S.R. Athreya and V.S. Sunder

Measure and Probability

May 31, 2007

Dedicated, with respect, to the memory of Paul Halmos

Preface

An initial set of notes - essentially containing three of the first four chapters of this book - was prepared when such a course was given by the second author at the 'Mathematics Training and Talent Search Programmee' at the Fergusson College, Pune in 1993. These notes were re-used subsequently in courses given to students at the 'NURTURE' programme at the Institute of Mathematical Sciences (IMSc), as well as in a 'core course' on real analysis at IMSc. Later, when the second author had to give a variant of this course to beginning graduate students at the Chennai Mathematical Institute, some 'analysis' - meaning some basic material about L^p spaces, absolute continuity, etc. - seemed to be called for; and this was the genesis of parts of the last chapter of this book.

When it was subsequently suggested that there should be a serious 'probability component' to the book, for it to be useful, the first author was approached with the plea that he might help modify the book so that this lacuna might be rectified to some extent; and chapters 3, 5 and 6 were the result. The 'Appendix' was included almost as a second thought, with the hope that it may be a source for the 'uniniated student' to fill in some possible gaps in the prerequisites needed for reading this book. This appendix has been 'lifted' verbatim from an appendix written for similar reasons for the book [Sun]. The authors plead guilty to having yielded to the temptation of the ease and convenience of the 'cut-and-paste option' offered by word-processing; more importantly, they are very grateful to Rajendra Bhatia and the TRIM series for kindly granting the permission necessary to so 'lift' this material.

It is hoped that this book could be read by anybody with a bachelor's degree in Mathematics. Even this 'requirement' is not necessary if the prospective reader is blessed with a 'modicum of mathematical maturity'; an appendix has been supplied for precisely such a reader. We believe the entire contents of this book could be very comfortably covered in a two-semester course, while a reasonable one-semester course can be fashioned by selecting appropriate parts of the book according to the need of the student (or taste of the instructor).

The first chapter sets as its goal the construction of Lebesgue measure. After a brief discussion of the need for restricting oneself to a suitable class of 'measurable sets', we get down to definitions and basic properties of abstract σ-algebras (as well as algebras and monotone classes) of subsets of a given universal set Ω, and pass to measures defined on such algebras; the fundamental Caratheodory extension theorem is then stated and proved. Finally the existence (and uniqueness) of Lebesgue measure is deduced, and the existence of 'non-Lebesgue-measurable' sets demonstrated.

The second chapter begins by discussing measurable functions, and then establishes the fundamental proposition regarding approximability of positive measurable functions by simple functions; and goes on to then define the 'Lebesgue' integral of appropriate functions, and derive such basic results as the monotone and dominated convergence theorems, Fatou's lemma, etc; the chapter ends with a brief discussion of the notion 'almost everywhere'.

The third chapter introduces the reader to the probabilistic terminology and approach: to start with, the terminology of a random variable (on an abstract probability space) and its distribution (the 'push-forward' probability measure defined on the Borel subsets of the real line) are discussed. The crucial notion of '(stochastic) independence' of events (or random variables) is introduced, and the Borel-Cantelli Lemma and Kolmogorov Zero-One Law' proved. Some of the more standard distributions - both discrete (Bernouilli, Binomial, Poisson, etc.) and continuous (Uniform, Normal, etc.) - are decsribed. And a final section is devoted to conditional expectations and probabilities.

The common link running through the fourth chapter is 'measures on product spaces'. Section 4.1 proves the existence of products of σ-finite measures on arbitrary spaces, and goes on to prove the fundamental theorems of Tonelli and Fubini. The next section connects product measures and independent random variables through the notion of their 'joint-distribution' (an appropriate 'push-forward measure'). Section 4.3 is devoted to (vacuously) applying the Caratherodory extension theorem to construct interesting examples of probability measured on arbitrary products of finite sets, and paves the way for Markov chains (at least in the case of finite state space). Section 4.4 is a quick discussion of Kolmogorov's consistency theorem, and the exercises here introduces the reader to 'transition probability measures' and 'Markov processes' in general, as well as the 'standard Brownian motion' in particular.

The fifth chapter is devoted to proving two cornerstones of probability theory, *viz.,* the Central Limit Theorem and the Law of Large Numbers. The first two sections pave the way with preliminary results - such as uniqueness and inversion theorem for 'characteristic functions (or Fourier transforms) of distributions', and theorems of Slutsky, Skorohod, Polya and Scheffe on relations between various modes of convergence. Section 5.3 proves the Central Limit Theorem (for i.i.d. random variables with finite variance), after preparing the ground with the conitnuity theorem which lists reformulations of convergence in distribution. The final Section 5.4 derives the Strong Law of Large Numbers from an auxiliary result to the effect that the 'sample means' of a sequence of stationary random variables with finite moment converge almost surely to the conditional expectation with respect to the ‘invariant σ-algebra’.

The sixth chapter focuses on discrete time Markov chains on countable state spaces. We begin with an introduction to the basic notions of aperiodicity, irreducibility, transience, reccurence and stationarity. We then proceed to prove the two main limit theorems in the area - namely, convergence to stationarity for aperiodic irreducible chains and the renewal theorem. We also explicitly exhibit, following [FW] the stationary measure for an irreducible finite state space Markov chain. We conclude with a discussion of recurrence and transience properties for birth-death and queueing chains.

The seventh chapter addresses various topics typically covered in early graduate courses in analysis. Section 7.1 addresses ‘finite complex measures’ (as well as finite real measures, traditionally called signed measures), and establishes that such measures are necessarily of ‘finite total variation’ and are expressible as linear combinations of finite positive measures. Section 7.2, devoted to L^p-spaces, establishes Hölder’s inequality and the fact that L^p-spaces are Banach spaces. The Radon Nikodym theorem is proved (following von Neumann) in Section 7.3, and then used to prove the duality among L^p-spaces corresponding to conjugate indices, as well as the Hahn decomposition of signed measures, and Lebesgue-Nikodym theorem. Section 7.4 is a brief digression into the ‘change of variables formula’, while Section 7.5 establishes classical results such as the fundamental theorem of calculus (for absolutely continuous functions). Finally, Section 7.6 is concerned with the Riesz Representation Theorem which is stated and proved in three flavours: the compact metric case, the general compact Hausdorff case, and finally the locally compact Hausdorff case.

Most of the topics covered here are ’standard fare’, but several proofs of ‘standard theorems’ are possibly unusual and not too ‘standard’. This comment might apply to our treatment of the following results: the Strong Law of Large Numbers (see Theorem 5.4.4), the Hahn decomposition of signed measures (Proposition 7.3.7), and the Riesz Representation Theorem (Theorems 7.6.1, 7.6.7, 7.6.9). To be entirely honest, however, we learnt later that V.S. Varadarajan had also constructed a proof of the Riesz Representation Theorem which, like ours, is based on the Hahn-Banach Theorem; unfortunately that proof appeared long ago and in a not very easily located journal. Similarly, our proof of SLLN was based on a lecture by Michael Keane which the first author heard, and the basic idea of the proof can be found in [Kea].

It is a pleasure to record here our gratitude to The Institute of Mathematical Sciences and the Indian Statistical Institute for the wonderful facilities and atmosphere they have been providing us for the last several years. In particular, we would also like to thank Rajeeva Karandikar, S. Kesavan, Rahul Roy, K. Ramamurthy, and S. Ramasubramanian for various conversations, and S. Sundar for pointing out various loopholes in earlier versions of some proofs.

Siva R. Athreya
V.S. Sunder
May 31, 2007

Contents

Measure Theoretic Probability offers a very generalized view of probability. Using sets rather than distributions represented by either discrete or continuous functions, it allows for complex problems to be understood more simply if you can get past the rigorous math! So today we start looking at Measure Theoretic Probability from a view that is much less rigorous than your average graduate textbook, but hopefully will allow you to take away some of the treasures this approach has. Measure for Measure. Back Cover. Probability and Measure. Third Edition. PATRICK BILLINGSLEY.Â Probability and measure / Pa#rÃ®tk_Billingsley. â€"3rd ed. p. cm. â€" (Wiley series in probability and mathematical. statistics. Probability and mathematical statistics). "A Wiley-Interscience publication." Includes bibliographical references and index. ISBN 0-471-00710-2. 1. Probabilities. 2. Measure theory. I. Title. In mathematics, a probability measure is a real-valued function defined on a set of events in a probability space that satisfies measure properties such as countable additivity. The difference between a probability measure and the more general notion of measure (which includes concepts like area or volume) is that a probability measure must assign value 1 to the entire probability space. Can someone explain probability measure in simple words? This term has been hunting me for my life. Today I came across Kullback-Leibler divergence. The KL divergence between probability measure P and Q is defined byÂ So, is a probability measure just a probability density but a broader and fancier saying? Am I overlooking a simple concept, or is this topic just that hard? probability measure-theory information-theory. 3. Probability Measure. This section contains the final and most important ingredient in the basic model of a random experiment. If you are a new student of probability, skip the technical detials. Definitions and Interpretations. Suppose that we have a random experiment with sample space \((S, \mathscr S) \), so that \(S \) is the set of outcomes of the experiment and \(\mathscr S \) is the collection of events.