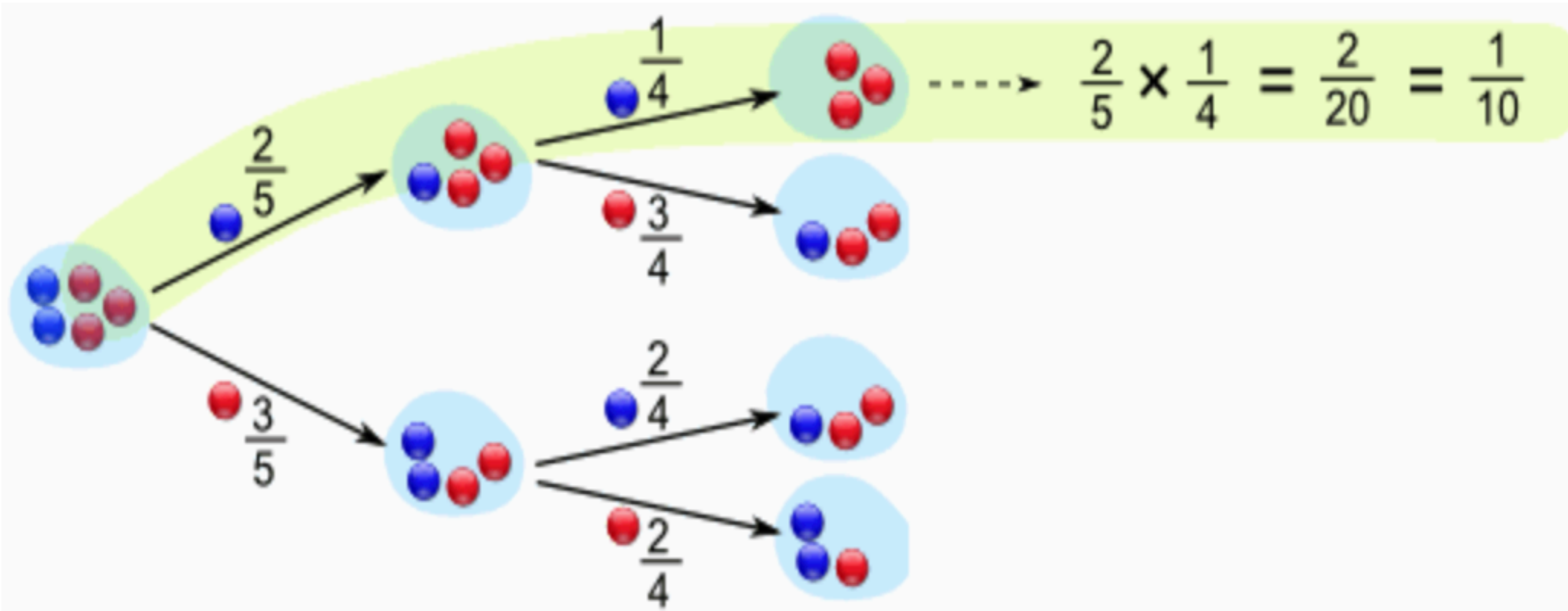
Stat 88: Probability and Mathematical Statistics in Data Science



Lecture 4: 1/24/2024

Bounds, Axioms, Intersections

Sections 1.2, 1.3, 2.1

Shobhana M. Stoyanov

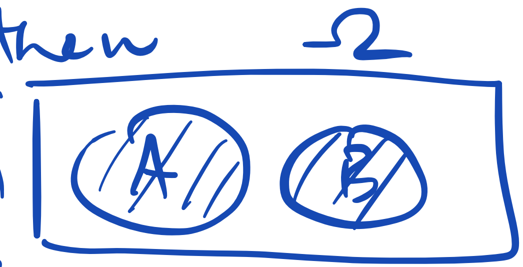
Warm up (hint: draw Venn diagrams)

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

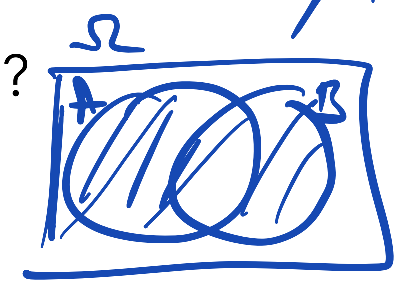
1) Can A and B be mutually exclusive?

If A & B were mutually exclusive then

$P(A \cup B) = P(A) + P(B) = 0.7 + 0.5 = 1.2$ X

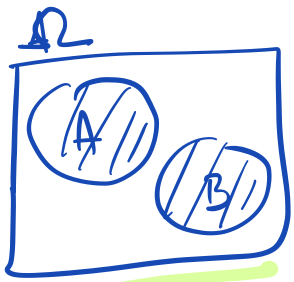


2) What can you say about $P(A \cup B)$?



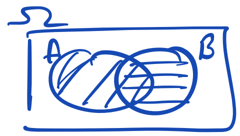
$\leq P(A \cup B) \leq 1$

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

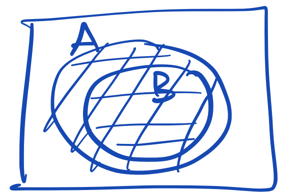


Extreme case 1

(either mutually exclusive or minimally overlap)



$P(A \cup B)$



Extreme case 2

maximal overlap.
when one of events is completely contained in other

In example above $P(A) + P(B) = 0.7 + 0.5 = 1.2$
so 0.2 is the excess above 1, so A & B must have at least that much overlap.

Minimal overlap gives the biggest union.

Ω = entire English alphabet : $\#(\Omega) = 26$

A = vowels $= \{a, e, i, o, u\}$ $\#(A) = 5$

$P(A) = \frac{5}{26}$ $\quad$ $D = \{i, o, u\}$ $\quad$ $D \subset A$ $\quad$ $P(A \cup D) = P(A) = 5/26$

$B = \{c, r, a, z, y\}$ $\quad$ $P(B) = \frac{5}{26} = \frac{\#(B)}{\#(\Omega)}$

Are A & B mutually exclusive

$A \cup B = \{a, e, i, o, u, c, r, z, y\}$

$P(A \cup B) = \frac{9}{26}$

$C = \{d, f, g, h, r\}$ $\quad$ $A \cap C = \phi$

$P(A \cup C) = P(A) + P(C) = \frac{10}{26}$ $\quad$ $P(\phi) = \frac{0}{26} = 0$

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.

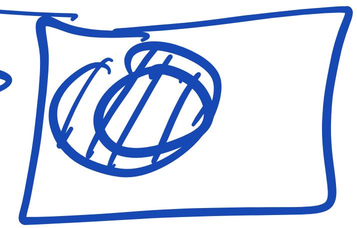
Let A be the event that you catch the bus to class instead of walking, P(A) = 70% and let B be the event that it rains, P(B) = 50%

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

means union

$0.7 \le P(A \cup B) \le 1$

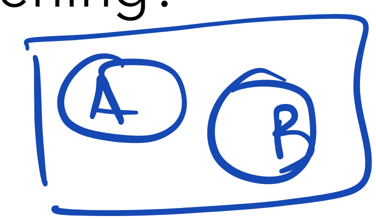
0.7

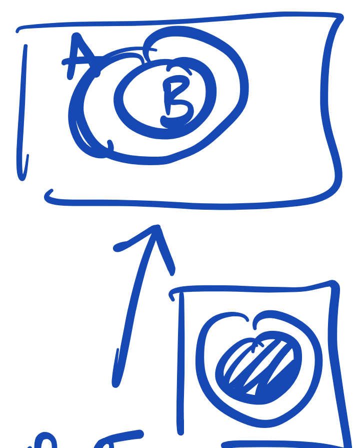


What is the chance of **both** of them happening?

intersection

$P(A \cap B)$





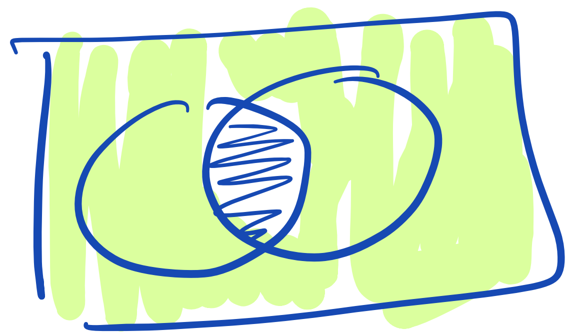
$0.2 \le P(A \cap B) \le 0.5$

0.5

Exercise: De Morgan's Laws

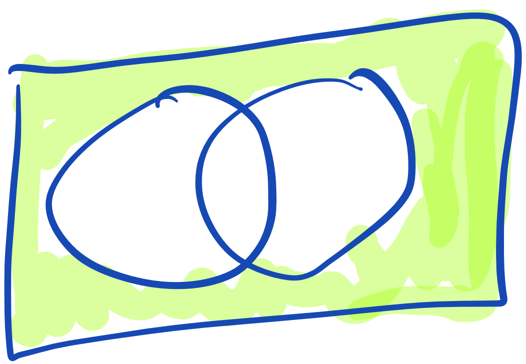
Exercise: Try to show these using Venn diagrams and shading:

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



$(A \cap B)^c$

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



$(A \cup B)^c$

§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.

First, some notation:

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:

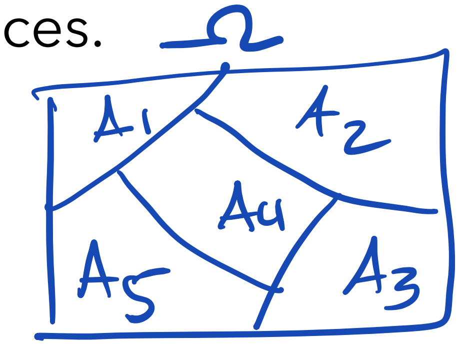
1. For every event $A \subseteq \Omega$, we have $P(A) \geq 0$
2. $P(\Omega) = 1$ (the outcome space is *certain*)
3. Addition rule: If two events are mutually exclusive, then the probability of their union is the sum of their probabilities:

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

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: $A_i \cap A_j = \varnothing$ 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.



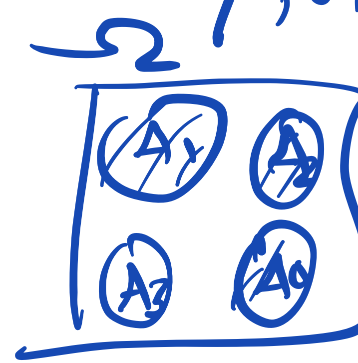
1. For every event $A \subseteq \Omega$, we have $P(A) \geq 0$
2. $P(\Omega) = 1$
3. If events $A_1, A_2, A_3 \ldots$ are mutually exclusive, then:

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

$A_1, A_2, \ldots A_n$, such that $A_i \cap A_j = \phi$, $i \neq j$

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$$

$$\sum_{i=1}^{n} P(A_i) = P(A_1) + P(A_2) + P(A_3) + \ldots + P(A_n)$$



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

Consider A, A^c

Know that $A \cup A^c = \Omega$

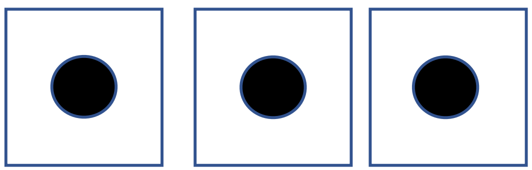
$$P(A \cup A^c) = 1 \Rightarrow P(A) + P(A^c) = 1$$

$$P(A^c) = 1 - P(A)$$

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".

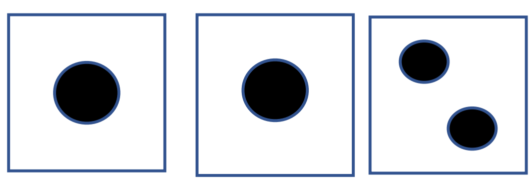
A = 

Exercise

If we roll a die twice

P(no $\boxed{\cdot}$ on either roll) = ?

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



$P(A^c) \neq P(\text{never } \boxed{\cdot})$

Consequences of the axioms

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

Exercise

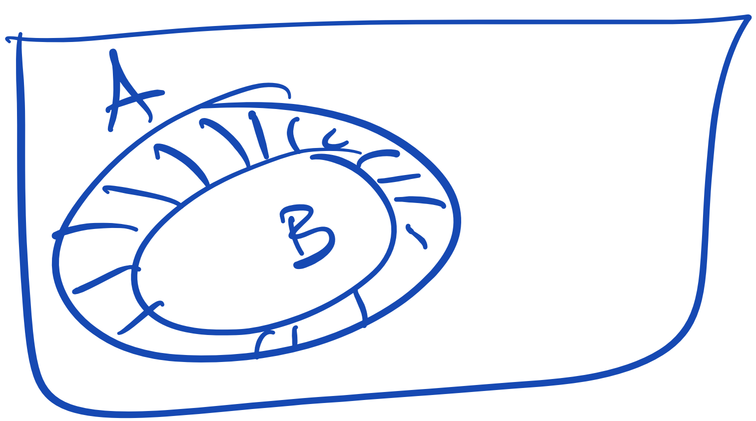
What is the probability of **not** rolling a pair of sixes in a roll of a pair of dice?

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.

A setminus B

$A \setminus B$

$= A \cap B^c$



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

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

we don't know if they overlap

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

Can't say =

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.

Division Rule.

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

$P(A) \neq 0$

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

$$\boxed{P(A \cap B) = P(A|B) \times P(B)}$$

2nd Exercise

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

Note : Is $P(A|B) = P(B|A)$, $P(A) > 0$, $P(B) > 0$