MATH 320: Probability

Lecture 3: Probability

Chapter 1: Probability (1.1)

Probability by counting equally likely outcomes

- Now we can update our original definition of probability using the counting concepts.

- Definition: Let A be an event from a sample space in which all outcomes are equally likely. The **probability of** A, denoted $P(A)$, is defined by:

$$\text{Probability of an event} = \frac{\textit{Number of outcomes in the event}}{\textit{Total number of possible outcomes}}$$

$$P(A) = \frac{n(A)}{n(S)}$$

- Examples: A standard 52 card deck is shuffled and one card is picked at random. Find the probabilities of the following events:

  (a) Club and King: $P(C \cap K) = \frac{1}{52}$

  (b) Club or King: $P(C \cup K) = \frac{13 + 4 - 1}{52} = \frac{16}{52}$

  (c) Not Club nor King: $P[\sim(C \cup K)] = \frac{52 - 16}{52} = \frac{36}{52}$

  (d) Club or Hearts: $P(C \cup H) = \frac{13 + 13}{52} = \frac{26}{52}$

- We will formalize these probability ideas soon.

More counting probability problems

- Now we can use all of the counting tools we've learned and our new probability knowledge to look at more interesting problems.
- *COUNTING STRATEGY*: Solve the numerator and denominator separately.
  - Numerator: "IS a condition" (selecting from restricted sample space).
  - Denominator: "NO condition" (selecting from unrestricted sample space).

$\text{prob} = \frac{\#\ \text{successes}}{\#\ \text{possibilities}}$

- Examples:
  1. A box of jerseys for a pick-up game of basketball contains 8 extra-large jerseys, 7 large jerseys, and 5 medium jerseys. If you are first to the box and grab 3 jerseys, what is the probability that you randomly grab 3 extra-large jerseys.

S: 20, XL: 8

Counting

(1) numerator
→ only selecting from XL jerseys
→ 3 from 8

(2) Denominator
→ select from ALL jerseys
→ 3 from 20

order ✗ ⇒ nCr

$P(3\,XL) = \frac{\binom{8}{3}}{\binom{20}{3}} = \frac{14}{285}$

→ only n changes
→ r is the same

Direct way

$\frac{8}{20} \times \frac{7}{19} \times \frac{6}{18} = \frac{14}{285}$

Selection 1, 2, 3
↳ w/o rep ⇒ drop # each time

  2. Suppose we have a shuffled deck and deal seven cards. What is the probability that we draw no queens?

S: 52, NQ: 48

Counting

(1) select 7 non-queens $\binom{48}{7}$
(2) Select 7 from ALL $\binom{52}{7}$

order ✗

$P(\text{No queens}) = \frac{\binom{48}{7}}{\binom{52}{7}} \approx 0.55$

Direct

$\frac{48}{52} \times \frac{47}{51} \times \cdots \times \frac{42}{46} \approx 0.55$

1, 2, 7

Not queen →

  3. Suppose we have a shuffled deck and deal three cards. What is the probability that we draw exactly one queen?

S: 52, Q: 4, NQ: 48

Counting

(1) Select 1 Q from 4 AND 2 NQ from 48
↳ = *

(2) Select 3 from ALL (52)

$P(1Q, 2NQ) = \frac{\binom{4}{1}\binom{48}{2}}{\binom{52}{3}} \approx 0.204$

→ $\sum r_{num} = r_{denom}$

Direct

$\frac{4}{52} \times \frac{48}{51} \times \frac{47}{50} \approx 0.068 \neq$ Counting

Q NQ NQ

3 sequences $\binom{3}{1}$

NQ Q NQ =

NQ NQ Q =

All same prob

seq * P(seq)

⇒ $3 * 0.068 \approx 0.204$ ✓

- Remember in order to use counting tools when finding probabilities, all outcomes need to be equally likely.

  This means the just the smallest possible results of the experiment (e.g. drawing a single card), not events (e.g. king or heart).

Generalizing probability

Motivation

- In real data studies, outcomes are rarely equally likely. So we need methods to work with probabilities in these scenarios as well.
- Example: A research study into the percentage of births which involve more than one child leads to the following probability table:

Number of children	1	2	3
Probability	0.9760	0.0311	0.0019

1

Intuitively we can easily find the probability of an event, such as a randomly selected birth involving more than one child, based on this table.

$P(\#\text{children} > 1) = P(2 \cup 3)$ 2

$= P(2) + P(3)$ 3

$= 0.0311 + 0.0019 = 0.033$

Let's break down what we implicitly did.

1. Assigned probabilities to each of the individual outcomes in the sample space (i.e. the table).
2. Wrote the event of interest in terms of the outcomes of interest.
3. Added probabilities of the mutually exclusive outcomes.

- This previous example illustrates a natural method for assigning probabilities to events of certain types of experiments.

Sample point method for calculated probabilities

- Theorem: Let $S = \{O_1, \ldots, O_n\}$ be a finite set, where all O_i are individual outcomes each with probability $P(O_i) \geq 0$ and $\sum P(O_i) = 1$. For any $A \in S$,

$$P(A) = \sum_{O_i \in A} P(O_i)$$

- Using this theorem, we have steps / techniques to find probability of any event.

- Example: When players A and B play tennis, the probability that A wins is 2/3. Suppose that A and B play two matches. What is the probability that A wins at least one match?

  Let AB denote the outcome that player A wins the first game and player B wins the second.

  1. What are the sample space and individual outcomes?

     $S = \{AA, AB, BA, BB\}$

  2. Find the probability of each individual outcomes.

     $P(AA) = \frac{2}{3}\frac{2}{3} = \frac{4}{9}$ $P(BA) = \frac{2}{9}$

     $P(AB) = \frac{2}{9}$ $P(BB) = \frac{1}{9}$

  3. Find the event of interest and express it as a union of individual outcomes.

     $W = AA \cup AB \cup BA \rightarrow P(W) = P(AA) + P(AB) + P(BA)$

     $= \{AA, AB, BA\}$ $= 8/9$

     OR in probability $\Longleftrightarrow$ Addition

General definition of probability

- Not all sample spaces are finite or easy to handle. So there are axioms that give general properties that an assignment of probabilities to events must have.
- **Axioms of Probability**: If you define a way to assign a probability $P(A)$ to any event A, the following axioms must be true:

  1. $P(A) \geq 0$ for any event A. $\rightarrow$ nonnegative probabilities
  2. $P(S) = 1$. $\rightarrow$ Total probability $= 1$
  3. Suppose $A_1, \ldots, A_n$ is a (possibly infinite) sequence of pairwise mutually exclusive events. Then

  $$P(A_1 \cup \cdots \cup A_n) = P(A_1) + \cdots + P(A_n)$$

  $\rightarrow$ sample point method = add probabilities if mutually exclusive

- Using this new definition, we have two important properites:

  1. **Any event can be expressed as a union of mutually exclusive outcomes.**
  2. **The probability of an event is the Sum of probabilities of the mutually exclusive outcomes.**

Theorems and their proofs

- Theorem: For probability assignment $P(\cdot)$ and any event A in the sample space S,

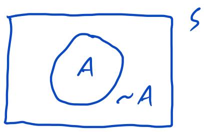
  (a) $P(\emptyset) = 0$ (b) $P(A) \leq 1$ (c) $P(\sim A) = 1 - P(A)$

  Complement probability

- Proofs (easier to prove out of order):

(c) $P(\sim A) = 1 - P(A)$

Strategy for proofs: Rewrite events we know and use axioms to simplify their probabilities.



$\Rightarrow S = A \cup \sim A$

$\Rightarrow P(S) = P(A \cup \sim A)$

$1 = P(A) + P(\sim A)$

$P(\sim A) \overset{\checkmark}{=} 1 - P(A)$

axioms

$\langle \underline{2} \quad \underline{3} \rangle$

$P(S) = 1 \quad A \cap \sim A = \emptyset$

$\Rightarrow ME$

(b) $P(A) \leq 1$ *HINT: Use part (c) of theorem*

$\rightarrow P(\sim A) = 1 - P(A)$ $\langle$ theorem (c) $\rangle$

$P(A) = 1 - \underbrace{P(\sim A)}_{\geq 0}$ by axiom 1

$\Rightarrow P(A) \overset{\checkmark}{\leq} 1$

Another way

$\rightarrow 0 \leq P(\sim A)$ (some event) $\langle$ axiom1 $\rangle$

$= 1 - P(A)$ $\langle$ theorem (c) $\rangle$

$P(A) \overset{\checkmark}{\leq} 1$

Combined with Axiom 1, this theorem gives us bounds on the probability of any event A: $\boxed{0 \leq P(A) \leq 1}$

(a) $P(\emptyset) = 0$ *HINT: Use part (c) of theorem*

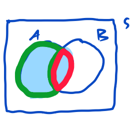
$\rightarrow P(\emptyset) = 1 - P(\sim \emptyset)$ $\langle$ theorem (c) $\rangle$

$= 1 - P(S)$ $\langle \sim\emptyset = S \rangle$

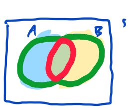
$= 1 - 1$ $\langle$ axiom 2 $\rangle$

$\overset{\checkmark}{=} 0$

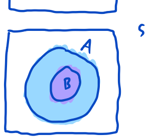
- Theorem: For probability assignment $P(\cdot)$ and any events A and B in the sample space S,



(a) $P(A \cap \sim B) =$ $P(A) - P(A \cap B)$



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



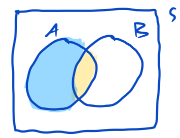
(c) [redacted]

If $B \subset A$, then $P(B) \leq P(A)$

- Proofs:

(a) $P(A \cap \sim B) = P(A) - P(A \cap B)$

HINT: Start with A *and use identity* $A = (A \cap B) \cup (A \cap \sim B)$



$\to A = (A \cap B) \cup (A \cap \sim B)$

$\to P(A) = P[(A \cap B) \cup (A \cap \sim B)]$ ‹ axiom 3 $(A \cap B)$ & $(A \cap \sim B)$ are ME ›

$= P(A \cap B) + P(A \cap \sim B)$

$P(A \cap \sim B) \overset{\checkmark}{=} P(A) - P(A \cap B)$

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

HINT: Use identity $(B \cup A) = (B \cup A) \cap \underbrace{(B \cup \sim B)}_{= S}$

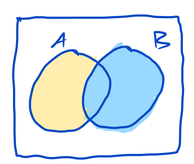
$\to B \cup A = (B \cup A) \cap (B \cup \sim B)$

$= B \cup (A \cap \sim B)$ ‹ by distributive law ›

$\to P(B \cup A) = P[B \cup (A \cap \sim B)]$ ‹ axiom 3, B & $(A \cap \sim B)$ are ME ›

$= P(B) + P(A \cap \sim B)$

$= P(B) + \{P(A) - P(A \cap B)\}$ ‹ theorem (a) ›



$P(A \cup B) \overset{\checkmark}{=} P(A) + P(B) - P(A \cap B)$

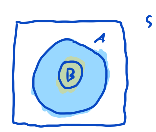
(c) If $B \subset A$, then $P(B) \leq P(A)$

HINT: Start with Axiom 1 and $P(A \cap \sim B)$

$\to 0 \leq P(A \cap \sim B)$ ‹ axiom 1 ›

$= P(A) - P(A \cap B)$ ‹ theorem (a) ›

$= P(A) - P(B)$ ‹ $B \subset A$ ›



$P(B) \overset{\checkmark}{\leq} P(A)$

More examples and concepts / theorems

1. A fair coin is flipped successively until the same face is observed on successive flips. Find the probability that it will take three or more flips of the coin to observe the same face on two consecutive flips.

   $\to S = \{\overset{\sim A}{2}, \overset{A}{\boxed{3, 4, \dots}}\}$

   $\to A = \{3, 4, \dots\}$

   $\to P(A) = 1 - P(\sim A)$

   Important strategy

   $= 1 - \frac{1}{2}$

   $= 1/2$

   2 flips $P(\text{same}) = \frac{2}{4}$

   sample point method $\{HH, HT, TH, HH\}$

   each $P(\cdot) = \frac{1}{4}$

2. Find the probability that in a room of 20 people, there are at least two people sharing the same birthday.

   $P(\text{At least 2 shared}) = 1 - P(\text{None shared})$

   $= 1 - \frac{P(365, 20)}{365^{20}}$ $= \frac{①}{②}$

   ① All different $\Rightarrow$ w/o rep: success $= \frac{365}{1} \times \frac{364}{2} \times \dots \times \frac{346}{20}$

   ② total $= \frac{365}{1} \times \frac{365}{2} \times \dots \times \frac{365}{20}$ w/ rep

   $\approx 1 - 0.589$

   ≈ 0.411

   - Note on order: When solving these types of problems, the numerator and the denominator must ALWAYS MATCH (be consistent) in terms of order.

     So should never have $\cancel{\frac{{}_nP_r}{{}_nC_r}}$

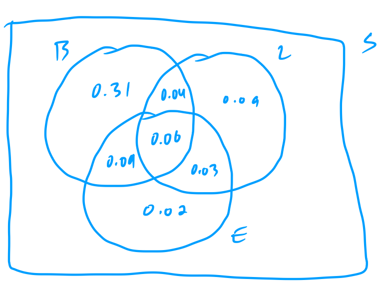
3. A cryptocurrency exchange sells Bitcoin, Litecoin and Ethereum.

   Let $B = \{\text{buys Bitcoin}\}$, $L = \{\text{buys Litecoin}\}$ and $E = \{\text{buys Ethereum}\}$.

   Based on past sales the exchange determines that for any new customer

$$P(B) = 0.50, P(L) = 0.22, P(E) = 0.20, P(B \cap L) = 0.10,$$
$$P(B \cap E) = 0.15, P(L \cap E) = 0.09, \text{ and } P(B \cap L \cap E) = 0.06.$$

   Find the probability that a new customer purchases at least one of these three currencies.



   $P(B \cup L \cup E) = \Sigma$ all individual section probs

   $= 0.64$

   start from inside and work out

   or solve with next theorem

- Theorem: For any events A, B, and C,

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

  The analogous counting rule for the union of three events is:

  $$n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(A \cap C) - n(B \cap C) + n(A \cap B \cap C)$$

Quick review

- Everything we are studying begins from experiments.
  From these, we obtain outcomes, which together make up the sample space.
- We want to know probabilities of events (subsets of S).
  To calculate this, we defined how to assign probabilities in general via the axioms.
- Then there are two methods to compute probabilities:
  1. sample point method $\rightarrow P(\text{Event}) = \sum P(\text{outcomes in event})$
  2. Counting tools $\rightarrow$ only when all outcomes are equally likely

Odds

- Odds of an event A are generally written as a ratio of two integers, such as 2:3 "2 to 3". The odds against A are given by the reverse ratio 3:2.
  Formally, odds are another way to represent probability (but the terms are not interchangeable).
- Definition: The **odds** for an event A are defined as the ratio $P(A)$ to $P(\sim A)$.

  Odds of A = $\dfrac{P(A)}{P(\sim A)} = \dfrac{P(A)}{1-P(A)} = \dfrac{P(\text{Event occurring})}{P(\text{Event not occurring})}$

  odds: for $\dfrac{P(W)}{P(L)}$ Against $\dfrac{P(L)}{P(W)}$
- Converting between odds and probability: $a:b \leftrightarrow P(A) = \dfrac{a}{a+b}$

  Example: Suppose a soccer team is in a playoff game.
  (a) If the probability winning is 0.20, what are the odds of winning?

  $0.2 = 1/5 \rightarrow$ 1 "part" win, 4 "parts" lose $\Rightarrow$ 1:4

  (b) If the odds of losing are 7:9, find the probability of winning.

  7 → L, 9 → W

  $P(W) = \dfrac{W}{L+W} = \dfrac{9}{16}$