Lecture 1, 6.10.2022

MATH70031/M4P70 MARKOV PROCESSES

Professor N. H. BINGHAM
Mathematical Finance Section, 702 Weeks Hall;
n.bingham@imperial.ac.uk; 7594-2085

Place: HUXLEY 139
Times:
Thursday 10 - 11 am, Weeks 1 - 10 (6 Oct - 8 Dec)
Friday 9 - 10 am, Weeks 1 - 10 (7 Oct - 9 Dec)
Monday, 2-3 pm, Weeks 2 - 10 (10 Oct - 12 Dec)
Office hour: TBA
Class representative: To be selected. First task, to arrange possible times for the office hour. Suggestion:
(a) Get a class list, and a timetable grid (day of week x time of day). Ask in the Departmental Office to have these printed for you, and tell them why (or, DIY).
(b) Delete Tuesdays and Wednesdays. Pass it round the class, asking each person to put a small cross in any hour that would clash for them.
(c) Delete 9 am (I'm doing it once already, but that's enough!) Take your pick from any uncrossed squares. If there are none, take your pick from the squares with the smallest number of crosses.
(d) Tell me, and I'll announce it and stick it on the course website.

Main texts
[Nor] NORRIS, J. R., *Markov chains*, CUP, 1997.
[Wil] WILLIAMS, David, *Probability with martingales*, CUP, 1991.

Also useful
[HaiL] HAIRER, Martin and LI, Xue-Mei, *Markov processes*, 204 p. See the course website. [This is designed for private study, not for use in the lecture room. I am teaching the course in Xue-Mei's absence.]
[GriS] GRIMMETT, Geoffrey R. and STIRZAKER, David R., *Probability and random processes*, 3rd ed., 2001/4th ed., 2020, OUP.

The classics
[Chu] CHUNG, Kai-Lai, *Markov chains with stationary transition probabilities*, 2nd ed., Grundlehren math. Wiss. 104, Springer, 1967 (1st ed. 1960).
[Dyn] DYNKIN, E. B., *Theory of Markov processes*, Dover, p/b, 2006 (repr. Eng. tr. 1961, Russ. 1959).
[Fel] FELLER, W., *An introduction to probability theory and its applications*, Volume I, 3rd ed., Wiley, 1968.
[MeyT] MEYN, S. and TWEEDIE, R. L., *Markov chains and stochastic stability*, 2nd ed., p/b, CUP, 2009 (1st ed. 1993).

Cited in the text
Ch. 0
[Bil] P. BILLINGSLEY, *Probability and measure*. Wiley, 1986;
[Kal] O. KALLENBERG, *Foundations of modern probability*, 2nd ed., Springer, 2002 (1st ed. 1997).
[KinT] J. F. C. KINGMAN and S. J. TAYLOR, *Introduction to measure and probability*, CUP, 1966;
[Rud] Walter RUDIN, *Real and complex analysis*, McGraw-Hill, 2nd ed. (1974)/3rd ed. (1987).

Ch. 1.
[BinK] N. H. BINGHAM and R. KIESEL, *Risk-neutral valuation: Pricing and hedging of financial derivatives*, 2nd ed., Springer, 2004 (1st ed. 1998).
[Rev] D. REVUZ, *Markov chains*, North-Holland, 1975

Ch. 2.
[Bil61] P. BILLINGSLEY, *Statistical inference for Markov processes*. U. Chicago Press, 1961 (75p).
[CoxM] D. R. COX and H. D. MILLER, *The theory of stochastic processes*. Chapman & Hall, 1965 (p/b, 1977).
[DoyS] P. G. DOYLE and J. L. SNELL, *Random walks and electric networks*. Carus Math. Monographs 22, Math. Assoc. America, 1984.
[Ewe] W. J. EWENS, *Mathematical population genetics*, Springer, 1979.
[Kel] F. P. KELLY, *Reversibility and stochastic networks*, Wiley, 1979.
[KemS] J. G. KEMENY and J. L. SNELL, *Finite Markov chains*. Van Nostrand, 1960.

Ch. 3. [Nor], above.

Ch. 4.
[Asm] S. ASMUSSEN, *Applied prob. and queues*, 2nd ed., Springer, 2003.
[Bre] L. BREIMAN, *Probability*. Addison-Wesley, 1968.
[Øks] B. ØKSENDAL, *Stochastic differential equations: An introduction with applications*, 6th ed., Universitext, Springer, 2003 [5th ed. 1998].
[RevY] D. REVUZ and M. YOR, *Continuous martingales and Brownian motion*, 3rd ed. Grundl. math. Wiss. 293, Springer, 1999
[RogW] L. C. G. ROGERS and D. WILLIAMS, *Diffusions, Markov processes and martingales. Volume 1: Foundations* 2nd ed., Wiley, 1994; *Volume 2: Itô calculus*, Wiley, 1987.
[Ste] J. M. STEELE, *Stochastic calculus and financial applications*, Springer, 2001 (BM, Ch. 3).
[StrV] D. W. STROOCK and S. R. S. VARADHAN, *Multidimensional diffusions.* Grundl. math. Wiss. 233, Springer, 1979.

Table of contents

Time-line: Who did what when

Bernoulli, Daniel (1700 - 1782), Bernoulli-Laplace model of diffusion, 1769
Bernoulli, Jacob (Jacques) (1654-1705), *Ars conjectandi*, 1713 (posth.)
Boltzmann, Ludwig (1844 - 1906), H-theorem, thermodynamics, 1896
Brown, Robert (1773 - 1858), Brownian motion, 1828
Chung, Kai-Lai (1917 - 2009), the classic *Markov chains with stationary transition probabilities*, 1960, 1967
Clausius, Rudolf (1822 - 1888), entropy, First and Second Laws of Thermodynamics, 1865
Cramér, Harald (1893 - 1985), Cramér-Lundberg collective-risk model, 1920s; Cramér estimate of ruin in insurance, 1930
Daniell, Percy John (1889 - 1946), Daniell-Kolmogorov theorem (1918)
Ehrenfest, Paul (1880 - 1933) and Tatiana, Ehrenfest urn, 1907
Fisher, R. A. (Sir Ronald) (1890 - 1962), Wright-Fisher model in genetics, 1930
Frobenius, Georg (1849 - 1917), Perron-Frobeius theorem, 1908
Haar, Alfred (1885 - 1910), Haar functions, 1910; Haar measure, 1933
Kolmogorov, Andrei Nikolaevich (1903 - 1987), Daniell-Kolmogorov theorem; the classic *Grundbegriffe der Wahrscheinlichkeitsrechnung*, (1933)
Laplace, P.-S. de (1749 - 1827), Bernoulli-Laplace model of diffusion, 1812
Lundberg, Filip (1876 - 1965), Cramér-Lundberg collective risk model, 1920s
Markov, A. A. (1856 - 1922), Markov chains, 1908 [book, Calculus of probabilities (Russian), 2nd ed., 1908 (1st ed., 1900); German transl. *Wahrscheinlichkeitsrechnung*, 1912]
Meyer, P.-A. (1934 - 2003), filtrations, 1970s
Neumann, John von (1903-57), construction of the natural numbers $\mathbb{N}$, $\mathbb{N}_0$ in 1923
Paley, Raymond (R. E. A. C.) (1907-1933), Paley-Wiener-Zygmund (PWZ) theorem, 1933
Perron, Oskar (1880 - 1975), Perron-Frobenius theorem, 1907
Poisson, Siméon-Denis (1781 - 1840), Poisson distribution, 1837
Pólya, George (1897 - 1985), Pólya's theorem (§2.9)
Schauder, Juliusz (1899 - 1943), Schauder functions, 1927
Stirling, James (1692 - 1770), Stirling's formula, 1730.
Wiener, Norbert (1894 - 1964), Wiener process, 1923; Paley-Wiener-Zygmund theorem, 1933
Wright, Sewall (1889 - 1988), Wright-Fisher model in genetics, 1931
Zygmund, Antoni (1900 - 1992), Paley-Wiener-Zygmund theorem, 1933

Theorem (Conditional Mean Formula. For $\mathcal{B}$ any σ-field,

$$E[E[X|\mathcal{B}]] = E[X].$$

Proof. In the tower property, take $\mathcal{C}$ the trivial σ-field $\{\emptyset, \Omega\}$. This contains no information, so an expectation conditioning on it is the same as an unconditional expectation. The the tower property now gives

$$E[E[X|\mathcal{B}]\ |\{\emptyset, \Omega\}] = E[X|\{\emptyset, \Omega\}] = E[X]. \qquad //$$

Theorem (Conditional Variance Formula).

$$var(Y) = E[var(Y|X)] + var(E[Y|X]).$$

Proof. Recall $varX := E[(X - EX)^2]$. Expanding the square,

$$varX = E[X^2 - 2X.(EX) + (EX)^2] = E(X^2) - 2(EX)(EX) + (EX)^2 = E(X^2) - (EX)^2.$$

Conditional variances can be defined in the same way. Recall that $E(Y|X)$ is constant when X is known ($= x$, say), so can be taken outside an expectation over X, E_X say. Then

$$var(Y|X) := E(Y^2|X) - [E(Y|X)]^2.$$

Take expectations of both sides over X:

$$E_X var(Y|X) = E_X[E(Y^2|X)] - E_X[E(Y|X)]^2.$$

Now $E_X[E(Y^2|X)] = E(Y^2)$, by the Conditional Mean Formula, so the right is, adding and subtracting $(EY)^2$,

$$\{E(Y^2) - (EY)^2\} - \{E_X[E(Y|X)]^2 - (EY)^2\}.$$

The first term is $varY$, by above. Since $E(Y|X)$ has E_X-mean EY, the second term is $var_X E(Y|X)$, the variance (over X) of the random variable $E[Y|X]$ (random because X is). Combining, the result follows. //

Interpretation.

$varY$ = total variability in Y,

$E_X var(Y|X)$ = variability in Y not accounted for by knowledge of X,

$var_X E(Y|X)$ = variability in Y accounted for by knowledge of X:

variance = mean of conditional variance + variance of conditional mean. This is extremely useful in Statistics, in breaking down uncertainty, or variability, into its contributing components. There is a whole area of Statistics devoted to such Components of Variance.

Measure Theory

We assume that the audience has taken MATH50006 Lebesgue measure and integration, currently taught by Dr P.-F. Rodriguez.

We list below the results from Measure Theory that we shall need:

The framework of Lebesgue measure theory; σ-additivity, σ-algebras
Borel sets, non-measurable sets; properties (regularity, invariance)
Convergence theorems:
Lebesgue monotone convergence theorem (monotone convergence, MCT);
Lebesgue dominated convergence theorem (dominated convergence, DCT);
Fatou's lemma
General measure theory (similar to the canonical case of Lebesgue; often based on Carathéodory's extension theorem)
Uniform integrability
Absolute continuity
Lebesgue decomposition; Lebesgue differentiation theorem
Radon-Nikodym theorem

Suggested texts:
[KinT] Kingman and Taylor (above), Ch. 1 - 9;
[Wil] Williams (above), Ch. 1 - 8 and their Appendices;
[Rud] Rudin (above), Ch. 1-8.

Ch. 0. Probability and Measure Theory: Prerequisites and revision

What is probability? (This is the title of Ch. 10 of the book Kingman & Taylor [KinT] cited above. I'm a probabilist, by the way.)

This is (in some sense) a silly question. Use short words rather than long words when one can. Let's ask instead: What is chance? First, we have to have a common language in which to communicate. Ours is (happens to be) English. So, one could (should?) reply: I'm speaking to you in plain English. If you don't understand plain English, that's your problem, not mine.

We deliberately do not '*define probability*' (or chance, if you prefer short words) – we are not dictionary compilers. In non-mathematical terms, we have as good an idea of what chance is/means as of any other ordinary word. In mathematical terms: we have in Measure Theory just the mathematics we need to do the job: *a* probability (measure) is just *a measure of mass* 1; a probability (of an event – a measurable set) is the measure that probability measure gives to the set. That's it.

This passage was suggested by a statement (H&L, p.13) 'The 'randomness' describes the lack of complete information about the system' (we could have used 'randomness' instead of probability or chance above). This statement is worth considering here, before we engage with the main mathematical content of the course. Several comments:
1. *Tossing a coin.* A coin is a rigid body. We learn in (Newtonian) Dynamics how the motion of rigid body is determined by its initial conditions (add air resistance, if you like). We have no way to predict or describe the fine detail of how a coin is tossed – which is why the toss of a coin is routinely used to break symmetry, on occasions such as the start of a football match. As both captains and the referee would agree, arguing about Newtonian Dynamics is beside the point here. Life's too short: stick to the point, toss the coin, and get the match started.
2. *Chance in your life.* A young person has their life ahead of them, and most of the big choices still to make: of career and partner; what job to take, where to live, how many kids to have, etc. The 'information' concerning your life is always going to be incomplete until your life is over. You/we/I experience the uncertainties of life as chance, or randomness. I doubt whether most people would happily describe this as 'not randomness, but incomplete information as to what will happen before we die'.
3. *Mortality* (leading on from this). 'Call no man happy till he take his

happiness down to a quiet grave' (Aeschylus, *Oedipus Rex*, last line). My paraphrases of this: 'While you're alive, your vulnerable' (know anyone who's been in a life-changing accident, or any old person suffering from dementia?) Or (football again): 'You don't know the score until the game is over'.

Returning to 2 above: the only viewpoint from which I regard 'not randomness but incomplete information' as natural when applied to individual lives (yours, mine etc.) is that of an obituarist (I'm sensitive to this: I happen to have written lot of obituaries!)

Moral: that's enough of 'philosophising' for this course. From now on, we're going to use good, proper Mathematics: that of *measures of mass* 1, which we will call *probability measures*, and call their values *probabilities*. One can avoid these terms and speak instead of 'measures of mass 1' and 'values of this measure of mass 1' instead (not recommended).

Probability and Statistics

Statistics is an eminently practical and very useful subject (and the area of one of the four Sections in this Department). One way to think about Probability is as the theoretical underpinning of Statistics (we have probabilists in all four Sections). Another is bringing Measure Theory – abstract Pure Mathematics – 'to life', by applying it in the real world all around us. Random variables (measurable functions, below) belong to Probability.

Sampling. Is data random?

When we sample the value of a random variable drawn from a distribution (usually to study the distribution, as this is the only way we can get at it), what we get is a number (datum), and if we do it lots of times (the more the better), what we get is data. Data are still random in some sense: they are realised values of random variables, and if you (or someone else) re-sampled, you (or they) would get different values. In another sense, they are not random – you have them written down, in front of you. They are *numbers*, preferably lots of them; you have them, in some format (list, array etc.), stored in some way (computer, paper etc.) The statistician's job is to extract as much information from it as possible about where it came from – the distribution from which it was sampled. Statistics depends on Probability; Statistics is eminently useful and widely applied. Its new frontiers with 'Big Data' – Markov chain Monte Carlo (MCMC), data handling, Computer Science, Machine Learning etc. – are constantly growing.