Probability: Theory and Examples

Rick Durrett

Edition 4.1, April 21, 2013

Typos corrected, three new sections in Chapter 8.

4th edition published by Cambridge University Press in 2010

i

Probability: Theory and Examples

Rick Durrett

Edition 4.1, April 21, 2013

Typos corrected, three new sections in Chapter 8.

4th edition published by Cambridge University Press in 2010

ii

Contents

iii

iv

CONTENTS

Chapter 1

Measure Theory

In this chapter, we will recall some definitions and results from measure theory. Our

purpose here is to provide an introduction for readers who have not seen these concepts

before and to review that material for those who have. Harder proofs, especially those

that do not contribute much to one's intuition, are hidden away in the appendix.

Readers with a solid background in measure theory can skip Sections 1.4, 1.5, and

1.7, which were previously part of the appendix.

1.1

Probability Spaces

Here and throughout the book, terms being defined are set in boldface. We begin

with the most basic quantity. A probability space is a triple (Ω, F , P) where Ω

is a set of "outcomes," F is a set of "events," and P : F → [0, 1] is a function that

assigns probabilities to events. We assume that F is a σ-field (or σ-algebra), i.e., a

(nonempty) collection of subsets of Ω that satisfy

(i) if $A \in F$ then $A^c \in F$, and

(ii) if $A_i \in F$ is a countable sequence of sets then $\cup_i A_i \in F$.

Here and in what follows, countable means finite or countably infinite. Since $\cap_i A_i =$

$(\cup_i A_i^c)^c$, it follows that a σ-field is closed under countable intersections. We omit the

i

last property from the definition to make it easier to check.

Without P , (Ω, F) is called a measurable space, i.e., it is a space on which we

can put a measure. A measure is a nonnegative countably additive set function; that is, a function $\mu : \mathcal{F} \to \mathbb{R}$ with

(i) $\mu(A) \geq \mu(\emptyset) = 0$ for all $A \in \mathcal{F}$, and

(ii) if $A_i \in \mathcal{F}$ is a countable sequence of disjoint sets, then

$\mu(\cup_i A_i) =$

$\mu(A_i)$

i

If $\mu(\Omega) = 1$, we call μ a probability measure. In this book, probability measures

are usually denoted by P.

The next result gives some consequences of the definition of a measure that we

will need later. In all cases, we assume that the sets we mention are in $\mathcal{F}$.

1

2

CHAPTER 1. MEASURE THEORY

Theorem 1.1.1. Let μ be a measure on $(\Omega, \mathcal{F})$

(i) monotonicity. If $A \subset B$ then $\mu(A) \leq \mu(B)$.

(ii) subadditivity. If $A \subset \cup^\infty$

A

$\mu(A$

m=1

m then $\mu(A) \leq$

∞

m=1

m).

(iii) continuity from below. If $A_i \uparrow A$ (i.e., $A_1 \subset A_2 \subset \ldots$ and $\cup_i A_i = A$) then

$\mu(A_i) \uparrow \mu(A)$.

(iv) continuity from above. If $A_i \downarrow A$ (i.e., $A_1 \supset A_2 \supset \ldots$ and $\cap_i A_i = A$), with

$\mu(A_1) < \infty$ then $\mu(A_i) \downarrow \mu(A)$.

Proof. (i) Let $B - A = B \cap A^c$ be the difference of the two sets. Using + to denote

disjoint union, $B = A + (B - A)$ so

$\mu(B) = \mu(A) + \mu(B - A) \geq \mu(A)$.

(ii) Let A = A

and for n > 1, B

$- \cup_{m=1}^{n-1} A_m$. Since the B

n

$_n \cap A$, $B_1 = A_1$

$_n = A_n$

m=1

m

$_n$ are

disjoint and have union A we have using (ii) of the definition of measure, $B_m \subset A_m$,

and (i) of this theorem

∞

∞

$\mu(A) =$

$\mu(B_m) \leq$

$\mu(A_m)$

m=1

m=1

(iii) Let $B_n = A_n - A_{n-1}$. Then the B_n are disjoint and have $\cup_{m=1}^{\infty} B$

m=1

$m = A$,

$\cup_n$

B

$m=1$

$m = A_n$ so

∞

n

$\mu(A) =$

$\mu(B_m) = \lim$

$\mu(B_m) = \lim \mu(A_n)$

$n \to \infty$

$n \to \infty$

$m=1$

$m=1$

(iv) $A_1 - A_n \uparrow A_1 - A$ so (iii) implies $\mu(A_1 - A_n) \uparrow \mu(A_1 - A)$. Since $A_1 \supset B$ we have

$\mu(A_1 - B) = \mu(A_1) - \mu(B)$ and it follows that $\mu(A_n) \downarrow \mu(A)$.

The simplest setting, which should be familiar from undergraduate probability, is:

Example 1.1.1. Discrete probability spaces. Let Ω = a countable set, i.e., finite

or countably infinite. Let F = the set of all subsets of Ω. Let

$P(A) =$

$p(\omega)$ where $p(\omega) \geq 0$ and

$p(\omega) = 1$

$\omega \in A$

$\omega \in \Omega$

A little thought reveals that this is the most general probability measure on this space. In many cases when Ω is a finite set, we have $p(\omega) = 1/|\Omega|$ where $|\Omega|$ = the number of points in Ω.

For a simple concrete example that requires this level of generality consider the astragali, dice used in ancient Egypt made from the ankle bones of sheep. This die could come to rest on the top side of the bone for four points or on the bottom for three points. The side of the bone was slightly rounded. The die could come to rest on a flat and narrow piece for six points or somewhere on the rest of the side for one point. There is no reason to think that all four outcomes are equally likely so we need probabilities p_1, p_3, p_4, and p_6 to describe P .

To prepare for our next definition, we need

Exercise 1.1.1. (i) If F_i, $i \in I$ are σ-fields then $\cap_{i \in I} F_i$ is. Here $I = \emptyset$ is an arbitrary index set (i.e., possibly uncountable). (ii) Use the result in (i) to show if we are given a set Ω and a collection A of subsets of Ω, then there is a smallest σ-field containing A. We will call this the σ-field generated by A and denote it by $\sigma(A)$.

Let R^d be the set of vectors $(x_1, . . . x_d)$ of real numbers and R^d be the Borel sets, the smallest σ-field containing the open sets. When $d = 1$ we drop the superscript.

Example 1.1.2. Measures on the real line. Measures on (R, R) are defined by giving probability a Stieltjes measure function with the following properties:

(i) F is nondecreasing.

(ii) F is right continuous, i.e. $\lim_{y \downarrow x} F(y) = F(x)$.

Theorem 1.1.2. Associated with each Stieltjes measure function F there is a unique measure μ on (R, R) with $\mu((a, b]) = F(b) - F(a)$

$\mu((a, b]) = F(b) - F(a)$

(1.1.1)

When $F(x) = x$ the resulting measure is called Lebesgue measure.

The proof of Theorem 1.1.2 is a long and winding road, so we will content ourselves to describe the main ideas involved in this section and to hide the remaining details in the appendix in Section A.1. The choice of “closed on the right” in $(a, b]$ is dictated by the fact that if $b_n \downarrow b$ then we have

$\cap_n (a, b_n] = (a, b]$

The next definition will explain the choice of “open on the left.”

A collection S of sets is said to be a semialgebra if (i) it is closed under intersection, i.e., $S, T \in S$ implies $S \cap T \in S$, and (ii) if $S \in S$ then S^c is a finite disjoint union of sets in S. An important example of a semialgebra is

Example 1.1.3. S_d = the empty set plus all sets of the form

$(a_1, b_1] \times \cdots \times (a_d, b_d] \subset R^d$

where $-\infty \le a_i < b_i \le \infty$

The definition in (1.1.1) gives the values of μ on the semialgebra S_1. To go from semialgebra to σ-algebra we use an intermediate step. A collection A of subsets of Ω is called an algebra (or field) if $A, B \in A$ implies A^c and $A \cup B$ are in A. Since $A \cap B = (A^c \cup B^c)^c$, it follows that $A \cap B \in A$. Obviously a σ-algebra is an algebra. An example in which the converse is false is:

Example 1.1.4. Let $\Omega = Z$ = the integers. A = the collection of $A \subset Z$ so that A or A^c is finite is an algebra.

Lemma 1.1.3. If S is a semialgebra then $\bar{S}$ = {finite disjoint unions of sets in S}

is an algebra, called the algebra generated by S.

Proof. Suppose A = +iSi and B = +jTj, where + denotes disjoint union and we

assume the index sets are finite. Then A ∩ B = +i,j Si ∩ Tj ∈ ¯

S. As for complements,

if A = +iSi then Ac = ∩iSc. The definition of S implies Sc ∈ ¯

S. We have shown

i

i

that ¯

S is closed under intersection, so it follows by induction that Ac ∈ ¯

S.

Example 1.1.5. Let Ω = R and S = S1 then ¯

S1 = the empty set plus all sets of the

form

∪k (a

i=1

i, bi]

where − ∞ ≤ ai < bi ≤ ∞

Given a set function μ on S we can extend it to ¯

S by

n

μ (+n A

μ(A

i=1

i) =

i)