# ABC of Malliavin calculus

Dario Gasbarra

16. toukokuuta 2013

# Gaussian vs Lebesgue measure

In probability theory , usually we work on an abstract measurable space $(\Omega, \mathcal{F})$ equipped with a probability measure $P$.

In analysis instead we usually work concretely with the euclidean space $\mathbb{R}^d$ equipped with Lebesgue measure.

The Lebesgue measure on $\mathbb{R}^d$ is $\sigma$-finite, meaning that $\mathbb{R}^d$ is covered by a countable union of unit cubes $(z + [0, 1]^d)$, $z \in \mathbb{Z}^d$.

On each finite dimensional unit cube the Lebesgue measure is a probability, i.e. integrates to 1.

By Kolmogorov consistency theorem, we can define the product Lebesgue measure on the infinite dimensional unit cube $[0,1]^{\mathbb{N}}$, However the infinite product $\mathbb{R}^{\mathbb{N}}$ cannot be covered by a countable union of unit cubes, $\mathbb{Z}^{\mathbb{N}}$ is not countable.
The infinite product of the Lebesgue measure on $\mathbb{R}^{\mathbb{N}}$ is not $\sigma$-finite.
On $\mathbb{R}^{\mathbb{N}}$ we work instead with Gaussian probability measures.
We start woking in finite dimension.

$$\gamma(x)dx = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) dx$$

is the standard Gaussian measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On $\mathbb{R}^d$, $\gamma^{\otimes d}(x) = \gamma(x_1)\dots\gamma(x_d)$ is the product density.

# Gaussian integration by parts

**Lemma**

*Let $G(\omega)$ a real valued gaussian random variable with $E(G) = 0$ and variance $E(G^2) = \sigma^2$, If $f, h$ are absolutely continuous functions,*

$$f(x) = f(0) + \int_0^x f'(y)dy, \quad h(x) = h(0) + \int_0^x h'(y)dy,$$

*with $f', h' \in L^2(\mathbb{R}, \gamma)$, (i.e. $f'(G), h'(G) \in L^2(\Omega)$) then $f(G), h(G) \in L^2(\Omega)$ and*

$$E(f'(G)h(G)) = E\left( f(G) \left( \frac{h(G)G}{E(G^2)} - h'(G) \right) \right)$$

# Proof

$P(G \in dx) = \gamma(x)dx$ with density

$$\gamma(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) .$$

Note that

$$\frac{d}{dx}\gamma(x) = -\frac{\gamma(x)x}{\sigma^2}$$

# Proof

$P(G \in dx) = \gamma(x)dx$ with density

$$\gamma(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) .$$

Note that

$$\frac{d}{dx}\gamma(x) = -\frac{\gamma(x)x}{\sigma^2}$$

Integrating by parts

$$\int_{-\infty}^{\infty} f'(x)h(x)\gamma(x)dx = -\int_{-\infty}^{\infty} f(x)\frac{d}{dx}\Big(h(x)\gamma(x)\Big)dx$$

$$= \int_{-\infty}^{\infty} f(x)\left(\frac{h(x)x}{\sigma^2} - h'(x)\right)\gamma(x)dx \;\Box$$

More precisely, it hold when $f(x)$ is supported on a finite interval $[a, b]$,

$$E_P(f'(G)h(G)) \int_{-a}^{b} f'(x)h(x)\gamma(x)dx =$$

$$f(b)h(b) - f(a)h(a) - \int_{a}^{b} f(x)\frac{d}{dx}\Big(h(x)\gamma(x)\Big)dx$$

Otherwise, take $f_n(x) = f(x)\eta_n(x)$ with $\eta_n(X)(1 - x/n)^+$, which as support on $[-n, n]$ and $\eta_n(x) \to 1 \ \forall x$ as $n \to \infty$. Note that $\eta_n'(x) = -\frac{1}{n}\text{sign}(x)$. By Lebesgue dominated convergence $h(G)\eta_n(G) \to h(G)$ in $L^2(\gamma)$.

$$E_P(f_n'(G)h(G)) = E_P(f'(G)\eta_n(G)h(G)) - \frac{1}{n}E_P(f(G)\text{sign}(G)h(G))$$

$$\longrightarrow E_P(f'(G)h(G))$$

Denote

$$\partial f(x) := f'(x) \text{ and } \partial^* h(x) := \left( \frac{h(x)x}{\sigma^2} - h'(x) \right)$$

Denote

$$\partial f(x) := f'(x) \text{ and } \partial^* h(x) := \left( \frac{h(x)x}{\sigma^2} - h'(x) \right)$$

Then

$$(\partial f, h)_{L^2(\mathbb{R},\gamma)} = (f, \partial^* h)_{L^2(\mathbb{R},\gamma)}$$

$\partial^*$ is the adjoint of the derivative operator in $L^2(\mathbb{R},\gamma)$.

Denote

$$\partial f(x) := f'(x) \text{ and } \partial^* h(x) := \left( \frac{h(x)x}{\sigma^2} - h'(x) \right)$$

Then

$$(\partial f, h)_{L^2(\mathbb{R},\gamma)} = (f, \partial^* h)_{L^2(\mathbb{R},\gamma)}$$

$\partial^*$ is the adjoint of the derivative operator in $L^2(\mathbb{R}, \gamma)$.

## Definition

*We say that $f \in L^2(\mathbb{R}, \gamma)$, has weak derivative $g \in L^2(\mathbb{R}, \gamma)$ in Sobolev sense if $\forall h$ with classical derivative $h'$ such that $\partial^* h \in L^2(\gamma)$,*

$$\int_{\mathbb{R}} g(x)h(x)\gamma(x)dx = \int_{\mathbb{R}} f(x)\partial^* h(x)\gamma(x)dx$$

*and we denote $\partial f = f' := g$.*

This definition extends the classical derivative. We introduce the weighted Sobolev space

$$W^{1,2}(\mathbb{R},\gamma) := \{f \in L^2 : f \text{ Sobolev differentiable } \}$$

as the $L^2$-closure of Domain($\partial$) with norm

$$\| f \|^2_{W^{1,2}(\gamma)} = \| f \|^2_{L^2(\gamma)} + \| \partial f \|^2_{L^2(\gamma)}$$

We extend also $\partial^*$ to the $L^2(\mathbb{R},\gamma)$ closure of Domain($\partial^*$)

## Lemma

*Let* $f_n \stackrel{L^2(\gamma)}{\to} 0$ *a sequence of smooth functions with* $\partial f_n \stackrel{L^2(\gamma)}{\to} g$. *Then* $g(x) = 0$ *almost everywhere.*

### Proof

For every $h \in L^2(\gamma)$ smooth with $\partial^* h \in L^2(\gamma)$,

$$E_P(\partial f_n(G)h(G)) \to E_P(g(G)h(G))$$
$$= E_P(f_n(G)\partial^* h(G)) \to 0$$

Therefore $E(g(G)h(G)) = 0$. For every $A \in \mathcal{B}(\mathbb{R})$ by smoothing $\mathbf{1}_A(x)$ we find smooth and uniformly bounded $h_\varepsilon(x) \to \mathbf{1}_A(x)$. By bounded convergence theorem it follows that $E(g(G)\mathbf{1}_A(G)) = 0$.

## Proposition

*The gaussian integration by parts formula*

$$E_P(\partial f(G)h(G)) = E_P(f(G)\partial^* h(G))$$

*extends to* $f \in W^{1,2}(\mathbb{R}, \gamma)$ $h \in Domain(\partial^*)$.

## Corollary

*For* $h(x) \equiv 1$, $f \in W^{1,2}(\mathbb{R}, \gamma)$

$$E(f'(G)) = \frac{E(f(G)G)}{E(G^2)}$$

# Linear regression

Let $X(\omega), Y(\omega) \in L^2(P)$ . Then

$$\widehat{X}(\omega) = \widehat{b} + \widehat{a} Y(\omega) \quad \text{with} \tag{0.1}$$

$$\widehat{a} = \frac{E(X(Y - E(Y)))}{E(Y^2) - E(Y)^2} \tag{0.2}$$

$$\widehat{b} = E(X) - \widehat{a} E(Y) \tag{0.3}$$

# Linear regression

Let $X(\omega), Y(\omega) \in L^2(P)$ . Then

$$\widehat{X}(\omega) = \widehat{b} + \widehat{a} Y(\omega) \quad \text{with} \tag{0.1}$$
$$\widehat{a} = \frac{E(X(Y - E(Y)))}{E(Y^2) - E(Y)^2} \tag{0.2}$$
$$\widehat{b} = E(X) - \widehat{a} E(Y) \tag{0.3}$$

is the $L^2$-projection of $X$ on the linear subspace generated by $Y$, such that

$$E((\widehat{X} - X)^2) = \min_{a,b \in \mathbb{R}} E\Big((a + bY - X)^2\Big)$$

In general $\widehat{X}(\omega) \neq E(X|\sigma(Y))(\omega)$, which is the projection of $X$ on the subspace $L^2(\Omega, \sigma(Y), P)$.

In general $\widehat{X}(\omega) \neq E(X|\sigma(Y))(\omega)$, which is the projection of $X$ on the subspace $L^2(\Omega, \sigma(Y), P)$.
When $(X, Y)$ is jointly gaussian, $\widehat{X} = E(X|\sigma(Y))$.

In general $\widehat{X}(\omega) \neq E(X|\sigma(Y))(\omega)$, which is the projection of $X$ on the subspace $L^2(\Omega, \sigma(Y), P)$.
When $(X, Y)$ is jointly gaussian, $\widehat{X} = E(X|\sigma(Y))$.
Let $W \sim \mathcal{N}(0, \sigma^2)$ and consider $F = f(W)$ for some non-linear function $f \in W^{1,2}(\mathbb{R}, \gamma)$. By 0.1 the best linear estimator of $f(W)$ given $W$ is

$$\widehat{f(W)} = E(f(W)) + \frac{E(f(W)W)}{E(W^2)} W$$

In general $\widehat{X}(\omega) \neq E(X|\sigma(Y))(\omega)$, which is the projection of $X$ on the subspace $L^2(\Omega, \sigma(Y), P)$.
When $(X, Y)$ is jointly gaussian, $\widehat{X} = E(X|\sigma(Y))$.
Let $W \sim \mathcal{N}(0, \sigma^2)$ and consider $F = f(W)$ for some non-linear function $f \in W^{1,2}(\mathbb{R}, \gamma)$. By 0.1 the best linear estimator of $f(W)$ given $W$ is

$$\begin{aligned}\widehat{f(W)} &= E(f(W)) + \frac{E(f(W)W)}{E(W^2)} W \\ &= E(f(W)) + E(f'(W))W\end{aligned}$$

$$f(W) = E_P(f(W)|\sigma(W)) = E(f(W)) + E(f'(W))W + M^f$$

Clearly $E(M^f) = 0$, but also

$$E(M^f W) = E\Big(\{f(W) - E(f(W)) - E(f'(W))W\}W\Big)$$
$$= E\left(\left\{f(W) - E(f(W)) - \frac{E(f(W)W)}{E(W^2)}W\right\}W\right) = 0$$

Clearly $E(M^f) = 0$, but also

$$E(M^f W) = E\Big( \{ f(W) - E(f(W)) - E(f'(W))W \} W \Big)$$
$$= E\left( \left\{ f(W) - E(f(W)) - \frac{E(f(W)W)}{E(W^2)} W \right\} W \right) = 0$$

The linearization error $M^f$ is uncorrelated with $W$.

## Lemma

*When the derivatives $f', f''$ are bounded and continuous,*

$$\frac{E\big((M^f)^2\big)}{\sigma^2} \to 0 \quad \text{as } \sigma \to 0$$

**Proof.** Let $G$ with $E_P(G) = 0$, $E_P(G^2) = 1$, not necessarily Gaussian. By Taylor expansion, as $\sigma \to 0$,

$$f(\sigma G) = f(0) + f'(0)\sigma G + \frac{1}{2} f''(0)\sigma^2 G^2 + o_P(1)\sigma^2$$
$$E\big(f(\sigma G)\big) = f(0) + \frac{1}{2} f''(0)\sigma^2 + o(1)\sigma^2$$
$$\text{Var}(f(\sigma G)) = f'(0)^2\sigma^2 + o(1)\sigma^2$$

$o_P(1)$ denotes a sequence of uniformly bounded random variables converging a.s. to 0 as $\sigma \to 0$.

As $\sigma \to 0$,

$$\frac{1}{\sigma^2}\Big\{\ f(0) - E\big(f(\sigma G)\big)\Big\}^2 \to 0, \quad \frac{1}{\sigma^2}\Big\{f'(0) - E\big(f'(\sigma G)\big)\Big\}^2 \to 0$$

By Cauchy Schwartz

$$\frac{1}{\sigma}E_P\left(\left\{f(\sigma G) - E_P(f(\sigma G)) - E_P(f'(\sigma G))\sigma G\right\}^2\right)^{1/2}$$
$$\leq \frac{1}{\sigma}E_P(\{f(\sigma G) - f(0) - f'(0)\sigma G\}^2)^{1/2}$$
$$+ \frac{\big|f(0) - E_P(f(\sigma)G)\big|}{\sigma} + \frac{\big|f'(0) - E_P(f(\sigma)G)\big|}{\sigma}\sigma E_P(G^2)^{1/2} \longrightarrow 0$$

When $G$ is standard Gaussian, integrating by parts

$$\frac{1}{\sigma^2} E_P\left(\left\{ f(\sigma G) - E_P(f(\sigma G)) - \frac{E_P(f(\sigma G)\sigma G)}{\sigma^2}\sigma G \right\}^2\right) \longrightarrow 0$$

as $\sigma \to 0$, in colour you have the linear regression coefficents.

# Multivariate case

Let $\Delta W_1, \dots, \Delta W_n$ i.i.d. Gaussian with $E(\Delta W_1) = 0$
$E(\Delta W_1^2) = \Delta T = T/n$.
These are consecutive increments of the random walk
$W_m = \sum_{k=1}^{m} \Delta W_k$.
Let

$$F(\omega) = f(\Delta W_1(\omega), \dots, \Delta W_n(\omega))$$

with $f(x_1, \dots, x_n) \in W^{1,2}(\mathbb{R}^n, \gamma^{\otimes n})$.
Introduce the $\sigma$-algebrae $\mathcal{F}_k = \sigma(\Delta W_1, \dots, \Delta W_k)$,
$k = 1, \dots, n$.

**Lemma**

*We have the martingale representation*

$$F = E_P(F) + \sum_{k=1}^{n} E(\partial_k f(\Delta W_1, \ldots, \Delta W_n) | \mathcal{F}_{k-1}) \Delta W_k + M_n$$

*where $M$ is a $(\mathcal{F}_k)$-martingale with $M_0 = 0$ and $\langle M, W \rangle = 0$.*

By induction it is enough to show that

$$E(F|\mathcal{F}_k) = E(F|\mathcal{F}_{k-1}) + E(\partial_k f(\Delta W_1, \dots, \Delta W_n)|\mathcal{F}_{k-1})\Delta W_k + \Delta M_k$$

with

$$E(\Delta M_k|\mathcal{F}_{k-1}) = 0, \; E(\Delta W_k \Delta M_k|\mathcal{F}_{k-1}) = 0 \tag{0.4}$$

Note that from independence,

$$E_P\left(\partial_k f(\Delta W_1, \dots, \Delta W_n)\middle|\mathcal{F}_{k-1}\right)(\omega) = \int_{\mathbb{R}^{n-k+1}} \partial_k f(\Delta W_1(\omega), \dots, \Delta W_{k-1}(\omega), x_k, \dots, x_n)\gamma^{\otimes(n-k+1)}(x)dx$$

Let's fix $k$ and consider the enlarged $\sigma$-algebra

$$\mathcal{G}_{k-1} = \sigma(\Delta W_1, \dots, \Delta W_{k-1}, \Delta W_{k+1}, \dots \Delta W_n) \supseteq \mathcal{F}_{k-1}$$

By fixing $(\Delta W_i, i \neq k)$, applying the 1-dimensional result to the $k$-the coordinate $\Delta W_k$

$$F = E(F|\mathcal{G}_{k-1}) + E(\partial_k(\Delta W_1, \dots \Delta W_k)|\mathcal{G}_{k-1})\Delta W_k + \Delta \widetilde{M}_k$$

By the independence of the increments

$$\begin{aligned}
& f(\Delta W_1, \dots \Delta W_n) \\
& = E\left( f(x_1, \dots, x_{k-1}, \Delta W_k, x_{k+1}, \dots x_n) \right)\Big|_{x_i = \Delta W_i,\ i \neq k} \\
& + E\left( \partial_k f(x_1, \dots, x_{k-1}, \Delta W_k, x_{k+1}, \dots x_n) \right)\Big|_{x_i = \Delta W_i,\ i \neq k} \Delta W_k \\
& + \Delta \widetilde{M}_k
\end{aligned}$$

By fixing $(\Delta W_i, i \neq k)$, applying the 1-dimensional result to the $k$-the coordinate $\Delta W_k$

$$F = E(F|\mathcal{G}_{k-1}) + E(\partial_k(\Delta W_1, \dots \Delta W_k)|\mathcal{G}_{k-1})\Delta W_k + \Delta \widetilde{M}_k$$

By the independence of the increments

$$\begin{aligned} & f(\Delta W_1, \dots \Delta W_n) \\ & = E\left( f(x_1, \dots, x_{k-1}, \Delta W_k, x_{k+1}, \dots x_n) \right)\Big|_{x_i = \Delta W_i,\ i \neq k} \\ & + E\left( \partial_k f(x_1, \dots, x_{k-1}, \Delta W_k, x_{k+1}, \dots x_n) \right)\Big|_{x_i = \Delta W_i,\ i \neq k} \Delta W_k \\ & + \Delta \widetilde{M}_k \end{aligned}$$

with

$$E(\Delta \widetilde{M}_k|\mathcal{G}_{k-1}) = 0, \quad E(\Delta \widetilde{M}_k \Delta W_k|\mathcal{G}_{k-1}) = 0,$$

which implies

$$E(\Delta \widetilde{M}_k|\mathcal{F}_{k-1}) = 0, \quad E(\Delta \widetilde{M}_k \Delta W_k|\mathcal{F}_{k-1}) = 0.$$

By taking conditional expectation w.r.t. $\mathcal{F}_k$ and using independence of increments

$$E(F|\mathcal{F}_k) =$$
$$E\left( f(x_1,\dots,x_{k-1},\Delta W_k,\Delta W_{k+1},\dots \Delta W_n)\right)\Big|_{x_i=\Delta W_i,\ i<k} +$$
$$E\left( \partial_k f(x_1,\dots,x_{k-1},\Delta W_k,\Delta W_{k+1},\dots \Delta W_n)\right)\Big|_{x_i=\Delta W_i,\ i<k} \Delta W_k$$
$$+\Delta M_k$$

where

$$\Delta M_k := E(\Delta \widetilde{M}_k|\mathcal{F}_k)$$

with

$$E(\Delta M_k|\mathcal{F}_{k-1}) = E(\Delta M_k \Delta W_k|\mathcal{F}_{k-1}) = 0 \quad \square$$

Assuming that the derivatives $\partial_k f, \partial^2_{kk} f$ are bounded and continuous, by Jensen's inequality for conditional expectation and lemma 3

$$E_P\Big((\Delta M_k)^2 \Big| \mathcal{F}_{k-1}\Big)(\omega) \le E_P\Big((\Delta \widetilde{M}_k)^2 \Big| \mathcal{F}_{k-1}\Big)(\omega) = o_P(1)\Delta t$$

where $o_P(1) \to 0$ a.s with bounded convergence as $\Delta t \to 0$ uniformly over $t$.

By the martingale property, when $\Delta t = T/n$ for $T$ fixed and $n \to \infty$

$$E_P\left(\left\{ \sum_{k=1}^{n} \Delta M_t \right\}^2\right) = \sum_{k=1}^{n} E_P\left(\left\{\Delta M_t\right\}^2\right) \leq o_P(1) T$$

**Definition**

*The (finite-dimensional) Malliavin derivative is the random gradient*

$$DF := \nabla f(\Delta W_1, \ldots, \Delta W_n) \in \mathbb{R}^n$$

## Definition

Brownian motion, $(W_t : t \in [0,T])$ is a gaussian process with $W_0 = 0$ and such that for every $n$, $0 = t_0 \le t_1 \le \cdots \le t_n = T$ the increments $(W_{t_i} - W_{t_{i-1}})$ are independent and gaussian with variances $(t_i - t_{i-1})$.

Brownian motion can be constructed as a random continuous function $[0,t] :\to \mathbb{R}$.

Suppose that we have a random variable $F(\omega)$ which is measurable with respect to the Brownian $\sigma$-algebra $\mathcal{F}_t^W = \sigma(W_s : 0 \le s \le t)$. When $E_P(F^2) < \infty$, by Doob's martingale convergence theorem and that this can be approximated a.s. and in $L^2(\Omega)$ by random variables of the form

$$F_n(\omega) := f_n\big(W_{t_1^{(n)}} - W_{t_0^{(n)}}, \ldots, W_{t_n^{(n)}} - W_{t_{n-1}^{(n)}}\big)$$

where $f_n(x_1, \ldots, x_n)$ is Borel measurable and $t_k^{(n)} := Tk/n$.

When $F(\omega)$ is Malliavin differentiable, $f_n(x_1, \dots, x_n)$ are smooth functions. As $n \to \infty$ the orthogonal linearization error in

$$F_n = E(F_n) + \sum_{k=1}^{n} E(\nabla_k F_n | \mathcal{F}_{k-1}^{(n)}) \Delta W_k^{(n)} + M_n^n$$

vanishes (in $L^2(P)$ sense ) and the limit is the Ito-Clark-Ocone martingale representation

$$F = E(F) + \int_0^T E(D_s F | \mathcal{F}_s^W) dW_s$$

where the Ito integral appears.

# Skorokhod integral

In $L^2(\mathbb{R}^n, \gamma^{\otimes n}(x)dx)$ the Malliavin derivative of $F = f(\Delta W_1, \dots, \Delta W_n)$ as the random gradient $DF = \nabla f(\Delta W_1, \dots, \Delta W_n)$, where $\Delta W_k$ are i.i.d. $\mathcal{N}(0, \Delta t)$
Let $u_k = u_k(\Delta W_1, \dots, \Delta W_n)$ for $k = 1, \dots, n$.
Let's introduce the scalar product

$$\langle u, v \rangle := \Delta t \sum_{k=1}^{n} u_k v_k$$

We give the $n$-dimensional generalization of the 1-dimensional integration by parts formula.
We need a random variable which we denote by $\delta(u)$ (the Skorokhod integral or divergence integral ) such that

$$E_P\big(\langle DF, u\rangle\big) = E_P\big(F\delta(u)\big)$$

for all smooth random variables $F$. This extends the one-dimensional Gaussian integration by parts formula

$$E_P\big(\partial f(G)h(G)\rangle\big) = E_P\big(f(G)\partial^*(G)\big)$$

Rewrite the left hand side

$$\Delta t \sum_{k=1}^{n} E\Big(u_k(\Delta W_1, \dots, \Delta W_n)\partial_k f(\Delta W_1, \dots, \Delta W_n)\Big)$$

by independence and the 1-dimensional gaussian integration by parts

$$= \Delta t \sum_{k=1}^{n} E(\partial_k^* u_k(\Delta W_1, \dots, \Delta W_n) f(\Delta W_1, \dots, \Delta W_n))$$

$$= E\left(F \Delta t \left(\sum_{k=1}^{n} \frac{u_k \Delta W_k}{\Delta t} - \sum_{k=1}^{n} \partial_k u_k\right)\right)$$

Rewrite the left hand side

$$\Delta t \sum_{k=1}^{n} E\big(u_k(\Delta W_1, \dots, \Delta W_n)\partial_k f(\Delta W_1, \dots, \Delta W_n)\big)$$

by independence and the 1-dimensional gaussian integration by parts

$$= \Delta t \sum_{k=1}^{n} E(\partial_k^* u_k(\Delta W_1, \dots, \Delta W_n) f(\Delta W_1, \dots, \Delta W_n))$$

$$= E\left( F \Delta t \left( \sum_{k=1}^{n} \frac{u_k \Delta W_k}{\Delta t} - \sum_{k=1}^{n} \partial_k u_k \right)\right)$$

so that

$$\delta(u) = \sum_{k=1}^{n} u_k \Delta W_k - \sum_{k=1}^{n} D_k u_k \Delta t$$

The first term is a Riemann sum, while the second term is called Malliavin trace.

When $u_k = u_k(\Delta W_1, \ldots, \Delta W_{k-1}, \Delta W_{k+1}, \ldots, \Delta W_n)$ does not depend on $\Delta W_k$ , the Malliavin trace vanishes.

When $u_k = u_k(\Delta W_1, \dots, \Delta W_{k-1}, \Delta W_{k+1}, \dots, \Delta W_n)$ does not depend on $\Delta W_k$ , the Malliavin trace vanishes.
For $F \equiv 1$, $DF \equiv 0$ and when exists $\delta(u) \in L^2(\Omega)$, necessarily

$$E(\delta(u)) = E(\langle u, 0 \rangle) = 0$$

When $u_k = u_k(\Delta W_1, \dots, \Delta W_{k-1}, \Delta W_{k+1}, \dots, \Delta W_n)$ does not depend on $\Delta W_k$ , the Malliavin trace vanishes.
For $F \equiv 1$, $DF \equiv 0$ and when exists $\delta(u) \in L^2(\Omega)$, necessarily

$$E(\delta(u)) = E(\langle u, 0 \rangle) = 0$$

In the continuous time case the Skorokhod integral with respect to the Brownian motion is given by

$$\delta(u) := \int_0^T u_s \delta W_s = \int_0^T u_s dW_s - \int_0^T D_s u_s ds$$

where $\int_0^T u_s dW_s$ is a *forward integral* defined as the limit in probability or $L^2(P)$-sense of the Riemann sums, and the last term is the Malliavin trace.
When $u$ is adapted, that is $u$ is $\mathcal{F}_s^W$-measurable for all $s$ the Malliavin trace vanishes and the Skorokhod integral coincides with the Ito integral.

Note that if $\varphi$ is smooth, $D\varphi(F) = \varphi'(F)DF$. We have also the product rule $D(FG) = G\ DF + F\ DG$.
Consider a process $u_k = u_k(\Delta W_1, \ldots, \Delta W_n)$

$$E\left(\delta\left(\frac{u}{\langle u, DF\rangle}\right)\varphi(F)\right) = E\left(\left\langle \frac{u}{\langle u, DF\rangle}, D\varphi(F)\right\rangle\right)$$
$$= E\left(\frac{\varphi'(F)}{\langle u, DF\rangle}\langle u, DF\rangle\right) = E(\varphi'(F))$$

This holds for all choices of $(u_k)$ and $\varphi$. By taking $u = DF$ we obtain

$$E(\varphi'(F)) = E\left(\varphi(F)\delta\left(\frac{DF}{\| DF \|^2}\right)\right)$$

where

$$\| DF \|^2 = \langle DF, DF\rangle = \Delta t \sum_k^n (D_k F)^2$$

# Computation of densities

Let $F = f(\Delta W_1, \dots, \Delta W_n)$ a random variable with Malliavin Sobolev derivative. For $a < b \in \mathbb{R}$ consider

$$\psi(x) = \int_a^b \mathbf{1}(r \le x) dr$$

with Sobolev derivative $\psi'(x) = \mathbf{1}_{[a,b]}(x)$.

$$P(a < F \le b) = \int_a^b p_F(r) dr \text{ (when } F \text{ has density )}$$

$$= E_P\big(\mathbf{1}(a < F \le b)\big) = E_P\big(\psi'(F)\big) = E_P\left(\psi(F)\delta\left(\frac{DF}{\| DF \|^2}\right)\right)$$

$$= E_P\left(\delta\left(\frac{DF}{\| DF \|^2}\right) \int_a^b \mathbf{1}(r \le F) dr\right) = \text{ (Fubini)}$$

$$= \int_a^b E_P\left(\mathbf{1}(r \le F)\delta\left(\frac{DF}{\| DF \|^2}\right)\right) dr$$

This implies

$$p_F(r) = E_P\left(\mathbf{1}(r \le F)\delta\left(\frac{DF}{\| DF \|^2}\right)\right) = E_P(\mathbf{1}(r \le F)Y)$$

with Malliavin weight

$$Y := \delta\left(\frac{DF}{\| DF \|^2}\right) =$$

$$\frac{1}{\| DF \|^2}\sum_{k=1}^{n} D_k F \Delta W_k - \sum_{k=1}^{n} D_k\left(\frac{D_k F}{\| DF \|^2}\right)\Delta t$$

$$= \frac{1}{\| DF \|^2}\sum_{k=1}^{n} D_k F \Delta W_k - \frac{1}{\| DF \|^2}\sum_{k=1}^{n} D^2_{kk} F \Delta t$$

$$+ \frac{2}{\| DF \|^4}\sum_{k=1}^{n}\sum_{h=1}^{n} D_k F \, D_h F \, D^2_{kh} F \, \Delta t \, \Delta t$$

For $F = f(\Delta W_1, \dots, \Delta W_n)$ we need that $f$ twice differentiable in Sobolev sense and integrability conditions. This extends to the infinite-dimensional case when $F$ is a smooth functional of the Brownian path.

For $F = f(\Delta W_1, \ldots, \Delta W_n)$ we need that $f$ twice differentiable in Sobolev sense and integrability conditions. This extends to the infinite-dimensional case when $F$ is a smooth functional of the Brownian path.
For $i \in \mathbb{N}$ let $(\Delta W_1^{(i)}, \ldots, \Delta W_n^{(i)})$, i.i.d copies of the gaussian vector, let

$$F^{(i)} := f(\Delta W_1^{(i)}, \ldots, \Delta W_n^{(i)}),$$
$$Y^{(i)} := Y(\Delta W_1^{(i)}, \ldots, \Delta W_n^{(i)})$$

We estimate $p_F(t)$ by Monte Carlo

$$\widehat{p}_F^{(M)}(r) = \frac{1}{M} \sum_{i=1}^{M} Y^{(i)} \mathbf{1}(F^{(i)} \geq r)$$

There are other choices for the Malliavin weight: for

$$u_k = \frac{1}{n\Delta t\ D_k F}$$

we obtain

$$\begin{aligned} E(\langle u, D\varphi(F)\rangle) &= \frac{1}{n\Delta t} E\big(\varphi'(F)\langle DF, (DF)^{-1}\rangle\big) = \\ &= \frac{1}{n\Delta t} E\left(\varphi'(F)\sum_{k=1}^{n}(D_k F)^{-1} D_k F \Delta t\right) \\ &= E(\varphi'(F)) = E\big(\varphi(F)U\big) \end{aligned}$$

with Malliavin weight

$$U = \frac{1}{n\Delta t}\delta\big((DF)^{-1}\big) = \frac{1}{n\Delta t}\sum_{k=1}^{n}\frac{1}{D_k F}\Delta W_k + \frac{1}{n}\sum_{k=1}^{n}\frac{D^2_{kk}F}{(D_k F)^2}$$

# Example: quadratic functional

Let

$$W_k = (\Delta W_1 + \cdots + \Delta W_k), \quad F = \sum_{k=1}^{n} W_k^2 \Delta t$$

$$D_h F = 2 \sum_{k=h}^{n} W_k \Delta t, \quad D^2_{h,k} F = 2(n - (h \vee k) + 1)\Delta t$$

We compute the Malliavin weight $U$

$$U = \frac{1}{n\Delta t}\left(\sum_{h=1}^{n} \frac{1}{D_h F} dW_h - \sum_{h=1}^{n} D_h\big((D_h F)^{-1}\big)\Delta t\right)$$

$$= \frac{1}{2n\Delta t}\sum_{h=1}^{n}\left(\sum_{k=h}^{n} W_k \Delta t\right)^{-1} \Delta W_h$$

$$+\frac{1}{n\Delta t}\sum_{h=1}^{n}\left(2\sum_{k=h}^{n} W_k \Delta t\right)^{-2} 2(n-h+1)(\Delta t)^2 =$$

$$\frac{1}{2n(\Delta t)^2}\left\{\sum_{h=1}^{n}\left(\sum_{k=h}^{n} W_k\right)^{-1} \Delta W_h + \right.$$

$$\left. + \sum_{h=1}^{n}\left(\sum_{k=h}^{n} W_k\right)^{-2} (n-h+1)\Delta t\right\}$$

# Counterexample: Maximum of gaussian random walk

Let $W_0 = 0$, $W_m = \sum_{k=1}^{m} \Delta W_k$ for $m = 1, \dots, n$ the gaussian random walk, and let

$$F = W_n^* := \max_{m=0,1,\dots,n} \{W_m\} = f(\Delta W_1, \dots, \Delta W_n)$$

Let

$$\tau_n = \tau_n(W_1, \dots, W_n) = \arg \max_{m=0,1,\dots,n} W_m$$

the random time where the maximum is achieved. Note that with positive probability $W_n^* = 0$ and $\tau_n = 0$ when the random walk stays on the negative side, so we know that there is point mass at 0, $W_n^*$ does not have a density.

Clearly for $k = 1, \dots, n$

$$D_k W_n^* = \partial_k f_n(\Delta W_1, \dots, \Delta W_n) = \mathbf{1}\big(\tau_n \geq k\big)$$
$$= \mathbf{1}\big(W_{k-1}^* < \max_{h=k,\dots,n} W_h\big) \qquad \text{a.s.}$$

The problem is that the indicator of a set is never Malliavin differentiable and the second order Malliavin derivative $D_{hk}^2 X_n^* = D_h \mathbf{1}\big(\tau_n \geq k\big)$ doesn't exist as random variables in $L^2$ and the Malliavin weights are not well defined.

# Skorohod integral with correlated Gaussian noise

Consider correlated Gaussian increments, with density

$$\gamma_K(\Delta Z_1, \dots, \Delta Z_n) = |K|^{-1/2} \pi^{-n/2} \exp\left( -\frac{1}{2} \Delta Z \, K^{-1} \Delta Z^\top \right)$$

with $E(\Delta Z_\ell) = 0$ and $K_{h\ell} = E\left( \Delta Z_h \Delta Z_\ell \right)$.

# Skorohod integral with correlated Gaussian noise

Consider correlated Gaussian increments, with density

$$\gamma_K(\Delta Z_1, \dots, \Delta Z_n) = |K|^{-1/2} \pi^{-n/2} \exp\left( -\frac{1}{2} \Delta Z \, K^{-1} \Delta Z^\top \right)$$

with $E(\Delta Z_\ell) = 0$ and $K_{h\ell} = E\left( \Delta Z_h \Delta Z_\ell \right)$.

In the correlated case, Gaussian integration by parts reads as

$$E_P\big(\partial_\ell f(\Delta Z_1, \dots, \Delta Z_n) g(\Delta Z_1, \dots, \Delta Z_n)\big) =$$
$$E_P\Bigg( f(\Delta Z_1, \dots, \Delta Z_n) \times$$
$$\Big\{ g(\Delta Z_1, \dots, \Delta Z_n) \sum_h K^{-1}_{h\ell} \Delta Z_h - \partial_\ell g(\Delta Z_1, \dots, \Delta Z_n) \Big\} \Bigg)$$

If $u_k = u_k(\Delta Z_1, \ldots, \Delta Z_n)$, we define the Skorokhod integral w.r.t. $Z_n$ as $\delta_Z(u)$ satisfying

$$E_P\big(\langle DF, u\rangle_K\big) = E_P\big(F\delta_Z(u)\big)$$

with the scalar product

$$\langle x, y\rangle_K = x\ K\ y^\top$$

for all random variables
$F(\omega) = f(\Delta Z_1, \ldots, \Delta Z_n) \in W^{1,2}(\mathbb{R}^n, \gamma_K)$, This gives

$$\delta_Z(u) = \sum_{h=1}^{n} u_k \Delta Z_k - \sum_{h=1}^{n}\sum_{\ell=1}^{n} K_{h\ell} D_h u_\ell$$

In the continuous case this gives

$$\delta_Z(u) := \int_0^T u_s \delta Z_s = \int_0^T u_s dZ_s - \int_0^T \int_0^T D_t u_s K(dt, ds)$$

where the first integral exists as the limit of Riemann sums in $L^2(P)$,

# Hermite polynomials

Let $\gamma(x)$ be the standard gaussian density in $\mathbb{R}$.

**Lemma**

*The polynomials are dense in $L^2(\mathbb{R}, \gamma)$.*

**Proof** Otherwise there is a random variable $F = f(G) \in L^2(P)$ with $E(f(G)G^n) = 0 \ \forall n \in \mathbb{N}$ where $G$ is standard gaussian. Consider the (signed) measure on $\mathbb{R}$

$$\mu(A) := E_P\big(f(G)\mathbf{1}_A(G)\big)$$

We show that $\mu \equiv 0$ which implies $f(G) = 0$ $P$ a.s.
The Fourier transform of $\mu$ is

$$\widehat{\mu}(t) := E_P(f(G)\exp(itG))$$

For $t = (\sigma + \tau i) \in \mathbb{C}$ with $\sigma, \tau \in \mathbb{R}$,

$$\widehat{\mu}(t) := E_P\big(f(G)\exp(i\sigma G)\exp(-\tau G)\big)$$

Since

$$\begin{aligned} & E_P\left(\left|\frac{\partial}{\partial\sigma}\left\{f(G)\exp(-\tau G)\exp(i\sigma G)\right\}\right|\right) \\ &= E_P\big(\big|f(G)\exp(-\tau G)iG\exp(i\sigma G)\big|\big) \\ &\le E_P\big(|f(G)G\exp(-\tau G)|\big) \\ &\le E_P\big(|f(G)G(\exp(-aG)+\exp(-bG))|\big) \end{aligned}$$

where $\exp(-\tau G) \le \exp(-aG) + \exp(-bG)$ $\forall \tau \in (a, b) \subseteq \mathbb{R}$.

By Cauchy-Schwartz inequality

$$\leq E_P(f(G)^2)^{1/2} E(G^2\{\exp(-aG)+\exp(-bG))^2)^{1/2}$$
$$= E_P(f(G)^2)^{1/2} \Big\{ E(G^2\exp(-2aG)) +$$
$$+ E(G^2\exp(-2bG)) + 2E(G^2\exp(-(a+b)G)) \Big\}^{1/2} < \infty$$

by Lebesgue's dominated convergence theorem we can change the order of derivation and integration (Theorem A 16.1 in Williams' book)

$$\frac{\partial}{\partial\sigma}\widehat{\mu}(\tau+i\sigma) = i\ E_P\Big(f(G)G\exp(i\sigma G)\exp(-\tau G)\Big)$$

Similarly

$$\frac{\partial}{\partial\tau}\widehat{\mu}(\tau+i\sigma) = -E_P\Big(f(G)G\exp(i\sigma G)\exp(-\tau G)\Big) = i\ \frac{\partial}{\partial\sigma}\widehat{\mu}(\tau+i\sigma)$$

$\widehat{\mu}:\mathbb{C}\to\mathbb{C}$ is analytic since satisfies the Cauchy-Riemann condition.

Therefore has the power series expansion

$$\widehat{\mu}(t) = \sum_{t=0}^{\infty} \widehat{\mu}^{(n)}(0)\frac{t^n}{n!}$$

$$\mu^{(n)}(t) = \frac{d^n}{dt^n}\widehat{\mu}(t) = i^n E_P(f(G)\exp(itG)G^n),$$

$$\widehat{\mu}^{(n)}(0) = i^n E_P(f(G)G^n) = 0 \;\forall n \in \mathbb{N}$$

where by adapting the previous argument we can take derivatives inside the expectation. Therefore $\widehat{\mu}(t) = 0$ and by Lévy inversion theorem $\mu(dx) = 0$, which implies $E_P(f(G)^2) = 0 \;\square$.

# Hermite polynomials in $L^2(\mathbb{R}, \gamma)$.

Let $G$ be a standard gaussian random variable with density $\gamma(x)$.
Define the (unnormalized) Hermite polynomials

- $$h_0(x) \equiv 1, \;\; h_n(x) = (\partial^* h_{n-1})(x) = (\partial^{*n} 1)(x)$$

  By using repeatedly the commutation relation
  $$\partial \partial^* f - \partial^* \partial f = f$$
  we get
  $$\partial \partial^{*n} f - \partial^{*n} \partial f = n \partial^{*(n-1)} f$$
  when $f(x) = 1$

- $$\partial h_n(x) = n h_{n-1}(x)$$

  $\partial$ and $\partial^*$ are **annihilation** and **creation** operators.

- $$h_n(x) = \exp(x^2/2)\frac{d^n}{dx^n}\exp(-x^2/2)$$

  Ex: $h_1(x) = x,\ h_2(x) = (x^2 - 1),\ h_3(x) = (x^3 - 3x),$
  $h_4(x) = x^4 - 6x^2 + 3,\ h_5(x) = (x^5 - 10x^3 + 15x)$

- $$E_P\big(h_n(G)h_m(G)\big) = E_P\big((\partial^{*n}1)(G)(\partial^{*m}1)(G)\big) =$$
  $$E_P\big((\partial^n\partial^{*m}1)(G)\ \mathbf{1}\big) = \delta_{n,m}n!$$

  (assuming $n \geq m$)

Since the polynomials are dense in $L^2(\mathbb{R},\gamma)$, the normalized Hermite polynomials

$$H_n(x) := \frac{h_n(x)}{\sqrt{n!}} \quad n \in \mathbb{N}$$

form an orthonormal basis in $L^2(\mathbb{R},\gamma)$: for $f(G) \in L^2(P)$,

$$f(G) = \sum_{n=0}^{\infty} E_P(f(G)H_n(G))H_n(G) = \sum_{n=0}^{\infty} E_P(f(G)h_n(G))\frac{h_n(G)}{n!}$$

and when $f(x)$ is infinitely differentiable in Sobolev sense

$$= \sum_{n=0}^{\infty} E_P(f(G)(\partial^{*n}1)(G))\frac{h_n(G)}{n!} = \sum_{n=0}^{\infty} E_P(\partial^n f(G))\frac{h_n(G)}{n!}$$

(one-dimensional Stroock formula)

the convergence is in $L^2(P)$ sense

$$E_P\left(\left\{f(G)-\sum_{n=1}^{N}E_P(f(G)H_n(G))H_n(G)\right\}^2\right)\to 0 \text{ as } N\uparrow\infty$$

the convergence is in $L^2(P)$ sense

$$E_P\left(\left\{f(G)-\sum_{n=1}^{N}E_P(f(G)H_n(G))H_n(G)\right\}^2\right)\to 0 \quad \text{as } N\uparrow\infty$$

Define the generating function

$$f(t,x):=\exp\bigl(tx-t^2/2\bigr)=\frac{\gamma(x-t)}{\gamma(x)}=\frac{d\mathcal{N}(t,1)}{d\mathcal{N}(0,1)}(x)$$

which is the density ratio for the gaussian shift $G\to(t+G)$
Note that $E_P(f(t,G))=1$. Since $f(t,x)\in C^\infty$, by Stroock formula

$$\exp\bigl(tx-t^2/2\bigr)=\sum_{n=0}^{\infty}E_P\left(\frac{d^n}{dx^n}f(t,G)\right)\frac{h_n(x)}{n!}$$

$$=\sum_{n=0}^{\infty}E_P\bigl(t^nf(t,G)\bigr)\frac{h_n(x)}{n!}=\sum_{n=0}^{\infty}h_n(x)\frac{t^n}{n!}$$

Note that

$$t^n = E_P\left( h_n(G)\exp\big(tG - t^2/2\big)\right) = E_P\big(h_n(t+G)\big)$$

where on the right side we have changed the measure.

# Hermite polynomials in $L^2(\mathbb{R}^n, \gamma^{\otimes n})$.

Let $G = (G_1, \dots, G_n)$ a random vector with indepedent standard gaussian coordinates.
Since $L^2(\mathbb{R}^n, \gamma^{\otimes n}) = \overline{\text{span} L^2(\mathbb{R}, \gamma)^n}$, which is the $L^2$-closure of the linear space containing the products $f_1(x_1) f_2(x_1) \dots f_n(x_n)$ with $f_i \in L^2(\mathbb{R}, \gamma)$,
the polynomials in the variables $x_1, \dots, x_n$ are dense in $L^2(\mathbb{R}^n, \gamma^{\otimes n})$.

**Definition**

$\alpha = (\alpha_1, \dots, \alpha_n)$ *with* $\alpha_i \in \mathbb{N}$ *is a multi-index.*
$\alpha! := \prod_{i=1}^n \alpha_i!$

For $x = (x_1, \ldots, x_n)$ define the unnormalized and normalized multivariate Hermite polynomials

$$h_\alpha(x) = \prod_{i=1}^{n} h_{\alpha_i}(x_i)$$

$$H_\alpha(x) = \prod_{i=1}^{n} H_{\alpha_i}(x_i) = \prod_{i=1}^{n} \frac{h_{\alpha_i}(x)}{\sqrt{\alpha_i!}} = \frac{h_\alpha(x)}{\sqrt{\alpha!}}$$

**Lemma**

*$\{H_\alpha(x) : \alpha$ multi-index$\}$ is an orthonormal basis in $L^2(\mathbb{R}^n, \gamma^{\otimes n})$*

**Proof** Let $\beta = (\beta_1, \ldots, \beta_n)$ $\beta_i \in \mathbb{N}$,

$$E_P(H_\alpha(G)H_\beta(G)) = E_P\left(\prod_{i=1}^n H_{\alpha_i}(G_i)\prod_{j=1}^n H_{\beta_j}(G_j)\right) =$$

$$\prod_{i=1}^n E_P\big(H_{\alpha_i}(G_i)H_{\beta_i}(G_i)\big) = \prod_{i=1}^n \delta_{\alpha_i,\beta_i} = \delta_{\alpha,\beta}$$

$$F(\omega) = f(G_1, \dots, \Delta G_n) = \sum_{\alpha} E_P\big(H_\alpha(G)F\big) H_\alpha(G) = \sum_{\alpha} c_\alpha H_\alpha(G)$$

with $F \in L^2(\mathbb{R}^n \gamma^{\otimes n}) \iff \sum_{\alpha} c_\alpha^2 < \infty$

# Infinite dimensional gaussian space

$L^2(\mathbb{R}^{\mathbb{N}}, \gamma^{\otimes \mathbb{N}})$ is the space of sequences $x = (x_i : i \in \mathbb{N})$.
On this space we use the product $\sigma$-algebra $\mathcal{B}(\mathbb{R}^{\mathbb{N}}) = \mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}}$ which is the smallest $\sigma$-algebra such that the coordinate evaluations $x \mapsto x_i$ are measurable.
The Borel $\sigma$-algebra $\mathcal{B}(\mathbb{R})$ is the smallest $\sigma$ algebra containing the open sets.
The product measure $\gamma^{\otimes \mathbb{N}}$ is such that $\forall n \in \mathbb{N}$, $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$

$$\gamma^{\otimes \mathbb{N}}\big(\{x : x_1 \in B_1, \dots, x_n \in B_n\}\big) = \prod_{i=1}^{n} \gamma(B_i)$$

### Definition

*$\alpha = (\alpha_i : i \in \mathbb{N})$ with $\alpha_i \in \mathbb{N}$ and*

$$|\alpha| := \sum_{i=1}^{\infty} \alpha_i < \infty$$

*is a multi-index*

### Definition

*A polynomial in the variables $(x_i : i \in \mathbb{N})$ is given by*

$$p(x) = c_0 + \sum_{i=1}^{\infty} c_i x_i^{\alpha_i}$$

*$c_i \in \mathbb{R}$, and $\alpha$ is a multiindex, $|\alpha| < \infty$, which depends on finitely many coordinates.*

$$L^2(\mathbb{R}^{\mathbb{N}}, \gamma^{\otimes \mathbb{N}}) = \bigoplus_{n \in \mathbb{N}} L^2(\mathbb{R}^n, \gamma^{\otimes n})$$

An orthonormal basis is given by

$$\left\{ H_\alpha(G) := \prod_{i=1}^{\infty} H_{\alpha_i}(G_i), \; \alpha \text{ multindex }, |\alpha| < \infty \right\}$$

where $(G_i : i \in \mathbb{N})$ is the canonical sequence of independent standard gaussian r.v.

**Lemma**

*If $(\xi_n : n \in \mathbb{N})$ are Gaussian random variables with $\xi_n \sim \mathcal{N}(\mu_n, \sigma_n^2)$, and $\xi_n \xrightarrow{d} \xi$ (in distribution), then $\xi$ has Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$ with $\mu = \lim_n \mu_n$ and $\sigma^2 = \lim_n \sigma_n^2$.*
*When $\sigma^2 = 0$, we agree that the constant random variable $\mu$ is Gaussian.*

**Corollary**

*If $(\xi_n : n \in \mathbb{N})$ are Gaussian and $\xi_n \xrightarrow{P} \xi$ in probability, since Gaussian variables have all moments it follows $(\xi_n : n\mathbb{N})$ is bounded in $L^p$ $\forall p < \infty$. and we have convergence also in $L^p(\Omega)$.*

# Random variables with values on a separable Banach space

Let $(E, \|\cdot\|)$ be a *separable Banach space*, and $E^*$ is the topological dual.
By separability we mean that there is $\{e_n : n \in \mathbb{N}\}$ which is dense in $E$.
The elements of $E^*$ are linear continuous functionals $\varphi$ with $|\varphi(x)| \le C \ \| x \|_E$.
We denote also $\varphi(x) = \langle \varphi, x \rangle_{E^*,E}$.

## Example

The space $C([0,1],\mathbb{R})$ of continuous functions with the norm

$$\|f\|_\infty = \sup_{t\in[0,1]} |f(t)|$$

is separable: by Bernstein's theorem which says that continuous functions can be approximated by polynomials uniformly on compacts. To obtain a dense countable set we take the polynomial functions with rational coefficients.
Its dual is the space of signed measures with finite total variation on $[0,1]$.

The topological dual $E^*$ is equipped with the strong operator norm

$$|\varphi|_{E^*} = \sup\{|\varphi(x)| : x \in E, \|x\| = 1\} .$$

The topological dual $E^*$ is equipped with the strong operator norm

$$|\varphi|_{E^*} = \sup\{|\varphi(x)| : x \in E, \|x\| = 1\} .$$

By using the duality we define the weak topology on $E$, where $x_n \stackrel{w}{\to} x$ weakly if $\varphi(x_n) \to \varphi(x)$ $\forall \varphi \in E^*$.
We define also the weak-$*$ topology on $E^*$, where $\varphi_n \stackrel{w-*}{\to} \varphi$ $*$-weakly if $\varphi_n(x) \to \varphi(x)$ $\forall x \in E$.

**Example**

*The weak topology is weaker than the $\|.\|$ norm topology in $E$ and the weak-$*$ topology is weaker than the $|.|_{E^*}$ norm topology in $E^*$.*

We have a probability space $(\Omega, \mathcal{F}, P)$ and a random variable $X$ which is measurable from $(\Omega, \mathcal{F})$ into $(E, \mathcal{B}(E))$. where $\mathcal{B}(E)$ is the Borel $\sigma$-algebra generated by the open sets.

**Definition**

*A simple $E$-valued random variable has the form*

$$X(\omega) = \sum_{i=1}^{N} x_i \mathbf{1}(A_i), \qquad \text{with } x_i \in E, \; A_i \in \mathcal{F} \; .$$

## Lemma

*Let $X$ be random variable defined on a probability $(\Omega, \mathcal{F}, P)$ space with values in $(E, \mathcal{B}(E))$.*
*There exist a sequence of simple $E$-valued random variables $\{X_n : n \in \mathbb{N}\}$ such that*

$$\| X \| \geq \| X - X_n \| \downarrow 0 \quad \textit{(monotonically), almost surely .}$$

**Proof**: Choose $X_n(\omega)$ as the element of $\{e_1, \dots, e_n\}$ which is closest to $X(\omega)$.

We will use this corollary of the Hahn-Banach Theorem:

**Lemma**

*For every $x \in E$ $\exists$ $\varphi \in E^*$ with $\parallel \varphi \parallel_{E^*} = 1$ and $\parallel x \parallel_E = \varphi(x)$*

.

**Theorem**

*If $E$ is a separable Banach space the Borel $\sigma$-algebra is generated by the sets*

$$\{x \in E : \varphi(x) \leq \alpha\}$$

*with $\varphi \in E^*$ and $\alpha \in \mathbb{R}$.*

# $E^*$ as a space of random variables

Note that $\varphi(X(\omega))$ for $\varphi \in E^*$ and $\|X(\omega)\|$ are real valued random variables, i.e. measurable functions from $(\Omega, \mathcal{F})$ into $(E, \mathcal{B}(E))$, since they are composition of a continuous and a measurable function.
For a simple $E$-valued r.v. $X(\omega) = \sum_{i=1}^{N} x_i \mathbf{1}(A_i)$ , with $x_i \in E$, $A_i \in \mathcal{F}$ we define the integral

$$\int_{\Omega} X(\omega) P(d\omega) = \sum_{i=1}^{N} x_i P(A_i)$$

# Bochner integral

Assume that $X$ is a $E$-valued r.v. and that

$$\int_\Omega \| X(\omega) \| P(d\omega) < \infty$$

Since $E$ is separable, we can approximate $X$ by a sequence of simple $E$-valued r.v. $\{X_n\}$ with
$\|X\| \geq \| X_n - X \| \downarrow 0$ (monotonically).

$$\left\| \int_\Omega X_n dP - \int_B X_m dP \right\| \leq \int_\Omega \| X_n - X_m \| dP$$
$$\leq \int_\Omega \| X - X_m \| dP + \int_\Omega \| X - X_m \| dP \to 0$$

By the monotone convergence theorem it follows that $\{\int_\Omega X_n dP\}$ is a Cauchy sequence in $E$, therefore since the space is complete it has a limit in $E$. By the same argument the limit does not depend on the choice of the approximating sequence, so that the Bochner integral of the r.v. $X$ is well defined.
Note that if $X$ is a $E$-valued r.v., to every $\varphi \in E^*$ corresponds a real valued r.v. $\varphi(\omega) := \varphi(X(\omega))$. We identify the r.v. and the element of $E^*$.

## Lemma

*If $\varphi \in E^*$ and $X(\omega)$ is Bochner integrable on $E$ under $P$,*

$$\varphi\left(\int_\Omega X(\omega)P(d\omega)\right) = \int_\Omega \varphi(X(\omega))P(d\omega)$$

**Proof** Let $X_n$ a sequence of simple $E$-valued r.v. with $\| X \| \geq \| X - X_n \| \downarrow 0$. Since $\varphi$ is linear the lemma holds for simple random variables, and by continuity

$$\left|\varphi\left(\int_\Omega XdP\right) - \int_\Omega \varphi(X)dP\right| =$$

$$\leq \left|\varphi\left(\int_\Omega XdP\right) - \varphi\left(\int_\Omega X_n)dP\right) + \int_\Omega \varphi(X_n)dP - \int_\Omega \varphi(X)dP\right|$$

$$\|\varphi\|_{E^*}\left\| \int X_n dP - \int XdP\right\| + \left| \int \varphi(X_n)dP - \int \varphi(X)dP\right| \to 0$$

## Definition

*If $\mu$ is a probability distribution on the Banach space $E$ we define the characteristic function as*

$$\widehat{\mu}(\phi) := \int_E \exp(i\,\psi(x))\mu(dx)$$

*where $\varphi \in E^*$.*

## Definition

*A cylynder set is of the form*

$$\{x \in E : (\varphi_1(x), \dots, \varphi_n(x)) \in B\}$$

*where $B \in \mathcal{B}(\mathbb{R}^n)$, $n \in \mathbb{N}$, $\varphi_i \in E^*$.*

Follows from theorem 12 that for separable Banach spaces the cylinder sets generate the Borel $\sigma$-algebra of $E$. In particular two measures on $(E, \mathcal{B}(E))$ coincide if they coincide on the

# Gaussian random variables on $E$

## Definition

*A measure $\gamma$ on the Banach space $E$ is (centered) gaussian iff for every $\varphi \in E^*$ the real valued r.v. $\varphi(x)$ is (centered) gaussian.*

## Lemma

*If $X$ is a $E$-valued r.v. with gaussian distribution, then for every $n$, $\varphi_1, \dots, \varphi_n \in E^*$, then the random variables $(\varphi_1(X), \dots, \varphi_n(X))$ are jointly gaussian.*

**Proof** Use the finite dimensional gaussian characterization with the characteristic function together with the linearity

$$E_P\left(\exp\left(i \sum_{i=1}^{n} \theta_i \varphi_i(X)\right)\right) = E_P\left(\exp\left(i\varphi\left(\left\{\sum_{i=1}^{n} \theta_i\right\} X\right)\right)\right)$$

The family $\{\varphi(X) : \varphi \in E^*\}$ is a gaussian process indexed by $E^*$.
Since Gaussian r.v. have all moments,

**Lemma**

*The embedding of $E^*$ into $L^p(E, \mathcal{B}(E), \gamma)$, $0 < p < \infty$ is continuous w.r.t. the weak-$*$ topology of $E^*$, (and therefore also in the $|\cdot|_{E^*}$ topology).*

**Proof** Let $\varphi_n, \varphi \in E^*$ with $\varphi_n \stackrel{w-*}{\to} \varphi$ in the weak-$*$ topology, that is for every fixed $x \in E$ $\varphi_n(x) \to \varphi(x)$.

In particular $(\varphi_n - \varphi) \to 0$, $\gamma(dx)$ a.s. Since $(\varphi_n - \varphi)(x)$ are centered gaussiam random variables, it follows from **??** that $\text{Var}(\varphi_n - \varphi) \to 0$, and by using gaussianity that for $p < \infty$

$$E_\gamma((\varphi_n(X) - \varphi(X))^p) \leq c_p E_\gamma((\varphi_n(X) - \varphi(X))^2)^{p/2} \to 0$$

, that is $\varphi_n(X) \to \varphi(X)$ in $L^p(E, \mathcal{B}(E), \gamma)$. $\square$.

### Definition

*We denote by* $\bar{E}^*$ *the closure of* $E^*$ *in* $L^2(\gamma)$.

Note that if $\varphi \in \bar{E}^*$ there is a sequence $\varphi_n \to \varphi$ in $L^2(\gamma)$. In case $\varphi \in \bar{E}^* \setminus E^*$ , $\varphi(x)$ is not defined pointwise but as a random variable for $\gamma$-almost every $x \in E$.

Note that in the one dimensional situation, if $X$ is centered gaussian with variance $\sigma^2$, then clearly $E(\exp(\lambda X^2)) < \infty$ for $\lambda < \left(2\sigma^2\right)^{-1}$. For the infinite-dimensional case we prove that the r.v. $\| X \|^2$ has exponential moment for some $\lambda > 0$.

**Theorem**

*(Fernique lemma) Let $\gamma$ be a centered gaussian measure on $(E, \mathcal{B})$ . If $\lambda > 0$, $r > 0$ such that*

$$\log\left(\frac{1-\gamma(\bar{B}(0,r))}{\gamma(\bar{B}(0,r))}\right) + 32\lambda r^2 \leq -1 \quad ,$$

*then*

$$\int_E \exp(\lambda \| x \|^2)\gamma(dx) \leq \exp(16\lambda r^2) + \frac{e^2}{e^2-1}$$

Since the r.v. $\| X \|^2$ has exponential moment for some $\lambda > 0$, we have $E_\gamma(\| X \|^p) < \infty$ , for all $p > 0$.

# The Kernel

Let $\gamma$ be a centered gaussian measure on a separable Banach space $E$.

## Definition

*The operator* $K : E^* \longrightarrow E$,

$$K\varphi := \int_E x\varphi(x)\gamma(dx) \text{ as Bochner integral}$$

is called **Kernel** .

Note that $K\varphi$ is in $E$ since

$$\| K\varphi \|_E \leq \int_E \| x\varphi(x) \| \gamma(dx) \leq \int_E \| x \| |\varphi(x)|\gamma(dx)$$

$$\leq \left(\int_E \| x \|^2 \gamma(dx)\right)^{1/2} \left(\int_E |\varphi(x)|^2 \gamma(dx)\right)^{1/2} < \infty$$

by Fernique lemma and since $E^*$ is imbedded in $L^2(\gamma)$.

Note that if $\varphi, \psi \in E^*$

$$\langle \psi, K\varphi \rangle = \int_E \psi(x)\varphi(x)\gamma(dx) = E_\gamma\big(\psi(X), \varphi(X)\big) = \langle \varphi, K\psi \langle$$

This map extends to $\overline{E}^*$, the closure in $L^2(\gamma)$ of $E^*$.

We introduce the Cameron-Martin's space.

$$H = \left\{ h = K\varphi : \ \varphi \in E^* \right\} \subseteq E$$

This is called Kernel or Cameron Martin space. It is an Hilbert space equipped with the scalar product

$$(h_1, h_2)_H = \langle \varphi_1, K\varphi_2 \rangle = E_\gamma\big(\varphi_1(X)\varphi_2(X)\big)$$

The scalar product

$$(h, x)_H = \varphi(x)$$

makes sense also when $h = K\varphi$ with $\varphi \in \overline{E}^*$ as a random variable in $L^2(\gamma)$.
We also have the **reproducing kernel property**:

$$\int_E \langle h, x \rangle_H \langle g, x \rangle_H \gamma(dx) = \langle h, g \rangle_H$$

# the Cameron-Martin space of Brownian motion

$\{B_t : t \in [0,1]\}$.
Let $E = C_0([0,1],\mathbb{R}) = \{x \in C([0,1]) : x(0) = 0\}$, and $E^*$ consists of signed measures $\mu$ on $[0,1]$ with finite variation, with the duality

$$\langle \mu, x \rangle := \int_0^1 x(s)\mu(ds)$$

which is defined as an usual Riemann-Stiletjes integral, since $x(\cdot)$ is continuous and $\mu$ has finite variation. We have the continuity property

$$|\langle \mu, x \rangle| \leq \|x\|_\infty \int_0^1 |\mu(ds)| \ .$$

The covariance is $E(B_s B_t) = E(B_s^2) + E(B_s(B_t - B_s)) = s$ for $s \le t$, so we can write $K(s,t) = (s \wedge t)$. By changing the order of integration and then using integration by parts

$$(K\mu)(t) = \int_E x(t)\langle \mu, x\rangle \Gamma(dx) = \int_E x(t)\left(\int_0^1 x(s)\mu(ds)\right)\Gamma(dx)$$

$$\int_0^1 K(t,s)\mu(ds) = \int_0^1 (t \wedge s)\mu(ds) =$$

$$\mu([0,1])t - \int_0^t \mu([0,s])ds = \int_0^t \mu((s,1])ds$$

which is an absolutely continuous function, since the function $s \mapsto \mu((s,1])$ is bounded.

We have that

$$E(\langle \mu, B\rangle\langle \nu, B\rangle) = \nu K \mu = \int_0^1 \left(\int_0^t \mu((s,1])ds\right)\nu(dt)$$

$$= \int_0^1 \nu((t,1])\mu((t,1])dt := (K\mu, K\nu)_H$$

By completing $K(E^*)$ w.r.t. the scalar product $(\cdot,\cdot)_H$ we obtain the Cameron-Martin space of Brownian motion

$$H = W^{1,2}([0,1], dt) = \left\{ h \in C_0([0,1]) : h(t) = \int_0^t \dot{h}(s)ds \text{ with} \right.$$

$$(h,g)_H = \int_0^1 \dot{h}(s)\dot{g}(s)ds = (\dot{h}, \dot{g})_{L^2([0,1],dt)}, \quad \text{for } h, g \in H$$

Note that we can extend the scalar product $(h,x)_H$ to the case where $h \in H$ and $x \in E$.
For $\mu \in E^*$ and the Brownian path $x(t) = B_t(\omega)$ we obtain

$$\langle \mu, B \rangle = (K\mu, B)_H := \int_0^1 B(s)\mu(ds) = \int_0^1 \mu((s,1])dB_s$$

and this can be extended to any $h \in H$

$$(h, B)_H := \int_0^1 \dot{h}(s) dB_s \ ,$$

which is the **Wiener integral** .

The reproducing Kernel property of Brownian motion reads as

$$(h, g)_H = E_P\left( \int_0^1 \dot{h}(s) dB_s \int_0^1 \dot{g}(s) dB_s \right) = E_P\left( \int_0^1 \dot{h}(s)\dot{g}(s) ds \right)$$

Let's fix $t$ and take $g(s) = K(t, s) = t \wedge s = E\big(B_t B_s\big)$ with $\frac{\partial}{\partial s} K(t, s) = \mathbf{1}(s \le t)$. We obtain

$$(h, K(t, \cdot))_H = E_P\left( B_t \int_0^t \dot{h}(s) dB_s \right) = \int_0^t \dot{h}(s) ds = h(t)$$