

Some Notes on σ -algebra and Random Variables

Ali Fele Paranj

January 25, 2025

1 Introduction

The probability theory is so natural that one can start thinking about some basic problems as soon as they start to know how to count! Because finding the answer to most of the problems boils down to counting the number of all possible scenarios, and then reporting the proportion of the outcomes of interest to the total number. The first shock in probability theory is experienced when you sit in a undergraduate level probability course. Then you will learn about the notion of a random variable that is a function! And even more shockingly, there is nothing random about a random variable. Surviving so many abuse of notations (like probability density function, probability function, distribution, law, etc) then you might get a feeling that you understand what probability is saying and you can imagine a peaceful life after your course. However, once one encounters the treatment of probability theory in a graduate level course, at first, you might even think that you are sitting in a wrong class (and continue thinking the same way for 4 more weak if not the whole term!). That is because you will hear about different notions like σ -algebra, measurable random variables, a filtration, adaptive random variable, and etc. The situation becomes even more scary when you hear for the first time that conditional expectation is a random variable! In this note, I will provide a more algebraic treatment of the notion of σ -algebra which will help to understand these notions better. Also, by the end of these notes you will have a more tangible sense of the damn sentence “we can think of σ -algebra as information available to us!”. Seriously.

2 σ -Algebra

The notion of σ -algebra is easiest to understand when considered in a finite setting. We are not going to review the definition of σ -algebra here. However, we want to highlight that we can look at a σ -algebra of a sample space Ω in a way that resembles a vector space. Then we can talk about subspaces of that vector spaces (which will be the notion of sub σ -fields). We start with an explicit example.

Let $\Omega = \{1, 2, 3\}$ be our sample space. Then for σ -algebra we have many options. One extreme we can have $\mathcal{F}_0 = \{\emptyset, \Omega\}$, and on the other extreme we can have

$$\mathcal{F}_3 = \{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\}, \Omega\}$$

We can consider any of these collections as a collection of functions rather than a collection of sets. For instance, we can define

$$\mathcal{F}_3 = \{f : \{1, 2, 3\} \rightarrow \mathbb{F}_2\},$$

where $\mathbb{F}_2 = \{0, 1\}$ is a field with characteristic 2. For instance the set $\{1\}$ will be the same as the function f where $f(1) = 1, f(2) = 0$, and $f(3) = 0$. And when the domain of those functions is finite, then the followings are equivalent representations of the same thing.

$$f(1) = 1, f(2) = 0, f(3) = 0, \quad (1, 0, 0), \quad \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 0 \end{pmatrix}, \quad 100,$$

and the list above is not an exhaustive one. Thus we can also represent the elements of \mathcal{F}_3 as

$$\mathcal{F}_3 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}.$$

It is easy to see that \mathcal{F}_3 forms a vector field on \mathbb{F}_2 . In other way of seeing this is to observe that

$$\mathcal{F}_3 = \bigoplus_{i=1}^3 \mathbb{F}_2,$$

i.e. \mathcal{F}_3 is a direct sum (equivalently direct product when we can dealing with finitely many of spaces) of the underlying field \mathbb{F}_2 . But what about the case that we choose

$$\mathcal{F}_1 = \{\{1\}, \{2, 3\}, \{1, 2, 3\}, \emptyset\}.$$

Then how we can show this collection as a vector space in above? It turns out that we can write it as

$$\mathcal{F}_1 = \{f : \{\{1\}, \{2, 3\}\} \rightarrow \mathbb{F}_2\}.$$

Observe that $\{\{1\}, \{2, 3\}\}$ is in fact a partition of Ω . So in general we can summarize as below.

Summary. Let Ω be a finite sample space, and $\mathcal{P} = \{P_1, P_2, \dots, P_n\}$ be any partition. Then the σ -algebra generated by \mathcal{P} , denoted as $\mathcal{F} = \sigma(\mathcal{P})$, is given as

$$\mathcal{F} = \left\{ \bigcup_{P \in A} P \mid A \subseteq \mathcal{P} \right\}.$$

Or equivalently

$$\mathcal{F} = \{f : \mathcal{P} \rightarrow \mathbb{F}_2\} = \bigoplus_{P \in \mathcal{P}} \mathbb{F}_2.$$

Observe that \mathcal{F} has a vector space structure. The elements of the partition \mathcal{P} is called the **atoms** of \mathcal{F} . Furthermore, the collection $\mathbb{B} = \{f_i\}_{i \in \mathcal{F}} \subset \mathcal{F}$, where $f_i(P_j) = \delta_{i,j}$ forms a basis for the \mathcal{F} .

We will follow the second interpretation above, as it provides a more clear algebraic structure. To understand the notion of measurable functions (i.e. random variable) we need to introduce the dual basis to \mathbb{B} . Consider the collection of functionals $\{\mathbb{1}_i\}_{i=1}^n$ defined as

$$\mathbb{1}_i : \mathcal{F} \rightarrow \mathbb{F}_2, \quad \text{where} \quad \mathbb{1}_i(f_j) = \delta_{i,j}.$$

This collection forms a dual basis for \mathcal{F} , i.e. it is a basis for \mathcal{F}' .

Summary. *In the same setting as the summary above, the collection $\{\mathbb{1}_i\}_{i=1}^n$ defined as*

$$\mathbb{1}_i : \mathcal{F} \rightarrow \mathbb{F}_2, \quad \text{where} \quad \mathbb{1}_i(f_j) = \delta_{i,j}$$

forms a dual basis for \mathcal{F} . With some abuse of notation, these functions are the similar to the indicator function $\{\mathbb{1}_P\}_{P \in \mathcal{P}}$ when we view the σ -algebra as $\mathcal{F} = \{\bigcup_{P \in A} P \mid A \subseteq \mathcal{P}\}$.

With this point of view, the notion of measurable functions becomes very easy to grasp. A measurable function is simply an element of \mathcal{F}' . I.e. $X : \Omega \rightarrow \mathbb{R}$ is measurable if and only if we can write it as a linear combination of dual basis. With slight abuse of notation (see the summary below for clarification) we can write:

$$X = \sum_{i=1}^n \alpha_i \mathbb{1}_i,$$

or with the notation of the indicator functions

$$X = \sum_{P \in \mathcal{P}} \alpha_i \mathbb{1}_P.$$

Summary. *X is \mathcal{F} -measurable if and only if $X \in \mathcal{F}'$ in the sense that we can write X as*

$$X = \sum_{i=1}^n \alpha_i \mathbb{1}_i \circ q$$

where $q : \Omega \rightarrow \mathcal{P}$ is the quotient map.

References