Syllabus for Advanced Probability II, Stochastic Processes 36-754

Cosma Shalizi

Spring 2006

This course is an advanced treatment of interdependent random variables and random functions, with twin emphases on extending the limit theorems of probability from independent to dependent variables, and on generalizing dynamical systems from deterministic to random time evolution. Familiarity with measure-theoretic probability (at the level of 36-752) is essential, but the emphasis will be on developing a sound understanding of the material, rather than on mathematical rigor.

The first part of the course will cover topics making up the basic toolkit of stochastic processes: random functions; stationary processes; Markov processes; the Wiener process; and the elements of stochastic calculus. These will be followed by slightly more unusual topics: ergodic theory, which extends the classical limit laws to dependent variables; the closely-related theory of Markov operators, including the stochastic behavior of deterministic dynamical systems (i.e., "chaos"); information theory, as it connects to statistical inference and to limiting distributions; large deviations theory, which gives rates of convergence in the limit laws; and multi-parameter (spatial or spatio-temporal) proccesses.

PREREQUISITES: Measure-theoretic probability, at the level of 36-752, is essential. So is familiarity with stochastic processes in elementary or intermediate probability.

GRADING: One to three problems will be assigned weekly. Some will be proofs, others simulation exercises based on the theorems. Students whose performance on homework is not adequate will have the opportunity to take an oral final exam in its place.

TEXTS The primary text will be the lecture notes. Kallenberg's *Foundations of Modern Probability* (2nd ed.) will also be used, especially as a source of problems. (It's on order at the CMU bookstore, but often cheaper online.) Students who took 36-752 last semester, when it used Ash and Doleans-Dade, will also find the last two chapters helpful.

Some material will also be drawn from the following books, free online:

- Robert M. Gray, *Probability, Random Processes, and Ergodic Properties*, `http://ee-www.stanford.edu/~gray/arp.html`

- Robert M. Gray, *Entropy and Information Theory*, `http://www-ee.stanford.edu/~gray/it.html`

- Sean P. Meyn and Richard L. Tweedie, *Markov Chains and Stochastic Stability*, `http://decision.csl.uiuc.edu/~meyn/pages/book.html`

- David M. Pollard, *Convergence of Stochastic Processes*, `http://www.stat.yale.edu/~pollard/1984book/`

Outline

Basics, 16–20 January Processes as indexed collections of random variables; random functions, sample paths, product σ-algebras and others; finite dimensional distributions, consistency; extension theorems, existence and uniqueness; "standard" spaces, functionals of sample paths

One-parameter Processes, Usually Functions of Time, 23–30 January Reminders about filtrations and stopping times; one-sided vs. two-sided processes; shift operators; time-evolution semi-groups; continuity of sample paths, continuous modifications, "cadlag" processes, "pure jump" processes; stationarity: strong, weak, conditional; functions of stationary processes; first measures of dependence and correlation; Gaussian processes; waiting times: hitting, first-passage, return/recurrence

Markov processes, 1–8 February Markov processes as generalizations of IID variables and of deterministic dynamical systems; the transition probability view; the evolution operator view; the generator of the time-evolution semi-group; the Markov property and the strong Markov property; the "martingale problem" and Feller processes; return times, recurrence, Harris recurrence; evolution equations for distributions (Fokker-Planck, Chapman-Kolmogorov, etc.); invariant distributions; eigenvalues and eigenvectors of evolution operators, the spectral gap; cadlag and pure-jump Markov processes (branching, birth-death, counting, Poisson); diffusions

The Wiener Process and Brownian Motion, 10–20 February The Wiener process defined by independent Gaussian increments and continuity; Wiener's construction of a probability measure on the space of continuous functions; nowhere differentiability of the sample paths; martingale property; Gaussian finite-dimensional distributions; weak convergence of random walks to Wiener processes, a.k.a. Donsker's theorem, a.k.a. the functional central limit theorem, a.k.a. the invariance principle; (*) the Wiener process and physical Brownian motion

A Foretaste of Empirical Process Theory, 22–24 February The Glivenko-Cantelli lemma; the empirical process; the Brownian bridge; the empirical central limit theorem for uniform random variables; quantile transformations; the empirical central limit theorem for general real-valued random variables

Stochastic Calculus and Diffusions, 27 February–6 March First approach to stochastic integrals via Euler's method; rigorous definition of integrals with respect to Wiener process; stochastic differential equations; "white noise"; diffusions as solutions of SDEs, generators of diffusions, evolution of probability density; Ito's formula; the Stratonovich interpretation of SDEs; some examples of SDEs (including Langevin equations and the Ornstein-Uhlenbeck process); SDEs and boundary-value ("Dirichlet") problems; Feynman-Kac formulae; (*) stochastic integrals with respect to other martingales

Ergodic Theory, 8–24 March Measure-preserving transformations and Markov operators; time averages; invariant sets and invariant measures; ergodic transformations; ergodic decomposition of non-ergodic processes; convergence of time averages; ergodicity as equality of time averages and expectations; the mean-square ergodic theorem, estimates of convergence rates; the maximal ergodic lemma; Birkhoff's almost-sure ergodic theorem; more on ergodic decompositions; the ergodic theorem for stationary processes; some necessary and sufficient conditions for ergodicity; (*) spectrum of a weakly stationary process; (*) Wiener-Khinchin theorem

Mixing, 27–31 March Definition and examples of mixing processes; mixing coefficients and rates, decay of correlations; deterministic mixing; central limit theorems and convergence of empirical measure resulting from mixing; exactness or asymptotic stability; correspondence between levels of distributional convergence and levels of ergodicity; (*) variations: geometric ergodicity, the Kingman-Liggett sub-additive ergodic theorem, the multiplicative ergodic theorem; (*) uniform ergodic theory

Information theory, 3–10 April Stochastic processes as information sources; Shannon entropy; relative entropy, a.k.a. Kullback-Leibler divergence; the Gibbs inequality; relative entropy, likelihood and Fisher information; the Rényi entropies and divergences; mutual information; the mutual information function as a measure of dependence; entropy and divergence rates; a glimpse of source coding; the asymptotic equipartition property (Shannon-MacMillan-Breiman theorem) for entropy and divergence; typical and jointly typical sequences; bounds on likelihood from AEP; partitions, symbolic dynamics, generating partitions; Kolmogorov-Sinai entropy (entropy rate) of a dynamical system; entropy rate as an invariant; isomorphism of Bernoulli processes and deterministic dynamics; behavior of the entropy under Markov operators; convergence of distributions in relative entropy

Large deviations, 12–19 April Large deviations principles and rate functions; IID large deviations, for empirical means, empirical measure, and for empirical process measures; the contraction principle; role of relative entropy as a rate function; application to hypothesis testing and parameter estimation; large deviations for Markov processes; (*) some applications of Markovian large deviations: information rates, effective action principles, the H-theorem of statistical mechanics; more general large deviations for stationary and ergodic processes; Gärtner-Ellis theorem; hypothesis testing more generally; (*) random perturbations of deterministic dynamics; (*) convergence of Markov processes to deterministic dynamics

Multi-parameter processes, 24–28 April Some general notions; Markov random fields and their key properties; a few examples of Markov random fields; Gibbs distributions and their key properties, including variational principles; random fields on graphs; the equivalence of the Gibbs and Markov properties; historical notes on the preceding theorem

Last week of lectures, 1–5 May Time permitting, we will use the last week to look briefly a more advanced topic, such as interacting particle systems, the use of recurrence times to learn the distribution of a process, or predictive representations, as dictated by class interest.