Some Notes on σ-algebra and Random Variables

Ali Fele Paranj

January 26, 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.
Things to be added or changed later: Notice that for any Ω the 2^Ω already has the same algebraic structure as discussed below, i.e. $2^\Omega = \oplus_{x\in\Omega}\mathbb{F}_2$. Then any σ-algebra will be a subspace of this space 2^Ω.

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.

2.1 Finite Case

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

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

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.

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

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

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

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

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

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

Summary 1. *Let Ω be a finite sample space, and $\mathcal{P} = \{P_1, P_2, \cdots, P_n\}$ be any partition. Then the σ-algebra generated by $\mathcal{P}$, denoted as $\mathcal{F} = \sigma(\mathcal{P})$, is given as*

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

Or equivalently

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

*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 $\mathbb{B} = \{f_i\} \subset \mathcal{F}$, 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

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

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

Summary 2. *In the same setting as the summary above, the collection $\{\mathbb{1}_i\}_{i=1}^n$ defined as*

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

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. $X : \Omega \to \mathbb{R}$ 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 3. *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 $q : \Omega \to \mathcal{P}$ is the quotient map.

The following example demonstrates our setup above. However, this example might sound quite silly, as we are over-killing our machinery for the purpose of demonstrating the tools we have in our disposal.

Example 1. Consider an experiment of tossing a dice. So the sample space will be

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Assume that we don't have enough technology to determine the exact outcome of an experiment. Instead, our measurement device can only report the outcome of the experiments in $\mod 2$, i.e. if the outcome is even or is odd. So the set of atomic events (i.e. the σ-algebra) will be[1]

$$\mathcal{A} = \{\{1, 3, 5\}, \{2, 4, 6\}\}.$$

And the σ-algebra of interest will be the sigma algebra generated by these atoms, i.e.

$$\mathcal{F} = \sigma(\mathcal{A}) = \{\{1, 3, 5\}, \{2, 4, 6\}, \Omega, \varnothing\}.$$

Or in a more algebraic point of view we have

$$\mathcal{F} = \bigoplus_{p \in \mathcal{A}} \mathbb{F}_2 = \mathbb{F}_2 \oplus \mathbb{F}_2 = \{f : \mathcal{A} \to \mathbb{F}_2\}.$$

Fixing a basis for this space (as discussed in Summary 1) we can write the elements of this vector space explicitly as

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

Now suppose that we did some upgrade in our measurement system (or wrote down some theory that enables more accurate measurements) our device can now measure the exact value of the outcome of the experiment. So the atoms for this measurement device, i.e. random variable will be

$$\mathcal{A}' = \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}\}.$$

So the generated σ-algebra will be

$$\mathcal{F}' = \sigma(\mathcal{A}') = 2^{\Omega},$$

i.e. the set of all subsets of of Ω. Or in the algebraic point of view we have

$$\mathcal{F}' = \mathbb{F}_2 \oplus \mathbb{F}_2 \oplus \mathbb{F}_2.$$

Remark 1. As we saw in the example above, it is usually the random variable that induces the σ-algebra rather than having a σ-algebra that determines which random variable is measurable and which is not. It is because of this reason that we are generally interested in the σ-algebra generated by a collection of random variables.

[1] Observe that $\mathcal{A}$ is a partition of the set Ω.

2.2 Infinite Case

When considering finite sample spaces, most of the notions above seems to be too much complexity for a system that can by analyzed easily. But the usefulness of the notions above becomes more evident when we consider sample spaces with infinite cardinality, specially the sample spaces associated to the cases where each outcome is the result of a countably many repeated experiments. For instance, the sample space for an infinite coin toss, which is the space of sequences of 0s and 1s. This system will be our running example for the rest of this section.

Example 2 (Running Example)*.* Imagine you tossing a coin infinitely many times. The sample space for this will be

$$\Omega = \{(1, 0, \dots), (0, 0, \cdots), (1, 1, \dots), (1, 0, \dots)\},$$

i.e. the set of all sequences composed of 1s and 0s. It is very hard to imagine this sample space in any practical and useful way. But we can identify the elements of this sample space with real numbers in $(0, 1]$. For each $\omega \in \Omega$ we identify it with $a \in (0, 1]$ where the binary digits of a in the non-terminating expansion matches with ω. For instance let $\omega_1 = (0, 1, 0, 1, 0, 1, \dots)$. The corresponding real number a_1 will be

$$a_1 = 0.01010101\dots.$$

And for $\omega_2 = (1, 0, 0, 0, \dots)$ the corresponding real number a_2 will be

$$a_2 = 0.01111111\dots.$$

Observe that we chose the non-terminating representation of $0.10000\dots$ (for more details see my lecture notes on the probability theory, or see Billingsley section 1.1). Now imagine that one outcome $\omega \in \Omega$ has occurred but we do not have enough machinery and technology to exactly say which one has occurred. Instead, our machinery can only determine what was the outcome of the coin flip on the fits toss. I.e. we have the following partition for the sample space

$$\mathcal{A} = \{A_0, A_1\},$$

where A_0 denotes the set of all outcomes whose first entry is 0 and A_1 denotes the set of all outcomes whose first entry if 1. So the σ-algebra will be

$$\mathcal{F}_1 = \sigma(\mathcal{A}) = \{A_0, A_1, \Omega, \varnothing\}.$$

Or for the algebraic point of view

$$\mathcal{F}_1 = \sigma(\mathcal{A}) = \mathbb{F}_1 \oplus \mathbb{F}_2.$$

Using Summary 2 the dual basis will be

$$\mathbb{B}' = \{\mathbb{1}_{A_0}, \mathbb{1}_{A_1}\}$$

Since it is hard to imagine the structure of Ω as above, we use our identification above to construct the similar notions for $\tilde{\Omega} = (0,1]$. With our identification of elements of Ω with real numbers it is easy to see that

$$\tilde{\mathcal{A}} = \{(0, \frac{1}{2}], (\frac{1}{2}, 1]\}.$$

So

$$\tilde{\mathcal{F}}_1 = \{(0, \frac{1}{2}], (\frac{1}{2}, 1], \Omega, \varnothing\}.$$

And the dual basis functions will be

$$\tilde{\mathbb{B}}' = \{\mathbb{1}_{(0,\frac{1}{2}]}, \mathbb{1}_{(\frac{1}{2},1]}\}.$$

In the setting of example above, we are interested in some special random variables that will be crucial for the rest of our analysis. So for the space $(\Omega, \mathcal{F}_1)$ as above we define

$$X_1 : \Omega \to \mathbb{R}, \qquad (x, \dots) \to x,$$

where ω_1 is its first entry. Or in the corresponding space $(\tilde{\Omega}, \tilde{\mathcal{F}}_1)$ we can write this random variable as

$$d_1 : \tilde{\Omega} \to \mathbb{R}, \qquad 0.x \cdots \to x.$$

We call d_1 as dyadic function.

Example 3. Assume that we have received some new measurement devices that can measure the second, and the third entry of each outcome respectively. Each of these devices will result in the following measurable spaces

$$(\Omega, \mathcal{F}_2), \quad (\Omega, \mathcal{F}_3),$$

where

$$\mathcal{F}_2 = \sigma(\mathcal{A}_2) = \sigma(\{A_{00}, A_{01}, A_{10}, A_{11}\}),$$

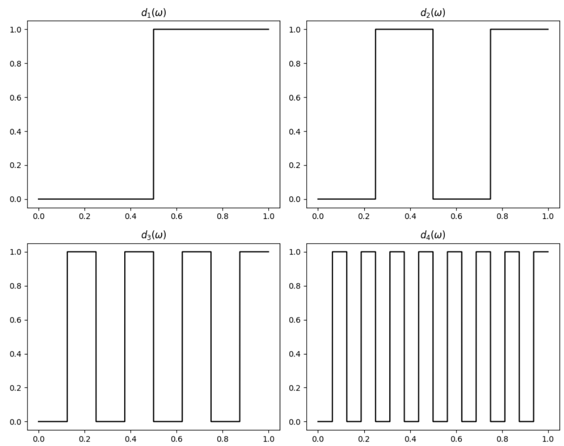
where A_{10} is the set of all outcomes that start with 10, and a similar definition for other sets. Similarly

$$\mathcal{F}_3 = \sigma(\mathcal{A}_3) = \sigma(\{A_{000}, A_{001}, A_{010}, A_{011}, A_{100}, A_{101}, A_{110}, A_{111}\}).$$

This becomes more interesting when we consider the $\tilde{\Omega}$, the real-interval counterpart of these sets. We are not going to list all of them here but for instance we have

$$A_{00} = (0, \frac{1}{4}], \quad A_{10} = (\frac{2}{2}, \frac{3}{4}], \quad A_{110} = (\frac{6}{8}, \frac{7}{8}], \quad A_{011} = (\frac{3}{8}, \frac{4}{8}], \quad \dots.$$

It is also very informative to look at the dyadic functions d_2, and d_3 associated with these σ-algebras.



First observe that

- d_1 is $\mathcal{F}_1$ measurable.
- d_2 is $\mathcal{F}_2$ measurable.
- d_3 is $\mathcal{F}_3$ measurable.
- d_4 is $\mathcal{F}_4$ measurable.

It is immediate to see the above as each of the dyadic functions above is simple a linear combination of the dual basis elements of their corresponding σ-algebra. It is also easy to see that these random variables are independent. For instance to check the independence between d_1 and d_2 we have

$$\mathbb{P}(d_1^{-1}(1)|d_2^{-1}(1)) = \frac{\mathbb{P}((1/2,1] \cap ((1/4,1/2] \cup (3/4,1]))}{\mathbb{P}((1/4,1/2] \cup (3/4,1]))} = \frac{1/4}{1/2} = 1/2 = \mathbb{P}(d_1^{-1}(1)).$$

$$\mathbb{P}(d_1^{-1}(0)|d_2^{-1}(0)) = \frac{\mathbb{P}((0,1/2] \cap ((0,1/4] \cup (2/4,3/4]))}{\mathbb{P}((0,1/4] \cup (2/4,3/4]))} = \frac{1/4}{1/2} = 1/2 = \mathbb{P}(d_1^{-1}(0)).$$

3 Conditional Expectation

The algebraic point of view above helps with understanding the notion of conditional expectation in a much more depth! The following summary covers most of the ideas we have developed above.

Summary 4. *For any sample space Ω we can associate a vector space given as*

$$2^\Omega = \prod_{x\in\Omega} \mathbb{F}_2,$$

where we can replace $\prod$ with $\bigoplus$ if Ω is finite. The dual basis of this space will be $\{f_x\}_{x\in\Omega}$ where $f_x(A) = \mathbb{1}_A(x)$. Given any partition of the space, e.g. $\mathcal{A}$, which are called the atoms, we will have an associated subspace (up to an isomorphism) of Ω given as

$$\mathcal{F} = \bigoplus_{P\in\mathcal{A}} \mathbb{F}_2.$$

The dual basis vectors for this space is $\{f_P\}_{P\in\mathcal{A}}$ where $f_P = \mathbb{1}_P$, i.e. assumes the value 1 on the partition P and 0 otherwise [2].

Let X be a $\mathcal{F}$-measurable random variable, and let $\mathcal{G} \subset \mathcal{F}$ be a sub σ-algebra. Then the conditional expectation of X with respect to $\mathcal{G}$ is really a projection of the functional X (that can be expressed in therms of the dual basis of $\mathcal{F}$) to a functional on $\mathcal{G}$. For the sale of concreteness consider the following example.

Example 4. Let Ω be a sample space and let $\mathcal{A}_n = \{P_1, \cdots, P_n\}$, $\mathcal{A}_m = \{P'_1, \cdots, P'_m\}$ be two partitions. Note that we can have the situation where $\mathcal{A}_n$ is a partition, i.e. $\mathcal{A}_n$ is the finer partition that partitions each element of $\mathcal{A}_m$ which is a courses partition. For instance you can assume the finer partition to be $\{(0, 1/4], (1/4, 2/4], (2/4, 3/4], (3/4, 1/2]\}$, and the coarser partition to be $\{(0, 1/2], (1/2, 1]\}$. Let $\mathcal{F}_n = \sigma(\mathcal{A}_n)$, and $\mathcal{F}_m = \sigma(\mathcal{A}_m)$ be the corresponding σ-algebras. Let X be a $\mathcal{A}_n$-measurable random variable. I.e. we can write

$$X = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{P_i}.$$

Then we define the conditional expectation of X as

$$\mathbb{E}[X|\mathcal{F}_n] = \sum_{i=1}^{m} \mathbb{E}[X|P'_i]\mathbb{1}_{P'_i}.$$

We can write this projection linear operator as a matrix by applying it on the dual basis vectors of $\mathcal{F}_n$. The matrix of this projection can be written as

$$M = \begin{bmatrix} \mathbb{E}[\mathbb{1}_{P_1}|P'_1] & \mathbb{E}[\mathbb{1}_{P_2}|P'_1] & \cdots & \mathbb{E}[\mathbb{1}_{P_n}|P'_1] \\ \mathbb{E}[\mathbb{1}_{P_1}|P'_2] & \mathbb{E}[\mathbb{1}_{P_2}|P'_2] & \cdots & \mathbb{E}[\mathbb{1}_{P_n}|P'_2] \\ \vdots & \vdots & \ddots & \cdots \\ \mathbb{E}[\mathbb{1}_{P_1}|P'_m] & \mathbb{E}[\mathbb{1}_{P_2}|P'_m] & \cdots & \mathbb{E}[\mathbb{1}_{P_n}|P'_m] \end{bmatrix}.$$

Observe that $\mathbb{E}[\mathbb{1}_{P_1}|P'_1] = \mathbb{P}(P_1|P'_1)$. So we can write

$$M = \begin{bmatrix} \mathbb{P}[P_1|P'_1] & \mathbb{P}[P_2|P'_1] & \cdots & \mathbb{P}[P_n|P'_1] \\ \mathbb{P}[P_1|P'_2] & \mathbb{P}[P_2|P'_2] & \cdots & \mathbb{P}[P_n|P'_2] \\ \vdots & \vdots & \ddots & \cdots \\ \mathbb{P}[P_1|P'_m] & \mathbb{P}[P_2|P'_m] & \cdots & \mathbb{P}[P_n|P'_m] \end{bmatrix}.$$

Now consider the example below for a more explicit calculation.

Example 5. Let $\Omega = \{a, b, c, d, e, f\}$. Consider the following partitions

$$\mathcal{A}_1 = \{\{a\}, \{b, c\}, \{d, e\}, \{f\}\}, \qquad \mathcal{A}_2 = \{\{a, b\}, \{c, d\}, \{e, f\}\},$$

[2]Note that in principle, the domain of f_P should be $\mathcal{F}$. So we really mean $f_P \circ q$ when we wrote f_P above, where q is the quotient map. See Summary 3 for more details

and let $\mathcal{F}_1 = \sigma(\mathcal{A}_1)$ and $\mathcal{F}_2 = \sigma(\mathcal{A}_2)$ be the corresponding σ-algebras. It is easy to check that the dual basis for these spaces is

$$\mathbb{B}_1 = \{\mathbb{1}_{\{a\}}, \mathbb{1}_{\{b,c\}}, \mathbb{1}_{\{d,e\}}, \mathbb{1}_{\{f\}}\}, \quad \mathbb{B}_2 = \{\mathbb{1}_{\{a,b\}}, \mathbb{1}_{\{c,d\}}, \mathbb{1}_{\{e,f\}}\}$$

Let X be a $\mathcal{F}_1$ measurable random variable given as

$$X \equiv \begin{pmatrix} a & b & c & d & e & f \\ 1 & 3 & 3 & 5 & 5 & 7 \end{pmatrix}.$$

We can write X as

$$X = 1\mathbb{1}_{\{a\}} + 3\mathbb{1}_{\{b,c\}} + 5\mathbb{1}_{\{d,e\}} + 7\mathbb{1}_{\{f\}} := \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \end{bmatrix}.$$

We want to compute the conditional expectation $\mathbb{E}[X|\mathcal{F}_2]$. First, we need to compute the transformation matrix as in the example above. For instance

$$(M)_{1,1} = \mathbb{P}(\{a\}|\{a,b\}) = \frac{\mathbb{P}(\{a\} \cap \{a,b\})}{\mathbb{P}(\{a,b\})} = \frac{1/6}{2/6} = \frac{1}{2}.$$

And the full matrix will be given as

$$M = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

So we can write

$$\mathbb{E}[X|\mathcal{F}_2] = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}.$$

So we can write

$$\mathbb{E}[X|\mathcal{F}_2] = 2\mathbb{1}_{\{a,b\}} + 4\mathbb{1}_{c,d} + 6\mathbb{1}_{e,f}.$$

Now let $\mathcal{A}_3$ be yet another partition given by

$$\mathcal{A}_3 = \{\{a,b,c,d\}, \{e,f\}\},$$

and let $\mathcal{F}_3 = \sigma(\mathcal{A}_3)$ be the generated σ-algebra by this partition. We can also easily calculate $\mathbb{E}[X|\mathcal{F}_3]$ by first computing the transformation matrix

$$M' = \begin{bmatrix} 1/4 & 1/2 & 1/4 & 0 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix}.$$

Thus we will have

$$\mathbb{E}[X|\mathcal{F}_3] = \begin{bmatrix} 1/4 & 1/2 & 1/4 & 0 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

Thus

$$\mathbb{E}[X|\mathcal{F}_3] = 3\mathbb{1}_{\{a,b,c,d\}} + 6\mathbb{1}_{\{e,f\}}.$$

Viewing the conditioning operation as the matrices above gives us a very powerful level of abstraction to think about these notions. Then we can utilize many different tools from linear operator theory and spectral theory for different purposes in the setting of probability theory. Also

> Considering the example above, it is possible to construct the matrix M as soon as we know the σ-algebras of interest. I.e. the transformation matrices are purely determined by the σ-algebras and do not depend on any random variable.

Example 6. (Two interesting extreme cases) Let X be a random variable (i.e. a measurable function with respect to $\sigma(X)$), and let $\mathcal{G}$ be a σ-algebra. Then if

$$\mathbb{E}[X|\mathcal{G}] = X,$$

then we say that X is $\mathcal{G}$-measurable. And if

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X],$$

then we say that X is independent of $[G]$. We will return to the coin tossing example to demonstrate the shape of the transformation matrices in these two extreme cases. Let

$$\mathcal{A}_1 = \{A_0, A_1\}, \qquad \mathcal{A}_2 = \{A_{00}, A_{01}, A_{10}, A_{11}\}$$

as in Example 2. Equivalently, we can write these partitions as the subsets of $(0,1]$, i.e.

$$\tilde{\mathcal{A}}_1 = \{(0,1/2],(1/2,1]\},$$

and

$$\tilde{\mathcal{A}}_2 = \{(0,1/4],(1/4,2/4],(2/4,3/4],(3/4,1]\}.$$

Then we can write $M : \sigma(\mathcal{A}_1) \to \sigma(\mathcal{A}_2)$ as

$$M = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

And $M' : \sigma(\mathcal{A}_2) \to \sigma(\mathcal{A}_1)$ as

$$M' = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix}.$$

Consider the dyadic functions d_1 that is $\mathcal{F}_1$ measurable, and d_2 that is $\mathcal{F}_2$ measurable. We can write

$$d_1 = \mathbb{1}_{(0,1/2]}, \qquad d_2 = \mathbb{1}_{(1/4,2/4]} + \mathbb{1}_{(3/4,1]}.$$

Then it is easy to check that d_2 is independent of $\mathcal{F}_1$ because

$$\mathbb{E}[d_2|\mathcal{F}_1] = M' \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = \mathbb{E}[d_2].$$

Furthermore, it is easy to see that d_1 is $\mathcal{F}_2$-measurable because

$$\mathbb{E}[d_1|\mathcal{F}_2] = M \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix},$$

which is the same as d_1 as both the RHS of the equation above and d_1 assumes the same values on all of the point in Ω.