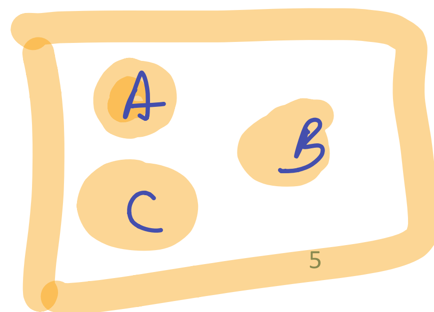


- (Draw a Venn diagram)

$a + ab + ac + abc + b + ab + bc + abc$
$+ c + ac + bc + abc - (ab + abc)$
$- (abc + ac) - (bc + abc) + abc$

- Of course, if A and B (or A and B and C) *don't* intersect, then the general addition rule becomes the **simple** addition rule of

$$P(A \cup B) = P(A) + P(B), \text{ or}$$
$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$



H_1, H_2, H_3, H_4, H_5

1st ♡ , 2nd ♡.

Examples

- Deal 5 cards from the top of a well shuffled deck. What is the probability that all are hearts? (Extend the multiplication rule)

$$\star \rightarrow P(H_1 H_2 H_3 H_4 H_5) = \underbrace{\frac{13}{52}}_{P(H_1)} \cdot \underbrace{\frac{12}{51}}_{P(H_2 \mid H_1)} \cdot \underbrace{\frac{11}{50}}_{P(H_3 \mid H_1 H_2)} \cdot \underbrace{\frac{10}{49}}_{P(H_4 \mid H_1 H_2 H_3)} \cdot \underbrace{\frac{9}{48}}_{P(H_5 \mid H_1 H_2 H_3 H_4)}$$

- Deal 5 cards, what is the chance that they are all the same suit? (flush)

A

$P(H_1) = \frac{13}{52}$ ♡ $P(H_2 \mid H_1)$ ♡

$P(A) = 4(\star)$

1st could be any suit $\frac{52}{52} \cdot \frac{12}{51} \cdot$

$P(\text{all same suit}) = \underbrace{P(\text{all }\heartsuit)}_{=\star} + \underbrace{P(\text{all }\diamondsuit)}_{\star} + \underbrace{P(\spadesuit)}_{\star} + \underbrace{P(\clubsuit)}_{\star}$

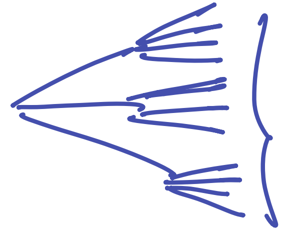
$= 4(\star)$

$\boxed{\text{Partition}}$ A into mutually exclusive events & compute prob. of each, use addition rule

Sec. 2.2: Symmetry in Simple Random Sampling

- One of the topics we will revisit many times is ***simple random sampling***.
- Sampling **without** replacement, each time with equally likely probabilities
- Example to keep in mind: dealing cards from a deck, or drawing tickets from a box, without replacement
- Sampling with replacement: We keep putting the sampled outcomes back before sampling again. (Can do this to simulate die rolls.)
- Need to count number of possible outcomes from repeating an action such as sampling, will use the product rule of counting.

Product rule of counting



12 outcomes

- If a set of actions (call them $A_1, A_2, \ldots, A_n$) can result, respectively, in $k_1, k_2, \ldots, k_n$ possible outcomes, then the entire set of actions can result in:

$$k_1 \times k_2 \times k_3 \times \cdots \times k_n \text{ possible outcomes}$$

- For example: toss a coin twice. Each toss can have 2 possible outcomes, therefore 2 tosses can have 4 possible outcomes.
- So we can count the outcomes for *each* action and multiply these counts to get the number of possible sequences of outcomes.

$2 \cdot 2 \cdot 2 \cdot \ldots \cdot 2 = 2^n$ possible sequences of H,T when tossing coin n times

How many ways to arrange…

- Consider the box that contains O R A N G E:
- How many ways can we rearrange these letters?

$\underline{6} \cdot \underline{5} \cdot \underline{4} \cdot \underline{3} \cdot \underline{2} \cdot \underline{1} = 6! \leftarrow$ "6 factorial"

# of arrangements/permutations of 6 items

- Now say we only want to choose **2 letters** out of the six: $\underline{6} \cdot \underline{5}$

$= 30$ arrangements when ORDER MATTERS.

Symmetries in cards

- Deal 2 cards from top of the deck.
  - How many possible sequences of 2 cards?
  - What is the chance that the second card is red?

- P(5th card is red)

- $P(R_{21} \cap R_{35})$ = (write it using conditional prob)

- P(7th card is a queen)

- $P(B_{52} \mid R_{21}R_{35})$

Section 2.3: Bayes' Rule:

- I have two containers: a jar and a box. Each container has five balls: The jar has three red balls and two green balls, and the box has one red and four green balls.
- Say I pick one of the containers at random, and then pick a ball at random. What is the chance that I picked the box, if I ended with a red ball?

Prior prob of box $= \frac{1}{2}$

New info: ball chosen is red.

If we know this, then we think $P(\text{box}) < \frac{1}{2}$ b/c jar has more red balls.

We will compute a new prob: POSTERIOR Prob.

"a priori" $P(box) = \frac{1}{2}$

Prior and Posterior probabilities

- The ***prior*** probability of drawing the box = $\frac{1}{2}$ (before we knew anything about the balls drawn)

- The ***posterior*** probability of drawing the box = ____ (this is after we *updated* our probability, *given* the information about which ball was drawn)

$$P(Box \mid R) = \frac{P(Box \,\&\, red)}{P(Red)}$$

$$P(Box \mid ball\ red)$$