Notes on Probability

November 11, 2019

Contents

Some conventions:

- $\mathbb{N} = \{0, 1, 2, \ldots\}$,
- A_+ with $A = \mathbb{R}, \mathbb{Q}, \mathbb{N}, \mathbb{Z}$, refers to the respective subset of positive elements,
- A_- is the same as above but for negative elements.

1 Combinatorics

Definition 1.1. Let n be a natural number. The *factorial* of n, denoted $n!$, is defined by

$$\begin{cases} 1 & \text{if } n = 0, \\ n(n-1)! & \text{otherwise.} \end{cases}$$

Proposition 1.1. *Let $A = \{x_1, \ldots, x_n\}$ be a set, such that $x_i \neq x_j$ whenever $i \neq j$. There are exactly $n!$ bijections from A to itself.*

Proof. □

Definition 1.2. Let $n \geq k \geq 0$ positive integers. The binomial coefficient is defined as

$$\binom{n}{k} := \frac{n!}{k!(n-k)!} \tag{1}$$

2 The Basics of the Theory of Probability

2.1 Axiomatization

Definition 2.1. Consider a non-empty set Ω together with a function $\mathbb{P} : \mathcal{E} \subset \mathcal{P}(\Omega) \to [0, 1]$. The triple $(\Omega, \mathcal{E}, \mathbb{P})$ is said to be a *probability space* if the following propositions are true:

1. $\mathcal{E}$ is closed under complementation and under countable unions,
2. $\Omega \in \mathcal{E}$;
3. $\forall E \in \mathcal{E}, \mathbb{P}(E) \geq 0$;
4. Let $(E_n \in \mathcal{E} : n \in \mathbb{N})$ be a pairwise disjoint, then

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

5. $\mathbb{P}(\Omega) = 1$.

The set Ω is conventionally called the *sample space*, a set $E \in \mathcal{E}$ is called an *event*, the function $\mathbb{P}$ is called a *measure of probability* on $\mathcal{E}$.

Remark 1. There are other equivalent axiomatizations, and later we will cover the same ground through the measure-theoretic perspective. **For now, assume that whenever we speak of a probability space, the set of events $\mathcal{E} = \mathcal{P}(\Omega)$, so that an event is simply a subset of Ω.**

We now state some elementary consequences of these axioms.

Theorem 2.1. *Let $(\Omega, \mathcal{E}, \mathbb{P})$ be a probability space and $A, B \in \mathcal{E}$, then*

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

2. *If $E_1, \ldots, E_n$ are pairwise disjoint events, then*

$$\mathbb{P}(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} E_i,$$

3. $\mathbb{P}(\Omega \setminus A) = 1 - \mathbb{P}(A)$,

4. *If $A \subset B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$,*

5. $0 \leq \mathbb{P}(A) \leq 1$,

6. $\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$,

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

8. $\mathbb{P}(A \cap B) \geq 1 - \mathbb{P}(\Omega \setminus A) - \mathbb{P}(\Omega \setminus B)$.

We may sometimes write $\Omega \setminus A$ as $A^{\complement}$ for the sake of brevity.

2.2 Independence and Conditional Probability

Definition 2.2. Let $(\Omega, \mathcal{E}, \mathbb{P})$ be a probability space, and let $A, B \in \mathcal{E}$, with $\mathbb{P}(B) > 0$. The *conditional probability of A given B* is

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

Definition 2.3. Events A and B are said to be independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

2.3 Random Variables

Definition 2.4. Let $(\Omega, \mathcal{E}, \mathbb{P})$ be a probability space.

- By a *random variable* X, we mean a function $X : \Omega \to \mathbb{R}$.
- If $X(\Omega)$ is a countable set, then X is said to be a *discrete* random variable.
- If X is a discrete random variable and $(x_i \in \mathbb{R} : i \in I)$ is an indexed list of the values it takes, the function $x_i \mapsto \mathbb{P}(X = x_i) = p_i$ is called the *probability mass function* of X. Furthermore it satisfies
$$\sum_{i \in I} p_i = 1 \text{ and } \forall i \in I : p_i \geq 0.$$
- The *cumulative distribution function* of X is the function $F_X : \mathbb{R} \to \mathbb{R}$ given by $x \mapsto \mathbb{P}(X \leq x)$.

The expression $\mathbb{P}(X = x)$ is merely an abbreviation of $\mathbb{P}(\{y \in \Omega : X(y) = x\})$. The set $\{y \in \Omega : X(y) = x\} = X^{-1}(\{x\})$ is simply the pre-image of $\{x\}$, which is clearly a subset of Ω. [1] Similarly $X \leq x$ may represent $\{y \in \Omega : X(y) \leq x\}$, we will henceforth make use of this convention whenever unambiguous.

Theorem 2.2 (Properties of cumulative distribution functions)**.** *Let F_X the cumulative distribution function of X, then*

1. $0 \leq F_X(x) \leq 1$.
2. $\lim_{x \to +\infty} F_X(x) = 1$.
3. $\lim_{x \to -\infty} F_X(x) = 0$.
4. $a < b \Rightarrow F_X(a) \leq F_X(n)$.

2.4 Expectation and Variance

Definition 2.5. Let X be a discrete random variable

3 Distributions

3.1 Binomial

3.2 Negative Binomial

3.3 Hypergeometric

3.4 Poisson

[1] But is it an element of $\mathcal{E}$?