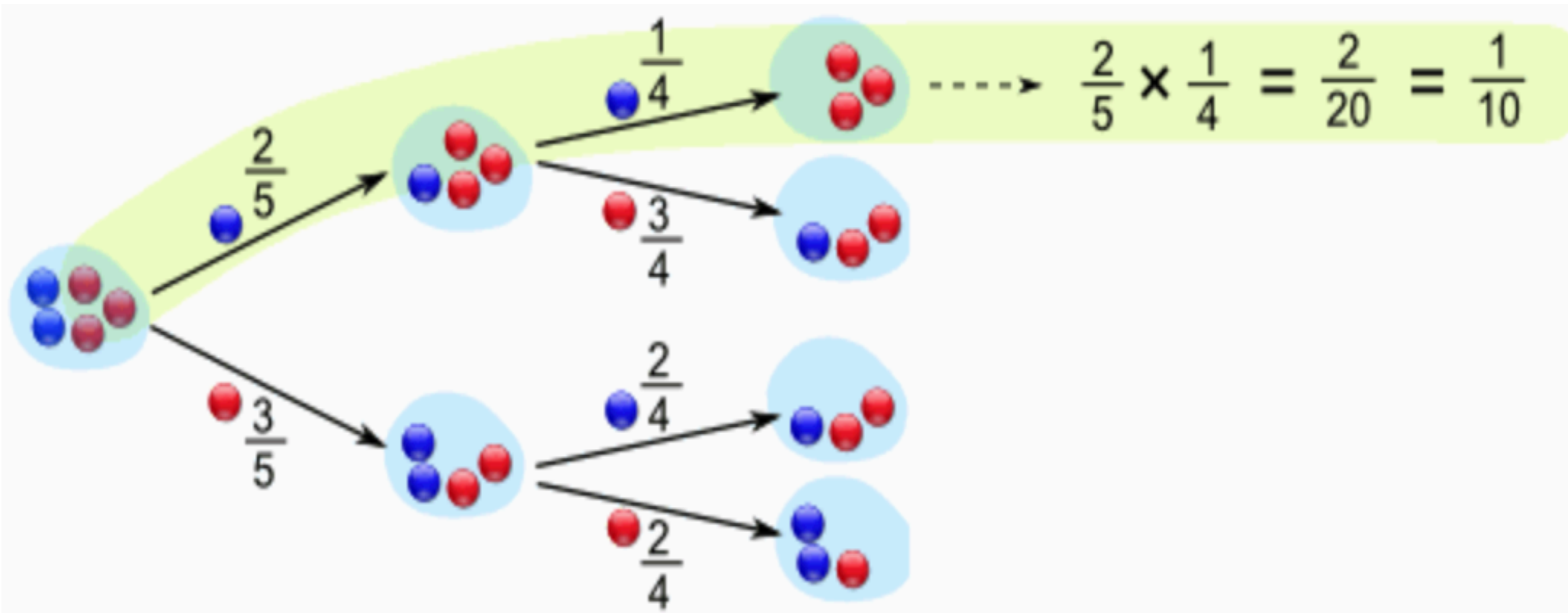
Stat 88: Probability and Mathematical Statistics in Data Science



Lecture 4: 1/24/2024

Bounds, Axioms, Intersections

Sections 1.2, 1.3, 2.1

Shobhana M. Stoyanov

Warm up (hint: draw Venn diagrams)

If we have events A and B such that $P(A) = 0.7$ and $P(B) = 0.5$,

1) Can A and B be mutually exclusive?

2) What can you say about $P(A \cup B)$?

3) What can you say about $P(A \cap B)$?

Agenda

- Bounds on intersections and unions of events
- Axioms of probability
- De Morgan’s laws (exercise)
- The multiplication rule
- Generalized Addition rule
- Inclusion Exclusion

Back to warm up problem, now with some context.

Let A be the event that you catch the bus to class instead of walking, P(A) = 70% and let B be the event that it rains, P(B) = 50%

What is the chance of **at least** one of these two events happening?

What is the chance of **both** of them happening?

Exercise: De Morgan's Laws

Exercise: Try to show these using Venn diagrams and shading:

1. $(A \cap B)^c = A^c \cup B^c$

2. $(A \cup B)^c = A^c \cap B^c$

§1.3: Fundamental Rules

Also called "Axioms of probability", first laid out by Kolmogorov

Recall Ω, the outcome space. Note that Ω can be finite or infinite.

<u>First, some notation:</u>

Events are denoted (usually) by $A, B, C\ldots$

Recall that Ω is itself an event (called the ***certain*** event) and so is the empty set (denoted $\varnothing$, and called the ***impossible*** event or the *empty set*)

The *complement* of an event A is everything ***else*** in the outcome space (all the outcomes that are *not* in A). It is called "not A", or the complement of A, and denoted by A^c

Rethinking the definition of $P(A)$

- So far, we have thought about the probability of an event A as the proportion of the outcomes in A. That is, if the outcome space Ω has n equally likely outcomes, each outcome will have probability $\frac{1}{n}$; and if the event A has k outcomes, then $P(A) = \frac{k}{n}$.
- Now we can rethink our definition to make it more general. We keep the idea of probability of an event A describing the relative *size* of A, and we will generalize the properties of proportions that we have seen so far, and used.
- Let's think of probability as a numerical ***function*** on ***events***, so the input into this function is an event A, and the output is $P(A)$, a number between 0 and 1 satisfying some natural axioms (rules).

The Axioms of Probability

$P(A)$ is a number between 0 and 1 satisfying the axioms below.

Formally, let $A \subset \Omega$, then for every such A, we have a number $P(A)$ such that:

1. For every event $A \subseteq \Omega$, we have $P(A) \geq 0$
2. $P(\Omega) = 1$ (the outcome space is *certain*)
3. Addition rule: If two events are mutually exclusive, then the probability of their union is the sum of their probabilities:

$$A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$$

The third axiom is actually more general and says: If we have infinitely many events that are *mutually exclusive* (*no* pair of them has an overlap- that is: $A_i \cap A_j = \emptyset$ for every pair A_i, A_j; $i \neq j$), then the probability of their union is the sum of their probabilities.

The Axioms of Probability

Let's restate them - they don't look like much, but the entire course is essentially studying the axioms and their consequences.

1. For every event $A \subseteq \Omega$, we have $P(A) \geq 0$
2. $P(\Omega) = 1$
3. If events $A_1, A_2, A_3 \ldots$ are mutually exclusive, then:

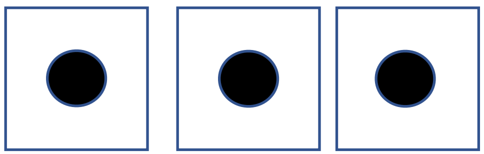
$$P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$$

- Now we can derive the complement rule from (2) and (3):

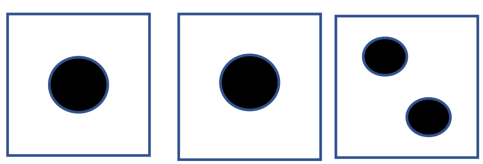
Example of complements

Roll a die 3 times, let A be the event that we roll an ace **each** time.

A^C = **not** A, or not ***all*** aces. It is **not equal** to “never an ace”.

A = 

What about “not A”? Here is an example of an outcome in that set.



Consequences of the axioms

1. Complement rule: $P\left(A^c\right) = 1 - P(A)$

What is the probability of **not** rolling a pair of sixes in a roll of a pair of dice?

2. Difference rule: If $B \subseteq A$, then $P(A \setminus B) = P(A) - P(B)$ where A \B refers to the *set difference between* A *and* B, *that* is, all the outcomes that are A but not in B.

Consequences of the axioms

3. **Boole's (and Bonferroni's) inequality:** generalization of the fact that the probability of the union of A and B is ***at most*** the sum of the probabilities.

We know that $P(A \cup B) \leq P(A) + P(B)$. We can extend this to unions of n events:

For all events $A_1, A_2, A_3, \ldots, A_n$, we have:

$$P(\bigcup_{i=0}^{n} A_i) \leq \sum_{i=0}^{n} P(A_i)$$

How do we solve problems like these:

- What is the probability that the top card in a standard 52 card deck is a queen ***and*** the bottom card is a queen?

- What is the probability that the top card in a standard 52 card deck is a queen ***or*** the bottom card is a queen?

Probability of an intersection

- Say we have three colored balls in an urn (red, blue, green), and we draw two balls ***without*** replacement.
- Find the probability that the first ball is red, and the second is blue (*Write down the outcome space and compute the probability*)

- We can also write it down in sequence: P(first red, then blue) = P(first drawing a red ball) $\times$ P(second ball is blue, given 1st was red)

Multiplication rule

- Conditional probability written as $P(B|A)$, read as "the probability of the event B, given that the event A has occurred"
- The probability that two things will both happen is the chance that the first happens, *multiplied* by the chance that the second will happen *given* that the first has happened.

- Let $A,\ B \subseteq \Omega,\ P(A) > 0, P(B) > 0$

- Multiplication rule:

$$P(A \cap B) = P(A \mid B) \times P(B)$$

$$P(A \cap B) = P(B \cap A) = P(B) \times P(A \mid B)$$

Multiplication rule

$$P\left(A \cap B\right) = P\left(A \middle| B\right) \times P(B)$$

- <u>Ex.</u>: Draw a card at random, from a standard deck of 52
    - P(King of hearts) =?

- Draw **2** cards one by one, **without** replacement.
    - P(1st card is K of hearts)=

    - P(2nd card is Q of hearts| 1st is K of hearts) =

    - P(1st card is K of hearts AND 2nd is Q of hearts) =

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

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

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)

- 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

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

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

Exercise:

- What is the probability that the top card in a standard 52 card deck is a queen ***and*** the bottom card is a queen?

- What is the probability that the top card in a standard 52 card deck is a queen ***or*** the bottom card is a queen?

Examples

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

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

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

- 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

- 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:

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

- 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.

How many ways to arrange…

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

- Now say we only want to choose **2 letters** out of the six: __ __

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(5^{th} card is red)

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

- P(7^{th} card is a queen)

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