LECTURE 4

Stochastic processes. A *stochastic process* is an indexed set of random variables. Usually the index set is $\mathbf{R}_+$ or $\mathbf{R}$ (continuous time) or $\mathbf{N}_+$ or $\mathbf{N}$ (discrete time). The plus just means non-negative real numbers or non-negative integers. If you have n random variables $X_1, \ldots, X_n$ then you can just think of it as a random n-vector or a random element of $\mathbf{R}^n$ and all the statistical information is contained in the distribution function

$$F(x_1, \ldots, x_n) = P(X_1 \leq x_1, \ldots, X_n \leq x_n).$$

Analogously a discrete stochastic process $X_1, X_2, \ldots$ can be thought of as a random sequence or a random element of $\prod_{i=1}^{\infty} \mathbf{R}$ and a continuous stochastic process $X_t,\ t \geq 0$ can be thought of as a random function or a random element of $\prod_{t\geq 0} \mathbf{R}$. The statistical information is contained in the family of finite dimensional distributions

$$F_{t_1,\ldots,t_n}(x_1, \ldots, x_n) = P(X_{t_1} \leq x_1, \ldots, X_{t_n} \leq x_n).$$

Another way of saying this is that the statistical distribution of a random sequence or a random function is determined by the collection of all distributions of finite subsets.

It is not always natural to have $\mathbf{R}$ as the state space. For example $X_t,\ t \geq 0$ could be the positions of a particle hopping around on some set S with a σ-field $\mathcal{F}$. Then the finite dimensional distributions can be written

$$\mu_{t_1,\ldots,t_n}(A_1, \ldots, A_n) = P(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n)$$

where the sets A_i are elements of $\mathcal{F}$.

Gaussian processes. A process is called Gaussian if all the finite dimensional distributions are Gausssians.

Martingales. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{F}_t,\ t \geq 0$ be a filtration (increasing family of sub-σ-fields), and $X_t,\ t \geq 0$ a stochastic process with $X_t \in L^1$ for all $t \geq 0$. X_t is a *martingale* with respect to $\mathcal{F}_t$ if for any $s \leq t$,

$$E[X_t \mid \mathcal{F}_s] = X_s \qquad a.s.$$

(If the equality is replaced by $\geq$ it is a submartingale and by $\leq$ it is a supermartingale.)

Note that it is automatic in the definition that X_t is measurable with respect to $\mathcal{F}_t$.

Markov chains. A sequence $X_0, X_1, \ldots$ is called a *Markov chain* if for any function g,

$$E[g(X_{n+1}) \mid X_1, \ldots, X_n) = E[g(X_{n+1}) \mid X_n].$$

In words: What you will do in the future depends only on where you are and not how you got there. Or: The future is independent of the past, given the present.

Examples. i. Random walk.

ii. (Ehrenfest model) There are two containers containing k balls. At each time step a ball is chosen at random and moved to the other container. X_i is the number of balls in the first container after the ith step.

iii. (Discrete time diffusions) A particle moves on the real line. If it is at x at time n then it makes a jump which is normally distributed with mean $\frac{1}{n}b(x)$ and variance $\frac{1}{n}a(x)$ where $b(x)$ and $a(x) \geq 0$ are some given functions.

Note that a Markov chain is completely determined by its *initial distribution*, i.e. the distribution of X_0, and the transition probabilities

$$P(X_{n+1} \in A \mid X_n = x).$$

This is because the finite dimensional distributions can be written in terms of these transition probabilities. In particular, by the Markov property, for any $n_0 \leq n_1 \leq n_2$ we have the *Chapman-Kolmogorov* equation,

$$P(X_{n_2} \in A \mid X_{n_0} = x_0) = \int P(X_{n_2} \in A \mid X_{n_1} = x_1) P(X_{n_1} \in dx_1 \mid X_{n_0} = x_0).$$

Applying this repeatedly we get

$$\begin{aligned} &P(X_{n_1} \in A_1, \ldots, X_{n_k} \in A_k) \\ &= \int_{A_{k-1}} \cdots \int_{A_1} \int P(X_0 \in dx_0) P(X_{n_1} \in dx_1 \mid X_0 = x_0) \cdots P(X_{n_k} \in A_k \mid X_{n_{k-1}} = x_{k-1}). \end{aligned}$$

If $P(X_{n+1} \in A \mid X_n = x)$ does not depend on n then we say the chain is *homogeneous* and we can write it as $p(x, A)$.

Discrete space. Note that all the integrations are summations if S is a discrete space. Also in this case the process is determined by all the $p(x, \{y\})$ which are then called p_{xy} and usually written as a matrix. Then the n step transition probabilities are given by $P(X_{m+n} = y \mid X_m = x) = p^{(n)}_{xy}$ where $p^{(n)}$ is the product of the matrix p with itself n times.

Suppose μ is a probability measure on S. Suppose we start the Markov chain with distribution μ as initial distribution, i.e. $P(X_0 = x) = \mu(\{x\}) = \mu_x$ and then look at the distribution of X_1. It is just μp or $p^*\mu$, i.e.

$$P(X_1 = x) = \sum_y \mu_y p_{yx}.$$

A distribution μ on S is called *invariant* if $\mu p = \mu$, i.e. for all $x \in S$,

$$\mu_x = \sum_y \mu_y p_{yx}.$$

An invariant measure is called *reversible* if for all $x, y \in S$,

$$\mu_x p_{xy} = \mu_y p_{yx}.$$

If you start a Markov chain with an invariant measure then the resulting process is *stationary*, i.e.

$$P(X_{n_1+m} \in A_1, \ldots, X_{n_k+m} \in A_k) = P(X_{n_1} \in A_1, \ldots, X_{n_k} \in A_k).$$

It just means you have no sense of absolute time.

If you start a Markov chain with a reversible measure then the resulting process looks the same backward and forward in time.

Warning. Not every Markov chain has an invariant measure let alone a reversible one. When they do, they may not be unique.

The *ergodic theorem* for finite state Markov chains says that if there is a number k such that $p^{(k)}_{x,y} > 0$ for all x and y, then there is a unique invariant measure μ and furthermore, for each y,

$$\lim_{n \to infty} p^{(n)}_{xy} = \mu_y.$$

What happens is that for a finite state Markov chain with a unique invariant measure it forgets in time what distribution it started in, and more and more looks like a stationary process starting from the invariant measure. In general though, things can get pretty complicated.

Suppose that X_n is a time homogeneous Markov chain with transition matrix p_{xy} and let g be a function. Then pg is just the function

$$pg_x = \sum_y p_{xy} g_y.$$

Then

$$M_n = g_{X_n} - \sum_{i=0}^{n-1} pg_{X_i} - g_{X_i}$$

is a martingale with respect to $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ since

$$E[g_{X_n} \mid \mathcal{F}_{n-1}] = E[g_{X_n} \mid X_{n-1}] = pg_{X_{n-1}}.$$

Markov processes. A process $X_t,\ t \geq 0$ is called a *Markov process* if for any function g and any $t \geq s$,

$$E[g(X_t) \mid X_u,\ 0 \leq u \leq s] = E[g(X_t) \mid X_s].$$

The process is uniquely determined by the initial distribution and the transition probabilities

$$p(s, x, t, A) = P(X_t \in A \mid X_s = x)$$

for $s < t$ with

$$\begin{aligned} &P(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n) \\ &= \int_{A_{n-1}} \cdots \int_{A_1} \int P(X_0 \in dx_0) p(0, x_0, t_1, dx_1) p(t_1, x_1, t_2, dx_2) \cdots p(t_{n-1}, x_{n-1}, t_n, A_n). \end{aligned}$$

The Chapman-Kolmogorov equations read,

$$p(s, x, t, A) = \int p(s, x, u, dy) p(u, y, t, A) \qquad \text{for } s \leq u \leq t.$$

If the process is time homogeneous then there is another way to write this. For any function u define

$$P_t u(x) = E_x[u(X_t)] = \int u(y)p(0,x,t,dy).$$

The subscript x on the expectation just notes that the process starts at x. All the P_t are *linear operators* which take functions to functions. It is clear that P_0 is just the identity. The Chapman-Kolmogorov equations become

$$P_{t+s} = P_t P_s.$$

Its called a *semigroup*.

Brownian motion. The special case where

$$p(s,x,t,dy) = \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{|x-y|^2}{2(t-s)}}$$

is called Brownian motion. What it means is that conditioned on X_u, $u \le s$, $X_t = X_s +$ an independent $\mathcal{N}(0, t-s)$.

Brownian motion and the heat equation. If we let

$$u(x,t) = P_t f(x)$$

then u solves the *heat equation*

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}, \qquad u(x,0) = f(x).$$

Independent increments. By construction Brownian motion has *independent increments*, i.e. if $0 < t_1 < \cdots < t_n$ then

$$X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$$

are independent.

Covariance. Clearly $E[X_t] = \int x \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} dx = 0$ and

$$E[(X_t)^2] = \int x^2 \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} dx = t.$$

The covariance is given by

$$E[X_t X_s] = E[(X_t - X_s)X_s] + E[(X_s)^2] = s$$

if $s \le t$ since $X_t - X_s$ is independent of X_s and mean 0.

Brownian motion as the limit of random walks. Let $Y_1, Y_2, \ldots$ be a simple symmetric random walk. Consider the process

$$X_t^n = \frac{1}{\sqrt{n}} Y_{\lfloor nt \rfloor}.$$

Here $\lfloor x \rfloor$ is the greatest integer $\leq x$. X_t^n is a random walk which jumps up or down $1/\sqrt{n}$ every time $1/n$. Let us look at the process at times $t_1, \ldots, t_m$. By the central limit theorem, for large n, the m-dimensional random vector

$$X_{t_1}^n, \ldots, X_{t_m}^n$$

is approximately Gaussian. What is its covariance? Well,

$$E[(X_t^n)^2] = E[(\frac{1}{\sqrt{n}} Y_{\lfloor nt \rfloor})^2] = t + O(1/n)$$

and

$$E[(X_t^n - X_s^n) X_s^n] = O(1/n)$$

so in the limit we have

$$\lim_{n \to \infty} E[X_t^n X_s^n] = \min(t, s)$$

which says that in the limit we have a Brownian motion. More precisely, what we showed is that if $F^n_{t_1,\ldots,t_m}(x_1, \ldots, x_m)$ are the finite dimensional distributions of X_t^n, $t \geq 0$, and $F_{t_1,\ldots,t_m}(x_1, \ldots, x_m)$ are the finite dimensional distributions of Brownian motion X_t, $t \geq 0$, then

$$\lim_{n \to \infty} F^n_{t_1,\ldots,t_m}(x_1, \ldots, x_m) = F_{t_1,\ldots,t_m}(x_1, \ldots, x_m).$$

PROBLEMS.

1. Show that for a Markov chain on a discrete state space the n step transition probabilities are obtained from the one step transition probabilities by multiplying the matrix by itself n times.
2. Let $X_1, X_2, \ldots$ be a Markov chain on the two state space $S = \{0, 1\}$. Let the transition matrix be $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Find all the invariant measures, check for reversibility and determine conditions so that the ergodic theorem holds .
3. Let X_i, $i = 1, 2, \ldots$ be a simple random walk on the integers $\{1, 2, \ldots, n\}$ with probability p to step $+1$ and $1-p$ to step -1. Write down the transition matrix assuming
   i. Absorbing barriers: When the walk gets to 1 or n it stays there forever.
   ii. Reflecting barriers: When the walk gets to 1 or n attempted jumps to 0 or $n+1$ are interpreted as staying put.
   iii. Periodic boundary conditions: We just pretend $n + 1 = 1$ and $0 = n$.
4. Write down the transition probabilities for the Ehrenfest model and show that the Binomial distribution is invariant and reversible.