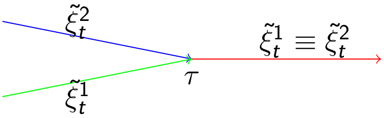
[1] Course “Introduction to Stochastic Processes”

Michael Lvovich Blank

HSE 2020 (winter)



Lectures: Michael Lvovich Blank – <blank@iitp.ru>
Seminars: Mauro Mariani – <mariani.mau@gmail.com>
Assistent: Nikita Puchkin – <npuchkin@gmail.com>

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

15:30–14:50 + 17:00–18:20 at 211

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

Questions:

(a) Does the game stops in finite time?

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

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

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: 101100. . . .
Corollary. A wins iff $\xi_1\xi_2\xi_3 = 000$.
Hence $\mathbf{P}(A - \text{wins}) = 2^{-3} = 1/8 < 7/8 = \mathbf{P}(B - \text{wins})$.

(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.
$100 \sim 011 > 000 \sim 111$, $001 \sim 110$, $101 \sim 010$, $100 - ? - 101$

Table of winning probabilities for the case $|A| = |B| = 3$ (courtesy of Anna Tutubalina and Martin Gardner):

A/B	000	001	010	011	100	101	110	111
000	–	1/2	2/5	2/5	1/8	5/12	3/10	1/2
001	1/2	–	2/3	2/3	1/4	5/8	1/2	7/10
010	3/5	1/3	–	1/2	1/2	1/2	3/8	7/12
011	3/5	1/3	1/2	–	1/2	1/2	3/4	7/8
100	7/8	3/4	1/2	1/2	–	1/2	1/3	3/5
101	7/12	3/8	1/2	1/2	1/2	–	1/3	3/5
110	7/10	1/2	5/8	1/4	2/3	2/3	–	1/2
111	1/2	3/10	5/12	1/8	2/5	2/5	1/2	–

"Best" patterns for the case of an unfair coin with $\mathbf{P}(1) = p \neq \frac{1}{2}$?

r.v. $\xi : (\Omega, \mathcal{F}, \mathbf{P}) \to (X, \mathcal{B})$ – a measurable map, $\forall x \in \mathcal{B}$.
$\xi \sim \eta \iff \mathbf{P}(\xi \neq \eta) := \mathbf{P}(\{\omega \in \Omega : \ \xi(\omega) \neq \eta(\omega)\}) = 0$ – equivalence. $\mathcal{F}_\xi := \sigma(\xi_{-1}\mathcal{B})$ – σ-algebra generated by ξ.
$\mathcal{M}(X) \ni \Phi_\xi(A) := \mathbf{P}(\xi \in A), \ A \in \mathcal{B}$ – distribution of ξ.
$\Phi_{\xi_1,\dots,\xi_n}(A) := \mathbf{P}((\xi_1, \dots, \xi_n) \in A), \ A \in \mathcal{B}^n$ – joint distribution.
$\xi_1, \dots, \xi_n$ are *independent* if $\Phi_{\xi_1,\dots,\xi_n} = \prod_i \Phi_{\xi_i}$.

$$\mathbf{P}(\xi \in A, \eta \in B) = \mathbf{P}(\xi \in A)\mathbf{P}(\eta \in B)$$

If $n = \infty$ independence if $\forall k < \infty$ of them 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.
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 $L^p(\Omega, \mathcal{F}, \mathbf{P})$: $||\xi_n - \xi||_p \overset{n\to\infty}{\longrightarrow} 0$.
- weak: $\Phi_{\xi_n} \overset{n\to\infty}{\longrightarrow} \Phi_\xi$ in the weak sense.

Example of independence in pairs but not jointly.
$\{A_i\}_{i=1}^4$ – independent events with $\mathbf{P}(A_i) = 1/4 \ \forall i$.
$\{B_j := \{A_j, A_4\}\}_{j=1}^3 \implies \mathbf{P}(B_j) = 1/4 \ \forall j$.
$\mathbf{P}(B_i \cap B_j) = \mathbf{P}(A_4) = 1/4 = \mathbf{P}(B_i)\mathbf{P}(B_j) \ \forall i \neq j$ (pairs).
$\mathbf{P}(\cap_{j=1}^3 B_j) = \mathbf{P}(A_4) = 1/4 \neq 1/64 = \prod_{j=1}^3 \mathbf{P}(B_j)$ (joint).

$\Omega := [0,1], \mathbf{P} := \mathrm{Leb} = m, \quad \xi, \eta : (\Omega, \mathrm{Bor}, m) \to (\mathbb{R}, \mathrm{Bor})$.
Check independence: $\quad \mathbf{P}(\xi \in A, \eta \in B) = \mathbf{P}(\xi \in A)\mathbf{P}(\eta \in B)$
(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) $\xi(\omega) := 1_I(\omega)\cos(2a\pi\omega), \ \eta(\omega) := 1_J(\omega)\cos(2b\pi\omega), \ I, J \in \mathrm{Bor}$

Let $\mathrm{Bin}(n,p) := \sum_{i=1}^{n} \xi_i,\ \xi_i \in \{0,1\},\ \ \mathbf{P}(\xi_i = 1) = p.$
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 $\pi_n = ((1-2p)^n + 1)/2 \overset{n\to\infty}{\longrightarrow} 1/2.$

$Ef(\xi) := \mathbf{P}(f \circ \xi) := \int\limits_{\Omega} f(\xi(\omega))\mathbf{P}(d\omega)$, where $f : X \to \mathbb{R}$.
Variance $Df(\xi) = E(f(\xi) - Ef(\xi))^2 = \mathbf{P}((f \circ \xi - Ef(\xi))^2)$,
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.
Claim. Let ξ, η be independent with densities f_ξ, f_η, then
$f_{\xi+\eta}(x) = f_\xi * f_\eta(x) := \int f_\xi(t) f_\eta(x - t)dt$.

(1) Let $\{\xi_i\}_{i=1}^n$ be independent r.v. with $E\xi_i = 0, D\xi_i < \infty$ and let $\eta_k := \sum_{i=1}^k \xi_i$. Prove/disprove that
$\mathbf{P}(\max_{k \le n} |\eta_k| \ge \varepsilon) \le 2\mathbf{P}(\eta_n \ge \varepsilon - \sqrt{2D\eta_n}) \quad \forall \varepsilon \in \mathbb{R}$.
(2) Let $\varphi(x) = \varphi(-x) \ge 0$ be a nonincreasing for $x \ge 0$ function, and let ξ, η be r.v. Prove/disprove
$\mathbf{P}(|\xi| \le \varepsilon) \ge \mathbf{P}(|\eta| \le \varepsilon) \quad \forall \varepsilon \ge 0 \Longrightarrow E\varphi(\xi) \ge E\varphi(\eta)$ and vice versa.
(3) Let $\{\xi_i\}_{i=1}^n$ be independent r.v. with the same distribution function $F(x)$. Let $\xi_- := \min_i \xi_i, \xi_+ := \max_i \xi_i$. Find the distribution function of the vector (ξ_-, ξ_+).
(4) Let ξ be a r.v. with the median m_ξ. Prove/disprove that
$m_{\varepsilon\xi} = \varepsilon m_\xi \quad \forall \varepsilon \in \mathbb{R}$.
(5) Let $\{\xi_i\}_{i=1}^n$ be independent r.v. with $E\xi_i = a, D\xi_i = \sigma^2 < \infty$ and let $\eta_n := \frac{1}{n}\sum_{i=1}^n \xi_i$. Find a function $\varphi(n)$ such that
$E\varphi(n)\sum_{i=1}^n (\xi_i - \eta_n)^2 = \sigma^2$.
(6) Let A be a $n \times n$ matrix with independent random entries a_{ij} with $Ea_{ij} \equiv 0, \ Da_{ij} \equiv \sigma^2$. Calculate $D(\det A)$.
(7) Let $\{\xi_i\}_{i=1}^n$ be iid r.v. with $0 < D\xi_i < \infty$. Find all possible values of the function $\varphi(x) := \lim_{n\to\infty} \mathbf{P}(\sum_{i=1}^n \xi_i < x), \quad x \in \mathbb{R}$.

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

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 $[0,1]$.

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

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

$\omega_4\ \omega_8\ \dots$

$\omega_7\ \dots$

$\dots$

$\xi_{nk}(\omega)$ is the (n,k) element of this triangle table.

Claim 3. $\{\xi_n := \sum_{k\geq 1} 2^{-k}\xi_{nk}\}_{n\geq 1}$ are uniformly distributed iid r.v. on $X := [0,1]$.

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 $\tilde{F}^{-1}(t) := \inf\{s : \; F(s) \geq t\}$ – a generalized inverse function.

Claim 5. Let $\{F_k\}_{k\geq 1}$ be an arbitrary sequence of distribution functions. Then there 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 $\{\xi_k\}$.

Independent events: $\mathbf{P}(A,B) = \mathbf{P}(A)\mathbf{P}(B)$. Possibilities:
$A \cap B = \emptyset, A = B, A \cap B \neq \emptyset$ – which is correct?
$\exists$ of independent events in $([0,1], \mathrm{Bor}, \mathbf{P})$?
$\mathbf{P}(0) = \frac{1}{2}, \mathbf{P}(\frac{1}{2}) = \frac{1}{3}, \mathbf{P}(1) = \frac{1}{6}$ – no independent events.

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

$E\xi = \sum_i \mathbf{P}(B_i)E(\xi|B_i)$ if $\sqcup_i B_i = \Omega$, $\Delta := \{B_i\} \in \mathcal{F}$.
$\mathbf{P}(A|\Delta)(\omega) := \sum_i \mathbf{P}(A|B_i) \cdot 1_{B_i}(\omega)$ – random variable:
(a) $A \cap B = \emptyset \Longrightarrow \mathbf{P}(A \cup B|\Delta) = \mathbf{P}(A|\Delta) + \mathbf{P}(A|\Delta)$
(b) $\mathbf{P}(A|\Omega) = \mathbf{P}(A)$, $(c)E(\mathbf{P}(A|\Delta)) = \mathbf{P}(A)$.
If $\#(\eta(\Omega)) < \infty$ then $\exists\Delta_\eta := \{B_i\}$ – 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) $E(a\xi + b\eta|\Delta) = aE(\xi|\Delta) + bE(\eta|\Delta)$ (b) $E(\xi|\Omega) = E(\xi)$
(c) $E(1_A|\Delta) = \mathbf{P}(A|\Delta)$ (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:
(a) $E(\xi|\mathcal{A})$ is $\mathcal{A}$-measurable, (b) $\mathbf{P}(\xi \cdot 1_A) = \int_A E(\xi|\mathcal{A}) d\mathbf{P}$.
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 $E(\xi|\mathcal{B}) = \sum_i z_i 1_{B_i} = \sum_i E(\xi|B_i) \cdot 1_{B_i} = E(\xi|\Delta)$. QED

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

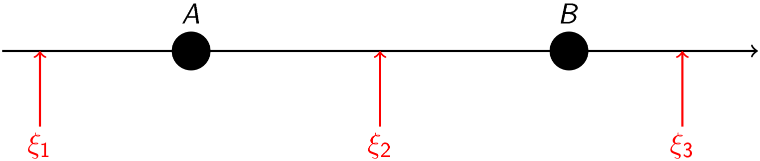
$E(\xi|\eta) := E(\xi|\sigma(\eta))$. Here $E(\xi|\eta = x)(x)$ – conditional ME.
Conditional density $p_\xi(x|\eta = y) = p_{\xi\eta}(x, y)/p_\eta(y)$:
$E(f(\xi, \eta)|\eta = y) = \int f(x, y) p_\xi(x|\eta = y) dx$ for $f \in L^1$.

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 prob. measure on $\mathbb{R}^2$ with $p > 0$ ($P(A) := m(1_A \cdot p)$ 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.
$E(\xi|\mathcal{B}_x) = m_y(\xi(x,\cdot)\rho(x,\cdot)) = \frac{\int \xi(x,y)p(x,y)dy}{\int p(x,y)dy}$.

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

Test 1

(1) $(\Omega, \mathcal{F}, \mathbf{P})$, $\Omega = [0,1]$, $\mathcal{F} = \mathrm{Bor}$, $\mathbf{P} = \mathrm{Leb}$,
$\{\xi_n := \omega^n\}_{n \in \mathbb{Z}_+}$, $\omega \in \Omega$. $A_n := \{\omega \in \Omega : \xi_n \leq 1/n\}$
Calculate: $\mathbf{P}(\cup_{n\geq 1} A_n) = 1$, $\mathbf{P}(\cap_{n\geq 1} A_n) = 1/\sqrt[3]{3}$.
Solution: $A_n = \{0 \leq \omega \leq 1/\sqrt[n]{n}\} \Longrightarrow A_1 = [0,1]$,
further use that $\sqrt[n]{n} \to 1$ and $\max_n \sqrt[n]{n} = \sqrt[3]{3}$.

(2) Find all $a, \varepsilon \geq 0$ such that $\mathbf{P}(\xi \geq a) \leq e^{-a-\varepsilon} E e^{\xi}$, $\forall \xi$.
$\mathbf{P}(\xi \geq a) \leq e^{-a} E(e^{\xi} \cdot 1_{\xi \geq a}) \leq e^{-a} E e^{\xi}$ (Chernoff inequality)
Answer: $\varepsilon = 0, \forall a \geq 0$, since for $\xi \equiv e^a \Longrightarrow \mathbf{P}(\xi \geq a) = 1 > e^{-\varepsilon}$.

(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: $\xi(\omega) := 1_{\{1,2\}}(\omega)$, $\eta(\omega) := 1_{\{2,3\}}(\omega)$
$\mathbf{P}(\xi = 1) = \mathbf{P}(\eta = 0) = \frac{1}{2}$,
$\mathbf{P}(\xi = 0, \eta = 1) = \mathbf{P}(\omega \in \{3\}) = \frac{1}{4} = \mathbf{P}(\xi = 0)\mathbf{P}(\eta = 1) \quad \ldots$

Idea of large deviations, from r.v. and up to random DS.
Chebyshev inequality with $\varphi \nearrow$ – nondecreasing:
$\mathbf{P}(\xi \geq t) = \mathbf{P}(\varphi(\xi) \geq \varphi(t)) \leq \frac{E\varphi(\xi)}{\varphi(t)}$. 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:
$M_\xi(t) := Ee^{t\xi} = 1 + tE\xi + \frac{t^2}{2}E\xi^2 + \dots + \frac{t^n}{n!}E\xi^n + \dots$
Generating (proizvodyaschaya) function: $Ez^\xi, \quad |z| < 1$,
Characteristic function: $Ez^{it\xi}$.

Properties of $M_\xi(t)$: (a) $E\xi^n = M_\xi^{(n)}(0)$ (if $\exists$ near 0)
(b) $M_\xi(t) = M_\eta(t) \quad |t| < \delta \Longrightarrow \xi = \eta$ (on distribution)
(c) ξ, η independent $\Longrightarrow M_{\xi+\eta} = M_\xi M_\eta$
Proof. $M_{\xi+\eta}(t) = Ee^{t(\xi+\eta)} = Ee^{t\xi}e^{t\eta} = M_\xi M_\eta$. QED

T1. ξ_i iid $\mathbf{P}(\xi_i = \pm 1) = \frac{1}{2} \Longrightarrow \mathbf{P}(|\sum_{k=1}^n \xi_k| > \varepsilon) < 2e^{-\varepsilon^2/(2n)}$.
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 $\mathbf{P}(S_n > \varepsilon) \le e^{-\varepsilon^2/(2n)}$.
QED

T2. 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 $t > np$.

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

T3. 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 $E\xi_i = 0$ and due to symmetry that $\mathbf{P}(S_n \ge \varepsilon\sigma) \le e^{-t\varepsilon\sigma/2}$ for $t = \min(\varepsilon/(2\sigma), 1)$.
$\sum^n D\xi_i = \sigma^2$, hence by the Lemma:
$\mathbf{P}(\xi < \varepsilon\sigma) \le e^{-t\varepsilon\sigma} \prod_{i=1}^n Ee^{t\xi_i} \le e^{-t\varepsilon\sigma} \prod_{i=1}^n e^{t^2 D\xi_i} = e^{-t\varepsilon\sigma + t^2\sigma^2}$.
Thus choosing $t \le \frac{\varepsilon}{2\sigma}$ we get the result. QED (Appl.: 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!

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

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

– $\mathcal{T}^{t+s}(\mu) = \mathcal{T}^s \circ \mathcal{T}^t(\mu)$ – 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: $F : (X, \mathcal{B}) \hookleftarrow \implies \mathcal{T}^n \mu(A) := \mu(F^{-n}A) \ \forall A$.

Random Examples:

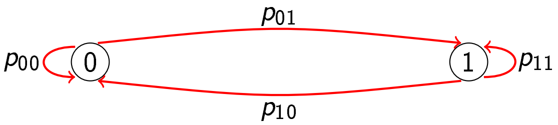
(a) *Random map*: $F_1, F_2 (X, \mathcal{B}) \hookleftarrow$, $0 < p < 1$,

$$\mathcal{T}^1 \mu(A) := p\mu(F_1^{-1}A) + (1-p)\mu(F_2^{-1}A).$$

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

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

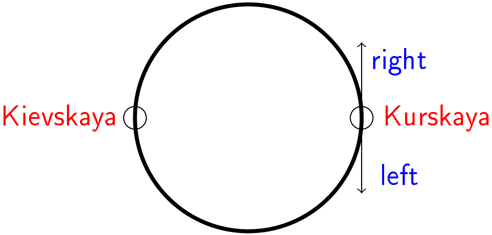
$\mathcal{T}\mu = \mu^* P.$

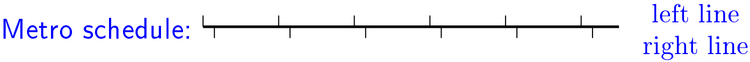


(c) iid $\xi_i \in \{0, 1\}$ Markov chain with $p_{ij} = 1/2$.
(d) General continuous time Markov chains: transition probabilities $P_s^t(x, A) := \mathbf{P}(\xi_{s+t} \in A | \xi_s = x)$.
How this corresponds to the traditional approach?

[23] Lecture 3. Asymmetry of subway rides.

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?





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* $F_{t_1,\dots,t_n}(x_1,\dots,x_n) = \mathbf{P}(\omega:\ \xi_{t_1} \le x_1,\dots,\xi_{t_n} \le x_n)$.

T(Kolmogorov) Let $F_{t_1,\dots,t_n}(x_1,\dots,x_n)$ be a given family of finite dimensional distributions, satisfying the following conditions of consistency: $F_{t_1,\dots,t_k,\dots,t_n}(x_1,\dots,\infty,\dots,x_n)$
$= F_{t_1,\dots,t_{k-1},t_{k+1},\dots,t_n}(x_1,\dots,x_{k-1},x_{k+1},\dots,x_n)$. 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:
– For fixed s, t, x the function $Q_s^t(x, \cdot)$ 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 the action on measures:

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

and the action on functions:

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

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

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

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. $\xi_t : (\Omega, \mathcal{F}, \mathbf{P}) \to (X, 2^X)$, $t \in \mathbb{Z}_+$, $X := \{0, 1\}$,

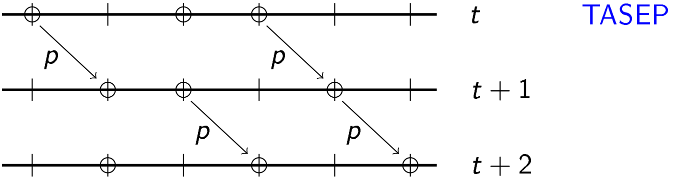
$$\mathbf{P}(\xi_t = 1) = p, \mathbf{P}(\xi_t = 0) = 1 - p.$$

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



The main problem in the analysis is an infinite number of simultaneous interactions between neighboring particles.

(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.

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.

L. Let $\Omega_n^k := \{\omega : \text{ a return to k occurs after 2n time steps}\} \Longrightarrow$ $\mathbf{P}(\cup_{n \geq 0} \Omega_n^k) = 1$ iff $\sum_{n \geq 0} \mathbf{P}(\Omega_n^k) = \infty$.

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 $p = q = 1/2$. see next slide

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

T. $v := \max_n \mathbf{P}(\xi_n = k | \xi_0 = k) = 1$ iff $\sum_{n\geq 1} p_{k,k}^{(n)} = \infty \quad \forall k \in \mathbb{Z}_+$.
Proof. Let $v_n := \mathbf{P}(\text{the 1st return to k occurs after n steps})$, 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, $\lim_{z\to 1} U(z) = \lim_{z\to 1} \sum_{m\geq 0} u_m z^m = \sum_{m\geq 0} u_m = \infty$. QED

(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 show that
$Z := \mathbf{P}(\{A\cos(\eta(t_1 + h) + \varphi), \ldots, A\cos(\eta(t_n + h) + \varphi)\} \in C)$
$= \mathbf{P}(\{A\cos(\eta t_1 + \varphi), \ldots, A\cos(\eta t_n + \varphi)\} \in C)$.
$B := \{(x, y, z) : x, y \geq 0, z \in [0, 2\pi),$
$\{x\cos(yt_1 + z), \ldots, x\cos(yt_n + z)\} \in C\}$ is a Borel set.
Denoting by $\{z\}_{2\pi}$ the fractional part of z mod 2π we get
$Z = \mathbf{P}((A, \eta, \{\varphi + yh\}_{2\pi}) \in B) = \mathbf{P}((A, \eta, \varphi) \in B)$.
(A, η) and φ are independent $\Longrightarrow \mu_{A,\eta,\varphi} = \mu_{A,\eta} \times \mu_\varphi$. Thus
$Z = \int_0^\infty \int_0^\infty \mu_{A,\eta}(dxdy)\ \mu_\varphi(C_1 := \{z :\ (x, y, \{z + yh\}_{2\pi}) \in B\})$
$= \int_0^\infty \int_0^\infty \mu_{A,\eta}(dxdy)\ \mu_\varphi(C_2 := \{z :\ (x, y, z + y) \in B\})$, since
C_1 is obtained from C_2 by the translation by yh and mod 2π. Now
μ_φ is uniform on $[0, 2\pi)$ and does not change under translations.
QED

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

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

B is the intersection of the events:

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

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

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

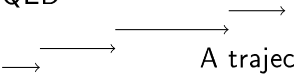
$$\mathbf{P}(C_N) \geq \prod_{k=0}^{2^n N-1} \mathbf{P}(\{\xi_{(k+1)\cdot k^{-n}} - \xi_{k\cdot k^{-n}} \in \{0,1\}\})$$

$$\geq (e^{-a2^{-n}} + a2^{-n}e^{-a2^{-n}})^{2^n N} \geq (1 - o(a2^{-n}))^{2^n N} \stackrel{n\to\infty}{\longrightarrow} 1,$$

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 $\mathbf{P}(Z) = 1$.

QED



A trajectory of ξ_t

(III) Wiener process w_t starting from 0 on $T := \mathbb{R}_+$:
(0) $w_0 = 0$.
(i) $\forall 0 \leq t_0 < t_1 < \cdots < t_n$ r.v. $\Delta w_{t_i, t_{i-1}} := w_{t_i} - w_{t_{i-1}}$ are independent.
(ii) r.v. $\Delta w_{t,s} := w_t - w_s,\ 0 \leq s \leq t$ are Gaussian $\mathcal{N}(0, t-s)$.
(iii) Trajectories of w_t are continuous.

T1. $\forall 0 \leq a \leq t_0 < t_1 < \cdots < t_n = b$

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

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

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

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

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

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

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. QED

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. For $t > s$ we have $E\frac{w_t - w_s}{|t-s|^\beta} = 0, \quad D\frac{w_t - w_s}{|t-s|^\beta} = \frac{t-s}{|t-s|^{2\beta}} = (t-s)^{1-2\beta} \stackrel{t-s\to 0}{\longrightarrow} 0$ 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 $EV(w_t, \Delta) = \sqrt{\frac{2}{\pi}(b-a)} \cdot \sqrt{n} \stackrel{n\to\infty}{\longrightarrow} \infty, \ DV(w_t, \Delta) \stackrel{n\to\infty}{\longrightarrow} b - a.$

Now again the Chebyshev inequality gives the result.

Statistics:

Claim. 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 $\mathbf{P}(|f_n(t) - t| \overset{n\to\infty}{\longrightarrow} 0) = 1$ uniformly on $[0, T]$.

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

For $n \geq m$ we have

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

Since the mathematical expectation is finite, the series converges by the Chebyshev inequality with probability 1. Hence, $f_n(t) - t \to 0$. QED

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. QED

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

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

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

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

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

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

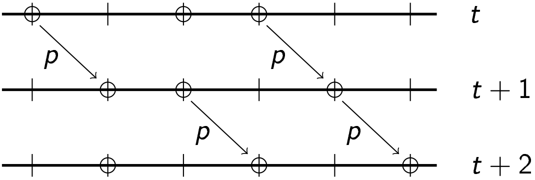
(III) Existence. Let ξ_n be a simple symmetric random walk on $\mathbb{Z}$ with $\xi_0 = 0$ and $\mathbf{P}(\xi_{n+1} = i+1 | \xi_n = i) = 1/2$. We interpolate it and rescale to $[0,1]$, namely $\forall n \in \mathbb{Z}_+$, $(i_0, i_1, \ldots, i_n) \in \mathbb{Z}^{n+1}$ define $h_{(i_0,i_1,\ldots,i_n)}(t) := \frac{1}{\sqrt{n}}(1 - nt + [nt])\xi_{[nt]} + (nt - [nt])\xi_{[nt]+1}$, which linearly interpolates the points $\{1/n, i_k/\sqrt{n}\}_k$. $\mu_n(\{f\}) := 2^{-n} 1_{h_{(0,i_1,\ldots,i_n)}}(f)$ for some $(i_1, \ldots, i_n) \in \mathbb{Z}^n$ is a measure on $\{f\} \in C([0,1])$. $\{\mu_n\}_n$ describes the random motion of a particle which performs many tiny steps during each time moment.

T. $\mu_n \stackrel{n\to\infty}{\longrightarrow} W$ weakly – probability measure on $C([0,1])$, called the Wiener measure. $W(\{f : \ f(t_{i+1}) - f(t_i) \in I_i, \ i = 1, 2, \ldots, n\})$

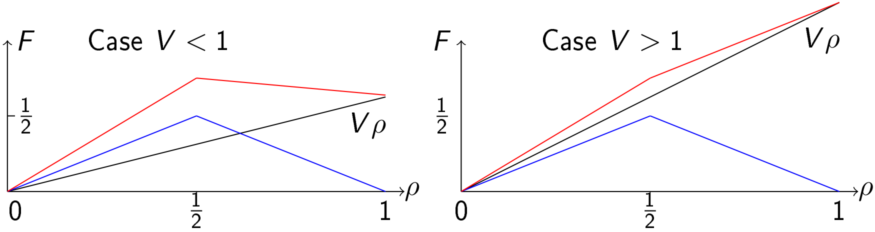
$$= \prod_{i=1}^n \frac{1}{\sqrt{2\pi(t_{i+1} - t_i)}} \int_{I_i} e^{\frac{x^2}{2(t_{i+1}-t_i)}} dx.$$

Here I_i are intervals from $[0,1]$. W corresponds to a random process called Brownian motion and satisfying all properties of the Wiener process.

TASEP, $p = 1$



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 $\frac{1+V}{2} > V \implies V < 1$.

Moments $(d=1)$: $E\xi_t = E(\xi_t - \xi_0) = 0$, $D\xi_t = D(\xi_t - \xi_0) = t$.
For $0 \le s \le t$ we have
$\operatorname{cov}(\xi_t, \xi_s) = E(w_t - w_s)w_s + Ew_s^2 = D(w_s - w_0) = s = t \wedge s$.

Continuity: $E(\xi_{t+h} - \xi_t) = 0, D(\xi_{t+h} - \xi_t) = h$. Therefore $\xi_t \to \xi_{t_0}$ as $t \to t_0 \quad \forall t_0$ 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 in a Brownian bridge are not independent.
Representations: $B_t = w_t - \frac{t}{T} w_T = \frac{T-t}{\sqrt{T}} w_{\frac{t}{T-t}}$.

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. $\forall 0 \le t_1 \le \cdots \le t_n$ 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,
$E(\xi_t - \xi_s)^2 = E\xi_t^2 + E\xi_s^2 - E\xi_t\xi_s = t + s - 2(t \wedge s) = t - s$. QED
Remark. $E|\xi_t - \xi_s| = \sqrt{\frac{2(t-s)}{\pi}}$.

Let τ_a be the 1st moment of time when $\xi_t = a > 0$.
Claim. $\mathbf{P}(\xi_t \geq a | \tau_a \leq t) = \frac{1}{2}$.
Proof. $\xi_t = a$ at $t = \tau_a$, while for $t > \tau_a$ by the symmetry. QED
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, we can find the maximal value of ξ_s during the time t:

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

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

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

[48] Lecture 6. Wiener processes 12

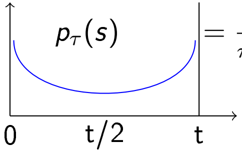
The density of the r.v. (τ, ξ) at a point $(\tau = s, \xi = a)$ coincides with the density of (τ_a, ξ) at the same point, since

$$p_{\tau,\xi}(s, a) = p_\tau(s|\xi = a)p_\xi(a) = p_{\tau_a}(s|\xi = a)p_\xi(a) = p_{\tau_a,\xi}(s, a).$$

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

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

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



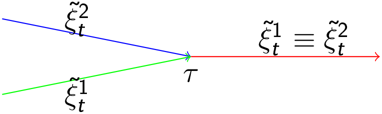
$$= \frac{1}{\pi\sqrt{s(t-s)}}$$

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

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
$\tilde{\mathbf{P}}(A^1 \times \Omega_2) = \mathbf{P}^1(A^1)$, $\tilde{\mathbf{P}}(\Omega^1 \times A^2) = \mathbf{P}^2(A^2)$ $\forall A^i \in \mathcal{F}^i$.
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.
Let $(\Omega^i, \mathcal{F}^i) = (\Omega, \mathcal{F})$, then the total variation distance
$||\mathbf{P}^1 - \mathbf{P}^2||_{tv} := \sup_{A \in \mathcal{F}} |\mathbf{P}^1(A) - \mathbf{P}^2(A)|$.
T1(Coupling inequality). Given r.v. ξ^i, $i = 1, 2$ with probability distributions $\mathbf{P}^i$ for any coupling $||\mathbf{P}^1 - \mathbf{P}^2||_{tv} \leq \tilde{\mathbf{P}}(\tilde{\xi}^1 \neq \tilde{\xi}^2)$.
Proof. $\mathbf{P}^1(\xi^1 \in A) - \mathbf{P}^2(\xi^2 \in A) = \tilde{\mathbf{P}}(\tilde{\xi}^1 \in A) - \tilde{\mathbf{P}}(\tilde{\xi}^2 \in A)$
$= \tilde{\mathbf{P}}(\tilde{\xi}^1 \in A, \tilde{\xi}^1 = \tilde{\xi}^2) + \tilde{\mathbf{P}}(\tilde{\xi}^1 \in A, \tilde{\xi}^1 \neq \tilde{\xi}^2)$
$-\tilde{\mathbf{P}}(\tilde{\xi}^2 \in A, \tilde{\xi}^1 = \tilde{\xi}^2) - \tilde{\mathbf{P}}(\tilde{\xi}^2 \in A, \tilde{\xi}^1 \neq \tilde{\xi}^2) \leq \tilde{\mathbf{P}}(\tilde{\xi}^1 \neq \tilde{\xi}^2)$. QED

see next slide

A *coupling* of r. processes ξ_t^i, $i = 1, 2$ on the same space $(\Omega, \mathcal{F}, \mathbf{P})$ is a new r. process $\tilde{\xi}_t := (\tilde{\xi}_t^1, \tilde{\xi}_t^2)$ on $(\Omega^1 \times \Omega^2, \mathcal{F}^1 \times \mathcal{F}^2, \tilde{\mathbf{P}})$.
$\tau := \inf\{t \in T : \ \xi_t^1 = \xi_t^2\}$ – the coupling time.
A coupling $\tilde{\mathbf{P}}$ is called successful if $\tilde{\mathbf{P}}(\tilde{\xi}_t^1 \neq \tilde{\xi}_t^2) = 0 \ \ \forall t \geq \tau$.



T2. $||\mathbf{P}^1(\xi_t^1 \in \cdot) - \mathbf{P}^2(\xi_t^2 \in \cdot)||_{tv} \leq \tilde{\mathbf{P}}(\tau > t) \ \ \forall t \in T$.
Proof. $\{\xi_t^1 \neq \xi_t^2\} \subseteq \{\tau \leq t\}$ by T1. QED

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.

(1) $\xi_t \in \mathbb{R}^d$ is called a *Gaussian* random function if all its finite dimensional distributions are Gaussian.
When the random sin oscillations: $\xi_t := A\cos(\eta t + \varphi)$ are Gaussian? ($\exists y_0 : \ \mathbf{P}(\eta = y_0) = 1$) and A has the density ae^{-ax}.)
(2) $\xi_t \in \mathbb{R}^d$ is called a process with *independent increments* if all its increments over non-intersecting time intervals are independent.
(2') 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 $\mathrm{cov}(\xi, \eta) := E(\xi - E\xi)(\eta - E\eta)$.
(3) $\xi_t \in \mathbb{R}^d$ is called *stationary* if all its finite dimensional distributions are shift-invariant: $\mu_{\bar{t}+h} = \mu_{\bar{t}}$.
(3') $\xi_t \in \mathbb{R}^1$ is called *stationary in 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 $E\xi_t = m, \quad K(t+h, s+h) = K(t-s)$.

(4) $\xi_t \in \mathbb{R}^1$ 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. Give an example:
$\xi_t := A\cos(\eta t + \varphi) + \alpha t + \beta$, where φ does not depend on (A, η, α, β) and is uniformly distributed on $[0, 2\pi)$.

(5) *Markov chains* – the future and the past are independent if the present state is fixed.
Discuss connections with the transition function.

T1. Let $E|\xi_t|^2 < \infty \ \forall t$. Then $\exists (L^2) \lim_{t\to t_0} \xi_t$ iff $\exists \lim_{t,s\to t_0} E\xi_t\xi_s$.
Proof. The necessity follows from the continuity of the scalar product, while the sufficient part follows from the Cauchy condition

$$\lim_{t,s\to t_0} E|\xi_t - \xi_s|^2 = \lim_{t,s\to t_0} [E|\xi_t|^2 - E\xi_t\xi_s - E\xi_s\xi_t + E|\xi_s|^2] = 0.$$

Problem 1. 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 $\eta_n := \sum_{i=1}^n \xi_i \Longrightarrow K_{\eta\eta}(n,m) = \sum\limits_{i\leq \min(n,m)} D\xi_i$. Now use T1 above.

T2. $(P)\lim_{t\to t_0}\xi_t$ exists iff exists the weak $\lim_{t,s\to t_0}\mu_{\xi_t,\xi_s} =: \mu$.

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
$\mathbf{P}(|\xi_t - \xi_s| \geq \varepsilon) \leq Ef_\varepsilon(\xi_t - \xi_s) = \int\int f_\varepsilon(x-y)\mu_{\xi_t,\xi_s}(dxdy)$
$\overset{t,s\to t_0}{\longrightarrow} \int\int f_\varepsilon(x-y)\mu(dxdy) = 0$ since $f_\varepsilon \in C^0$ and μ is supported by the diagonal $\{x=y\}$. Hence the sequence in fundamental in probability. QED

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.

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. Even that realizations of the Poisson process are discontinuous. Why?
Answer: $\mathbf{P}$(a discontinuity happens at a given point)=0.

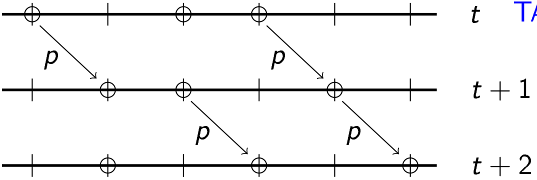
Problem. Prove/disprove that if ξ_t are independent $\forall t$ and has the same density $p(x)$, then ξ_t is stochastically discontinuous $\forall t$.
Proof. $\mathbf{P}(|\xi_t - \xi_{t_0}| \geq \varepsilon) = \int \int_{|x-y| \geq \varepsilon} p(x)p(y)dxdy$
$\xrightarrow{\varepsilon \to 0} \int \int_{x \neq y} p(x)p(y)dxdy = \int \int p(x)p(y)dxdy = 1.$
Hence $\exists \varepsilon_0 > 0$ such that $\mathbf{P}(|\xi_t - \xi_{t_0}| \geq \varepsilon) > 1/2 \Longrightarrow$ there is no convergence in probability. QED

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

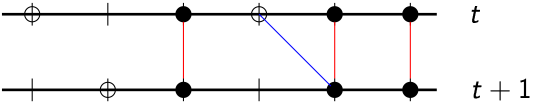
Problem. Prove/disprove that ξ_t is (a) stochastically continuous on T iff μ_{ξ_t,ξ_s} is weakly continuous on $(t,s)\in T\times T$; and is (b) stochastically continuous in L^2 iff $E\xi_t\bar{\xi}_s$ is continuous.
(a) Follows from T1; (b) from T2 (about continuity).

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.

Lecture 7. Coupling for TASEP



TASEP, $0 < p \leq 1$



Pairing

(1) Let $w_t,\ t \geq 0$ be a standard Wiener process. Find ALL increasing functions $\varphi(t)$ and constants $a, b > 0$ such that $\xi_t := \varphi^{-1}(at)w_{\varphi(bt)}$ is a stationary process and calculate its correlation function.
(2) 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 $(P)\lim_{n\to\infty}\eta_n$.
(3) Find ALL stationary processes $\xi_t,\ t \geq 0$ such that $\exists (P)\lim_{t\to\infty}\xi_t$.
(4) Let $w_t,\ t \geq 0$ be a standard Wiener process. Find a joint distribution of w_t and $\int_0^t w_s ds$.
(5) Prove/disprove existence of a Gaussian process $\xi_t,\ 0 \leq t \leq 1$ 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.
(6) Let $\xi_t,\ 0 \leq t \leq 1$ be a Cauchy process and let $\eta_t := \xi_t + Ct$. Find ALL constants $C \in \mathbb{R}$ such that distributions of these processes are absolutely continuous wrt each other.
(7) Let w_t be a standard Wiener process. Find a distribution of a r.v. $\eta := \max_{0\leq t\leq T}(w_t + Ct)$ as a function of $C \in \mathbb{R}$.

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.
T1. Let $E|\xi_t|^2 < \infty \ \forall t$. Then $\exists (L^2) \lim_{t \to t_0} \xi_t$ iff $\exists \lim_{t,s \to t_0} E\xi_t \xi_s$.
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.

Wiener process has no derivative even in probability. Why?
$w_t - w_s$ is Gaussian $\mathcal{N}(0, \frac{1}{|t-s|})$, i.e. $(w_t)' := \frac{d}{dt} w_t$ diverges.

Poisson process has a derivative in probability, but not in $L^p, p \geq 1$.
Why?
$(P) \lim_{s \to t} (\xi_t - \xi_s)/(t-s) = 0$. Hence the L^p limit if it exists should be equal to 0 almost surely. However, $E|\xi_t - \xi_s|/|t-s| \geq \frac{\mathbf{P}(\xi_t \neq \xi_s)}{|t-s|^p} \sim a|t-s|^{1-p}$, 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. $\forall \varepsilon > 0, s \in [a,b] \ \exists O_s : \ |\xi_t - \xi_s| \le \varepsilon |t-s| \ \forall t \in O_s$.

Assume that $s := \liminf\{t \in [a,b] : \ \ \xi_t \ne \xi_a\} > a$.

$\Longrightarrow \exists \varepsilon > 0 : \ |\xi_t - \xi_a| > \varepsilon|t-a| \ \ \forall t > s$ and (by continuity of ξ_t)

$|\xi_s - \xi_a| = \varepsilon|s-a| \Longrightarrow$ for $O_s \ni t > s$ we have

$|\xi_t - \xi_a| \le |\xi_t - \xi_s| + |\xi_s - \xi_a| \le \varepsilon|t-s| + \varepsilon|s-a| = \varepsilon|t-a|,$

which contradicts to the definition of s. QED

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 $(a,b)^2$. (Follows from standard analysis).

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

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

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. (Standard analysis + uniform continuity.)

Claim 2. Let τ be a r.v. uniformly distributed on $T := [0,1]$. Then the process $\xi_t := (1-\tau)^{-1} 1_{t>\tau}$ is stochastically continuous on T, but $\int_0^1 \xi_t dt$ does not exist in L^2-sense.

One can differentiate the integral over lower and upper limits and we have the Newton-Leibniz formula.
Application: if ξ_t – Poisson process, then $\frac{d}{dt}\left((L^p)\int_a^b \xi_t dt\right) = \xi_t$.

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.

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

(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, $J(\omega) := \int_a^b \xi_t(\omega)dt$ is the same limit $\forall\omega$. Now a.e. convergence implies the convergence in probability. Moreover, by from the Riemann-integrable assumption $J(\omega)$ is measurable. QED

(b) If the realization 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 $\left((L^p)\int_a^b \xi_t dt\right)(\omega) \sim \int_a^b \xi_t(\omega)dt \quad (*)$.
Proof. It is enough to consider the case $p=1$ (other convergences will follow). By Fubini's theorem the rhs of (*) exists and $\int_a^b \int_\Omega |\xi_t(\omega)| dt\mathbf{P}(d\omega) \leq (b-a)\max_{t\in[a,b]} E|\xi_t| < \infty$.
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 $\exists \delta > 0 \ \ \forall |t-s| < \delta \Longrightarrow E|\xi_t - \xi_s| < \varepsilon$.
Therefore if $\mathrm{diam}\{t_i\} < \delta$ the math. expectation between the integral sums corresponding to lhs and rhs of (*) is $\leq \varepsilon(b-a)$.
Since $\varepsilon > 0$ is arbitrary the integrals in both senses coincide a.e.
QED

Computing of moments of integrals is rather simple.
Claim. Let ξ_t be continuous in L^2-sense and $E|\xi_t|^2 < \infty$. Then
$E\int_a^b \xi_t dt = \int_a^b E\xi_t dt, \quad \mathrm{cov}(\int_a^b \xi_t dt, \xi_s) = \int_a^b K_{\xi\xi}(t,s)dt,$
$\mathrm{cov}(\int_a^b \xi_t dt, \int_c^d \xi_s ds) = \int_a^b \int_c^d K_{\xi\xi}(t,s)dtds.$
Follows from the continuity of the scalar product on its arguments:
Indeed, $E\int_a^b \xi_t dt = E(L^2)\lim\sum_i (t_{i+1} - t_i)\xi_{s_i}$
$= (E(L^2)\lim\sum_i (t_{i+1} - t_i)\xi_{s_i},\ 1) = (L^2)\lim(E\sum_i (t_{i+1} - t_i)\xi_{s_i},\ 1)$
$= (L^2)\lim\sum_i (t_{i+1} - t_i)E\xi_{s_i} = \int_a^b E\xi_t dt.$
This proves the 1-st equality. Two others for homework.

(1) Let ξ_t be a L^2-continuous stationary process with $E\xi_t \neq 0$. 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 $\equiv 0$.

(2) Let ξ_t be a L^2-continuous stationary process and let its correlation function $K(t) \to 0$ as $t \to \infty$. Calculate
$(L^2)\lim_{s\to\infty} \frac{1}{s} \int_t^{t+s} \xi_u du$.
Answer: $E\xi_t =: m$. Indeed,
$E|\frac{1}{s} \int_t^{t+s} \xi_u du - m| = \frac{2}{s} \int_0^s (s-u)K(u)du \to 0$ iff
$\lim_{s\to 0} \int_0^s K(u)du = 0$.

(3) Let a sequence of independent r.v. $\{\xi_n\}$ converge with probability 1. Prove/disprove existence of a number C such that $\mathbf{P}(\lim_{n\to\infty} \xi_n = C) = 1$.
Answer: Consider $\varphi(x) := \mathbf{P}(\lim_{n\to\infty} \xi_n \leq x)$. Then $\varphi(x) \in \{0, 1\}$ and C is the point, where φ makes a jump.