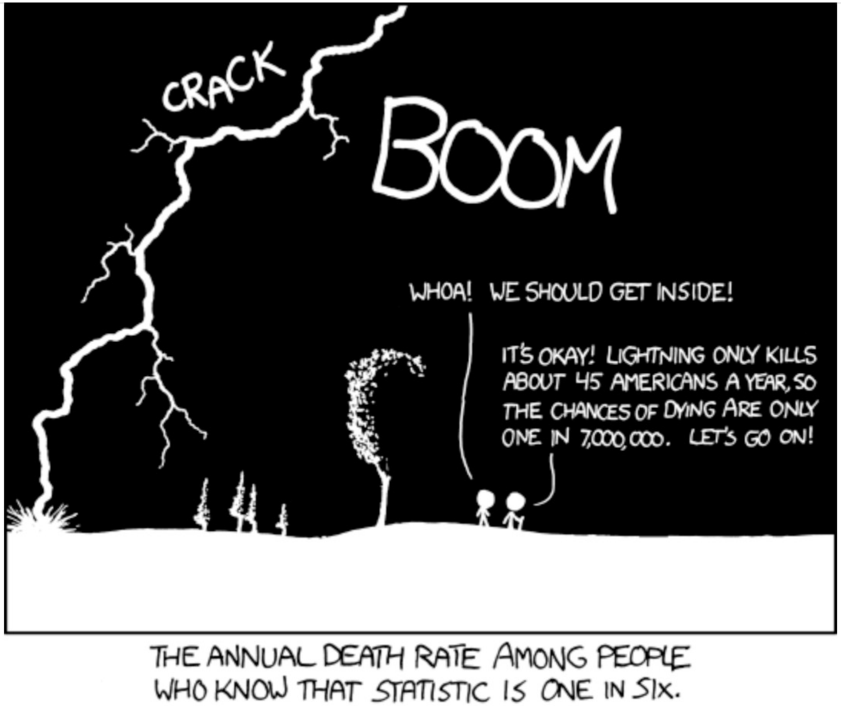
Stat 88: Probability and Statistics in Data Science



https://xkcd.com/795/

Lecture 3: 1/25/2022

Axioms of Probability, Intersections,

Sections 1.3, 2.1

Shobhana M. Stoyanov

Agenda

- Section 1.3: Fundamental Rules (the Axioms of Probability)
  - Notation
  - Axioms
  - Consequences of the axioms
  - De Morgan's Law

- Section 2.1: The Probability of Intersections
  - Conditioning
  - Multiplication rule

- Section 2.2 : Symmetries in Sampling

So far:

$\Omega = \{1, 2, 3, 4, 5, 6\}$ $A = \{1\}$

- Defined ***random experiments***, and their ***outcomes***, ***the outcome space*** (aka *the sample space* Ω), ***events***, ***complements of events***, ***the certain event*** (Ω), ***the impossible event*** $\emptyset$

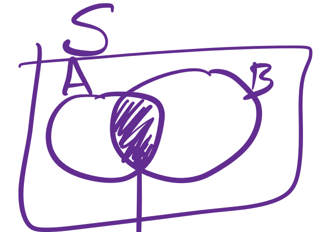
$A^c = \{2,3,4,5,6\}$

- If all the possible outcomes are ***equally likely***, then each outcome has probability $1/n$, where $n = \#(\Omega)$ and $P(A) = \frac{\#(A)}{\#(\Omega)}$, $A \subseteq \Omega$
- Sum of the probabilities of all the distinct outcomes should add to 1
- $0 \leq P(A) \leq 1, A \subseteq \Omega$
- Venn diagrams

$\Omega = \{1,2,3,4,5,6\}$

$P(\text{roll a } 4 \mid \text{roll} > 2) = \frac{1}{\#\{3,4,5,6\}} = \frac{1}{4}$

- A *distribution* of the outcomes over different categories is when each outcome appears in one and only one category.
- When we get some information about the outcome or event whose probability we want to figure out, our outcome space ***reduces***, incorporating that information. We now call the probabilities that we compute ***conditional probabilities***



Notation review: Intersections and Unions

- When two events A ***and*** B ***both*** happen, we call this the ***intersection*** of A and B and write it as

$$A\ and\ B = \mathrm{A} \cap B \text{ (also written as } AB\text{)}$$

~~P(A) ∩ P(B)~~

A∩B

→ AB P(AB)

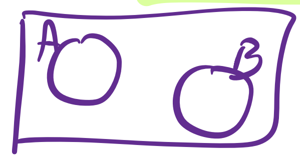
event · a number

- When either A ***or*** B happens, we call this the ***union*** of A and B and write it as

$$A\ or\ B\ =\ A\ \cup B$$

at least one of A or B

- If two events A and B ***cannot both occur*** at the same time, we say that they are *mutually exclusive or disjoint.*



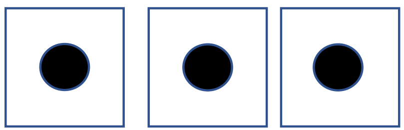
$$\mathrm{A} \cap B = \emptyset$$

$$\cancel{P(A \cap B) = \emptyset}$$

0

symbol for empty set.

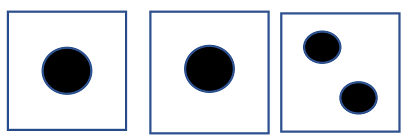
Example of complements

- 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 = {  }

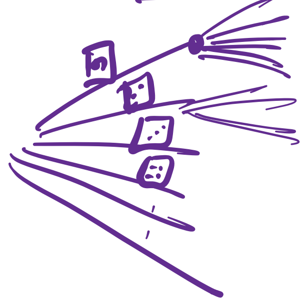
$$P(A) = \frac{1}{216} = \frac{\#(A)}{\#(\Omega)}$$

$$P(A^c) = \frac{215}{216}$$

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



$$6 \cdot 6 \cdot 6 = 216$$



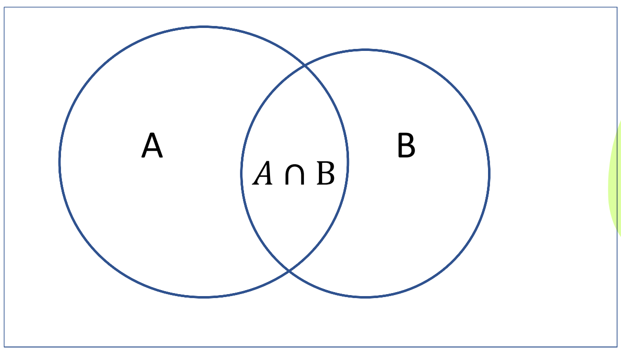
Bounds

A^c, $\sim A$

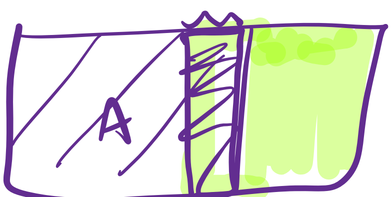
- When we get some information about the outcome or event whose probability we want to figure out, our outcome space reduces, incorporating that information.
- $P(A \cup B)$ for mutually exclusive events

- **Bounds** on probabilities of unions and intersections when events are ***not*** mutually exclusive.

$P(A \cup B) = P(A) + P(B) = 0.7 + 0.5 = \cancel{1.2} > 1$



- P(A) = 0.7, P(B) = 0.5
- $\underline{0.7} \leq P(A \cup B) \leq \underline{1}$
- $\underline{0.2} \leq P(A \cap B) \leq \underline{0.5}$

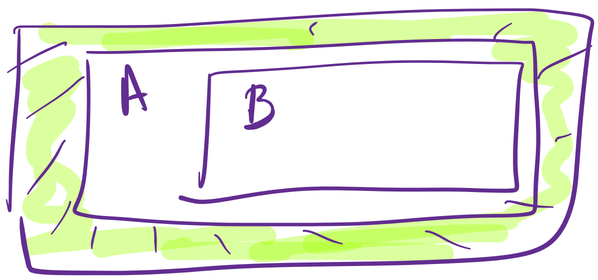


$P(A \cap B) = P(A|B)\,P(B)$

$P(A) + P(B - P(A \cap B \leq 1$

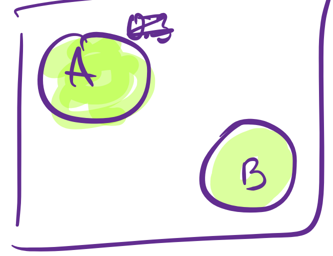
$P(A)=0.7$

$P(A^c)=0.3$

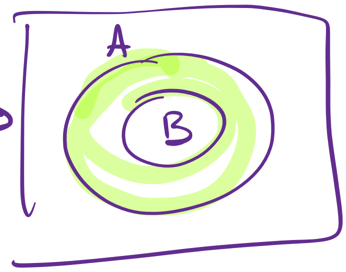


$\underline{0.4} \le \underline{P(A\cup B)} \le \underline{0.7}$

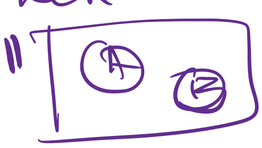
$P(A)=0.4$ $P(B)=0.3$



Smallest value of $A\cup B$ is when the intersection is the greatest (overlap) $\max\{P(A), P(B)\}$



The greatest value of $A\cup B$ is when the intersection is the least



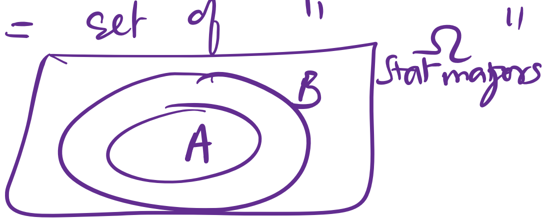
The smallest value of $A\cap B$ is when the over lap is least

The greatest is the min of $P(A)$, $P(B)$ (max. overlap)

Let A = set of all stat majors that have taken Stat 150

B = set of " " that have taken 135

$A \Rightarrow B$ ("implies")



))

$0.2 = 0.7 + 0.5 - 1$

$P(A \cap B) \geq P(A) + P(B) - 1$

$P(A) = 0.3$, $P(B) = 0.4$
Still true

Exercise from Thursday

- A ten-sided fair die is rolled twice:
  - If the first roll lands on 1, what is the chance that the second roll lands on a number bigger than 1?
  - Find the probability that the second number is greater than the *twice* the first number. Let this be A

Ω = set of all outcomes when we roll a 10-sided die twice

$\Omega = \{(1,1), (1,2), \ldots, (1,10), (2,1), (2,2) \ldots\}$

$\#(\Omega) = 100$

Given the first roll is 1

how many such outcomes are there = 10

new $\Omega = \{(1,1), (1,2), (1,3), (1,4) \ldots (1,10)\}$ Ω'

reduced outcome space given first roll is 1

$P(2^{nd} \text{ roll} > 1 \mid 1^{st} \text{ roll is } 1) = 9/10$

$\#(\text{first roll} = 1; 2^{nd} \text{ roll greater} > 2) = 8$

$\#(1^{st} \text{ roll} = 2, 2^{nd} > 4) = 6$

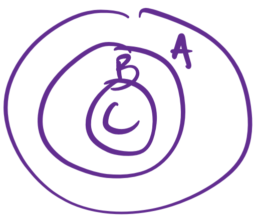
$\#(1^{st} \text{ roll} = 3, 2^{nd} > 6) = 4$

$\#(1^{st} \text{ roll} = 4, 2^{nd} > 8) = 2$

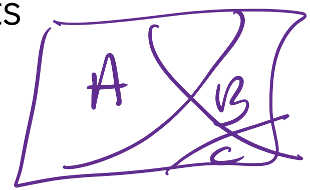
$$P(A) = \frac{\#(A)}{\#(\Omega)} = \frac{20}{100} \leftarrow = \frac{8+6+4+2}{100}$$

Example with bounds

- Let A be the event that you catch the bus to class instead of walking, P(A) = 70%
- Let B be the event that it rains, P(B) = 50%
- Let C be the event that you are on time to class, P(C) = 10%
- What is the chance of **at least one** of these three events happening?

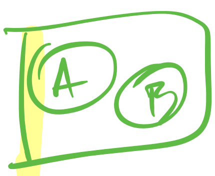


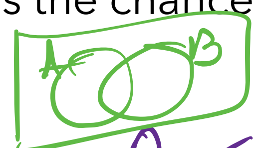
$$0.7 \le P(A \cup B \cup C) \le 100\%$$



$$P(A \cup B \cup C) \le P(A) + P(B) + P(C) = 130\%$$

- What is the chance of **all three** of them happening?





$$0 \le P(A \cap B \cap C) \le 0.1$$

$P(A) = 0.3$
$P(B) = 0.2$
$P(C) = 0.1$
$P(A \cup B \cup C) \le 0.6$
$P(A \cup B \cup C) \le 1$

If $P(A|B) = P(A)$, $P(B|A) = P(B)$ then we say that A & B are INDEPENDENT

$P(A \cap B) = P(A) \cdot P(B)$

Consider a sample space S, & let $A \subseteq S$.
(Note that A might be the empty set ($A = \emptyset$) or A might be S itself.)

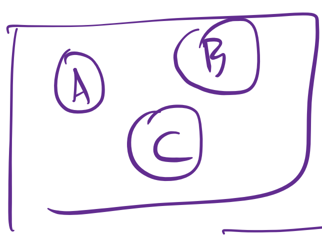
Now consider the probability of A, which is a real number. What properties should the prob. of $A = P(A)$ satisfy?

① $0 \leq P(A) \leq 1$

② If $A \cap B = \emptyset$ then $P(A \cup B) = P(A) + P(B)$

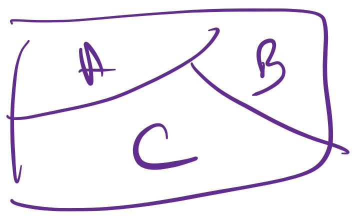
$$P(A \cup B) \leq P(A) + P(B)$$

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



If A, B, C are (pairwise) mutually exclusive

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



§1.3: Fundamental Rules

MacTutor

- 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 $\mathrm{A, B, C} \ldots$
- Recall that Ω is itself an event (called the ***certain*** event) and so is the empty set (denoted $\emptyset$, 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

The Axioms of Probability

Think about probability as a ***function*** on ***events***, so put in an event A, and output $P(A)$, a number between 0 and 1 satisfying the axioms below.

Formally: $A \subseteq \Omega, P(A) \in [0,1]$ such that

1. For every event $A \subseteq \Omega$, we have $P(A) \geq 0$
2. The outcome space is certain, that is: $P(\Omega) = 1$
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 overlap), then the chance of their union is the sum of their probabilities.

$A_1, A_2, \ldots$ s.t. $A_i \cap A_j = \emptyset$

$P(A_1 \cup A_2 \cup \ldots \cup A_n \cup \ldots) = \sum_{i=1}^{\infty} A_i$

Consequences of the axioms

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

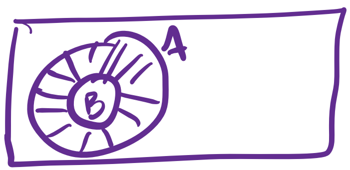
1. **Complement rule**: $P(A^c) = 1 - P(A)$

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

$\frac{35}{36}$

A set minus B

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



everything in A not in B

$A \cap B^c$

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.

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

$P(A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n) \leq P(A_1) +$

$P(A_2) + \dots$
$+ P(A_n)$

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