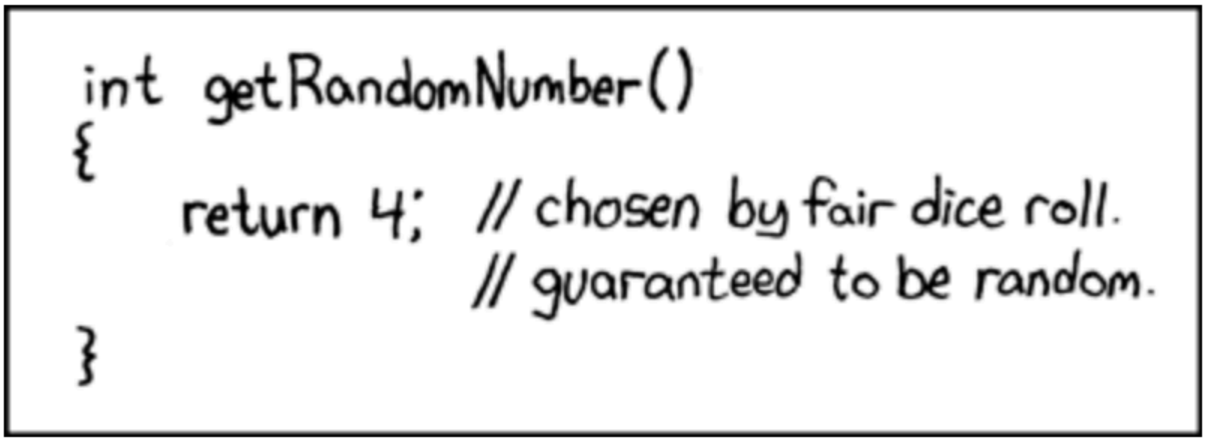
# Stat 88: Probability and Mathematical Statistics in Data Science




https://xkcd.com/221/

Lecture 3: 1/25/2021

Axioms of Probability, Intersections

Sections 1.3, 2.1

Shobhana M. Stoyanov

# Agenda

- The Basics:
  - 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

# Last time:

- 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.

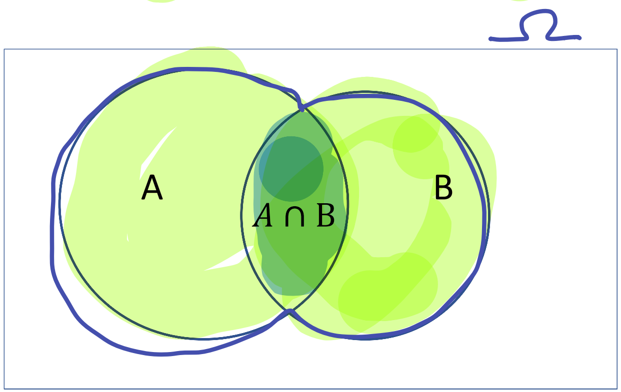
$A \cup B \Leftrightarrow A$ "OR" $B$

$A \cap B \Leftrightarrow A$ "and" $B$

$0.7 + 0.5 = \underline{\underline{1.2}} > 1$

Take diff

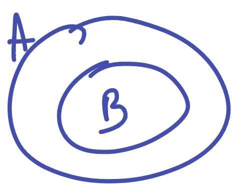
$1.2 - 1 = \min$ $P(A \cap B)$ must be




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

A & B are each $\subseteq A \cup B$

So $P(A \cup B) \geq \max(P(A), P(B))$




$A \cap B \subseteq$ both A & B

So $P(A \cap B) \leq \min(P(A), P(B))$

# Section 1.3: Fundamental Rules

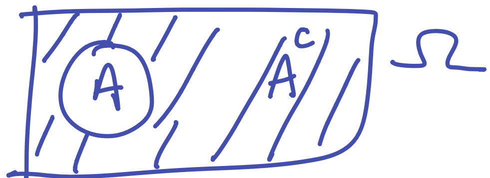
← MacTutor

- Also called "Axioms of probability", first laid out by Kolmogorov
- Recall $\Omega$, the outcome space. Note that $\Omega$ can be finite or infinite.

MacTutor Math biographies

$\Omega = [0, 1]$
any real # b/w 0, 1


- First, some notation:
- Events are denoted (usually) by $A, B, C$ …
- Note that $\Omega$ 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 $A^c$



$\Omega = \{H, T\}$ $P(\Omega) = 1$

# 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

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

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

$$A \text{ or } B = A \cup B$$

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

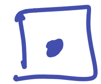
$$A \cap B = \emptyset$$

$\Omega = \{H, T\}$, fair coin

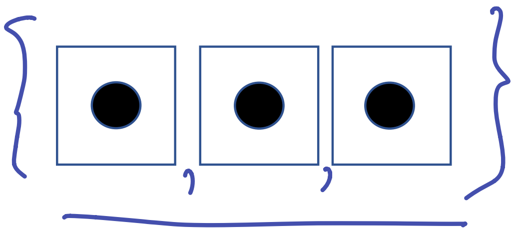
$\phi$, $\{H\}$, $\{T\}$, $\{H, T\}$

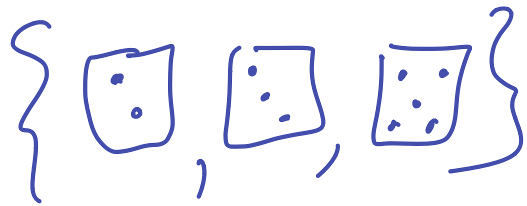
$P(\phi) = 0, \quad P(\{H\}) = \frac{1}{2}, \quad P(\{T\}) = \frac{1}{2}, \quad P(\Omega) = 1$

# Example of complements

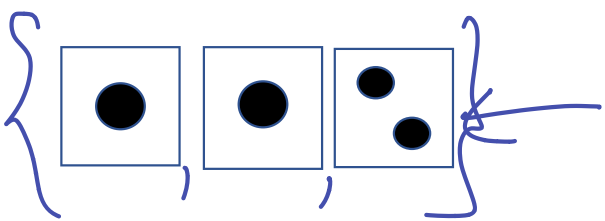


- Roll a die 3 times, let A be the event that we roll an ace **each** time.
- $A^C$ = **not** A, or not *all* aces. It is **not equal** to "never an ace".
- A =





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





A-Complement = everything in $\Omega$ that is not in A

# The Axioms of Probability

P(.) is a SET FUNCTION

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

"is in"

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

$$\boxed{0 \leq P(A) \leq 1}$$

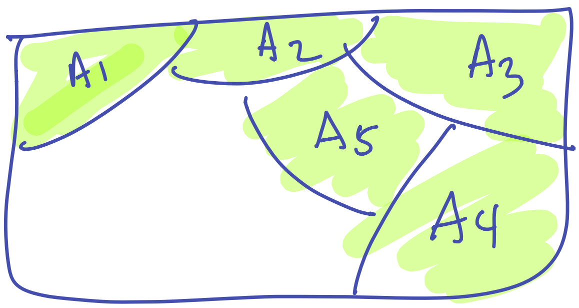
1. For every event $A \subseteq \Omega$, we have $P(A) \geq 0$
2. The outcome space is certain, that is: $P(\Omega) = 1$
3. Addition rule: If two events are mutually exclusive, then the probability of their union is the sum of their probabilities:

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

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.

Let $A_1, A_2, A_3 \ldots A_n$ such that $A_i \cap A_j = \emptyset$ for all $i, j$,

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




Addition rule for unions.

$$P(A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5)$$
$$= P(A_1) + P(A_2) + P(A_3) + P(A_4) + P(A_5)$$

Even if $n \to \infty$, this rule still holds

$A_1, A_2, A_3, \ldots\ldots$ such that $A_i$ are mutually exclusive $P(A_i \cap A_j) = 0$

then $P(A_1 \cup A_2 \cup \ldots) = \underbrace{\sum_{i=1}^{\infty} P(A_i)}_{P(A_1) + P(A_2) + \ldots}$

$\phi = \{\ \}$ emptyset

---

Toss a fair coin twice: Prob of both coins land up the same?

$A_1$ = event both land H

$A_2$ = " " both land T

$A_1 \cap A_2 = \phi$

$\Omega = \{HH, HT, TH, TT\}$

$P(A_1) = 1/4 \quad P(A_2) = \frac{1}{4}$

$P(A_1 \cup A_2) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

$P(A_1 \cup A_2 \cup A_3 \cup \dots \cup A_n) = P(A_1) + P(A_2) + P(A_3) + \dots + P(A_n)$

# Consequences of the axioms

$A \cap A^c = \phi$


$P(A \cup A^c) = 1 = P(\Omega)$

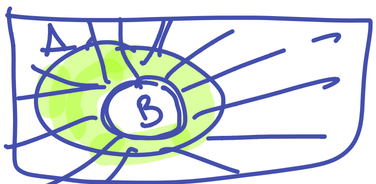
1. $\boxed{\textbf{Complement rule}: P(A^c) = 1 - P(A)}$

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

A = event of not rolling a pair a sixes $\quad P(A^c) = \frac{1}{36}$

$A^c$ = " " rolling pair of sixes $\quad P(A) = 1 - \frac{1}{36} = \frac{35}{36}$

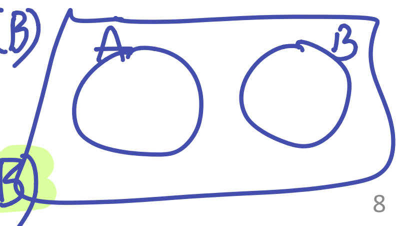
1. **Difference rule**: If $B \subseteq A$, then $P(A \setminus B) = P(A) - P(B)$ where $A \setminus B$ refers to the *set difference between A and B, that* is, all the outcomes that are *A* but not in *B*.

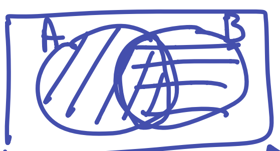



$A \setminus B = A \cap B^c$

2. **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.

if $A \cap B = \phi$, $P(A \cup B) = P(A) + P(B)$




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

Boole's law

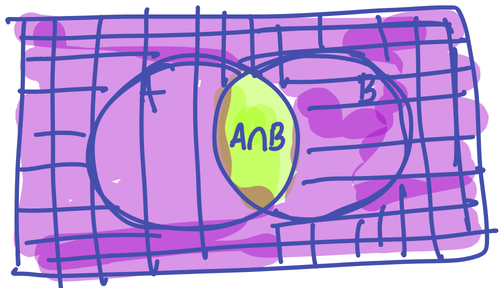
$P(A_1 \cup A_2 \cup \dots \cup A_n) \leq P(A_1) + P(A_2) + \dots + P(A_n)$

Boole's / Bonferroni's

# De Morgan's Laws

- Try to show these using Venn diagrams and shading:

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




$\equiv A^c$

$||| \; B^c$

Exercise: Show using Venn diagrams

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

# Example (Exercise 1.4.5)

- Here's a question from Quora: "If a student applies to ten colleges with a 20% chance of being accepted to each, what are the chances that he will be accepted by at least one college?" Without making any further assumptions, what can you say about this chance?

$A_1$ = accepted by $1^{st}$ college

$\rightarrow A_i$ = " " $i^{th}$ college, $1 \leq i \leq 10$

$$P(A_1 \cup A_2 \cup A_3 \ldots \cup A_{10}) \leq \sum_{i=1}^{10} P(A_i) = 2$$

$$0.2 \leq P\left(\bigcup_{i=1}^{10} A_i\right) \leq 1$$

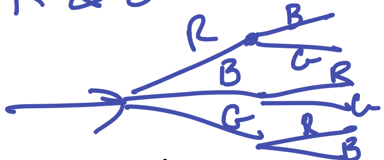
$$A_1 \subseteq \bigcup_{i=1}^{10} A_i$$

# Probability of an intersection

2 balls w/repl.

- 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

$P(1^{st} \text{ is } R \ \& \ 2^{nd} \text{ is blue})$ $\quad \Omega = \{RB, RG, BR, BG, GR, GB\}$




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

this is visualized using a tree diagram


$P(RB) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$

$P(1^{st} R) = \frac{1}{3}$

$P(2^{nd} \text{ is } B \mid 1^{st} \text{ is } R) = \frac{1}{2}$

$P(1^{st} R \ \& \ 2^{nd} B) = P(1^{st} R) \cdot P(2^{nd} B \mid 1^{st} R)$

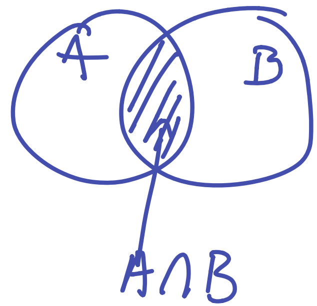
$P(\text{Intersection}) = \frac{1}{3} \cdot \frac{1}{2}$

# Multiplication rule

- Conditional probability written as P(B|A), read as "the probability of the event B, given that the event A has occurred"
- Chance 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.

$A \cap B = AB$

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




- Multiplication rule:

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

$$= P(A) * P(B|A)$$

$$P(AB) = P(BA) = P(B) \times P(A|B)$$

# Multiplication rule

Exercise

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

- Ex.: 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) =