

Econ 8107 Recitation 1

Zach Swaziek

Sections adapted from Alex Wurdinger's and Amol Amol's notes,
as well as Sergio Ocampo-Diaz's math camp notes

January 17, 2022

Measure Theory Brush Up

In the following I will give a brief brush up on Measure Theory. To this end, I rely heavily on the first half of Chapter 7 of SLP together with some other sources¹. Understanding measure theory will be important in order to understand Stochastic Dynamic Programming, Markov processes and recursive competitive equilibria.

As an example, let's start with a simple stochastic growth model. Output is produced with $zf(k)$, where z is an i.i.d. random variable. The functional equation is then

$$v(k, z) = \max_{0 \leq k' \leq zf(k)} U(zf(k) - k') + \beta E[v(k', z')]$$

Our goal is to be able to deal with $E[v(k', z')]$ and importantly to be consistent about it. To see what this means, consider two examples.

- **Example 1:** Productivity z takes values in a finite or countably infinite set, i.e. $z \in Z = \{z_1, z_2, \dots, z_i, \dots\}$, with probabilities $(\pi_1, \pi_2, \dots, \pi_i, \dots)$ and $\pi_i \geq 0 \forall i$ as well as $\sum_i \pi_i = 1$. Then

$$v(k, z) = \max_{0 \leq k' \leq zf(k)} U(zf(k) - k') + \beta \sum_{i=1}^{\infty} \pi_i v(k', z'_i)$$

- **Example 2:** Productivity z takes values in the interval $z \in Z = [a, b]$. The probabilities are assigned to subsets by the continuous density function $\pi(z)$. Thus for any $[c, d] \subseteq [a, b]$ we have $Pr\{z \in [c, d]\} = \int_c^d \pi(z) dz$ with $\pi(z) \geq 0 \forall z \in [a, b]$ and $\int_a^b \pi(z) dz = 1$. Then

$$v(k, z) = \max_{0 \leq k' \leq zf(k)} U(zf(k) - k') + \beta \int_c^d v(k', z') \pi(z') dz'$$

¹Alex suggested "A User's Guide to Measure Theoretic Probability" by David Pollard. The first couple chapters of "Measures, Integrals, and Martingales" by Rene Schilling and Efe Ok's chapter B of "Measure and Probability Theory with Economic Applications" are also references for my additions.

The Dynamic Programming we have learned from chapter 4 in SLP can be adjusted to both of them. But neither is a special case of the other. We therefore want a unified theory that can deal with both (and more complicated cases, like e.g. a mix of both continuous and discrete distributions). We will thus be dealing with probability measures on state spaces, which includes exogenous and endogenous states.

Going back to example 1 in a more general way, let again the state space be finite or countably infinite $s \in S = \{s_1, s_2, \dots, s_i, \dots\}$. The probability distribution is then an assignment of numbers $(\pi_1, \pi_2, \dots)$ to the points $(s_1, s_2, \dots)$ with $\pi_i \geq 0 \forall i$ as well as $\sum_i \pi_i = 1$. Thus we have $\pi_i = Pr\{s = s_i\}$. So we are assigning probabilities to *elements* of S , which naturally also assigns probabilities to *subsets* of S : For any set $A \subseteq S$ define μ by

$$\mu(A) = \sum_{i \in I_A} \pi_i = Pr\{s \in A\} \quad \forall A \subseteq S,$$

where $I_A = \{i : s_i \in A\}$. Here, μ is an example of a *measure*, which we will define later. Let's look at the properties of μ in this example.

1. Domain: family of subsets including $\emptyset$ and S .
2. $\mu(A) \geq 0 \forall A \subseteq S$
3. $\mu(\emptyset) = 0$
4. $\mu(S) = 1$
5. For any $A_1, A_2, \dots$ of pairwise disjoint subsets

$$\mu(\cup_i A_i) = \sum_i \mu(A_i)$$

6. For any real valued function f , the expected value of f w.r.t. μ is

$$E(f) = \sum_i \mu(s_i) f(s_i)$$

If S is uncountable, however, this does not work that easily. We cannot define a function on the class of *all* subsets of S that has the additive property 5, since we want a countable additive measure. To hold onto this property, we have to restrict the class of subsets, which is done in the following.

Measurable Space

Given set S , on what collection of subsets $\mathcal{S}$ should we define measures (this includes probability measures)?

Definition 1 Let S be a set and $\mathcal{S}$ a collection of subsets. $\mathcal{S}$ is a σ -algebra if

1. $\emptyset, S \in \mathcal{S}$
2. $A \in \mathcal{S}$ implies $A^c = S \setminus A \in \mathcal{S}$
3. $A_n \in \mathcal{S}, n = 1, 2, \dots$, implies $\cup_{n=1}^{\infty} A_n \in \mathcal{S}$

In practice, we can think of the set S as a “sample space,” which contains all possible outcomes of an experiment. Any element A of the σ -algebra $\mathcal{S}$ can be thought of as an “event.”

Sergio Ocampo-Diaz’s math camp notes includes an intuitive explanation of why the idea of a σ -algebra are useful: “A σ -algebra imposes certain consistency to the family of sets under consideration. The way to interpret it is that only subsets of the σ -algebra can be known, hence measured.” Each property of the definition is important:

1. It is possible to know if none or at least one of the outcomes in the sample space S occurred.
2. If there is an event that can happen, it is also possible to determine if it doesn’t happen.
3. If multiple individual events are possible, then it is also possible to determine if at least one or all of the events occurred.

Now, let us go through a few examples.² Suppose we have a four-sided die. Then our set and sample space is $S = \{1, 2, 3, 4\}$. A trivial σ -algebra is simply:

$$\mathcal{S}_0 = \{\emptyset, S\}$$

Clearly, properties (1), (2), and (3) in the definition are satisfied.

The power set of S , $\mathcal{S}_1 = 2^S$, is also a σ -algebra (it is actually the “maximal” one).

$$\mathcal{S}_1 = 2^S = \left\{ \begin{array}{l} \{1\}, \{2\}, \{3\}, \{4\}, \{2, 3, 4\}, \{1, 3, 4\}, \{1, 2, 4\}, \{1, 2, 3\}, \\ \{1, 2\}, \{2, 3\}, \{3, 4\}, \{1, 3\}, \{1, 4\}, \{2, 4\}, \emptyset, S \end{array} \right\} \quad (1)$$

It is clear that property 1 and 3 from the definition of a σ -algebra is satisfied. Now let’s check property 2 for a few sets in $\mathcal{S}_1$:

²This example is adapted from Sergio Ocampo-Diaz’s math camp notes.

- Let $A_1 = \{1\} \in \mathcal{S}_1$. Then $A_1^c = S \setminus A_1 = \{2, 3, 4\}$. Clearly $A_1^c \in \mathcal{S}$. A similar argument can be made for $\{2\}, \{3\}, \{4\}$.
- Let $A_2 = \{2, 3, 4\}$. Then $A_2^c = S \setminus A_2 = \{1\}$. Clearly $A_2^c \in \mathcal{S}_1$. A similar argument can be made for the other sets in $\mathcal{S}_1$ with 3 elements.
- Let $A_3 = \{1, 2\}$. Then $A_3^c = S \setminus A_3 = \{3, 4\}$. Clearly $A_3^c \in \mathcal{S}_1$. A similar argument can be made for the other sets in $\mathcal{S}_1$ with 2 elements.

We see that property 2 is satisfied and that $\mathcal{S}_1$ is clearly a σ -algebra.

Does a σ -algebra necessarily have to be the power set? No! Consider the following collection of subsets of S :

$$\mathcal{S}_2 = \left\{ \{1, 3\}, \{2, 4\}, \emptyset, S \right\} \quad (2)$$

This could be interpreted as if the observer can only determine if the die rolled an even or an odd number. In general, for any $A \subset S$, $\{A, A^c, \emptyset, S\}$ is a σ -algebra.

Next, consider a collection of subsets of S that is **not** a σ -algebra:

$$\mathcal{S}_3 = \left\{ \{1\}, \{2\}, \{3\}, \{4\}, \emptyset, S \right\} \quad (3)$$

$\mathcal{S}_2$ fails to satisfy property 2 and 3. Now that we have a better idea of what a σ -algebra is, we can define a measurable space.

Definition 2 A pair $(S, \mathcal{S})$, where $\mathcal{S}$ is a σ -algebra, is a **measurable space**. Any set $A \in \mathcal{S}$ is an $\mathcal{S}$ -measurable set, or just a **measurable set** if $\mathcal{S}$ is known.

Examples of Measurable Spaces:

- Let S be any set and $\mathcal{S}$ consist only of $\emptyset$ and S .
- Let S be any set and $\mathcal{S}$ a collection of all subsets of S (the power set of S).
- Let S be any set and $\mathcal{S}$ consist of $A, A^c, \emptyset$, and S , where $A \subset S$.

Notice that the second example (power set) is a well defined σ -algebra even if S is uncountable (e.g. the interval $[a, b]$ from earlier). However, it is not useful for where we are going when S is uncountable (the power set is “too big a class to be useful” in these cases). We want to start building a more useful (and “smaller”) class $\mathcal{S}$.

For any collection of subsets $\mathcal{A}$ of S , consider the σ -algebras containing $\mathcal{A}$. At least one σ -algebra which contains $\mathcal{A}$ exists; the power set of S . Furthermore, the intersection of all σ -algebras containing $\mathcal{A}$ is a σ -algebra containing $\mathcal{A}$. This is the smallest σ -algebra that contains $\mathcal{A}$, and is called the **σ -algebra generated by $\mathcal{A}$** .

Definition 3 Let S be a nonempty set and $\mathcal{A}$ a subset of 2^S . The smallest σ -algebra on S that contains $\mathcal{A}$ (in the sense that this σ -algebra is included in any other σ -algebra that contains $\mathcal{A}$) is called the **σ -algebra generated by $\mathcal{A}$** .

Take $\mathbb{R}^1$. Let $\mathcal{A}$ be the collection of all open intervals of the form $(-\infty, b)$, (a, b) , $(a, +\infty)$, $(-\infty, +\infty)$. The smallest σ -algebra containing $\mathcal{A}$ is the Borel algebra for $\mathbb{R}^1$: $\mathcal{B}^1$. One can show that $\mathcal{B}^1$ is also the σ -algebra generated by all the closed or half open intervals. Finally, any set in $\mathcal{B}^1$ is a *Borel set*. Consider the following examples:

- Every open set U in $\mathbb{R}$ is a Borel set
- Take a closed set C . Then $\mathbb{R} \setminus C$ is open, $\mathbb{R} \setminus C \in \mathcal{B}$ and thus $C \in \mathcal{B}$
- Every countable union of open sets is a Borel set
- Every countable intersection of closed sets is a Borel set
- A single point in $\mathbb{R}$ is a Borel set, since

$$\{a\} = \cap_{n=1}^{\infty} [a, a + \frac{1}{n}]$$

When talking about any set S in $\mathbb{R}$, like $[a, b]$, we still want to use the notion of Borel sets. So we want $\mathcal{S}$ to be the appropriate restriction of Borel sets

$$\mathcal{B}_S = \{A \in \mathcal{B}^1 : A \subseteq S\}.$$

These are the Borel sets of S . We can also define a Borel algebra for any other metric space:

Definition 4 For any metric space (S, ρ) , the Borel algebra is the smallest σ -algebra containing the open balls, i.e. containing all sets of the form $A = \{s \in S : \rho(s, s_0) < \delta\}$, where $s_0 \in S$ and $\delta > 0$. In other words, it is the σ -algebra generated by the set of all open balls.

Measures

We now want to assign values of size or probability to all sets in $\mathcal{S}$ and again we want this to be consistent.

Definition 5 Let $(S, \mathcal{S})$ be a measurable space. A measure is a function $\mu : \mathcal{S} \rightarrow \mathbb{R}_+$ such that

1. $\mu(\emptyset) = 0$
2. $\mu(A) \geq 0 \quad A \in \mathcal{S}$
3. If $\{A_n\}_{n=1}^{\infty}$ is a sequence of countable, disjoint subsets in $\mathcal{S}$, then

$$\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$$

Definition 6 A measure space is a triplet $(S, \mathcal{S}, \mu)$

Note that if $\mu(S) = 1$, then μ is a probability measure and $(S, \mathcal{S}, \mu)$ is a probability space. A measurable set $A \in \mathcal{S}$ is an event and $\mu(A)$ is the probability of the event.

Examples of probability spaces from introduction:

- $(S, \mathcal{S}, \mu)$ where $S = \{s_1, s_2, \dots\}$, $\mathcal{S} = 2^S$, and measure μ defined by using probabilities $\{\pi_1, \pi_2, \dots\}$ on elements of S .
- $(S, \mathcal{S}, \mu)$ where $S = [a, b]$, $\mathcal{S} = \mathcal{B}_S$ (the Borel subsets of S), and measure μ is given by (in open, half-open, and closed intervals):

$$\mu([c, d]) = \mu([c, d)) = \mu((c, d]) = \mu((c, d)) = \int_c^d \pi(s) ds, \quad \forall a \leq c \leq d \leq b$$

Measurable Functions

Ultimately, we still want to talk about expectations of real valued functions, $E(f)$. All functions $f : S \rightarrow \mathbb{R}$ is too large to work with. Instead...

Definition 7 Let $(S, \mathcal{S})$ and $(T, \mathcal{T})$ be measurable spaces. Then the function $f : S \rightarrow T$ is measurable if the inverse image of every measurable set is measurable, i.e.

$$f^{-1}(A) = \{s \in S : f(s) \in A\} \in \mathcal{S} \quad \forall A \in \mathcal{T}$$

If $(T, \mathcal{T}) = (\mathbb{R}^1, \mathcal{B}^1)$ we can write

$$\{s \in S : f(s) \leq a\} \in \mathcal{S} \quad \forall a \in \mathbb{R}^1.$$

Note that if $f : S \rightarrow \mathbb{R}$ and the space is a probability space, then f is a random variable. Also note that any continuous function on $\mathbb{R}^l$ is $\mathcal{B}^l$ -measurable (Exercise 7.13).

In addition I present a definition and theorem needed for the stochastic Principle of Optimality in chapter 9.1.

Definition 8 Let $(S, \mathcal{S})$ and $(T, \mathcal{T})$ be measurable spaces, let Γ be a correspondence from S to T . Then the function $h : S \rightarrow T$ is a measurable selection from Γ if h is measurable and $h(s) \in \Gamma(s)$ $s \in S$

Theorem 1 (Measurable Selection Theorem) Let $(S \subseteq \mathbb{R}^l$ and $(T \subseteq \mathbb{R}^m$ be Borel sets with Borel subsets $\mathcal{S}, \mathcal{T}$. Let $\Gamma : S \rightarrow T$ be a non-empty, compact valued, usc correspondence, then there exists a measurable selection from Γ .

Markov Processes

This section is adapted from Amol Amol's notes.

Nearly every stochastic process that we examine in economics can be characterized as a Markov Process. Even i.i.d. is a special case of a Markov process. In words, a markov process is a stochastic process in which the probability of some future event depends only on the current realization, rather than the entire history of the process. This is a convenient characteristic to include in models when we want to preserve the recursive structure. The key feature of a Markov process is a transition matrix, or more generally a transition function.

Using markov processes allows us to add serial dependence to the shocks of our models. So instead of only working with functional equations of the form

$$v(x, z) = \sup_{y \in \Gamma(x, z)} \left[F(x, y, z) + \beta \int_Z v(y, z') \lambda(dz') \right]$$

where λ is a probability measure on $(Z, \mathcal{Z})$ that is independent of the current shock z , we can use functional equations of the form

$$v(x, z) = \sup_{y \in \Gamma(x, z)} \left[F(x, y, z) + \beta \int_Z v(y, z') Q(z, dz') \right]$$

where $Q(z, \cdot)$ is a probability measure on $\mathcal{Z}$ for each $z \in Z$. In this case, the current shock z affects the probability measure over next period's shock z' . We call the function Q a transition function.

Definition 9 Let $(Z, \mathcal{Z})$ be a measurable space. A **transition function** is a function $Q : Z \times \mathcal{Z} \rightarrow [0, 1]$, such that:

1. for each $z \in Z$, $Q(z, \cdot)$ is a probability measure on $(Z, \mathcal{Z})$,
2. for each $A \in \mathcal{Z}$, $Q(\cdot, A)$ is a $\mathcal{Z}$ -measurable function.

Definition 10 A stationary stochastic process $\{x_t\}$ is a markov process if for $n \geq 1$,

$$Pr(x_{t+n} | x_t, x_{t-1}, \dots, x_0) = Pr(x_{t+n} | x_t)$$

For a markov process with a transition function Q ,

$$Pr(x_{t+1} \in X | x_t) = Q(x_t, X)$$

In a discrete setting, the transition function is often characterized by a transition

matrix P where an element of the matrix p_{ij} is such that,

$$Pr(x_{t+1} = j | x_t = i) = p_{ij}$$

Now, let us define the expected value of a measurable function.

Definition 11 Let $(Z, \mathcal{Z})$ be a measurable space and Q a transition function. For any $\mathcal{Z}$ -measurable function f , define Tf by

$$(Tf)(z) = \int f(z')Q(z, dz'), \quad \text{all } z \in Z$$

We interpret $(Tf)(z)$ as the expected value of the function f next period, given that the current state is z .

The following two definitions describe stronger properties that we sometimes want to impose on T

Definition 12 A transition function Q on $(Z, \mathcal{Z})$ has the **Feller Property** if the associated operator T maps the space of bounded continuous functions on Z onto itself, i.e. $T : C(Z) \rightarrow C(Z)$

Definition 13 A transition function Q on $(Z, \mathcal{Z})$ is **monotone** if the associated operator T has the property that for every nondecreasing function $f : Z \rightarrow \mathbb{R}$, the function Tf is also nondecreasing.