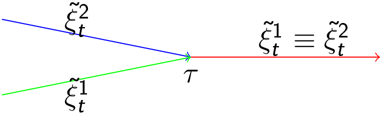
Michael Lvovich Blank   blank@iitp.ru



“*Human mathematics is a sort of dance around an unwritten formal text, which if written would be unreadable.*” (David Ruelle, 1998)

Result = 0.4*(class work) + 0.6*Exam

14:50-16:10 on Mondays (room 108)

http://iitp.ru/ru/userpages/74/233.htm

Main topics

- Antagonistic and cooperative games/strategies.
- The concept of a random process.
- Elements of random analysis.
- Correlation theory of random processes.
- Markov processes.
- Wiener and Poisson processes.
- Stochastic integral. Ito's formula.
- (sub/super) martingales.
- Infinitesimal semi-group operator.
- Large deviations and nonlinear Markov processes.

- D. Stirzaker. Elementary probability, Cambrige University Press, 2003
- А.Д. Вентцель. Курс теории случайных процессов. М.: Наука. 1996
- N.V. Krylov. Introduction to the theory of random processes. 2002
- Б. Оксендаль. Стохастические дифференциальные уравнения, 2003
- А.Н. Ширяев. Вероятность, 2 т. МЦНМО, 2007.

An antagonistic or zero-sum game: one player's victory means the other players' loss (boxing, tennis). A cooperative game: players win as a team (football), hence the choice of a cooperative strategy.

Problem

Players A and B independently (and without discussion with the other) toss a proper coin and try to predict the result of the other player's toss. Antagonistic: a correct prediction means victory (otherwise, loss). Cooperative: players win as a team if at least one of them makes the correct prediction (otherwise they lose).

Antagonistic: the probability to win is equal to $1/2$.
Cooperative: it seems that the probability is $1-(1/2)^2 = 3/4$.
Strategy: the player A uses his result as a prediction for B, while the player B reverts his result: Pred(A,B):=(1-B,A).
Claim. Under this strategy the team always win.
Proof. $00 \to 10 \quad 01 \to 00 \quad 10 \to 11 \quad 11 \to 01$.

Problem

$N = 2n \gg 1$ numbered prisoners must find their own numbers in one of N boxes in order to survive, opening at most half of boxes.

Antagonistic strategy: the probability to survive $= 2^{-N}$.
Cooperative strategy: each prisoner opens the box labeled with his own number, then the one whose number is found in the box, etc.
Claim. The strategy ensures that the correct number is eventually found along the cycle, regardless of its length.
Calculation. For $k > n$ there are C_N^k ways to select the numbers of such the k-cycle, which can be arranged in $(k-1)!$ ways. Since the remaining numbers can be arranged in $(N-k)!$ ways, the number of permutations of the numbers 1 to N with a cycle of length $k > n$ is equal to $C_N^k(k-1)!(N-k)! = \frac{N!}{k}$.
Finally, the probability, that a uniformly distributed random permutation contains no cycle of length greater than n is given by $1 - \frac{1}{N!}\sum_{i=1}^{n} \frac{N!}{n+i} \approx 0.3$.

Definition

A *random variable* (r.v.) $\xi : (\Omega, \mathcal{F}, \mathbf{P}) \to (X, \mathcal{B})$ is an arbitrary measurable map, provided $\forall x \in \mathcal{B}$.

Equivalence:

$\xi \sim \eta \iff \mathbf{P}(\xi \neq \eta) := \mathbf{P}(\{\omega \in \Omega : \ \xi(\omega) \neq \eta(\omega)\}) = 0.$

σ-algebra generated by ξ: $\mathcal{F}_\xi := \sigma(\xi^{-1}\mathcal{B})$.

Distribution of ξ: $\Phi_\xi(A) := \mathbf{P}(\xi \in A) \in \mathcal{M}(X), \ \forall A \in \mathcal{B}$.

Joint distribution: $\Phi_{\xi_1,\dots,\xi_n}(A) := \mathbf{P}((\xi_1, \dots, \xi_n) \in A), \ A \in \mathcal{B}^n$.

Definition

R.v. $\xi_1, \dots, \xi_n$ are *independent* if $\Phi_{\xi_1,\dots,\xi_n} = \prod_i \Phi_{\xi_i}$.

Example: $\mathbf{P}(\xi \in A, \eta \in B) = \mathbf{P}(\xi \in A)\mathbf{P}(\eta \in B)$.
$n = \infty$: independence means that $\forall k < \infty$ are independent.
σ-algebras $\mathcal{F}_\alpha \subseteq \mathcal{F}$ are independent if $\forall A_\alpha \in \mathcal{F}_\alpha$ are independent for different α.

1. Independence in pairs but not jointly.
   $X := \bigsqcup_{i=1}^{4} A_i$, $\mathbf{P}(A_i) = 1/4\ \forall i$.
   $\{B_j := A_j \cup A_4\}_{j=1}^{3} \implies \mathbf{P}(B_j) = 1/2\ \forall j$.
   $\mathbf{P}(B_i \cap B_j) = \mathbf{P}(A_4) = 1/4 = \mathbf{P}(B_i)\mathbf{P}(B_j)\ \forall i \neq j$ (pairs).
   $\mathbf{P}(\cap_{j=1}^{3} B_j) = \mathbf{P}(A_4) = 1/4 \neq 1/8 = \prod_{j=1}^{3} \mathbf{P}(B_j)$ (joint).
2. $X := \{1, 2, 3\}, \mathcal{B} := \{1, \{2, 3\}, \emptyset, X\}, \mathbf{P}(1) := \mathbf{P}(\{2, 3\}) = \frac{1}{2}$.
   Find all independent events. Answer: $(\emptyset, X), (\emptyset, \emptyset), (X, X)$.
3. Existence of independent random variables? Any constants
4. $\Omega := [0, 1], \mathbf{P} := \mathrm{Leb} = m, \quad \xi, \eta : (\Omega, \mathrm{Bor}, m) \to (\mathbb{R}, \mathrm{Bor})$.
   Check independence: $\mathbf{P}(\xi \in A, \eta \in B) = \mathbf{P}(\xi \in A)\mathbf{P}(\eta \in B)$
   (a) $\xi(\omega) := a + b\omega,\ \eta(\omega) := c + d\omega$
   (b) $\xi(\omega) := a\omega,\ \eta(\omega) := b\omega^3$
   (c) $\xi(\omega) := a \sin(2\pi\omega),\ \eta(\omega) := b \cos(2\pi\omega)$
   (e) $\xi(\omega) := 1_I(\omega) \cos(2a\pi\omega),\ \eta(\omega) := 1_J(\omega) \cos(2b\pi\omega)$,
   $I, J \in \mathrm{Bor}$

Binomial sum:
$\mathrm{Bin}(n,p) := \sum_{i=1}^{n} \xi_i,\ \xi_i \in \{0,1\},\ \ \mathbf{P}(\xi_i = 1) = p.$
Calculate $\pi_n := \mathbf{P}(B_n := \{\mathrm{Bin(n,p) = 2k \ - \ even}\}).$

$$\pi_{n+1} = \mathbf{P}(B_{n+1}) = \mathbf{P}(B_{n+1} \cap B_n) + \mathbf{P}(B_{n+1} \cap B_n^c)$$
$$= \mathbf{P}(\xi_{n+1} = 0 | B_n)\mathbf{P}(B_n) + \mathbf{P}(\xi_{n+1} = 1 | B_n^c)\mathbf{P}(B_n^c)$$
$$= (1-p)\pi_n + p(1-\pi_n) = (1-2p)\pi_n + p.$$

How to solve this difference equation with $\pi_0 = 1$?
Solution: $\pi_n := a(1-2p)^n + b \Longrightarrow a = b = 1/2.$
Finally $\pi_n = ((1-2p)^n + 1)/2 \overset{n\to\infty}{\longrightarrow} 1/2.$

If (X, ρ) is a metric space, we consider convergences:

- in probability (P) $\lim\limits_{n\to\infty} \xi_n = \xi$ if $\mathbf{P}(\rho(\xi_n, \xi) \geq \varepsilon) \overset{n\to\infty}{\longrightarrow} 0.$
- in average $L^p(\Omega, \mathcal{F}, \mathbf{P})$: $||\xi_n - \xi||_p \overset{n\to\infty}{\longrightarrow} 0.$
- weak: $\Phi_{\xi_n} \overset{n\to\infty}{\longrightarrow} \Phi_\xi$ in the weak sense.

Mathematical expectation: $Ef(\xi) := \mathbf{P}(f \circ \xi) := \int\limits_{\Omega} f(\xi(\omega))\mathbf{P}(d\omega)$.

Variance: $Df(\xi) = E(f(\xi) - Ef(\xi))^2 = E(f(\xi))^2 - (Ef(\xi))^2$,

Covariance:

$\mathrm{cov}(f(\xi), g(\eta)) := E\left((f(\xi) - Ef(\xi))(g(\eta) - Eg(\eta))^*\right)$.

Chebyshev ineq-ty: $\mathbf{P}(f(\xi) \geq \varepsilon) \leq Ef(\xi)/\varepsilon \qquad f \geq 0, \varepsilon > 0.$
$\mathbf{P}(|\xi - E\xi| \geq \varepsilon) \leq D\xi/\varepsilon^2 \quad \xi \in \mathbb{R}^1.$

i-th *marginal distribution* of $\xi := (\xi_1, \ldots, \xi_n)$ is Φ_{ξ_i}.
Characteristic function $\varphi_\xi(t) := Ee^{i(t,\xi)}$ for $X = \mathbb{R}$.

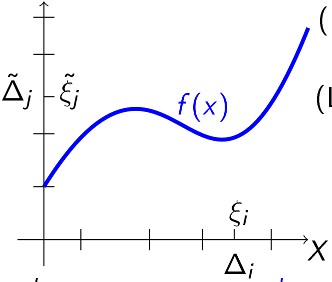
$\xi \in \mathbb{R}^n$ is *Gaussian* $\mathcal{N}(a, A)$ if $\varphi_\xi(t) := e^{i(t,a) - \frac{1}{2}(At,t)}$, $A \geq 0$
with density $f(x) := \sqrt{(2\pi)^{-n} \det A^{-1}}\, e^{-(A^{-1}(x-a),(x-a))/2}$.

Claim. $\xi_i \in \xi \in \mathcal{N}(a, A)$ are independent iff A is diagonal.
Question. Is it true that uncorrelated Gaussian r.v. are independent? No

Claim. Let ξ, η be independent with densities f_ξ, f_η. Then
$f_{\xi+\eta}(x) = f_\xi * f_\eta(x) := \int f_\xi(t) f_\eta(x - t)dt$.

[9] Lecture 2. Integrals and weak convergence

Let $\{X, \mathcal{B}, \mu\}$ be a *measure space*, $f : X \to \mathbb{R}$ – measurable.



$$\text{(R)}\int f\ d\mu := \lim_{\mathrm{diam}\Delta\to 0} \sum_i f(\xi_i)\mu(\Delta_i)$$

$$\text{(L)}\int f\ d\mu := \lim_{\mathrm{diam}\tilde{\Delta}\to 0} \sum_j \tilde{\xi}_j \mu(f^{-1}\tilde{\Delta}_j)$$

$$= \lim_{\mathrm{diam}\tilde{\Delta}\to 0} \sum_j \tilde{\xi}_j f \mu(\tilde{\Delta}_j) =: \mu(f)$$

$\exists \int_a^b f'(x)\ dx = ?$ $(R) \int_a^b f'(x)\ dx = f|_a^b$, $(L) \int_a^b f'(x)\ dx \neq f|_a^b$

Definition

The *weak* topology* on $\mathcal{M}(X)$ is defined so that, $\mu_n \xrightarrow{n\to\infty} \mu \iff \mu_n(\varphi) \xrightarrow{n\to\infty} \mu(\varphi) \quad \forall \varphi \in C^0_{\mathrm{bounded}}(X)$. The convergence in the weak* topology is called the *weak convergence*.

If X is a compact set, then there is a metric generating this convergence $\mathrm{dist}(\mu, \nu) := \sup_{|\varphi|+|\varphi'|\leq 1} (\mu(\varphi) - \nu(\varphi))$.

List 1 – deadline 26.02

1. Let $\xi : \Omega \to \mathbb{R}$ be a r.v. with the median m_ξ. Prove/disprove that $m_{a\xi+b} = am_\xi + b \;\; \forall a, b \in \mathbb{R}$.
2. Let $\xi, \eta, \zeta : \Omega \to \mathbb{R}$ be r.v. and let (ξ, η) and (ξ, ζ) be independent. Prove/disprove that $(\xi, \eta + \zeta)$ are independent.
3. Let $\{\xi_i\}_{i=1}^n$ be iid positive r.v. Prove/disprove that $\forall k$ $E\left(\frac{\sum_{i=1}^k \xi_i}{\sum_{i=1}^n \xi_i}\right) = \frac{k}{n}, \quad 1 \le k \le n.$
4. Let $(X := \mathbb{R}, \mathcal{B} := \mathrm{Bor}, \mathbf{P})$ be a probability space, $\xi : \Omega \to X$ - a r.v. with distribution $\mathbf{P}$, and let $f : \mathbb{R} \to \mathbb{R}$ be an absolutely continuous function with $f(\pm\infty) = a$, $\int_{-\infty}^{\infty} f(x)\, d\mathbf{P}(x) = b$. Calculate $Q := \int_{-\infty}^{\infty} f'(x)\, \mathbf{P}(\xi \ge x)\, dx$ in terms of a, b only.
5. Let $\{\xi_i\}_{i=1}^n$ be independent r.v. with the same distribution function $F(x)$. Find the distribution function of the vector $(\min_i \xi_i, \max_i \xi_i)$.
6. Let $\{\xi_i\}_{i=1}^n$ be iid r.v. with $0 < D\xi_i < \infty$. Find all possible values of the function $\varphi(x) := \lim_{n\to\infty} \mathbf{P}(\sum_{i=1}^n \xi_i < x), \quad x \in \mathbb{R}$.

Do not wait until the deadline, and send written solutions (preferably in LaTex) by e-mail.

Definition

Let $\{\xi_n\}$ be a sequence of iid real valued r.v. with a finite $E\xi_n = a$. We say that this sequence satisfies the Law of Large Numbers (LLN) if $\tilde{\xi}_n := \frac{1}{n}\sum_{k=1}^n \xi_k \overset{n\to\infty}{\longrightarrow} a$. In case of the convergence in probability this is called the weak LLN, while in case of a.s. convergence: $\mathbf{P}(\lim_{n\to\infty}\tilde{\xi}_n = a) = 1$ this is called the strong LLN.

Kolmogorov Theorem (LLN)

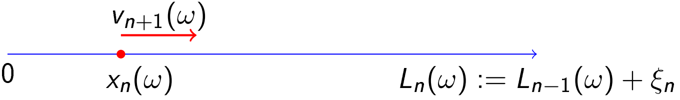
If $\{\xi_n\}$ are iid real valued r.v., then $\tilde{\xi}_n - E\tilde{\xi}_n \overset{n\to\infty}{\longrightarrow} 0$ a.s.

Counterexample

Let $\{\eta_n\}$ be iid r.v. such that $\eta_n = \pm 1$ with probability $\frac{1}{2}$. Then (provided $\log\log\log(N) > 0$) the sequence $\{\xi_n := \eta_n\sqrt{(n+N)/\log\log\log(n+N)}\}$ satisfies the weak LLN, but the strong LLN fails.

[12] Lecture 2. Ant on a Rubber Rope

Initially, the ant stands on the left end of a rubber rope $L_0 = \ell = 1$ meter long and crawls along the rope with a speed $v_n \ll 1$ at time n. At the n-th minute of time, the speed of the ant instantly changes from v_n to v_{n+1}, and the rope stretches uniformly from L_n to $L_{n+1} := L_n + \xi_n$. We assume that $\{v_n\}$ and $\{\xi_n\}$ are nonnegative independent random variables with $0 < Ev_n = v \ll E\xi_n = \ell < \infty$.



$x_0(\omega) \equiv 0, \quad L_0(\omega) \equiv \ell$

Question: Will the ant ever reach the right end of the rubber rope?

$$x_{n+1} = \frac{L_{n+1}}{L_n} x_n + v_{n+1} \Longrightarrow \frac{x_n}{L_n} = \frac{x_{n-1}}{L_{n-1}} + \frac{v_n}{L_n} = \sum_{k=1}^n \frac{v_k}{L_k}.$$

$$E\left[\frac{x_n}{L_n}\right] = \sum_{k=1}^n \frac{Ev_k}{EL_k} = \sum_{k=1}^n \frac{v}{k\ell} = \frac{v}{\ell} \sum_{k=1}^n \frac{1}{k} \xrightarrow{n\to\infty} \infty.$$

This argument hints that the answer should be yes, but does not really prove anything. $\Longrightarrow$

Claim

Let $\{v_n\}$ be positive, independent, and identically distributed random variables (iid r.v.) with $Ev_n = v > 0$, and let $\{\xi_n\}$ be positive iid r.v. with $E\xi_n = \ell < \infty$. Let $L_n := L_{n-1} + \xi_n$, $L_0 := 1 \;\; \forall n \geq 1$. Then $\sum_{k\geq 0} \frac{v_k}{L_k} = \infty$ almost surely.

Proof. Let Ω be the sample space. By the strong law of large numbers $\exists \Omega'$ with $\mathbf{P}(\Omega') = 1$, such that
$\frac{1}{n}\sum_{k=1}^{n} v_k \overset{n\to\infty}{\longrightarrow} v$ and $\frac{1}{n}\sum_{k=1}^{n} \xi_k \overset{n\to\infty}{\longrightarrow} \ell \;\; \forall \omega \in \Omega'$.

Hence $\forall \omega \in \Omega'$, $\varepsilon > 0 \;\; \exists N$:
$|\frac{1}{n}\sum_{k=1}^{n} v_k - v| < \varepsilon, \quad |\frac{1}{n}\sum_{k=1}^{n} \xi_k - \ell| < \varepsilon \quad$ if $n > N$.

Thus $\forall p \in \mathbb{N}$ we have:
$|\frac{1}{N}\sum_{k=(p-1)N}^{pN} v_k - v| = |\frac{1}{N}\sum_{k=1}^{pN} v_k - \frac{1}{N}\sum_{k=1}^{(p-1)N} v_k - v|$
$= |p\frac{1}{pN}\sum_{k=1}^{pN} v_k - pv + (p-1)v - (p-1)\frac{1}{(p-1)N}\sum_{k=1}^{(p-1)N} v_k|$
$\leq p\varepsilon + (p-1)\varepsilon = (2p-1)\varepsilon.$ $\Longrightarrow$

Hence

$$\begin{aligned}\sum_{n\geq 1}\frac{v_n}{L_n} &\geq \sum_{k=1}^{N}\frac{v_k}{L_k}+\sum_{k=N+1}^{2N}\frac{v_k}{L_k}+\ldots+\sum_{k=(p-1)N+1}^{pN}\frac{v_k}{L_k}\\ &= \sum_{q=1}^{p}\sum_{k=(q-1)N+1}^{qN}\frac{v_k}{L_k}\\ &\geq \sum_{q=1}^{p}\sum_{k=(q-1)N+1}^{qN}\frac{v_k}{L_{qN}}\\ &= \sum_{q=1}^{p}\frac{1}{q}\frac{\sum_{k=(q-1)N+1}^{qN}\frac{v_k}{N}}{\frac{L_{qN}}{qN}}\\ &\geq \sum_{q=1}^{p}\frac{1}{q}\frac{v-\varepsilon}{\ell+\varepsilon}\overset{\varepsilon\to 0+}{\longrightarrow}\frac{v}{\ell}\sum_{q=1}^{p}\frac{1}{q}\overset{p\to\infty}{\longrightarrow}\infty.\end{aligned}$$

□

Question:

Will the ant ever reach the right end of the rubber rope if

1. $L_{n+1} := L_n + (\frac{L_n}{2024})^{\alpha_n}$, where $\{\alpha_n\}$ are iid r.v. with $E\alpha_n = \alpha \in (0, 2)$.
2. $L_n(\omega) := L_{n-1}(\omega)\cdot\lambda_n(\omega)$, where $\{\lambda_n\}$ are iid r.v. with $E\lambda_n = \lambda \in (1, 2)$.

 Test 1

1. $(\Omega, \mathcal{F}, \mathbf{P})$, $\Omega = [0,1]$, $\mathcal{F} = \mathrm{Bor}$, $\mathbf{P} = \mathrm{Leb}$,
$\{\xi_n := \omega^n\}_{n\in\mathbb{Z}_+}$, $\omega \in \Omega$. $A_n := \{\omega \in \Omega : \xi_n \le 1/n\}$
Calculate: $\mathbf{P}(\cup_{n\ge1} A_n) = 1$, $\mathbf{P}(\cap_{n\ge1} A_n) = 1/\sqrt[3]{3}$.
Solution: $A_n = \{0 \le \omega \le 1/\sqrt[n]{n}\} \implies A_1 = [0,1]$,
further use that $\sqrt[n]{n} \to 1$ and $\max_n \sqrt[n]{n} = \sqrt[3]{3}$.
2. Find all $a, \varepsilon \ge 0$ such that $\mathbf{P}(\xi \ge a) \le e^{-a-\varepsilon} E e^{\xi}$, $\forall \xi$.
$\mathbf{P}(\xi \ge a) \le e^{-a} E(e^{\xi} \cdot 1_{\xi \ge a}) \le e^{-a} E e^{\xi}$ (Chernoff inequality)
Answer: $\varepsilon = 0, \forall a \ge 0$, since for
$\xi \equiv e^a \implies \mathbf{P}(\xi \ge a) = 1 > e^{-\varepsilon}$.
3. Let $\Omega := \{1,2,3,4\}, \mathcal{F} := 2^{\Omega}, \mathbf{P}(\{i\}) = \frac{1}{4}$. Prove/disprove $\exists$
of iid non-constant r.v. $\xi, \eta : (\Omega, \mathcal{F}, \mathbf{P}) \to (\mathbb{R}, \mathrm{Bor})$.
Solution: $\xi(\omega) := 1_{\{1,2\}}(\omega)$, $\eta(\omega) := 1_{\{2,3\}}(\omega)$
$\mathbf{P}(\xi = 1) = \mathbf{P}(\eta = 0) = \frac{1}{2}$,
$\mathbf{P}(\xi = 0, \eta = 1) = \mathbf{P}(\omega \in \{3\}) = \frac{1}{4} = \mathbf{P}(\xi = 0)\mathbf{P}(\eta = 1) \quad \ldots$

Solution: (problem number) Answer. Short proof.

$$\Omega := [0,1], \mathcal{F} := \mathrm{Bor}, \mathbf{P} = \mathrm{Leb},\ \Omega \ni \omega = \sum_{k\geq 1} 2^{-k}\omega_k,\ \omega_k \in \{0,1\}.$$

Claim 1

$\{\xi_n(\omega) := \omega_n\}_{n\geq 1}$ are iid Bernoulli(1/2) r.v. with $\mathbf{P}(\xi_n = 0) = \mathbf{P}(\xi_n = 1) = \frac{1}{2}$.

Claim 2

Let $\{\xi_n\}_{n\geq 1}$ be iid Bernoulli(1/2) r.v., then $\eta(\omega) := \sum_{k\geq 1} 2^{-k}\xi_k(\omega)$ is uniformly distributed r.v. on $[0,1]$.

Let $\tilde{\xi}_{nk}(\omega)$ be the (n,k) element in the following triangle table:

$\omega_1\ \omega_3\ \omega_6\ \omega_{10}\ \dots$

$\omega_2\ \omega_5\ \omega_9\ \dots$

$\omega_4\ \omega_8\ \dots$

$\omega_7\ \dots$

$\dots$

Claim 3

$\{\zeta_n := \sum_{k\geq 1} 2^{-k}\tilde{\xi}_{nk}\}_{n\geq 1}$ are uniformly distributed iid on $[0,1]$.

Claim 4

Let ξ be a uniformly distributed r.v. on $[0,1]$ and let F be an arbitrary distribution function. Then $\eta := \tilde{F}^{-1}(\xi)$ is a r.v. with the distribution $F_\xi = F$.

Here $\tilde{F}^{-1}(t) := \inf\{s : \ F(s) \geq t\}$ is the generalized inverse function (for non srictly monotone setting).

Claim 5

Let $\{F_k\}_{k\geq 1}$ be an arbitrary sequence of distribution functions. Then $\exists$ a sequence of independent r.v. $\{\eta_k\}_{k\geq 1}$ with $F_{\eta_k} = F_k$.

Construction:

$\eta_k := \tilde{F}_k^{-1}(\xi_k)$ with iid uniformly distributed $\{\xi_k\}$.

Wrt an event $B \in \mathcal{F}$, $\mathbf{P}(B) > 0$: $\quad E(\xi|B) := E(\xi \cdot 1_B)/\mathbf{P}(B)$.
Observation: $E(\xi = 1_A|B) = E(1_{A\cap B})/\mathbf{P}(B) = \frac{\mathbf{P}(A\cap B)}{\mathbf{P}(B)} = \mathbf{P}(A|B)$.
$E(\xi|B) = \frac{\mathbf{P}(\xi\cdot 1_B)}{\mathbf{P}(B)} = \int_B \xi(\omega)\frac{d\mathbf{P}(\omega)}{\mathbf{P}(B)} = \int_B \xi d\mathbf{P}(\omega|B) = \int_\Omega \xi d\mathbf{P}(\omega|B)$.

$E\xi = \sum_i \mathbf{P}(B_i)E(\xi|B_i)$ if $\sqcup_i B_i = \Omega$, $\quad \Delta := \{B_i\} \in \mathcal{F}$.

$\mathbf{P}(A|\Delta)(\omega) := \sum_i \mathbf{P}(A|B_i)\cdot 1_{B_i}(\omega)$ – random variable:
(a) $A \cap B = \emptyset \Longrightarrow \mathbf{P}(A \cup B|\Delta) = \mathbf{P}(A|\Delta) + \mathbf{P}(B|\Delta)$,
(b) $\mathbf{P}(A|\Omega) = \mathbf{P}(A)$, $\quad (c)E(\mathbf{P}(A|\Delta)) = \mathbf{P}(A)$.

If $\#(\eta(\Omega)) < \infty$ then $\exists\Delta_\eta := \{B_i\}$ – a partition generated by η and $\mathbf{P}(A|\eta) := \mathbf{P}(A|\Delta_\eta)$.

$E(\xi|\Delta)(\omega) := \sum_i E(\xi|B_i)\cdot 1_{B_i}(\omega)$ – random variable:
(a) $E(a\xi + b\eta|\Delta) = aE(\xi|\Delta) + bE(\eta|\Delta)$ $\quad$ (b) $E(\xi|\Omega) = E(\xi)$
(c) $E(1_A|\Delta) = \mathbf{P}(A|\Delta)$ $\quad$ (d) $E(E(\xi|\Delta)) = E(\xi)$
(e) $\eta := \sum_i z_i 1_{B_i} \Longrightarrow E(\xi\eta|\Delta)(\omega) = \eta(\omega)E(\xi|\Delta)(\omega)$
(f) $\omega \in B_i \Longrightarrow E(\xi\eta|\Delta)(\omega) = E(\xi\eta|B_i) = z_i E(\xi|B_i)$
$$= \eta(\omega)E(\xi|B_i) = \eta(\omega)E(\xi|\Delta)(\omega).$$

$E(\xi|\mathcal{A}), \mathcal{A} \subseteq \mathcal{F}$ is a random variable in $\mathbb{R} \cup \pm\infty$ such that:
– $E(\xi|\mathcal{A})$ is $\mathcal{A}$-measurable; – $\mathbf{P}(\xi \cdot 1_A) = \int_A E(\xi|\mathcal{A}) d\mathbf{P}$, $A \in \mathcal{A}$.

Properties: linearity, monotonicity +
– $E(\xi\eta|\mathcal{A}) = \xi \cdot E(\eta|\mathcal{A})$ if ξ is $\mathcal{A}$-measurable.
– $E(E(\xi|\tilde{\mathcal{A}})|\mathcal{A}) = E(\xi|\mathcal{A})$ if $\mathcal{A} \subset \tilde{\mathcal{A}} \subseteq \mathcal{F}$. Equalities are $\mathbf{P}$-a.e.

$E(\xi|\eta) := E(\xi|\mathcal{B}_\eta)$, where $\mathcal{B}_\eta := \sigma(\eta)$. $\mathbf{P}(A|\eta) = \mathbf{P}(1_A|\mathcal{B}_\eta)$.

T1. Let $\Delta_\eta := \{B_i\}_1^n$ be a partition of $(\Omega, \mathcal{F})$, $\mathcal{B} := \sigma(\Delta)$ and $|E\xi| < \infty \Longrightarrow E(\xi|\mathcal{B}) = E(\xi|\Delta)$ with probability 1.
Proof. By $\mathcal{B}$-measurability $\mathbf{P}(E(\xi|_{B_i}|\mathcal{B}) = z_i = \mathrm{const}) = 1$. Hence $E(\xi|\mathcal{B}) = \sum_i z_i 1_{B_i} = \sum_i E(\xi|B_i) \cdot 1_{B_i} = E(\xi|\Delta)$. □

$\exists E\xi < \infty$ implies $\exists!$ $E(\xi|\mathcal{A})(\omega)$.

$E(\xi|\eta) := E(\xi|\sigma(\eta))$. Here $E(\xi|\eta = x)(x)$ – conditional ME.
Conditional density $p_\xi(x|\eta = y) = p_{\xi\eta}(x, y)/p_\eta(y)$:
$E(f(\xi, \eta)|\eta = y) = \int f(x, y) p_\xi(x|\eta = y) dx$ for $f \in L^1$.

In general $E(\xi|\mathcal{B})$ cannot be calculated explicitly, however in some simple cases this is still possible.

Let $dP(x,y) = p(x,y)dxdy$ be a probability measure on $\mathbb{R}^2$ with $p(x,y) > 0$ $(P(A) := m(1_A \cdot p)$ for the Lebesgue $m)$.
Consider σ-algebras $\mathcal{B}_x$ generated by the coordinate function x and let P_x be the projection (marginal distribution) of P to the x-coordinate with $p_x(x) = m_y(p(x,\cdot))$.
The conditional measure P^x on $\ell_x := \{(x,y) : y \in \mathbb{R}\}$ has the density $p^x(y) = p(x,y)/p_x(y) = p(x,y)/\int p(x,y)dy$ – prob. measure.
$E(\xi|\mathcal{B}_x) = m_y(\xi(x,\cdot)\rho(x,\cdot)) = \frac{\int \xi(x,y)p(x,y)dy}{\int p(x,y)dy}$.

Problem

2 gamers play a symmetric coin $\mathbf{P}(\{0/1\}) = 1/2$. Each has a winning pattern A and B respectively (a finite number of consecutive binary digits). The game stops when one of the patterns appears.

Questions:

(a) Does the game stop after a finite time?

(b) Let the pattern length (i) $|A| = |B|$, (ii) $|A| < |B|$.

Who will win?

$$\text{(a) } \mathbf{P}\left(\underbrace{\underbrace{\ldots}_{k}\underbrace{\ldots}_{k}\cdots\underbrace{\ldots}_{k}}_{n}\right) \leq (1-2^{-k})^n \overset{n\to\infty}{\longrightarrow} 0 \quad k = |A|,$$

each k-pattern $\neq A$.

(b. i) **Counterexample**. $A := 000, B := 100 \Longrightarrow |A| = 3 = |B|$.
$\xi_1, \xi_2, \ldots, \xi_n, \ldots, \quad \xi_i \in \{0, 1\}$.
Claim. If $\xi_1 = 1$ then B will show up before A: $\quad 101100\ldots$
Corollary. A wins iff $\xi_1\xi_2\xi_3 = 000$.
Hence $\mathbf{P}(A-\mathrm{wins}) = 2^{-3} = 1/8 < 7/8 = \mathbf{P}(B-\mathrm{wins})$.

(b. ii) **Counterexample**. Let
$A = 000, B = 1000 \Longrightarrow |A| = 3 < |B|$. To compensate the length difference the winning counts from the beginning of the pattern.

(c) $\exists$ of the "best" pattern of a given length? Mirror symmetry.
$100 \sim 011 > 000 \sim 111$, $001 \sim 110$, $101 \sim 010$, $100 - ? - 101$

Idea of large deviations, from r.v. and up to random DS.

Chebyshev inequality with $0 < \varphi \nearrow$ – nondecreasing:

$\mathbf{P}(\xi \geq t) \leq \mathbf{P}(\varphi(\xi) \geq \varphi(t)) \leq \frac{E\varphi(\xi)}{\varphi(t)}$. Set $\varphi(x) := e^{tx}$

$\mathbf{P}(\xi \geq \varepsilon) \leq \frac{Ee^{t\xi}}{e^{t\varepsilon}}, \quad \mathbf{P}(\xi \leq -\varepsilon) = \mathbf{P}(e^{-t\xi} \geq e^{t\varepsilon}) \leq \frac{Ee^{-t\xi}}{e^{t\varepsilon}}$.

Chernoff's idea is to find the value of t minimizing r.h.s.

Moment generating function:

$M_\xi(t) := Ee^{t\xi} = 1 + tE\xi + \frac{t^2}{2}E\xi^2 + \cdots + \frac{t^n}{n!}E\xi^n + \ldots$

Generating (proizvodyaschaya) function: $Ez^\xi, \quad |z| < 1$,

Characteristic function: $Ee^{it\xi}$.

Properties of $M_\xi(t)$: (a) $E\xi^n = \frac{d^n}{dt^n}M_\xi(0)$ (if $\exists$ near 0)

(b) $M_\xi(t) = M_\eta(t) \ \ |t| < \delta \implies \xi = \eta$ (on distribution)

(c) ξ, η independent $\implies M_{\xi+\eta} = M_\xi M_\eta$

Proof. $M_{\xi+\eta}(t) = Ee^{t(\xi+\eta)} = Ee^{t\xi}e^{t\eta} = M_\xi M_\eta$. □

Theorem 1

ξ_i iid $\mathbf{P}(\xi_i = \pm 1) = \frac{1}{2} \implies \mathbf{P}(|\sum_{k=1}^n \xi_k| > \varepsilon) < 2e^{-\varepsilon^2/(2n)}$.

Proof. We check that $\mathbf{P}(S_n := \sum_{k=1}^n \xi_k > \varepsilon) < e^{-\varepsilon^2/(2n)}$.
$Ee^{t\xi_k} = (e^t + e^{-t})/2 = \cosh(t) \le e^{t^2/2}$. To prove the last inequality we compare the corresponding Taylor series:
$\cosh(t) = (e^t + e^{-t})/2 = \sum_{k\ge 0} \frac{t^{2k}}{(2k)!}$ (odd terms cancel) and
$e^{t^2/2} = \sum_{k\ge 0} \frac{t^{2k}}{2^k k!}$

$$(2k)! = \underbrace{(2k)(2k-1)\dots(k+1)}_{\ge 2^k} k! \ge 2^k k!$$

Now since $Ee^{tS_n} = \prod_{k=1}^n Ee^{t\xi_k} = \cosh^n(t) \le e^{nt^2/2}$ we get
$\mathbf{P}(S_n > \varepsilon) \le e^{\frac{nt^2}{2}}/e^{t\varepsilon} = e^{\frac{nt^2}{2} - t\varepsilon}$.
Choosing $t = \varepsilon/n$ (minimizing rhs) we get
$\mathbf{P}(S_n > \varepsilon) \le e^{-\varepsilon^2/(2n)}$. $\square$

Theorem 2

Let $\mathrm{Bin}(n,p) := \sum_{i=1}^{n} \xi_i$, $\mathbf{P}(\xi_i = 1) = p$. Then
$\mathbf{P}(|\mathrm{Bin}(n,p) - np| > t) < 2e^{-t^2/(3np)}$ if $0 \le t \le np$.
$\mathbf{P}(|\mathrm{Bin}(n,p) - np| > t) < 2e^{-np/3}$ if $t > np$.

Lemma

Let $|\xi| \le 1, E\xi = 0 \Longrightarrow M_\xi(t) \le e^{t^2 D\xi} \quad \forall t \in [-1,1]$.

Proof. $|t\xi| \le 1, E\xi = 0 \Longrightarrow e^{t\xi} \le 1 + t\xi + (t\xi)^2 \Longrightarrow$
$\Longrightarrow Ee^{t\xi} \le 1 + t^2 E\xi^2 = 1 + t^2 D\xi \le e^{t^2 D\xi}$. $\square$

Theorem 3

Let ξ_i be independent with $|\xi_i - E\xi_i| \le 1 \ \forall i$.
Set $S_n := \sum_{i=1}^n \xi_i$, $\sigma := \sqrt{DS_n}$
$\Longrightarrow \mathbf{P}(|S_n - ES_n| \ge \varepsilon\sigma) \le 2\max(e^{-\varepsilon^2/4}, e^{-\varepsilon\sigma/2})$.

Proof. It is enough to consider $E\xi_i = 0$. Due to symmetry we get
$\mathbf{P}(S_n \ge \varepsilon\sigma) \le e^{-t\varepsilon\sigma/2}$ for $t = \min(\varepsilon/(2\sigma), 1)$.
$\sum^n D\xi_i = \sigma^2$, hence by the Lemma:
$\mathbf{P}(S_n \ge \varepsilon\sigma) \le e^{-t\varepsilon\sigma} \prod_{i=1}^n Ee^{t\xi_i} \le e^{-t\varepsilon\sigma} \prod_{i=1}^n e^{t^2 D\xi_i} = e^{-t\varepsilon\sigma + t^2\sigma^2}$.
Thus choosing $t \le \frac{\varepsilon}{2\sigma}$ we get the result. □

Application: coin tossing:
$\xi_i \in \{0,1\}, p = \frac{1}{2}, ES_n = \frac{n}{2}, DS_n = \frac{n}{4}$.
Chebyshev: $\mathbf{P}(|S_n - ES_n| \ge \delta ES_n) \le \frac{DS_n}{\delta^2 (ES_n)^2} = \frac{1}{\delta^2 n}$
Chernoff: $\mathbf{P}(|S_n - ES_n| \ge \delta ES_n) \le 2e^{-\delta^2 ES_n/3} = 2e^{-\delta^2 n/6}$

much better!

$(X,\mathcal{B})$ – a measurable space, $\mathcal{M}=\mathcal{M}(X,\mathcal{B})$ – probabilistic measures on X. *Markov chain* $\mathcal{T}^t:\mathcal{M}\to\mathcal{M}$, $t\in\mathbb{Z}$ or $\mathbb{R}$ is a family of operators such that $\forall\mu,\nu\in\mathcal{M}$:

– $\mathcal{T}^t(a\mu+(1-a)\nu)=a\mathcal{T}^t(\mu)+(1-a)\mathcal{T}^t(\nu)$, $\quad 0\le a\le 1$

– $\mathcal{T}^{t+s}(\mu)=\mathcal{T}^s\circ\mathcal{T}^t(\mu)$ – semigroup or Markov property.

If $\mathcal{T}^t\delta_x=\delta_y$ $\forall x\in X$ and some $y=y(x)\in X \implies$ *deterministic* and *random* otherwise.

Deterministic: $F:(X,\mathcal{B})\to(X,\mathcal{B}) \implies \mathcal{T}^n\mu(A):=\mu(F^{-n}A)\ \forall A$.

Random Examples:

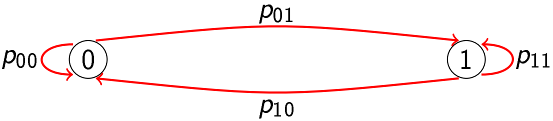
(a) *Random map*: $F_1,F_2:(X,\mathcal{B})\to(X,\mathcal{B})$, $0<p<1$,
$$\mathcal{T}^1\mu(A):=p\mu(F_1^{-1}A)+(1-p)\mu(F_2^{-1}A).$$

(b) *Finite state Markov chain*: $X:=\{0,1\}$, $P:=\begin{pmatrix} p_{00} & p_{01}\\ p_{10} & p_{11}\end{pmatrix}$,

$$\mathcal{T}^1\mu:=\mu^*P:\ \begin{pmatrix}\mu(0)\to p_{00}\mu(0)+p_{10}\mu(1)\\ \mu(1)\to p_{01}\mu(0)+p_{11}\mu(1)\end{pmatrix},\ \mathcal{T}^n\mu=\mu^*P^n.$$

[28] Lecture 4. Markov chains (unorthodox approach) 2

$\mathcal{T}\mu = \mu^* P.$



(c) iid $\xi_i \in \{0,1\}$ Markov chain with $p_{ij} = 1/2$.
(d) General continuous time Markov chains: transition probabilities $P_s^t(x, A) := \mathbf{P}(\xi_{s+t} \in A | \xi_s = x)$.
How this corresponds to the traditional approach?

Stochastic function is a a family of r.v. $\{\xi_t\}_{t\in T}$, or $\xi_t(\omega) : (\Omega, \mathcal{F}, \mathbf{P}) \times T \to (X, \mathcal{B})$ – $\forall t$ measurable on ω.
When $T = \mathbb{Z}^d$ or $\mathbb{R}^d$ we identify t with time and speak about *stochastic processes*. $\xi_t(\bullet)$ – r.v. for a fixed t.
$\xi_\bullet(\omega)$ – *realization or trajectory* – nonrandom for a given ω.
$\xi_t \sim \eta_t$ if $\mathbf{P}(\xi_t \neq \eta_t) = 0 \ \ \forall t \in T$ – *equivalence*.
$\Phi_{t_1,\dots,t_n}(A) := \mathbf{P}((\xi_{t_1}, \dots, \xi_{t_n}) \in A)$ – *a finite dimensional distribution*.
$\xi_t \sim \eta_t \implies \Phi^\xi = \Phi^\eta$ (but not vice versa).
Question: what about realizations? - No.
Example: Let $T := [0,1]$ and a r.v. $\tau \in (0,1)$ have a continuous distribution. Set $\xi_t \equiv 0$, $\eta_t := \begin{cases} 0 & \text{if } t \neq \tau \\ 1 & \text{otherwise} \end{cases}$. $\quad \xi_t \sim \eta_t$ since $\mathbf{P}(\xi_t \neq \eta_t) = \mathbf{P}(\tau = t) = 0$, however each trajectory of ξ_t is identically 0, while each trajectory of η_t has a "jump" at time τ.

Let $F(x)$ be a probability distribution. Is there a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and a r.v. $\xi(\omega)$ such that $\mathbf{P}(\omega : \ \xi(\omega) \le x) = F(x)$? Set $\Omega := \mathbb{R}, \mathcal{F} := \mathrm{Bor}(\mathbb{R})$. Then $\exists! \mathbf{P} : \ \mathbf{P}((a, b]) = F(b) - F(a)$ and thus for the r.v. $\xi(\omega) := \omega$ we have $F_\xi(x) \equiv F(x)$.

Now we consider the same problem for a random process $\xi_t, \ \ t \in T \subseteq \mathbb{R}$ with *finite dimensional distributions* $F_{t_1,\dots,t_n}(x_1, \dots, x_n) = \mathbf{P}(\omega : \ \xi_{t_1} \le x_1, \dots, \xi_{t_n} \le x_n)$.

Theorem (Kolmogorov)

Let $F_{t_1,\dots,t_n}(x_1, \dots, x_n)$ be a given family of finite dimensional distributions, satisfying the following consistency conditions:

$$\begin{aligned} &F_{t_1,\dots,t_{k-1},t_k,t_{k+1},\dots,t_n}(x_1, \dots, x_{k-1}, \infty, x_{k+1}, \dots, x_n) \\ = {}&F_{t_1,\dots,t_{k-1}, \quad t_{k+1},\dots,t_n}(x_1, \dots, x_{k-1}, \quad\; x_{k+1}, \dots, x_n). \end{aligned}$$

Then $\exists(\Omega, \mathcal{F}, \mathbf{P})$ and a random process $\xi_t, \ \ t \in T$ such that $\mathbf{P}(\omega : \ \xi_{t_1} \le x_1, \dots, \xi_{t_n} \le x_n) = F_{t_1,\dots,t_n}(x_1, \dots, x_n)$.

A random process $\xi_t : (\Omega, \mathcal{F}, P) \to (X, \mathcal{B}, m)$ acting on a Borel $(X, \mathcal{B})$ space with a finite reference measure m $(\neq m(\xi_t))$ is a *Markov chain* defined by *transition probabilities*

$$Q_s^t(x, A) := P(\xi_{s+t} \in A | \xi_s = x), \;\; A \in \mathcal{B},$$

with standard properties:

- $Q_s^t(x, \cdot)$ is a probability measure on the σ-algebra $\mathcal{B}$.
- For fixed s, t, A the function $Q_s^t(\cdot, A)$ is $\mathcal{B}$-measurable.
- For $t = 0 \quad Q_s^t(x, A) = \delta_x(A)$.
- For each $s, 0 \leq t \leq t'$ and $A \in \mathcal{B}$ we have

$$Q_s^{t'}(x, A) = \int_X Q_s^t(x, dy) Q_t^{t'-t}(y, A).$$

The process ξ_t induces an action on measures:

$$Q_s^t\mu(A) := \int Q_s^t(x, A)\ d\mu(x)$$

and an action on functions:

$$Q_s^t\varphi(x) := \int \varphi(y) Q_s^t(x, dy).$$

Definition

A Borel measure μ is said to be *invariant* or *stationary* for the Markov chain ξ_t if it is a solution to the equation

$$Q_s^t\mu = \mu \quad \forall s, t.$$

A system is *deterministic* iff $Q_s^t\delta_x$ is a δ-measure $\forall x, s, t$. This agrees with the discussion of the unorthodox approach.

(1) iid r.v. $\xi_t : (\Omega, \mathcal{F}, \mathbf{P}) \to (X, 2^X)$, $t \in \mathbb{Z}_+$, $X := \{0, 1\}$,
$$\mathbf{P}(\xi_t = 1) = p, \mathbf{P}(\xi_t = 0) = 1 - p.$$
Let $b := \{b_k\}$ be a binary sequence and let $S^N(b)$ be its left shift by N positions, i.e. $(S^N(b))_i := b_{i+N}$. $W(b, n) := (b_1, b_2, \ldots, b_n)$.

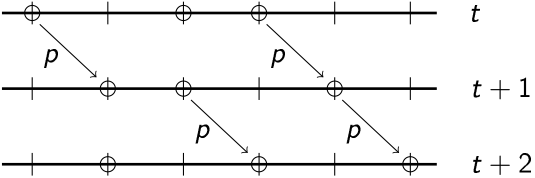
Definition

A sequence b is strongly recurrent if $\forall n_0, n \in \mathbb{Z}_+$ there exists $N = N(b, n_0, n)$ such that $W(S^{n_0} b, n) = W(S^{n_0+N} b, n)$; and uniformly strongly recurrent if $\exists$ an infinite sequence of shifts $\{N_k\}$, such that $\sup_k |N_{k+1} - N_k| < \infty$.

Calculate: $\mathbf{P}(\xi_t \text{ is s.recurrent})$, $\mathbf{P}(\xi_t \text{ is uniformly s.recurrent})$.

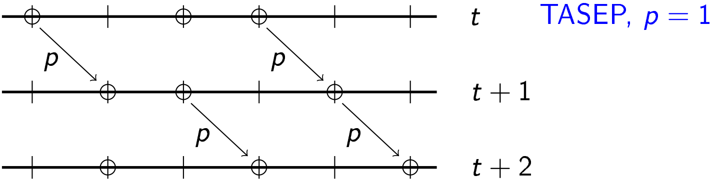
(2) Simple random walk: $\eta_t : (\Omega, \mathcal{F}, \mathbf{P}) \to (X, 2^X)$, $t \in \mathbb{Z}_+$, $X := \mathbb{Z}$, $\eta_{t+1} := \eta_t + \xi_t$, where $\xi_t \in \{-1, 1\}$ are iid with $\mathbf{P}(\xi_t = 1) = p, \mathbf{P}(\xi_t = -1) = 1 - p$.

(3) Collective random walks – the exclusion process EP.
A configuration $\zeta_t := (\dots, \zeta_t^{-1}, \zeta_t^{0}, \zeta_t^{1}, \dots), \quad \zeta_t^i \in \mathbb{Z}$ describes positions of "particles" on the lattice $\mathbb{Z}$ at time t. Each particle performs the random walk if it does not interfere with other particles.

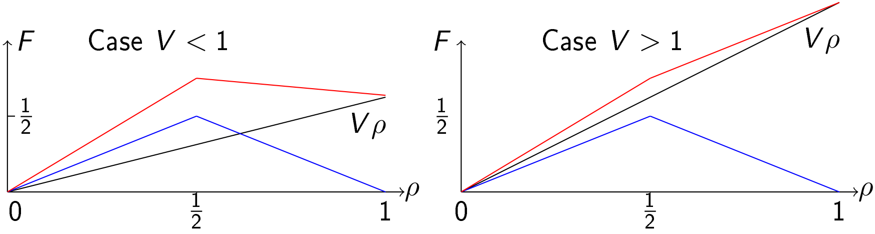


TASEP

The main problem in the analysis of such systems is an infinite number of simultaneous interactions between neighboring particles. An example will be discussed on the next slide.



Let V be the escalator's velocity, ρ – the density of passengers.
Then the passengers flow $F(\rho, V) := (1-|1-2\rho|)/2 + V\rho$.



$F(\frac{1}{2}, V) > F(1, V)$ iff $\frac{1+V}{2} > V \Longrightarrow V < 1$.

Let $\eta_t[p]$ be a random walk on $\mathbb{Z}$ with $\mathbf{P}(\eta_{t+1} - \eta_t = 1) = p$. A sequence $\{b_k\}, b_k \in \mathbb{Z}$ is *recurrent* if $\forall i\ \exists n = n(i) > 0 : \ b_i = b_{i+n}$.

Find all values of the parameter $p \in [0, 1]$ such that
(a) $\eta_t[p]$ is recurrent,
(b) $\eta_t[p]$ is strongly recurrent,
(c) $\eta_t[p]$ is uniformly strongly recurrent.

Claim

Let $\Omega_n^k := \{\omega : \ \text{a return to k occurs after 2n time steps}\} \Longrightarrow$ $\mathbf{P}(\cup_{n\ge0}\Omega_n^k) = 1$ iff $\sum_{n\ge0}\mathbf{P}(\Omega_n^k) = \infty$.

$\Longrightarrow$ see a more general statement on the next slide.

Discussion. Set $q := 1 - p$. We have
$\mathbf{P}(\Omega_n^k) = C_{2n}^n (pq)^n = \frac{(2n)!(pq)^n}{n!n!} \sim \frac{(4pq)^n}{\sqrt{\pi n}}$ (by the Stirling formula).
Thus recurrence occurs iff $p = q = 1/2$.

General framework. Let ξ_n be a Markov chain on $\mathbb{Z}^+$ with transition probabilities $p_{i,j}^{(n)}$.

Theorem

$v := \max_n \mathbf{P}(\xi_n = k | \xi_0 = k) = 1$ iff $\sum_{n \geq 1} p_{k,k}^{(n)} = \infty \quad \forall k \in \mathbb{Z}_+$.

Proof. Let $v_n := \mathbf{P}(\text{the 1st return to k occurs after n steps})$, and let $v := \sum_{n \geq 1} v_n$. By the formula of total probability we have

(*) $\; p_{i,i}^{(n)} = \sum_{j=0}^{n} p_{i,i}^{(j)} v_{n-j}$.

Set additionally $u_n := p_{i,i}^{(n)}$ and introduce the generating functions $U(z) := \sum_{m \geq 0} u_m z^m$, $V(z) := \sum_{m \geq 0} v_m z^m$, which are analytic for $|z| \leq 1$. Then (*) is equivalent to

$U(z) - u_0 = U(z) V(z), u_0 = 1 \Longrightarrow U(z) = \frac{1}{1 - V(z)}$.

$\lim_{z \to 1} U(z) = \lim_{z \to 1} \frac{1}{1 - V(z)} = \frac{1}{1 - v} = \infty$ if $v = 1$. On the other hand, $\lim_{z \to 1} U(z) = \lim_{z \to 1} \sum_{m \geq 0} u_m z^m = \sum_{m \geq 0} u_m = \infty$. $\square$

(0) Random sin oscillations: $\xi_t := A\cos(\eta t + \varphi)$, r.v. $A, \eta \geq 0, \varphi$.
φ is uniformly distributed on $[0, 2\pi)$ and does not depend on A, η.

(I) Poisson process ξ_t with the parameter $a > 0$ on $T := \mathbb{R}_+$:
(0) $\xi_0 = 0$.
(i) $\forall 0 \leq t_0 < t_1 < \cdots < t_n$ r.v. $\Delta\xi_{t_i,t_{i-1}} := \xi_{t_i} - \xi_{t_{i-1}}$ independent.
(ii) r.v. $\Delta\xi_{t,s} := \xi_t - \xi_s$, $0 \leq s \leq t$ are Poisson distributed:
$\mathbf{P}(\Delta\xi_{t,s} = k) = (a(t-s))^k e^{-a(t-s)}/k!$, $k \in \mathbb{Z}_+$.
(iii) Trajectories of ξ_t are right continuous.

(II) Cauchy process: (0) + (i) +
(ii') r.v. $\Delta\xi_{t,s} := \xi_t - \xi_s$, $0 \leq s \leq t$ are Cauchy distributed with the density $p(x) = \pi^{-1}(t-s)/((t-s)^2 + x^2)$.

(III) Wiener process w_t: (0) + (i) +
(ii'') r.v. $\Delta w_{t,s} := w_t - w_s$, $0 \leq s \leq t$ are Gaussian $\mathcal{N}(0, t-s)$.
(iii') Trajectories of w_t are continuous.

Test 2

1. Prove that stochastically equivalent processes (i.e. $P(\xi(t) \neq \eta(t)) = 0 \ \forall t$) have the same finite dimensional distributions. $P(\xi(t_i) \in B_i \ \ i \in \{1, \dots, n\}) = \ \ (\forall t_i, B_i \in \mathcal{B})$
$= P(\cap_i \{\xi(t_i) \in B_i\} \ \cap \ \cap_i \{\xi(t_i) = \eta(t_i)\})$
$= P(\cap_i \{\eta(t_i) \in B_i\} \ \cap \ \cap_i \{\xi(t_i) = \eta(t_i)\})$
$= P(\eta(t_i) \in B_i \ \ i \in \{1, \dots, n\})$.
2. Construct stochastically non-equivalent processes, having the same finite dimensional distributions. $\xi, \eta : \Omega \to \{-1, 1\}$, $\mathbf{P}(\xi = \pm 1) = \mathbf{P}(\eta = \pm 1) = \frac{1}{2}$, $\xi(\omega) = -\eta(\omega) \ \forall \omega \in \Omega$.
3. Let ξ have the standard normal distribution $\mathcal{N}(0, 1)$. Check if $P(\xi \geq a) \leq e^{-a^2/2} \ \ \forall a \geq 0$.
$P(\xi \geq a) = \frac{1}{\sqrt{2\pi}} \int_a^\infty e^{-x^2/2} dx \leq \frac{1}{\sqrt{2\pi}} \int_a^\infty \frac{x}{a} e^{-x^2/2} dx$
$= \frac{1}{a\sqrt{2\pi}} e^{-a^2/2} \leq e^{-a^2/2}$ if $a\sqrt{2\pi} \geq 1$. Otherwise, if $a\sqrt{2\pi} < 1$:
$P(\xi \geq a) \leq P(\xi \geq 0) = \frac{1}{2} \leq e^{-\frac{1}{2}(\sqrt{2\pi})^{-2}} \leq e^{-a^2/2}$
since $\ln 2 \sim 0.69 > 1/(4\pi)$. [Wrong if $a < 0$]

Solution: (problem number) Answer. Short proof.

(0) Random sin oscillations: $\xi_t := A\cos(\eta t + \varphi)$, r.v. $A, \eta \geq 0, \varphi$.
φ is uniformly distributed on $[0, 2\pi)$ and does not depend on A, η.

Claim

Finite dimensional distributions of $\xi_t, t \in T := \mathbb{R}$ are translationally invariant: $\mu_{\bar{t}+h} = \mu_{\bar{t}}$ $\forall \bar{t} = (t_1, \ldots, t_n), h \in \mathbb{R}$.

Proof. We need to prove the following equality:

(*) $Z := \mathbf{P}(\{A\cos(\eta(t_1+h)+\varphi), \ldots, A\cos(\eta(t_n+h)+\varphi)\} \in C)$
$= \mathbf{P}(\{A\cos(\eta t_1+\varphi), \ldots, A\cos(\eta t_n+\varphi)\} \in C)$.

$B := \{(x, y, z) : x, y \geq 0, z \in [0, 2\pi),$
$\{x\cos(yt_1+z), \ldots, x\cos(yt_n+z)\} \in C\}$ is a Borel set.

Denoting by $\{z\}_{2\pi}$ the fractional part of z mod 2π, from (*) we get
$Z = \mathbf{P}((A, \eta, \{\varphi+\eta h\}_{2\pi}) \in B) = \mathbf{P}((A, \eta, \varphi) \in B)$.

(A, η) and φ are independent $\Longrightarrow \mu_{A,\eta,\varphi} = \mu_{A,\eta} \times \mu_\varphi$. Thus
$Z = \int_0^\infty \int_0^\infty \mu_{A,\eta}(dxdy)\ \mu_\varphi(C_1 := \{z : (x, y, \{z+yh\}_{2\pi}) \in B\})$
$= \int_0^\infty \int_0^\infty \mu_{A,\eta}(dxdy)\ \mu_\varphi(C_2 := \{z : (x, y, z) \in B\})$, since C_1 is obtained from C_2 by the translation by yh and taking mod 2π. Finally, μ_φ is uniform on $[0, 2\pi)$ and does not change under translations. $\square$

(I) Poisson process ξ_t with the parameter $a > 0$ on $T := \mathbb{R}_+$:
(0) $\xi_0 = 0$.
(i) $\forall 0 \le t_0 < t_1 < \cdots < t_n$ r.v. $\Delta\xi_{t_i,t_{i-1}} := \xi_{t_i} - \xi_{t_{i-1}}$ independent.
(ii) r.v. $\Delta\xi_{t,s} := \xi_t - \xi_s$, $0 \le s \le t$ are Poisson distributed:
$\mathbf{P}(\Delta\xi_{t,s} = k) = (a(t-s))^k e^{-a(t-s)}/k!$, $k \in \mathbb{Z}_+$.
(iii) Trajectories of ξ_t are right continuous.

Claim

a.a. trajectories are non-decreasing integer valued functions with jumps of size 1.

Proof. Main idea. Show that probabilities of the events
$A := \{\xi_t \in \mathbb{Z}\ \forall t = k2^{-n}\}$, $B := \{\xi_s \le \xi_t\ \forall s \le t = k2^{-n}\}$,
$C_N := \{\forall k \in \mathbb{Z} \cap [0, \xi_N]\ \exists t = k2^{-n} \in [0, N] : \xi_t = k\}$ are equal 1.
To this end one approximates them by events depending only on a finite number of values ξ_t.

see next slide

Proof. The event $A = \bigcap_{t=k\cdot 2^{-n}} (A_t := \{\xi_t \in \mathbb{Z}_+\})$

$$\mathbf{P}(\xi_t \in \mathbb{Z}_+) = \mathbf{P}(\xi_t - \xi_{t_0} \in \mathbb{Z}_+) = \sum_{i=0}^{\infty} \mathbf{P}(\xi_t = i) = 1 = \mathbf{P}(A).$$

B is the intersection of the events:

$B_n := \{\xi_0 \le \xi_{1\cdot 2^{-n}} \le \dots \xi_{k\cdot 2^{-n}} \le \dots\} = \bigcap_k \{\xi_{k\cdot 2^{-n}} \le \xi_{(k+1)\cdot 2^{-n}}\}$.

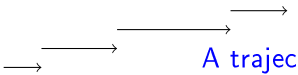
Since $\mathbf{P}(\xi_{k\cdot 2^{-n}} \le \xi_{(k+1)\cdot 2^{-n}}) = 1$, we have $1 = \mathbf{P}(B_n) = \mathbf{P}(B)$.

$C_N \supseteq \bigcap_{k=0}^{2^n N-1} \{\xi_{(k+1)\cdot 2^{-n}} - \xi_{k\cdot 2^{-n}} \in \{0,1\}\} \Longrightarrow$ by (i)+(ii)

$$\mathbf{P}(C_N) \ge \prod_{k=0}^{2^n N-1} \mathbf{P}(\{\xi_{(k+1)\cdot 2^{-n}} - \xi_{k\cdot 2^{-n}} \in \{0,1\}\})$$
$$\ge (e^{-a2^{-n}} + a2^{-n}e^{-a2^{-n}})^{2^n N} \ge (1 - o(a2^{-n}))^{2^n N} \xrightarrow{n\to\infty} 1,$$

since $e^{-x} + xe^{-x} = 1 - o(x)$ as $x \to 0 \Longrightarrow \mathbf{P}(C_N) = 1$.

Finally, the event that the jumps are equal to 1 coincides (by the right continuity) with the event $Z := AB \cap_N C_N$ with $\mathbf{P}(Z) = 1$. $\square$



A trajectory of ξ_t

(III) Wiener process w_t starting from 0 on $T := \mathbb{R}_+$:

(0) $w_0 = 0$.

(i) $\forall 0 \le t_0 < t_1 < \dots < t_n$ r.v. $\Delta w_{t_i, t_{i-1}} := w_{t_i} - w_{t_{i-1}}$ are independent.

(ii) r.v. $\Delta w_{t,s} := w_t - w_s$, $0 \le s \le t$ are Gaussian $\mathcal{N}(0, t-s)$.

(iii) Trajectories of w_t are continuous.

Theorem 1

$\forall 0 \le a \le t_0 < t_1 < \dots < t_n = b$

$$(L_2) \lim_{\mathrm{diam}\{t_i\} \to 0} \sum_{i=0}^{n-1} (w_{t_{i+1}} - w_{t_i})^2 = b - a.$$

Proof. Let $Z := \sum_{i=0}^{n-1} (w_{t_{i+1}} - w_{t_i})^2$. Then by independence

$$EZ = \sum_{i=0}^{n-1} E(w_{t_{i+1}} - w_{t_i})^2 = \sum_{i=0}^{n-1} D(w_{t_{i+1}} - w_{t_i})$$
$$= \sum_{i=0}^{n-1} (t_{i+1} - t_i) = b - a.$$

(see next slide)

Similarly

$$DZ = \sum_{i=0}^{n-1} D(w_{t_{i+1}} - w_{t_i})^2 \quad \text{(evaluaing } \int x^4 e^{-\frac{x^2}{2\sigma^2}} dx \text{ by parts } u = x^3$$

$$= \sum_{i=0}^{n-1} [E(w_{t_{i+1}} - w_{t_i})^4 - (E(w_{t_{i+1}} - w_{t_i})^2)^2] \qquad dv = xe^{-\frac{x^2}{2\sigma^2}} dx)$$

$$= \sum_{i=0}^{n-1} [3(t_{i+1} - t_i)^2 - (t_{i+1} - t_i)^2] = 2 \sum_{i=0}^{n-1} (t_{i+1} - t_i)^2$$

$$\leq 2 \max(t_{i+1} - t_i) \times \sum_{i=0}^{n-1} (t_{i+1} - t_i) = 2(b-a)\mathrm{diam}\{t_i\} \to 0.$$

Thus $E(\sum_{i=0}^{n-1} (w_{t_{i+1}} - w_{t_i})^2 - (b-a))^2 = D \sum_{i=0}^{n-1} (w_{t_{i+1}} - w_{t_i})^2 \to 0$, which implies the convergence in L_2. □

Important observation. An increment of a smooth function is of the same order as the increment of its argument, while the sum of squares of increments goes to 0. In the case of w_t the situation is rather different.

Statistics:

Holder exponent for the Wiener process w_t. Estimate $\frac{w_t - w_s}{|t-s|^\beta}$.

For $t > s$ we have

$E\frac{w_t - w_s}{|t-s|^\beta} = 0, \quad D\frac{w_t - w_s}{|t-s|^\beta} = \frac{t-s}{|t-s|^{2\beta}} = (t-s)^{1-2\beta} \overset{t-s\to 0}{\longrightarrow} 0$ iff $0 < \beta < 1/2$. Further one applies the Chebyshev inequality.

Variation $\mathrm{var}(w_t)$. For $\Delta := \{[t_i, t_{i+1})\}_i^n \subset [a,b]$ denote $V(w_t, \Delta) := \sum_i^n |w_{t_i} - w_{t_{i+1}}|$. Find $(E/D)V(w_t,\Delta) = ?$

$$E|w_{t+h} - w_t| = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} |x| e^{-x^2/2\sigma^2} dx \qquad (\sigma^2 = h)$$

$$= \frac{2\sigma^2}{\sigma\sqrt{2\pi}} \int_0^\infty e^{-x^2/2\sigma^2} d\frac{x^2}{2\sigma^2} = \sqrt{\frac{2h}{\pi}}.$$

Thus for $|t_i - t_{i+1}| = \frac{1}{n}$, $a = t_0 < \cdots < t_n = b$ we have

$$EV(w_t,\Delta) = \sqrt{\frac{2}{\pi}(b-a)} \cdot \sqrt{n} \overset{n\to\infty}{\longrightarrow} \infty, \quad DV(w_t,\Delta) \overset{n\to\infty}{\longrightarrow} b-a.$$

Now again the Chebyshev inequality gives the result.

Claim [Statistics]

Let $f_n(t)$ be piecewise linear with with vertices at points $\{k2^{-n}, \sum_{i=0}^{k-1}(w_{(i+1)2^{-k}} - w_{i2^{-k}})^2\}$. Then $\mathbf{P}(|f_n(t) - t| \xrightarrow{n\to\infty} 0) = 1$ uniformly on $[0, T]$.

Proof. The functions $f_n(t)$ are nondecreasing $\Longrightarrow$ it is enough to prove the convergence on a dense set, say for all $t = k2^{-m}$. Why?

For $n \geq m$ we have

$$E(f_n(t) - t)^2 = 2t2^{-n},\ E\sum_{n\geq 0}(f_n(t) - t)^2 = \sum_{n\geq 0} E(f_n(t) - t)^2 < \infty.$$

Hence by the Chebyshev inequality the series converges with probability 1. Thus $f_n(t) - t \to 0$. $\square$

Theorem (Continuity of trajectories:)

Let $\xi_t, t \in T = [a, b]$ be a random process such that $\exists \alpha, \varepsilon, C > 0$: $E|\xi_t - \xi_s|^\alpha \leq C|t - s|^{1+\varepsilon}$ $\forall t, s \in T$. Then $\exists$ a modification of ξ_t with continuous trajectories.

(III) Multidimensional Wiener process. $w_t := (w_t^1, \dots, w_t^d) \in \mathbb{R}^d$, $t \in \mathbb{R}_+$, $w_0 = x_0$. The definition is exactly the same as in the 1D case, except that the increments $w_t - w_s$ have the covariation matrix diag(t-s) instead of a single number.

Claim

The events from $\mathcal{F}_{w_t^i}$ are independent, which implies that the d-dimensional Wiener process is simply a collection of d independent 1D processes.

Proof. $\forall 0 \le t_1 < \cdots < t_n$ consider random vectors $W^i := (w_{t_1}^i, \dots, w_{t_n}^i)$. Their joint distribution is Gaussian. Hence for independence it is enough to observe that the coordinates of W^i and W^j for $i \neq j$ are uncorrelated. $\square$

The density of the joint distribution of $(w_t^1, \dots, w_t^d)$ is

$$p_{w_t^1,\dots,w_t^n}(x_1, \dots, x_n) = \prod_{i=1}^{n} \frac{1}{(2\pi(t_i - t_{i-1})^2)} \exp\Big(-\sum_{i=1}^{n} \frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})}\Big).$$

Statistics:

Claim

Let $a = t_0 < \cdots < t_n = b$. Then

$$\sum_{i=0}^{n-1}(w^1_{t_{i+1}} - w^1_{t_i})(w^2_{t_{i+1}} - w^2_{t_i}) \xrightarrow{n\to\infty} 0 \text{ in } L_2.$$

Proof. $\tilde{w}_t := (w^1_t + w^2_t)/\sqrt{2}$ is again the Wiener process. Thus

$$\lim \sum_{i=0}^{n-1}(w^1_{t_{i+1}} - w^1_{t_i})(w^2_{t_{i+1}} - w^2_{t_i})$$

$$= \frac{1}{2}[\lim \sum_{i=0}^{n-1} 2(\tilde{w}_{t_{i+1}} - \tilde{w}_{t_i})^2 - \sum_{j=1}^{2} \lim \sum_{i=0}^{n-1}(w^j_{t_{i+1}} - w^j_{t_i})^2]$$

$$= \frac{1}{2}[2(b-a) - (b-a) - (b-a)] = 0. \qquad \square$$

Existence. Let ξ_n be the simple symmetric random walk on $\mathbb{Z}$ with $\xi_0 = 0$ and $\mathbf{P}(\xi_{n+1} = i \pm 1 | \xi_n = i) = 1/2$. We interpolate it and rescale to $[0, 1]$, namely $\forall n \in \mathbb{Z}_+$, $(i_0, i_1, \ldots, i_n) \in \mathbb{Z}^{n+1}$ define
$Z_t^{(n)} := \frac{1}{\sqrt{n}}(1 - nt + [nt])\xi_{[nt]} + \frac{1}{\sqrt{n}}(nt - [nt])\xi_{[nt]+1}$,
which linearly interpolates the points of the rescaled random walk. Let Φ_n be the finite dimensional distribution of the process $Z_t^{(n)}$.

Theorem

$\Phi_n \overset{n\to\infty}{\longrightarrow} \mu^w$ weakly, which is a probability measure on $C([0, 1])$, called the Wiener measure.

$$\mu^w(\{f : \ f(t_i) \in B_i, \ i = 1, 2, \ldots, k\})$$
$$= \prod_{i=1}^k \frac{1}{\sqrt{2\pi(t_{i+1} - t_i)}} \int_{B_i} e^{\frac{x^2}{2(t_{i+1} - t_i)}} dx.$$

Here B_i are measurable sets from $[0, 1]$. The measure μ^w corresponds to a random process called Brownian motion, which satisfies all properties of the Wiener process.

Let ξ_n be the simple symmetric random walk on $\mathbb{Z}$ with $\xi_0 = 0$.

Theorem

A piecewise constant r.p. $Z_t^{(n)} := n^{-1/2}\xi_{[nt]}$, $t \in [0, \infty)$ converges in distribution as $n \to \infty$ to the r.process satisfying the conditions (0-ii) of the Wiener process w_t.

Proof. By the CLT the normalized symmetric random walk $n^{-1/2}\xi_n$ converges in distribution to $\mathcal{N}(0, 1)$ as $n \to \infty$.

Let us check that at a fixed time $t > 0$, the r.p. $Z_t^{(n)}$ converges as $n \to \infty$ to a r.v. with the distribution to $\mathcal{N}(0, t)$:

$$Z_t^{(n)} := n^{-1/2}\xi_{[nt]} = \frac{\xi_{[nt]}}{\sqrt{[nt]}} \frac{\sqrt{[nt]}}{\sqrt{n}},$$

which converges in distribution to a r.v. distributed as $\mathcal{N}(0, t)$. Independence of increments over non-intersecting time intervals follows from the construction. $\square$

Only continuity of trajectories is under question.

Theorem

Let ξ_n be iid r.v. with the standard normal distribution $\mathcal{N}(0,1)$. Then $Z_t := \sqrt{\frac{2}{\pi}} \sum_{k=1}^{\infty} \frac{\sin(kt)}{k} \xi_k$ is the Wiener process on $[0,\pi]$.

Theorem (Feynman-Kac)

The solution of the diffusion equation $\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}$, $\quad u(x,0) := f(x)$ can be represented as $u(x,t) = Ef(x+w_t)$, provided $f \in C^2$.

Proof. Due to the independence of $w_{t+s} - w_t$ and w_t, we have

$$u(x,t+s) = Ef(x+w_{t+s}) = Ef(x+(w_{t+s}-w_t)+w_t)$$
$$= Eu(x+(w_{t+s}-w_t),t) \equiv Eu(x+w_s,t).$$

Therefore,

$$\frac{\partial u}{\partial t}(x,t) = \lim_{s\to 0}\frac{1}{s}(u(x,t+s)-u(x,t))$$
$$= \lim_{s\to 0}\frac{1}{s}E(u(x+w_s,t)-u(x,t))$$
$$= \lim_{s\to 0}\frac{1}{s}\left(\frac{\partial u}{\partial x}Ew_s + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}Ew_s^2 + o(s)\right) = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}.$$

The result follows by noting that $Ew_s = 0$, $Ew_s^2 = s$. □

Moments($d=1$): $Ew_t = E(w_t - w_0) = 0$, $Dw_t = D(w_t - w_0) = t$.
For $0 \le s \le t$ we have
$\operatorname{cov}(w_t, w_s) = E(w_t - w_s)w_s + Ew_s^2 = D(w_s - w_0) = s = t \wedge s$.

Continuity: $E(w_{t+h} - w_t) = 0, D(w_{t+h} - w_t) = h$. Therefore $w_t \to w_{t_0}$ as $t \to t_0$ $\forall t_0$ in probability.

A Brownian bridge is a process B_t whose law is the conditional probability distribution of a Wiener process on $[0, T]$ subject to the condition $w_T = 0$, i.e. $B_t := (w_t | w_T = 0)$, $t \in [0, T]$. Then $EB_t \equiv 0$, but $DB_t = \frac{t(T-t)}{T} \Longrightarrow$ the most uncertainty is in the middle. $\operatorname{cov}(B_t, B_s) = \frac{s(T-t)}{T}$ if $s < t$.
Remark. The increments of a Brownian bridge are not independent.

Representation of the Brownian bridge:

$$B_t = w_t - \frac{t}{T} w_T = \frac{T-t}{\sqrt{T}} w_{\frac{t}{T-t}}.$$

Definition

A d-dimensional random process ξ_t is *Gaussian* if all its finite dimensional distributions are Gaussian, i.e. they are defined by 2 functions $m_t := E\xi_t$ and $R_{s,t} := E(\xi_s - m_s)(\xi_t - m_t)$.

Let $\xi_t, t \in \mathbb{R}_+$ be Gaussian and (0) $\xi_0 = 0$, (a) $E\xi_t = 0$, (b) $E\xi_t\xi_s = \min(t,s) = t \wedge s$, (c) ξ_t is continuous on t a.e.

Claim

ξ_t is a Wiener process.

Proof. $\forall 0 \le t_1 \le \dots \le t_n$ r.v. $(\xi_{t_{i+1}} - \xi_{t_i})$ have a joint Gaussian distribution. (b) implies that the increments are uncorrelated, and the Gaussian distribution implies their independence. Finally,
$E(\xi_t - \xi_s)^2 = E\xi_t^2 + E\xi_s^2 - E\xi_t\xi_s = t + s - 2(t \wedge s) = t - s.$ $\square$

Remark. $E|\xi_t - \xi_s| = \sqrt{\frac{2(t-s)}{\pi}}$.

Claim

$\mathbf{P}(\xi_t \geq a | \tau_a \leq t) = \frac{1}{2}$, where $\xi_{\tau_a} = a > 0,\ \xi_t < a\ \forall t < \tau_a$.

Proof. The event $\xi_t \geq a$ is a subset of $\tau_a \leq t$ and $\mathbf{P}(\xi_t \geq a | \tau_a \leq t) = \frac{\mathbf{P}(\xi_t \geq a)}{\mathbf{P}(\tau_a \leq t)}$. By the symmetry, the probability, that after starting at the point a, to be to the right of a at time t is the same as to be to the left of it. The result follows. $\square$

Corollary. For $t > 0$

$$\mathbf{P}(\tau_a \leq t) = \frac{\mathbf{P}(\xi_t \geq a)}{\mathbf{P}(\xi_t \geq a | \tau_a \leq t)} = 2\mathbf{P}(\xi_t \geq a) = \sqrt{\frac{2}{\pi}} \int_{a/\sqrt{t}}^{\infty} e^{-x^2/2} dx. \quad (*)$$

Hence $\mathbf{P}(\tau_a < \infty) = 1$. Moreover:

$$\mathbf{P}(\max_{0 \leq s \leq t} \xi_s \geq x) = \mathbf{P}(\tau_x \leq t) = \sqrt{\frac{2}{\pi}} \int_{x/\sqrt{t}}^{\infty} e^{-y^2/2} dy$$

$$= \sqrt{\frac{2}{\pi t}} \int_{x}^{\infty} e^{-y^2/(2t)} dy = 2\mathbf{P}(\xi_t \geq x)$$ – the doubled normal law.

Similarly for the minimum value. Observe also that

$$\mathbf{P}(\max_{0 \leq s \leq t} \xi_s > 0) = \mathbf{P}(\min_{0 \leq s \leq t} \xi_s < 0) = 1.$$

Claim The arcsin law for the maximum of ξ_s:

$$\mathbf{P}(\tau_{\max} \le s) = \int_0^s \frac{dy}{\pi\sqrt{y(t-y)}} = \frac{2}{\pi} \arcsin \sqrt{\frac{s}{t}} \quad \text{for } 0 \le s \le t.$$

Proof. After the moment τ_a the process obeys the same laws as when starting from 0. Therefore $\xi_{\max} := \max\limits_{0 \le u \le t} \xi_u \equiv \max\limits_{s \le u \le t} \xi_u$ if $\tau_a = s \le t$, and $\xi_{\max}$ has the same probability distribution as $a + \max\limits_{0 \le u \le t-s} \xi_u$. According to (*) this r.v. has the following conditional probability density:

$$p_{\xi_{\max}}(x|\tau_a = s) = \sqrt{\frac{2}{\pi(t-s)}} \exp\left(-\frac{(x-a)^2}{2(t-s)}\right), \quad a \le x < \infty. \text{ Hence}$$

$$p_{\tau_a,\xi_{\max}}(s,x) = p_{\tau_a}(s) p_{\xi_{\max}}(x|\tau_a = s) = \frac{1}{\pi\sqrt{s(t-s)}} \frac{a}{s} e^{-\frac{a^2}{2s}} e^{-\frac{(x-a)^2}{2(t-s)}}.$$

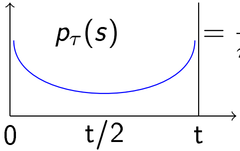
Denote by τ and ξ the (time) position and the value of the global maximum of ξ_u on the interval $[0, t]$. see next slide

The density of the r.v. (τ, ξ) at a point $(\tau = s, \xi = a)$ coincides with the density of (τ_a, ξ) at the same point, since
$p_{\tau,\xi}(s,a) = p_\tau(s|\xi = a)p_\xi(a) = p_{\tau_a}(s|\xi = a)p_\xi(a) = p_{\tau_a,\xi}(s,a)$.

$\Longrightarrow p_{\tau,\xi}(s,a) = \frac{1}{\pi\sqrt{s(t-s)}} \frac{a}{s} e^{-\frac{a^2}{2s}}$ for $0 < s < t$, $\quad 0 < a < \infty$; and

$\Longrightarrow p_\tau(s) = \int_0^\infty p_{\tau,s}(s,x)dx = \frac{1}{\pi\sqrt{s(t-s)}} \int_0^\infty \frac{x}{s} e^{-\frac{x^2}{2s}} dx = \frac{1}{\pi\sqrt{s(t-s)}}$.

Therefore $\mathbf{P}(\tau \le s) = \int_0^s \frac{du}{\pi\sqrt{u(t-u)}} = \frac{2}{\pi} \arcsin\sqrt{\frac{s}{t}}$. $\square$



$= \frac{1}{\pi\sqrt{s(t-s)}}$

Thus the maximum is near one of the end-points.

List 2 – deadline 22.04

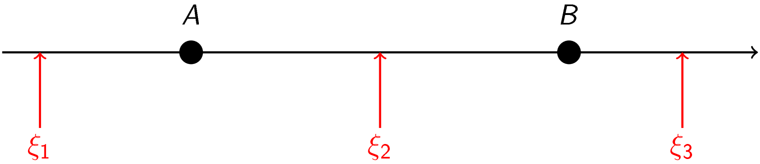
1. Construct uncorrelated but dependent normally distributed r.v.
2. Prove/disprove existence of a Gaussian process $\xi_t,\ 0 \le t \le 1$ with $E\xi_t \equiv 0$ and a correlation function $K(t,s) := t \wedge s - ts$, such that almost all its realizations are continuous.
3. Let w_t be a standard Wiener process, and let $t_i := \frac{i}{n},\ 0 \le i \le n$. Calculate $\lim\limits_{n\to\infty} \mathbf{P}(\sum_{i=0}^{n-1} |w_{t_{i+1}} - w_{t_i}| > n^\alpha)$ as a function of $\alpha \in \mathbb{R}$.
4. Let $\{\xi_i\}_{i=1}^n$ be iid r.v. with $E\xi_i = 0,\ D\xi_i = 1$. Let $\eta_n := \sqrt{n}\frac{\sum_{i=1}^n \xi_i}{\sum_{i=1}^n \xi_i^2}$. Prove that η_n is asymptotically normal as $n \to \infty$.
5. Let $\{\xi_i\}_{i=1}^n$ be iid r.v. and let $\frac{1}{n}\sum_{i=1}^n \xi_i \xrightarrow{n\to\infty} 1$ almost surely. Prove that $E|\xi_1| < \infty$ and calculate $E\xi_1$.
6. Let $\{\xi_i\}_{i=1}^n$ be iid r.v. with $E\xi_i = 0, D\xi_i = \sigma^2 > 0$ and let $\eta_n := \frac{1}{\sigma\sqrt{n}}\sum_{i=1}^n \xi_i$. Prove/disprove existence of $(P)\lim_{n\to\infty} \eta_n$.
7. Find ALL stationary processes $\xi_t,\ t \ge 0$ such that $\exists (P)\lim_{t\to\infty} \xi_t$.

Do not wait until the deadline, and send written solutions (preferably in LaTex) by e-mail.

[58] Lecture 7. Choice of the largest unknown number

The numbers $A \neq B$ are in closed envelopes. I take one at random (say A) and read it. Is it possible to construct an algorithm (deterministic or random) answering the question if the second (unknown) number is larger?

Algorithm. Let ξ be a Gaussian r.v. If $\xi > A$ I decide that $B > A$ and vice versa. How this helps?



The probability to win

$$= \frac{1}{2} \cdot (1 - \mathbf{P}((A-\xi)(B-\xi) < 0)) + 1 \cdot \mathbf{P}((A-\xi)(B-\xi) < 0)$$

$$= \frac{1+\mathbf{P}((A-\xi)(B-\xi)<0)}{2} > \frac{1}{2}.$$

Definition

A *coupling* of measures $\mathbf{P}^i$ on $(\Omega^i, \mathcal{F}^i)$, $i = 1, 2$ is a new measure $\tilde{\mathbf{P}}$ on $(\tilde{\Omega} := \Omega^1 \times \Omega^2, \tilde{\mathcal{F}} := \mathcal{F}^1 \times \mathcal{F}^2)$ such that $\tilde{\mathbf{P}}(A^1 \times \Omega_2) = \mathbf{P}^1(A^1)$, $\tilde{\mathbf{P}}(\Omega^1 \times A^2) = \mathbf{P}^2(A^2)$ $\forall A^i \in \mathcal{F}^i$.

Definition

A *coupling* of r.v. ξ^i, $i = 1, 2$ is a new r.v. $\tilde{\xi} := (\tilde{\xi}^1, \tilde{\xi}^2)$ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ such that its distribution is the coupling of the distributions of ξ^i.

Remark. Couplings are not uniquely defined.

Definition

Let $(\Omega^i, \mathcal{F}^i) = (\Omega, \mathcal{F})$, then the total variation distance $||\mathbf{P}^1 - \mathbf{P}^2||_{tv} := \sup_{A \in \mathcal{F}} |\mathbf{P}^1(A) - \mathbf{P}^2(A)|$.

see next slide

Theorem 1 Coupling inequality

Given r.v. ξ^i, $i = 1, 2$ with probability distributions $\mathbf{P}^i$ for any coupling $\tilde{\mathbf{P}}$ we have $||\mathbf{P}^1 - \mathbf{P}^2||_{tv} \le \tilde{\mathbf{P}}(\tilde{\xi}^1 \neq \tilde{\xi}^2)$.

Proof. $\mathbf{P}^1(\xi^1 \in A) - \mathbf{P}^2(\xi^2 \in A) = \tilde{\mathbf{P}}(\tilde{\xi}^1 \in A) - \tilde{\mathbf{P}}(\tilde{\xi}^2 \in A)$
$= \tilde{\mathbf{P}}(\tilde{\xi}^1 \in A, \tilde{\xi}^1 = \tilde{\xi}^2) + \tilde{\mathbf{P}}(\tilde{\xi}^1 \in A, \tilde{\xi}^1 \neq \tilde{\xi}^2)$
$-\tilde{\mathbf{P}}(\tilde{\xi}^2 \in A, \tilde{\xi}^1 = \tilde{\xi}^2) - \tilde{\mathbf{P}}(\tilde{\xi}^2 \in A, \tilde{\xi}^1 \neq \tilde{\xi}^2) \le \tilde{\mathbf{P}}(\tilde{\xi}^1 \neq \tilde{\xi}^2)$. $\square$

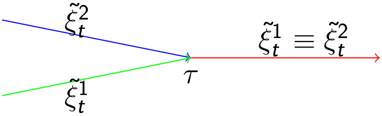
Definition

A *coupling* of r. processes ξ^i_t, $i = 1, 2$ on the same space $(\Omega, \mathcal{F}, \mathbf{P})$ is a new r. process $\tilde{\xi}_t := (\tilde{\xi}^1_t, \tilde{\xi}^2_t)$ on $(\Omega^1 \times \Omega^2, \mathcal{F}^1 \times \mathcal{F}^2, \tilde{\mathbf{P}})$.
$\tau := \inf\{t \in T : \xi^1_t = \xi^2_t\}$ – the coupling time.

see next slide

Definition

A coupling $\tilde{\mathbf{P}}$ is called successful if $\tilde{\mathbf{P}}(\tilde{\xi}_t^1 \neq \tilde{\xi}_t^2) = 0 \quad \forall t \geq \tau$.



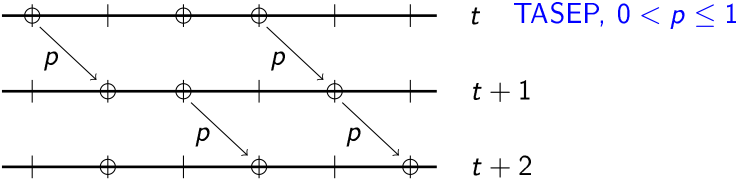
Theorem 2

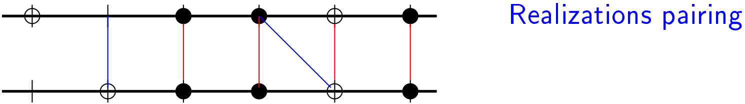
$||\mathbf{P}^1(\xi_t^1 \in \cdot) - \mathbf{P}^2(\xi_t^2 \in \cdot)||_{tv} \leq \tilde{\mathbf{P}}(\tau > t) \quad \forall t \in T$.

Proof. $\{\xi_t^1 \neq \xi_t^2\} \subseteq \{\tau \leq t\}$ by Theorem 1. □

Application: convergence of Markov chains.

Problem. Let $\xi_t^i := a^i w_t^i + b^i$, $i = 1, 2$ and let w_t^i be independent Wiener processes on $\mathbb{R}^1$. Check existence of the successful coupling.





Unpaired particles (empty circles) from different realizations become paired (filled circles) when they share the same position.

Definition

$\xi_t \in \mathbb{R}^d$ is called a *Gaussian* random function if all its finite dimensional distributions are Gaussian.

Problem. When the random sin oscillation $\xi_t := A\cos(\eta t + \varphi)$ is a Gaussian random function?

Definition

$\xi_t \in \mathbb{R}^d$ is called a process with *independent increments* if all its increments over non-intersecting time intervals are independent.

Definition

A similar notion in the broad sense – a process with *uncorrelated increments*: $\mathrm{cov}(\xi_{t_2} - \xi_{t_1},\ \xi_{t_4} - \xi_{t_3}) = 0$ for $t_1 \le t_2 \le t_3 \le t_4$.
Recall that $\mathrm{cov}(\xi, \eta) := E(\xi - E\xi)(\eta - E\eta)$.

Definition

$\xi_t \in \mathbb{R}^d$ is called *stationary* if all its finite dimensional distributions are translationally invariant: $\mu_{\bar{t}+h} = \mu_{\bar{t}}$.

Definition

$\xi_t \in \mathbb{R}^1$ is called *stationary in a broad sense* if the first two moments exist and
$E\xi_{t+h} = E\xi_t, \quad K(t+h, s+h) = K(t,s) := \mathrm{cov}(\xi_t, \xi_s)$.
This is equivalent to $E\xi_t = m, \quad K(t+h, s+h) = K(t-s)$.

Definition

$\xi_t \in \mathbb{R}^1$ is called a process with *stationary increments* if joint distributions of its increments are shift invariant.

Obviously all stationary processes have stationary increments, but not all of them have independent increments. Counterexample: $\xi_t := A\cos(\eta t + \varphi) + \alpha t + \beta$, where φ does not depend on (A, η, α, β) and is uniformly distributed on $[0, 2\pi)$.

Theorem 1

Let $E|\xi_t|^2 < \infty \ \forall t$. Then $\exists (L^2) \lim\limits_{t\to t_0} \xi_t$ iff $\exists \lim\limits_{t,s\to t_0} E\xi_t\xi_s$.

Proof. The necessity follows from the continuity of the scalar product, while the sufficient part follows from the Cauchy condition $\lim\limits_{t,s\to t_0} E|\xi_t - \xi_s|^2 = \lim\limits_{t,s\to t_0} [E|\xi_t|^2 - E\xi_t\xi_s - E\xi_s\xi_t + E|\xi_s|^2] = 0$. □

Claim

Let $\{\xi_n\}$ be uncorrelated. Then $\exists (L^2) \lim \sum_{n\geq 1} \xi_n$ iff the series $\sum_{n\geq 1} E\xi_n$ and $\sum_{n\geq 1} D\xi_n$ converge.

Proof. Let $\eta_n := \sum_{i=1}^n \xi_i \implies K_{\eta\eta}(n,m) = \sum\limits_{i\leq \min(n,m)} D\xi_i$.

Use Theorem 1 above to get the claim. □

Theorem 2

$(P)\lim_{t\to t_0}\xi_t$ exists iff exists the weak $\lim_{t,s\to t_0}\mu_{\xi_t,\xi_s} =: \mu$.

Proof. (a) Necessity. $(P)\lim_{t,s\to t_0}(\xi_t,\xi_s)=(\eta,\eta)$. Hence the 2-dim. distributions μ_{ξ_t,ξ_s} converge weakly.
(b) Adequacy. $\lim_{t,s\to t_0}\mu_{\xi_t,\xi_s}$ is supported by the diagonal (since (ξ_t,ξ_t) is there).
Let $f_\varepsilon\in C^0$, $f_\varepsilon(0)=0$ and $f_\varepsilon(x)=1$ for $|x|>\varepsilon$.
Then by the Chebyshev inequality
$\mathbf{P}(|\xi_t-\xi_s|\geq\varepsilon)\leq Ef_\varepsilon(\xi_t-\xi_s)=\int\int f_\varepsilon(x-y)\mu_{\xi_t,\xi_s}(dxdy)$
$\xrightarrow{t,s\to t_0}\int\int f_\varepsilon(x-y)\mu(dxdy)=0$
since $f_\varepsilon\in C^0$ and μ is supported by the diagonal $\{x=y\}$.
Hence the sequence is fundamental in probability. □

Problem. Prove/disprove that if ξ_t is stationary and $\mathbf{P}(\xi_t=\mathrm{const})=0$, then $(P)\lim_{t\to t_0}\xi_t$ does not exist.

Theorem (Kolmogorov)

Let $F_{t_1,\dots,t_n}(x_1,\dots,x_n)$ be a given family of finite dimensional distributions, satisfying the following consistency conditions:
(a) $F_{t_1,\dots,t_n}(A_1 \times A_2 \cdots \times A_n) = F_{t_{i_1},\dots,t_{i_n}}(A_{i_1} \times \cdots \times A_{i_n})$,
(b) $F_{t_1,\dots,t_{k-1},t_k,t_{k+1},\dots,t_n}(x_1,\dots,x_{k-1},\infty,x_{k+1},\dots,x_n)$
$= F_{t_1,\dots,t_{k-1},\quad t_{k+1},\dots,t_n}(x_1,\dots,x_{k-1},\quad x_{k+1},\dots,x_n)$.
Then $\exists(\Omega,\mathcal{F},\mathbf{P})$ and a random process ξ_t, $t \in T$ such that $\mathbf{P}(\omega:\ \xi_{t_1} \le x_1,\dots,\xi_{t_n} \le x_n) = F_{t_1,\dots,t_n}(x_1,\dots,x_n)$.

Proof. Necessity follows from the definition of the finite dimensional distribution.
Adequacy. $\exists F$ on $(\mathbb{R}^T,\mathcal{B}^T)$, corresponding to the given family of finite dimensional distributions (nontrivial and we skip its proof). Now we choose a new probability space $(\mathbb{R}^T,\mathcal{B}^T,F)$, where each elementary event ω is a function $z_*: T \to \mathbb{R}$. Then the random variable $\xi_t(\omega)$ defined as $\xi_t(\omega = z_*) = z_t$ satisfies all required conditions. $\square$

(I) Existence of a sequence of independent r.v. with given distributions $\{F_n\}$. It is enough to set $F_{t_1,\dots,t_n} := F_{i_1} \times \cdots \times F_{i_n}$.
(II) Existence of a Gaussian random process.

Claim

For any function $m : T \to \mathbb{R}$ and any non-negative definite function $K(t,s) := \sum_{i,j} c_i c_j K(t_i, t_j)$ there exists a Gaussian process ξ_t with $E\xi_t = m(t)$ and $\operatorname{cov}(\xi_t, \xi_s) = K(t,s)$.

Proof. Choose $F_{t_1,\dots,t_n}$ as the n-dimensional Gaussian distribution with the vector of mathematical expectations $(m(t_1), \dots, m(t_n))$ and the covariance matrix $(K(t_i, t_j))$.
To check the consistency conditions, observe that
(i) the Gaussian distribution is completely determined by m, K,
(ii) each sub-vector of a Gaussian vector is again Gaussian. □
Corollary. Existence of the Wiener process follows from this result with $m \equiv 0$, $K(t,s) := t \wedge s$.
Homework: K is non-negative definite?

Definition

ξ_t is *stochastically continuous* at $t_0 \in T$ if $(P)\lim_{t\to t_0} \xi_t = \xi_{t_0}$.

This property is defined by 2-dim. distributions. All above examples of random processes are stochastically continuous. Despite that the realizations of the Poisson process are discontinuous. Explain?
Answer: **P**(a discontinuity happens at a given point)=0.

Claim

Prove/disprove that if ξ_t are independent $\forall t$ and has the same density $p(x)$, then ξ_t is stochastically discontinuous $\forall t$.

Proof. $\mathbf{P}(|\xi_t - \xi_{t_0}| \geq \varepsilon) = \int\int_{|x-y|\geq\varepsilon} p(x)p(y)dxdy$

$$\xrightarrow{\varepsilon\to 0} \int\int_{x\neq y} p(x)p(y)dxdy = \int\int p(x)p(y)dxdy = 1.$$

Hence $\exists\varepsilon > 0$ such that $\mathbf{P}(|\xi_t - \xi_{t_0}| \geq \varepsilon) > 1/2 \Longrightarrow$ there is no convergence in probability. $\square$

Lecture 8. Stochastic continuity in L^p

Definition

ξ_t is *stochastically continuous in L^p* if $(L^p)\lim_{t\to t_0}\xi_t = \xi_{t_0}$.

Problem

Prove/disprove that ξ_t is (a) stochastically continuous on T iff μ_{ξ_t,ξ_s} is weakly continuous on $(t,s)\in T\times T$; and is
(b) stochastically continuous in L^2 iff $E\xi_t\bar{\xi}_s$ is continuous.

(a) Follows from Theorem 1; (b) from Theorem 2 (about continuity).

Problem

Prove/disprove that if ξ_t is stochastically continuous in L^p, $p\geq 1$ on a compact set A, then (a) it is uniformly continuous;
(b) $\sup_{t\in A} E|\xi_t|^p<\infty$.

Follows from standard mathematical analysis arguments.

Definition

A derivative of ξ_t at $t \in T$ is $(\xi_t)' := \lim_{s\to t} \frac{\xi_t - \xi_s}{t-s}$ in various senses.

Theorem T1

Let $E|\xi_t|^2 < \infty$ $\forall t$. Then $\exists (L^2) \lim_{t\to t_0} \xi_t$ iff $\exists \lim_{t,s\to t_0} E\xi_t\xi_s$.

Theorem T2

$(P) \lim_{t\to t_0} \xi_t$ exists iff exists the weak $\lim_{t,s\to t_0} \mu_{\xi_t,\xi_s} =: \mu$.

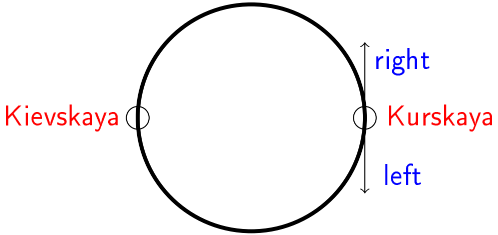
Conditions of the differentiation in probability and in L^2 are given by T1 and T2. Hence the differentiability is defined by finite dimensional distributions of the process of order ≤ 3.

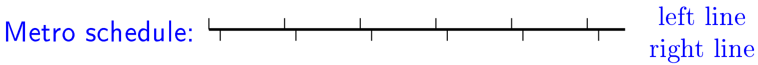
Claim

Wiener process has no derivative even in probability.

Proof. $\frac{\xi_t - \xi_s}{t-s}$ is Gaussian $\mathcal{N}(0, \frac{1}{|t-s|}) \implies (w_t)' := \frac{d}{dt} w_t$ diverges.

Every morning you drive from home to work along the metro ring line from Kurskaya to Kievskaya. Since the distance in both directions is almost the same, you choose the first train in any direction. After a while, you find that you choose the right direction 5 times more often. How can this be explained?





Claim

Poisson process has a derivative in probability, but not in $L^p, p \geq 1$.

Proof. $(P) \lim_{s \to t}(\xi_t - \xi_s)/(t-s) = 0$. Hence if the L^p limit exists, it should be equal to 0 almost surely. However, $E\frac{|\xi_t - \xi_s|^p}{|t-s|^p} \geq \frac{\mathbf{P}(\xi_t \neq \xi_s)}{|t-s|^p} \sim a|t-s|^{1-p}$, which does not vanish as $t - s \to 0$. □

A random function is not uniquely defined by its derivative in probability, while in L^2 the situation is much better.

Claim 1

If $\exists \xi'_t \in L^p$ and $\xi'_s \equiv 0 \ \forall s \in [a,b] \Longrightarrow \xi_t \equiv \xi_a \ \forall t \in [a,b]$.

Proof. $\forall \varepsilon > 0, s \in [a,b] \ \exists O_s : \ |\xi_t - \xi_s| \leq \varepsilon|t-s| \ \forall t \in O_s$.
Assume that $s := \liminf\{t \in [a,b] : \ \xi_t \neq \xi_a\} > a$.
$\Longrightarrow \exists \varepsilon > 0 : \ |\xi_t - \xi_a| > \varepsilon|t-a| \ \ \forall t > s$ and (by continuity of ξ_t)
$|\xi_s - \xi_a| = \varepsilon|s-a| \Longrightarrow$ for $O_s \ni t > s$ we have
$|\xi_t - \xi_a| \leq |\xi_t - \xi_s| + |\xi_s - \xi_a| \leq \varepsilon|t-s| + \varepsilon|s-a| = \varepsilon|t-a|$,
which contradicts to the definition of s. □

Claim 2

$(\xi_t)' \in C^1$ in L^2-sense on (a, b) iff $E\xi_t\xi_s$ has a continuous derivative $\frac{\partial^2 E\xi_t\xi_s}{\partial t \partial s}$ on $(a, b)^2$.

Proof. Follows from standard analysis.

Corollary.

$\frac{\partial^2 K(t,s)}{\partial t \partial s}$ is the correlation function of $(\xi_t)'$, while the joint correlation function of ξ_t and $(\xi_t)'$ is:

$$\begin{pmatrix} K_{\xi\xi} & K_{\xi\xi'} \\ K_{\xi'\xi} & K_{\xi'\xi'} \end{pmatrix}(t, s) = \begin{pmatrix} K_{\xi\xi} & \frac{\partial K_{\xi\xi}}{\partial s} \\ \frac{\partial K_{\xi\xi}}{\partial t} & \frac{\partial^2 K_{\xi\xi}}{\partial t \partial s} \end{pmatrix}(t, s).$$

If $\xi_t \in C^0$ then $\int_a^b \xi_t dt$ can be defined as $\lim \sum_i^{n-1}(t_{i+1} - t_i)\xi_{s_i}$, where $a = t_0 < \cdots < t_n = b$ and non-random points $s_i \in [t_i, t_{i+1}]$. Again everything is ok in $L^p, p \geq 1$-sense but not in probability.

Claim 1

If $\xi_t \in C^0$ in $L^p([a,b])$, then $\exists \int_a^b \xi_t dt$ in terms of L^p-convergence.

Proof. Standard analysis + uniform continuity.

Question

Let τ be a r.v. uniformly distributed on $T := [0,1]$. Check if the process $\xi_t := (1-\tau)^{-1}1_{t>\tau}$ is stochastically continuous on T, and if $\int_0^1 \xi_t dt$ exists in L^2-sense.

One can differentiate the integral over lower and upper limits which yields the Newton-Leibniz formula.

Application: if ξ_t is a Poisson process $\Longrightarrow \frac{d}{dt}\left((L^p)\int_a^t \xi_s ds\right) = \xi_t$.

Realizations of L^p-integrable ξ_t need not be integrable and one needs to distinguish the integral as a r.v. and as a function $\int_a^b \xi_t(\omega)dt$ for $\omega \in \Omega$, which might not be even measurable.

Question

Under which conditions $\left((L^p)\int_a^b \xi_t dt\right)(\omega) \sim \int_a^b \xi_t(\omega)dt$?

Answers:

(a) If all realizations are Riemann-integrable.

Indeed, the L^p-integral is the limit of sums on average, and hence on probability. On the other hand, $J(\omega) := \int_a^b \xi_t(\omega)dt$ has the same limit $\forall\omega$. Now a.e. convergence implies the convergence in probability. Moreover, from the Riemann-integrable assumption $J(\omega)$ is measurable. □

(b) If the realizations are only Lebesgue-integrable the situation is more complex. See next slide

Claim

Let ξ_t is measurable in $[a,b]$ and continuous in $L^p, p \geq 1$-sense. Then $\left((L^p)\int_a^b \xi_t dt\right)(\omega) \sim \int_a^b \xi_t(\omega)dt$ (*).

Proof. It is enough to consider the case $p = 1$ (other convergences will follow). By Fubini's theorem the rhs of (*) exists, since $\int_a^b \int_\Omega |\xi_t(\omega)| dt \mathbf{P}(d\omega) \leq (b-a) \max_{t\in[a,b]} E|\xi_t| < \infty$.
Consider the lhs of (*) as the limit of integral sums while the rhs as the limit of sums $\sum_i \int_{t_{i-1}}^{t_i} \xi_t dt$. Then

$$E|(t_{i+1} - t_i)\xi_{s_i} - \int_{t_{i-1}}^{t_i} \xi_t dt| = E|\int_{t_{i-1}}^{t_i} (\xi_t - \xi_{s_i}) dt|$$
$$\leq \int_{t_{i-1}}^{t_i} E|\xi_t - \xi_{s_i}| dt \leq (t_i - t_{i-1}) \cdot \max_{t\in[t_{i-1},t_i]} E|\xi_t - \xi_{s_i}|.$$

The continuity on average implies the uniform continuity on average. Hence $\exists \delta > 0 \ \ \forall |t-s| < \delta \Longrightarrow E|\xi_t - \xi_s| < \varepsilon$.
Thus $\operatorname{diam}\{t_i\} < \delta \Longrightarrow$ the math. expectation of the difference of integral sums corresponding to lhs and rhs of (*) $\leq \varepsilon(b-a)$. Since $\varepsilon > 0$ is arbitrary, the integrals in both senses coincide a.e. $\square$

Computation of moments of integrals is rather simple.

Claim

Let ξ_t be continuous in L^2-sense and $E|\xi_t|^2 < \infty$. Then
$E \int_a^b \xi_t dt = \int_a^b E\xi_t dt$, $\quad \mathrm{cov}(\int_a^b \xi_t dt, \xi_s) = \int_a^b K_{\xi\xi}(t,s)dt$,
$\mathrm{cov}(\int_a^b \xi_t dt, \int_c^d \xi_s ds) = \int_a^b \int_c^d K_{\xi\xi}(t,s)dtds$.

Proof. Follows from the continuity of the scalar product $\langle \cdot, \cdot \rangle$ on its arguments. Indeed,

$$\begin{aligned} E \int_a^b \xi_t dt &= E\ (L^2) \lim \textstyle\sum_i (t_{i+1} - t_i)\xi_{s_i} \\ &= \langle (L^2) \lim \textstyle\sum_i (t_{i+1} - t_i)\xi_{s_i},\ 1 \rangle \\ &= \lim \langle \textstyle\sum_i (t_{i+1} - t_i)\xi_{s_i},\ 1 \rangle \\ &= \lim \textstyle\sum_i (t_{i+1} - t_i) E\xi_{s_i} \\ &= \int_a^b E\xi_t dt. \end{aligned}$$

This proves the 1-st equality. Two others for homework. $\square$

$K_{\xi\xi}(t,s) = K(t,s) := E(\xi_t - E\xi_t)(\xi_s - E\xi_s)^* \leq \sqrt{K(t,t)\cdot K(s,s)}$.
$K(t,s) = (K(s,t))^*$, $\quad K(t,t) \geq 0$, $\quad \sum_{i,j} K(t_i,s_j) z_i (z_j)^* \geq 0$.
Let us check the last claim for the case $E\xi_t \equiv 0$:

$$\sum_{i,j} K(t_i,s_j) z_i (z_j)^* = E[\sum_{i,j} \xi_{t_i} (\xi_{s_j})^* z_i (z_j)^*]$$
$$= E[\sum_i \xi_{t_i} z_i \cdot \sum_j (\xi_{s_j})^* (z_j)^*] = E|\sum_i \xi_{t_i} z_i|^2 \geq 0$$

In fact this is a characteristic property of the class of correlation functions.

Claim 1

Let $E\xi_t^2 < \infty$. Then ξ_t is L^2-continuous at t_0 iff $K(s,t)$ is continuous at (t_0,t_0).

Proof. For simplicity we assume that $E\xi_t \equiv 0$.

Adequacy:
$$E|\xi_t - \xi_{t_0}|^2 = E(\xi_t - \xi_{t_0})(\xi_t - \xi_{t_0})^*$$
$$= E\xi_t(\xi_t)^* - E\xi_t(\xi_{t_0})^* - E\xi_{t_0}(\xi_t)^* + E\xi_{t_0}(\xi_{t_0})^*$$
$$= K(t,t) - K(t,t_0) - K(t_0,t) + K(t_0,t_0) \xrightarrow{t\to t_0} 0.$$

$\Longrightarrow$ see next slide

Necessity: $|K(t,s)-K(t_0,s_0)|$
$= |E\xi_t(\xi_t-\xi_s)^* + E(\xi_t-\xi_{t_0})(\xi_{t_0})^* + E\xi_{t_0}(\xi_t-\xi_{t_0})^*|$
$\le E|\xi_t-\xi_{t_0}|\cdot|\xi_s-\xi_{t_0}| + E|\xi_t-\xi_{t_0}|\cdot|\xi_{t_0}| + E|\xi_{t_0}|\cdot|\xi_s-\xi_{t_0}|$
$\le \sqrt{E|\xi_t-\xi_{t_0}|^2}\sqrt{E|\xi_s-\xi_{t_0}|^2} + \sqrt{E|\xi_t-\xi_{t_0}|^2}\sqrt{E|\xi_{t_0}|^2}$
$+\sqrt{E|\xi_{t_0}|^2}\sqrt{E|\xi_s-\xi_{t_0}|^2} \xrightarrow{t,s\to t_0} 0.$ $\square$

Corollary

If $K(t,s)$ is continuous on the diagonal $t=s$, then it is continuous $\forall t,s$.

Claim 2

Let $E\xi_t^{(n)} \equiv 0$ and $K_{\xi^{(n)}\xi^{(n)}}(t,s) \xrightarrow{t\to t_0, s\to t_0} 0$. Then $\xi_t^{(n)}$ converges in probability to 0 at t_0.

Proof. Follows from $E|\xi_t^{(n)}|^2 = K_{\xi^{(n)}\xi^{(n)}}(t,t)$. $\square$

1. Let $\xi_n, \xi_0 = x$ and $\tilde{\xi}_n, \tilde{\xi}_0 = \tilde{x}$ be nearest neighbor random walks on $\mathbb{Z}$ with transition probabilities q (left), p (right), $r = 1 - p - q$ (on place) and $(\tilde{p}, \tilde{q}, \tilde{r})$. Find ALL combinations of $(x, \tilde{x}, p, q, \tilde{p}, \tilde{q})$ admitting the successful coupling of $\xi_n, \tilde{\xi}_n$.
   $p = \tilde{p}, q = \tilde{q}$ (a) $(x - \tilde{x} = 2k, pq > 0)$,
   (b) $(x - \tilde{x} = 2k + 1, r(1 - r) > 0)$.
   (a) $\exists \tau : \mathbf{P}(\xi_\tau = \tilde{\xi}_\tau) > 0$. If $pq = 0$ this is not the case.
   (b) If $r(1 - r) = 0$ then $\xi_n - \tilde{\xi}_n$ is always odd $\Longrightarrow$ no intersection. Otherwise as in the case (a).
2. Let $\xi_t, t \in [0, 1]$ be a stochastically continuous random process and let $\varphi : \mathbb{R} \to \mathbb{R}$ be a continuous function. Check the stochastic continuity of $\varphi \circ \xi_t$. Answer: Yes.
3. Let τ be a random variable uniformly distributed on $[0, 1]$ and let $\xi_t := \frac{1_{(\tau,1]}(t)}{t-\tau}$. Check stochastic continuity of ξ_t and existence of $(P) \int_0^1 \xi_t dt$ (i.e. in the probability sense). Answer: Yes (similar to the Poisson process), No (even math. exp. $\nexists$).

Solution: (problem number) Answer. Short proof

1. Answer: $r(1-r)>0$. All of you have a serious problem with the notion "successful coupling". Roughly speaking this means that after a (random) finite time the coupled processes will be "glued together" (become equal). In the situation under study this happens iff $r(p+q)>0$.
2. Answer: Yes. We need to show: $\forall\varepsilon>0\ \exists\sigma>0$ such that $|t-s|\le\sigma \Longrightarrow \mathbf{P}(|\varphi\circ\xi_t-\varphi\circ\xi_s|\ge\varepsilon)\le\varepsilon$.
   For $\varepsilon>0$ choose $\delta>0$ such that $|x-y|\le\delta \Longrightarrow |\varphi(x)-\varphi(y)|\le\varepsilon$. Now we choose $\sigma>0$ such that $|t-s|\le\sigma \Longrightarrow \mathbf{P}(|\xi_t-\xi_s|\ge\delta)\le\varepsilon$ (by stochastic continuity). Thus $|t-s|\le\sigma \Longrightarrow$
   $\Longrightarrow \mathbf{P}(|\varphi\circ\xi_t-\varphi\circ\xi_s|\ge\varepsilon)\le\mathbf{P}(|\xi_t-\xi_s|\ge\delta)\le\varepsilon.$ $\square$
3. Answer: Yes (similar to the Poisson process), No (even math. exp. $\nexists$). Indeed, the probability of the discontinuity at a given time is zero.

$\xi_t : (\Omega, \mathcal{F}, \mathbf{P}) \to (X, \mathcal{B}) := (\mathbb{R}^d, \mathrm{Bor})$ or $= (\mathbb{C}^d, \mathrm{Bor})$.
In the space $L^2(\Omega, \mathcal{F}, \mathbf{P})$ we define the scalar product $\langle \xi_t, \xi_s \rangle := E\xi_t\xi_s^*$. This function defines ξ_t uniquely up to an isometric linear transformation. A centered version of this scalar product is called a *correlation function*

$$K_\xi(t,s) := \mathrm{cov}(\xi_t, \xi_s) := E\xi_t\xi_s^* - E\xi_t(E\xi_s)^*$$
$$= \langle \xi_t, \xi_s \rangle - \langle \xi_t, 1 \rangle \cdot \langle 1, \xi_s \rangle.$$

For a pair ξ_t, η_s one defines a *cross correlation function* $K_{\xi\eta}(t,s) := \mathrm{cov}(\xi_t, \eta_s)$ and a matrix-valued *mutual correlation function* $\begin{pmatrix} K_{\xi\xi} & K_{\xi\eta} \\ K_{\eta\xi} & K_{\eta\eta} \end{pmatrix}(t,s)$.

In these terms one studies the correlation theory of random processes, being curves in a Hilbert space. A serious restriction is that only linear transformations of random processes can be studied this way. To emphasize that only two first moments are taken into consideration, we speak about a theory in the "broad sense".

see next slide

(1) $\xi_t := A\cos(t\eta+\varphi)$, where φ is uniformly distributed on $[0,2\pi)$ and does not depend on A,η. We have $E\xi_t = EA\cdot E\cos\varphi = 0$.
$K_\xi(t) = E\xi_{s+t}\xi_s = EA^2\cos((s+t)\eta+\varphi)\cos(s\eta+\varphi)$
$= \frac{1}{2}[EA^2\cos(t\eta) + EA^2\cos((2s+t)\eta+2\varphi)]$ (blue term=0)
$= \frac{1}{2}\int_0^\infty\int_0^\infty x^2\cos(ty)\Phi_{A\eta}(dxdy) = \int_0^\infty \cos(ty)d\mu(y)$,
where $\mu(B) := \frac{1}{2}\int_0^\infty\int_B x^2\Phi_{A\eta}(dxdy) = \frac{1}{2}EA^2 1_B(\eta)\ \ \forall B\in\mathcal{B}$.
Let ν be a symmetrization on $\mathbb{R}$ of the measure μ, i.e.

$$\nu([0,a]) := \begin{cases}\frac{1}{2}\mu([0,a]) & \text{if } a\ge 0\\ \frac{1}{2}\mu([0,-a]) & \text{otherwise.}\end{cases}$$

Then $K_\xi(t) = \int_{-\infty}^{\infty} e^{ity}d\nu(y)$ is the Fourier transform of the measure ν. For example, if $d\nu(y) := \frac{\pi^{-1}dy}{1+y^2} \Longrightarrow K(t) = e^{-|t|}$.

see next slide

A linear differential operator $P(\frac{d}{dt}) := \sum_{k=0}^{n} a_k \frac{d^k}{dt^k}$ being applied to a stationary ξ_t transforms it to a stationary process $\eta_t := P(\frac{d}{dt})\xi_t$.
$E\eta_t = P(\frac{d}{dt})E\xi_t = a_0 E\xi_t$.
$K_{\eta\eta}(t-s) = P(\frac{d}{dt})P^*(\frac{d}{ds})K_{\xi\xi}(t-s) \Longrightarrow K_{\eta\eta}(t) = P(\frac{d}{dt})P^*(-\frac{d}{dt})K_{\xi\xi}(t)$.
$K_{\eta\xi}(t-s) = P(\frac{d}{dt})K_{\xi\xi}(t-s) \Longrightarrow K_{\eta\xi}(t) = P(\frac{d}{dt})K_{\xi\xi}(t)$.

Examples:

1. $P(x) := x \Longrightarrow K_{\xi'\xi'}(t) = -K''_{\xi\xi}(t)$ and
$$K_{\xi\xi'} = \begin{pmatrix} K_{\xi\xi} & -K'_{\xi\xi} \\ K'_{\xi\xi} & -K''_{\xi\xi} \end{pmatrix}$$
2. $P(x) := 1 + x + x^2 \Longrightarrow P^*(-x) = P(-x) = 1 - x + x^2$,
$P(x)P^*(-x) = 1 + x^2 + x^4 \Longrightarrow \eta_t := P\xi_t = \xi_t + \xi'_t + \xi''_t \Longrightarrow$
$K_{\eta\eta}(t) = K_{\xi\xi}(t) + K''_{\xi\xi}(t) + K^{(4)}_{\xi\xi}(t)$.

Definition

$(S)\int_a^b f(x)\ dg(x) := \lim\limits_{\max|x_i - x_{i-1}|\to 0} \sum_i f(z_i)(g(x_i)-g(x_{i-1})),\ x_{i-1}<z_i\le x_i.$

1. Estimate from above $Q := (S)\int_a^b f(x)\ dg(x)$.
   $Q \le \sup|f| \cdot \lim\limits_{\max|x_i - x_{i-1}|\to 0} \sum_i |g(x_i)-g(x_{i-1})|$
   $= \sup|f| \cdot V_a^b g.$
2. Integration by parts: $\int_a^b f\ dg = fg|_a^b - \int_a^b g\ df$?
   $a = x_0 < x_1 < \cdots < x_n = b,\ a \le z_0 < z_1 < \cdots < z_n \le b$
   $\sum_i f(z_i)(g(x_i)-g(x_{i-1})) = f(b)g(b) - f(a)g(a)$
   $-\sum_i g(x_{i-1})(f(z_i)-f(z_{i-1}))$
3. $\varphi(x) := \mu((-\infty, x]) \in C^0$. Calculate $Q := (L)\int_a^b x\ d\mu$
   $Q = (S)\int_a^b x\ d\varphi(x) = b\varphi(b) - a\varphi(a) - \int_a^b \varphi(x)\ dx$
4. $g_i \in$BV. $\int_a^b f\ dg_1 + \int_a^b f\ dg_2 = \int_a^b f\ d(g_1+g_2)$?

1. Let $K(x)$ be the Cantor staircase on $[0,1]$, i.e. on intervals of the n-th rank it is equal to $2^{-n}, 3\cdot 2^{-n}, 5\cdot 2^{-n}, \dots, 2^{n-1}\cdot 2^{-n}$. Calculate $Q_n := \int_0^1 x^n \, dK(x)$.
Self-similarity: $K(\frac{x}{3}) = \frac{1}{2}K(x)$, $K(\frac{x}{3}+\frac{2}{3}) = \frac{1}{2}K(x)+\frac{1}{2}$.
$Q_n = \int_0^{\frac{1}{3}} x^n \, dK(x) + \int_{\frac{2}{3}}^1 x^n \, dK(x)$
$= 3^{-n}\cdot\frac{1}{2}\left(\int_0^1 y^n \, dK(y) + \int_0^1 (2+y)^n \, dK(y)\right)$
$= 3^{-n}Q_n + \frac{1}{2}3^{-n}\sum_{k=1}^n C_n^k 2^k Q_{n-k}$. $Q_0 = 1$, $Q_1 = \frac{1}{2}$, $Q_3 = \frac{3}{8}$
2. $g \in$BV[a,b], $N := \{x: \ g(x) \neq 0\}$ не более, чем счетно. $f \in C^0[a,b]$. Вычислите $Q := \int_a^b f \, dg$? $= 0$ (Homework)

If $\exists(\xi_t)' \Longrightarrow J(f) := \int_a^b f(t)d\xi_t = \int_a^b f(t)(\xi_t)'dt$.
In general one cannot use this argument. In what follows we assume that ξ_t has uncorrelated increments (hence is not differentiable) and $E\xi_t \equiv 0$. The idea is to represent ξ_t as a sum of infinite number of infinitesimal addends.

Claim

$\exists F(t) \nearrow$ (nondecreasing): $D(\xi_t - \xi_s) = F(t) - F(s)$.

Proof. For an arbitrary $t_0 \in T$ we set $F(t_0) = 0$ and

$$F(t) := \begin{cases} E|\xi_t - \xi_{t_0}|^2 & \text{if } t > t_0 \\ -E|\xi_t - \xi_{t_0}|^2 & \text{if } t \le t_0 \end{cases} = F(s) + E|\xi_t - \xi_s|^2 \Longrightarrow \nearrow.$$

To demonstrate this we consider only the case $t_0 \le s \le t$, then

$$F(t) = E|\xi_t - \xi_{t_0}|^2 = E|\xi_s - \xi_{t_0}|^2 + E(\xi_s - \xi_{t_0})(\xi_t - \xi_s)^* + E(\xi_s - \xi_{t_0})^*(\xi_t - \xi_s) + E|\xi_t - \xi_s|^2 = F(s) + E|\xi_t - \xi_s|^2.$$

(uncorrelated increments!)
ξ_t with uncorrelated increments is L^2-continuous iff $F \in C^0$, and $\exists \lim_{t\to\pm\infty} \xi_t$ iff $|F(\pm\infty)| < \infty$. □

Claim

ξ_t with uncorrelated increments admits L^2-limits at $\forall t$.

Proof. It is enough to show that $\lim\limits_{s,u\to t_-} E(\xi_s-\xi_t)(\xi_u-\xi_t)^*$. $\square$

$F(t)$ generates a (σ)finite measure $\mu((s,t]) := F(t) - F(s)$.
Our aim now is to define the stochastic integral $J(f) := \int_a^b f(t)d\xi_t$ for non-random functions $f \in L^2(dF)$.
We start with piecewise constant (PC) functions on intervals $f|_{[t_i,t_{i+1})} \equiv f_{t_i}$. Then $J(f) := \sum_{i=0}^n (\xi_{t_{i+1}} - \xi_{t_i}) f_{t_i}$.

Claim

$J(af+bg) = aJ(f) + bJ(g)$. Moreover, $EJ(f) = 0$,
$E|J(f)|^2 = \sum_{i=0}^n |f_{t_i}|^2 (F(t_{i+1}) - F(t_i)) = \int_a^b |f(t)|^2 dF(t)$.

see next slide

Claim

$EJ(f)(J(g))^* = \int_a^b f(t)(g(t))^* dF(t).$

Proof. $EJ(f)(J(g))^* = E\sum_i(\xi_{t_{i+1}} - \xi_{t_i})f_{t_i} \times (\sum_j(\xi_{t_{j+1}} - \xi_{t_j})g_{t_j})^*$
$= \sum_{i,j} E(\xi_{t_{i+1}} - \xi_{t_i})(\xi_{t_{j+1}} - \xi_{t_j})^* f_{t_i}(g_{t_j})^*$
$= \sum_i (F(t_{i+1}) - F(t_i))f_{t_i}(g_{t_i})^* = \int_a^b f(t)(g(t))^* dF(t).$ □

Hence we have an isometry between PC-functions in $L^2(dF)$ and a subset of $L^2(\mathbf{P})$, which can be extended by continuity to their closures. Thus $J(f) = (L^2)\lim J(f_n)$ (by the Cauchy principle). Moreover this limit does not depend of the approximation $\{f_n\}$.

Claim 1

The extension coincides with the entire $L^2(dF)$.

Properites: (a) $J(af + bg) = aJ(f) + bJ(g)$, (b) $EJ(f) \equiv 0$,
(c) $E|J(f)|^2 = \int_a^b |f|^2 dF$, (d) $EJ(f)(J(g))^* = \int_a^b f(g)^* dF$.

see next slide

Claim 2

Let $f \in C^0([a,b])$ and let F be right-continuous. Then $J(f) = (L^2) \lim \sum_i (\xi_{t_{i+1}} - \xi_{t_i}) f(s_i)$, where $t_i \le s_i \le t_{i+1}$.

Proof. Set $\tilde{f}(t) := f(t_i)$, $t \in [t_i, t_{i+1})$. Then

$$E|J(f) - J(\tilde{f})|^2 = E|J(f - \tilde{f})|^2 = \int_a^b |f - \tilde{f}|^2 dF$$
$$\le (F(b) - F(a)) \cdot \max_i \mathrm{Osc}(f, [t_i, t_{i+1}))^2 \xrightarrow{\mathrm{diam} \to 0} 0. \quad \square$$

Integration by parts

Let f be continuously differentiable. Then $J(f) := \int_a^b f(t) d\xi_t = f(b)\xi_b - f(a)\xi_a - (L^2) \int_a^b \xi_t f'(t) dt$.

Proof.
$$J(f) = (L^2) \lim \sum_i (\xi_{t_{i+1}} - \xi_{t_i}) f(t_i)$$
$$= (L^2) \lim [f(b)\xi_b - f(a)\xi_a - \sum_{i=0}^n \xi_{t_{i+1}} (f(t_{i+1}) - f(t_i))]. \quad \square$$

(Homework) Prove. Let $f \in L^2([a,b]), \eta_t := \int_{t_0}^t f(s) d\xi_s$. Then η_t is also a process with independent increments, $E\eta_t \equiv 0$, and $G(s) := D(\eta_{\tau+s} - \eta_\tau) = \int_{t_0}^t |f(u)|^2 dF(u)$. Moreover, $\int_a^b g \; d\eta_t = \int_a^b gf \; d\xi_t$ if $g \in L^2([a,b])$.

Problem: Bus waiting time paradox

You are waiting for a regular bus: it seems that buses in the opposite direction go more often, since during your wait, usually several buses pass in the opposite direction.

How to construct a proper mathematical model and to prove this observation?

Heuristics: in reality, buses run in clusters (several in a row - then a long break). Therefore, if the time spread between the buses is large, then most people will wait for a long time.

see next slide

We measure the time t from the moment of departure of the previous bus. Denote by ξ a r.v. which is equal to the waiting time until the next bus, and its density by $f(t)$. Let us find the probability density of the waiting time $\rho(t)$. The probability that a passenger arrives during the time t is proportional to its length. The quantities in question are independent, so $\rho(t) = Ctf(t)$. Normalization: $1 = \int Ctf(t)dt = CE\xi \Longrightarrow C = 1/E\xi$. Therefore the waiting time averaged upon random arrival of passengers is equal to half of the waiting time from the point of view of an individual passenger

$$ET = \frac{1}{2E\xi} \int_0^\infty t^2 f(t)dt = \frac{(E\xi)^2 + D\xi}{2E\xi} > \frac{1}{2}\ E\xi$$

under the assumption that $D\xi > 0$.
Hence, on average, we wait more than half the average time between buses.

Definition

Filtration is a non-decreasing family of σ-algebras $\mathcal{F}_t \subseteq \mathcal{F}, \ t \in T$.

Definition

A *natural filtration* of the random process $(\xi_t, t \in T)$ is $\mathcal{F}_{\le t} := \sigma\{\xi_s, \ s \le t\}$.

Definition

$\tau(\omega) \in T$ is a *Markov moment* wrt $\{\mathcal{F}_t\}$ if $\{\tau \le t\} \in \mathcal{F}_t \ \ \forall t \in T$.

In short, a Markov moment is a random event that you can learn about when you don't know what will happen after that moment. Another name for such objects is a r.v. independent of the future.

Claim

For $(\xi_n, n \in \mathbb{N})$ on $(X, \mathcal{B}) \implies \tau_B := \min(n : \ \xi_n \in B)$ is a Markov moment wrt the natural filtration $\forall B \in \mathcal{B}$.

Proof. $\{\tau_B \le n\} = \cup_{k=1}^n \{\xi_k \in B\} \in \mathcal{F}_{\le n}$. $\square$

Counterexample

$\tau_B := \min(n: \ \xi_{n+k} \in B)$ with $k \geq 1$ is not a Markov moment.

Proof. $\{\xi_{n+k} \in B\}$ needs not belong to $\mathcal{F}_{\leq n}$. $\square$

Definition

$\mathcal{F}_\tau := \{A \in \mathcal{F} : \ \{\tau \leq t\} \cap A \in \mathcal{F}_t \ \forall t \in T\}$ is called a σ-algebra wrt the Markov moment τ.

Properties:

- $\mathcal{F}_\tau$ is a σ-algebra, and $\mathcal{F}_\tau \in \mathcal{F}$.
- the r.v. τ is $\mathcal{F}_\tau$-measurable.
- If $\tau = t = \mathrm{const}$, then $\mathcal{F}_\tau = \mathcal{F}_t$.

Definition

A random process $(\xi_t, t \in \mathbb{R}_+)$ is called a *Levi process* if it has independent stationary increments and $\xi_0 = 0$.

Theorem (strong Markov property)

Let $(\xi_t, t \in \mathbb{R}_+)$ be a Levi process with realizations continuous from the right, and let τ be a Markov moment wrt its natural filtration. Then $\eta_t := \xi_{t+\tau} - \xi_\tau$ has the same finite dimensional distributions and is independent wrt $\mathcal{F}_\tau$.

If $\{\tau_i\}$ - Markov moments $\Longrightarrow \min \tau_i, \ \max \tau_i, \ \lim \tau_i$ are Markov.

Example

ξ_n - symmetric r.w. on $\mathbb{Z}$, $\xi_0 = 0$. $\mathcal{F}_n := \sigma\{\xi_k, \ k \leq n\}$, i.e.
$\mathcal{F}_0 := \{\Omega, \emptyset\}$, $\mathcal{F}_1 := \sigma\{\xi_1 = -1 \cup \xi_1 = 1\}$, ... Then
$\tau := \min\{i : \ |\xi_i| = 2\}$ – the 1st moment when $\xi_n = 2$ is Markov.
$\tau_m := \max\{i \leq 4 \ |\xi_i| = 2\}$ - the last moment is non Markov.

Indeed, $\tau_m \leq 2$ is equivalent to the event $\{|\xi_4| \neq 2\} \notin \mathcal{F}_2$.

Question. Let $\tau(\omega) \leq \tilde{\tau}(\omega)$ be Markov moments $\Longrightarrow \mathcal{F}_\tau \subseteq \mathcal{F}_{\tilde{\tau}}$?

Recall that a non-decreasing family of σ-algebras $\mathcal{F}_t \subseteq \mathcal{F}$ is a *filtration*. The natural filtration wrt ξ_t we denote by $\mathcal{F}_t^{\xi}$.

Definition

$(\xi_t, t \in T)$ is a *martingale* wrt $\mathcal{F}_t$, if
(a) ξ_t is $\mathcal{F}_t$-measurable $\forall t \in T$;
(b) $E|\xi_t| < \infty$ $\forall t \in T$;
(c) $\xi_s = E(\xi_t | \mathcal{F}_s)$ a.e. $\forall s \leq t \in T$

$$\sim \int_A \xi_s d\mathbf{P} = \int_A \xi_t d\mathbf{P} \quad \forall A \in \mathcal{F}_s;$$

and sub/sup martingales if $\xi_s {\leq}{\geq} E(\xi_t | \mathcal{F}_s)$.

Theorem

Let $(\xi_t, t \in T)$ has independent increments. Then ξ_t is a martingale wrt its natural filtration iff $E\xi_t = \mathrm{const}$.
$E\xi_t$ is non-decreasing/increasing, then it is a sub/super-martingale.

Corollaries

- $S_n := \sum_{k=1}^{n} \xi_k$ with independent ξ_k such that $E|\xi_k| < \infty$ is a martingale iff $E\xi_k = 0\ \forall k$.
- $Z_n := \prod_{k=1}^{n} \xi_k$ with independent ξ_k such that $E|\xi_k| < \infty$ is a martingale iff $E\xi_k = 1\ \forall k$.
- Wiener process is a martingale.
- Poisson process ξ_t with the intensity $a > 0$ and $\xi_0 = 0$ is NOT a martingale, but $\xi_t - at$ is a martingale.

Comments:

(1) Let S_n be a biased random walk: $p = \mathbf{P}(+1) \neq \frac{1}{2}$. Then $\eta_n := (q/p)^{S_n}$ is a martingale wrt S_n. Indeed:

$$E(\eta_{n+1}|S_1, \ldots, S_n) = p(q/p)^{S_n+1} + q(q/p)^{S_n-1}$$
$$= q(q/p)^{S_n} + p(q/p)^{S_n} = (q/p)^{S_n} = \eta_n.$$

(2) Direct proof for the Wiener process:

$$E(w_t|\mathcal{F}_{\le s}) = w_s + E(w_t - w_s|\mathcal{F}_{\le s}) = w_s + E(w_t - w_s) = w_s,$$

since $w_t - w_s$ does not depend on events from $\mathcal{F}_{\le s}$.

Claim (Levi martingales)

Let ξ be a r.v. with $E|\xi| < \infty$, and let $\mathcal{F}_t$ be a filtration. Then $\eta_t := E(\xi|\mathcal{F}_t)$ is a martingale.

Proof. For $t > s$ we have
$E(\eta_t|\mathcal{F}_s) = E(E(\xi|\mathcal{F}_t)|\mathcal{F}_s) = E(\xi|\mathcal{F}_s) = \eta_s$ a.e. □

Claim

Let $(\xi_t, \mathcal{F}_t)$ be a martingale, and let g be a convex function, such that $E|g(\xi_t)| < \infty$. Then $(g(\xi_t), \mathcal{F}_t)$ be a sub-martingale.

The convexity implies that g is measurable. Now we use Jensen inequality for the conditional mathematical expectation:
$E(g(\xi_t)|\mathcal{F}_s) \geq g(E(\xi_t|\mathcal{F}_s)) = g(\xi_s)$ a.e. □

If time is discrete, it is enough to check the martingale property only for neighboring time moments, i.e. $E(\xi_n|\mathcal{F}_{n-1}) = \xi_{n-1}$ a.e.
Indeed, $E(\xi_n|\mathcal{F}_k) = E(E(\xi_n|\mathcal{F}_{n-1})|\mathcal{F}_k) = E(\xi_{n-1}|\mathcal{F}_k) = \cdots = \xi_k$.

Definition

$(\xi_n, n \in \mathbb{N})$ is called predictable wrt the filtration $\mathcal{F}_n$, if $\xi_n \quad \forall n$ is measurable wrt $\mathcal{F}_{n-1}$.

Theorem (Doob-Meyer decomposition)

Let $(\xi_n, n \in \mathbb{N})$ with $E|\xi_n| < \infty$ agrees with the filtration $\mathcal{F}_n$. Then there exists a unique representation $\xi_n = M_n + Q_n$, where $(M_n, \mathcal{F}_n)$ is a martingale, while $(Q_n, \mathcal{F}_n)$ is a predictable process.

Claim

Let ξ_t be a martingale wrt $\mathcal{F}_t$ and $E|\xi_t|^2 < \infty$. Then ξ_t is a process with uncorrelated increments.

Proof. Let $s < t < u$, then

$$E(\xi_t - \xi_s)(\xi_u - \xi_t)^* = E(E((\xi_t - \xi_s)(\xi_u - \xi_t)^* | \mathcal{F}_t))$$
$$= E((\xi_t - \xi_s) \cdot [E(\xi_u - \xi_t | \mathcal{F}_t)]^*) = 0.$$

Since $E(\xi_u - \xi_t | \mathcal{F}_t) = 0$ by the definition of a martingale. $\square$

Let $T := \{1, 2, \ldots, N\}$, let ξ_n with $n \in T$ be a sub-martingale wrt σ-algebras $\mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_N$ and let $\tau \in T$ be a Markov moment.

Claim 1

$\xi_\tau \leq E(\xi_N|\mathcal{F}_\tau)$ a.e.

Proof. ξ_τ is measurable wrt $\mathcal{F}_\tau$. Thus it is enough to prove
$\int_A \xi_\tau d\mathbf{P} \leq \int_A \xi_N d\mathbf{P} \quad \forall A \in \mathcal{F}_\tau$.
$A_n := A \cap \{\tau = n\} \in \mathcal{F}_n \Longrightarrow \int_{A_n} \xi_\tau d\mathbf{P} = \int_{A_n} \xi_n d\mathbf{P} \leq \int_{A_n} \xi_N d\mathbf{P}$.
Summing up over $n \in T$ we get the result. $\square$

Claim 2

$E\xi_1 \leq E\xi_\tau$.

Proof. $E\xi_\tau = \sum_{n\in T} \int_{\{\tau=n\}} \xi_n d\mathbf{P}$

$$= \int_\Omega \xi_1 d\mathbf{P} - \sum_n \int_{\{\tau>n\}} (\xi_n - \xi_{n+1}) d\mathbf{P} + \int_{\{\tau>N-1\}} \xi_N d\mathbf{P}.$$

$\{\tau > n-1\} = \Omega \setminus \{\tau \leq n-1\} \in \mathcal{F}_{n-1} \Longrightarrow$
$\int_{\{\tau>n-1\}} \xi_n d\mathbf{P} = \int_{\{\tau>n-1\}} E(\xi_n|\mathcal{F}_{n-1}) d\mathbf{P} \geq \int_{\{\tau>n-1\}} \xi_{n-1} d\mathbf{P}$.
Returning to the 1-st formula we get the result. $\square$

Claim 3 (Homework)

$\xi_\tau \leq E(\xi_\sigma|\mathcal{F}_\tau)$ a.e. $\forall$ Markov moments $\tau \leq \sigma \Longrightarrow E\xi_\tau \leq E\xi_\sigma$.

Is it possible to pass these results to continuous time? The idea is to approximate continuous time Markov moments by finite valued Markov moments $\tau_k \to \tau$, $\sigma_k \to \sigma$. Then, using continuity of the realizations ξ_t, prove that $\xi_{\tau_k} \to \xi_\tau$, $\xi_{\sigma_k} \to \xi_\sigma$, etc.
To some surprise, this is not enough! Let w_t be the Wiener process, and let $\tau := \min\{t : \ w_t = -1\}$. Then $\mathbf{P}(\tau < \infty) = 1$ (Prove).
Nevertheless, $0 = w_0 \neq E(w_\tau|\mathcal{F}_{\leq 0}) = Ew_\tau = -1$.

Claim 3

Let T be a finite segment, and let ξ_t be a martingale with realizations continuous from the right, and let $\tau \leq \sigma$ be Markov moments. Then $\xi_\tau \leq E(\xi_\sigma|\mathcal{F}_\tau)$ a.e.

Claim

Let ξ_t have independent increments, $\xi_0 = 0$, $E\xi_t = 0\ \forall t$ and let $\exists F(t):\ E(\xi_t - \xi_s)^2 = F(t) - F(s)\ \forall s \leq t$. Then $(\xi_t^2 - F(t), \mathcal{F}_{\leq t})$ is a martingale.

Proof. $E(\xi_t^2 - F(t)|\mathcal{F}_{\leq s}) = E((\xi_s + \xi_t - \xi_s)^2 - F(t)|\mathcal{F}_{\leq s})$
$= E(\xi_s^2 + 2\xi_s(\xi_t - \xi_s) + (\xi_t - \xi_s)^2 - F(t)|\mathcal{F}_{\leq s})$
$= \xi_s^2 + E(\xi_t - \xi_s)^2 - F(t) = \xi_s^2 + F(t) - F(s) - F(t) = \xi_s^2 - F(s)$. □

Claim

Let ξ_n with $E\xi_1 = 0$, $E\xi_n^2 < \infty\ \forall n$ be a martingale. Then $\mathbf{P}(\max_{k\leq n} |\xi_k| \geq a) \leq \frac{1}{a^2} E\xi_n^2\ \ \forall a > 0$.

Proof. Let $\tau_n := \min(k \leq n : |\xi_k| \geq a)$ and $\tau_n := n$ if this level was not achieved. This is a Markov moment wrt the natural filtration. Then $\mathbf{P}(\max_{k\leq n} |\xi_k| \geq a) = \mathbf{P}(|\xi_{\tau_n}| \geq a) \leq \frac{1}{a^2} E\xi_{\tau_n}^2$ (Chebyshev). ξ_n is a martingale $\Longrightarrow E\xi_{\tau_n}^2 = E\xi_n^2$. (For a sub-martingale $\leq$.) □

1. Find ALL r.v. $\xi : (\Omega, \mathcal{B}, \mathbf{P}) \to (\mathbb{R}, \mathrm{Bor})$, such that the pair of r.v. ξ, ξ are independent. Answer: $\xi = \mathrm{const}$. If $\xi(\Omega) = A \sqcup B$, then $\mathbf{P}(\xi \in A, \xi \in B) = 0 \neq \mathbf{P}(\xi \in A)\mathbf{P}(\xi \in B)$.
2. Let ξ_t be a L^2-continuous stationary process with $E\xi_t \neq 0 \ \ \forall t$. Prove/disprove existence of a nontrivial r.v. η, such that $\int_0^{t\eta} \xi_s ds$ is a stationary process.
   No. The derivative of a (broad sense) stationary process $\equiv 0$.
3. Calculate the correlation function for the Poisson process with the parameter $a > 0$.
   Answer: $K(t,s) = a \min(t,s)$. Indeed, for $s \leq t$ we have:
   $$\begin{aligned} K(t,s) &= \mathrm{cov}(\xi_t, \xi_s) = \mathrm{cov}(\xi_t - \xi_s + \xi_s, \xi_s) \\ &= \mathrm{cov}(\xi_t - \xi_s, \xi_s - \xi_0) + \mathrm{cov}(\xi_s, \xi_s) \\ &= 0 + \mathrm{cov}(\xi_s, \xi_s) = as. \end{aligned}$$

Solution: (problem number) Answer. Short proof.

Let ξ_t be a random process on $(X, \mathcal{B})$. Introduce σ-algebras
$\mathcal{F}_{\le t} := \sigma\{\xi_s, \quad s \le t \in T\}, \quad \mathcal{F}_{\ge t} := \sigma\{\xi_s, \quad s \ge t \in T\},$
$\mathcal{F}_{[s,t]} := \sigma\{\xi_u, \quad s \le u \le t \in T\}, \quad \mathcal{F}_{=t} := \sigma\{\xi_t, \quad t \in T\}.$
ξ_t is a *Markov* process if $\mathbf{P}(AB|\mathcal{F}_{=t}) = \mathbf{P}(A|\mathcal{F}_{=t})\mathbf{P}(B|\mathcal{F}_{=t})$ a.e.(*)
$\forall t \in T, \; A \in \mathcal{F}_{\le t}, B \in \mathcal{F}_{\ge t}$. By definition $\mathbf{P}(A|\mathcal{F}_{=t}) \equiv \mathbf{P}(A|\xi_t)$.
Given present, past and future are independent.

Claim

$t \to -t$ preserves (*) and it is equivalent to each of
(a) $\mathbf{P}(B|\mathcal{F}_{\le t}) = \mathbf{P}(B|\mathcal{F}_{=t}) \; \forall B \in \mathcal{F}_{\ge t}$
(b) $\mathbf{P}(A|\mathcal{F}_{\ge t}) = \mathbf{P}(A|\mathcal{F}_{=t}) \; \forall A \in \mathcal{F}_{\le t}$

Proof. To derive (a) from (*) we need to show that
$\mathbf{P}(AB) = \int_A \mathbf{P}(B|\mathcal{F}_{=t}) d\mathbf{P} \quad \forall A \in \mathcal{F}_{\le t}, B \in \mathcal{F}_{\ge t}$ (**).
By (*) we have $\mathbf{P}(AB) = E\mathbf{P}(AB|\mathcal{F}_{=t}) = E[\mathbf{P}(A|\mathcal{F}_{=t})\mathbf{P}(B|\mathcal{F}_{=t})]$.

$$\int_A \mathbf{P}(B|\mathcal{F}_{=t}) d\mathbf{P} = E(\mathbf{P}(B|\mathcal{F}_{=t}) \cdot 1_A) = EE(\mathbf{P}(B|\mathcal{F}_{=t}) \cdot 1_A|\mathcal{F}_{=t}))$$
$$= E(\mathbf{P}(B|\mathcal{F}_{=t}) E(1_A|\mathcal{F}_{=t})) = \mathbf{P}(AB).$$

We use that a value measurable wrt $\mathcal{F}_{=t}$ can be taken out of $E()$. $\Longrightarrow$

To derive (*) from (a) we need to show that
$\mathbf{P}(ABC) = \int_C \mathbf{P}(A|\mathcal{F}_{=t})\mathbf{P}(B|\mathcal{F}_{=t})d\mathbf{P} \quad \forall A \in \mathcal{F}_{\le t}, B \in \mathcal{F}_{\ge t}, C \in \mathcal{F}_{=t}$
The rhs $= E[\mathbf{P}(A|\mathcal{F}_{=t})\mathbf{P}(B|\mathcal{F}_{=t}) \cdot 1_C]$.
The lhs $= EE(1_A 1_B 1_C|\mathcal{F}_{\le t}) = E1_A 1_C \mathbf{P}(B|\mathcal{F}_{\le t}) = E1_A 1_C \mathbf{P}(B|\mathcal{F}_{=t})$
since $A \in \mathcal{F}_{\le t}$, $C \in \mathcal{F}_{=t} \subseteq \mathcal{F}_{\le t}$ and due to (a).
Rewriting again as $E(\cdot|\mathcal{F}_{=t})$ we get
$\mathbf{P}(ABC) = EE(1_A 1_C \mathbf{P}(B|\mathcal{F}_{=t})|\mathcal{F}_{=t}) = E(1_C \mathbf{P}(B|\mathcal{F}_{=t})E(1_A|\mathcal{F}_{=t}))$,
which coincides with the rhs (see above).
The proof of the property (b) is similar (homework). □

Problems. Prove/disprove that

1. $\mathbf{P}(\xi_t \in A|\mathcal{F}_{\le s}) = P_s^{t-s}(\xi_s, A)$.
2. $\mathbf{P}(\xi_{t+1} \in A|\xi_t \in B, \xi_{t-1} \in C) = \mathbf{P}(\xi_{t+1} \in A|\xi_t \in B)$
   for a Markov chain ξ_t and $\forall A, B, C \in \mathcal{B}$, $t \ge 1$.
3. w_t, w_{-t} are Markov and calculate transition probabilities.

So far we were considering Markov processes with fixed initial states. Let $\xi_t(\omega) : T \times \Omega \to (X, \mathcal{B})$ be an arbitrary map (this is not a Markov chain yet !). Define 4 σ-algebras:
$\mathcal{F}_T := \sigma\{\xi_t,\ t \in T\}$, $\mathcal{F}_{\le t} := \sigma\{\xi_s,\ s \le t \in T\}$, $\mathcal{F}_{\ge t}$, $\mathcal{F}_{[s,t]}$;
and consider $\forall s \in T, x \in X$ a measure $\mathbf{P}_{s,x}$ on $\mathcal{F}_{\ge s}$.

Definition

A pair $(\xi_t, \mathbf{P}_{s,x})$ is called a *Markov family* with the transition probabilities $P_t^u(\xi_t, A) := \mathbf{P}_{s,x}(\xi_{t+u} \in A | \mathcal{F}_{[s,t]})\ \ \forall s \le t$ if $P_t^u(\cdot, \cdot)$ is a transition probability and $\mathbf{P}_{s,x}(\xi_s = x) = 1$.

i.e., $\mathbf{P}_{s,x}(AB|\xi_t) \stackrel{a.e.}{=} \mathbf{P}_{s,x}(A|\xi_t)\mathbf{P}_{s,x}(B|\xi_t)$, $\ A \in \mathcal{F}_{[s,s+t]}, B \in \mathcal{F}_{\ge s+t}$.

Claim

If $(\xi_t, \mathbf{P}_{s,x})$ is a Markov family, then $\mathbf{P}_{s,x}(B|\mathcal{F}_{[s,t]}) \stackrel{a.e.}{=} \mathbf{P}_{t,\xi_t}(B)$ $\forall s \le t, B \in \mathcal{F}_{\ge t}$ wrt $\mathbf{P}_{s,x}$.

Thus if we fix a random process up to time t, then its behavior after t is the same as if it starts at time t from the point $\xi_t(\omega)$.

Let $\mathbf{P}_{s,x}$ be a probability measure on $\mathcal{F}_{\geq s}$ and let $P_s^u(x, A)$ be a transition probability. Then a pair $(\xi_t, \mathbf{P}_{s,x})$ is a Markov family with this transition function iff finite dimensional distributions of ξ_t wrt $\mathbf{P}_{s,x}$ with $s \leq t_i$ satisfy $\mathbf{P}_{s,x}(\xi_{t_i} \in A_i,\ 1 \leq i \leq n) =$

$$= \int_{A_1} P_s^{t_1 - s}(x, dy_1) \int_{A_2} \dots \int_{A_n} P_{t_{n-1}}^{t_n - t_{n-1}}(y_{n-1}, dy_n).$$

Coordination of finite-dimensional distributions:

Claim (Chapman-Kolmogorov equation)

For $s \leq t \leq u$

$$P_s^{u-s}(x, A) = \int_X P_s^{t-s}(x, dy) P_t^{u-t}(y, A).$$

Proof. By definition $\mathbf{P}_{s,x}(\xi_u \in A) = P_s^{u-s}(x, A)$ for $s \leq t \leq u$. On the other hand, $\mathbf{P}_{s,x}(\xi_t \in X, \xi_u \in A) = \int_X P_s^{t-s}(x, dy) P_t^{u-t}(y, A)$. This implies the result. □

Rewriting in terms of densities we get:

$$p_s^{u-s}(x, z) = \int_X p_s^{t-s}(x, y) p_t^{u-t}(y, z) dy.$$

$P^t(\cdot,\cdot) \equiv P_s^t(\cdot,\cdot) \quad \forall s,t \in T$:
(i) $P^t(x,\cdot) \in \mathcal{M}(X,\mathcal{B})$, (ii) $P^t(\cdot,A)$ - a measurable function,
(iii) $P^0(x,A) = \delta_x(A)$, (iv) $P^{t+s}(x,A) = \int_X P^t(x,dy)P^s(y,A)$.
Correspondingly the operators depend only on one parameter
$P^t f(x) = \int_X P^t(x,dy)f(y)$, $\quad \nu P^t(A) = \int_X \nu(dx)P^t(x,A)$.
(i) P^t - linear contractions in the cone of nonnegative functions,
(ii) $P^t 1 = 1$, $\nu P^t(X) = \nu(X)$, (iii) $P^0 = I$, (iv) $P^{t+s} = P^t P^s$.

Invariant measure is any solution to the equation $\mu = \mu P^t \quad \forall t$.
A pair $(P^t, \mu \equiv \mu P^t)$ is called a *stationary Markov chain*.
Q: $\exists$ a finite (σ-finite) invariant measure for the Wiener process?
For a compact $A \in \mathbb{R}^d$ we have $P^t(x,A) \stackrel{t\to\infty}{\longrightarrow} 0 \Longrightarrow \mu P^t \stackrel{t\to\infty}{\longrightarrow} 0$ if $\mu(X) < \infty \Longrightarrow \mu_{inv} \equiv 0$. If $m = \mathrm{Leb} \Longrightarrow mP^t(A) = \int dx P^t(x,A)$
$= \int dx(\int_A p^t(x,y)dy) = \int_A(\int p^t(x,y)dx)dy = \int_A dy = m(A)$.

Claim

Each one-dimensional distribution Φ_t of a stationary process ξ_t is an invariant measure.

Proof. $\Phi_t(A) = \mathbf{P}(\xi_t \in A) = \mu_{inv}(A)$ $\square$

The idea is that, knowing that $P^t(x, A) = \cdots + o(t)$, we can restore the whole function under certain regularity conditions.

Definition

An *infinitesimal operator* for the semi-group of operators P^t, $P^0 = I$, $t \geq 0$ is $Af := \lim\limits_{t \to 0_+} \frac{P^t f - f}{t} = \frac{d^+}{dt} P^t f|_{t=0}$.

1. Shift to the right defines a semi-group $P^t f(x) := f(x+t)$. $Af(x) = \lim\limits_{t \to 0_+} \frac{f(x+t) - f(x)}{t} = \frac{d^+ f(x)}{dx}$.
2. $\frac{dx_t}{dt} = b(x_t)$, $x_0 = x \in \mathbb{R}^d$. P^t is a shift along solutions. $f(x_t) = f(x + b(x)t + o(t)) = f(x) + \sum_i \frac{\partial f}{\partial x^i} \cdot b^i(x) \cdot t + o(t)$ for $t \to 0$ uniformly on x. Hence $Af(x) := \sum_i b^i(x) \frac{\partial f}{\partial x^i}(x)$.
3. $P^t := e^{tA} := \sum_{n \geq 0} \frac{1}{n!} t^n A^n$, where A is a bounded linear operator on $f \in E$. Then A is the infinitesimal operator. $||\frac{1}{t}(e^{tA} f - f) - Af|| = ||\sum_{n \geq 2} \frac{1}{n!} t^{n-1} A^n f||$ $\leq \sum_{n \geq 2} t^{n-1} ||A^n f|| \leq t||A||^2 ||f|| \left(\sum_{n \geq 0} t^n ||A||^n\right) \xrightarrow{n \to 0} 0.$

Diffusion processes are Markov families $(\xi_t, \mathbf{P}_x)$ on $(\mathbb{R}^d, \mathrm{Bor})$ with continuous realizations such that their infinitesimal operator on C_b^2 (bounded uniformly C^2-continuous functions) is

$Af(x) = Lf(x) = \frac{1}{2}\sum_{i,j} a_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_i b_i(x)\frac{\partial f}{\partial x_i}$, where

$a_{ij}(x), b_i(x) \in C^0$ and (a_{ij}) - symmetric and nonnegative definite.

The differential operator L is called a *generating* operator of ξ_t.

Remark. If we compactify $\mathbb{R}^d$, then $Lf(x)$ allows to prove existence of continuous realizations, but this does not allow to do the same on $\mathbb{R}^d$, i.e. trajectories may go to and return from infinity.

Theorem (A=L)

Let (ξ_t, P_x) be a Markov family on $(\mathbb{R}^d, \mathrm{Bor})$, such that $\forall \varepsilon > 0$ uniformly on x we have:

(i) $P^t(x, \mathbb{R}^d \setminus B_\varepsilon(x)) = o(t)$

(ii) $\int_{B_\varepsilon(x)} (y_i - x_i) P^t(x, dy) = b_i(x)t + o(t)$

(iii) $\int_{B_\varepsilon(x)} (y_i - x_i)(y_j - x_j) P^t(x, dy) = a_{ij}(x)t + o(t)$

Then the infinitesimal operator $Af(x) = Lf(x)$ (as above) on C_b^2.

The meaning of these conditions:
(i) gives sufficient conditions for the existence of a Markov family with a given transition function and continuous trajectories.
(ii) and (iii) describe "truncated" 1st and 2nd moments. Therefore b_i and a_{ij} are called local means and covariations (a transfer vector and a diffusion matrix).

Example 1. Consider a homogeneous Gauss Markov process with the transition density $p^t(x,y) = \frac{1}{\sqrt{2\pi}\sigma_t} e^{-(y-m_t x)^2/(2\sigma_t^2)}$, defined by $m_{t+s} = m_t m_s$, $\sigma_{t+s}^2 = m_t^2\sigma_s^2 + \sigma_s^2$.
Then a general C^0 solution is $m_t = e^{bt/2}$, $\sigma_t^2 = \frac{a}{b}(e^{bt} - 1)$ for $b \neq 0$, and $\sigma_t^2 = at$ otherwise.
Local mean: $\lim_{t\to 0} \frac{1}{t} E_x(\xi_t - x) = \lim_{t\to 0} \frac{1}{t}(m_t * x - x) = \frac{b}{2}x$.
Local variance: $\lim_{t\to 0} \frac{1}{t}\sigma_t^2 = a$. $\Longrightarrow Lf(x) = \frac{a}{2}f''(x) + \frac{b}{2}xf'(x)$.

Example 2. $(w_t,\ \xi_0 + \int_0^t w_s ds)$ with $w_0 = x, \xi_0 = y$. Then at time t this is a Gaussian process with the mean $(x, y + tx)$ and the covariation matrix $\begin{pmatrix} t & t^2/2 \\ t^2/2 & t^3/3 \end{pmatrix}$. This gives $Lf = \frac{1}{2}\frac{\partial^2 f}{\partial x^2} + x\frac{\partial f}{\partial y}$
- a degenerated elliptic operator, since $a_{ij} = 0\ \forall i, j \neq 1$.

Idea of the proof of Theorem (A=L).

According to Taylor's formula

$$f(y) = f(x) + \sum_i \frac{\partial f}{\partial x_i}(y_i - x_i) + \frac{1}{2}\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(y_i - x_i)(y_j - x_j) + o(|x-y|^2).$$

$$P^t f(x) = \int_{B_\varepsilon(x)} P^t(x,dy) f(y) + \int_{\mathbb{R}^d \setminus B_\varepsilon(x)} P^t(x,dy) f(y).$$

By means of (II,III) the 1st integral is equal to

$$\sum_i b_i(x) \frac{\partial f}{\partial x_i} t + \frac{1}{2}\sum_{i,j} a_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j} t + o(t),$$

while the 2nd integral is of order $||f||\, o(t)$, which yields the result.

$J(f) := \int_0^T f(t,\omega)dw_t$ – a principal limitation is a *non-anticipation*, namely $f(t,\omega)$ and $(w_{t+s} - w_t)$ should be independent $\forall t, s > 0$. First we define $J(f)$ for step functions, i.e. for some non-random $0 < t_1 < \cdots < t_n \le T$ we have $f(t,\omega) = f(t_i,\omega)$ for $t_i \le t < t_{i+1}$, thus $f(t,\omega) = \sum_{i=0}^{n-1} \eta_i 1_{[t_i,t_{i+1})}(t)$ with independent η_i. Assume $E\int_0^T |f(t,\omega)|^2 dt < \infty$ and set $J(f) := \sum_i f(t_i,\omega)(w_{t_{i+1}} - w_{t_i})$. It is easy to check that $J(f)$ is correctly defined, i.e. it does not change if we add new points and it is linear on step functions.

Claim

$J(f)$ is isometric, i.e. $E|J(f)|^2 = E\int_0^T |f(t,\omega)|^2 dt$.

Proof.
$$E|J(f)|^2 = E\sum_i |f(t_i,\omega)|^2(w_{t_{i+1}} - w_{t_i})^2 + 2Re\ E\sum_{j<i} f(t_i,\omega)(w_{t_{j+1}} - w_{t_j})(f(t_j,\omega))^*(w_{t_{i+1}} - w_{t_i}).$$
$$E\sum_i |f(t_i,\omega)|^2(w_{t_{i+1}} - w_{t_i})^2 = \sum_i E|f(t_i,\omega)|^2(t_{i+1} - t_i) = E\int_0^T |f(t,\omega)|^2 dt,$$
while the 2nd addend is equal to 0. $\square$

see next slide

The isometric linear transformation $J(f)$ can be extended by continuity from the set of step functions to its closure in $L^2([0,T]\times\Omega)$ preserving the linear isometry. Indeed, if a sequence $\{f_n\}\to f$ is fundamental, then the sequence $\{J(f_n)\}$ is fundamental in L^2 (since distances between the elements are the same).
Now, L^2 is complete, hence $\exists J(f) := (L^2)\lim J(f_n)$.
To finalize the construction one needs to show that the closure of the set of random step functions coincides with L^2. (Homework)
Example. $f(t,\omega)=w_t$. We have $\int_0^T Ew_t^2 dt = \int_0^T t dt = \frac{T^2}{2} < \infty$.
For $0=t_0<t_1<\cdots<t_n=T$ set $f_n(t,\omega)=w_{t_i}$ for $t_i\le t<t_{i+1}$.

$$\int_0^T E[f_n(t,\omega)-w_t]^2 dt = \sum_i \int_{t_i}^{t_{i+1}} E[w_{t_i}-w_t]^2 dt$$

$$= \sum_i (t_{i+1}-t_i)^2 \xrightarrow{diam\{t_i\}\to 0} 0.$$

Hence $\int_0^T w_t dw_t = (L^2)\lim \sum_{i=0}^{n-1} w_{t_i}(w_{t_{i+1}}-w_{t_i})$. We have

$$w_T^2 \equiv \sum_i (w_{t_{i+1}}-w_{t_i})^2 + 2\sum_i (w_{t_{i+1}}-w_{t_i})w_{t_i} \xrightarrow{(L^2)} T + 2\int_0^T w_t dw_t.$$

$$\Longrightarrow \int_0^T w_t dw_t = \frac{1}{2}(w_T^2 - T).$$

Definition

$d\xi_t = f(t,\omega)dw_t + g(t,\omega)dt$ is called a *stochastic differential* of a random process $\xi_t, \;\; t \geq 0$ with values in $(\mathbb{R}^1, \mathcal{B})$ if

– a.a. trajectories ξ_t are continuous,

– f, g satisfies the non-anticipation condition, i.e.
  $f(t,\omega), \;\; g(t,\omega)$ and $(w_{t+s} - w_t)$ are independent $\forall t \geq 0, \;\; s > 0$,

– $f \in L^2((0, T] \times \Omega) \;\; \forall T < \infty$, $g \in L^1_{loc}$ for a.a. ω,

– $\xi_t \overset{a.e.}{=} \xi_0 + \int_0^t f(s,\omega)dw_s + \int_0^t g(s,\omega)ds, \quad t \geq 0.$

How to understand this? $d\xi_t := Lin[\xi_{t+dt} - \xi_t] = gdt + fdw_t$.

Theorem (*Ito formula*)

Let $F(t,x)$ be continuously differentiable on $t \geq 0$, and twice differentiable on x with bounded partial derivatives. Then

$$\begin{aligned} dF(t,\xi_t) &= F'_t dt + F'_x d\xi_t + \tfrac{1}{2} F''_{xx} (d\xi_t)^2 \\ &= \frac{\partial F}{\partial x}(t,\xi_t) f(t,\omega) dw_t \\ &\quad + \left[\frac{\partial F}{dt}(t,\xi_t) + \frac{\partial F}{\partial x}(t,\xi_t) g(t,\omega) + \frac{1}{2} \frac{\partial^2 F}{\partial x^2}(t,\xi_t) f^2(t,\omega) \right] dt. \end{aligned}$$

Remark. fdw_t is of order $\sqrt{dt}$, since $(w_{t+dt} - w_t)$ is Gaussian $\mathcal{N}(0, dt)$, while gdt is of order dt.

A sketch of the proof:

$F(t+dt, \xi_t + d\xi_t) - F(t, \xi_t) = \frac{\partial F}{dt}dt + \frac{\partial F}{dx}d\xi_t + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}(d\xi_t)^2$.
$(d\xi_t)^2 = (fdw_t + gdt)^2 = f^2 \cdot (dw_t)^2 + \ldots$, where the remaining terms are of higher order (since $(dw_t)^2 \sim dt$) and can be neglected.

A real proof is based on the approximation of the functions f, g by step functions $f^{(n)}$, $g^{(n)}$ and the analysis of
$\xi_t^{(n)} := \xi_0 + \int_0^t f^{(n)}(s,\omega)dw_s + \int_0^t g^{(n)}(s,\omega)ds$. (See later)

The multidimensional Ito formula:

$$dF(t, \xi_t) =$$
$$\sum_j \left(\sum_i \frac{\partial F}{dx_i} f_{ij}\right) dw_t^j + \left[\frac{\partial F}{dt} + \sum_i \frac{\partial F}{dx_i} g_i + \sum_{i,j} \frac{\partial^2 F}{\partial x_i \partial x_j} \sum_k f_{ik} f_{jk}\right] dt,$$

where $d\xi_t^i = \sum_j f_{ij}(t,\omega)dw_t^j + g_i(t,\omega)dt$.

This formula should remind you the generating operator of a diffusion process and explain connections with diffusion processes.

$d\xi_t = f(t,\xi_t)dw_t + g(t,\xi_t)dt,\ \xi_0 = C_0.$

Theorem

If f, g are Lipschitz continuous (may be weaken) then the local in time solution of this equation exists and is unique.

The proof is based on a successive approximation method:
$\xi_t^{(n+1)} := \xi_t^{(n)} + \int_0^t f(s,\xi_s^{(n)})dw_s + \int_0^t g(s,\xi_s^{(n)})ds.$ (See later)
Ornstein-Uhlenbeck process – an idea: $ma = m\frac{dv}{dt} + m\dot{w}.$

$d\xi_t = -\xi_t dt + dw_t,\ \xi_0 = C_0.$
Solution:
$d(e^t\xi_t) = e^t\xi_t dt + e^t d\xi_t = e^t\xi_t dt + e^t(-\xi_t dt + dw_t) = e^t dw_t.$
$\Longrightarrow e^t\xi_t = C_0 + \int_0^t e^s dw_s \Longrightarrow \xi_t = C_0 e^{-t} + e^{-t}\int_0^t e^s dw_s.$
This is a Gaussian process with $E\xi_t = C_0 e^{-t}$ and the covariation

$$E(\xi_t - E\xi_t)(\xi_s - E\xi_s) = e^{-t-s}E\left(\int_0^t e^u dw_u \cdot \int_0^s e^u dw_u\right)$$
$$= e^{-t-s}\int_0^{t\wedge s} e^{2u}du = e^{-t-s}(e^{t\wedge s} - 1).$$

We use here the isometric property of the stochastic integral:
$E(\int_0^t F dw_s)^2 = E\int_0^t F^2 dt.$

The drift term $-\xi_t dt$ in the Ornstein-Uhlenbeck process is negative which implies the stability of the zero solution for the unperturbed system.

Let us study the unstable case: $d\xi_t = \xi_t dt + b\xi_t dw_t$

To this end we try to find a solution of the type $\xi_t := e^{cw_t - at}$:

$$d\xi_t = -ae^{cw_t-at}dt + ce^{cw_t-at}dw_t + \frac{1}{2}c^2 e^{cw_t-at}dt$$
$$= (-a + \tfrac{c^2}{2})\xi_t dt + c\xi_t dw_t$$
$$\Longrightarrow -a + \frac{c^2}{2} = 1 \Longrightarrow c = \sqrt{2(1+a)}.$$

$d\xi_t = \xi_t dt + \sqrt{2(1+a)}\xi_t dw_t$. Its solution

$\xi_t := C_0 e^{\sqrt{2(1+a)}w_t - at} \xrightarrow{\to\infty} 0$ if $a > 0$.

Thus for $b \geq \sqrt{2(1+a)} > \sqrt{2}$ the system becomes stochastically stable.

Definition

A solution $\xi_t^{(x)}(\omega)$ with $\xi_0^{(x)} = x$ is said to be stochastically stable if for each $\varepsilon > 0$ we have $\lim_{y\to x} \mathbf{P}(\sup_{t>0} |\xi_t^{(y)} - \xi_t^{(x)}| > \varepsilon) = 0$.

$d\xi_t = f(t,\xi_t)dw_t + g(t,\xi_t)dt,\ \xi_0.$

Definition

A continuous process ξ_t is a *strong solution* to the above SDE if $\xi_t = \xi_0 + \int_0^t g(s,\xi_s)ds + \int_0^t f(s,\xi_s)dw_s$ for a.e. $t \in T$.

Theorem

Let f, g be Lipschitz continuous:
$|f(t,x) - f(t,y)| + |g(t,x) - g(t,y)| \leq C|x-y|,$
$|f(t,x)| + |g(t,x)| \leq C(1+|x|)$ and $D\xi_0 < \infty$.
Then the strong solution of the SDE exists and is unique.

Proof. We use a successive approximation method: $\xi_t^{(n+1)} := \xi_t^{(n)} + \int_0^t f(s,\xi_s^{(n)})dw_s + \int_0^t g(s,\xi_s^{(n)})ds,\ \xi_t^{(0)} := \xi_0$. One needs to check that the rhs is well defined for all $n \geq 1$. We do this for $n = 1$ (the general case follows by induction argument). The functions f, g are measurable on (t,x) and ξ_0 is measurable on ω.

$\Longrightarrow$ see next slide

Hence the functions $f(t,\xi_0(\omega))$ and $g(t,\xi_0(\omega))$ are measurable on (t,ω). By Fubini theorem $g(t,\xi_0(\omega))$ is measurable on t for $\forall\omega$.
Additionally, $|g(t,\xi_0(\omega))| \le C(1+|\xi_0(\omega)|)$.
Hence $\exists \tilde{C}(\omega)$ such that $\int_0^t |g(s,\xi_0(\omega))|ds \le t\tilde{C} < \infty$.
Similarly, $\int_0^t Ef^2((s,\xi_0(\omega))|ds \le 2C^2t(1+E\xi_0^2) < \infty$.
Let us estimate the second moment of $\xi_t^{(1)}$:
$[\xi_t^{(1)}]^2 \le 3[\xi_0^2 + (\int_0^t |g(s,\xi_0(\omega))|ds)^2 + (\int_0^t f(s,\xi_0(\omega))dw_s)^2]$.
We make the estimation by parts

$$E(\int_0^t |g(s,\xi_0(\omega))|ds)^2 \le tE\int_0^t g^2(s,\xi_0(\omega))ds$$
$$\le tE\int_0^t (1+\xi_0(\omega))^2 ds \le 2tE\int_0^t (1+\xi_0^2(\omega))ds \le 2t^2(1+E\xi_0^2).$$

Using the isometric property of the stochastic integral, we get
$E(\int_0^t f(s,\xi_0(\omega))dw_s)^2 = tEf^2(s,\xi_0(\omega)) \le 2t(1+E\xi_0^2)$.
This gives
$E[\xi_t^{(1)}]^2 \le 3[E\xi_0^2 + 2t^2(1+E\xi_0^2) + 2t(1+E\xi_0^2)]$,
which is bounded on any finite time interval.

$\Longrightarrow$ see next slide

To prove the convergence of the sequence $\xi_t^{(n)}$ we estimate

$$E(\xi_t^{(n)} - \xi_t^{(n-1)})^2 \leq 2E(\int_0^t (g(s,\xi_s^{(n-1)}) - g(s,\xi_s^{(n-2)}))ds)^2 + 2E(\int_0^t (f(s,\xi_s^{(n-1)}) - f(s,\xi_s^{(n-2)}))dw_s)^2.$$

The 1st addend can be estimated from above by

$C^2 t \int_0^t E(\xi_s^{(n)} - \xi_s^{(n-1)})^2 ds.$

For the 2nd addend using again the isometric property, we get the same estimate but without the t factor. Thus

$$E(\xi_t^{(n)} - \xi_t^{(n-1)})^2 \leq 2C(1+t)\int_0^t E(\xi_s^{(n)} - \xi_s^{(n-1)})^2 ds.$$

Let us check that the function $\varphi_n(t) := E(\xi_t^{(n)} - \xi_t^{(n+1)})^2$ goes down on n fast enough.

Denoting $a := 1 + E\xi_0^2$, we obtain:

$\varphi_0(t) \leq 2C^2 at(t+1) \leq 2C^2 a(t+1)^2, \ldots$

$\varphi_n(t) \leq 2C^2(t+1)\int_0^t \varphi_{n-1}(s)ds \leq \frac{b^n(t+1)^{2n+2}}{(2n+1)!!},$

where $b = b(C,a)$ is a new constant.

Thus the series $\sum_n \sqrt{\varphi_n(t)}$ converges uniformly on t.

It remains to show that the limit of our construction is the strong solution of the SDE.

$\Longrightarrow$ see next slide

We already know that $\xi_t^{(n)} \stackrel{n\to\infty}{\longrightarrow} \xi_t$ in L^2 sense. Now we need to check the convergence of the rhs's.

$\frac{1}{2}E(\int_0^t (g(s,\xi_s^{(n)}) - g(s,\xi_s))ds + \int_0^t (f(s,\xi_s^{(n)}) - f(s,\xi_s))dw_s)^2$
$\leq E(\int_0^t (g(s,\xi_s^{(n)}) - g(s,\xi_s))ds)^2 + E(\int_0^t (f(s,\xi_s^{(n)}) - f(s,\xi_s))dw_s)^2$
$\leq C^2 t \int_0^t E(\xi_s^{(n)} - \xi_s)^2 ds + C^2 \int_0^t E(\xi_s^{(n)} - \xi_s)^2 ds.$

All estimates are similar to the ones in the first part of the proof. Therefore the rhs also converges to the integral representation of the strong solution. $\square$

$d\xi_t = f(t,\xi_t)dw_t + g(t,\xi_t)dt,\ \xi_0.$

Theorem (uniqueness)

Under the Lipschitz assumption the strong solution is unique.

Proof. Let us check that any strong solutions ξ_t, η_t coincide a.s. $\xi_t - \eta_t = \int_0^t (f(s,\xi_s) - f(s,\eta_s))dw_s + \int_0^t (g(s,\xi_s) - g(s,\eta_s))ds.$ Denoting by $Z_t^{(K)}$ a r.v. being equal to 1 if $|\xi_t|, |\eta_t| \le K$ and to 0 otherwise, we get:

$$\begin{aligned} EZ_t^{(K)}|\xi_t - \eta_t|^2 \le 2EZ_t^{(K)} &\left(\int_0^t Z_s^{(K)}(f(s,\xi_s) - f(s,\eta_s))dw_s\right)^2 \\ +2EZ_t^{(K)} &\left(\int_0^t Z_s^{(K)}(g(s,\xi_s) - g(s,\eta_s))dw_s\right)^2 \\ \le \tilde{C} &\int_0^t EZ_t^{(K)}|\xi_s - \eta_s|^2 ds. \end{aligned}$$

Now we use the classical Gronwall inequality:
$x(t) \le C\int_0^t x(s)ds + h(t) \Longrightarrow x(t) \le h(t) + C\int_0^t e^{C(t-s)}h(s)ds.$
Hence for $h(t) \equiv 0$ we get $EZ_t^{(K)}|\xi_t - \eta_t|^2 = 0$, which implies the result due to the continuity of the processes ξ_t, η_t. □

Definition

A solution $\xi_t^{(x)}(\omega)$ with $\xi_0^{(x)} = x$ is said to be stochastically stable if for each $\varepsilon > 0$ we have $\lim_{y \to x} \mathbf{P}(\sup_{t>0} |\xi_t^{(y)} - \xi_t^{(x)}| > \varepsilon) = 0$.

Consider the Ito SDE in $\mathbb{R}^d$ with the generating operator $LV(t,x) := \frac{\partial V}{\partial t} + \sum_{i=1}^d b_i(t,x) \frac{\partial V}{\partial x_i} + \frac{1}{2} \sum_{i,j}^d a_{ij}(t,x) \frac{\partial^2 V}{\partial x_i \partial x_j}$.

Theorem

Let $b_i(0,x) = bx_i$, $a_{ij}(0,x) = C^2 x_i x_j$ and let the Lipshitz assumptions hold true. Then $\xi_t \equiv 0$ is a trivial solution of the SDE. Assume also that the exists a (Lyapunov) function $V(t,x) \geq \tilde{V}(x) > 0$ for $x \neq 0$ and $V(t,0) = 0$, satisfying the inequality $LV(t,x) \leq 0$. Then the trivial solution of the SDE is stochastically stable.

Example: $d\xi_t = b\xi_t dt + C\xi_t dw_t$.

Different types of stochastic integrals: $\int_0^T f(s,\omega)dw_s$

$J(f) := \lim \sum_i f(t_i,\omega)(w_{t_{i+1}} - w_{t_i})$

$\tilde{J}(f) := \lim \sum_i f(t_{i+1},\omega)(w_{t_{i+1}} - w_{t_i})$

Will they coincide?

If $f(t,\omega) := w_t(\omega)$, then

$E(\tilde{J}(f) - J(f)) = \lim \sum_i E(w_{t_{i+1}} - w_{t_i})^2 = T \neq 0$

$J_S(f) := \lim \sum_i f(\frac{t_i+t_{i+1}}{2},\omega)(w_{t_{i+1}} - w_{t_i})$ - Stratonovich integral.

$d\xi_t = A\xi_t dt + B\xi_t dw_t$

$F(t,x) := \log x \Longrightarrow F'_t = 0, \quad F'_x = 1/x, \quad F''_x = -1/x^2.$

Then by Ito formula we get:

$$d\log\xi_t = \frac{1}{\xi_t}A\xi_t dt + \frac{1}{\xi_t}B\xi_t dw_t - \frac{1}{2\xi_t^2}B^2\xi_t^2 dt$$
$$= (A - B^2/2)dt + Bdw_t.$$

$\Longrightarrow \xi_t = \xi_0 e^{(A-B^2/2)t+Bw_t}.$

$d\xi_t = dt + 2\sqrt{\xi_t}dw_t$

Then by Ito formula we get:

$dF = (F'_t + F'_x + 2\xi_t F''_x)dt + 2\sqrt{\xi_t}F'_x dw_t$

Set $F(t,x) := p(t)\sqrt{x} + q(t)$.

Hence the 1st term (without dt) becomes

$$\sqrt{\xi_t}p'_t + q'_t + \frac{p}{2\sqrt{\xi_t}} - 2\xi_t\frac{p}{2\xi_t\sqrt{\xi_t}} = \sqrt{\xi_t}p'_t + q'_t$$

Choose $p = 1$, $q = 0$. Then $F = \sqrt{\xi_t} \Longrightarrow dF = dw_t$.

Finally, $\sqrt{\xi_t} = \xi_0 + w_t \Longrightarrow \xi_t = (w_t + \sqrt{\xi_0})^2$.