Chapter 2

Axioms of Probability

Jingchen (Monika) Hu

Vassar College

MATH 241

Outline

Sample space

Definition

A sample space S is the set of all possible outcomes of an experiment.

Examples:	3 coin tosses	S $= \{$HHH, HHT, HTH, HTT, THH, THT, TTH, TTT$\}$
	One die roll	S $= \{$1,2,3,4,5,6$\}$
	Sum of two rolls	S $= \{$2,3,...,11,12$\}$
	Seconds waiting for bus	S $= [0, \infty)$

Examples of sample spaces

- Experiment is playing five rounds of Russian roulette
  Sample space is $\{D, LD, LLD, LLLD, LLLLD, LLLLL\}$.

  Why is this different than coin flipping?

Examples of sample spaces

- Experiment is playing five rounds of Russian roulette
  Sample space is $\{D, LD, LLD, LLLD, LLLLD, LLLLL\}$.

  Why is this different than coin flipping?

- Experiment is sequencing three nucleotides
  Sample space is $\{AAA, CCC, GGG, TTT, AAC, AAT, AAG, ...\}$.

Examples of sample spaces

- Experiment is playing five rounds of Russian roulette
  Sample space is $\{D, LD, LLD, LLLD, LLLLD, LLLLL\}$.

  Why is this different than coin flipping?

- Experiment is sequencing three nucleotides
  Sample space is $\{AAA, CCC, GGG, TTT, AAC, AAT, AAG, ...\}$.

How big is this sample space? (Hint: there are four types of nucleotides.)

Examples of sample spaces

- Experiment is playing five rounds of Russian roulette
  Sample space is $\{D, LD, LLD, LLLD, LLLLD, LLLLL\}$.

  Why is this different than coin flipping?

- Experiment is sequencing three nucleotides
  Sample space is $\{AAA, CCC, GGG, TTT, AAC, AAT, AAG, ...\}$.

How big is this sample space? (Hint: there are four types of nucleotides.)

$$4 \times 4 \times 4 = 64$$

An event

Definition

An event E is any subset of the sample space S.

$$E \subset S$$

An event

Definition

An event E is any subset of the sample space S.

$$E \subset S$$

Examples:	2 heads	E $= \{$HHT, HTH, THH$\}$
	Even number	E $= \{$2,4,6$\}$
	$<$ 2 minutes	E $= [0, 120)$

An event

Definition

An event E is any subset of the sample space S.

$$E \subset S$$

Examples:	2 heads	E $= \{$HHT, HTH, THH$\}$
	Even number	E $= \{$2,4,6$\}$
	$<$ 2 minutes	E $= [0, 120)$

- Impossible event: empty set $\emptyset \subset S$
- $S \subset S$

Set theory

Let A, B be two events.

Definition

1. **Intersection** $A \cap B$: *implies the event that both A and B occur*
2. **Union** $A \cup B$: *implies the event that at least one of A or B occur*
3. *The* **complement** *of an event A denoted A^c (also notated A' or $\bar{A}$):* $A^c = S \backslash A$ *- the event that A does not occur*
4. $A \subset B$ *implies that the occurrence of A implies the occurrence of B*

Venn diagram

More set theory

Definition

Two events A and B are **mutually exclusive** *or* **disjoint** *if they have no outcomes in common, i.e. $A \cap B = \emptyset$.*

More set theory

Definition

Two events A *and* B *are* **mutually exclusive** *or* **disjoint** *if they have no outcomes in common, i.e.* $A \cap B = \emptyset$.

Examples of exclusive events?

Some rules

1. Commutative laws

$$A \cup B = B \cup A, \quad A \cap B = B \cap A$$

Some rules

1. Commutative laws

$$A \cup B = B \cup A, \quad A \cap B = B \cap A$$

2. Associative laws

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

$$(A \cap B) \cap C = A \cap (B \cap C)$$

Some rules

1. Commutative laws

$$A \cup B = B \cup A, \quad A \cap B = B \cap A$$

2. Associative laws

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

$$(A \cap B) \cap C = A \cap (B \cap C)$$

3. Distributive laws

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

DeMorgan's Laws

$$(A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c$$

DeMorgan's Laws

$$(A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c$$

- Outline of proof (to the first equation): two steps
  Left $\subset$ Right $\Longleftrightarrow$ For any $x \in (A \cup B)^c$, then $x \in A^c \cap B^c$.
  Right $\subset$ Left $\Longleftrightarrow$ For any $x \in A^c \cap B^c$, then $x \in (A \cup B)^c$.

DeMorgan's Laws

$$(A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c$$

- Outline of proof (to the first equation): two steps
  Left $\subset$ Right $\Longleftrightarrow$ For any $x \in (A \cup B)^c$, then $x \in A^c \cap B^c$.
  Right $\subset$ Left $\Longleftrightarrow$ For any $x \in A^c \cap B^c$, then $x \in (A \cup B)^c$.

- DeMorgan's laws can be generalized to n events $A_1, \ldots, A_n$:

$$\left(\bigcup_{i=1}^{n} A_i\right)^c = \bigcap_{i=1}^{n} A_i^c, \quad \left(\bigcap_{i=1}^{n} A_i\right)^c = \bigcup_{i=1}^{n} A_i^c$$

DeMorgan's Laws

$$(A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c$$

- Outline of proof (to the first equation): two steps
  Left $\subset$ Right $\iff$ For any $x \in (A \cup B)^c$, then $x \in A^c \cap B^c$.
  Right $\subset$ Left $\iff$ For any $x \in A^c \cap B^c$, then $x \in (A \cup B)^c$.

- DeMorgan's laws can be generalized to n events $A_1, \ldots, A_n$:

$$\left(\bigcup_{i=1}^{n} A_i\right)^c = \bigcap_{i=1}^{n} A_i^c, \quad \left(\bigcap_{i=1}^{n} A_i\right)^c = \bigcup_{i=1}^{n} A_i^c$$

https://www.youtube.com/watch?v=oOx7FSzSav4

Recap

- Sample space S, event E
- Set operations: intersection, union, complement, subset, disjoint/mutually exclusive
- Rules: commutative, associative, distributive, DeMorgan's laws
- Hint: Venn Diagram is helpful

Outline

What is probability?

Given an experiment and a sample space S, the objective of probability is to assign to each event A a number $P(A)$, called the probability of the event A, which will give a precise measure of the chance that A will occur.

What is probability?

Given an experiment and a sample space S, the objective of probability is to assign to each event A a number $P(A)$, called the probability of the event A, which will give a precise measure of the chance that A will occur.

What is $P(A)$ of the following events (A's)?

1. $A =$ Someone in this class room wins the MegaMillion.
2. $A =$ You spin a quarter and it comes up heads.
3. $A =$ You spin a quarter and it stands up.
4. $A =$ At least two people in this class room have the same birthday.

Probability: Frequentist interpretation

- There are several possible interpretations of probability.
- There is no agreement, at all, in how probabilities should be interpreted.
- There is (nearly) complete agreement on the mathematical rules probability must follow (axioms).

Probability: Frequentist interpretation

- There are several possible interpretations of probability.
- There is no agreement, at all, in how probabilities should be interpreted.
- There is (nearly) complete agreement on the mathematical rules probability must follow (axioms).
- *Frequentist* interpretation: The probability of event A is the proportion of times (frequency) that A occurs in an infinite sequence (or very long run) of separate tries of the experiment.

$$P(A) = \lim_{n\to\infty} \frac{\#\text{ times A happens}}{n}.$$

Probability: Frequentist interpretation

- There are several possible interpretations of probability.
- There is no agreement, at all, in how probabilities should be interpreted.
- There is (nearly) complete agreement on the mathematical rules probability must follow (axioms).
- *Frequentist* interpretation: The probability of event A is the proportion of times (frequency) that A occurs in an infinite sequence (or very long run) of separate tries of the experiment.

$$P(A) = \lim_{n \to \infty} \frac{\#\text{ times A happens}}{n}.$$

- Often associated with Jerzy Neyman and Egon Pearson who described the logic of statistical hypothesis testing.

Probability: Bayesian interpretation

- A *Bayesian* can pick whatever number they prefer for $P(A)$, based on their own personal experience and intuition, provided that number is consistent with all of the other probabilities they choose in life.

Probability: Bayesian interpretation

- A *Bayesian* can pick whatever number they prefer for $P(A)$, based on their own personal experience and intuition, provided that number is consistent with all of the other probabilities they choose in life.
- A Bayesian interprets probability as a subjective degree of belief: *For the same event, two separate people could have differing probabilities.*

Probability: Bayesian interpretation

- A *Bayesian* can pick whatever number they prefer for $P(A)$, based on their own personal experience and intuition, provided that number is consistent with all of the other probabilities they choose in life.
- A Bayesian interprets probability as a subjective degree of belief: *For the same event, two separate people could have differing probabilities.*
- The Bayesian's view must: (1) conform to all other personal opinions; (2) change as new data arise according to *Bayes' Rule*.

Probability: Bayesian interpretation

- A *Bayesian* can pick whatever number they prefer for $P(A)$, based on their own personal experience and intuition, provided that number is consistent with all of the other probabilities they choose in life.
- A Bayesian interprets probability as a subjective degree of belief: *For the same event, two separate people could have differing probabilities.*
- The Bayesian's view must: (1) conform to all other personal opinions; (2) change as new data arise according to *Bayes' Rule*.
- Bayesian interpretations of probabilities avoid some of the philosophical difficulties of frequency interpretations.

Probability: Bayesian interpretation

- A *Bayesian* can pick whatever number they prefer for $P(A)$, based on their own personal experience and intuition, provided that number is consistent with all of the other probabilities they choose in life.
- A Bayesian interprets probability as a subjective degree of belief: *For the same event, two separate people could have differing probabilities.*
- The Bayesian's view must: (1) conform to all other personal opinions; (2) change as new data arise according to *Bayes' Rule*.
- Bayesian interpretations of probabilities avoid some of the philosophical difficulties of frequency interpretations.
- Named after the 18^{th} century Presbyterian minister and mathematician Thomas Bayes.
- Largely popularized by revolutionary advance in computational technology and methods during the last twenty years.

Axiomatic Foundations of Probability

- The foundations of modern probability theory, the axiomatic basis, are laid by Andrey Kolmogorov in 1933.

Axiomatic Foundations of Probability

- The foundations of modern probability theory, the axiomatic basis, are laid by Andrey Kolmogorov in 1933.

- The axiomatic approach is not concerned with the interpretation of probabilities.
- Concerned only that probabilities are defined by a function satisfying the axioms.

Axiomatic Foundations of Probability

- The foundations of modern probability theory, the axiomatic basis, are laid by Andrey Kolmogorov in 1933.

- The axiomatic approach is not concerned with the interpretation of probabilities.
- Concerned only that probabilities are defined by a function satisfying the axioms.
- Kolmogorov (1903-1987) was one of the greatest mathematicians of the 20th century. This axiomatization was one of his "trivial" accomplishment.

Axioms of probability

Definition

Let P be a function that assigns a nonnegative real number to each event E of a sample space S. We call P a probability *if*

Axioms of probability

Definition

Let P be a function that assigns a nonnegative real number to each event E of a sample space S. We call P a probability if

1. *Axiom 1: non-negative*

$$0 \leq P(E) \leq 1$$

Axioms of probability

Definition

Let P be a function that assigns a nonnegative real number to each event E of a sample space S. We call P a probability if

1. *Axiom 1: non-negative*
$$0 \leq P(E) \leq 1$$
2. *Axiom 2: total one*
$$P(S) = 1$$

Axioms of probability

Definition

Let P be a function that assigns a nonnegative real number to each event E of a sample space S. We call P a probability if

1. *Axiom 1: non-negative*
$$0 \leq P(E) \leq 1$$
2. *Axiom 2: total one*
$$P(S) = 1$$
3. *Axiom 3: countable addition*

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i), \text{ if } E_i \cap E_j = \emptyset \text{ for } i \neq j$$

Axioms of probability

Definition

Let P be a function that assigns a nonnegative real number to each event E of a sample space S. We call P a probability if

1. *Axiom 1: non-negative*
$$0 \leq P(E) \leq 1$$
2. *Axiom 2: total one*
$$P(S) = 1$$
3. *Axiom 3: countable addition*

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i), \text{ if } E_i \cap E_j = \emptyset \text{ for } i \neq j$$

In particular, for k **disjoint** events $E_1, \ldots, E_k$,

$$P\left(\bigcup_{i=1}^{k} E_i\right) = \sum_{i=1}^{k} P(E_i)$$

Outline

Propositions

☞ Complement Rule:

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

Proof (hint: use Axiom 2 and Axiom 3)

Propositions

☞ Complement Rule:

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

Proof (hint: use Axiom 2 and Axiom 3)

- $P(\emptyset) = 0$

☞ Difference Rule:

$$P(B \cap A^c) = P(B) - P(A), \text{ if } A \subseteq B$$

Proof (hint: use Axiom 3)

☞ Difference Rule:

$$P(B \cap A^c) = P(B) - P(A), \text{ if } A \subseteq B$$

Proof (hint: use Axiom 3)

- $P(B) \geq P(A)$, if $A \subseteq B$

Proof (hint: use Axiom 1)

☞ Inclusion-Exclusion: two events A, B (not necessarily disjoint)

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

Proof (hint: use Axiom 3 and Venn-diagram)

Inclusion-Exclusion example

Suppose that for a randomly selected student in a probability class,

- $P(\text{live Eastern Time Zone at home}) = 63\%$.
- $P(\text{senior}) = 41\%$.
- $P(\text{live Eastern Time Zone at home and senior}) = 31\%$.

Find the probability that a student is either live Eastern Time Zone at home or a senior.

Inclusion-Exclusion example

Suppose that for a randomly selected student in a probability class,

- P(live Eastern Time Zone at home) $= 63\%$.
- P(senior) $= 41\%$.
- P(live Eastern Time Zone at home and senior) $= 31\%$.

Find the probability that a student is either live Eastern Time Zone at home or a senior.

1. Conditions: event $A = \{$live Eastern Time Zone at home$\}$, $B = \{$senior$\}$,

$$P(A) = 0.63, P(B) = 0.41, P(A \cap B) = 0.31$$

Inclusion-Exclusion example

Suppose that for a randomly selected student in a probability class,

- P(live Eastern Time Zone at home) $= 63\%$.
- P(senior) $= 41\%$.
- P(live Eastern Time Zone at home and senior) $= 31\%$.

Find the probability that a student is either live Eastern Time Zone at home or a senior.

1. Conditions: event $A = \{$live Eastern Time Zone at home$\}$, $B = \{$senior$\}$,
$$P(A) = 0.63, P(B) = 0.41, P(A \cap B) = 0.31$$
2. Question: find $P(A \cup B)$.

Inclusion-Exclusion example

Suppose that for a randomly selected student in a probability class,

- P(live Eastern Time Zone at home) $= 63\%$.
- P(senior) $= 41\%$.
- P(live Eastern Time Zone at home and senior) $= 31\%$.

Find the probability that a student is either live Eastern Time Zone at home or a senior.

1. Conditions: event $A = \{$live Eastern Time Zone at home$\}$, $B = \{$senior$\}$,

$$P(A) = 0.63, P(B) = 0.41, P(A \cap B) = 0.31$$

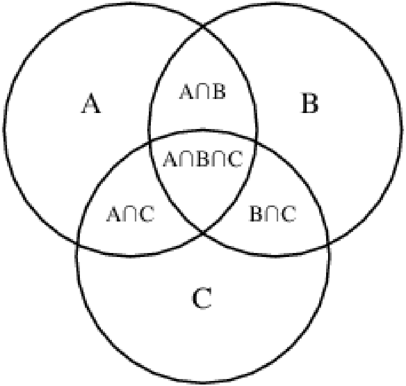
2. Question: find $P(A \cup B)$.
3. Formula: inclusion-exclusion

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.63 + 0.41 - 0.31 = 0.73$$

Propositions

☞ Inclusion-Exclusion: three events A, B, C (not necessarily disjoint)

$$\begin{aligned} P(A \cup B \cup C) =& P(A) + P(B) + P(C) - P(A \cap B) \\ &- P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) \end{aligned}$$



Recap

Three axioms of probability P

1. $0 \leq P(E) \leq 1$
2. $P(S) = 1$
3. $P\left(\bigcup\limits_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$, if $E_i \cap E_j = \emptyset$ for $i \neq j$

Recap

Three axioms of probability P

1. $0 \leq P(E) \leq 1$
2. $P(S) = 1$
3. $P\left(\bigcup\limits_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$, if $E_i \cap E_j = \emptyset$ for $i \neq j$

Propositions of probability

- $P(A^c) = 1 - P(A)$
- $P(B \cap A^c) = P(B) - P(A)$, if $A \subseteq B$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A \cup B \cup C) =$ $P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)$

Outline

Sample spaces with equally likely outcomes

- ☞ Suppose a sample space has N equally likely outcomes $\{1\}, \ldots, \{N\}$, then

  $$S = \{1\} \cup \ldots \cup \{N\}$$

  Disjointness gives

  $$1 = P(S) = P(\{1\}) + \ldots + P(\{N\}) = NP(\{i\}),$$

  so for each $1 \leq i \leq N$,

  $$P(\{i\}) = \frac{1}{N}$$

Sample spaces with equally likely outcomes

- Suppose a sample space has N equally likely outcomes $\{1\}, \ldots, \{N\}$, then

  $$S = \{1\} \cup \ldots \cup \{N\}$$

  Disjointness gives

  $$1 = P(S) = P(\{1\}) + \ldots + P(\{N\}) = NP(\{i\}),$$

  so for each $1 \leq i \leq N$,

  $$P(\{i\}) = \frac{1}{N}$$

- For event E in a sample space S with equally likely outcomes,

  $$P(E) = \frac{\#(E)}{\#(S)}$$

  Notation:

  Cardinality - $\#(E) =$ number of elements in set E

Probability of rolling an even number with a six sided die?

Probability of rolling an even number with a six sided die?

$$E = \{2, 4, 6\} \text{ and } S = \{1, 2, 3, 4, 5, 6\}$$

$$P(E) = 3/6 = 1/2$$

Probability of rolling an even number with a six sided die?

$$E = \{2, 4, 6\} \text{ and } S = \{1, 2, 3, 4, 5, 6\}$$

$$P(E) = 3/6 = 1/2$$

A couple has two kids, what is the probability that they are not both girls?

Probability of rolling an even number with a six sided die?

$$E = \{2, 4, 6\} \text{ and } S = \{1, 2, 3, 4, 5, 6\}$$
$$P(E) = 3/6 = 1/2$$

A couple has two kids, what is the probability that they are not both girls?

$$E = \{BB, GB, BG\} \text{ and } S = \{BB, GG, GB, BG\}$$
$$P(E) = 3/4$$

Example: birthday problem

Ignoring leap years, and assuming birthdays are equally likely to be any day of the year, what is the chance of no tie in birthdays among n students?

Example: birthday problem

Ignoring leap years, and assuming birthdays are equally likely to be any day of the year, what is the chance of no tie in birthdays among n students?

$$\begin{aligned}
\#(\text{birthdays of } n \text{ people}) &= 365^n \\
\#(\text{no match}) &= 365 \times 364 \times \cdots \times (365 - n + 1) \\
P(\text{no match}) &= \frac{365!}{(365 - n)!\, 365^n}
\end{aligned}$$

Example: birthday problem

Ignoring leap years, and assuming birthdays are equally likely to be any day of the year, what is the chance of no tie in birthdays among n students?

$$\#(\text{birthdays of } n \text{ people}) = 365^n$$

$$\#(\text{no match}) = 365 \times 364 \times \cdots \times (365 - n + 1)$$

$$P(\text{no match}) = \frac{365!}{(365-n)!\,365^n}$$