Plan for the week

- M: Introduction to Probability
- W: Probability, cont'd
  - Random Variables
  - Mean & variance, revisited
  - Binomial, normal distributions
  - Law of large numbers
  - Central limit theorem
- F: No section this week!
  - Classical Statistics

Introduction to Probability

What is probability?

- Frequentists define the probability of an event as its long-term relative frequency
    - Probability is undefined for a single trial.
    - Nothing is uncertain. (Probabilities are averages.)
- Bayesians define probability as a degree of belief about an event occurring, or about a hypothesis
    - Probabilities reflect knowledge and experience
    - Certain beliefs are expressed using probabilities of 0 and 1
- Regardless of your philosophical stand, all definitions of probability rely on Kolmogorov’s axioms

Set Theory Definitions

- A set is a collection of elements, e.g., {1, 2, 3, 4, five}
- A subset of a set **A** is a set **B** that contains 0 or more elements of **A**.
  - If A = {1, 2, 3, 4, five}, then B = {1, 2, 3} is a subset of A.
- The cardinality of a set is the number of elements in the set.
  - E.g., the cardinality of A is 5. This is written as $|A| = 5$.

Basic Definitions (grounded in set theory)

- An (random) experiment is the process of observing something uncertain.
    - A coin flip, a roll of a die, today's weather, etc.
- An outcome is a result of a random experiment.
- The sample space, or universe, is the set of all possible outcomes, denoted Ω.

Toss a fair coin	$\Omega = \{H, T\}$
Toss two fair coins	$\Omega = \{HH, HT, TH, TT\}$

Roll a fair die	$\Omega = \{1, 2, 3, 4, 5, 6\}$
Roll a loaded die (25% chance of 1)	$\Omega = \{1, 2, 3, 4, 5, 6\}$
Roll two dice	$\Omega = \{(1,1), (1,2), \ldots, (1,6), (2,1), (2,2), \ldots, (2,6), \ldots, (6,1), (6,2), \ldots, (6,6)\}$

Event Probabilities

- An event is any subset of the universe (i.e., a set of outcomes).
- The probability of an event is its likelihood of occurring.
- How do you compute probabilities?
  - First, find the size of the event: i.e., number of outcomes the event comprises.
  - Then, divide by the size of the universe: i.e., the total number of outcomes.

Toss a fair coin	$\Omega = \{H, T\}$	$\lvert\{H\}\rvert \,/\, \lvert\{H, T\}\rvert = 1/2$
Toss two fair coins	$\Omega = \{HH, HT, TH, TT\}$	$\lvert\{HH, HT, TH\}\rvert \,/\, \lvert\{HH, HT, TH, TT\}\rvert = 3/4$

Toss a fair die	$\Omega = \{1, 2, 3, 4, 5, 6\}$	$\lvert\{1\}\rvert \,/\, \lvert\{1, 2, 3, 4, 5, 6\}\rvert = 1/6$
Toss a loaded die (25% chance of 1)	$\Omega = \{1, 2, 3, 4, 5, 6\}$	$p(1) = .25$, $p(\{2, 3, 4, 5, 6\}) = .75$ $p(2) = \lvert\{2\}\rvert \,/\, \lvert\{2, 3, 4, 5, 6\}\rvert\ (.75) = .15$

Event Probabilities

- An event is any subset of the universe (i.e., a set of outcomes).
- The probability of an event is its likelihood of occurring.
- How do you compute probabilities?
  - First, find the size of the event: i.e., number of outcomes the event comprises.
  - Then, divide by the size of the universe: i.e., the total number of outcomes.

Toss a fair coin	Ω = {H, T}	|{H}| / |{H, T}| = 1/2
Toss two fair coins	Ω = {HH, HT, TH, TT}	|{HH, HT, TH}| / |{HH, HT, TH, TT}| = 3/4

Toss a fair die	Ω = {1, 2, 3, 4, 5, 6}	|{1}| / |{1, 2, 3, 4, 5, 6}| = 1/6
Toss a loaded die (25% chance of 1)	Ω = {1a, 1b, 1c, 1d, 1e, 2a, 2b, 2c, 3a, 3b, 3c, 4a, 4b, 4c, 5a, 5b, 5c,	p(1) = |{1a, 1b, 1c, 1d, 1e}| / |Ω| = .25 p(2) = |{2a, 2b, 2c}| / |Ω| = .15 ...

Fair Coin Tosses

- What is the probability of seeing a head when we flip a fair coin?
  - There are two outcomes in the universe, heads and tails ({H, T}).
  - The (sole) outcome we are interested in is heads.
  - Thus, the probability is ½.
- What is the probability of seeing a head when we flip two fair coins?
  - There are four outcomes in the universe: {HH, HT, TH, TT}.
  - The outcomes (plural!) we are interested in are HH, HT, and TH.
  - Thus, the probability is ¾.

Toss a fair coin, See heads	Ω = {H, T}	|{H}| / |{H, T}| = 1/2
Toss two fair coins See heads	Ω = {HH, HT, TH, TT}	|{HH, HT, TH}| / |{HH, HT, TH, TT}| = 3/4

Rolling the dice

- What is the probability of seeing a 1 when you roll a fair die?
    - There are six outcomes in the universe: {1, 2, 3, 4, 5, 6}.
    - The outcome we are interested in is 1.
    - Thus, the probability is ⅙.
- What is the probability of seeing a 2 when you roll a loaded die that favors 1?
    - There are six outcomes in the universe: {1, 2, 3, 4, 5, 6}.
    - The probability of seeing a 1 is .25. The probability of seeing something else is .75.
    - The outcome we are interested in is 2.
    - Thus, the probability is .15.

Toss a fair die	Ω = {1, 2, 3, 4, 5, 6}	|{1}| / |{1, 2, 3, 4, 5, 6}| = 1/6
Toss a loaded die (25% chance of 1)	Ω = {1, 2, 3, 4, 5, 6}	p(1) = .25, p({2, 3, 4, 5, 6}) = .75 |{2}| / |{2, 3, 4, 5, 6}| (.75) = .15

Rolling the dice

- What is the probability of seeing a 1 when you roll a fair die?
  - There are six outcomes in the universe: {1, 2, 3, 4, 5, 6}.
  - The outcome we are interested in is 1.
  - Thus, the probability is ⅙.
- What is the probability of seeing a 2 when you roll a loaded die that favors 1?
  - There are 20 outcomes in the universe:
    {1a, 1b, 1c, 1d, 1e, 2a, 2b, 2c, 3a, 3b, 3c, 4a, 4b, 4c, 5a, 5b, 5c, 6a, 6b, 6c}.
  - The event we are interested in is 2, represented by {2a, 2b, 2c}.
  - Thus, the probability is $p(2) = |\{2a, 2b, 2c\}| / |\Omega| = .15$.

Toss a fair die	$\Omega = \{1, 2, 3, 4, 5, 6\}$	$\|\{1\}\| / \|\{1, 2, 3, 4, 5, 6\}\| = 1/6$
Toss a loaded die (25% chance of 1)	$\Omega = \{1a, 1b, 1c, 1d, 1e,$ $2a, 2b, 2c, 3a, 3b, 3c,$ $4a, 4b, 4c, 5a, 5b, 5c,$	$p(1) = \|\{1a, 1b, 1c, 1d, 1e\}\| / \|\Omega\| = .25$ $p(2) = \|\{2a, 2b, 2c\}\| / \|\Omega\| = .15$ …

ICA

- In the game of Monopoly, a player rolls two dice. The player then adds their rolls together, and moves that many spaces.
- There is a *Free Parking* space, and if a player lands on it, they get a payout of some amount of (Monopoly) money.
- It is your turn, and you are 8 spaces away from the *Free Parking* space.
- What is the probability you land on *Free Parking* on your next turn?

Enumerating the Sample Space

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Finding the Event of Interest

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Solution

- There are 36 outcomes in all.
- We can roll two dice that sum to 8 in five different ways.
  - The event of interest is {(6,2), (5,3), (4,4), (3,5), (2,6)}.
- Hence, the probability that you will land on *Free Parking* is 5/36!

Axioms and Laws of Probability

Kolmogorov's Axioms of Probability

- Probability is a branch of mathematics. As such, it is an axiomatic science.
- An axiom is a statement that is regarded as self-evidently true.
- In probability, there are 3 basic axioms:
  - Probabilities cannot be negative.
  - Probabilities cannot exceed 1.
    - Axioms 1 and 2 imply that all probabilities are bounded between 0 and 1, inclusive.
  - The probability of the union of two disjoint events, A and B, is the sum of their respective probabilities: i.e., P(A or B) = P(A) + P(B)
    - In particular, the sum of the probabilities of all outcomes is 1.
- We will check that our definition of probabilities satisfies these basic axioms: $P(A) = |A|/|\Omega|$.

Axiom 1

Probabilities cannot be negative:

- Recall how we calculate probabilities
  - In the monopoly example, we counted how many outcomes an event comprises, and then we divided that number by the size of the universe (i.e., the total number of outcomes).
- At worst, an event occurs 0 times (e.g., rolling a 7 on a six-sided die)!
- Since cardinalities are bounded below by 0, so too are probabilities.

Axiom 2

Probabilities cannot exceed 1:

- $P(\Omega) = |\Omega|/|\Omega| = 1$.
- The universe Ω includes all possible outcomes.
- The probability that an event outside the universe occurs is 0.
- Therefore, the probability that an event inside the universe occurs is 1!

Axiom 3

The probability of the union of two disjoint events, A and B, is the sum of their respective probabilities: i.e., P(A or B) = P(A) + P(B):

- $P(A \text{ or } B) = |A \cup B|/|\Omega|$
- $P(A) + P(B) = |A|/|\Omega| + |B|/|\Omega|$
- Must show: $|A \cup B|/|\Omega| = |A|/|\Omega| + |B|/|\Omega|$
- But $|A|/|\Omega| + |B|/|\Omega| = (|A| + |B|)/|\Omega| = |A \cup B|/|\Omega|$, because A and B are disjoint

Example of Axiom 3

- If A and B are mutually exclusive, then
  - P(A or B) = P(A) + P(B)
- Example: Roll two fair coins
  - Event A: “two heads”
  - Event B: “two tails”
  - What is the probability of A or B?
  - P(A or B) = ¼ + ¼ = ½

Event A	{HH}	|{HH}| / |{HH, HT, TH, TT}| = 1/4
Event B	{TT}	|{TT}| / |{HH, HT, TH, TT}| = 1/4
Event A or B	{HH, TT}	|{HH, TT}| / |{HH, HT, TH, TT}| = 1/2

Corollary of Axiom 3

The sum of the probabilities of all outcomes must be 1:

- The universe is the union of all outcomes.
- Outcomes do not overlap.
  - For example, you cannot roll both a 1 and a 2 at the same time!
- So, if we add up the probabilities of all the outcomes, we should get the probability of the universe.
- But we already know that the probability of the universe is 1: i.e., $P(\Omega) = 1$.

Laws of Probability

Definition of probability:

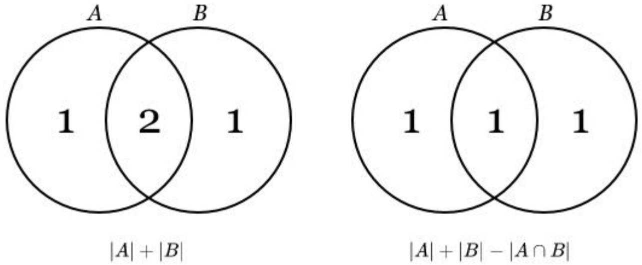
- $P(A) = |A|/|\Omega|$
- $P(|\Omega|) = 1$

Laws implies by the axioms:

- $P(\emptyset) = 0$
- $P(\Omega \setminus A) = 1 - P(A)$
- If $A \subseteq B$, then $P(A) \leq P(B)$
- Inclusion-exclusion principle

Inclusion-Exclusion Principle

- If A and B are **not** mutually exclusive, then
  - P(A or B) = P(A) + P(B) - P(A and B)
- Example: Flip two fair coins
  - Event A: “Heads on the first coin”
  - Event B: “Tails on the second”
  - What is the probability of A or B?
  - P(A or B) = ½ + ½ - ¼ = ¾



Event A	{HH, HT}	|{HH, HT}| / |{HH, HT, TH, TT}| = 1/2
Event B	{HT, TT}	|{HT, TT}| / |{HH, HT, TH, TT}| = 1/2
Event A and B	{HT}	|{HT}| / |{HH, HT, TH, TT}| = 1/4
Event A or B	{HH, HT, TT}	|{HH, HT, TT}| / |{HH, HT, TH, TT}| = 3/4

ICA

Assume the following:

- There is a 60% chance of rain today
- There is a 75% chance of rain tomorrow
- There is a 20% chance of no rain either day

What is the probability it will rain today and tomorrow?

	Rain tomorrow (.75)	No rain tomorrow (.25)
Rain today (.6)	?	?
No rain today (.4)	?	.2

Define the relevant events

A: Rain today

B: Rain tomorrow

A or B: Rain today or tomorrow

A and B: Rain today and tomorrow

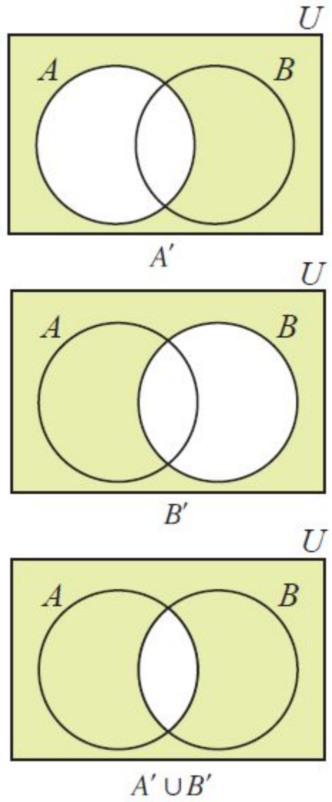
DeMorgan's Law

¬A: No rain today

¬B: No rain tomorrow

¬A and ¬B: No rain today and no rain tomorrow

¬(¬A and ¬B) = A or B: Rain today or rain tomorrow



DeMorgan's Law

¬A: No rain today
¬B: No rain tomorrow

¬A and ¬B: No rain today and no rain tomorrow
¬(¬A and ¬B) = A or B: Rain today or rain tomorrow

A	B	A or B	¬A	¬B	¬A and ¬B	¬(¬A and ¬B)
1	1	1	0	0	0	1
1	0	1	0	1	0	1
0	1	1	1	0	0	1
0	0	0	1	1	1	0

Truth Table

The probability of rain today and tomorrow

P(A and B)

= P(A) + P(B) - P(A or B)

= P(Rain today) + P(Rain tomorrow) - P(Rain today or tomorrow)

= .6 + .75 - ?

The probability of rain today or tomorrow

P(A or B)

= P(Rain today or rain tomorrow)

= 1 - P(No rain today and no rain tomorrow)

= 1 - .2

= .8

The probability of rain today and tomorrow

P(A and B)

= P(A) + P(B) - P(A or B)

= P(Rain today) + P(Rain tomorrow) - P(Rain today or tomorrow)

= .6 + .75 - .8

= .55

ICA

Assume the following:

- There is a 60% chance of rain today
- There is a 75% chance of rain tomorrow
- There is a 20% chance of no rain either day

What is the probability it will rain today and tomorrow?

	Rain tomorrow (.75)	No rain tomorrow (.25)
Rain today (.6)	.55	?
No rain today (.4)	?	.2

ICA

Assume the following:

- There is a 60% chance of rain today
- There is a 75% chance of rain tomorrow
- There is a 20% chance of no rain either day

What is the probability it will rain today and tomorrow?

	Rain tomorrow (.75)	No rain tomorrow (.25)
Rain today (.6)	.55	.05
No rain today (.4)	.2	.2

Conditional Probability

Conditional Probability

- In 2020, Brown had an acceptance rate of 6%.
  - Does this mean that all applicants had an identical 6% chance of getting in?
  - Probably not! If you won the Nobel Peace Prize in your senior year of high school, then you probably had a higher likelihood of getting in.
- The probability of gaining weight if you don’t exercise is probably higher than the probability of gaining weight if you do.
- Conditional probabilities are probabilities, given an event, or “condition”.
- We write $P(A|B)$, and say “the probability of A given B”

A Worked Example

- Experimental setup: Roll two die, and record their sum.
- Given that the first roll is even, what is the probability the sum is greater than or equal to 8?

Enumerating the Sample Space

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Conditioning the Sample Space

	1	2	3	4	5	6
~~1~~	~~2~~	~~3~~	~~4~~	~~5~~	~~6~~	~~7~~
2	3	4	5	6	7	8
~~3~~	~~4~~	~~5~~	~~6~~	~~7~~	~~8~~	~~9~~
4	5	6	7	8	9	10
~~5~~	~~6~~	~~7~~	~~8~~	~~9~~	~~10~~	~~11~~
6	7	8	9	10	11	12

Finding the Event of Interest

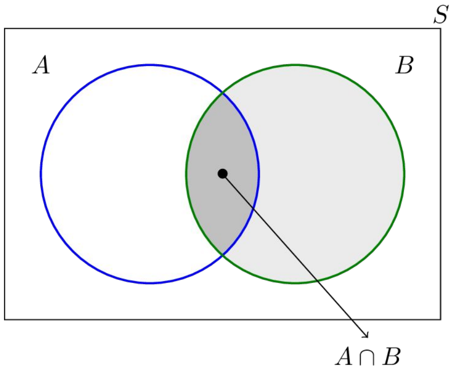
	1	2	3	4	5	6
~~1~~	~~2~~	~~3~~	~~4~~	~~5~~	~~6~~	~~7~~
2	3	4	5	6	7	8
~~3~~	~~4~~	~~5~~	~~6~~	~~7~~	~~8~~	~~9~~
4	5	6	7	8	9	10
~~5~~	~~6~~	~~7~~	~~8~~	~~9~~	~~10~~	~~11~~
6	7	8	9	10	11	12

Computing the Conditional Probability

- After conditioning on the first roll being even, there are 18 outcomes.
- Among these outcomes *only*, we can roll two dice whose sum is at least 8 in nine different ways.
  - The event of interest is {(2,6), (4,4), (4,5), (4,6), (6,2), (6,3), (6,4), (6,5), (6,6)}.
- Hence, the conditional probability is 9/18 = ½!

Formula for Conditional Probability

- P(A|B) = |A and B|/|B| = P(A and B)/P(B)
- Why? Because |A and B|/|B| = (|A ∩ B|/|Ω|) / (|B|/|Ω|)
- Notice, this implies the probability of P(A and B) = P(A|B)P(B)
- Moreover, P(A and B) = P(A|B)P(B) = P(B|A)P(A)



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

Image source

Independent vs. Dependent Events

Example 1

- Let A be the event that you get an A in this class
- Let B be the event that you take another computer science course
- Do these sound like independent events?
- Not so much

Example 2

- Let A be the event that you get an A in this class
- Let B be the event that I go to East Side Pockets for lunch
- These sound like independent events

Independent Events

- Let A be the event that you get an A in this class
  - Suppose this event has probability .8
- Let B be the event that I go to East Side Pockets for lunch
  - Suppose this event has probability .4
- What is the probability that you get an A in this class, given that I go to East Side Pockets for lunch?
- What I eat for lunch does not affect your grade, so the probability of you getting an A remains .8, regardless of where I eat lunch.

Independent Events (cont’d)

- Recall the following, from the definition of conditional probability:
  - P(A and B) = P(A|B)P(B)
  - P(A and B) = P(B|A)P(A)

- Mathematically, two events are independent if:
  - P(A|B) = P(A)
  - P(B|A) = P(B)

- Consequently, two events are independent if P(A and B) = P(A)P(B).

Independent vs. Dependent Events

INDEPENDENT	Rain tomorrow (.75)	No rain tomorrow (.25)
Rain today (.6)	.45 = (.6)(.75)	.15 = .6 - .45
No rain today (.4)	.3 = .4 - .3	.1

DEPENDENT	Rain tomorrow (.75)	No rain tomorrow (.25)
Rain today (.6)	.55	.05 = .6 - .55
No rain today (.4)	.2 = .4 - .2	.2