MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.436J/15.085J Fall 2008
Lecture 13 10/22/2008

PRODUCT MEASURE AND FUBINI'S THEOREM

Contents

In elementary math and calculus, we often interchange the order of summation and integration. The discussion here is concerned with conditions under which this is legitimate.

1 PRODUCT MEASURE

Consider two probabilistic experiments described by probability spaces $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ and $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$, respectively. We are interested in forming a probabilistic model of a "joint experiment" in which the original two experiments are carried out independently.

1.1 The sample space of the joint experiment

If the first experiment has an outcome ω_1, and the second has an outcome ω_2, then the outcome of the joint experiment is the pair (ω_1, ω_2). This leads us to define a new sample space $\Omega = \Omega_1 \times \Omega_2$.

1.2 The σ-field of the joint experiment

Next, we need a σ-field on Ω. If $A_1 \in \mathcal{F}_1$, we certainly want to be able to talk about the event $\{\omega_1 \in A_1\}$ and its probability. In terms of the joint experiment, this would be the same as the event

$$A_1 \times \Omega_1 = \{(\omega_1, \omega_2) \mid \omega_1 \in A_1,\ \omega_2 \in \Omega_2\}.$$

Thus, we would like our σ-field on Ω to include all sets of the form $A_1 \times \Omega_2$, (with $A_1 \in \mathcal{F}_1$) and by symmetry, all sets of the form $\Omega_1 \times A_2$ (with $(A_2 \in \mathcal{F}_2)$. This leads us to the following definition.

Definition 1. *We define $\mathcal{F}_1 \times \mathcal{F}_2$ as the smallest σ-field of subsets of $\Omega_1 \times \Omega_2$ that contains all sets of the form $A_1 \times \Omega_2$ and $\Omega_1 \times A_2$, where $A_1 \in \mathcal{F}_1$ and $A_2 \in \mathcal{F}_2$.*

Note that the notation $\mathcal{F}_1 \times \mathcal{F}_2$ is misleading: this is not the Cartesian product of $\mathcal{F}_1$ and $\mathcal{F}_2$!

Since σ-fields are closed under intersection, we observe that if $A_i \in \mathcal{F}_i$, then $A_1 \times A_2 = (A_1 \times \Omega_2) \cap (\Omega_1 \cap A_2) \in \mathcal{F}_1 \times \mathcal{F}_2$. It turns out (and is not hard to show) that $\mathcal{F}_1 \times \mathcal{F}_2$ can also be defined as the smallest σ-field containing all sets of the form $A_1 \times A_2$, where $A_i \in \mathcal{F}_i$.

1.3 The product measure

We now define a measure, to be denoted by $\mathbb{P}_1 \times \mathbb{P}_2$ (or just $\mathbb{P}$, for short) on the measurable space $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$. To capture the notion of independence, we require that

$$\mathbb{P}(A_1 \times A_2) = \mathbb{P}_1(A_1)\mathbb{P}_2(A_2), \qquad \forall\, A_1 \in \mathcal{F}_1,\ A_2 \in \mathcal{F}_2. \tag{1}$$

Theorem 1. *There exists a unique measure $\mathbb{P}$ on $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$ that has property (1).*

Theorem 1 has the flavor of Carathéodory's extension theorem: we define a measure on certain subsets that generate the σ-field $\mathcal{F}_1 \times \mathcal{F}_2$, and then extend it to the entire σ-field. However, Caratheodory's extension theorem involves certain conditions, and checking them does take some nontrivial work. Various proofs can be found in most measure-theoretic probability texts.

1.4 Beyond probability measures

Everything in these notes extends to the case where instead of probability measures $\mathbb{P}_i$, we are dealing with general measures μ_i, under the assumptions that the measures μ_i are σ-**finite**. (A measure μ is called σ-finite if the set Ω can be partitioned into a countable union of sets, each of which has finite measure.)

The most relevant example of a σ-finite measure is the Lebesgue measure on the real line. Indeed, the real line can be broken into a countable sequence of intervals $(n, n+1]$, each of which has finite Lebesgue measure.

1.5 The product measure on $\mathbb{R}^2$

The two-dimensional plane $\mathbb{R}^2$ is the Cartesian product of $\mathbb{R}$ with itself. We endow each copy of $\mathbb{R}$ with the Borel σ-field $\mathcal{B}$ and one-dimensional Lebesgue measure. The resulting σ-field $\mathcal{B} \times \mathcal{B}$ is called the Borel σ-field on $\mathbb{R}^2$. The resulting product measure on $\mathbb{R}^2$ is called two-dimensional Lebesgue measure, to be denoted here by λ_2. The measure λ_2 corresponds to the natural notion of area. For example,

$$\lambda_2([a,b] \times [c,d]) = \lambda([a,b]) \cdot \lambda([c,d]) = (b-a) \cdot (d-c).$$

More generally, for any "nice" set of the form encountered in calculus, e.g., sets of the form $A = \{(x,y) \mid f(x,y) \leq c\}$, where f is a continuous function, $\lambda_2(A)$ coincides with the usual notion of the area of A.

Remark for those of you who know a little bit of topology – otherwise ignore it. We could define the Borel σ-field on $\mathbb{R}^2$ as the σ-field generated by the collection of open subsets of $\mathbb{R}^2$. (This is the standard way of defining Borel sets in topological spaces.) It turns out that this definition results in the same σ-field as the method of Section 1.2.

2 FUBINI'S THEOREM

Fubini's theorem is a powerful tool that provides conditions for interchanging the order of integration in a double integral. Given that sums are essentially special cases of integrals (with respect to discrete measures), it also gives conditions for interchanging the order of summations, or the order of a summation and an integration. In this respect, it subsumes results such as Corollary 1 at the end of the notes for Lecture 12.

In the sequel, we will assume that $g : \Omega_1 \times \Omega_2 \to \mathbb{R}$ is a measurable function. This means that for any Borel set $A \subset \mathbb{R}$, the set $\{(\omega_1, \omega_2) \mid g(\omega_1, \omega_2) \in A\}$ belongs to the σ-field $\mathcal{F}_1 \times \mathcal{F}_2$. As a practical matter, it is enough to verify that for any scalar c, the set $\{(\omega_1, \omega_2) \mid g(\omega_1, \omega_2) \leq c\}$ is measurable. Other than using this definition directly, how else can we verify that such a function g is measurable? The basic tools at hand are the following:

(a) continuous functions from $\mathbb{R}^2$ to $\mathbb{R}$ are measurable;

(b) indicator functions of measurable sets are measurable;

(c) combining measurable functions in the usual ways (e.g., adding them, multiplying them, taking limits, etc.) results in measurable functions.

Fubini's theorem holds under two different sets of conditions: (a) nonnegative functions g (compare with the MCT); (b) functions g whose absolute value has a finite integral (compare with the DCT). We state the two versions separately, because of some subtle differences.

The two statements below are taken verbatim from the text by Adams & Guillemin, with minor changes to conform to our notation.

Theorem 2. *Let $g : \Omega_1 \times \Omega_2 \to \mathbb{R}$ be a nonnegative measurable function. Let $\mathbb{P} = \mathbb{P}_1 \times \mathbb{P}_2$ be a product measure. Then,*

(a) For every $\omega_1 \in \Omega_1$, $g(\omega_1, \omega_2)$ is a measurable function of ω_2.

(b) For every $\omega_2 \in \Omega_2$, $g(\omega_1, \omega_2)$ is a measurable function of ω_1.

(c) $\int_{\Omega_2} g(\omega_1, \omega_2)\, d\mathbb{P}_2$ is a measurable function of ω_1.

(d) $\int_{\Omega_1} g(\omega_1, \omega_2)\, d\mathbb{P}_1$ is a measurable function of ω_2.

(e) We have

$$\begin{aligned}\int_{\Omega_1} \left[\int_{\Omega_2} g(\omega_1, \omega_2)\, d\mathbb{P}_2 \right] d\mathbb{P}_1 &= \int_{\Omega_2} \left[\int_{\Omega_1} g(\omega_1, \omega_2)\, d\mathbb{P}_1 \right] d\mathbb{P}_2 \\ &= \int_{\Omega_1 \times \Omega_2} g(\omega_1, \omega_2)\, d\mathbb{P}.\end{aligned}$$

Note that some of the integrals above may be infinite, but this is not a problem; since everything is nonnegative, expressions of the form $\infty - \infty$ do not arise.

Recall now that a function is said to be **integrable** if it is measurable and the integral of its absolute value is finite.

Theorem 3. *Let $g : \Omega_1 \times \Omega_2 \to \mathbb{R}$ be a measurable function such that*

$$\int_{\Omega_1 \times \Omega_2} |g(\omega_1, \omega_2)|\, d\mathbb{P} < \infty,$$

where $\mathbb{P} = \mathbb{P}_1 \times \mathbb{P}_2$.

(a) For almost all $\omega_1 \in \Omega_1$, $g(\omega_1, \omega_2)$ is an integrable function of ω_2.

(b) For almost all $\omega_2 \in \Omega_2$, $g(\omega_1, \omega_2)$ is an integrable function of ω_1.

(c) There exists an integrable function $h : \Omega_1 \to \mathbb{R}$ such that $\int_{\Omega_2} g(\omega_1, \omega_2)\, d\mathbb{P}_2 = h(\omega_1)$, a.s. (i.e., except for a set of ω_1 of zero $\mathbb{P}_1$-measure for which $\int_{\Omega_2} g(\omega_1, \omega_2)\, d\mathbb{P}_2$ is undefined or infinite).

(d) There exists an integrable function $h : \Omega_2 \to \mathbb{R}$ such that $\int_{\Omega_1} g(\omega_1, \omega_2)\, d\mathbb{P}_1 = h(\omega_2)$, a.s. (i.e., except for a set of ω_2 of zero $\mathbb{P}_2$-measure for which $\int_{\Omega_1} g(\omega_1, \omega_2)\, d\mathbb{P}_1$ is undefined or infinite).

(e) We have

$$\begin{aligned}\int_{\Omega_1} \Big[\int_{\Omega_2} g(\omega_1, \omega_2)\, d\mathbb{P}_2 \Big]\, d\mathbb{P}_1 &= \int_{\Omega_2} \Big[\int_{\Omega_1} g(\omega_1, \omega_2)\, d\mathbb{P}_1 \Big]\, d\mathbb{P}_2 \\ &= \int_{\Omega_1 \times \Omega_2} g(\omega_1, \omega_2)\, d\mathbb{P}.\end{aligned}$$

We repeat that all of these results remain valid when dealing with σ-finite measures, such as the Lebesgue measure on $\mathbb{R}^2$. This provides us with conditions for the familiar calculus formula

$$\int \int g(x, y)\, dx\, dy = \int \int g(x, y)\, dy\, dx.$$

In order to apply Theorem 3, we need a practical method for checking the integrability condition

$$\int_{\Omega_1 \times \Omega_2} |g(\omega_1, \omega_2)|\, d\mathbb{P} < \infty.$$

in Theorem 3. Here, Theorem 2 comes to the rescue. Indeed, by Theorem 2, we have

$$\int_{\Omega_1 \times \Omega_2} |g(\omega_1, \omega_2)|\, d\mathbb{P} = \int_{\Omega_1} \int_{\Omega_2} |g(\omega_1, \omega_2)|\, d\mathbb{P}_2\, d\mathbb{P}_1,$$

so all we need is to work with the right hand side, and integrate one variable at a time, possibly also using some bounds on the way.

Finally, let us note that all the hard work goes into proving Theorem 2. Theorem 3 is relatively easy to derive once Theorem 2 is available: Given a function g, decompose it into its positive and negative parts, apply Theorem 2 to each part, and in the process make sure that you do not encounter expressions of the form $\infty - \infty$.

3 Some cautionary examples

We give a few examples where Fubini's theorem does not apply.

3.1 Nonnegative and Integrability

Suppose both of our sample spaces are the nonnegative integers: $\Omega_1 = \Omega_2 = \{1, 2, \ldots, \}$. The σ-fields $\mathcal{F}_1$ and $\mathcal{F}_2$ will be all subsets of Ω_1 and Ω_2, respectively. Then, $\sigma(F_1 \times F_2)$ will be composed of all subsets of $\{1, 2, \ldots, \}^2$. Both P_1 and P_2 will be the counting measure, i.e. $P(A) = |A|$. This means that

$$\int_A g dP_1 = \sum_{a \in A} f(a), \quad \int_B h dP_2 = \sum_{b \in B} h(b), \quad \int_C f dP_1 \times P_2 = \sum_{c \in C} f(c).$$

Consider the function f defined by $f(m, m) = 1$, $f(m, m+1) = -1$, and $f = 0$ elsewhere. It is easier to visualize f with a picture:

$$\begin{array}{ccccc} 1 & -1 & 0 & 0 & \cdots \\ 0 & 1 & -1 & 0 & \cdots \\ 0 & 0 & 1 & -1 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{array}$$

So

$$\int_{\Omega_1} \int_{\Omega_2} f dP_1 dP_2 = \sum_n \sum_m f(n, m) = 0 \neq 1 = \sum_m \sum_n f(n, m) = \int_{\Omega_2} \int_{\Omega_1} f dP_2 dP_1$$

The problem is that the function we are integrating is neither nonnegative nor integrable.

3.2 σ-finiteness

Let $\Omega_1 = (0,1)$, and let $\mathcal{F}_1$ be the Borel sets, and P_1 be the Lebesgue measure. Let $\Omega_2 = (0,1)$ and $\mathcal{F}_2$ be the set of all subsets of $(0,1)$, and let P_2 be the counting measure.

Define $f(x,y) = 1$ if $x = y$ and 0 otherwise. Then,

$$\int_{\Omega_1}\int_{\Omega_2} f(x,y)dP_2(y)dP_1(x) = \int_{\Omega_1} 1dP_1(y) = 1,$$

but

$$\int_{\Omega_2}\int_{\Omega_1} f(x,y)dP_1(x)dP_2(y) = \int_{\Omega_2} 0dP_2(y) = 0.$$

The problem is that the counting measure on $(0,1)$ is not σ-finite.

4 An application

Let's apply Fubini's theorem to prove a generalization of a familiar relation from a beginning probability course.

Let X be a nonnegative integer-valued random variable. Then,

$$E[X] = \sum_{i=1}^{\infty} P(X \geq i).$$

This is usually proved as follows:

$$\begin{aligned} E[X] &= \sum_{i=1}^{\infty} ip(i) \\ &= \sum_{i=1}^{\infty}\sum_{k=1}^{i} p(i) \\ &= \sum_{k=1}^{\infty}\sum_{i=k}^{\infty} p(i) \\ &= \sum_{k=1}^{\infty} P(X \geq k) \end{aligned}$$

where the sum exchange is typically justified by an appeal to nonnegativity.

Let's rigourously prove a justification of this relation in the most general case. We will show that if X is a nonnegative random variable, then

$$E[X] = \int_0^{\infty} P(X \geq x)dx.$$

Proof: Define $A = \{(w, x) \mid 0 \leq x \leq X(w)\}$. Intuitively, if $\Omega = R$, then A would be the region under the curve $X(w)$. We argue that

$$E[X] = \int_\Omega X(w) dP = \int_\Omega \int_0^\infty 1_A(w, x) dx dP,$$

and now let's postpone the technical issues for a moment and interchange the integrals to get

$$\begin{aligned} E[X] &= \int_0^\infty \int_\Omega 1_A(w, x) dP dx \\ &= \int_0^\infty P(X \geq x) dx. \end{aligned}$$

Now let's consider the technical details necessary to make the above argument work. The integral interchange can be justified on account of the funciton 1_A being nonnegative, so we just need to show that all the functions we deal with are measurable. In particular we need to show that:

1. For fixed x, $1_A(w, x)$ is a measurable functions of w.
2. For fixed w, $1_A(w, x)$ is a measurable function of x.
3. $X(\omega)$ is a measurable function of ω.
4. $P(X \geq x)$ is a measurable function of x.
5. $1_A(w, x)$ is a measurable function of w and x.

and we do this as follows:

1. For fixed x, $1_A(w, x)$ is the indicator function of the set $X \geq x$, so it must be measurable.
2. For fixed w, $1_A(w, x)$ is the indicator function of the interval $[0, X(w)]$, so it is lebesgue measurable.
3. X is measurable since its a random variable.
4. Using the notation $Z(x) = P(X \geq x)$, observe that if $a \in \{Z \geq z\}$, then so is every number below a. It follows that the set $\{Z \geq z\}$ is always an interval, so it is Lebesgue measurable.

5. To show that 1_A is measurable, we argue that A is measurable.Indeed, the function $g : \Omega \times R \to R$ defined by $g(w, x) = X(w)$ is measurable, since for any Borel set B, $g^{-1}(B) = X^{-1}(B) \times (-\infty, +\infty)$. Similarly, $h : \Omega \times R \to R$ defined as $h(w, x) = x$ is measurable for the same reason. Since

$$A = \{g \geq h\} \bigcap \{h \geq 0\},$$

it follows that A is measurable.