

Exercise Walkthrough: Basic Properties of Probability and Measure

Justin Lanfermann

25. June of 2025

Abstract

This document provides a detailed, step-by-step walkthrough for an exercise on the fundamental properties of probability and measure spaces. Each part of the exercise is analyzed, with justifications based on definitions and theorems from the "Discrete Probability Theory" script. The goal is to build understanding from the ground up, clarifying the reasoning behind each conclusion. For key concepts, clickable references are provided for more in-depth explanations at the end of the document.

1 Introduction

We will now go through each statement of the exercise, determine if it is true or false, and provide a clear, step-by-step justification. The reasoning for each step will be explained, and we will reference the relevant definitions from the course script to ensure our arguments are rigorous.

Exercise (i)

Statement: For a probability space (Ω, \mathcal{A}, P) we have $P(\Omega \setminus A) = 1 - P(A)$ for all $A \in \mathcal{A}$.

Verdict: True.

Justification: This statement deals with the probability of the complement of an event. Let's break down the proof using the fundamental axioms of a probability space [1].

1. **Decomposition of the Sample Space:** The sample space Ω can always be partitioned into two [disjoint sets](#): any event A and its complement, $A^c = \Omega \setminus A$. By definition, their union is the entire sample space: $A \cup (\Omega \setminus A) = \Omega$. Also, their intersection is the empty set: $A \cap (\Omega \setminus A) = \emptyset$.
2. **Using Additivity:** A probability measure P is a specific type of measure. According to Definition 1.16(iii) from the script, a measure must be σ -additive [3]. This property implies that for any finite collection of pairwise disjoint sets, the measure of their union is the sum of their measures. Since A and $\Omega \setminus A$ are disjoint, we can write:

$$P(\Omega) = P(A \cup (\Omega \setminus A)) = P(A) + P(\Omega \setminus A)$$

3. **Using the Axiom $P(\Omega) = 1$:** According to Definition 1.18 (probability measure), the measure of the entire sample space must be 1, i.e., $P(\Omega) = 1$. Substituting this into our equation gives:

$$1 = P(A) + P(\Omega \setminus A)$$

4. **Conclusion:** By rearranging the terms, we arrive at the desired statement:

$$P(\Omega \setminus A) = 1 - P(A)$$

This confirms that the probability of an event not happening is one minus the probability that it does.

Exercise (ii)

Statement: Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, $A, B \in \mathcal{A}$ such that $\mu(A \cap B) < \infty$. Then $\mu(B \setminus A) = \mu(B) - \mu(A \cap B)$.

Verdict: True.

Justification: This property is sometimes called the "difference rule" for measures.

1. **Decomposition of B:** Let's consider the set B . We can split it into two disjoint parts: the part of B that also lies in A , and the part of B that does not lie in A .

- The part of B in A is the intersection $A \cap B$.
- The part of B not in A is the set difference $B \setminus A$.

These two sets are disjoint and their union reconstructs B . We can write this formally as:

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

2. **Additivity of the Measure:** Since $(\Omega, \mathcal{A}, \mu)$ is a measure space [2], the measure μ is additive for disjoint sets. Applying this to our decomposition of B :

$$\mu(B) = \mu(B \setminus A) + \mu(A \cap B)$$

3. **The Role of the Finiteness Condition:** The problem states that $\mu(A \cap B) < \infty$. This is a crucial condition. It ensures that $\mu(A \cap B)$ is a real number that can be subtracted. If $\mu(A \cap B)$ were ∞ , an expression like $\mu(B) - \infty$ could be undefined (e.g., if $\mu(B)$ were also ∞).

4. **Conclusion:** Since $\mu(A \cap B)$ is finite, we can subtract it from both sides of the equation:

$$\mu(B) - \mu(A \cap B) = \mu(B \setminus A)$$

This proves the statement. This rule is listed in Proposition 1.17(ii) of the script.

Exercise (iii)

Statement: For a probability space (Ω, \mathcal{A}, P) and a sequence $(A_i)_{i \in \mathbb{N}}$ in \mathcal{A} , we have $P(\bigcup_{i \in \mathbb{N}} A_i) = \sup_{i \in \mathbb{N}} P(A_i)$.

Verdict: False.

Justification: The statement claims that the probability of a union of events is the [supremum](#) of their individual probabilities. This is not generally true. A simple counterexample can demonstrate this. This property, known as continuity from below (Proposition 1.17(iv)), holds only if the sequence of sets is increasing (i.e., $A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots$).

1. **Construct a Counterexample:** Let's use a simple probability space. Consider a single fair coin toss, where the outcomes can be heads (H) or tails (T).

- Sample Space: $\Omega = \{H, T\}$
- Event Space: $\mathcal{A} = \mathcal{P}(\Omega) = \{\emptyset, \{H\}, \{T\}, \{H, T\}\}$
- Probability Measure: $P(\{H\}) = 1/2, P(\{T\}) = 1/2$

2. **Define a Sequence of Events:** Let's define a sequence of events $(A_i)_{i \geq 1}$ that are *not* increasing:

$$A_1 = \{H\}, \quad A_2 = \{T\}, \quad A_n = \emptyset \text{ for } n \geq 3$$

3. **Calculate the Left-Hand Side (LHS):** The LHS is the probability of the union of all events in the sequence:

$$\bigcup_{i=1}^{\infty} A_i = A_1 \cup A_2 \cup A_3 \cup \dots = \{H\} \cup \{T\} \cup \emptyset \cup \dots = \{H, T\} = \Omega$$

Therefore, the probability is:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = P(\Omega) = 1$$

4. **Calculate the Right-Hand Side (RHS):** The RHS is the supremum of the probabilities of the individual events. Let's find the set of these probabilities:

$$\{P(A_1), P(A_2), P(A_3), \dots\} = \{P(\{H\}), P(\{T\}), P(\emptyset), \dots\} = \{1/2, 1/2, 0, 0, \dots\}$$

The supremum (least upper bound) of this set of values is $1/2$.

$$\sup_{i \in \mathbb{N}} P(A_i) = \sup\{1/2, 1/2, 0, 0, \dots\} = 1/2$$

5. **Compare and Conclude:** We found that LHS = 1 and RHS = $1/2$. Since $1 \neq 1/2$, the original statement is false.

Exercise (iv)

Statement: Let (Ω, \mathcal{A}, P) be a probability space and $A, B, C \in \mathcal{A}$. Then $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$.

Verdict: True.

Justification: This formula is a direct application of the Principle of Inclusion-Exclusion for $n = 3$ events, as stated in Theorem 1.20 of the script.

1. **Principle of Inclusion-Exclusion:** Theorem 1.20 provides the general formula for n events:

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k})$$

2. **Applying for n=3:** Let's expand this formula for our case with $A_1 = A, A_2 = B, A_3 = C$.

- For $k = 1$: We sum the probabilities of single events. The term is $(-1)^{1-1}[P(A) + P(B) + P(C)] = P(A) + P(B) + P(C)$.
- For $k = 2$: We subtract the probabilities of all pairwise intersections. The term is $(-1)^{2-1}[P(A \cap B) + P(A \cap C) + P(B \cap C)] = -[P(A \cap B) + P(A \cap C) + P(B \cap C)]$.
- For $k = 3$: We add the probability of the three-way intersection. The term is $(-1)^{3-1}[P(A \cap B \cap C)] = P(A \cap B \cap C)$.

3. **Conclusion:** Combining these terms gives the exact formula stated in the exercise:

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

This confirms the statement is a special case of the general [Inclusion-Exclusion Principle](#) [6].

Exercise (v)

Statement: The triple consisting of \mathbb{N} , $\mathcal{P}(\mathbb{N})$, and the set function $\mu : \mathcal{P}(\mathbb{N}) \rightarrow [0, \infty]$, $A \mapsto \begin{cases} 0 & A \text{ finite} \\ \infty & A \text{ infinite} \end{cases}$ form a measure space.

Verdict: False.

Justification: For $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ to be a measure space, μ must satisfy the properties of a measure from Definition 1.16, most importantly σ -additivity. We can show that this property is violated.

1. **The σ -additivity Requirement:** A function μ is σ -additive if for any *countable sequence of pairwise disjoint sets* $(A_n)_{n \in \mathbb{N}}$, the following equality holds:

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

2. **Construct a Disjoint Sequence:** Let's choose a sequence of disjoint sets whose union is \mathbb{N} . The simplest choice is to let each set contain a single natural number:

$$A_n = \{n\} \quad \text{for each } n \in \mathbb{N} = \{1, 2, 3, \dots\}$$

These sets are pairwise disjoint (e.g., $A_1 \cap A_2 = \{1\} \cap \{2\} = \emptyset$).

3. **Calculate the LHS:** First, we find the measure of the union.

$$\bigcup_{n=1}^{\infty} A_n = \{1\} \cup \{2\} \cup \{3\} \cup \dots = \mathbb{N}$$

The set \mathbb{N} is infinite. According to the definition of our function μ , its measure is:

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

4. **Calculate the RHS:** Next, we find the sum of the measures of the individual sets. For each $n \in \mathbb{N}$, the set $A_n = \{n\}$ is finite (it contains only one element). Therefore, according to the definition of μ :

$$\mu(A_n) = \mu(\{n\}) = 0 \quad \text{for all } n$$

The sum is therefore:

$$\sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} 0 = 0 + 0 + 0 + \dots = 0$$

5. **Conclusion:** We have $\text{LHS} = \infty$ and $\text{RHS} = 0$. Since $\infty \neq 0$, the σ -additivity property fails. Therefore, μ is not a measure, and the triple is not a measure space.

Exercise (vi)

Statement: We consider the two-time fair coin flip, where we interpret the two sides to count as 0 or 1. We are only interested in the sum of the two outcomes. Then the cdf is given by

$$F(x) = \begin{cases} 0 & x < 0 \\ 1/4 & x \in [0, 1) \\ 3/4 & x \in [1, 2) \\ 1 & x \geq 2 \end{cases}$$

Verdict: True.

Justification: Let's derive the Cumulative Distribution Function (CDF) [7] from scratch and compare it to the given function.

1. **Model the Experiment:** The experiment is two fair coin flips. Let's denote the outcomes as (flip1, flip2). The sample space is $\Omega = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$. Since the flips are fair and independent, each of these four elementary outcomes has a probability of $1/4$.
2. **Define the Random Variable:** We are interested in the sum of the outcomes. Let S be this random variable. The possible values for S are:
 - $S = 0$ for outcome $(0, 0)$.
 - $S = 1$ for outcomes $(0, 1)$ and $(1, 0)$.
 - $S = 2$ for outcome $(1, 1)$.

The set of possible values for S is $\Omega_S = \{0, 1, 2\}$.

3. **Determine the Probability Mass Function (pmf) of S:** We calculate the probability for each possible value of S .
 - $P(S = 0) = P(\{(0, 0)\}) = 1/4$.
 - $P(S = 1) = P(\{(0, 1), (1, 0)\}) = P(\{(0, 1)\}) + P(\{(1, 0)\}) = 1/4 + 1/4 = 1/2$.
 - $P(S = 2) = P(\{(1, 1)\}) = 1/4$.

(Check: $1/4 + 1/2 + 1/4 = 1$. The pmf is valid.)

4. **Derive the CDF $F(x) = P(S \leq x)$:** We build the CDF by considering different ranges for x .
 - **For $x < 0$:** There are no values $s \in \Omega_S$ such that $s \leq x$. So, $P(S \leq x) = P(\emptyset) = 0$.
 - **For $0 \leq x < 1$:** The only value $s \in \Omega_S$ that satisfies $s \leq x$ is $s = 0$. So, $F(x) = P(S \leq x) = P(S = 0) = 1/4$.
 - **For $1 \leq x < 2$:** The values $s \in \Omega_S$ satisfying $s \leq x$ are $s = 0$ and $s = 1$. So, $F(x) = P(S \leq 1) = P(S = 0) + P(S = 1) = 1/4 + 1/2 = 3/4$.
 - **For $x \geq 2$:** All possible values of S satisfy $s \leq x$. So, $F(x) = P(S \leq 2) = P(S = 0) + P(S = 1) + P(S = 2) = 1/4 + 1/2 + 1/4 = 1$.
5. **Conclusion:** The derived CDF is identical to the one given in the statement. Thus, the statement is true.

In-Depth Explanations

[1] **Probability Space (Def 1.18):** A probability space is the mathematical foundation for modeling a random process. It is a triple (Ω, \mathcal{A}, P) where:

- Ω is the **sample space**, the set of all possible outcomes of an experiment (e.g., for a die roll, $\Omega = \{1, 2, 3, 4, 5, 6\}$).
- \mathcal{A} is the **event space** (a σ -algebra on Ω), which is a collection of subsets of Ω . These subsets are called "events". It must contain Ω itself and be closed under complement and countable unions. For discrete spaces, we often use the power set $\mathcal{P}(\Omega)$, which is the set of all possible subsets.
- P is a **probability measure**, a function that assigns a probability (a number in $[0, 1]$) to every event in \mathcal{A} . It must satisfy $P(\Omega) = 1$ and be **σ -additive**.

[2] **Measure Space (Def 1.16):** A measure space is a more general concept than a probability space. It's a triple $(\Omega, \mathcal{A}, \mu)$ where Ω and \mathcal{A} are the same as above, but μ is a **measure**. A measure is a function $\mu : \mathcal{A} \rightarrow [0, \infty]$ that assigns a non-negative "size" to each set in \mathcal{A} . It must satisfy $\mu(\emptyset) = 0$ and be **σ -additive**. Probability is just a special case where the total "size" of the space is 1.

[3] **σ -additivity (Def 1.16(iii)):** This is a crucial property of any measure (including probability measures). It states that if you have a *countable* sequence of *pairwise disjoint* events A_1, A_2, A_3, \dots (meaning $A_i \cap A_j = \emptyset$ for $i \neq j$), then the measure of their union is equal to the sum of their individual measures:

$$\mu \left(\bigcup_{i=1}^{\infty} A_i \right) = \sum_{i=1}^{\infty} \mu(A_i)$$

This property is what allows us to break down complex events into simpler, disjoint pieces and sum their probabilities.

[4] **Disjoint Sets:** Two sets are disjoint if they have no elements in common. Their intersection is the empty set, \emptyset . For example, the set of even numbers and the set of odd numbers are disjoint.

[5] **Supremum (sup):** The supremum of a set of real numbers S is its **least upper bound**. It's the smallest number that is greater than or equal to every number in S . For the set $\{0.9, 0.99, 0.999, \dots\}$, the supremum is 1. For a finite set like $\{1, 5, 2\}$, the supremum is simply the maximum, 5.

[6] **Inclusion-Exclusion Principle (Thm 1.20):** This principle provides a way to calculate the probability of the union of several events. The basic idea for two events is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. We subtract $P(A \cap B)$ because it was counted twice (once in $P(A)$ and once in $P(B)$). The formula for three or more events extends this logic, alternately adding and subtracting the probabilities of intersections of increasing size.

[7] **Cumulative Distribution Function (CDF) (Def 1.21):** The CDF of a random variable X , denoted $F_X(x)$, gives the probability that X will take a value less than or equal to x . Formally, $F_X(x) = P(X \leq x)$. The CDF completely describes the distribution of a random variable. For a discrete variable, the CDF is a step function that jumps at each possible value of the variable. The height of each jump corresponds to the probability of that specific value.