MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.436J/15.085J Fall 2008
Lecture 11 10/15/2008

ABSTRACT INTEGRATION — I

Contents

The material in these notes can be found in practically every textbook that includes basic measure theory, although the order with which various properties are proved can be somewhat different in different sources.

1 PRELIMINARIES

The objective of these notes is to define the integral $\int g\,d\mu$ [sometimes also denoted $\int g(\omega)\,d\mu(\omega)$] of a measurable function $g : \Omega \to \overline{\mathbb{R}}$, defined on a measure space $(\Omega, \mathcal{F}, \mu)$.

Special cases:

(a) If $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $X : \Omega \to \overline{\mathbb{R}}$ is measurable (i.e., an extended-valued random variable), the integral $\int X\,d\mathbb{P}$ is also denoted $\mathbb{E}[X]$, and is called the expectation of X.

(b) If we are dealing with the measure space $(\mathbb{R}, \mathcal{B}, \lambda)$, where $\mathcal{B}$ is the Borel σ-field and λ is the Borel measure, the integral $\int g\,d\lambda$ is often denoted as $\int g(x)\,dx$, and is meant to be a generalization of the usual integral encountered in calculus.

The program: We will define the integral $\int g\, d\mu$ for progressively general classes of measurable functions g:

(a) Finite nonnegative functions g that take finitely many values ("simple functions"). In this case, the integral is just a suitably weighted sum of the values of g.

(b) Nonnegative functions g. Here, the integral will be defined by approximating g from below by a sequence of simple functions.

(c) General functions g. This is done by decomposing g in the form $g = g_+ - g_-$, where g_+ and g_- are nonnegative functions, and letting $\int g\, d\mu = \int g_+\, d\mu - \int g_-\, d\mu$.

We will be focusing on the integral over the entire set Ω. The integral over a (measurable) subset B of Ω will be defined by letting

$$\int_B g\, d\mu = \int (1_B g)\, d\mu.$$

Here 1_B is the indicator function of the set g, so that

$$(1_B g)(\omega) = \begin{cases} g(\omega), & \text{if } \omega \in B, \\ 0, & \text{if } \omega \notin B. \end{cases}$$

Throughout, we will use the term "almost everywhere" to mean "for all ω outside a zero-measure subset of Ω." For the special case of probability measures, we will often use the alternative terminology "almost surely," or "a.s." for short. Thus, if X and Y are random variables, we have $X = Y$, a.s., if and only if $\mathbb{P}(X = Y) = \mathbb{P}\big(\{\omega : X(\omega) \neq Y(\omega)\}\big) = 0$.

In the sequel, an inequality $g \leq h$ between two functions will be interpreted as "$g(\omega) \leq h(\omega)$, for all ω." Similarly, "$g \leq h$, a.e.," means that "$g(\omega) \leq h(\omega)$, for all ω outside a zero-measure set." The notation "$g_n \uparrow g$" will mean that for every ω, the sequence $g_n(\omega)$ is monotone nondecreasing and converges to $g(\omega)$. Finally, "$g_n \uparrow g$, a.e.," will mean that the monotonic convergence of $g_n(\omega)$ to $g(\omega)$ holds for all ω outside a zero-measure set.

2 THE MAIN RESULT

Once the construction is carried out, integrals of nonnegative functions will always be well-defined. For general functions, integrals will be left undefined only when an expression of the form $\infty - \infty$ is encountered.

The following properties will turn out to be true, whenever the integrals or expectations involved are well-defined. On the left, we show the general version;

on the right, we show the same property, specialized to the case of probability measures. In property 8, the convention $0 \cdot \infty$ will be in effect when needed.

1.	$\int 1_B \, d\mu = \mu(B)$	$\mathbb{E}[1_B] = \mathbb{P}(B)$
2.	$g \geq 0 \ \Rightarrow \ \int g \, d\mu \geq 0$	$X \geq 0 \ \Rightarrow \mathbb{E}[X] \geq 0$
3.	$g = 0$, a.e. $\Rightarrow \ \int g \, d\mu = 0$	$X = 0$, a.s. $\Rightarrow \ \mathbb{E}[X] = 0$
4.	$g \leq h \ \Rightarrow \ \int g \, d\mu \leq \int h \, d\mu$	$X \leq Y \ \Rightarrow \mathbb{E}[X] \leq \mathbb{E}[Y]$
4′	$g \leq h$, a.e. $\Rightarrow \ \int g \, d\mu \leq \int h \, d\mu$	$X \leq Y$, a.s. $\Rightarrow \mathbb{E}[X] \leq \mathbb{E}[Y]$
5.	$g = h$, a.e. $\Rightarrow \ \int g \, d\mu = \int h \, d\mu$	$X = Y$, a.s. $\Rightarrow \mathbb{E}[X] = \mathbb{E}[Y]$
6.	[$g \geq 0$, a.e., and $\int g \, d\mu = 0$] $\Rightarrow \ g = 0$, a.e.	[$X \geq 0$, a.s., and $\mathbb{E}[X] \geq 0$] $\Rightarrow \ X = 0$, a.s.
7.	$\int (g + h) \, d\mu = \int g \, d\mu + \int h \, d\mu$	$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
8.	$\int (ag) \, d\mu = a \int g \, d\mu$	$\mathbb{E}[aX] = a\mathbb{E}[X]$
9.	$0 \leq g_n \uparrow g \Rightarrow \int g_n \, d\mu \uparrow \int g \, d\mu$	$0 \leq X_n \uparrow X, \Rightarrow \ \mathbb{E}[X_n] \uparrow \mathbb{E}[X]$
9′.	$0 \leq g_n \uparrow g$, a.e. $\Rightarrow \int g_n \, d\mu \uparrow \int g \, d\mu$	$0 \leq X_n \uparrow X$, a.s. $\Rightarrow \ \mathbb{E}[X_n] \uparrow \mathbb{E}[X]$
10.	$g \geq 0 \Rightarrow \ \nu(B) = \int_B g \, d\mu$ is a measure	[$f \geq 0$ and $\int f \, d\mathbb{P} = 1$] $\Rightarrow \ \nu(B) = \int_B f \, d\mathbb{P}$ is a probability measure

Property 9, and its generalization, property 9′, is known as the Monotone Convergence Theorem (MCT), and is a cornerstone of integration theory.

3 THE RIEMANN INTEGRAL

Before proceeding, it is worth understanding why the traditional integral encountered in calculus is not adequate for our purposes. Let us recall the definition of the (Riemann) integral $\int_a^b g(x) \, dx$. We subdivide the interval $[a, b]$ using a finite sequence $\sigma = (x_1, x_2, \ldots, x_n)$ of points that satisfy $a = x_1 < x_2 < \cdots < x_n = b$, and define

$$U(\sigma) = \sum_{i=1}^{n-1} \Big(\max_{x_i \leq x < x_{i+1}} g(x) \Big) \cdot (x_{i+1} - x_i),$$

$$L(\sigma) = \sum_{i=1}^{n-1} \Big(\min_{x_i \leq x < x_{i+1}} g(x) \Big) \cdot (x_{i+1} - x_i).$$

Thus, $U(\sigma)$ and $L(\sigma)$ are approximations of the "area under the curve g," from above and from below, respectively. We say that the integral $\int_a^b g(x)\,dx$ is well-defined, and equal to a constant c, if

$$\sup_{\sigma} L(\sigma) = \inf_{\sigma} U(\sigma) = c.$$

In this case, we also say that g is Riemann-integrable over $[a, b]$. Intuitively, we want the upper and lower approximants $U(\sigma)$ and $L(\sigma)$ to agree, in the limit of very fine subdivisions of the interval $[a, b]$.

It is known that if $g : \mathbb{R} \to \mathbb{R}$ is Riemann-integrable over every interval $[a, b]$, then g is continuous almost everywhere (i.e., there exists a set S of Lebesgue measure zero, such that g is continuous at every $x \notin S$). This is a severe limitation on the class of Riemann-integrable functions.

Example. Let Q be the set of rational numbers in $[0, 1]$. Let $g = 1_Q$. For any $\sigma = (x_1, x_2, \ldots, x_n)$, and every i, the interval $[x_i, x_{i+1})$ contains a rational number, and also an irrational number. Thus, $\max_{x_i \le x < x_{i+1}} g(x) = 1$ and $\min_{x_i \le x < x_{i+1}} g(x) = 0$. It follows that $U(\sigma) = 1$ and $L(\sigma) = 0$, for all σ, and $\sup_\sigma L(\sigma) \neq \inf_\sigma U(\sigma)$. Therefore 1_Q is not Riemann integrable. On the other hand if we consider a uniform distribution over $[0, 1]$, and the binary random variable 1_Q , we have $\mathbb{P}(1_Q = 1) = 0$, and we would like to be able to say that $\mathbb{E}[X] = \int_{[0,1]} 1_Q(x)\,dx = 0$. This indicates that a different definition is in order.

4 THE INTEGRAL OF A NONNEGATIVE SIMPLE FUNCTION

A function $g : \Omega \to \mathbb{R}$ is called **simple** if it is measurable, finite-value, and takes finitely many different values. In particular, a simple function can be written as

$$g(\omega) = \sum_{i=1}^{k} a_i 1_{A_i}(\omega), \qquad \forall\, \omega, \tag{1}$$

where k is a (finite) nonnegative integer, the coefficients a_i are nonzero and finite, and the A_i are measurable sets.

Note that a simple function can have several representations of the form (1). For example, $1_{[0,2]}$ and $1_{[0,1]} + 1_{(1,2]}$ are two representations of the same function. For another example, note that $1_{[0,2]} + 1_{[1,2]} = 1_{[0,1)} + 2 \cdot 1_{[1,2]}$. On the other hand, if we require the a_i to be distinct and the sets A_i to be disjoint, it is not hard to see that there is only one possible representation, which we will call the **canonical** representation. More concretely, in the canonical representation, we let $\{a_1, \ldots, a_k\}$ be the range of g, where the a_i are distinct, and $A_i = \{\omega \mid g(\omega) = a_i\}$.

Definition 1. *If g is a simple function, of the form* (1)*, its integral is defined by*
$$\int g\,d\mu = \sum_{i=1}^{k} a_i \mu(A_i).$$

Before continuing, we need to make sure that Definition 1 is sound, in the following sense. If we consider two alternative representations of the same simple function, we need to ensure that the resulting value of $\int g\,d\mu$ is the same. Technically, we need to show the following:

$$\text{if} \quad \sum_{i=1}^{k} a_i 1_{A_i} = \sum_{i=1}^{m} b_i 1_{B_i}, \quad \text{then} \quad \sum_{i=1}^{k} a_i \mu(A_i) = \sum_{i=1}^{m} b_i \mu(B_i).$$

This is left as an exercise for the reader.

Example. We have $1_{[0,2]} + 1_{[1,2]} = 1_{[0,1)} + 2 \cdot 1_{[1,2]}$. The first representation leads to $\mu([0,2]) + \mu([1,2]$, the second to $\mu([0,1)) + 2\mu([1,2])$. Using the fact $\mu([0,2]) = \mu([0,1)) + \mu([1,2])$ (finite additivity), we see that the two values are indeed equal.

For the case where the underlying measure is a probability measure $\mathbb{P}$, a simple function $X : \Omega \to \mathbb{R}$ is called a simple random variable, and its integral $\int X\,d\mathbb{P}$ is also denoted as $\mathbb{E}[X]$. We then have

$$\mathbb{E}[X] = \sum_{i=1}^{k} a_i \mathbb{P}(A_i).$$

If the coefficients a_i are distinct, equal to the possible values of X, and by taking $A_i = \{\omega \mid X(\omega) = a_i\}$, we obtain

$$\mathbb{E}[X] = \sum_{i=1}^{k} a_i \mathbb{P}(\{\omega \mid X(\omega) = a_i\}) = \sum_{i=1}^{k} a_i \mathbb{P}(X = a_i),$$

which agrees with the elementary definition of $\mathbb{E}[X]$ for discrete random variables.

Note, for future reference, that the sum or difference of two simple functions is also simple.

4.1 Verification of various properties for the case of simple functions

For the various properties listed in Section 2, we will use the shorthand "property S-A" and "property N-A", to refer to "property A for the special case of simple functions" and "property A for nonnegative measurable functions," respectively.

We note a few immediate consequences of the definition. For any $B \in \mathcal{F}$, The function 1_B is simple and $\int 1_B \, d\mu = \mu(B)$, which verifies **property 1.** In particular, when Q is the set of rational numbers and μ is Lebesgue measure, we have $\int 1_Q \, d\mu = \mu(Q) = 0$, as desired. Note that a nonnegative simple function has a representation of the form (1) with all a_i positive. It follows that $\int g \, d\mu \geq 0$, which verifies **property S-2**.

Suppose now that a simple function satisfies $g = 0$, a.e. Then, it has a canonical representation of the form $g = \sum_{i=1}^{k} a_i 1_{A_i}$, where $\mu(A_i) = 0$, for every i. Definition 1 implies that $\int g \, d\mu = 0$, which verifies **property S-3**.

Let us now verify the linearity **property S-7**. Let g and h be nonnegative simple functions. Using canonical representations, we can write

$$g = \sum_{i=1}^{k} a_i 1_{A_i}, \qquad h = \sum_{j=1}^{m} b_j 1_{B_j},$$

where the sets A_i are disjoint, and the sets B_j are also disjoint. Then, the sets $A_i \cap B_j$ are disjoint, and

$$g + h = \sum_{i=1}^{k} \sum_{j=1}^{m} (a_i + b_j) 1_{A_i \cap B_j}.$$

Therefore,

$$\begin{aligned}
\int (g+h) \, d\mu &= \sum_{i=1}^{k} \sum_{j=1}^{m} (a_i + b_j) \mu(A_i \cap B_j) \\
&= \sum_{i=1}^{k} a_i \sum_{j=1}^{m} \mu(A_i \cap B_j) + \sum_{j=1}^{m} b_j \sum_{i=1}^{k} \mu(A_i \cap B_j) \\
&= \sum_{i=1}^{k} a_i \mu(A_i) + \sum_{j=1}^{m} b_j \mu(B_j) \\
&= \int g \, d\mu + \int h \, d\mu.
\end{aligned}$$

(The first and fourth equalities follow from Definition 1. The third equality made use of finite additivity for μ.)

Property S-8 is an immediate consequence of Definition 1. We only need to be careful for the case where $\int g \, d\mu = \infty$ and $a = 0$. Using the convention $0 \cdot \infty$, we see that $ag = 0$, so that $\int (ag) \, d\mu = 0 = 0 \cdot \infty = a \int g \, d\mu$, and the property indeed holds. By combining properties S-7 and S-8, with $a = -1$ we see that, for simple functions g and h we have $(g - h)\, d\mu = \int g\, d\mu - \int h\, d\mu$.

We now verify **property S-4′**, which also implies **property S-4** as a special case. Suppose that $g \leq h$, a.e. We then have $h = g + q$, for a simple function q such that $q \geq 0$, a.e. In particular, $q = q_{-}q_{-}$, where $q_{+} \geq 0$, $q_{-} \geq 0$, and $q_{-} = 0$, a.e. Thus, $h = g + q_{-}q_{-}$. Note that q, q_{+}, and q_{-} are all simple functions. Using the linearity property S-7and then properties S-3, S-2, we obtain

$$\int h\, d\mu = \int g\, d\mu + \int q_{+}\, d\mu - \int q_{-}\, d\mu = \int g\, d\mu + \int q_{+}\, d\mu \geq \int g\, d\mu.$$

We next verify **property S-5**. If $g = h$, a.e., then we have both $g \leq h$, a.e., and $h \leq g$, a.e. Thus, $\int g\, d\mu \leq \int h\, d\mu$, and $\int g\, d\mu \geq \int h\, d\mu$, which implies that $\int g\, d\mu = \int h\, d\mu$.

We finally verify **property S-6**. Suppose that $g \geq 0$, a.e., and $\int g\, d\mu = 0$. We write $g = g_{+} - g_{-}$, where $g_{+} \geq 0$ and $g_{-} \geq 0$. Then, $g_{-} = 0$, a.e., and $\int g_{-}\, d\mu = 0$. Thus, using property S-7, $\int g_{+}\, d\mu = \int g\, d\mu + \int g_{-}\, d\mu = 0$. Note that g_{+} is simple. Hence, its canonical representation is of the form $g_{+} = \sum_{i=1}^{k} a_i 1_{A_i}$, with $a_i > 0$. Since $\sum_{i=1}^{k} a_i \mu(A_i) = 0$, it follows that $\mu(A_i) = 0$, for every i. From finite additivity, we conclude that $\mu(\cup_{i=1}^{k} A_i) = 0$. Therefore, $g_{+} = 0$, a.e., and also $g = 0$, a.e.

5 THE INTEGRAL OF A NONNEGATIVE FUNCTION

The integral of a nonnegative function g will be defined by approximating g from below, using simple functions.

Definition 2. *For a measurable function $g : \Omega \to [0, \infty]$, we let $S(g)$ be the set of all nonnegative simple (hence automatically measurable) functions q that satisfy $q \leq g$, and define*

$$\int g\, d\mu = \sup_{q \in S(g)} \int q\, d\mu.$$

We will now verify that with this definition, properties N-2 to N-10 are all satisfied. This is easy for some (e.g., property N-2). Most of our effort will be devoted to establishing properties N-7 (linearity) and N-9 (monotone convergence theorem).

The arguments that follow will make occasional use of the following continuity property for monotonic sequences of measurable sets B_i: If $B_i \uparrow B$, then

$\mu(B_i) \uparrow \mu(B)$. This property was established in the notes for Lecture 1, for the special case where μ is a probability measure, but the same proof applies to the general case.

5.1 Verification of some easy properties

Throughout this subsection, we assume that g is measurable and nonnegative.

Property N-2: For every $q \in S(g)$, we have $\int q\, d\mu \geq 0$ (property S-2). Thus, $\int g\, d\mu = \sup_{q \in S(g)} \int q\, d\mu \geq 0$.

Property N-3: If $g = 0$, a.e, and $0 \leq q \leq g$, then $q = 0$, a.e. Therefore, $\int q\, d\mu = 0$ for every $q \in S(g)$ (by property S-3), which implies that $\int g\, d\mu = 0$.

Property N-4: Suppose that $0 \leq g \leq h$. Then, $S(g) \subset S(h)$, which implies that

$$\int g\, d\mu = \sup_{q \in S(g)} \int q\, d\mu \leq \sup_{q \in S(h)} \int q\, d\mu = \int h\, d\mu.$$

Property N-5: Suppose that $g = h$, a.e. Let $A = \{\omega \mid g(\omega) = h(\omega)\}$, and note that the complement of A has zero measure, so that $q = 1_A q$, a.e., for any function q. Then,

$$\begin{aligned} \int g\, d\mu &= \sup_{q \in S(g)} \int q\, d\mu = \sup_{q \in S(g)} \int 1_A q\, d\mu \leq \sup_{q \in S(1_A g)} \int q\, d\mu \\ &\leq \sup_{q \in S(h)} \int q\, d\mu = \int h\, d\mu. \end{aligned}$$

A symmetrical argument yields $\int h, d\mu \leq \int g\, d\mu$.

Exercise: Justify the above sequence of equalities and inequalities.

Property N-4′: Suppose that $g \leq h$, a.e. Then, there exists a function g' such that $g' \leq h$ and $g = g'$, a.e. Property N-5 yields $\int g\, d\mu = \int g'\, d\mu$. Property N-4 yields $\int g'\, d\mu \leq \int h\, d\mu$. These imply that $\int g\, d\mu \leq \int h\, d\mu$.

Property N-6: Suppose that $g \geq 0$ but the relation $g = 0$, a.e., is not true. We will show that $\int g\, d\mu > 0$. Let $B = \{\omega \mid g(\omega) > 0\}$. Then, $\mu(B) > 0$. Let $B_n = \{\omega \mid g(\omega) > 1/n\}$. Then, $B_n \uparrow B$ and, therefore, $\mu(B_n) \uparrow \mu(B) > 0$. This shows that for some n we have $\mu(B_n) > 0$. Note that $g \geq (1/n)1_{B_n}$. Then, properties S-4, S-8, and 1 yield

$$\int g\, d\mu \geq \int \frac{1}{n} \cdot 1_{B_n}\, d\mu = \frac{1}{n} \int 1_{B_n}\, d\mu = \frac{1}{n} \mu(B_n) > 0.$$

Property N-8, when $a \geq 0$: If $a = 0$, the result is immediate. Assume that $a > 0$. It is not hard to see that $q \in S(g)$ if and only if $aq \in S(ag)$. Thus,

$$\int (ag)\,d\mu = \sup_{q \in S(ag)} \int q\,d\mu = \sup_{aq \in S(ag)} \int (aq)\,d\mu = \sup_{q \in S(g)} \int (aq)\,d\mu = a \int q\,d\mu.$$

5.2 Proof of the Monotone Convergence Theorem

We first provide the proof of property 9, for the special case where g is equal to a simple function q, and then generalize.

Let q be a nonnegative simple function, represented in the form $q = \sum_{i=1}^{k} a_i 1_{A_i}$, where the a_i are finite positive numbers, and the sets A_i are measurable and disjoint. Let g_n be a sequence of nonnegative measurable functions such that $g_n \uparrow q$. We distinguish between two different cases, depending on whether $\int q\,d\mu$ is finite or infinite.

(i) Suppose that $\int q\,d\mu = \infty$. This implies that $\mu(A_i) = \infty$ for some i. It follows that $\mu(A_i) = \infty$ for some i. Fix such an i, and let

$$B_n = \{\omega \in A_i \mid g_n(\omega) > a_i/2\}.$$

For every $\omega \in A_i$, there exists some n such that $g_n(\omega) > a_i/2$. Therefore, $B_n \uparrow A_i$. From the continuity of measures, we obtain $\mu(B_n) \uparrow \infty$. Now, note that $g_n \geq (a_i/2)1_{B_n}$. Then, using property N-4, we have

$$\int g_n\,d\mu \geq \frac{a_i}{2}\mu(B_n) \uparrow \infty = \int q\,d\mu.$$

(ii) Suppose now that $\int q\,d\mu < \infty$. Then, $\mu(A_i) < \infty$, for all $i \in S$. Let $A = \cup_{i=1}^{k} A_i$. By finite additivity, we have $\mu(A) < \infty$. Let us fix a positive integer r such that $1/r < a$. Let

$$B_n = \big\{\omega \in A \mid g_n(\omega) \geq q(\omega) - (1/r)\big\}.$$

We observe that $B_n \uparrow A$ and, by continuity, $\mu(B_n) \uparrow \mu(A)$. Since $\mu(A) = \mu(B_n) + \mu(A \setminus B_n)$, and $\mu(A) < \infty$, this also yields $\mu(A \setminus B_n) \downarrow 0$.

Note that $1_A q = q$, a.e. Using properties, S-5 and S-7, we have

$$\int q\,d\mu = \int 1_A q\,d\mu = \int 1_{B_n} q\,d\mu + \int 1_{A \setminus B_n} q\,d\mu. \qquad (2)$$

For $\omega \in B_n$, we have $g_n(\omega) + (1/r) \geq q(\omega)$. Thus, $g_n + (1/r)1_{B_n} \geq 1_{B_n} q$. Using properties N-4 and S-7, together with Eq. (2), we have

$$\begin{aligned} \int g_n \, d\mu + \int \frac{1}{r} 1_{B_n} \, d\mu \geq \int 1_{B_n} q \, d\mu &= \int q \, d\mu - \int 1_{A \setminus B_n} q \, d\mu \\ &\geq \int q \, d\mu - a\mu(A \setminus B_n). \end{aligned}$$

By taking the limit as $n \to \infty$, we obtain

$$\lim_{n\to\infty} \int g_n \, d\mu + \frac{1}{r}\mu(A) \geq \int q \, d\mu.$$

Since this is true for every $r > 1/a$, we must have

$$\lim_{n\to\infty} \int g_n \, d\mu \geq \int q \, d\mu.$$

On the other hand, we have $g_n \leq q$, so that $\int g_n \, d\mu \leq \int q \, d\mu$, and $\lim_{n\to\infty} \int g_n \, d\mu \leq \int q \, d\mu$.

We now turn to the general case. We assume that $0 \leq g_n \uparrow g$. Suppose that $q \in S(g)$, so that $0 \leq q \leq g$. We have

$$0 \leq \min\{g_n, q\} \uparrow \min\{g, q\} = q.$$

Therefore,

$$\lim_{n\to\infty} \int g_n \, d\mu \geq \lim_{n\to\infty} \int \min\{g_n, q\} \, d\mu = \int q \, d\mu.$$

(The inequality above uses property N-4; the equality relies on the fact that we already proved the MCT for the case where the limit function is simple.) By taking the supremum over $q \in S(g)$, we obtain

$$\lim_{n\to\infty} \int g_n \, d\mu \geq \sup_{q \in S(g)} \int q \, d\mu = \int g \, d\mu.$$

On the other hand, we have $g_n \leq g$, so that $\int g_n \, d\mu \leq \int g \, d\mu$. Therefore, $\lim_{n\to\infty} \int g_n \, d\mu \leq \int g \, d\mu$, which concludes the proof of property N-9.

To prove property $9'$, suppose that $g_n \uparrow g$, a.e. Then, there exist functions g_n' and g', such that $g_n = g'n$, a.e., $g = g'$, a.e., and $g_n' \uparrow g'$. By combining properties N-5 and N-9, we obtain

$$\lim_{n\to\infty} \int g_n \, d\mu = \lim_{n\to\infty} \int g_n' \, d\mu = \int g' \, d\mu = \int g \, d\mu.$$

5.3 Approximating g from below using "special" simple functions

Let g be a nonnegative measurable function. From the definition of $\int g\,d\mu$, it follows that there exists a sequence $q_n \in S(g)$ such that $\int q_n\,d\mu \to \int q\,d\mu$. This does not provide us with much information on the sequence q_n. In contrast, the construction that follows provides us with a concrete way of approximating $\int g\,d\mu$.

For any positive integer r, we define a function $g_r : \Omega \to \mathbb{R}$ by letting

$$g_r(\omega) = \begin{cases} r, & \text{if } g(\omega) \geq r \\ \dfrac{i}{2^r}, & \text{if } \dfrac{i}{2^r} \leq g(\omega) < \dfrac{i+1}{2^r}, \qquad i = 0, 1, \ldots, r2^r - 1 \end{cases}$$

In words, the function g_r is a quantized version of g. For every ω, the value of $g(\omega)$ is first capped at r, and then rounded down to the nearest multiple of 2^{-r}.

We note a few properties of g_r that are direct consequences of its definition.

(a) For every r, the function g_r is simple (and, in particular, measurable).

(b) We have $0 \leq g_r \uparrow g$; that is, for every ω, we have $g_r(\omega) \uparrow g(\omega)$.

(c) If g is bounded above by c and $r \geq c$, then $|g_r(\omega) - g(\omega)| \leq 1/2^r$, for every ω.

Statement (b) above gives us a transparent characterization of the set of measurable functions. Namely, a nonnegative function is measurable if and only if it is the monotonic and pointwise limit of simple functions. Furthermore, the MCT indicates that $\int g_r\,d\mu \uparrow \int g\,d\mu$, for this particular choice of simple functions g_r. (In an alternative line of development of the subject, some texts start by defining $\int g\,d\mu$ as the limit of $\int g_r\,d\mu$.)

5.4 Linearity

We now prove linearity **(property N-7)**. Let g_r and h_r be the approximants of g and h, respectively, defined in Section 5.3. Since $g_r \uparrow g$ and $h_r \uparrow h$, we have $(g_r + h_r) \uparrow (g + h)$. Therefore, using the MCT and property S-7 (linearity for simple functions),

$$\begin{aligned} \int (g+h)\,d\mu &= \lim_{r\to\infty} \int (g_r + h_r)\,d\mu \\ &= \lim_{r\to\infty} \Big(\int g_r\,d\mu + \int h_r\,d\mu \Big) \\ &= \lim_{r\to\infty} \int g_r\,d\mu + \lim_{r\to\infty} \int h_r\,d\mu \\ &= \int g\,d\mu + \int h\,d\mu. \end{aligned}$$

6 THE GENERAL CASE

Consider now a measurable function $g : \Omega \to \overline{\mathbb{R}}$. Let $A_+ = \{\omega \mid g(\omega) > 0\}$ and $A_- = \{\omega \mid g(\omega) < 0\}$; note that these are measurable sets. Let $g_+ = g \cdot 1_{A_+}$ and $g_- = -1_{A_-} g$; note that these are nonnegative measurable functions. We then have $g = g_+ - g_-$, and we define

$$\int g \, d\mu = \int g_+ \, d\mu - \int g_- \, d\mu.$$

The integral $\int g \, d\mu$ is well-defined, as long as we do not have both $\int g_+ \, d\mu$ and $\int g_- \, d\mu$ equal to infinity.

With this definition, verifying properties 3-8 is not too difficult, and there are no surprises. We decompose the functions g and h into negative and positive parts, and apply the properties already proved for the nonnegative case. The details are left as an exercise.

In order to prove the last property (property 10) one uses countable additivity for the measure μ, a limiting argument based on the approximation of g by simple functions, and the MCT. The detailed proof is left as an exercise for the reader.

6.436J / 15.085J Fundamentals of Probability
Fall 2008