Lecture 1.

One of the problems in describing and studying properties of special classes of stochastic processes is to find a convenient way of parametrizing them. The general way of describing stochastic processes by a consistent family of finite dimensional distributions, satisfying suitable additional conditions to provided regularity of sample paths is useful only in special cases. The finite dimensional distributions have to be provided. The Gaussian ones are natural and a Gaussian process can be specified by its mean and covariance. The only other large class is diffusion processes for which the finite dimensional distributions can be specified in terms of fundamental solutions of certain parabolic partial differential equations.

A convenient way of describing a discrete time stochastic process is through succesive conditional distributions, i.e. $\{P_n(dx_n|x_0, x_1, \ldots, x_{n-1})\}$. This has the advantage that if the index set is really time, this decribes a model for the evolution of the process.

The continuos time analog of this is not so obvious. In the deterministic case a continuos evolution can be described in the simplest case by an ordinary differential equation, the discrete anlog of which is a recurrence relation. If one thinks of the ODE $\dot{x} = b(x)$ as describing a an approximate recurrence of the form $x(t+h) \simeq x(t) + hb(x(t))$, then in the stochastic case we are looking for an approximate way of normalizing the conditional distribution $\mu_{t,h,\omega}(A) = P[x(t+h) - x(t) \in A | x(s) : 0 \le s \le t]$. If one thinks of the ODE as describing a vector filed that is tangent to the curve, then one has to define some sort of tangent to a stochastic process. Since the tangent is a blowup of a small difference we need to blow up the samll distribution $\mu_{t,h,\omega}$. This is done by convoluting it with itself $[\frac{1}{h}]$ times. Note that in the deterministic case this is essentially the same as dividing by h or adding it to itself $[\frac{1}{h}]$ times. In the limit the high convolution, if it has a limit will converge to an infinitely divisible distribution $D_{t,\omega}$ which can be called the tangent.

Example 1. If $x(t)$ is Brownian motion, then $\mu_{h,t}$ is Gaussian with mean 0 and vaiance h and the limiting tangent is the normal distribution with zero mean and variance one. It is constant, i.e. it is independent of t and ω. Processes with constant tangent are like staright lines and are in fact processes with homogeneous independent increments. Another example is the Posson process $N(t)$ for which the tangent is of course the Poisson distribution.

Example 2: Markov processes have conditionals $\mu_{h,t,\omega}$ that depend on ω only through $x(t) = x(t,\omega)$ and there is a map $(t,x) \to D_{t,x}$ that defines the tangent. $D_{t,x}$ is an infinitely divisible distribution. Homogeneous transition probabilities correspond to $D_{t,x} = D_x$ depending only on x.

Example 3: Continuous sample paths correspond to $\mu_{h,t,\omega}[|x| \ge \epsilon] = o(h)$. This is of course the same as the Lindeberg condition that is needed for the limit to be Gaussian. In other words $D_{t,\omega}$ being Gaussian corresponds to continuos paths.

Example 4: Finally, if the process is Markov and has continuos paths then $D_{t,\omega}$ is Gaussian with mean $b(t, x(t))$ and variance $a(t, x(t))$. It is defined in terms of the two functions, $a(t,x)$ and $b(t,x)$. Homogeneous case corresponds to functions that depend only on x.

Example 5: The Gaussian processes are characterized by linear regression and nonrandom conditional variance. In this case the Gaussian measure $D_{t,\omega}$ with mean $b(t,\omega)$ which is a linear functional of the path $x(s) : 0 \leq s \leq t$ and variance $a(t)$ which is purely a function of t. If $b(t,\omega) = b_1(t) + b_2(t)x(t)$ then we should have a Gauss Markov processes.

The question now is how to describe in mathematically precise terms the relationship between the process P and its tangent $D_{t,\omega}$. Let us look at Brownian motion as an example.

$$E^P\big[e^{\theta[x(t+h)-x(t)]}|x(s) : 0 \leq s \leq t\big] = e^{\frac{\theta^2 h}{2}}$$

or

$$E^P\big[e^{\theta[x(t+h)-x(t)]-h\frac{\theta^2}{2}}|x(s) : 0 \leq s \leq t\big] = 1$$

Equivalently

$$\exp[\theta[x(t) - x(0) - \frac{\theta^2}{2}t]$$

is a martingale w.r.t P and the natural filtration $\mathcal{F}_t = \sigma\{x(s) : 0 \leq s \leq t\}$. The converse that the only process P with respect to which the martingale property is valid for all $\theta \in R$ is Brownian motion is not hard to prove. From the martingale relation it follows that

$$E^P\big[\exp[\theta[x(t) - x(s)]|\mathcal{F}_s\big] = e^{\frac{\theta^2(t-s)}{2}}$$

which implies that $x(t) - x(s)$ is conditionally independent of $\mathcal{F}_t$ and has a Gaussian distribution with mean zero and variance $t - h$. This makes it Brownian motion.

Lecture 2.

We will now discuss processes with independent increments. We are familiar with Brownian motion whose increments over intervals of lenth τ are normally distributed with mean 0 and variance τ. More generally the distribution of the increments of a procees with independent increments over an interval of length τ will be a family μ_τ of probability distributions such that $\mu_\tau * \mu_s = \mu_{\tau+s}$ where $*$ denotes convolution. Such distributions are of course infinitely divisible and possess the Levy-Khintchine representation for their charecteristic function.

$$\widehat{\mu}_\tau(\xi) = \int e^{i\xi x}\mu_\tau(dx) = \exp\Big[\tau\big[ia\xi - \frac{\sigma^2\xi 2}{2} + \int [e^{i\xi y} - 1 - \frac{i\xi y}{1+y^2}]M(dy)\big]\Big] \tag{2.1}$$

Here $M(dy)$ is a sigma-finite mesure supported on $R\backslash 0$, which can be infinite only near 0. In other words $M[y : |y| \geq \delta] < \infty$ for any $\delta > 0$. More over M cannot be too singular near 0. It must integrate y^2.

$$\int_{|y|\leq 1} y^2 M(dy) < \infty$$

This allows the expression in the exponent of equation (2.1) to be well defined. It is not hard to see that if we add two processes with indpendent increments which are themselves independent of each other the sum is again a process with independent increments. Since the characteristic function of an indpendent sum is the product of the charecteristic functions, the Levy-Khintchine representation of the new process, which is the sum of the two original processes is easily obtained by adding the exponents in (2.1). In partcular we can try to understand the process that goes with μ_τ by breaking up the exponent into pieces.

If $\widehat{\mu}_\tau(\xi) = e^{i\tau a\xi}$ then the process is deterministic with $x(t) \equiv at$ with probability 1. It is a straightline with slope a. The slope is often called the "drift". But it does not necessaily imply that in general $E[x(t)] = at$.

If

$$\widehat{\mu}_\tau(\xi) = e^{-\frac{\tau\sigma^2\xi^2}{2}}$$

then process $x(t)$ is seen to be Brownian motion with variance $\sigma^2 t$. It can be represented as $\sigma\beta(t)$ in terms of the canonical Brownian motion with variance 1.

Before we turn to the final component that involves M let us look at the canonical Poisson process $N(t)$. This is a process with independent increments such that the distribution of $N(t) - N(s)$ is Poisson with parameter $t - s$. Since the increments are all nonnegative integers, this is a process $N(t)$ which increases by jumps that are integers. In fact the jumps are always of size 1. This is not so obvious and needs a calculation. Let us split the interval $[0, 1]$ into n intervals of length $\frac{1}{n}$ and ask what is the probability that at leat one increment is atleat k. This can be evaluated as

$$1 - [\sum_{0\le j\le k-1} e^{-\frac{1}{n}}[\frac{1}{n}]^j \frac{1}{j!}]^n \simeq 1 - (1 - \frac{1}{n^k k!})^n \to 0$$

if $k \ge 2$. This implies that the jumps are all of size 1. The times between jumps are all independent and have exponential distributions. One can then visualize the Poisson process as waiting for indpendent events with exponential distributions and counting the number of events upto time t. Therefore

$$\begin{aligned} P[N(t) \le k] &= P[\tau_1 + \tau_2 + \cdots + \tau_{k+1} \ge t] \\ &= \int_{\tau_1+\tau_2+\cdots+\tau_{k+1}\ge t} e^{-(\tau_1+\tau_2+\cdots+\tau_{k+1})} d\tau_1 d\tau_2 \cdots d\tau_{k+1} \\ &= \frac{1}{k!}\int_t^\infty e^{-\tau}\tau^k d\tau \end{aligned}$$

One can make the Poisson process more complex by building on top of it. Let us keep a sequence of i.i.d.random variables $X_1, X_2, \ldots, X_n, \ldots$ ready and define

$$X(t) = X_1 + X_2 + \cdots + X_{N(t)}$$

so that instead of just counting each time the event occurs we add and indpendent X to the sum. We then get the sum of a random number of indpendent random variables. The characteristic function of $X(t)$ is easy to compute.

$$E[e^{i\xi X(\tau)}] = E[[\phi(\xi)]^N(\tau)] = \sum_{j=0}^\infty [\phi(\xi)]^k \frac{e^{-\tau}t^k}{k!} = \exp[\tau[\phi(\xi) - 1]]$$

This is easily seen to lead to a Levy-Khintchine formula with exponent

$$\exp[\int (e^{i\xi y} - 1) d\alpha(y)]$$

where α is the distribution of X_i. While slightly diffrerent from the earlier form it can be put in that form by writing it as

$$\exp[i\xi \int \frac{y}{1+y^2} d\alpha(y) + \int [e^{i\xi y} - 1 - \frac{i\xi y}{1+y^2}] d\alpha(y)]$$

One should think of the Levy-Khintchine form as the centered form where the centering is not done by the expected value which will be by $\int y d\alpha(y)$. This may not in general be defined. Instead the centering is done by a truncated mean $\int \frac{y}{1+y^2} d\alpha(y)$. There is nothin sacred about $\frac{y}{1+y^2}$. One could have used any function $\theta(y)$ that looks sufficiently like y near the origin and remains bounded near infinity. Notice that the Poisson process $N(t)$ that enters the definition of $x(t)$ can have intensity λ different from 1, and nothing significant would change except the final formula

$$\widehat{\mu}_t(\xi) = \exp[t\lambda \int (e^{i\xi y} - 1) d\alpha(y)] = \exp[t \int (e^{i\xi y} - 1) dM(y)]$$

where now M is a measure with total mass λ. These are called compound Poisson processes.

We can always decompose the σ-finite measure M as an infinite sum $\sum_j M_j$ of finite measures. then for the process $X(t)$ with characteristic function

$$\widehat{\mu}_t(\xi) = \exp[t \int (e^{i\xi y} - 1 - \frac{i\xi y}{1+y^2}) dM(y)]$$

we have a representation as the sum

$$X(t) = \sum_j [X_j(t) - a_j\, t]$$

where $X_j(t)$ is the compound Poisson process with Levy measure M, and the constants $a_j = \int \frac{y}{1+y^2} dM_j(y)$ are centering constants that may be needed. Kolmogorov's three series theorem will guarantee that for the convergence of $\sum_j [X_j(t) - a_j\, t]$ it is necessary and sufficient

$$\sum_j \int_{-1}^{1} y^2 dM_j(y)$$

and

$$\sum_j M_j[y : |y| \geq 1]$$

converge. Equivalently

$$\sum_j \int \frac{y^2}{1+y^2} dM_j(y) < \infty$$

With out the centering the series may diverge and so it is not possible to separate the two sums $\sum_j X_j(t)$ and $\sum_j a_j\, t$ unless

$$\int_{-1}^{1} |y| dM(y) < \infty$$

Remarks:. If we add several mutually independent processes with independent increments their jumps cannot coincide. They just pile up at different times. Therefore for any process with independent increments $X(t)$ the levy measure M has the interpretation that for any set A not containing the origin the number of jumps $N_A(t)$ in the interval $[0, t]$ that are from the set A is a Poisson process with parameter $M(A)$ and for disjoint sets $\{A_\alpha\}$, the processes $\{N_{A_\alpha}\}$ are mutually independent. Except for the centering that may be needed this gives a complete picture of Poisson type processes with independent increments. What is left is a Process with independent increments with no jumps, which is of course Brownian motion with some variance $\sigma^2 t$ and drift at.

Generators and Semigroups. A process with independent increments is a one parameter semigroupμ_t of convolutions, i.e for $t, s \geq 0$, $\mu_t * \mu_s = \mu_{t+s}$. On the space $C(R)$ of bounded continuos function on R this defines a one parameter semi group of bounded operators

$$(T_t f)(x) = \int f(x+y)\mu_t(dy)$$

that satisfy $T_{t+s} = T_t T_s$. Their infinitesimal generator $\mathcal{A}$ is defined as

$$(\mathcal{A}f)(x) = \lim_{t\to 0} \frac{(T_t f)(x) - f(x)}{t} = \{\frac{d(T_t f)(x)}{dt}\}_{t=0}$$

The general theory of semigroups of linear operator outlines how to recover T_t from. Basically $T_t = e^{tA} = \sum_j \frac{t^j A^j}{j!}$. let us look at this for the Poisson process.

$$(T_t f)(x) = \sum_j f(x+j) e^{-t} \frac{t^j}{j!}$$

Simple differentiation gives

$$\mathcal{A}(f) = f(x+1) - f(x) = Sf - I$$

the difference operator, where S is the shift by 1. Then

$$e^{tA} = e^{t(S-I)} = e^{-t} e^{tS} = e^{-t} \sum \frac{t^j}{j!} S^j$$

giving us the Poisson semigroup. For the Brownian motion semigroup

$$(T_t f) = \frac{1}{\sqrt{2\pi t}} \int f(x+y) e^{-\frac{y^2}{2t}} dy = \frac{1}{\sqrt{2\pi}} \int f(x + y\sqrt{t}) e^{-\frac{y^2}{2}} dy$$

by a Taylor expansion of f, it is easy to see that for smooth f

$$(\mathcal{A}f)(x) = \frac{1}{2} f_{xx}(x)$$

and for the deterministic process $X(t) = at$

$$(\mathcal{A}f)(x) = \lim_{t \to 0} \frac{f(x+at) - f(x)}{t} = a f_x$$

We can put all the pieces together and write for a general process with independent increments represented in its Levy-Khintchine formula by $[a, \sigma^2, M]$, the infinitesimal generator is given by

$$(\mathcal{A}f)(x) = a f_x(x) + \frac{\sigma^2}{2} f_{xx}(x) + \int [f(x+y) - f(x) - \frac{y\, f_x(x)}{1+y^2}] M(dy)$$

Among all process that commute with translations these operators are singled out because of two properties. One, $\mathcal{A}1 = 0$, beacuse μ_t are probability measures and $T_t 1 = 1$. Since $T_t f$ is nonnegative when ever f is, at the global minimum x_0 of $f(x)$, since $f(x) \geq f(x_0)$, it follows that $(T_t f)(x_0) \geq f(x_0)$ for all $t > 0$. This means $(\mathcal{A}f)(x_0)$, which is the maximum principle. Any translation invariant operator with maximum principle satisfying $\mathcal{A}1 = 0$ is given by a Levy-Khintchine formula.

The situation is much less precise if we allow $D_{t,x}$ to depend on t and $x(t)$. Now instead of one Levy measure we have a whole family

$$[a(t,x), \sigma^2(t,x), M_{t,x}(dy)]$$

and infinitesimal generators that depend on t and do not commute with translations.

$$((\mathcal{A}_t f)(x) = a(t,x) f_x(x) + \frac{\sigma^2(t,x)}{2} f_{xx}(x) + \int [f(x+y) - f(x) - \frac{y\, f_x(x)}{1+y^2}] M(t,x,dy)$$

One of the more convenient ways of exploring the relationship between the process and the objects that occur in its representation is through a natural class of functionals that can be constructed from $[a, \sigma^2, M]$ and P is characterized as the measure with respect to which these functionals are martingales.

References.

Ito, Kiyosi. On stochastic differential equations. Mem. Amer. Math. Soc. 1951, (1951). no. 4, 51 pp.

Stroock, Daniel W.; Varadhan, S. R. Srinivasa. Multidimensional diffusion processes. Grundlehren der Mathematischen Wissenschaften 233. Springer-Verlag, Berlin-New York, 1979. xii+338 pp. ISBN: 3-540-90353-4

Stroock, Daniel W. Markov processes from K. Itô's perspective. Annals of Mathematics Studies, 155. Princeton University Press, Princeton, NJ, 2003. xviii+267 pp.

Varadhan, S. R. S. Stochastic processes. Notes based on a course given at New York University during the year 1967/68 Courant Institute of Mathematical Sciences, New York University, New York 1968 v+190 pp.

Varadhan, S. R. S. Lectures on diffusion problems and partial differential equations. With notes by Pl. Muthuramalingam and Tara R. Nanda. Tata Institute of Fundamental Research Lectures on Mathematics and Physics, 64. Tata Institute of Fundamental Research, Bombay, 1980. iii+315 pp.

Stroock, Daniel W. Diffusion processes associated with Lvy generators. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 32 (1975), no. 3, 209–244.