B. Oksendal, Stochastic Dif. Eqs.

L.C. Evans, An Introduction to Stoch. Dif. Eqs.

---

# Measure & probability theory

① $\Omega$ : set of ~~all~~ possible elementary outcomes or sample space, $\Omega \neq \emptyset$.

result of experiment or trial.

Ex 1 dice $\Omega = \{1, 2, \dots 6\}$

Ex 2 real line $\Omega = \mathbb{R}$

② $\mathcal{F}$: event space.

Each event $A \in \mathcal{F}$ is a subset $A \subset \Omega$.

$\mathcal{F}$ must be a $\sigma$-algebra

Def A $\sigma$-algebra $\mathcal{F}$ on $\Omega$ is a family of subsets $A \subset \Omega$ with the properties:

(i) $\emptyset \in \mathcal{F}$

(ii) $A \in \mathcal{F} \Rightarrow A^c := \Omega \setminus A \in \mathcal{F}$ (complement)

(iii) $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow A = \bigcup_{i=1} A_i \in \mathcal{F}$.

(countable unions)

__Obs__ $\Omega = \Omega \setminus \phi \in \mathcal{F}$

$\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$ (ex.)

__Ex 1__ $\Omega = \{1, \ldots, 6\}$

$\mathcal{F} = \{A \subset \Omega\}$ all subsets

This is the case for any finite set $\Omega$.

__Ex 2__ $\Omega = \mathbb{R}$

$\mathcal{F}$ can be a Borel $\sigma$-algebra on $\mathbb{R}$.

__Def__ If $\Omega$ is a topological space, then the Borel $\sigma$-algebra on $\Omega$ is a ~~[illegible]~~ smallest $\sigma$-algebra containing all open sets of $\Omega$.

(each $\omega \in \Omega$ is formed by operations of countable union, contable intersections, complements, etc.)

~~[illegible] $\mathbb{R}$~~

~~Construction~~ measure theory.

Def of the other side

Obs Given any family $\mathcal{U}$ of subsets of $\Omega$, there is a smallest $\sigma$-algebra $\mathcal{H}_{\mathcal{U}}$ containing $\mathcal{U}$:

$$\mathcal{H}_{\mathcal{U}} = \bigcap \{ \mathcal{H};\ \mathcal{H} \text{ is a } \sigma\text{-alg. of } \Omega,\ \mathcal{U} \in \mathcal{H} \}.$$

$\mathcal{H}_{\mathcal{U}}$ is called the $\sigma$-algebra generated by $\mathcal{U}$.

③ $P : \mathcal{F} \mapsto [0,1]$ : probability measure such that (Def)

(a) $P(\Omega) = 1$

(b) $P$ is countably additive:

if $A_1, A_2, \ldots \in \mathcal{F}$ are pairwise disjoint ($A_i \cap A_j = \emptyset$), then $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$.

Obs $P(\emptyset) = 0$. $A \subseteq B \Rightarrow P(A) \leq P(B)$. (e

~~Def Probability space is a triplet $(\Omega, \mathcal{F}, P)$.~~

Ex 1 : $\Omega = \{1, \ldots, 6\}$

$$P(A) = \# A / \# \Omega$$

cardinality

$P(\{1\}) = 1/6$, $P(\{2,4,6\}) = 1/2$

Ex 2 : $\Omega = \mathbb{R}$

$$P(A) = \frac{1}{2\pi} \int_A e^{-x^2/2} \, dx.$$

(normal distribution)

**Def** Probability space is a triple $(\Omega, \mathcal{F}, P)$.

**Def** $Y : \Omega \to \mathbb{R}^n$ is a measurable (or $\mathcal{F}$-measurable) function if $Y^{-1}(\mathcal{U}) := \{\omega \in \Omega : Y(\omega) \in \mathcal{U}\} \in \mathcal{F}$ for all open sets $\mathcal{U} \in \mathbb{R}^n$.

**Obs** This is automatically true for all Borel sets $\mathcal{U}$.

**Def** For any function $X : \Omega \to \mathbb{R}^n$, the $\sigma$-alg. $\mathcal{H}_X$ generated by $X$ is the smallest $\sigma$-alg. on $\Omega$ containing the sets $X^{-1}(\mathcal{U})$, $\mathcal{U} \in \mathbb{R}^n$ open.

**Prop** $\mathcal{H}_X = \{X^{-1}(B) : B \in \mathcal{B}\}$ (ex.)

$\mathcal{B}$ ← Borel $\sigma$-alg. on $\mathbb{R}^n$.

**Lemma (Doob-Dynkin)**

Let $X, Y : \Omega \to \mathbb{R}^n$. Then $Y$ is $\mathcal{H}_X$-measurable iff (if and only if) $\exists$ Borel-measurable function $g : \mathbb{R}^n \to \mathbb{R}^n$ such that $Y = g(X)$.

A <u>random variable</u> is an $\mathcal{F}$-measurable function

$$X : \Omega \to \mathbb{R}^n$$

<u>Distribution of X</u> is a probability measure

$$\mu_X = X_{\#} P \quad , \quad \mu_X(B) = P(X^{-1}(B)) , \quad B \in \mathcal{B}.$$

(Borel in $\mathbb{R}^n$)

<u>Expectation</u>

~~Recalling [illegible]~~
If $X$ is integrable $\left( \int_\Omega |X(\omega)| \, dP(\omega) < \infty \right)$,

then $$E[X] = \int_\Omega X(\omega) \, dP(\omega) = \int_{\mathbb{R}^n} x \, d\mu_X(x).$$

If $f : \mathbb{R}^n \to \mathbb{R}^n$ is Borel measurable and

$\int_\Omega |f(X(\omega))| \, dP(\omega) < \infty$, Then

$$E[f(X)] = \int_\Omega f(X(\omega)) \, dP(\omega) = \int_{\mathbb{R}^n} f(x) \, d\mu_X(x).$$

<u>Examples</u>: $X = \chi_A(\omega) := \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$ (indicator f.)

$\cancel{E[X] = P(A)}$

$E[X] = P(A)$.

or $X = \sum_{i=1}^m a_i \chi_{A_i}$ for $A, A_1, A_2, \dots \in \mathcal{F}$

(simple functions). $E(X) = \sum_{i=1}^m a_i P(A_i)$ if $A_1, A_2, \dots$ are disjoint.

Variance $\quad V[X] = \int_\Omega |X - E[X]|^2 \, dP$

Obs $\quad V(X) = E[|X|^2] - |E(X)|^2.$ $\quad$ (ex.)

Distribution function

If $X : \Omega \to \mathbb{R} \implies F_X(x) = P(X \le x), \quad x \in \mathbb{R}.$

If $X : \Omega \to \mathbb{R}^n \implies$ —— " ——, where $X \le x$ means $X_i \le x_i$ for all $i = 1, .., n.$

Density function

If $\mu_X$ is absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}^n$, then $\exists$ a ~~positive~~ (non-negative) measurable function $p_X : \mathbb{R}^n \to \mathbb{R}$ such that

$$P(X \in B) = \int_B p_X(x) \, dx \quad \text{for all } B \in \mathcal{B}.$$

Example (Normal distribution) or (Gaussian distr.)

$$p_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left( - \frac{|x-m|^2}{2\sigma^2} \right), \quad x \in \mathbb{R}$$

$E[X] = m, \quad V[X] = \sigma^2.$

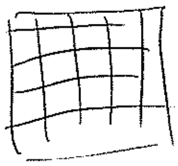
Obs Notation : $X \sim N(m, \sigma^2).$

Ex Normal (or Gaussian) distribution in $\mathbb{R}^n$ is given by

$$p_X(x) = \frac{1}{\sqrt{(2\pi)^n \det C}} \exp\left[-\frac{1}{2}(x-m)^T C^{-1}(x-m)\right], \quad x \in \mathbb{R}^n.$$

Here $m \in \mathbb{R}^n$ and $C$ is a (positive definite, symmetric) covariance matrix.

$$E[X] = m, \qquad E[(X_j - m_j)(X_k - m_k)] = c_{jk}.$$

$m \longrightarrow$  $\qquad C \longrightarrow$ 

Lemma Let $X$ be rand. variable with prob. density $p_X$. If $g : \mathbb{R}^n \to \mathbb{R}$ is integrable, then

$$E[g(X)] = \int_{\mathbb{R}^n} g(x)\, p_X(x)\, dx.$$

In particular

$$E[X] = \int_{\mathbb{R}^n} x\, p_X(x)\, dx, \quad V[X] = \int_{\mathbb{R}^n} |x - E[X]|^2 p_X(x)\, dx.$$

## Summary

**Probability space:** $(\Omega, \mathcal{F}, P)$

sample space (set) → $\Omega$

$\mathcal{F}$ ← $\sigma$-algebra

$P$ ← probability measure $P(\Omega) = 1$.

**Random variable** : $X : \Omega \to \mathbb{R}^n$, $\mathcal{F}$-measurable

**Distribution** : $\mu_X = X_{\#} P$, prob. measure on $\mathbb{R}^n$.

**Density function** : $p_X : \mathbb{R}^n \to \mathbb{R}$ non-negative measurable

$$\text{t.q.} \quad P(X \in B) = \int_B p_X(x)\, dx, \quad B \in \mathcal{B}.$$

**Lemma** If $g : \mathbb{R}^n \to \mathbb{R}$ is integrable & ~~that~~

and $X$ is a random variable with pr. density $p_X$,

then
$$E[g(X)] = \int_{\mathbb{R}^n} g(x)\, p_X(x)\, dx.$$

In particular,

$$E[X] = \int_{\mathbb{R}^n} x\, p_X(x)\, dx$$

$$V[X] = \int_{\mathbb{R}^n} |x - E[X]|^2\, p_X(x)\, dx.$$

# Independence

**Def** Two events $A, B \in \mathcal{F}$ are independent if

$$P(A \cap B) = P(A) \cdot P(B).$$

*Intuition:* $P(A)$, $P(B)$ are fractions of "samples" belonging to $\omega \in A$ and $B$.

**Def** A collection $\{\mathcal{H}_i, i \in I\}$ of families $\mathcal{H}_i$ of measurable sets is independent if

$$P(H_{i_1} \cap \dots \cap H_{i_k}) = P(H_{i_1}) \cdots P(H_{i_k})$$

for any $i_1, \dots, i_k$ and $H_{i_1} \in \mathcal{H}_{i_1}, \dots, H_{i_k} \in \mathcal{H}_{i_k}$.

**Def** A collection of random variables $\{X_i, i \in I\}$ is independent if the collection of ~~respective~~ generated $\sigma$-algebras $\{\mathcal{H}_{X_i}, i \in I\}$ is independent.

~~Ex If $X_1, \dots, X_n$ are real-valued independent and $E[|X_i|] < \infty$, then $E[X_1 \cdots X_n] = E[X_1] \cdots E[X_n]$.~~

**Obs.** $\{X_i, i \in I\}$ are idep. iff. $P(X_{i_1} \in B_1, \dots, X_{i_k} \in B_k) = \prod_{j=1}^{k} P(X_{i_j} \in B_j)$ (ex)
for all $X_{i_1} \dots X_{i_k}$ and $B_1, \dots B_k \in \mathcal{B}$.

Example Take $\Omega = [0,1)$, $\mathcal{F}$ = Borel on $\Omega$,
$\mathbb{P}$ Lebesgue measure.

Define $X_n(\omega) = \begin{cases} 1, & [2^n \omega] \text{ even} \\ -1, & [2^n \omega] \text{ odd}. \end{cases}$ (integer part)

$n = 1, 2, \dots$

Obs: related to binary representation
$(0,1) \to (1,-1)$

$\omega = .01101\dots$
$\downarrow \downarrow (n)$
$\dots 1 \; -1 \dots$

---

$X_n$ are called Rademacher functions.

They $\{X_n, n = 1, 2, \dots\}$ are independent.

$$P(X_1 = e_1, X_2 = e_2, \dots, X_k = e_k) = \prod_{i=1}^{k} P(X_i = e_i) \left[= 2^{-k}\right]$$

for any $e_1, \dots, e_k \in \{-1, 1\}$. (ex).

Th If $X_1 \dots X_{m+n}$ are independent, then
$Y = f(X_1, \dots X_m)$ and $Z = g(X_{n+1}, \dots, X_{n+m})$
are independent for any measurable functions $f, g$.

~~Th $X_1, \dots, X_m$ are independent iff~~

~~**Th** Let $X_1, \ldots, X_m$ be rand. variables with prob. densities $p_{X_1}, \ldots, p_{X_m}$. Then, $X_1, \ldots X_m$ are independent iff~~

~~$$\underbrace{p_{X_1, \ldots X_m}}_{\substack{\text{prob. density of} \\ X = (X_1, \ldots, X_m)}} (x_1, \ldots, x_m) = p_{X_1}(x_1) \cdots p_{X_m}(x_m).$$~~

~~(ex).~~

**Th** If $X_1, \ldots, X_m$ are independent, real-valued with $E[|X_i|] < \infty$ for all $i = 1, \ldots, m$, then

$$E[X_1 \cdots X_m] = E[X_1] \cdots E[X_m].$$

**Th** If $X_1, \ldots X_m$ are indep., real-valued with $V(X_i) < \infty$ for $i = 1, \ldots, m$, then

$$V(X_1 + \cdots + X_m) = V(X_1) + \cdots + V(X_m). \qquad \text{(ex.)}$$

Conditional probability (Oks → App. B, Evans).

$A, B \in \mathcal{F}$, $P(B) > 0$.

What means $P(A|B)$ = the probability of $A$ given $B$.

Intuition : We know that $\omega \in B$.

Take $\tilde{\Omega} = B$, $\tilde{\mathcal{F}} = \{ C \cap B : C \in \mathcal{F} \}$,

$$\tilde{P} = \frac{P}{P(B)} . \quad \Rightarrow \quad \tilde{P}(\tilde{\Omega}) = 1 .$$

Then, probability of $\omega \in A$ is defined as

$$\tilde{P}(A \cap B) = \frac{P(A \cap B)}{P(B)} .$$

Def Conditional probability is

$$P(A|B) := \frac{P(A \cap B)}{P(B)} \quad \text{provided} \quad P(B) > 0 .$$

Obs If $A$ and $B$ are independent, then

$$P(A \cap B) = P(A)P(B) \Rightarrow$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = P(A) \longrightarrow \text{no dependence on } B.$$

Similarly, using $\ddot{P}$ we define $E[X|B]$ as $E[X]$.

**Def** $$E[X|B] := \frac{1}{P(B)} \int_B X \, dP \quad \left( \begin{array}{c} \text{mean of } X \\ \text{over } B \end{array} \right)$$

~~Next step: define E[X|Y] given two r.v. X and Y.~~

Extend $B \to \mathcal{H} \subset \mathcal{F}$, $\sigma$-algebra.

**Def** $\mathbb{F} = E[X|\mathcal{H}]$ is the function $\mathbb{F}: \Omega \to \mathbb{R}^n$ s.t.

(a) $\mathbb{F}$ is $\mathcal{H}$-measurable

(b) $\int_H \mathbb{F} \, dP = \int_H X \, dP$ for all $H \in \mathcal{H}$.

**Obs** $\mathbb{F}$ is a random variable on $(\Omega, \mathcal{H}, P|\mathcal{H})$.

It is a.s. unique.

**Proof** Let $\mu(H) = \int_H X \, dP$, $H \in \mathcal{H}$, measure on $\mathcal{H}$.

$\mu$ is absolutely cont. w.r.t. $P|\mathcal{H}$ $\Rightarrow$

$\Rightarrow \exists$ $P|\mathcal{H}$ unique, $\mathcal{H}$-measurable function $F$ s.t.

$\mu(H) = \int_H F \, dP$ for $H \in \mathcal{H}$. $\blacksquare$

**Remark** (Def) $\mu$ is abs. cont. w.r.t. $\lambda$ ($\mu \ll \lambda$)

if $\lambda(A) = 0 \Rightarrow \mu(A) = 0$.

~~Obs~~ **Th** Suppose $X, Y : \Omega \to \mathbb{R}^n$ random variables $L^1$ and $\mathbb{E}[|Y|] < \infty$. Then

(a) $E[aX + bY \mid \mathcal{H}] = aE[X|\mathcal{H}] + bE[Y|\mathcal{H}]$, for $a, b \in \mathbb{R}$

(b) $E[E[X|\mathcal{H}]] = E[X]$.

(c) $E[X|\mathcal{H}] = X$ if $X$ is $\mathcal{H}$-measurable.

(d) $E[X|\mathcal{H}] = E[X]$ if $X$ is independent of $\mathcal{H}$.

(e) $E[\underbrace{Y \cdot X}_{\text{scalar product}} | \mathcal{H}] = Y \cdot E[X|\mathcal{H}]$ if $Y$ is $\mathcal{H}$-measurable

__Pr. (d)__ (ex)

$$\int_H X dP = \int_\Omega X \chi_H dP = \left(\int_\Omega X dP\right)\left(\int_\Omega \chi_H dP\right) =$$

$$= E(X) P(H) \Rightarrow F \equiv E[X] \text{ in Def.} \blacksquare$$

**Th** If $\mathcal{G}, \mathcal{H}$ are $\sigma$-algebras and $\mathcal{G} \subset \mathcal{H}$. Then

$$E[X|\mathcal{G}] = E[E[X|\mathcal{H}]|\mathcal{G}].$$

**Pr** $G \in \mathcal{G} \Rightarrow G \in \mathcal{H} \Rightarrow \underbrace{\int_G E[X|\mathcal{H}] dP}_{E[E[X|\mathcal{H}]|G]} = \underbrace{\int_G X dP}_{E[X|G]}$

for any $G \in \mathcal{G}$. $\blacksquare$

Ex. (Oks.)
__2.1 – 2.7__

<u>Def</u> ~~...~~ Let $A_1, A_2, \ldots \in \mathcal{F}$ be events.

The $A_n$ infinitely often ($A_n$ i.o.) event is

$$\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m = \{\omega \in \Omega : \omega \text{ belongs to infinitely many of } A_n\}.$$

<u>Lemma</u> (Borel - Cantelli).

If $\sum_{n=1}^{\infty} P(A_n) < \infty \implies P(A_n \text{ i.o.}) = 0.$

<u>Pr</u> $P(A_n \text{ i.o.}) \le P\left(\bigcup_{m=n}^{\infty} A_m\right) \le \sum_{m=n}^{\infty} P(A_m) \xrightarrow{n\to\infty} 0.$ ∎

<u>Convergence of random variables</u>: $X_1, X_2, \ldots \to X$.

<u>Def</u> (Convergence in <u>distribution</u> (<u>in law</u> / <u>weakly</u>))

$X_n \to X$ if $\mu_{X_n} \to \mu_X$ weakly.

$\iff P(X_n \in A) \to P(X \in A)$ for any <u>continuity</u> set

$(\mu_X(\partial A) = 0) \longleftarrow$ ($A \in \mathcal{B}$, $\partial A$ is zero ~~...~~ $\mu_X$ ~~...~~ measure)

$\iff P(X_n \le x) \to P(X \le x)$ for all continuity points of $x \mapsto P(X \le x)$

$\iff E[f(X_n)] \to E[f(X)]$ for all bounded continuous functions $f: \mathbb{R}^n \to \mathbb{R}$.

etc.

Def $X_n \to X$ in probability if $P(|X_n - X| \geq \varepsilon) \to 0$ for any $\varepsilon > 0$.

Def $X_n \to X$ a.s. (also called strongly, with probability 1, etc.)

if $P(\lim_{n\to\infty} X_n = X) = 1$.

Obs We mean $P(\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1$

Th If $X_n \to X$ in probability, then there exists an a.s. convergent subsequence: $X_{n_j} \to X$ a.s.

Pr Take $n_1 < n_2 < \dots$ such that

$$P(|X_{n_j} - X| > \tfrac{1}{j}) \leq \frac{1}{j^2}.$$

Since $\sum \frac{1}{j^2} < \infty \overset{\text{lemma}}{\underset{\text{B-C.}}{\Longrightarrow}} P(A_j \text{ i.o.}) = 0$ for $A_j = \{|X_{n_j} - X| > \frac{1}{j}\}$

$\Rightarrow |X_{n_j}(\omega) - X(\omega)| \leq \frac{1}{j}$ for $j \geq J(\omega)$ a.s.

$\Rightarrow X_{n_j}(\omega) \to X(\omega)$. a.s. ∎

Characteristic function (probabilistic Fourier transform)

$\cancel{\Phi_X : \mathbb{R}^n \to \mathbb{C}}$

Def Given a rand. var. $X : \Omega \to \mathbb{R}^n$ we define the characteristic function $\varphi_X : \mathbb{R}^n \to \mathbb{C}$ as

$$\varphi_X(u) = E\left[ e^{iu\cdot X} \right]. \qquad (u\cdot X = u_1 X_1 + \dots + u_n X_n.)$$

$$\cancel{= \int_{\mathbb{R}^n} e^{iu\cdot X} P[X}$$

Obs
$$\varphi_X(u) = \int_{\mathbb{R}^n} e^{iu\cdot x}\, d\mu_X(x) = \int_{\mathbb{R}^n} e^{iu\cdot x}\, P[X \in dx].$$

Th $\varphi_X = \varphi_Y \iff X \overset{law}{=} Y$ ($\mu_X = \mu_Y$, equal distributi

(See prob. Th.)

Examples $X = N(m, \sigma^2)$ 1D normal r.v.

$$\Rightarrow \varphi_X(u) = \exp\left( imu - \frac{u^2\sigma^2}{2} \right), \quad u \in \mathbb{R}.$$

$X = N(m, C)$ nD normal r.v.

$$\Rightarrow \varphi_X(u) = \exp\left( i m\cdot u - \frac{1}{2} u^T C u \right), \quad u \in \mathbb{R}^n$$

(ex.)

**Th** $X = (X_1, \dots, X_n)$ is normal iff

$Y = \lambda_1 X_1 + \dots + \lambda_n X_n$ is normal for all $\lambda_1, \dots, \lambda_n \in \mathbb{R}$.

**Pr** $(\Rightarrow)$

$$\begin{aligned}
\varphi_Y(u) &= E[\exp(iu(\lambda_1 X_1 + \dots + \lambda_n X_n))] \\
&= E[\exp(i(u\lambda_1 X_1 + \dots + u\lambda_n X_n))] \\
&= E[\exp(i\,\mathcal{U}\cdot X)], \quad \left(\begin{array}{l} \mathcal{U} = (u\lambda_1, \dots, u\lambda_n) \\ \mathcal{U} = u\lambda, \quad \lambda = (\lambda_1, \dots, \lambda_n) \end{array}\right) \\
&= \varphi_X(\mathcal{U}). \\
&= \exp\Big(i \underbrace{m\cdot\lambda}_{\tilde{m}}\, u - \frac{u^2}{2} \underbrace{\lambda^T C \lambda}_{\tilde{\sigma}^2}\Big)
\end{aligned}$$

is a char. function of $N(\tilde{m}, \tilde{\sigma}^2) \Rightarrow Y$ is normal.

$(\Leftarrow)$ ex.: compute $\varphi_X$ using $\varphi_Y$.

**Th** Let $X_k : \Omega \to \mathbb{R}^n$ be noral for $k = 1, 2, \dots$

If $E[|X_k - X|^2] \to 0$ as $k \to \infty$, then $X$ is normal.

**Pr** ~~Show that~~ $|e^{iu\cdot x} - e^{iu\cdot y}| < |u|\,|x-y|$ (ex.)

Take $\varphi_{X_n} = e^{iu\cdot X_n}$, $\varphi_X = e^{iu\cdot X}$ for any given $u \in \mathbb{R}^n$.

Then $E[|\varphi_{X_n} - \varphi|^2] < |u|^2 E[|X_n - X|^2] \to 0$ by cond. of Th. ($L_2$-conv.).

$\xrightarrow{(L_2 \Rightarrow L_1)}$ $E[\varphi_{X_n} - \varphi] \to 0$. $\Rightarrow E[Z]$

<u>Pr</u> Show that $\left| e^{iu\cdot x} - e^{iu\cdot y} \right| < |u|\,|x-y|$ (ex.)

Take $\Phi_n(u) = e^{iu\cdot X_n}$, $\Phi(u) = e^{iu\cdot X}$ $\quad \left( \varphi_{X_n}(u) = E[\Phi_n(u)] \right)$

$\left( \varphi_X(u) = E[\Phi(u)] \right)$

Then $E\left[ \left| \Phi_n(u) - \Phi(u) \right|^2 \right] < |u|^2 E\left[ |X_n - X|^2 \right] \to 0$

for any $u \in \mathbb{R}$.

$\Longrightarrow$ $E\left[ \left| \Phi_n(u) - \Phi(u) \right| \right] \to 0$

($L_2$ conv. $\Rightarrow$ $L_1$ conv.)

$\Rightarrow$ $E\left[ \Phi_n(u) - \Phi(u) \right] \to 0$ $\Rightarrow$ $\varphi_{X_n}(u) \to \varphi_X(u)$.

$\Rightarrow$ $\varphi_X(u)$ is a normal ch. function with

~~$m(E[X]) = \lim E[X_n]$ and $C = \lim_{n\to\infty} C_n$.~~

$m = \lim_{n\to\infty} m_n$ and $C = \lim_{n\to\infty} C_n$. ∎

<u>Ex</u>: Show that $\varphi_X^{(k)}(0) = i^k E[X^k]$ for $k = 0, 1, \ldots$

(alternative)

for any random variable $X : \Omega \to \mathbb{R}$.

**Def** Random variables $X_\alpha$, $\alpha \in I$, are identically distributed if ~~they have identical dis~~ their distributions are equal: $\mu_{X_\alpha} = \mu_{X_\beta}$, $\alpha, \beta \in I$.

We usually consider independent identically distributed (i.i.d.) random numbers $X_\alpha$, $\alpha \in I$.

**Th** (Strong law of large numbers)

Let $X_1, X_2, \dots$ be i.i.d. integrable random numbers defined on the same prob. space, with $m = E[X_n]$. Then

$$P\left( \lim_{n \to \infty} \frac{X_1 + \dots + X_n}{n} = m \right) = 1$$

**Ob** In other words ~~the limit = m a.s.~~ the mean converges to $m$ a.s.

**Proof** (see Evans.)

**Th** (Central limit theorem)

~~Let $X_1, X_2, \ldots$ be i.i.d. real valued~~

Let $X_n : \Omega \to \mathbb{R}$, $n = 1, 2, \ldots$, be i.i.d. rand. var.

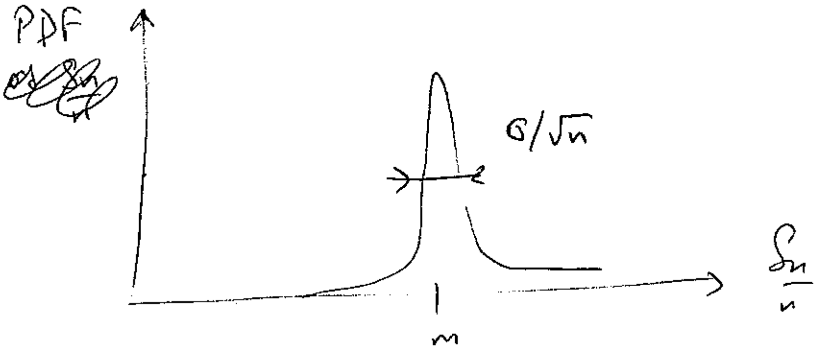
with $m = E[X_n]$ and $\sigma^2 = V[X_n] > 0$.

Let $S_n = X_1 + \cdots + X_n$. Then, for any $-\infty < a < b < \infty$,

$$\lim_{n \to \infty} P\left( a \le \frac{S_n - nm}{\sqrt{n}\,\sigma} \le b \right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}}\, dx.$$

**Proof** (see Evans).

**Interpretation**: For large $n$, the distribution of $S_n/n$ is approximated by $N\left[m, \frac{\sigma}{\sqrt{n}}\right]$.




Thus, fluctuations of $S_n/n$ about the mean value are of order $\sigma/\sqrt{n}$.

Def A stochastic process is a parametrized collection of random variables

$$\{X_t\}_{t \in T}$$

defined on a prob. space $(\Omega, \mathcal{F}, P)$ with values in $\mathbb{R}^n$.

Obs For SDE we use $T = [0, +\infty) = \mathbb{R}_+$ or subintervals. Other options are $T = \mathbb{Z}_+$ for Markov chains or $T = \mathbb{R}^d$ for SPDEs.

Given "time" $t \in T$,

$\omega \to X_t(\omega)$ is a random variable at $t$.

Given ~~the~~ "experiment" / sample $\omega \in \Omega$,

$t \to X_t(\omega)$ is a path of $X_t$.

The function $X : T \times \Omega \to \mathbb{R}^n$ is usually assumed to be jointly measurable in $(t, \omega) \in T \times \Omega$.

<u>Product space</u> (space of paths) $(\mathbb{R}^n)^T$

$(\mathbb{R}^n)^T = \{$ all functions $T \to \mathbb{R}^n\}$.

<u>Product $\sigma$-algebra</u> $\mathcal{F}$ is a $\sigma$-algebra generated by all <u>cylinder sets</u>

$\{\omega \in (\mathbb{R}^n)^T : \omega(t_1) \in F_1, \ldots, \omega(t_k) \in F_k\}$

for some $0 \le t_1 < \ldots < t_k < \infty$ and Borel sets $F_1, \ldots, F_k \subset \mathbb{R}^n$.

The function $X : \omega \to \{X_t(\omega)\}_{t \in T} \in (\mathbb{R}^n)^T$ $(\Omega \to (\mathbb{R}^n)^T)$

is measurable, and one can define a distribution of $X$ on $(\mathbb{R}^n)^T$ as $X_\# P$.

This construction corresponds to the so-called <u>product topology</u>.

One can also consider $\Omega = C(\mathbb{R}_+, \mathbb{R}^n)$ a space of continuous functions with a corresp. topology (uniform conv. on compact spaces) $\to$ Borel $\sigma$-alg.

However, the usual way is to consider <u>finite-dimensional distributions of $X$</u>: the measures

$\mu_{t_1, \ldots, t_k}$ defined on $\mathbb{R}^{nk}$ $(k = 1, 2, \ldots)$ by.

$$\mu_{t_1,\dots,t_k}(F_1 \times \dots \times F_k) = P[X_{t_1} \in F_1, \dots, X_{t_k} \in F_k]$$

for times $0 \le t_1 < \dots < t_k < \infty$ and Borel sets $F_1, \dots, F_k \subset \mathbb{R}$

~~Obs: $\nu_{t_1,\dots,t_k} = (p_{t_1,\dots,t_k})_{\#} P$, push forward by the projection $p_{t_1,\dots,t_k} : \omega \to (\omega(t_1), \dots, \omega(t_k)) \in \mathbb{R}^k$.~~

<u>Consistency conditions</u>

(i) $\nu_{t_{\sigma(1)},\dots,t_{\sigma(k)}}(F_1 \times \dots \times F_k) = \nu_{t_1,\dots,t_k}(F_{\sigma^{-1}(1)} \times \dots F_{\sigma^{-1}(k)})$

for all permutations $\sigma$ on $\{1, \dots, k\}$.

(ii) $\nu_{t_1,\dots,t_k}(F_1 \times \dots, F_k) = \nu_{t_1,\dots,t_k,t_{k+1},\dots,t_{k+m}}(F_1 \times \dots \times F_k \times \mathbb{R}^n \times \dots \times \mathbb{R}^n)$.

for all $m = 1, 2, \dots$

<u>Th</u> Given a collection of finite dimensional measures $\nu_{t_1,\dots,t_k}$ satisfying the consistency conditions, there exists a probability space $(\Omega, \mathcal{F}, P)$ and a stoch. process $\{X_t\}_{t \in T}$ such that these measures are finite-dim. distributions of $X$. This is called the <u>Kolmogorov extension/consistency theorem</u>.

**Obs:** Taking $(\Omega, \mathcal{F})$ to be a product space $\Omega = (\mathbb{R}^n)^T$ with product $\sigma$-algebra, and $X_t : \omega \to \omega_t$ the canonical process, we then the measure $P$ from Kolm. ext. Th. exists and is unique. ( ).
Thus, in product topology, equality of ~~measures~~ distribution $\Leftrightarrow$ eq. of fin.dim. dists

**Def** Let $\{X_t\}$ and $\{Y_t\}$ be stoch. processes on $(\Omega, \mathcal{F}, P)$. We say that $\{X_t\}$ is a version (or a modification) of $\{Y_t\}$ if

$$P(\{\omega : X_t(\omega) = Y_t(\omega)\}) = 1 \quad \text{for all } t.$$

**Obs:** Finite-dim. distributions of $\{X_t\}$ and $\{Y_t\}$ are equal $\Rightarrow$ processes are equal in law on $(\mathbb{R}^n)^T$.
However, pathwise properties may be different.
(Ex. 2.9.: $t \mapsto X_t(\omega)$ is cont. for all $\omega$, $t \mapsto Y_t(\omega)$ is discont. for all $\omega$).

(page 20).

Example Let $(\Omega, \mathcal{F}, P) = (\mathbb{R}_+, \mathcal{B}, \mu)$,

where $\mu$ is an absolutely cont. measure on $\mathbb{R}_+$, $\mu(\mathbb{R}_+) = 1$. Define

$$X_t(\omega) = \begin{cases} 1, & t = \omega \\ 0, & t \neq \omega \end{cases}. \qquad \left(\begin{matrix}\text{discontinuous} \\ \text{for all } \omega\end{matrix}\right)$$

and

$$Y_t(\omega) \equiv 0. \qquad (\text{continuous for all } \omega)$$

Then $\{\omega : X_t(\omega) = Y_t(\omega)\} = \{\omega \in \mathbb{R}_+, \omega \neq t\}$

$\Rightarrow P(\{\dots\}) = 1 \Rightarrow Y$ is a version of $X$.

Also, $\{\omega : X_{t_1} \in F_1, \dots X_{t_k} \in F_k\}$ differs from

$\{\omega : Y_{t_1} \in F_1, \dots, Y_{t_k} \in F_k\}$ by at most

the discrete values $\omega \in \{t_1, \dots, t_k\}$.

$\Rightarrow$ all finite-dim. distributions are the same.
(abs. cont.)

Th (Kolmogorov continuity theorem)

Let $\{X_t\}$ be a stoch. process. Assume that, for any $T>0$ there exist $\alpha, \beta, D > 0$ such that

$$E[\,|X_t - X_s|^\alpha\,] \le D\,|t-s|^{1+\beta}, \quad 0 \le s,t \le T$$

Then, there exists a continuous version $Y$ of $X$, i.e.,

~~$t \mapsto Y_t(\omega)$ is a continuous function for all $\omega \in$~~

all paths $t \mapsto Y_t(\omega)$, $\omega \in \Omega$, are continuous functions.

Obs: In fact, all paths of $Y$ are locally $\gamma$-Hölder-continuous for every $0 < \gamma < \beta/\alpha$.

~~( $|Y_t(\omega) - Y_s(\omega)| / |t-s|^\gamma$ is bounded for each $\omega$ and $t,s$ for finite time~~

$$\left( \sup_{t,s \in I} \frac{|Y_t(\omega) - Y_s(\omega)|}{|t-s|^\gamma} < \infty \text{ for each } \omega \in \Omega \text{ and finite time interval } I \right)$$

# Brownian motion / Wiener process

**Def** A real valued stoch. process ~~{}~~

$\{B_t\}_{t\geq 0}$ is called a Brownian motion (or Wiener process) starting at 0 if

(i) $B_0 = 0$ a.s.

(ii) $B_t - B_s$ is $N(0, t-s)$ for all $t \geq s \geq 0$

(iii) for all $0 < t_1 < t_2 < \dots < t_n$, the random variables $B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}$ are independent (independent increments)

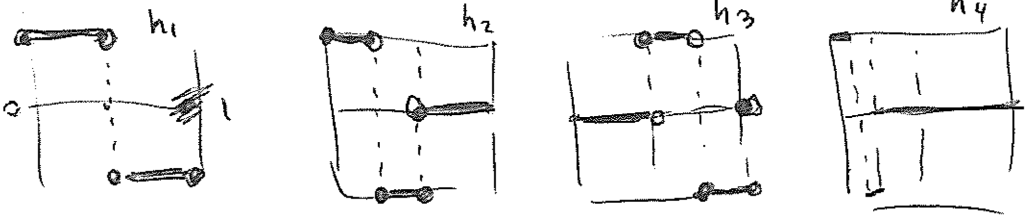
**Obs** Independent ~~kicks~~ disturbances, C.L.T $\to$ Gaussianity.

**Obs** Similarly, define Wiener process starting at $x \in \mathbb{R}$

(i) $\to$ $B_0 = x$ a.s.

## Explicit construction

The family of Haar function (Haar wavelet) on the unit interval $0 \leq t \leq 1$:




etc.

$$h_0(t) \equiv 1, \qquad h_1(t) = \begin{cases} 1, & 0 \le t \le 1/2 \\ -1, & 1/2 < t \le 1 \end{cases}$$

For $k = 2^n, \dots, 2^{n+1}-1$, $n = 1, 2, \dots$ take

$$h_k(t) = \begin{cases} 2^{n/2} & \text{for } \frac{k-2^n}{2^n} \le t \le \frac{k-2^n+1/2}{2^n} \\ -2^{n/2} & \text{for } \frac{k-2^n+1/2}{2^n} < t \le \frac{k-2^n+1}{2^n} \\ 0 & \text{otherwise.} \end{cases}$$

<u>Lemma 1</u> $\{h_k\}_{k=0}^{\infty}$ form a complete, orthonormal basis in $L^2(0,1)$.

<u>Pr.</u> <u>Normality</u>: $\int_0^1 h_k^2 dt = 2^n\left(\frac{1}{2^{n+1}} + \frac{1}{2^{n+1}}\right) = 1$.

<u>Orthogonality</u>: if $l > k \Rightarrow$ either $h_k h_l \equiv 0$ or $h_k = \text{const}$ on $t \in \operatorname{supp} h_l$. Since $\int h_l \, dt = 0 \Rightarrow \int h_k h_l \, dt = 0$.

<u>Completeness</u>: Let $f \in L^2(0,1)$, $\int_0^1 f h_k \, dt = 0$, $k = 0, 1, 2, \dots$
We must prove that $f \equiv 0$ almost everywhere.

$$\left.\begin{array}{l} n = 0 \Rightarrow \int_0^1 f\,dt = 0 \\ n = 1 \Rightarrow \int_0^{1/2} f\,dt - \int_{1/2}^1 f\,dt = 0 \end{array}\right\} \Rightarrow \begin{array}{l} \int_0^{1/2} f\,dt = 0 \\ \int_{1/2}^1 f\,dt = 0 \end{array}$$

$n = 2, 3 \Rightarrow \int_0^{1/4} \quad \int_{1/4}^{2/4} \quad \int_{2/4}^{3/4} \quad \int_{3/4}^{1} = 0$, etc. (by induction)

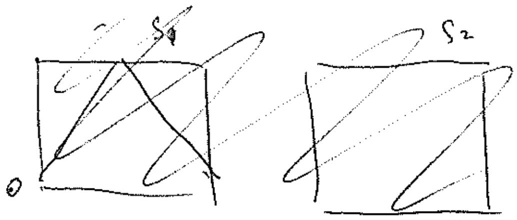
[30]

$\Rightarrow \int_s^r f\,dt = 0$ for all dyadic rationals $r, s$. $(k/2^n)$.

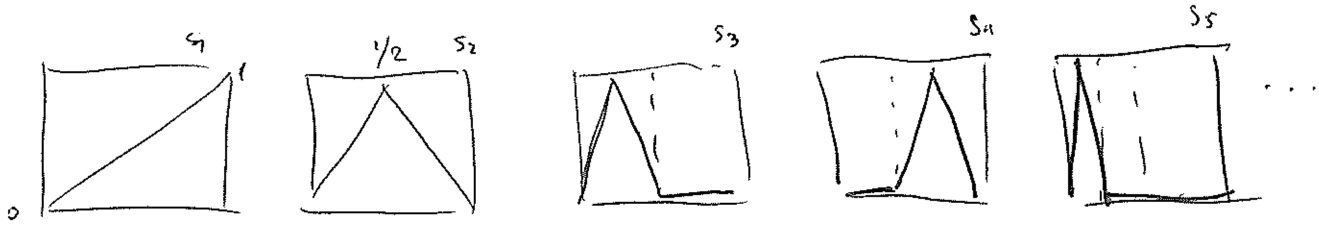
(cont.) $\Rightarrow \int_s^r f\,dt = 0$ for all ~~[illegible]~~ $0 \le s \le r \le 1$.

~~[illegible]~~ $\Rightarrow f = 0$ a.e. (ex.) ∎.

<u>Schauder functions</u>: $S_k(t) = \int_0^t h_k(s)\,ds$, $0 \le t \le 1$.




$\max_t S_k(t) = 2^{-n/2-1}$, $k = 2^n, \dots, 2^{n+1}-1$.




<u>Lemma 2</u> For each $0 \le t, s \le 1$, $\sum_{k=0}^{\infty} S_k(s)\, S_k(t) = t \wedge s$ $(\min\{t,s\})$.

<u>Pr.</u> ex. ∎

<u>Lemma 3</u> Let $\{a_k\}_{k=0}^{\infty}$ be a sequence of real numbers s.t. $|a_k| \le C k^{\delta}$ $(k = 1, 2, \dots)$ for some $C > 0$ and $0 \le \delta < 1/2$.

Then the series $\sum_{k=0}^{\infty} a_k S_k(t)$ converges uniformly ~~[illegible]~~ for $t \in [0,1]$.

Pr. Fix $\varepsilon>0$. For $k=2^n,\dots,2^{n+1}-1$, only one term in the series is nonzero for any $t$. Let

$$b_n = \max_{k=2^n,\dots,2^{n+1}-1} |a_k| \le C(2^{n+1})^\delta.$$

$$\Rightarrow \sum_{k=2^m}^\infty |a_k S_k(t)| \le \sum_{n=m}^\infty b_n \max_{k=2^n,\dots,2^{n+1}-1} |S_k(t)|$$

$$\le C\sum_{n=m}^\infty (2^{n+1})^\delta\, 2^{-n/2-1} = C\,2^{\delta-1}\sum_{n=m}^\infty 2^{n\overbrace{(\delta-\frac{1}{2})}^{<0}} < \varepsilon$$

for $m$ suff. large. ∎

<u>Lemma 4</u> Let $\{A_k\}_{k=1}^\infty$ be independent, $N(0,1)$ random variables. Then, for almost every $\omega$,

$$|A_k(\omega)| = O(\sqrt{\log k}) \text{ as } k\to\infty$$

Pr. For all $x>0$, $k=2,3,\dots$

$$P(|A_k|>x) = \frac{2}{\sqrt{2\pi}}\int_x^\infty e^{-s^2/2}\,ds \le \frac{2}{\sqrt{2\pi}}e^{-x^2/4}\int_x^\infty e^{-s^2/4}\,ds$$

$$\le C e^{-x^2/4} \quad \text{for some } C>0.$$

Taking $x=4\sqrt{\log k}$, we have

$$P(|A_k|\ge 4\sqrt{\log k}) \le Ce^{-4\log k} = C/k^4.$$

Since $\sum 1/k^4 < \infty \implies$ (L. Borel - Cantelli)

$$P\left( |A_k| \geq 4\sqrt{\log k} \quad i.o. \right) = 0$$

$\Rightarrow \quad |A_k(\omega)| \leq 4\sqrt{\log k}$ for a.e. $\omega$ and $k \geq K(\omega)$. ∎

<u>Th</u> (construction on $t \in [0,1]$)

Let $\{A_k\}_{k=0}^{\infty}$ be a sequence of independent, $N(0,1)$ random variables ~~of~~ defined on the same prob. space

Then, the sum $B_t(\omega) = \sum_{k=0}^{\infty} A_k(\omega) s_k(t)$

converges uniformly in $t \in [0,1]$ for a.e. $\omega$.

Also, (i) $B_t$ is a Brownian motion for $t \in [0,1]$.

(ii) for a.e. $\omega$, the sample path $t \mapsto B_t(\omega)$ is continuous.

<u>Pr</u> Use ~~[illegible]~~ Lemmas 3, 4 with $\sqrt{\log k} \leq c k^{1/4}$.

$\Rightarrow$ uniform convergence.

$\Rightarrow$ ~~[illegible]~~ limit is continuous. (all $s_k$ are continuous)

- $B_0 = 0$ because $s_k(0) = 0$.

- Let us show that $B_t - B_s$ is $N(0, t-s)$, $0 \leq s \leq t \leq$

Characteristic function:

$$E\left[ e^{i\lambda (B_t - B_s)} \right] = E\left[ e^{i\lambda \sum_{k=0}^{\infty} A_k (S_k(t) - S_k(s))} \right]$$

$$= \prod_{k=0}^{\infty} E\left[ e^{i\lambda A_k (S_k(t) - S_k(s))} \right] \qquad \text{(independence)}$$

$$= \prod_{k=0}^{\infty} e^{-\frac{\lambda^2}{2}(S_k(t) - S_k(s))^2} \qquad (A_k \sim N(0,1))$$

$$= \exp\left( -\frac{\lambda^2}{2} \sum_{k=0}^{\infty} (S_k(t) - S_k(s))^2 \right)$$

$$= \exp\left( -\frac{\lambda^2}{2} \sum_{k=0}^{\infty} \left[ S_k^2(t) - 2 S_k(t) S_k(s) + S_k^2(s) \right] \right)$$

$$\overset{L.2}{=} \exp\left( -\frac{\lambda^2}{2} (t - 2s + s) \right) = \underbrace{e^{-\frac{\lambda^2}{2}(t-s)}}_{\text{char. funct. of } N(0,t-s)}$$

$$\Rightarrow \quad B_t - B_s \sim N(0, t-s).$$

- Independence of increments.

Take $0 = t_0 < t_1 < \dots < t_m \le 1$.

Chor. function of $(B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \dots, B_{t_m} - B_{t_{m-1}})$

$$E\left[ e^{i \sum_{j=1}^{m} \lambda_j (B_{t_j} - B_{t_{j-1}})} \right] = \prod_{j=1}^{m} e^{-\frac{\lambda_j^2}{2}(t_j - t_{j-1})}.$$

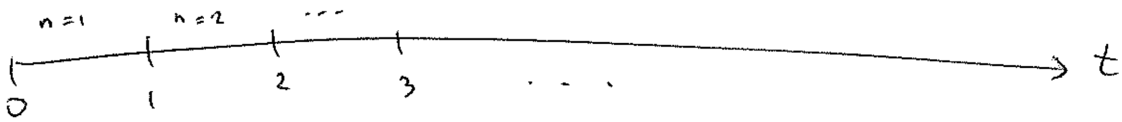
similar computation (ex.)
using Lemma 2.

This is the charact. function of independent Gaussian rand. variables $N(0, t_j - t_{j-1})$. $\Rightarrow$ independence

Th (Existence of Brownian motion)

Let $(\Omega, \mathcal{F}, P)$ be a prob. space with countably many independent, $N(0,1)$ random variables $\{A_n\}_{n=1}^{\infty}$. Then there exists a one-dimensional Brownian motion (Wiener process) $B_t(\omega)$, $\omega \in \Omega, t \geq 0$.

Pr. Divide time line into integer sections.




Take independent $\{\tilde{A}_{n1}, \tilde{A}_{n2}, \dots\}$ by relabling $\{A_n\}$, for all $n = 1, 2, \dots$ $\Rightarrow$ Build independent $B_t^{(n)}$ for all $n$.

Iteratively define: $B_t = B_t^{(1)}$, $t \in [0, 1]$

$$B_t = B_{n-1} + B^{(n)}_{t-(n-1)}, \quad t \in [n-1, n].$$

This is a Brownian motion on $t \geq 0$.

Obs All we need are countably many independent $N(0,1)$ variables!

# Brownian motion (Wiener process)

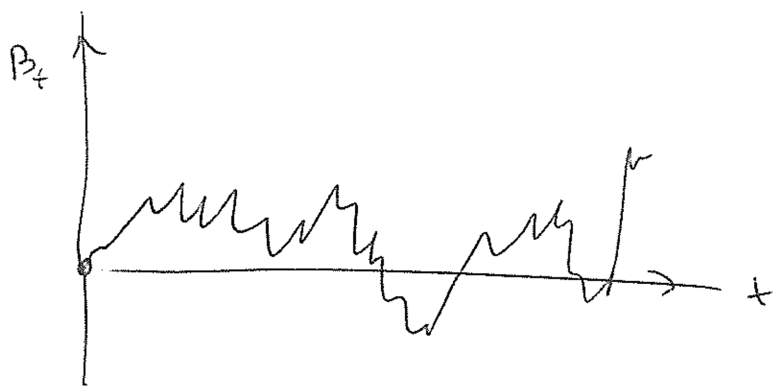
Let $(\Omega, \mathcal{F}, P)$ be a prob. space.

<u>Def</u> A real-valued stochastic process $B_t$, $t \geq 0$, is called a Br. m. / W. pr. (starting at $x \equiv 0$) if

(i) $B_0 = 0$ a.s.

(ii) $B_t - B_s$ is $N(0, t-s)$ for all $t \geq s \geq 0$.

(iii) For any $0 < t_1 < \dots < t_n$, the random variables $B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}$ are independent (independent increments).




$B_t$ is $N(0,t)$ with prob. density

$$\nu_t(F) = \frac{1}{\sqrt{2\pi t}} \int_F e^{-x^2/2t}\, dx \quad , F \in \mathcal{B} \text{ Borel in } \mathbb{R}.$$

## Finite-dimensional distributions

$$\nu_{t_1,\dots,t_n}(F_1 \times \dots \times F_n) = P(B_{t_1} \in F_1, \dots, B_{t_n} \in F_n)$$

**Th** For any ~~measurable~~ integrable function $f : \mathbb{R}^n \to \mathbb{R}$ and ~~all~~ $0 = t_0 < t_1 < \dots < t_n$ we have

$$E[f(B_{t_1}, \dots, B_{t_n})] = \int_{\mathbb{R}^n} f(x_1, \dots, x_n)\, p(t_1, 0, x_1)\, p(t_2 - t_1, x_1, x_2) \dots p(t_n - t_{n-1}, x_{n-1}, x_n)\, dx_1 \dots dx_n$$

for $p(t, x, y) = \underbrace{\frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t}\right)}_{N(0,t) \text{ w.r.t. } x-y.}$.

**Pr** Take $X_i = B_{t_i}$, $Y_i = X_i - X_{i-1}$,

$\cancel{k(y_1, y_2, \dots, y_n) = f(y_1, y_1 + y_2, \dots, y_1 + y_2 + \dots + y_n)}$

Then $X_1 = Y_1$, $X_2 = Y_1 + Y_2$, $\dots$, $X_n = Y_1 + \dots + Y_n$.

Variables $Y_i$ are independent and $N(0, \underbrace{t_i - t_{i-1}}_{\Delta t_i})$.

We have

$$E[f(B_{t_1},\dots,B_{t_n})] = E[f(X_1,\dots,X_n)] =$$

$$= E[f(Y_1, Y_1+Y_2, \dots, Y_1+\dots+Y_n)]$$

$$= \int_{\mathbb{R}^n} f(y_1, y_1+y_2, \dots, y_1+\dots+y_n) \prod_{i=1}^n \frac{e^{-y_i^2/2\Delta t_i}}{\sqrt{2\pi \Delta t_i}} \, dy_1 \dots dy_n \quad (=)$$

change of coordinates $x_1 = y_1$, $x_2 = y_1+y_2, \dots, x_n = y_1+\dots+y_n$

$$y_i = x_i - x_{i-1} \qquad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & & & 0 \\ 1 & 1 & & \\ 1 & 1 & 1 & \\ \dots & & & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \Rightarrow \det = 1$$

$$(=) \int_{\mathbb{R}^n} f(x_1,\dots,x_n) \prod_{i=1}^n \underbrace{\frac{\exp\left(-\frac{x_i - x_{i-1}}{2\Delta t_i}\right)}{\sqrt{2\pi \Delta t_i}}}_{p(\Delta t_i, x_{i-1}, x_i)} \, dx_1 \dots dx_n$$

In particular, take

$$f(x_1,\dots,x_n) = \chi_{F_1}(x_1) \dots \chi_{F_n}(x_n) = \begin{cases} 1, & x_1 \in F_1, \dots, x_n \in F_n \\ 0 & \text{otherwise} \end{cases}$$

$$\Rightarrow \nu_{t_1,\dots,t_n}(F_1 \times \dots \times F_n) = \int_{F_1 \times \dots \times F_n} p(t_1, 0, x_1)\, p(t_2 - t_1, x_1, x_2) \dots p(t_n - t_{n-1}, x_{n-1}, x_n) \, dx_1 \dots dx_n.$$

Ex: Check that finite-dim distributions satisfy the continuity conditions of K. extension Th.

$\Rightarrow$ existence of Br. motion follows from K. ext. Th.

## Continuity of Br. motion $\quad t \to B_t(\omega)$.

**Def** $f : [0,T] \mapsto \mathbb{R}$ is (uniformly) Hölder continuous with exponent $0 < \gamma \le 1$ if $\exists K > 0$ s.t.

$$|f(t) - f(s)| \le K\,|t-s|^{\gamma} \quad \text{for all } s,t \in [0,T].$$

**Obs** $C^1 \Rightarrow$ H. cont. with $\gamma = 1$.

H. cont. with $\gamma > 0 \Rightarrow C$.

**Def** $f$ is $\gamma$-Hölder cont. at point $s$ if $\exists K > 0$ s.t. $|f(t) - f(s)| \le K|t-s|^{\gamma}$ for $t \in$ some neigh. of $s$.

**Th** (Kolmogorov continuity)

Let $X_t$ be a stochastic process. Assume that for any $T > 0$ there exist $\alpha, \beta, D > 0$ s.t.

$$E\left[\,|X_t - X_s|^{\alpha}\right] \le D\,|t-s|^{1+\beta} \quad \text{for } 0 \le s,t \le T.$$

Then, there exists a version $Y_t$ of $X_t$

( $X_t = Y_t$ a.s. for any $t$) such that

& $Y_t$ is sample-continuous ( $Y_t(\omega) \in C$ for any $\omega$).

Moreover, paths $Y_t(\omega)$ are $\gamma$-Hölder continuous for any $t \in [0,T]$ and $0 < \gamma < \beta/\alpha$.

For Brownian motion:

$$E[\,|B_t - B_s|^{2m}] = \frac{1}{\sqrt{2\pi r}} \int x^{2m} e^{-x^2/2r} dx, \quad r = t-s>0$$

$$= \frac{r^m}{\sqrt{2\pi}} \int y^{2m} e^{-y^2/2} dy \quad \left( y = \frac{x}{\sqrt{r}} \right) = D\, r^m = D|t-s|^m.$$

$\Rightarrow$ Kolmogorov's th. holds with $\alpha = 2m$, $\beta = m-1$.

$\Rightarrow$ <u>Brownian motion has a continuous version</u>

(we always assume, e.g. see construction from previous lecture).

# Hölder exponents

$$\frac{\beta}{\alpha} = \frac{m-1}{2m} \quad \text{for any } m \geq 1, \quad \frac{\beta}{\alpha} \to \frac{1}{2} \text{ as } m \to \infty.$$

Hence, Brownian motion is $\gamma$-Hölder continuous for any $\underline{0 < \gamma < 1/2}$.

~~Th (i) For a.e. $\omega$, paths $B_t(\omega)$ are nowhere Hölder continuous with exponent for any $\gamma \in [\frac{1}{2}, 1]$.~~

**Th** For any $\gamma \in [\frac{1}{2}, 1]$ and almost every $\omega$, the path $B_t(\omega)$ is nowhere Hölder continuous.

**Cor** For almost every $\omega$, the path $t \to B_t(\omega)$ is nowhere differentiable and is of infinite variation on each subinterval.

**Obs** Finite variation implies differentiability a.e.

(See proof in Evans.)

<u>Comment</u> about 1/2 - Hölder continuity

$t$ is called a slow point of $B_t(\omega)$ if

$$\overline{\lim_{h \to 0}} \frac{|B_{t+h} - B_t|}{|h|^{1/2}} < \infty \quad \left( \begin{array}{l} 1/2 \text{ Hölder} \\ \text{at point } t \end{array} \right)$$

- The set of slow points has zero measure
- The set of slow points has full (=1) Hausdorff dimension.

(Between "almost never" and "almost always").

## <u>Generalizations</u>

- Brownian motion starting at point $x \in \mathbb{R}$: replace $B_0 = 0$ a.s. by $B_0 = x$ a.s.

- Brownian motion in $\mathbb{R}^n$, $B_t = (B_t^1, \dots, B_t^n)$, starting at $x \in \mathbb{R}^n$.

<u>Def (version Evans)</u>

- $B_0 = x$ ~~a.s.~~ $\in \mathbb{R}^n$ a.s.

- Every component $B_t^k$ is a one-dimensional Brownian motion starting at $x_k$.

- Independence of components:
  σ-algebras $\mathcal{H}^k$ generated by $\{ B_t^k : t \geq 0 \}$ are independent.

Def (version Oksendal)

- $B_0 = x \in \mathbb{R}^n$ a.s.
- Finite-dimensional distributions are

$$\mu_{t_1,\dots,t_n}(F_1 \times \dots \times F_n) := P(B_{t_1} \in F_1, \dots, B_{t_n} \in F_n)$$

$$= \int_{F_1 \times \dots \times F_n} p(t_1, x, x_1)\, p(t_2 - t_1, x_1, x_2) \dots p(t_n - t_{n-1}, x_{n-1}, x_n)\, dx_1 \dots dx_n.$$

with $F_1, \dots, F_n$ are Borel sets in $\mathbb{R}^n$ and

$$p(t, x, y) = (2\pi t)^{-n/2} \exp\left( - \frac{|x-y|^2}{2t} \right), \quad t \geq 0,\ x, y \in \mathbb{R}^n.$$

Ex: Show equivalence.

Obs Canonical Br. motion is defined by these finite-dimensional distributions on $C([0,\infty), \mathbb{R}^n)$.

So, $B_t(\omega) = \omega(t)$, $\omega \in C([0,\infty), \mathbb{R}^n)$.

Properties:

(a) $\underset{\text{(mean)}}{E[B_t]} = \int y \, p(t, x, y) \, d^n y \qquad (y = B_t)$

$$= \int y \, (2\pi t)^{-n/2} \exp\left(-\frac{|x-y|^2}{2t}\right) d^n y \qquad z = x - y$$

$$= \int (x + z) \, (2\pi t)^{-n/2} \, \cancel{\exp} \, e^{-|z|^2/2t} \, d^n z$$

$$= x \underbrace{\int (2\pi t)^{-n/2} e^{-|z|^2/2t} d^n z}_{1} + \int \underbrace{z \, (2\pi t)^{-n/2} e^{-|z|^2 dt} d^n z}_{\text{odd function} \Rightarrow 0.}$$

$$\Rightarrow \quad E[B_t] = x.$$

(b) $\cancel{\underset{\text{(variance)}}{E[|B_t - x|^2]} = \sum_{k=1}^{n} E[(B_t^k - x_k)^2]}$

$$E[(B_t^k - x_k)(B_t^l - x_l)] = \underbrace{\delta_{kl} \, t.}_{\text{covariance matrix}}$$

$$B_t \text{ is } N[x, t\underset{\uparrow\ n \times n \text{ identity matrix.}}{I}]$$

In particular,

$$E[|B_t - x|^2] = \sum_{k=1}^{n} E[(B_t^k - x_k)^2] = nt. \qquad \text{(variance)}$$

(c) At different times s, t:

$$E[(B_t - x)\cdot(B_s - x)] = \quad (\text{assume } 0 < s < t)$$

$$= E[((B_t - B_s) + (B_s - x))\cdot(B_s - x)] =$$

$$= \underbrace{E[(B_t - B_s)(B_s - x)]}_{\substack{\text{independent with}\\ \text{zero mean}}} + E[|B_s - x|^2] = ns.$$

Generally, $E[(B_t - x)\cdot(B_s - x)] = n \underbrace{\min(s,t)}_{s \wedge t}$.

Similarly, $E[|B_t - B_s|^2] = n(t-s)$, $t \geq s$.

For a collection $Z = (B_{t_1}, \dots, B_{t_k}) \in \mathbb{R}^{nk}$:

Gaussian $N(M, C)$, where

$M = E[Z] = (x, x, \dots, x) \in \mathbb{R}^{nk}$.

$$C = \begin{pmatrix} t_1 I & t_1 I & t_1 I & \cdots \\ t_1 I & t_2 I & t_2 I & \cdots \\ t_1 I & t_2 I & t_3 I & \cdots \\ \vdots & \vdots & & \end{pmatrix}$$

(d) Time-scaling: $\frac{1}{c} B_{c^2 t}$ is Brownian for any $c > 0$.

Pr Denote $\tilde{B}_t = \frac{1}{c} B_{c^2 t}$.

Increments are still independent.

We have to show that $\tilde{B}_t - \hat{B}_s$ is $N(0, t-s)$ for $t > s > 0$.

$$\tilde{B}_t - \hat{B}_s = \underbrace{\frac{1}{c}}_{\substack{\text{changes}\\ \text{variance}\\ \text{by } 1/c^2}} \underbrace{\left( B_{c^2 t} - B_{c^2 s} \right)}_{N(0,\, c^2(t-s))} \stackrel{?}{=} N(0, t-s).$$

(c) Stationarity:

A stochastic process $X_t$ is stationary if it has the same distribution (finite-dim. distributions) as $\tilde{X}_t = X_{t+h}$ for any fixed $h > 0$.

Increments of Brownian motion w.r.t.

~~$X_t = B_{t+\Delta t} - B_t$ are stationary for any $\Delta t$.~~
a ~~fixed time are stationary.~~

$$\xi_t = \frac{B_{t+a} - B_t}{a} \quad \text{for fixed } a > 0.$$

In the limit $a \to 0$, this is called the <u>white noise</u>.

(f) White noise (heuristics)

$$\xi_t = \underbrace{\mathring{B}_t}_{\text{non-differentiable!}} \approx \frac{B_{t+a} - B_t}{a} \quad \text{for small } a > 0.$$

$$r(\tau) = E[\xi(t)\, \xi(\underbrace{t-\tau}_{s})] \quad \text{autocorrelation function}$$

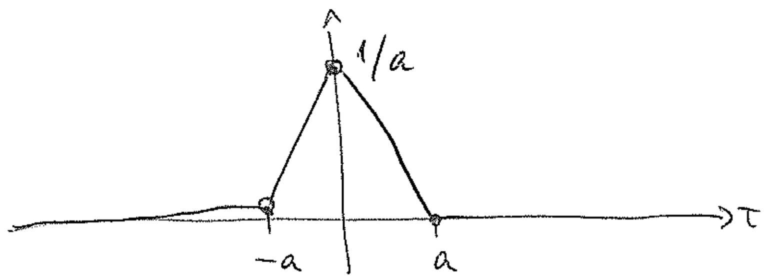
$$= E\left[ \frac{B_{t+a} - B_t}{a} \, \frac{B_{s+a} - B_s}{a} \right]$$

$$= \frac{1}{a^2} \Big( E[B_{t+a} B_{s+a}] - E[B_{t+a} B_s] - E[B_t B_{s+a}] +$$

$$+ E[B_t B_s] \Big)$$

$$= \frac{1}{a^2} \big( (t+a) \wedge (s+a) - (t+a) \wedge s - t \wedge (s+a) + t \wedge s \big)$$

$$= \frac{1}{a^2} \big( a \wedge (a-\tau) - (a \wedge (-\tau)) - 0 \wedge (a-\tau) + 0 \wedge (-\tau) \big)$$




As $a \to 0$ we have $r(\tau) \to \delta(\tau)$.

Power spectrum density (power in each frequency $f$)

is $S(f) = \hat{r}(f) \equiv 1$. ← white! (Analog of Parseval $\int |x(t)|^2 dt = \int |\hat{x}(f)|^2 df$)

(Wiener-Khinchin Th.)