Honors Statistical Inference I - APMA 1655

Milan Capoor

Spring 2023

1 Lecture 1, Jan 25: Random outcomes & sample spaces

Features of random events:

- There is more than one possible outcome.
- Before doing or observing the experiment of interest, you do not know which outcome you will see.
- Some possible outcomes are likely, and some other outcomes are quite unlikely e.g., winning one billion dollars by buying a lottery ticket.

Sample Spaces

Sample space: the set of all possible outcomes or results of that experiment (usually denoted Ω) Examples:

- Coin toss ($\Omega = \{H, T\}$)
- Schrodinger's cat ($\Omega = \{\text{alive}, \text{dead}\}$)
- The lifespan of a tree ($\Omega = \{1, 2, ...100, ...\} = \mathbb{Z}_+$)

Also note that not all elements/subsets of Ω are equally likely - hence, "probability"

2 Lecture 2, Jan 27: Events, event operations, and infinite operations

Suppose Ω is a sample space. Then:

- *Event:* each subset E of Ω
- *Impossible event:* the empty set $\emptyset$

Example: Final exam scores

1. The sample space: $\Omega = \{0, 1, 2, ..., 100\}$
2. The event "score is higher than 50": $E = \{51, 52, ..., 100\} \subset \Omega$
3. The event "score is a negative number": $\emptyset$

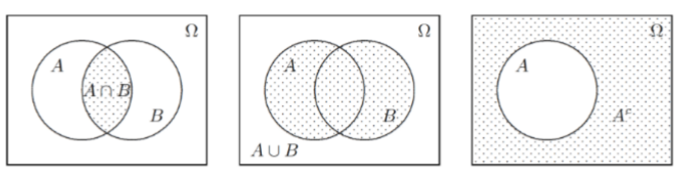
Set/Event Operations

Suppose Ω is a sample space and A and B are events $\{\omega \in \Omega : \omega \in A \text{ and } \omega \in B\}$:

- *Intersection:* both A and B occur; the collection of elements that are in sets A AND B
$$A \cap B$$
- *Union:* either A or B occurs; the collection of elements in A or B
$$A \cup B$$
- *Complement:* the collection of elements that are not in A; the opposite event of A
$$A^c$$

Note that

$$\Omega^c = \emptyset$$
$$\emptyset^c = \Omega$$



Intersection $A \cap B$ Union $A \cup B$ Complement A^c

De Morgan's Laws: for any two events A and B we have the following

$$(A \cup B)^c = A^c \cap B^c \tag{1}$$
$$(A \cap B)^c = A^c \cup B^c \tag{2}$$

Infinite Sets

Suppose $A_1, A_2, A_3, ...A_n, A_{n+1}, ...$ are events. Some of them may be identical and some of them may be empty

Infinite Operations:

- *Infinite intersection:* the collection of events that are in ALL the sets $A_1, ..., A_n$; i.e. "all the events A_n for $n = 1, 2, ...$ happen"

$$\bigcap_{n=1}^{\infty} A_n = \{\omega \in \Omega : \omega \in A_n \forall n = 1, 2, 3, ...\}$$

- *Infinite union:* the collection of elements in at least one of the sets; "at least one of these events happen"

$$\bigcup_{n=1}^{\infty} A_n = \{\omega \in \Omega : \exists n' | \omega \in A_{n'}\}$$

  ("there exists at least one n' such that omega is in the set)

3 Lecture 3, Jan 30: Probability space & properties of probability

Disjoint: for two events A and B, they are disjoint if $A \cap B = \emptyset$

Mutually disjoint: if all pairwise intersections of $A_1, A_2, ..., A_n$ are empty ($A_n \cup A_m = \emptyset$ if $n \neq m$)

Definition of Probability: from the following definition we can derive everything in probability theory.

Let Ω be a sample space. Suppose $\mathbb{P}$ is a real-valued function of subsets of Ω

$$\mathbb{P} : \{\text{subsets of } \Omega\} \rightarrow \mathbb{R}, \quad A \mapsto \mathbb{P}\{A\}$$

where A is an input and $\mathbb{P}A$ is the corresponding output. If $\mathbb{P}$ satisfies the following three axioms, the pair $(\Omega, \mathbb{P})$ is a *probability space*

1. $\mathbb{P}(A) \geq 0$ for any subset $A \subset \Omega$ (the probability of an event must be non-negative)

2. $\mathbb{P}(\Omega) = 1$
3. For any sequence of disjoint subsets $\{A_i\}_{i=1}^{\infty}$ (i.e. $A_i \cap A_j) = \emptyset$ we have

$$\mathbb{P}\{\bigcup_{i=1}^{\infty} A_i\} = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

The map $\mathbb{P}$ is called a *probability*

We can define this as a specific function

$$\mathbb{P}(A) := \frac{\#A}{n} \quad A \subset \Omega$$

where #A is the number of elements in A and $\Omega = \{1, 2, ..., n\}$ with a large n.

Lecture 4, Feb 1: Properties of Probability

Let $(\Omega, \mathbb{P})$ be a probability space. Then,

1. $\mathbb{P}(\emptyset) = 0$

   Note: while this implies that the probability of an impossible even is 0, there can be zero-probability events which are not themselves impossible
2. if two events E_1 and E_2 satisfy $E_1 \cap E_2 = \emptyset$, then

$$\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2)$$

3. if $A, B \subset \Omega$ and $A \subset B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$

   (Intuitively, if A happens, B must also happen so B is more likely)
4. $0 \leq \mathbb{P}(A) \leq 1$ for $A \subset \Omega$
5. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
6. for any $A, B \subset \Omega$

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

7. for any countable collection of subsets

$$\mathbb{P}\{\bigcup_{n=1}^{\infty} A_n\} \leq \sum_{n=1}^{\infty} \mathbb{P}\{A_n\}$$

   Note: the equality is obvious from axiom three in the case where all events are mutually disjoint. in the case of intersections, though, rule 6 must be generalized to account for overlap, hence the less than or equal to

4 Lecture 5, Feb 3: Conditional Probability

Part I - Motivating Problem

We know that a family has two children.

$$\Omega = \{(g,g),(b,b),(g,b),(b,g)\}$$
$$\mathbb{P}(A) = \frac{\#A}{\#\Omega} = \frac{\#A}{4} \quad A \subset \Omega$$

Event 1: $A = \{(g,g),(b,g),(g,b)\}$ ("at least one is girl")

$$\mathbb{P}(A) = \frac{3}{4}$$

Now suppose we get further information that the family has at least one boy:

$B = \{(b,g),(b,b),(g,b)\}$:

$$\mathbb{P}(A|B) = \frac{\#(A \cap B)}{\#B} = \frac{\#\{(b,g),(g,b)\}}{\#\{(b,g),(b,b),(g,b)\}} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{1}{2}$$

("knowing that event B occurs, what is the updated likelihood of A?")

Part II - A more rigorous definition

Let $A, B \subset \Omega$ such that

1. if $\mathbb{P}(B) > 0$ we call the following "the *conditional probability of A given B*"

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

2. otherwise, if $\mathbb{P}(B) = 0$ then $\mathbb{P}(A|B)$ is not well defined in the scope of this course (see real analysis)

Theorem: Let $(\Omega, \mathbb{P})$ be a probability space. Suppose $B \subset \Omega$ and $\mathbb{P}(B) > 0$. Then let

$$\tilde{\mathbb{P}}(A) := \mathbb{P}(A|B) \quad \forall A \subset \Omega$$

Then, $\tilde{\mathbb{P}}$ is another probability defined on Ω such that $\tilde{\mathbb{P}}$ also satisfies the 3 axioms

Part III - Properties of Conditional Probabilities

Assuming $\mathbb{P}(B) > 0$, and $A, B \subset \Omega$:

1. Multiplication Law

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

2. Let $B_1, B_2, ..., B_n \subset \Omega$. We say the events provide a *partition* of Ω if they are mutually disjoint and they satisfy[1]

$$\bigcup_{i=1}^{n} B_i = \Omega$$

3. *The law of total probability*

   Let $B_1, B_2, ..., B_n$ be events and they provide a partition of Ω and $\mathbb{P}(B_i) > 0$. Then,

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)$$

   Proof:

$$\Omega = \bigcup_{i=1}^{n} B_o \implies A = A \cap \Omega$$

   Then, by the laws above, $A \cap B_1, A \cap B_2, ...A \cap B_n$ is mutually disjoint. So

$$\mathbb{P}(A) = \mathbb{P}(\bigcup_{i=1}^{n} A \cap B_i) = \sum_{i=1}^{n} \mathbb{P}(A \cap B_i)$$

[1]Note that for $B_i = \{B_1, ...B_n\}$,

$$A \cap \left(\bigcup_{i=1}^{n} B_i\right) = A \cap (B_1 \cup B_2 \cup ... \cup B_n) = \bigcup_{i=1}^{n} A \cap B_i$$

Then from the multiplication law, $\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)$ and the desired result follows

4. Corollary: Let B be an event with $0 < \mathbb{P}(B) < 1$. Then we have

$$\mathbb{P}(A) = \mathbb{P}(A|B) \cdot \mathbb{P}(B) + \mathbb{P}(A|B^c) \cdot \mathbb{P}(B^c)$$

   Proof: Let $B_1 = B,\ B_2 = B^c,\ n = 2$. B_1 and B_2 obviously partition Ω and

$$P(B_1) = \mathbb{P}(B) > 0, \quad \mathbb{P}(B_2) = \mathbb{P}(B^c) = 1 - \mathbb{P}(B) > 0$$

   Then the corollary follows from the law of total probability

5 Lecture 6, Feb 6: Bayes' formula

Part I - Bayes' Rule

Suppose $B_1, ..., B_n$ provide a partition of Ω. In addition, $\mathbb{P}(B_i) > 0 \quad \forall i \in [1, n]$. Let A be any event such that $\mathbb{P}(A) > 0$. Then,

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)}{\sum_{j=1}^{n} \mathbb{P}(A|B_j) \cdot \mathbb{P}(B_j)} \quad i = 1, 2, ..., n$$

Proof: The definition of conditional probability implies

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

The multiplication law implies

$$\mathbb{P}(A \cap B_i) = \mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)$$

The law of total probability implies

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)$$

Then the desired result follows.

Though this proof is trivial, the formula is quite meaningful in that it allows an exchange between the conditions and results. Note that this is not the general Bayes formula.

6 Lecture 7, Feb 8: Independence

Independent events do not affect each other's outcomes (e.g. flipping two coins).That is

$$\begin{cases} \mathbb{P}(A|B) = \mathbb{P}(A) \\ \mathbb{P}(B|A) = \mathbb{P}(B) \end{cases}$$

In other words, knowing A does not help predict B.

Together with those equations and the multiplication law, we have

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

(for independent events)

Definition: Let $(\Omega, \mathbb{P})$ be a probability space.

1. Suppose A and B are two events. They are *independent* if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

2. Suppose $A_1, A_2, ...A_n$ are a sequence of events. The sequence is *mutually independent* if

$$\mathbb{P}(A_m \cap A_n) = \mathbb{P}(A_m) \cdot \mathbb{P}(A_n) \quad m \neq n$$

Theorem: Let $(\Omega, \mathbb{P})$ be a probability space with $A, B \subset \Omega$ such that $\mathbb{P}(A) > 0$ and $\mathbb{P}(B) > 0$ (this is necessary to have the conditional probabilities well-defined). Then the following three equations are equivalent:

1. $\mathbb{P}(A|B) = \mathbb{P}(A)$
2. $\mathbb{P}(B|A) = \mathbb{P}(B)$
3. $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$

Note that we choose this third equation to be the definition of independence stated above specifically because it does not defend on the positive probability condition of the other two

Example: Suppose a family has 3 children of unknown gender

$$\Omega = \{(g,g,g),(g,g,b),(g,b,g),(b,g,g),(g,b,b),(b,g,b),(b,b,g),(b,b,b)\}$$

$$\#\Omega = 2^3 = 8 \Longrightarrow \mathbb{P}(A) := \frac{\#A}{8}$$

Consider the events A ("the family has boys and girls") and B ("the family has at most one girl")

Question: Are A and B independent?

$$\begin{aligned}
\mathbb{P}(A) &= \frac{6}{8} \\
\mathbb{P}(B) &= \frac{4}{8} \\
\mathbb{P}(A) \cdot \mathbb{P}(B) &= \frac{3}{8} \\
\mathbb{P}(A \cap B) &= \frac{3}{8} = \mathbb{P}(A) \cdot \mathbb{P}(B)
\end{aligned}$$

So the events are independent

7 Lecture 8, Feb 10: Random Variable

Probability theory has 3 building blocks:

1. Sample space (Ω)
2. Probability ($\mathbb{P}$) Together, these two give us probability space $(\Omega, \mathbb{P})$
3. Random Variable

Motivating Example: Let Ω be the collection of all undergraduate students at Brown. Then $\mathbb{P}(A) := \frac{\#A}{\#\Omega}$. And, for each student $\omega \in \Omega$,

$$X(\omega) = \text{the SAT score of the given student} \quad \{X : \Omega \to \mathbb{R},\ \omega \mapsto X(\omega)\}$$

but where ω is unknown (say to protect anonymity) so it can be any value in Ω. Note, however, that if ω is unknown, then X is also uncertain.

Definition: Let $(\Omega, \mathbb{P})$ be a probability space. Suppose X is a real-valued function defined on Ω,

$$\begin{aligned}
X : \Omega &\to \mathbb{R} \\
\omega &\mapsto X(\omega)
\end{aligned}$$

We thus call X a *random variable*

8 Cumulative Distribution Functions

Motivating Example: Ω is all the undergrads at Brown, $\omega \in \Omega$ is a student at Brown, $X(\omega)$ is a random variable denoting the SAT score of a Brown student

Let A_{100} be the event "the SAT score of a Brown student is ≤ 100. Then,

$$A_{100} = \{\omega \in \Omega : X(\omega) \leq 100\}$$

Definition: Let $(\Omega, \mathbb{P})$ be a probability space and X be a random event. Then for any real number $X \in \mathbb{R}$, we define the event A_x by

$$A_x = \{\omega \in \Omega : X(\omega \leq x)\}$$

We define a real-valued function F on $\mathbb{R}$ by

$$F(x) := \mathbb{P}(A_x) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \leq x\})$$

This function F is the *cumulative distribution function (CDF)* of the random variable X, usually written F_X

9 Lecture 9, Feb 13: Cumulative distribution function

Part I - Review

A random variable X is a function defined on a sample space Ω

For each real number x, we define an event

$$A_x = \{\omega \in \Omega : X(\omega) \leq x\}$$

We then define a function by

$$F_X : \begin{array}{c} \mathbb{R} \to [0,1] \\ x \mapsto \mathbb{P}(A_x) \end{array}$$

Note that $F_X = \mathbb{P}(X \leq x)$ and this function is called *the cumulative distribution function*

Part II - The CDF

Example 1: Flipping a coin

- $\Omega = \{H, T\}$
- $\mathbb{P}(A) := \frac{\#A}{\#\Omega} = \frac{\#A}{2} \quad (A \subset \Omega)$
- $$\begin{cases} X(H) = 1 \\ X(T) = 0 \end{cases}$$

Claim: the CDF of X is

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2} & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

Proof:

1. When $x < 0$, then
$$A_x = \{\omega \in \Omega : X(\omega) \leq x\} = \emptyset$$
(this is impossible because X is only 0 or 1 so not negative). So $F_X(x) = \mathbb{P}(A_x) = 0$
2. When $0 \leq x < 1$,
$$A_x = \{\omega \in \Omega : X(\omega) \leq x\} = \{\omega \in \Omega : X(\omega) = 0\} = \{T\}$$
so $F_X(x) = \mathbb{P}(\{T\}) = \frac{1}{2}$
3. When $x \geq 1$
$$A_x = \{\omega \in \Omega : X(\omega) \leq x\} = \{\omega \in \Omega : X(\omega) \leq 1\} = \{T, H\}$$
so $P(A_x) = 1$

Example 2: Flipping a biased coin

- $\Omega = \{H, T\}$
- $$\mathbb{P}(A) := \begin{cases} p & A = \{H\} \\ 1 - p & A = \{T\} \end{cases}$$
for some $p \in [0, 1]$

- $X(H) = 1, X(T) = 0$

Claim: the CDF here is

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - p & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

Proof: same as above

Note: this example actually refers to the Bernoulli Distribution **Definition:** Let X be a random variable on $(\Omega, \mathbb{P})$ If the CDF of X is

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - p & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

for some $o \in [0, 1]$, we say "X follows the Bernoulli Distribution with success probability p" and denote $X \sim \text{Bernoulli}(p)$

Theorem: Let $(\Omega, \mathbb{P})$ be a probability space. Suppose X is a random variable defined on Ω and its CDF is F_X

F_X is non-decreasing ($F_X(x_1) \leq F_X(x_2) \quad x_1 \leq x_2$) for any CDF.

Proof:

$$A_{x_1} = \{\omega \in \Omega : X(\omega) \leq x_1\}$$
$$A_{x_2} = \{\omega \in \Omega : X(\omega) \leq x_2\}$$
$$x_1 \leq x_1 \implies Ax_1 \subset Ax_2$$
$$F_X(x_1) = \mathbb{P}(A_{x_1}) \leq \mathbb{P}(A_{x_2}) \leq F_X(x_2)$$

Lecture 10, Feb 15:

Part I - Review

Let X be a random variable defined on a probability space $(\Omega, \mathbb{P})$. For any real number x,

$$A_x = \{\omega \in \Omega : X(\omega) \leq x\}$$

Then the cumulative distribution function of X is

$$F_X(x) = \mathbb{P}(A_x)$$

Part II - Properties of the CDF

Let X be any random variable on any probability space $(\Omega, \mathbb{P})$ with $F_X(x)$ as the corresponding CDF

1. Any CDF is non-decreasing $(F_X(x_1) \leq F_X(x_2) \quad x_1 \leq x_2)$
2. $$\lim_{x \to -\infty} F_X(x) = 0$$
3. $$\lim_{x \to \infty} F_X(x) = 1$$
   (Note the rigorous proof is needs real analysis)
4. F_X is right-continuous, i.e.
   $$\forall x_0 \in \mathbb{R}, \quad F_X(x_0) = \lim_{x \to x_0^+} F_X(x)$$
   Note that the "right-continuous" property is implied from the $\leq$ sign. With a strict less-than inequality, the CDF becomes left-continuous
5. For any $x_0 \in \mathbb{R}$,
   $$\mathbb{P}(\{\omega \in \Omega : X(\omega) = x_0\}) = F_X(x_0) - \lim_{x \to x_0^-} F_X(x)$$
   Note that this is zero if the CDF is continuous

Lecture 11, Feb 17

Part I - Review

The building blocks of probability together define the CDF:

$$\begin{cases} \Omega \\ \mathbb{P} \\ X \end{cases} \implies F_X(x) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \leq x\})$$

Part II - Moving backwards

Question: Given a function F (satisfying some conditions), do there exist a sample space, a probability, and a random variable corresponding to the given F?

Theorem: Suppose we have a $F : \mathbb{R} \to [0,1]$ satisfying

- F is non-decreasing
- With end behavior

$$\lim_{x \to -\infty} F(x) = 0, \quad \lim_{x \to \infty} F(x) = 1$$

- F is right-continuous

Then, there exist a sample space, a probability, and a random variable such that

$$F(x) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \leq x\})$$

Proof: far beyond the scope of this course

Part III - Classification of Random Variables

Types of random variables:

- Continuous
- Discrete
- Neither continuous nor discrete

Definition: a function F is continuous if

$$\lim_{x \to x_0^-} F(x) = \lim_{x \to x_0^+} F(x) = F(x_0) \quad \forall\, x_0$$

Definition: a random variable X is a continuous random variable if the CDF $F_X : \mathbb{R} \to [0,1]$ is a continuous function

Theorem: Let X be a random variable on $(\Omega, \mathbb{P})$. If X is a continuous random variable,

$$\mathbb{P}(\{\omega \in \Omega : X(\omega) = x_0\}) = 0 \quad \forall\, x_0 \in \mathbb{R}$$

Proof:

$$\mathbb{P}(\{\omega \in \Omega : X(\omega) = x_0\}) = F_X(x_0) - \lim_{x \to x_0^-} F_X(x) = F_X(x_0) - F_X(x_0) = 0$$

Lecture 12, Feb 22: Continuous Random Variables

Part I - A "theorem"

"Theorem": Let F_X be the CDF of a continuous random variable X. Then, F_X is differentiable.

Remark: the true and rigorous version of this "theorem" requires lots of pure math so this version does have some edge cases such that "differentiable" really means "piecewise differentiable"

Example:

$$F(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

This is not technically differentiable because of the sharp points at $x = \{0, 1\}$ but we can use a generalized derivative to cheat:

$$F'(x) = \begin{cases} 0 & x < 0 \text{ or } x > 1 \\ 1 & 0 < x < 1 \\ k & x = 0 \text{ or } x = 1 \end{cases}$$

where k is any value whatsoever.

Definition: Let F_X be the CDF of a continuous random variable X.

$$p_X(x) = F'_X(x)$$

We call $p_X(x)$ the *probability density function (PDF)* of the random variable X

Part II - The Rigorous Treatment (optional)

Definition: X is a continuous random variable if F_X is absolutely continuous

Theorem: X is a continuous random variable. Then its CDF F_X is differentiable "almost everywhere with respect to the Lebesgue measure"

Definition: $p_X(x) = F'_X(x)$ is the probability density function for a continuous random variable X

Part III - The Probability Density Function

Less rigorously,

$$p_X(x) = F'_X(x)$$

is the probability density function and is the piecewise derivative of F_X for the continuous random variable X

Theorem: Let X be a continuous RV with CDF $F_X(x)$ and PDF $p_X(x)$. Then,

$$F_X(x) = \int_{-\infty}^{x} p_X(t)\ dt$$

Remarks:

- $p_X(x) := F'_X(x)$ so CDF determines PDF
- $F_X(x) = \int_{-\infty}^{\infty} p_X(t)\ dt$ so PDF determines CDF

Part IV - Examples of continuous random variables

Definition: Let X be a RV. If the CDF od X is

$$F_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{x-b} & a \leq x \leq b \\ 1 & x > b \end{cases}$$

(for $a < b$), then X follows the uniform distribution between a and b.

This is continuous so we can take the piecewise derivative as follows:

$$p_X(x) = \begin{cases} 0 & x < a \text{ or } x > b \\ \frac{1}{b-a} & a \leq x \leq b \end{cases}$$

Notice that the graph of this function is a rectangle with base $b - a$ and height $1/(b - a)$.

Experiment: Randomly select a number between 0 and 1

- $\Omega = (0, 1)$
- $X : \Omega \to \mathbb{R}, \quad X(\omega) = \omega, \quad \omega \in \Omega = (0, 1)$
- Because we randomly select numbers, we assume $X \sim \text{Unif}(0, 1)$

- Let $E = \{0.5\} = \{\omega \in \Omega : X(\omega) = 0.5\} = \emptyset$ be the event "the selected number is exactly 0.5"

So,

$$\mathbb{P}(E) = \mathbb{P}(X = 0.5) = F_X(0.5) - \lim_{x \to 0.5^-} F_X(x) = 0$$

because F_X is continuous so the limit is equal to the value