

Module 1

Matias Salibian Barrera and Daniel J. McDonald

22 Aug 2025

1 Welcome & “House rules”

- Weird schedule (Mondays and Fridays)
- Many opportunities to work and get help
- Weekly “routine”:
 - Class meetings: 3 hours
 - Pre-class reading: 1 to 2 hours
 - WebWork & assignments: 2 to 3 hours
 - Office hour
 - Total weekly hours: 7 to 9 hours
- Problem solutions are dangerous!!
- Online “solutions”: you must prove whatever you claim

2 Syllabus - Expectations - Grades

3 Basics of probability

- Formal treatment of randomness
- Key to this are models
- Defining a probability requires the following:
 - Random experiment
 - Sample space
 - Event
 - Rules to combine events (set operations)

3.1 Experiments

3.1.1 Definition

An action undertaken to make a discovery, test a hypothesis, or confirm a known fact.

3.1.2 Example

Release your pen from 4.9 meters above the ground.

3.1.3 Predicted outcomes

- The pen will fall to the ground.
- It will take about 1 sec to reach the ground.

3.2 Actual observations

- The pen did touch the ground
- Less sure if it took exactly 1 second to do so

3.3 Uncertainty

The outcome of some experiments cannot be determined beforehand.

Example

- Roll a die: which side will show?
- Draw a card from a well-shuffled deck: which one you will get?
- How many students will be in the classroom today?

3.4 Probability theory

- Even though die rolls are random, patterns emerge when we repeat the experiment many times.
- Probability Theory describes such patterns via mathematical models.
- It is a branch of Mathematics, and is based on a set of Axioms.
- Axioms: Statements or propositions accepted to hold true
- Theorems: Propositions which are established to hold true using sound logical reasoning.

3.5 Sample Space

Definition

Sample space is the set of all possible outcomes of a random experiment.

We denote it by Ω , and a generic outcome, also called sample point, by ω (i.e. $\omega \in \Omega$).

Example

- Roll a die: $\Omega = \{1, 2, 3, 4, 5, 6\} \subset \mathbb{N}$.
- Draw a card from a poker deck: $\Omega = \{2\spadesuit, 2\diamondsuit, \dots, A\clubsuit, A\heartsuit\}$
- Wind speed at YVR (km/h): $\Omega = [0, \infty) \subset \mathbb{R}$.
- Wait time for R4 at UBC (min): $\Omega = [0, 720) \subset \mathbb{R}$

3.6 Events

Definition

An event is a subset of the sample space Ω .

Notation: We commonly use upper case letters (A, B, C, \dots) for events.

Events are sets:

- $\omega \in A$ means “ ω is an element of A ”.
- $C \subset D$ means “ C is a subset of D ”.

3.7 Examples

Events are often formed by outcomes sharing some property. It's a good idea to practice listing explicitly the sample points of events described with words.

Example

- Roll a dice:
 - $A = \text{"roll an even number"} = \{2, 4, 6\}$
 - $B = \text{"roll a 3 or less"} = \{1, 2, 3\}$
 - $C = \text{"roll an even number no higher than 3"} = \{2\}$
- Bus wait time: $C = \text{"wait is less than half an hour"} = [15, 30]$
- Max-wind-speed: $A = \text{"wind is over 80 km/hour"} = (80, \infty)$

3.8 Set Operations

Suppose A, B are events (subsets of Ω).

- **Union:** $A \cup B$

$$\omega \in A \cup B \Leftrightarrow \omega \in A \text{ or } \omega \in B$$

- **Intersection:** $A \cap B$

$$\omega \in A \cap B \Leftrightarrow \omega \in A \text{ and } \omega \in B$$

- **Complement:** A^c

$$\omega \in A^c \Leftrightarrow \omega \notin A$$

- **Symmetric difference:** $A \triangle B$

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

3.9 Properties of set operations

- Commutative:

$$- A \cup B = B \cup A$$

$$- A \cap B = B \cap A$$

- Associative:

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

- Distributive:

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

- $A = B \Leftrightarrow A \subseteq B$ and $B \subseteq A$

3.10 De Morgan's Laws

Theorem

For any two events (sets) A and B , we have

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

To prove the theorem it is sufficient to show that

$$(A \cup B)^c \subseteq A^c \cap B^c$$

and that

$$A^c \cap B^c \subseteq (A \cup B)^c$$

3.11 Proof of De Morgan's Laws

Proof

Let us show $(A \cup B)^c \subseteq A^c \cap B^c$:

1. Take any $\omega \in (A \cup B)^c$, then we must have $\omega \notin A \cup B$;
2. This implies that $\omega \notin A$ and $\omega \notin B$ (because if either $\omega \in A$ or $\omega \in B$ then we'd have $\omega \in A \cup B$);
3. Hence: $\omega \in A^c$ and $\omega \in B^c$;
4. Which is the same as $\omega \in A^c \cap B^c$;
5. Thus, we showed that $\omega \in (A \cup B)^c$ implies $\omega \in A^c \cap B^c$. In other words, that

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

We leave it to you to prove that $A^c \cap B^c \subseteq (A \cup B)^c$.

3.12 Some Useful Identities

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

Proof

- First: $A \subseteq (A \cap B) \cup (A \cap B^c)$:
 1. Take $\omega \in A$, then either $\omega \in B$ or $\omega \notin B$.
 2. In the first case: $\omega \in A \cap B$. In the second case: $\omega \in A \cap B^c$.
 3. Thus $\omega \in (A \cap B) \cup (A \cap B^c)$.
- Also: $(A \cap B) \cup (A \cap B^c) \subseteq A$:
 1. Take $\omega \in (A \cap B) \cup (A \cap B^c)$.
 2. If $\omega \in (A \cap B)$ then $\omega \in A$, and if $\omega \in (A \cap B^c)$ then $\omega \in A$.
 3. Thus, we always have $\omega \in A$.

- $A \cup B = A \cup (B \cap A^c)$
 - Prove it!

3.13 Power set

The power set of Ω (denoted 2^Ω) is the set of all possible subsets of Ω .

For example, if $\Omega = \{1, 2, 3\}$ then:

$$2^\Omega = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$$

- The symbol \emptyset denotes the empty set.
- If Ω has n elements, then 2^Ω has 2^n elements. In symbols:

$$\#(2^\Omega) = 2^{\#\Omega}$$

How can we prove this when $\#\Omega = n \in \mathbb{N}$?

3.14 Counting / listing

List the n elements of Ω as $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$.

Any event $A \subseteq \Omega$ can be uniquely represented by a sequence of 0's and 1's:

$$x_i = \begin{cases} 1 & \text{if } \omega_i \in A, \\ 0 & \text{if } \omega_i \notin A. \end{cases} \quad 1 \leq i \leq n.$$

To each event A we associate the string (x_1, x_2, \dots, x_n) :

$$A \longrightarrow (x_1, x_2, \dots, x_n)$$

And, each sequence corresponds to an event: $(0, 1, 1, \dots, 0, 1) \longrightarrow B$.

There are 2^n distinct such n -long sequences. Each of them corresponds uniquely to one event, and every event corresponds to one such sequence. Hence there are the same number of events as there are sequences (2^n).

3.15 Counting / listing

Example

How many license plates can you create using a triplet of one letter, one digit, and one letter (e.g. R2A)?

4 Probability

4.1 Probability of an event

- Even though random outcomes cannot be predicted, in some cases we have an idea about the chance that an outcome occurs.
 - If you toss a fair coin, the chance of observing a head is the same as that of observing a tail.
 - If you buy a lottery ticket, the chance of winning is very small.
- A probability function \mathbb{P} quantifies these chances.
- Probability functions are computed on events $A \in \mathcal{B}$. We calculate $\mathbb{P}(A)$. Mathematically / formally, we have:

$$\mathbb{P} : \mathcal{B} \rightarrow [0, 1]$$

where \mathcal{B} is a collection of possible events.

- Probability functions need to do this “coherently”

4.2 Probability Axioms

Let Ω be a sample space and \mathcal{B} be a collection of events (i.e. subsets of Ω).

Definition

A probability function is a function \mathbb{P} with domain B such that

1. **Axiom 1:** $\mathbb{P}(\Omega) = 1$;
2. **Axiom 2:** $\mathbb{P}(A) \geq 0$ for any $A \in B$;
3. **Axiom 3:** If $\{A_n\}_{n \geq 1}$ is a sequence of **disjoint** events, then

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

Note: $\{A_n\}_{n \geq 1}$ is a sequence of **disjoint** events when $A_i \cap A_j = \emptyset$ if $i \neq j$

4.3 Probability

- Kolmogorov showed how one can construct such functions, and that a probability function only needs to satisfy those three properties to be a “coherent” probability function
- In other words: every desirable property of a probability \mathbb{P} can be shown to hold using only Axioms 1, 2, and 3 (and logic).
- Alternatively: any function \mathbb{P} that satisfies Axioms 1, 2, and 3 is a “proper” probability function.

4.4 Properties of the probability function

In general, A, B, C , etc. denote arbitrary events. Ω is the sample space.

- Probability of the complement:

$$\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$$

- Monotonicity:

$$A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$$

- Probability of the union:

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

- Boole’s inequality:

$$\mathbb{P}\left(\bigcup_{i=1}^m A_i\right) \leq \sum_{i=1}^m \mathbb{P}(A_i)$$

4.5 Proof of formula for complement

We need to show that if \mathbb{P} satisfies Axioms 1, 2, and 3, and A is an arbitrary event, then necessarily $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

Proof

1. $\Omega = A \cup A^c$ and A and A^c are disjoint.

2. Using Axiom 1

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

3. Axiom 3 says that $\mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c)$

4. Putting the last 2 statements together we get $1 = \mathbb{P}(\Omega) = \mathbb{P}(A) + \mathbb{P}(A^c)$.

Since $\mathbb{P}(A)$ and $\mathbb{P}(A^c)$ are just numbers, simple arithmetic gives

$$1 - \mathbb{P}(A) = \mathbb{P}(A^c)$$

4.6 Monotonicity

4.6.1 $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$

Proof

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

2. Since we know that $A \subset B$, then $B \cap A = A$ (prove it!). Hence the equation above is

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

3. Since A and $B \cap A^c$ are disjoint (why?), Axiom 3 says $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c)$

4. Axiom 2 ensures that $\mathbb{P}(B \cap A^c) \geq 0$. Thus, the last equation satisfies

$$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c) \geq \mathbb{P}(A)$$

4.7 Probability of a union of two events

4.7.1 $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

Proof

1. First prove that $A \cup B = A \cup (B \cap A^c)$. Since A and $B \cap A^c$ are disjoint events, Axiom 3 implies that

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c)$$

2. “Slice” B into the part it shares with A and the rest: $B = (B \cap A) \cup (B \cap A^c)$, where the events on the right-hand side are disjoint. Apply Axiom 3 to get $\mathbb{P}(B) = \mathbb{P}(B \cap A) + \mathbb{P}(B \cap A^c)$
3. The last equation can be re-written as $\mathbb{P}(B \cap A^c) = \mathbb{P}(B) - \mathbb{P}(B \cap A)$
4. Replace this expression for $\mathbb{P}(B \cap A^c)$ in 2. to get

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(B \cap A)$$

4.8 Probability of a union of two events

- Ven diagrams are not proofs. They are sometimes useful, but a diagram is not a proof or a logical argument.
- This is not a proof:

4.9 Proof of Boole’s inequality

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

Proof

- We will prove it by induction.
- For $n = 2$:

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$$

$$\leq \mathbb{P}(A_1) + \mathbb{P}(A_2)$$

because $\mathbb{P}(A_1 \cap A_2) \geq 0$

4.10 Boole's inequality (continued)

Proof

- Assume now that it holds for n . We then have

$$\begin{aligned}\mathbb{P}\left(\bigcup_{i=1}^{n+1} A_i\right) &= \mathbb{P}\left[\left(\bigcup_{i=1}^n A_i\right) \cup A_{n+1}\right] && \text{associative prop.} \\ &\leq \mathbb{P}\left[\bigcup_{i=1}^n A_i\right] + \mathbb{P}(A_{n+1}) && \text{case } n=2 \\ &\leq \sum_{i=1}^n \mathbb{P}(A_i) + \mathbb{P}(A_{n+1}) && \text{induction} \\ &= \sum_{i=1}^{n+1} \mathbb{P}(A_i).\end{aligned}$$

4.11 Example of applying these formulas

John borrows 2 books.

Suppose there is a 0.5 probability he likes the first book, 0.4 that he likes the second book, and 0.3 that he likes both.

What is the probability that he will **NOT** like either of the 2 books?

4.12 Solution

Introduce two notations for events:

$$A = \{ \text{John likes book 1} \}, B = \{ \text{John likes book 2} \}.$$

We are told that $\mathbb{P}(A) = 0.5$, $\mathbb{P}(B) = 0.4$, $\mathbb{P}(A \cap B) = 0.3$.

We are asked to calculate $\mathbb{P}(A^c \cap B^c)$.

Note that $A^c \cap B^c = (A \cup B)^c$. Hence,

$$\mathbb{P}(A^c \cap B^c) = \mathbb{P}[(A \cup B)^c] = 1 - \mathbb{P}(A \cup B).$$

Making use of another formula:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = 0.6.$$

We get $\mathbb{P}(A^c \cap B^c) = 1 - 0.6 = 0.4$.

General line of approach:

Connect $A^c \cap B^c$ with events A , B and $A \cap B$, because the probabilities of latter are provided.

4.13 Example of applying formulas

Jane must take two tests, call them T_1 and T_2 .

Suppose the probability she passes test T_1 is **0.8**, test T_2 is **0.7** and both tests is **0.6**.

Calculate the probability that:

- (a) She passes at least one test.
- (b) She passes at most one test.
- (c) She fails both tests.
- (d) She passes only one test.

4.14 Solutions

Notation:

$$A = \{\text{Jane passes test } T_1\}$$

$$B = \{\text{Jane passes test } T_2\}$$

We are told that

$$\mathbb{P}(A) = 0.80, \mathbb{P}(B) = 0.70, \text{ and } \mathbb{P}(A \cap B) = 0.6.$$

Our task is to use these information to figure out the probabilities of various events.

4.15 Solution to (a) Jane passes at least one test

We note that

$$\{\text{Passes at least one test}\} = A \cup B.$$

Hence,

$$\begin{aligned}\mathbb{P}(A \cup B) &= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \\ &= 0.80 + 0.70 - 0.60 \\ &= 0.90.\end{aligned}$$

The probability that Jane passes at least 1 test is 0.9.

4.16 Solution to (b) Jane passes at most one test

We notice that

$$\begin{aligned}\{\text{Passes at most one test}\} &= \{\text{Fails at least one test}\} \\ &= A^c \cup B^c.\end{aligned}$$

And we notice

$$A^c \cup B^c = (A \cap B)^c \quad \text{De Morgan rule}$$

which gives one way of computing its probability.

$$\begin{aligned}\mathbb{P}(\{\text{Passes at most one test}\}) &= \mathbb{P}((A \cap B)^c) \\ &= 1 - \mathbb{P}(A \cap B) \\ &= 1 - 0.60 = 0.40.\end{aligned}$$

The probability that Jane passes at most 1 test is 0.4.

4.17 Solution to (c) She fails both tests

Similarly, we observe by De Morgan rule,

$$\begin{aligned}\{\text{Fails both tests}\} &= A^c \cap B^c \\ &= (A \cup B)^c\end{aligned}$$

which gives one way of computing its probability.

$$\begin{aligned}
\mathbb{P}(\{\text{Fails both tests}\}) &= \mathbb{P}((A \cup B)^c) \\
&= 1 - \mathbb{P}(A \cup B) \\
&= 1 - 0.90 \\
&= 0.10 \text{ from Part (a)}.
\end{aligned}$$

Therefore, the probability that Jane fails both tests is 0.1.

4.18 Solution to (d) She passes only one test

We first decompose the event into the union of two disjoint events:

$$\{\text{Passes only one test}\} = (A \cap B^c) \cup (A^c \cap B)$$

We hence have

$$\begin{aligned}
\mathbb{P}(\{\text{Passes only one test}\}) &= \mathbb{P}(A \cap B^c) + \mathbb{P}(A^c \cap B) \\
&= [\mathbb{P}(A) - \mathbb{P}(A \cap B)] + [\mathbb{P}(B) - \mathbb{P}(A \cap B)] \\
&= \mathbb{P}(A) + \mathbb{P}(B) - 2 \times \mathbb{P}(A \cap B) \\
&= 0.80 + 0.70 - 2 \times 0.60 = 0.30.
\end{aligned}$$

The probability that Jane passes only 1 test is 0.3.

4.19 Remarks

- It is typical in introductory probability theory course to give a story first, followed by specifying events verbally.
- In these cases, your answer should be a sentence, not just with the context included.
- The best mathematical approach is to **define** some events, and express other events of interest using those events.
- Rather than relying on algebra as in our examples, it can be easier to use Venn diagram to have these events connected.
- There can be many ways to connect them, all lead to a viable probability calculation.

4.20 Exercise without a story

Exercise 2

(a) Suppose that $\mathbb{P}(A) = 0.85$ and $\mathbb{P}(B) = 0.75$. Show that

$$\mathbb{P}(A \cap B) \geq 0.60.$$

(b) More generally, prove the Bonferroni inequality:

$$\mathbb{P}(\cap_{i=1}^n A_i) \geq \sum_{i=1}^n \mathbb{P}(A_i) - (n-1).$$

4.21 Solution to 2 (a)

We know that $\mathbb{P}(A) = 0.85$ and $\mathbb{P}(B) = 0.75$ and wish to show that

$$\mathbb{P}(A \cap B) \geq 0.60.$$

Note that $A \cap B = (A^c \cup B^c)^c$ (De Morgan rule)

Hence:

$$\mathbb{P}(A \cap B) = 1 - \mathbb{P}(A^c \cup B^c)$$

Also

$$\mathbb{P}(A^c \cup B^c) \leq \mathbb{P}(A^c) + \mathbb{P}(B^c)$$

Therefore

$$\begin{aligned} \mathbb{P}(A \cap B) &= \mathbb{P}((A^c \cup B^c)^c) \\ &= 1 - \mathbb{P}(A^c \cup B^c) \\ &\geq 1 - [\mathbb{P}(A^c) + \mathbb{P}(B^c)] \\ &= 1 - [1 - \mathbb{P}(A)] - [1 - \mathbb{P}(B)] \\ &= \mathbb{P}(A) + \mathbb{P}(B) - 1 \\ &= 0.85 + 0.75 - 1 = 0.60. \end{aligned}$$

4.22 Solution to Example 2 (b) Prove the Bonferroni inequality

$$\mathbb{P} \left(\bigcap_{i=1}^n A_i \right) \geq \sum_{i=1}^n \mathbb{P}(A_i) - (n-1).$$

First note that

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

Therefore

$$\begin{aligned} \mathbb{P} \left(\bigcap_{i=1}^n A_i \right) &= \mathbb{P} \left(\left(\bigcup_{i=1}^n A_i^c \right)^c \right) \\ &= 1 - \mathbb{P} \left(\bigcup_{i=1}^n A_i^c \right) \\ &\geq 1 - \sum_{i=1}^n \mathbb{P}(A_i^c) \\ &= 1 - \sum_{i=1}^n (1 - \mathbb{P}(A_i)) && \text{complement rule} \\ &= 1 - \sum_{i=1}^n 1 + \sum_{i=1}^n \mathbb{P}(A_i) = 1 - n + \sum_{i=1}^n \mathbb{P}(A_i) \\ &= \sum_{i=1}^n \mathbb{P}(A_i) - (n-1). \end{aligned}$$

4.23 Refresh your memory of the probability axioms

Kolmogorov's axioms tell us that if we have (a) sample space, (b) collection of events, and wish to create a probability function $\mathbb{P}(\cdot)$; what properties this function should have.

The last a few examples show that if such a function $\mathbb{P}(\cdot)$ has been given, how one may derive the value of $\mathbb{P}(A)$ from the probabilities of other events.

We next suggest a way to set up (a) sample space, (b) collection of events, and (c) probability function $\mathbb{P}(\cdot)$ for a special type of experiments.

5 Building probability functions

5.1 Equally likely outcomes

Suppose the sample space Ω is finite.

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$$

In many applications, we trust that these distinct outcomes are equally likely so that we wish to make

$$\mathbb{P}(\{\omega_i\}) = a, \quad a \in (0, 1].$$

For such an experiment, we get a natural probability function.

5.2 Equally likely outcomes

Notice that $\Omega = \bigcup_{i=1}^n \{\omega_i\}$, and $\{\omega_i\}_{i=1}^n$ are disjoint events.

The probability theory Axioms require

$$1 = \mathbb{P}(\Omega) = \mathbb{P}\left(\bigcup_{i=1}^n \{\omega_i\}\right) = \sum_{i=1}^n \mathbb{P}(\{\omega_i\}) = \sum_{i=1}^n a = na.$$

Our only option is to let $a = 1/n$.

5.3 Equally likely outcomes

The axioms further require that for any event $A \subseteq \Omega$:

$$\begin{aligned} \mathbb{P}(A) &= \sum_{\omega_i \in A} \mathbb{P}(\{\omega_i\}) = \sum_{\omega_i \in A} \frac{1}{n} = \frac{\#A}{\#\Omega} \\ &= \frac{\#\{\text{favorable outcomes}\}}{\#\{\text{possible outcomes}\}} \end{aligned}$$

Being “favourable” means those in the event of interest: A in this example.

5.4 Probability calculation examples

- To calculate probabilities, we count the number of sample points in the event of interest.
- Counting the number of sample points in a set can be mathematically surprisingly complicated.
- **Combinatorial theory** deals with this problem.
- We will learn some basic combinatorial rules and techniques for this purpose.

5.5 Counting Example I

A die is rolled repeatedly until we see an outcome being 6.

- (a) Specify/describe the sample space.
- (b) Let E_n denote the event that the number of rolls is exactly n ($n = 1, 2, \dots$). Describe the event E_n .
- (c) Describe the event $E_1 \cup E_2$ and $(\bigcup_{n=1}^{\infty} E_n)^c$.

5.6 Solution to (a) Describe the sample space

The sample space consists of all sequences

$$(x_1, x_2, \dots, x_{n-1}, x_n)$$

with

$$1 \leq x_i \leq 5 \text{ for } 1 \leq i \leq n-1, \text{ and } x_n = 6.$$

For example:

$$\begin{aligned} (4, 1, 5, 1, 6) & \text{ with } n = 5 \\ (5, 2, 6) & \text{ with } n = 3 \\ (6) & \text{ with } n = 1 \\ (2, 4, 5, 1, 3, 4, 6) & \text{ with } n = 7 \end{aligned}$$

5.7 Solution to (b) Describe the event...rolls is n

$$\begin{aligned} E_1 &= \{(6)\} \\ E_2 &= \{(x_1, 6) : 1 \leq x_1 \leq 5\} \\ E_3 &= \{(x_1, x_2, 6) : 1 \leq x_1, x_2 \leq 5\} \\ &\vdots \\ E_n &= \{(x_1, x_2, \dots, x_{n-1}, 6) : 1 \leq x_1, x_2, \dots, x_{n-1} \leq 5\} \end{aligned}$$

5.8 Solution (c) $E_1 \cup E_2$ and $(\bigcup_{n=1}^{\infty} E_n)^c$

Verbally interpreting $E_1 \cup E_2$, $\bigcup_{n=1}^{\infty} E_n$, and $(\bigcup_{n=1}^{\infty} E_n)^c$:

Suggestive answers:

$$\begin{aligned} E_1 \cup E_2 &= \left\{ 6 \text{ appears before the } 3^{rd} \text{ roll} \right\} \\ \bigcup_{n=1}^{\infty} E_n &= \left\{ 6 \text{ eventually appears} \right\} \\ \left(\bigcup_{n=1}^{\infty} E_n \right)^c &= \left\{ 6 \text{ never appears} \right\} \end{aligned}$$

5.9 Counting example II

A system has **5 components**, which can either be **working** or have **failed**.

The experiment consists of observing the current status (W/F) of the 5 components.

- (a) Describe the sample space for this experiment.
- (b) How much is $\#\Omega$?
- (c) Suppose that the system will work if either components (1 and 2), or (3 and 4), or (1, 3 and 5) are working.

List the outcomes in the event $D = \{\text{The system works}\}$?

- (d) Let $A = \{\text{components 4 and 5 have failed}\}$. How much is $\#A$?
- (e) List the outcomes in $A \cap D$.

5.10 Answer to (a) Describe the sample space

The outcomes of the experiment are sequences

$$(x_1, x_2, x_3, x_4, x_5)$$

where we could have $x_i = W$ or $x_i = F$ depending on when component i is working or has failed.

One typical outcome is (W, W, F, W, F) which indicates that components **1**, **2** and **4** are working and **3** and **5** have failed,

5.11 Answer to (b) How many outcomes are there in Ω ?

NOTE: the number of elements of a set A (aka its cardinal number) is denoted by $\#A$. or $|A|$ or (sometimes) $\|A\|$

We have $\#\Omega = 2^5 = 32$.

5.12 Answer to (c) When does it work?

(c) The system works if (1 and 2), or (3 and 4) or (1, 3 and 5) work.

Let $D = \{\text{The system works}\}$. Let us count it.

Here is an exhaustive list of D :

1 and 2 work	3 and 4 work	1, 3 and 5 work
(W, W, W, W, W)	(F, F, W, W, W)	(W, F, W, F, W)
(W, W, F, W, W)	(F, W, W, W, W)	
(W, W, W, F, W)	(W, F, W, W, W)	
(W, W, W, W, F)		
(W, W, W, F, F)	(F, F, W, W, F)	
(W, W, F, W, F)	(F, W, W, W, F)	
(W, W, F, F, W)	(W, F, W, W, F)	
(W, W, F, F, F)		

Apparently, $\#D = 15$.

5.13 Answer to (d) How big is $|A|$?

(d) $A = \{\text{4 and 5 have failed}\}$. How many outcomes are there in A ?

The outcomes in A are of the form (x_1, x_2, x_3, F, F) with $x_i = W$ or $x_i = F$ for $i = 1, 2, 3$.

Hence, we have $2^3 = 8$ different outcomes:

$$\#A = 2^3 = 8.$$

5.14 Answer to (e) Describe all the outcomes in $A \cap D$

Recall the list of D :

1 and 2 work	3 and 4 work	1, 3 and 5 work
(W, W, W, W, W)	(F, F, W, W, W)	(W, F, W, F, W)
(W, W, F, W, W)	(F, W, W, W, W)	
(W, W, W, F, W)	(W, F, W, W, W)	
(W, W, W, W, F)		
(W, W, W, F, F)	(F, F, W, W, F)	
(W, W, F, W, F)	(F, W, W, W, F)	
(W, W, F, F, W)	(W, F, W, W, F)	
(W, W, F, F, F)		

Clearly, we have

$$W \cap A = \{(W, W, W, F, F), (W, W, F, F, F)\}$$

and $\#(W \cap A) = 2$.

5.15 Example III

Two dice have two sides painted red, two painted black, one painted yellow, and the other painted white.

When this pair of dice is rolled, what is the probability that both dice show the same color facing up?

Remark: If not explicitly declared, this type of problem assumes “equal likely outcomes”.

5.16 Answer to Example III

We answer this question with brute-force counting.

An exhaustive list of sample space is presented as follows:

	R1	R2	B3	B4	Y5	W6
R1	X	X				
R2	X	X				
B3			X	X		
B4			X	X		
Y5					X	
W6						X

The size of sample space is $\#\Omega = 36$.

The number of favourable outcomes is 10.

Hence, under the equal probability model:

$$\mathbb{P}\{\text{Same color}\} = 10/36 = 5/18.$$

5.17 Example IV

A small community consists of 20 families.

Four of them have 1 child, 8 have 2 children, 5 have 3 children, 2 have 4 children, and 1 has 5 children.

- (a) What is the probability that a randomly chosen family has i children, for each $1 \leq i \leq 5$?
- (b) What is the probability that a randomly chosen child comes from a family with i children, for each $1 \leq i \leq 5$?

Jargon: By “randomly chosen”, it says that the **specific unit** is equally likely selected.

Applying the equal probability experiment/model for (a): family; (b): child.

5.18 Answer to Example IV

We organize the information as follows:

i	Families with i children	Children from families w/ i children
1	4	4
2	8	16
3	5	15
4	2	8
5	1	5
Total	20	48

Catch: There are 20 families, 48 children in this community.

5.19 Answer to (a)

If one family is chosen at random, what is the probability it has i children, $i = 1, 2, 3, 4, 5$?

i	Families with i children	$\mathbb{P}(\text{family has } i \text{ children})$
1	4	4/20
2	8	8/20
3	5	5/20
4	2	2/20
5	1	1/20
Total	20	1.00

5.20 Answer to (b)

If one child is randomly chosen, what is the probability that it comes from a family having i children, $i = 1, 2, 3, 4, 5$?

i	Children from families with i children	$\mathbb{P}(\text{child comes from family with } i \text{ children})$
1	4	4/48
2	16	16/48
3	15	15/48
4	8	8/48
5	5	5/48
Total	48	1.00

Remark Part (a) and Part (b) are probability calculations under **Two Different** experiments.