AN INTRODUCTION TO STOCHASTIC CALCULUS

thanks to Sanz

0. PRELIMINARIES

- Probability theory / Measure theory.
- Characteristic function.
- Conditional expectation. (Notes of Sanz)

Bibliography

- Nualart - Sanz, Curs de Probabilitat
- Alabert, Mesura i Probabilitat
- Chung, A course in probability theory
- Gut, An intermediate course on probability theor
- Jacob - Protter, Probability essentials.

Characteristic function

Introduced by P. Lévy

Let $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m), P)$ be a probability espace. We will say that $\varphi: \mathbb{R}^m \longrightarrow \mathbb{C}$ is a characteristic function if

$$\varphi(t) = \int_{\mathbb{R}^m} e^{i\langle t, x\rangle} P(dx) = \int_{\mathbb{R}^m} \cos\langle t, x\rangle P(dx) + i \int_{\mathbb{R}^m} \sin\langle t, x\rangle P(dx)$$

Observation: φ is the Fourier transform of P.

Let X be an $\mathbb{R}^m$-valued random variable. Its characteristic function φ_X defined on $\mathbb{R}^m$ is

$$\varphi_X(t) = E\left[e^{i\langle t, X\rangle}\right] = E\left(e^{it^* X}\right) = \int_{\mathbb{R}^m} e^{i\langle t, x\rangle} P_X(dx).$$

Properties:

1) $\varphi_X(0) = 1$, $|\varphi_X(t)| \leq 1$;
2) $\varphi_X(-t) = \overline{\varphi_X(t)}$
3) φ is uniformly continuous
4) A $m \times m$, $b \in \mathbb{R}^m$

$$\varphi_{AX+b}(t) = e^{i\langle t, b\rangle} \varphi_X(A^* t)$$

5) P_X symmetrical $\Leftrightarrow$ φ_X is real
6) $X_1, \dots, X_m$ are independent $\Leftrightarrow$ $\varphi_X(t) = \prod_{j=1}^{m} \varphi_{X_j}(t_j)$. $t \in \mathbb{R}^m$, $t = (t_1, \dots, t_m)$
7) X, Y independent $\Rightarrow$ $\varphi_{X+Y}(t) = \varphi_X(t)\,\varphi_Y(t)$.

theorem: Let X be an $\mathbb{R}$ valued random variable with $\mathbb{E}(|X|^m) < +\infty$ for some integer $m \geq 1$. Then, the characteristic function φ_X has continuous derivatives up to order m

$$\varphi_X^{(k)}(t) = i^k \int_{\mathbb{R}} x^k e^{itx} P_X(dx), \quad k = 1, \dots, m.$$

In particular,

$$\varphi_X^{(k)}(0) = i^k \mathbb{E} X^k.$$

theorem: Let X be an $\mathbb{R}$ valued random variable. If the characteristic function φ_X is k times differentiable around 0; then X has moments of order $2m$, with $2m \leq k$.

Example $N(0,1) \Rightarrow \varphi_X(t) = e^{-t^2/2}$;
$N(\mu, \sigma^2) \Rightarrow \varphi_X(t) = e^{it\mu - t^2\sigma^2/2}$

theorem: the characteristic function of a probability measure P on $\mathbb{R}$ characterizes P, that is, if two probabilities on $\mathbb{R}$ admit the same characteristic function, they are equal

Observation: $\varphi_X \equiv \varphi_Y \iff P_X \equiv P_Y$.

GAUSSIAN RANDOM VARIABLES (the normal and the Multivariate normal distrib

Let us recall that a Normal random variable with parameters $(\mu, \sigma^2$ where $\mu \in \mathbb{R}$ and $\sigma^2 > 0$, is a random variable whose density is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ - \frac{(x-\mu)^2}{2\sigma^2} \right\}, \quad x \in \mathbb{R}. \qquad (1)$$

We know that

$$\mathbb{E} X = \mu \quad ; \quad \mathrm{Var}(X) = \sigma^2.$$

Observation: For a real-valued random variable X the definition of $N(\mu, \sigma^2)$ is clear

$X \equiv (1)$ if $\sigma^2 > 0$,
$X \equiv \mu$ if $\sigma^2 = 0$.

For an $\mathbb{R}^n$- valued r.v. the definition is more subtle.

Definition: An $\mathbb{R}^n$-valued random variable $X = (X_1, \dots, X_n)$ is Gaussian (or multivariate normal) if every linear combination $\sum_{j=1}^{n} a_j X_j$ has a one-dimensional Normal distribution.

Obs: it can be degenerate, for example, $N(0,0)$ - when $a_j = 0 \ \forall j$.

Theorem: X is an $\mathbb{R}^n$-valued random variable if and only if its characteristic function has the form:

$$\varphi_X(t) = e^{i\langle t, \mu \rangle - \frac{1}{2} t^* \Lambda t},$$

where $\mu \in \mathbb{R}^n$ and Λ is an $n \times n$ symmetric nonnegative semi-definite matrix ($x^* \Lambda x \geq 0 \ \forall x \in \mathbb{R}^n$).

Observation:
- $t^* \Lambda t = \langle t, \Lambda t \rangle$
- Λ is the covariance matrix of X and μ is the mean of X_i.

Properties of Gaussian distributions

1) Let $X \sim N(\mu, \Lambda)$. Let C be an orthogonal matrix ($C^* C = C C^* = I$) such that $\Lambda = C^* D C$ where D is a diagonal matrix with $d_1, \dots, d_n$ the eigenvalues, then the random variable $Y = C(X - \mu)$ has independent marginals $N(0, d_i)$, $d_i \geq 0$.

2) $X \sim N(\mu, \Lambda) \Rightarrow EX = \mu, \quad E((X-\mu)(X-\mu)^*) = \Lambda$

$$E(X_i X_j) = \mu_i \mu_j + \Lambda_{ij}.$$

3) $X \sim N(\mu, \Lambda)$, A $r \times n$ matrix $\Rightarrow AX \sim N(A\mu, A \Lambda A^*)$.

4) $X \sim \mu(\mu, \Lambda)$. the components X_i are independent if and only if the covariance matrix Λ is diagonal.

5) An $\mathbb{R}^n$-valued Gaussian random variable X has a density on $\mathbb{R}^n$ if and only if the covariance matrix Λ is non-degenerate (there doesn't not exist $a \in \mathbb{R}^n$ s.t $\Lambda a = 0$ or $\det \Lambda \neq 0$).

6) If Λ is regular ($\det \Lambda > 0$) $\Rightarrow N(\mu, \Lambda)$ is not degenerate

$$\Rightarrow f_X(x) = \frac{1}{\sqrt{(2\pi)^n \det \Lambda}} \exp\left\{-\frac{1}{2}(x-\mu)^* \Lambda^{-1} (x-\mu)\right\}.$$

theorem (Central limit theorem):

Let $(X_i)_{i \geq 1}$ be i.i.d $\mathbb{R}^d$-valued random variables. Let $\mu = \mathbb{E}X_j$ and let Λ denote the covariance matrix $\Lambda = (\lambda_{k,\ell})_{1 \leq k,\ell \leq d}$, where $\lambda_{k,\ell} = \mathrm{Cov}(X_i^k, X_i^\ell)$, where X_i^k is the k^{th} component of the $\mathbb{R}^d$-valued random variable X_i. Then

$$\frac{S_n - n\mu}{\sqrt{n}} \xrightarrow{\mathcal{L}} Z, \quad Z \sim N(0, \Lambda),$$

with $S_n = X_1 + \cdots + X_n$.

---

1 STOCHASTIC PROCESSES

Let $(\Omega, \mathcal{F}, P)$ be a probability space.

Definition: A stochastic process with state space S is a family $\{X_i, i \in I\}$ of random variables $X_i : \Omega \longrightarrow S$ indexed by a set I.

Observation: the set S is called the states space. it has associated a σ-field of parts of S denoted by $\mathcal{S}$. Usually, $S = \mathbb{R}^n$ or $\mathbb{C}^n$, and $I = \mathbb{N}, \mathbb{Z}_+, \mathbb{R}_+$.

A stochastic process indexed by $\mathbb{R}_+$ is an application

$$\begin{array}{ccc} \mathbb{R}^+ & \longrightarrow & L^0 \equiv L^0(\Omega, \mathcal{F}, P, \mathbb{R}) \\ t & \longrightarrow & X_t \end{array}$$

where L^0 denote the vectorial space of all random variables. We can consider a stochastic procès as an application

$$\begin{array}{ccc} \mathbb{R}_+ \times \Omega & \xrightarrow{X} & \mathbb{R} \\ (t, \omega) & \longrightarrow & X_t(\omega) \end{array}$$

For any $\omega \in \Omega$, the process defines a function

$$\begin{array}{rccc} X_\cdot(\omega) : & \mathbb{R}_+ & \longrightarrow & \mathbb{R} \\ & t & \longrightarrow & X_t(\omega) \end{array}$$

called the process path.

1.1 the law of a stochastic process

Definition: the finite-dimensional joint distributions of the process $\{X_i, i \in I\}$ consist of the collection of the multi-dimensional probability laws of any finite family of random vectors $X_{i_1}, \dots, X_{i_m}$, where $i_1, \dots, i_m \in I$ and $m \in \mathbb{N}$.

Observation: We can deal with the stochastic processes from other points of view.

Definition: A stochastic process $\{X_t, t \in I\}$ is said to be Gaussian if its finite-dimensional joint distributions are Gaussian laws, that is, for any choice of $m \in \mathbb{N}$ and $t_1, \dots, t_m \in I$ the m-dimensional random vector $(X_{t_1}, \dots, X_{t_m})$ has a m-dimensional normal (or Gaussian) law.

In particular, $\mu_{t_1, \dots t_m} = (E(X_{t_1}), \dots, E(X_{t_m}))$

$$\Lambda_{t_1, \dots t_m} = (\mathrm{Cov}(X_{t_i}, X_{t_j}))_{1 \le i,j \le m}.$$

If $\det \Lambda_{t_1, \dots, t_m} > 0 \Rightarrow (X_{t_1}, \dots, X_{t_m})$ has a density.

Now, we consider a stochastic process as a generalized rand variable. A stochastic process $\{X_t, t \ge 0\}$ can be viewed "informally" as a random vector

$$X : \Omega \longrightarrow \mathbb{R}^I$$

this can be turned into a "formal" statement by putting the appropriate σ-field of events in $\mathbb{R}^I$, say $\mathcal{B}(\mathbb{R}^I)$. $(\mathcal{B}(\mathbb{R})^I)$

Definition: the law of a stochastic process $\{X_t, t \in \mathbb{R}\}$ defined on $(\Omega, \mathcal{F}, P)$ is the law of the random variable $X : \Omega \longrightarrow \mathbb{R}^{\mathbb{R}^+}$ defined by

$$\begin{array}{rccc} X : & \Omega & \longrightarrow & \mathbb{R}^{\mathbb{R}_+} \\ & w & \longrightarrow & X_\cdot(w) \end{array}$$

where the σ-field $\mathbb{R}^{\mathbb{R}_+}$ is the σ-field product $\mathcal{E}$

Observation: If P_X is the law of the process X, there exist Mathematical results from measure theory such that P_X is defined by a procedure of extension of measures on cylinder sets given the elements a

all possible finite-dimensional joint distributions.

Kolmogorov's theorem: Consider a family

$$\{P_{t_1,\dots,t_n},\ t_1<\dots<t_n,\ t_i\in I,\ n\geq 1\} \qquad (1)$$

where

1) $P_{t_1,\dots,t_n}$ is a probability on $\mathbb{R}^n$,
2) if $\{t_{i_1}<\dots<t_{i_m}\}\subseteq\{t_1<\dots<t_n\}$, the probability law $P_{t_{i_1},\dots,t_{i_m}}$ is the marginal distribution of $P_{t_1,\dots,t_n}$.

Then, there exists a stochastic process $\{X_t,\ t\in I\}$ defined in some probability space, such that its finite-dimensional joint distributions are given by (1). That is, the law of the random vector $(X_{t_1},\dots,X_{t_n})$ is $P_{t_1,\dots,t_n}$.

We give an application of this theorem.

Existence of Gaussian processes.

Let $k: I\times I\longrightarrow\mathbb{R}$ be a symmetric, nonnegative definite function, that is, a function such that:

- for any $s,t\in I$, $k(t,s)=k(s,t)$;
- for any natural number n and arbitrary $t_1,\dots,t_n\in I$, and $x_1,\dots,x_n\in\mathbb{R}$,

$$\sum_{i,j=1}^{n} k(t_i,t_j)\,x_i x_j\geq 0$$

Remark: Covariance functions of random vectors provide an example of nonnegative definite functions.

$(Y_1,\dots,Y_n)$ $EY_i=0$, $EY_i^2<\infty$ (for simplicity)

Fix $x_1,\dots,x_n\in\mathbb{R}$

$$\sum_{i,j=1}^{n} x_i x_j\,\mathrm{Cov}(Y_i,Y_j)=\sum_{i,j=1}^{n} x_i x_j\,E(Y_iY_j)$$

$$=E\Big(\sum_{i=1}^{n} x_iY_i\Big)^2\geq 0.$$

Proposition: Let $k: I\times I\longrightarrow\mathbb{R}$ be a symmetric, nonnegative definite function. Then, there exist a Gaussian process $\{X_t,\ t\geq 0\}$ such that $E(X_t)=0$ for any $t\in I$ and $\mathrm{Cov}(X_{t_i},X_{t_j})=k(t_i,t_j)$, for any $t_i,t_j\in I$.

Proof. We apply Kolmogorov's theorem. Fix $t_1, \dots, t_m \in I$ and we consider

$$\mu_{t_1,\dots,t_m} = (0, \dots, 0)$$

$$\Lambda_{t_1,\dots,t_m} = (K(t_i, t_j))_{1 \le i,j \le m},$$

$$P_{t_1,\dots,t_m} \sim N(0, \Lambda_{t_1,\dots,t_m}).$$

that means that $P_{t_1,\dots,t_m}$ is a m-dimensional gaussian law, mean zero and covariance matrix $\Lambda_{t_1,\dots,t_m}$.

We denote by $(X_{t_1}, \dots, X_{t_m})$ a random vector with law $P_{t_1,\dots,t_m}$. For any subset $\{t_{i_1}, \dots, t_{i_m}\}$ of $\{t_1, \dots, t_m\}$, we have

$$A(X_{t_1}, \dots, X_{t_m}) = (X_{t_{i_1}}, \dots, X_{t_{i_m}})$$

with

$$A = \begin{pmatrix} \delta_{t_1 t_{i_1}} & \cdots & \delta_{t_m t_{i_1}} \\ \vdots & & \\ \delta_{t_1 t_{i_m}} & \cdots & \delta_{t_m t_{i_m}} \end{pmatrix},$$

where $\delta_{s,t}$ denotes the kronecker Delta function.

By the properties on linear transformations of Gaussian random vectors, we know that the random $(X_{t_1}, \dots, X_{t_m})$ has an m-dimensional normal distribution, zero mean and covariance matrix $A \Lambda_{t_1,\dots,t_m} A^T$. By the definition of A, it is immediate to check that

$$A \Lambda_{t_1,\dots,t_m} A^T = (K(t_{i_\ell}, t_{i_k}))_{1 \le \ell, k \le m}.$$

Hence, the assumptions of theorem Kolmogorov's hold true and the results follows. ∎

Examples of stochastic processes

- Poisson process (Alabert)
- Cauchy process
- Ornstein-Uhlenbeck process

Examples of Gaussian processes

- Brownian motion
- Brownian bridge

Brownian motion

A one-dimensional Brownian $\{B_t, t \geq 0\}$ is a Gaussian stochastic process, with zero mean and covariance function given by

$$E(B_t B_s) = s \wedge t = \min(s, t).$$

The existence of such process is ensured by the previous proposition. Indeed, it suffices to check that

$$(s, t) \longmapsto \Gamma(s, t) = s \wedge t$$

is nonnegative definite function on $\mathbb{R}_+$. That means, for any $t_i \geq 0$ and any real numbers a_i,

$$\sum_{i,j=1}^{m} a_i a_j \Gamma(t_i, t_j) \geq 0.$$

Indeed,

$$s \wedge t = \int_0^\infty \mathbb{1}_{[0,s]}(r)\, \mathbb{1}_{[0,t]}(r)\, dr,$$

because $s \wedge t$ is the length of the intersection of the intervals $[0, s]$ and $[0, t]$. Hence,

$$\sum_{i,j=1}^{m} a_i a_j (t_i \wedge t_j) = \sum_{i,j=1}^{m} a_i a_j \int_0^\infty \mathbb{1}_{[0,t_i]}(r)\, \mathbb{1}_{[0,t_j]}(r)\, dr$$

$$= \int_0^\infty \left(\sum_{i=1}^{m} a_i \mathbb{1}_{[0,t_i]}(r) \right)^2 dr \geq 0.$$

Notice also that, since $E(B_0^2) = 0$, the random variable B_0 is zero almost surely. Brownian motion is also called Wiener process.

Brownian bridge

A Brownian bridge is a mean-zero Gaussian proces $\{X_t, t \in [0,1]\}$ with covariance

$$E(X_s X_t) = s \wedge t - st, \quad s, t \in [0, 1].$$

We can argue as before (Jensen's inequality). We also can prove the existence of this process by checking that

$$B_t - t B_1, \quad t \in [0, 1],$$

has the same law as X_t. Observe that $\{B_t - t B_1, t \in [0,1]\}$ is a gaussian process.

Some probabilistic properties of stochastic processes

Definition: Let $\{X_t, t \in I\}$ be a stochastic process. Assume that the set of indices I is a subset of $\mathbb{R}^d$ endowed with the usual order defined coordinatewise.

i) the stochastic process has independent increments if for any choice of $t_1 < t_2 < \cdots < t_k$, $k \geq 4$, the random variables

$$X_{t_2} - X_{t_1}, \cdots, X_{t_k} - X_{t_{k-1}},$$

are independents.

ii) the stochastic process has stationary increments if for any $t_1 < t_2$ the law of the random variable $X_{t_2} - X_{t_1}$ depends only on $t_2 - t_1$, but not on the particular value of t_1 and t_2.

- the Brownian motion will be stationary and have independent increments.
- If $X_{t_2} - X_{t_1} \sim X_{t_2 - t_1} \rightarrow X$ stationary

there exists some stronger notion of stationarity. For example L^2 - stationarity (mean-square stationarity). Consider X a stochastic process such that all its random variables X_t have finite second order moments, $E(X_t^2) < +\infty$. In this case, X is said to be stationary in L^2 if for any $0 \leq s \leq t$

$$E\left[(X_t - X_s)^2\right] = \varphi(t-s),$$

for some function φ.

1.2 Sample paths

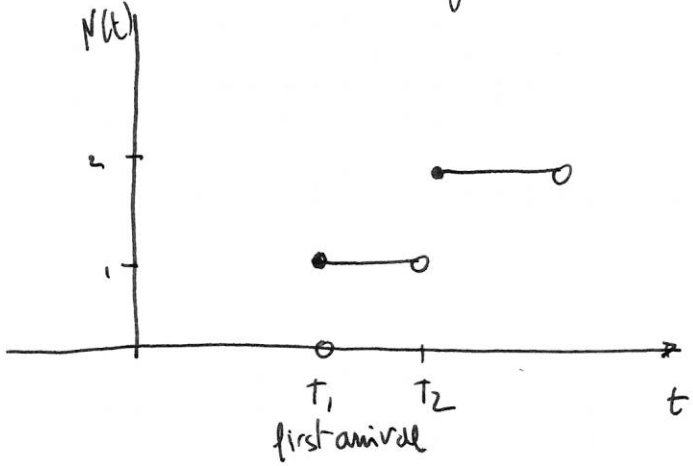
In the previous section, stochastic process are seen as random vectors. In the context of modelling, the observed values of the process are the relevant data. We fix values $\omega \in \Omega$ and we look it as a collection of their observations

Definition: the sample paths of a stochastic process $\{X_t, t \in I\}$ are the family of functions indexed by $\omega \in \Omega$, $X(\omega) : I \longrightarrow S$, defined by $X(\omega)(t) = X_t(\omega)$.

sample paths are also called trajectories.

Examples

1) Consider random arrivals of customers at a store. We set the clock at zero and record the time between two consecutive arrivals. They are random variables $X_1, X_2, \dots$



$X_j = T_j - T_{j-1}$. We assume $X_j > 0$, a.s. Let $t_0 = 0$ and $T_n = \sum_{j=1}^{n} X_j$, $n \geq 1$. The random T_n is the time of the n-th arrival. Define N_t as the random number of customers who have visited the store during the time interval $(0, t)$, $t \geq 0$.

Clearly, $N_0 = 0$, and for $t > 0$, $N_t = k$ if only if $T_k \leq t < T_{k+1}$. The process $\{N_t, t \geq 0\}$ takes values on $\mathbb{Z}_+$. Its sample paths are increasing right continuous functions, with jumps at the random times T_n, $n \geq 1$, of size one.

It is a particular case of a "counting process". Sample paths of counting processes are always increasing right continuous functions, their jumps are natural numbers.

In the particular case where $\{X_i, i \geq 1\}$ is a sequence of i.i.d random variables with law $\exp(-\lambda)$, $\lambda > 0$, then the process $\{N_t, t \geq 0\}$ is the Poisson process.

2) Perhaps, the Geometric Brownian motion that is used as a model for evolution of stock prices