36710-36752 Advanced Probability Overview Fall 2020

2. Measurable Functions, Random Variables, and Integration

Instructor: Alessandro Rinaldo

Associated reading: Sec 1.5 of Ash and Doléans-Dade; Sec 1.3 and 1.4 of Durrett.

1 Measurable Functions

1.1 Measurable functions

Measurable functions are functions that we can integrate with respect to measures in much the same way that continuous functions can be integrated "dx". Recall that the Riemann integral of a continuous function f over a bounded interval is defined as a limit of sums of lengths of subintervals times values of f on the subintervals. The measure of a set generalizes the length while elements of the σ-field generalize the intervals. Recall that a real-valued function is continuous if and only if the inverse image of every open set is open. This generalizes to the inverse image of every measurable set being measurable.

Definition 1 (Measurable Functions). *Let $(\Omega, \mathcal{F})$ and $(S, \mathcal{A})$ be measurable spaces. Let $f : \Omega \to S$ be a function that satisfies $f^{-1}(A) \in \mathcal{F}$ for each $A \in \mathcal{A}$. Then we say that f is* $\mathcal{F}/\mathcal{A}$-measurable. *If the σ-field's are to be understood from context, we simply say that f is measurable.*

Example 2. *Let $\mathcal{F} = 2^{\Omega}$. Then every function from Ω to a set S is measurable no matter what $\mathcal{A}$ is.*

Example 3. *Let $\mathcal{A} = \{\varnothing, S\}$. Then every function from a set Ω to S is measurable, no matter what $\mathcal{F}$ is.*

Proving that a function is measurable is facilitated by noticing that inverse image commutes with union, complement, and intersection. That is, $f^{-1}(A^C) = [f^{-1}(A)]^C$ for all A, and for arbitrary collections of sets $\{A_\alpha\}_{\alpha \in \aleph}$,

$$
\begin{aligned}
f^{-1}\left(\bigcup_{\alpha \in \aleph} A_\alpha\right) &= \bigcup_{\alpha \in \aleph} f^{-1}(A_\alpha), \\
f^{-1}\left(\bigcap_{\alpha \in \aleph} A_\alpha\right) &= \bigcap_{\alpha \in \aleph} f^{-1}(A_\alpha).
\end{aligned}
$$

Exercise 4. *Is the inverse image of a σ-field is a σ-field? That is, if $f : \Omega \to S$ and if $\mathcal{A}$ is a σ-field of subsets of S, then $f^{-1}(\mathcal{A})$ is a σ-field of subsets of Ω.*

Proposition 5. *If f is a continuous function from one topological space to another (each with Borel σ-field's) then f is measurable.*

The proof of this makes use of Lemma 7.

Definition 6 (σ-filed generated by measurable functions). *Let $f : \Omega \to S$, where $(S, \mathcal{A})$ is a measurable space. The σ-field $f^{-1}(\mathcal{A})$ is called the* σ-field generated by f. *The σ-field $f^{-1}(\mathcal{A})$ is sometimes denoted $\sigma(f)$.*

It is easy to see that $f^{-1}(\mathcal{A})$ is the smallest σ-field $\mathcal{C}$ such that f is $\mathcal{C}/\mathcal{A}$-measurable. We can now prove the following helpful result.

Lemma 7. *Let $(\Omega, \mathcal{F})$ and $(S, \mathcal{A})$ be measurable spaces and let $f : \Omega \to S$. Suppose that $\mathcal{C}$ is a collection of sets that generates $\mathcal{A}$. Then f is measurable if $f^{-1}(\mathcal{C}) \subseteq \mathcal{F}$.*

Exercise 8. *Prove Lemma 7.*

Exercise 9 (A common method to prove measurability). *Prove the following. Let $S = I\!R$ in Lemma 7. Let D be a dense subset of $I\!R$, and let $\mathcal{C}$ be the collection of all intervals of the form $(-\infty, a)$, for $a \in D$. To prove that a real-valued function is measurable, one need only show that $\{\omega : f(\omega) < a\} \in \mathcal{F}$ for all $a \in D$. Similarly, we can replace $< a$ by $> a$ or $\leq a$ or $\geq a$.*

Exercise 10. *Show that a monotone increasing function is measurable.*

Example 11. *Suppose that $f : \Omega \to \overline{I\!R}$ takes values in the extended reals. Then $f^{-1}(\{-\infty, \infty\}) = [f^{-1}((-\infty, \infty))]^C$. Also*

$$f^{-1}(\{\infty\}) = \bigcap_{n=1}^{\infty} \{\omega : f(\omega) > n\},$$

and similarly for $-\infty$. In order to check whether f is measurable, we need to see that the inverse images of all semi-infinite intervals are measurable sets. If we include the infinite endpoint in these intervals, then we don't need to check anything else. If we don't include the infinite endpoint, and if both infinite values are possible, then we need to check that at least one of $\{\infty\}$ or $\{-\infty\}$ has measurable inverse image.

Here are some simple properties of measurable functions.

Theorem 12 (Properties of Measurable Functions). *Let $(\Omega, \mathcal{F})$, $(S, \mathcal{A})$, and $(T, \mathcal{B})$ be measurable spaces.*

1. *If f is an extended real-valued measurable function and a is a constant, then af is measurable.*

2. *If $f : \Omega \to S$ and $g : S \to T$ are measurable, then $g(f) : \Omega \to T$ is measurable.*

3. *If f and g are measurable real-valued functions, then $f + g$ and fg are measurable.*

Proof: For $a = 0$, part 1 is trivial. Assume $a \neq 0$. Because

$$\{\omega : af(\omega) < c\} = \begin{cases} \{\omega : f(\omega) < c/a\} & \text{if } a > 0, \\ \{\omega : f(\omega) > c/a\} & \text{if } a < 0, \end{cases}$$

we see that af is measurable.

For part 2, just notice that $[g(f)]^{-1}(B) = f^{-1}(g^{-1}(B))$.

For part 3, let $h : \mathbb{R}^2 \to \mathbb{R}$ be defined by $h(x, y) = x + y$. This function is continuous, hence measurable. Then $f + g = h(f, g)$. We now show that $(f, g) : \Omega \to \mathbb{R}^2$ is measurable, where $(f, g)(\omega) = (f(\omega), g(\omega))$. To see that (f, g) is measurable, look at inverse images of sets that generate $\mathcal{B}^2$, namely sets of the form $(-\infty, a] \times (-\infty, b]$, and apply Lemma 7. We see that

$$(f, g)^{-1}((-\infty, a] \times (-\infty, b]) = f^{-1}((-\infty, a]) \cap g^{-1}((-\infty, b]),$$

which is measurable. Hence, (f, g) is measurable and $h(f, g)$ is measurable by part 2. Similarly fg is measurable. ■

You can also prove that f/g is measurable when the ratio is defined to be an arbitrary constant when $g = 0$. Similarly, part 3 can be extended to extended real-valued functions so long as care is taken to handle cases of $\infty - \infty$ and $\infty \times 0$.

Theorem 13. *Let $f_n : \Omega \to \mathbb{R}$ be measurable for all n. Then the following are measurable:*

1. $\limsup_{n\to\infty} f_n$*,*

2. $\liminf_{n\to\infty} f_n$*,*

3. $\{\omega : \lim_{n\to\infty} f_n$ *exists*$\}$.

4. $f = \begin{cases} \lim_{n\to\infty} f_n & \textit{where the limit exists,} \\ 0 & \textit{elsewhere.} \end{cases}$

Exercise 14. *Prove Theorem 13.*

2 Random Variables and Induced Measures

Definition 15 (Random Variables). *If $(\Omega, \mathcal{F}, P)$ is a probability space and $X : \Omega \to \overline{I\!R}$ is measurable, then X is called a* random variable*. In general, if $X : \Omega \to S$, where $(S, \mathcal{A})$ is a measurable space, we call X a* random quantity*.*

Example 16. *Let $\Omega = (0, 1)$ with the Borel σ-field, and let μ be Lebesgue measure, a probability. Let $Z(\omega) = \omega$ and $X(\omega) = \lfloor 2\omega \rfloor$. Both X and Z are random variables. X takes only two values, 0 and 1. It is easy to see that $\mu(\{\omega : X(\omega) = 1\}) = 1/2$. It is also easy to see that $\mu(\{\omega : Z(\omega) \le c\}) = c$ for $0 \le c \le 1$.*

Each measurable function from a measure space to another measurable space induces a measure on its range space.

Lemma 17 (Induced Measure). *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and let $(S, \mathcal{A})$ be a measurable space. Let $f : \Omega \to S$ be a measurable function. Then f induces a measure on $(S, \mathcal{A})$ defined by $\nu(A) = \mu(f^{-1}(A))$ for each $A \in \mathcal{A}$.*

Proof: Clearly, $\nu \ge 0$ and $\nu(\varnothing) = 0$. Let $\{A_n\}_{n=1}^{\infty}$ be a sequence of disjoint elements of $\mathcal{A}$. Then

$$\begin{aligned}
\nu\left(\bigcup_{n=1}^{\infty} A_n\right) &= \mu\left(f^{-1}\left[\bigcup_{n=1}^{\infty} A_n\right]\right) \\
&= \mu\left(\bigcup_{n=1}^{\infty} f^{-1}[A_n]\right) \\
&= \sum_{n=1}^{\infty} \mu(f^{-1}[A_n]) \\
&= \sum_{n=1}^{\infty} \nu(A_n).
\end{aligned}$$

■

The measure ν in Lemma 17 is called the *measure induced on $(S, \mathcal{A})$ from μ by f*. This measure is only interesting in special cases. First, if μ is a probability then so is ν.

Definition 18 (Probability Distribution). *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $(S, \mathcal{A})$ be a measurable space. Let $X : \Omega \to S$ be a random quantity. Then the measure induced on $(S, \mathcal{A})$ from P by X is called the* distribution of *X.*

We typically denote the distribution of X by μ_X. In this case, μ_X is a measure on the space $(S, \mathcal{A})$.

Exercise 19. *Consider the random variables in Example 16. The distribution of X is the Bernoulli distribution with parameter 1/2. The distribution of each Z is the uniform distribution on the interval* $(0,1)$*. Write down the corresponding* $(S,\mathcal{A},\nu)$ *in Lemma 17 for X and Z, respectively.*

If μ is infinite and f is not one-to-one, then the induced measure may be of no interest at all.

Exercise 20. *Either prove or create a counterexample to the following conjecture: If μ is a σ-finite measure on some measurable space $(\Omega,\mathcal{F})$, then for any measurable function f from Ω to S, the induced measure is also σ-finite.*

Example 21 (Jacobians). *If $\Omega = S = \mathbb{R}^k$ and f is one-to-one with a differentiable inverse, then ν is the measure you get from the usual change-of-variables formula using Jacobians.*

We have just seen how to construct the distribution from a random variable. Oddly enough, the opposite construction is also available. First notice that every probability ν on $(\mathbb{R},\mathcal{B}^1)$ has a distribution function F defined by $F(x) = \nu((-\infty,x])$. Now, we can construct a probability space $(\Omega,\mathcal{F},P)$ and a random variable $X : \Omega \to \mathbb{R}$ such that $\nu = P(X^{-1})$.[1] Indeed, just let $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}^1$, $P = \nu$, and $X(\omega) = \omega$.

3 Simple Functions

Definition 22 (Simple Functions). *A measurable function that takes at most finitely many values is called a* simple function.

Example 23. *Let $(\Omega,\mathcal{F})$ be a measurable space and let $A_1,\ldots,A_n$ be disjoint elements of $\mathcal{F}$, and let $a_1,\ldots,a_n$ be real numbers. Then $f = \sum_{i=1}^n a_i I_{A_i}$ defines a simple function since $f^{-1}((-\infty,a))$ is a union of at most finitely many measurable sets.*

Definition 24 (Canonical Representation of A Simple Function). *Let f be a simple function whose distinct values are $a_1,\ldots,a_n$, and let $A_i = \{\omega : f(\omega) = a_i\}$. Then $f = \sum_{i=1}^n a_i I_{A_i}$ is called the* canonical representation of f.

Lemma 25 (Monotone Approximation). *Let f be a nonnegative measurable extended real-valued function from Ω. Then there exists a sequence $\{f_n\}_{n=1}^\infty$ of nonnegative (finite) simple functions such that $f_n \le f$ for all n and $\lim_{n\to\infty} f_n(\omega) = f(\omega)$ for all ω.*

[1] **Notation:** When X is a random quantity and B is a set in the space where X takes its values, we use the following two symbols interchangeably: $X^{-1}(B)$ and $X \in B$. Both of these stand for $\{\omega : X(\omega) \in B\}$. Finally, for all B,

$$\mu_X(B) = \Pr(X \in B) = P(X^{-1}(B)).$$

Proof: For each n, define $A_{n,k} = f^{-1}([k/n, (k+1)/n))$ for $k = 0, 1, \ldots, n^2 - 1$ and $A_{n,\infty} = f^{-1}([n, \infty])$. Define $f_n(\omega) = \frac{1}{n} \sum_{k=0}^{n^2-1} k I_{A_{n,k}}(\omega) + n I_{A_{n,\infty}}(\omega)$. The proof is easy to complete now. ■

Lemma 25 says that each nonnegative measurable function f can be approximated arbitrarily closely from below by simple functions. It is easy to see that if f is bounded the approximation is uniform once n is greater than the bound.

Many theorems about real-valued functions are easier to prove for nonnegative measurable functions. This leads to the common device of splitting a measurable function f as follows.

Definition 26 (Splitting Measurable Functions). *Let f be a real-valued function. The* positive part *f^+ of f is defined as $f^+(\omega) = \max\{f(\omega), 0\}$. The* negative part *$f^-$ of f is $f^-(\omega) = -\min\{f(\omega), 0\}$.*

Notice that both the positive and negative parts of a function are nonnegative. It follows easily that $f = f^+ - f^-$. It is easy to prove that the positive and negative parts of a measurable function are measurable.

4 Integration

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. The definition of integral is done in three stages. We start with simple functions.

Definition 27 (Integral of Simple Functions). *Let $f : \Omega \to \overline{I\!R}^{+0}$ be a simple function with canonical representation $f(\omega) = \sum_{i=1}^n a_i I_{A_i}(\omega)$ The* integral of f with respect to μ *is defined to be $\sum_{i=1}^n a_i \mu(A_i)$. The integral is denoted variously as $\int f d\mu$, $\int f(\omega)\mu(d\omega)$, or $\int f(\omega) d\mu(\omega)$.*

The values $\pm\infty$ are allowed for an integral.

We use the following convention whenever necessary in defining an integral: $\pm\infty \times 0 = 0$. This applies to both the case when the function is 0 on a set of infinite measure and when the function is infinite on a set of 0 measure.

Proposition 28. *If $f \leq g$ and both are nonegative and simple, then $\int f d\mu \leq \int g d\mu$.*

Definition 29 (Integrable Functions). *We say that f is* integrable with respect to μ *if $\int f d\mu$ is finite.*

Example 30. *A real-valued simple function is always integrable with respect to a finite measure.*

The second step in the definition of integral is to consider nonnegative measurable functions.

Definition 31 (Integral for General Nonnegative Functions). *For nonnegative measurable f, define the* integral of f with respect to μ *by*

$$\int f d\mu = \sup_{\text{nonnegative finite simple } g \leq f} \int g d\mu.$$

That is, if f is nonnegative and measurable, $\int f d\mu$ is the least upper bound (possibly infinite) of the integrals of nonnegative finite simple functions $g \leq f$. Proposition 28 helps to show that Definition 27 is a special case of Definition 31, so the two definitions do not conflict when they both apply.

Finally, for arbitrary measurable f, we first split f into its positive and negative parts, $f = f^+ - f^-$.

Definition 32 (Integral of General Measurable Functions). *Let f be measurable. If either f^+ or f^- is integrable with respect to μ, we define the* integral of f with respect to μ *to be $\int f^+ d\mu - \int f^- d\mu$, otherwise the integral* does not exist.

It is easy to see that Definition 31 is a special case of Definition 32, so the two definitions do not conflict when they both apply. The reason for splitting things up this way is to avoid ever having to deal with $\infty - \infty$.

One unfortunate consequence of this three-part definition is that many theorems about integrals must be proven in three steps. One fortunate consequence is that, for most of these theorems, at least some of the three steps are relatively straightforward.

Definition 33 (Integration Over A Set). *If $A \in \mathcal{F}$, we define $\int_A f d\mu$ by $\int I_A f d\mu$.*

Proposition 34 (Monotonicity of Integral). *If $f \leq g$ and both integrals are defined, then $\int f d\mu \leq \int g d\mu$.*

Example 35 (Integration with Counting Measure). *Let μ be counting measure on a set Ω. (This measure is not σ-finite unless Ω is countable.) If $A \subseteq \Omega$, then $\mu(A) = \#(A)$, the number of elements in A. If f is a nonnegative simple function, $f = \sum_{i=1}^n a_i I_{A_i}$, then*

$$\int f d\mu = \sum_{i=1}^n a_i \#(A_i) = \sum_{\text{All } \omega} f(\omega).$$

It is not difficult to see that the equality of the first and last terms above continues to hold for all nonnegative functions, and hence for all integrable functions.

4.1 Riemann and Lebesgue Integrals

Before we study integration in detail, we should note that integration with respect to Lebesgue measure is the same as the Riemann integral in many cases.

Theorem 36. *Let f be a continuous function on a closed bounded interval $[a, b]$. Let μ be Lebesgue measure. Then the Riemann integral $\int_a^b f(x)dx$ equals $\int_{[a,b]} f d\mu$.*

Exercise 37. *Prove Theorem 36.*

Example 38. *A case in which the Riemann integral differs from the Lebesgue integral is that of "improper" Riemann integrals. These are defined as limits of Riemann integrals that are each defined in the usual way. For example, integrals of unbounded functions and integrals over unbounded regions cannot be defined in the usual way because the Riemann sums would always be ∞ or undefined. Consider the function $f(x) = \sin(x)/x$ over the interval $[1, \infty)$. It is not difficult to see that neither f^+ nor f^- is integrable with respect to Lebesgue measure. Hence, the integral that we have defined here does not exist. However, the improper Riemann integral is defined as $\lim_{T\to\infty} \int_1^T f(x)dx$, if the limit exists. In this case, the limit exists.*

Definition 39 (Expectation and Variance of Random Variables). *If P is a probability and X is a random variable, then $\int X dP$ is called the* mean *of X,* expected value *of X, or* expectation *of X and denoted $\mathrm{E}(X)$. If $\mathrm{E}(X) = \mu$ is finite, then the* variance *of X is $\mathrm{Var}(X) = \mathrm{E}[(X - \mu)^2]$.*

4.2 Properties of Integrals

Definition 40 (Almost sure/almost everywhere). *Suppose that some statement about elements of Ω holds for all $\omega \in A^C$ where $\mu(A) = 0$. Then we say that the statement holds* almost everywhere*, denoted a.e. $[\mu]$. If P is a probability, then almost everywhere is often replaced by* almost surely*, denoted a.s. $[P]$.*

Some simple properties of integrals include the following:

- For c a constant, $\int cf d\mu = c \int f d\mu$ if the latter exists.
- If $f \geq 0$, then $\int f d\mu \geq 0$.
- if $f = g$ for all $\omega \in \Omega$ except on a set with zero measure, and if either $\int f d\mu$ or $\int g d\mu$ exists, then so does the other, and they are equal. Similarly, if one of the integrals doesn't exist, then neither does the other.

Intuitively, the following two results seem obvious. In order to rigorously prove them, we need some additional tools on integrals and limits.

Theorem 41 (Additivity). *$\int (f + g)d\mu = \int f d\mu + \int g d\mu$ whenever at least two of them are finite.*

Theorem 42 (Change of Variable). *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and let $(S, \mathcal{A})$ be a measurable space. Let $f : \Omega \to S$ be a measurable function. Let ν be the measure induced on $(S, \mathcal{A})$ by f from μ. (See Definition 18.) Let $g : S \to \mathbb{R}$ be $\mathcal{A}/\mathcal{B}^1$ measurable. Then*

$$\int g d\nu = \int g(f) d\mu, \tag{1}$$

if either integral exists.

Exercise 43. *Let f and g be nonnegative simple functions defined on a measure space $(\Omega, \mathcal{F}, \mu)$. Then $\int (f+g) d\mu = \int f d\mu + \int g d\mu$.*