Honors Single Variable Calculus 110.113

September 21, 2023

Contents

1 Project Homework

> *... in mathematics you don't understand things. You just get used to them* - von Neumann

Learning Objectives

A recurring theme that you would see throughout your study of more "theoretical" sciences is

- Making *good* definitions.
- *Working* with definitions

This project aims to familiarize you with the foundations of probability theory as set up by A. Komolgorov. In pure mathematics and its applications, it is desirable to have a foundation where one can discuss non deterministic statements, which we will refer as *events*, and non deterministic values, which are *random variables*. The project will proceed in the following order:

1. Probability space, 1.1.
2. How one can use this as a language for *modeling*, such as making inferences.
3. You will then have two choices to explore:
    - probability theoretic: we will explore foundational results as the strong law of large numbers.

- statistical theoretic: this is "more applied; we will explore application in the realm of inference and language models.

Current status of the available content: only 1. is available. There will be compulsory problems from both 1 and 2. In 3. you have one of two options depending on your taste.

1.1 Defining a probability space following A. Kolmogorov

Definition 1.1. A *measure space* consists of a pair $(\Omega, \mathcal{E})$ where Ω is a set, and $\mathcal{E}$ is a σ-algebra on Ω.

- elements $E \in \mathcal{E}$ are referred as *events* or *measurable sets.*

Definition 1.2. Let $(\Omega, \mathcal{E})$ be a measure space. A *(finite) probability measure* is a map $\mathbb{P} : \mathcal{E} \to \mathbb{R}_{\geq 0}$ satisfying

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

2. Finitely additivity. Let $\{A_i\}_{i \in I}$ be a finite (that is $|I| = n$ for some $n \in \mathbb{N}$) collection of disjoint (def. **??**) elements in $\mathcal{E}$ [1]. Then

$$\mathbb{P}\left(\bigcup_{i=0}^{N} A_i\right) = \sum_{i=0}^{N} \mathbb{P}(A_i)$$

Once we have learnt the definition of series, we will add in another axiom called *countable additivity.*

Definition 1.3. A *probability space* is the datum of $(\Omega, \mathcal{E}, \mathbb{P})$, where $\mathbb{P}$ is a probability measure.

Example

The discrete case. Let Ω be a finite discrete set.

1. $\mathcal{E} := 2^{\Omega}$ is the set of all subsets of Ω. This is a σ-algebra.

2. Let p_w be any finite collection of real numbers such that $\sum_{w \in \Omega} p_w = 1$. Then there is a map

$$\mathbb{P} : 2^{\Omega} \to [0, 1]$$

uniquely extending the condition

$$\mathbb{P}(\{w\}) = p_w \quad w \in \Omega$$

[1] Remember, these are subsets of 2^{Ω}.

Example

Modeling n tosses of a fair coin. We define $(\Omega_n, \mathcal{E}, \mathbb{P})$.

- Ω_n is the set of all n consecutive ordered sets of letters which are either H or T.[a]
- $\mathcal{E}$ is the set of all subsets of Ω_n. One event can be

$$E_{\geq k} := \{\omega \in \Omega_n \,:\, \text{at least } k \text{ heads appear in the } n \text{ toss}\}$$

  This is the set of all sequences with at least k Hs.
- Set $\mathbb{P}(\{\omega\}) = \frac{1}{2^n}$ for all singleton subsets $\{\omega\} \in \mathcal{E}$ where $\omega \in \Omega_n$. This uniquely extends to a function (why?)

$$\mathbb{P} : \mathcal{E} \to \mathbb{R}_{\geq 0}$$

[a] Of course, from our language of set theory, this is not a valid set. But we can equally use 0 or 1 to model this, in this case, this follows from the axioms.

1.1.1 Problems on n toss of a fair coin

The following problems are related to the model described above on n-tosses of a fair coin.

1. List out the elements in the event space of

$$\Omega_i, E_{\geq i}$$

   for $i = 1, 2$ and 3. Prove Ω_n has 2^n elements for $n \in \mathbb{N}_{\geq 1}$[2]

2. For a fix choice of n, give a formula for

$$\mathbb{P}(E_{\geq i})$$

   for $0 \leq i \leq n$.

3. Consider now Ω_{2n}. How many elements are the event

$$E := \{\text{exactly } n \text{ heads appear}\}$$

   Prove that

$$\mathbb{P}(E) = \frac{1}{2^{2n}}\binom{2n}{n}$$

   [3]

[2] This will be a short hand for positive integer.
[3] One can apply *Stirling's* formula to show that this is $\sim \frac{1}{\sqrt{\pi n}}$ as $n \to \infty$.

1.2 Conditional Expectation

Let us consider the discrete case for warm-up, once we have learnt integration, we will repeat the same story for density functions. The result below is referred as Baye's rule.

Definition 1.4. A *partition of a* X is a collection of subsets X_i, indexed by a set $i \in I$ such that

1. $\bigcup_{i \in I} X_i = X$
2. The sets X_is are pairwise disjoint: for any $i, j \in I$, the intersection (Def. **??**) of X_i and X_j is empty, $X_i \cap X_j = \emptyset$.

1.2.1 Problems

1. (**) Let I be a finite set. Let $\{B_1, B_2, \ldots\}_{i \in I}$ be a finite partition, 1.4, of Ω and $\mathbb{P}(B_i) > 0$ for all $i \in I$. Prove that

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A|B_i)\mathbb{P}(B_i)$$

using the additivity axiom.

2. (**) By conditioning on something, we would expect that we get a *new* probability space. If $B \in \mathcal{E}$ such that $\mathbb{P}(B) > 0$ show that $\mathbb{Q} : \mathcal{E} \to \mathbb{R}$ given by $\mathbb{Q}(A) = \mathbb{P}(A|B)$ defines a probability space $(\Omega, \mathcal{E}, \mathbb{Q})$.

References