Stein's method and Malliavin calculus

Ciprian A. Tudor

Université de Lille 1

International Colloquim on Stein's method, Concentration Inequalities and Malliavin calculus
Missillac, France June 2014

The purpose of the so-called Stein method is to measure the distance between two probability distributions.

This distance, denoted by d, can be defined in several ways :

- the Kolmogorov distance
- the Wasserstein distance
- the total variation distance
- the Fortet-Mourier distance.

Concretely, let X, Y be two random variables.

The distance between the law of X and the law of Y is usually defined by ($\mathcal{L}(F)$ denotes the law of F)

$$\mathbf{d}(\mathcal{L}(\mathbf{X}), \mathcal{L}(\mathbf{Y})) = \sup_{\mathbf{h} \in \mathcal{H}} |\mathbf{Eh}(\mathbf{X}) - \mathbf{Eh}(\mathbf{Y})|$$

where $\mathcal{H}$ is a suitable class of functions.

For example, if $\mathcal{H}$ is the set of indicator function

$$1_{(-\infty,\mathbf{z}]}, \mathbf{z} \in \mathbb{R}$$

we obtain the Kolmogorov distance

$$\mathbf{d_K}(\mathcal{L}(\mathbf{X}), \mathcal{L}(\mathbf{Y})) = \sup_{\mathbf{z}\in\mathbb{R}} |\mathbf{P}(\mathbf{X} \leq \mathbf{z}) - \mathbf{P}(\mathbf{Y} \leq \mathbf{z})|.$$

If $\mathcal{H}$ is the set of $1_{\mathbf{B}}$ with $\mathbf{B}$ a Borel set, one has the total variation distance

$$\mathbf{d_{TV}}(\mathcal{L}(\mathbf{X}), \mathcal{L}(\mathbf{Y})) = \sup_{\mathbf{B}\in\mathcal{B}(\mathbb{R})} |\mathbf{P}(\mathbf{X} \in \mathbf{B}) - \mathbf{P}(\mathbf{Y} \in \mathbf{B})|.$$

If

$$\mathcal{H} = \{\mathbf{h}; \|\mathbf{h}\|_{\mathbf{L}} \leq 1\}$$

($\|\cdot\|_{\mathbf{L}}$ is the Lipschitz norm) one has the Wasserstein distance.

Other examples of distances between the distributions of random variables exist, e.g. the Fortet-Mourier distance.

An important particular case : compute the distance between the law of an arbitrary r.v. F and the standard normal law

Useful for many applications.

For instance, in statistics, if an estimator is asymptotically normal, one needs to know how fast it converges to the normal distribution

Let Z a r.v. with law N(0,1). How to estimate

$$\mathbf{d(F,Z)} = \mathbf{sup}_{\mathbf{h}\in\mathcal{H}}|\mathbf{Eh(F)} - \mathbf{Eh(Z)}|$$

In particular, how to compute the Kolmogorov distance

$$\sup_{\mathbf{z}\in\mathbb{R}} |\mathbf{P}(\mathbf{X} \le \mathbf{z}) - \mathbf{P}(\mathbf{Y} \le \mathbf{z})|$$

The starting point to compute the distance between the law of F (an arbitrary r.v.) and the law of Z is the **Stein equation**

$$\mathbf{h}(\mathbf{x}) - \mathbf{Eh}(\mathbf{Z}) = \mathbf{f}'(\mathbf{x}) - \mathbf{xf}(\mathbf{x}).$$

- h is given
- one needs to find the function f which is the solution of the Stein equation.
- in the case of the Kolmogorov distance, $\mathbf{h}(\mathbf{x}) = 1_{(-\infty,\mathbf{z}]}(\mathbf{x})$. Need to find $\mathbf{f} = \mathbf{f_z}$ that satisfies the Stein's equation for every x .

(F is arbitrary, $\mathbf{Z} \sim \mathbf{N}(0,1)$) In the case of the Kolmogorov distance (in the Stein equation, put $x = F$ and then take the expectation)

$$\sup_{\mathbf{z} \in \mathbb{R}} \left| \mathbf{P}(\mathbf{F} < \mathbf{z}) - \mathbf{P}(\mathbf{Z} < \mathbf{z}) \right| = \mathbf{sup}_{\mathbf{z} \in \mathbb{R}} \left| \mathbf{E} \mathbf{f}'_{\mathbf{z}}(\mathbf{F}) - \mathbf{F} \mathbf{f}_{\mathbf{z}}(\mathbf{F}) \right|$$

where $\mathbf{f_z}$ is the solution of the Stein equation

$$1_{(-\infty,\mathbf{z})}(\mathbf{x}) - \mathbf{P}(\mathbf{Z} < \mathbf{z}) = \mathbf{f}'(\mathbf{x}) - \mathbf{f}(\mathbf{x}), \mathbf{x} \in \mathbb{R}$$

Key fact : the solution of the Stein equation is "nice" (for example its derivative is bounded)

Recall : we need to compute

$$\mathbf{E}\mathbf{f}'(\mathbf{F}) - \mathbf{E}\mathbf{F}\mathbf{f}(\mathbf{F})$$

Idea : use some integration by parts to write

$$\mathbf{E}\mathbf{F}\mathbf{f}(\mathbf{F}) = \mathbf{E}\mathbf{f}'(\mathbf{F})\mathbf{G}_{\mathbf{F}}$$

Then

$$\mathbf{E}\mathbf{f}'(\mathbf{F}) - \mathbf{E}\mathbf{F}\mathbf{f}(\mathbf{F}) = \mathbf{E}\mathbf{f}'(\mathbf{F})(1 - \mathbf{G}_{\mathbf{F}})$$

and use the fact that $\mathbf{f}'$ is "nice"

How to express $\mathbf{G_F}$?

The Malliavin calculus comes into the play ! The fundamental formula : if F is centered, then

$$\mathbf{F} = \delta \mathbf{D}(-\mathbf{L})^{-1}\mathbf{F}$$

where D is the Malliavin derivative, L the Ornstein-Uhlenbeck operator, δ the divergence (Skorohod) integral

How are these operators defined ?
Let 's undesrtand how they act on multiple stochastic integrals

Let $(\mathbf{W_t})_{\mathbf{t}\in[0,1]}$ a standard Wiener process and $\mathbf{I_n}$ the multiple integral of order $\mathbf{n}$ w.r.t. $\mathbf{W}$.
$\mathbf{I_n}$ is an isometry from $\mathbf{L}^2[0,1]^{\mathbf{n}}$ onto $\mathbf{L}^2(\Omega)$

$$\mathbf{E}\mathbf{I_n}(\mathbf{f})^2 = \mathbf{n}!\|\tilde{\mathbf{f}}\|^2_{\mathbf{L}^2[0,1]^{\mathbf{n}}}$$

where $\tilde{\mathbf{f}}$ is the symmetrization of $\mathbf{f}$

$\mathbf{I_n}(\mathbf{f})$ is also an iterated Itô integral
If $\mathbf{f}$ is symmetric,

$$\mathbf{I_n} = \mathbf{n}!\int_0^1 \mathbf{dW_{t_n}} \ldots\ldots\ldots \int_0^{\mathbf{t}_2} \mathbf{dW_{t_1}} \mathbf{f}(\mathbf{t}_1, \ldots, \mathbf{t_n})$$

Wiener chaos decomposition :
any random variable $\mathbf{F} \in \mathbf{L}^2(\Omega, \mathcal{F}, \mathbf{P})$ ($\mathcal{F}$ is the sigma-algebra generated by $\mathbf{W}$) can be written as

$$\mathbf{F} = \sum_{\mathbf{n} \geq 0} \mathbf{I_n}(\mathbf{f_n})$$

with $\mathbf{f_n} \in \mathbf{L^2_S}[0,1]^{\mathbf{n}}$ (uniquely determined by $\mathbf{F}$)

the subset of $\mathbf{L}^2(\Omega)$ generated by $\mathbf{I_n}(\mathbf{f}), \mathbf{f} \in\in \mathbf{L^2_S}[0,1]^{\mathbf{n}}$ is called the Wiener chaos of order $\mathbf{n}$

In Malliavin calculus, the multiple integrals are very useful (fit well with the theory)

The Malliavin operators on Wiener chaos :

$$\mathbf{D_s I_n(f) = n I_{n-1} f(\cdot, s)}$$

$$\mathbf{(-L)^{-1} I_n(f) = \frac{1}{n} I_n(f)}$$

$$\delta \mathbf{I_n f(\cdot, t) = I_{n+1}(\tilde{f})}.$$

Easy to see that

$$\mathbf{F} = \delta \mathbf{D}(-\mathbf{L})^{-1}\mathbf{F}$$

if $\mathbf{F = I_n(f)}$

Since

$$\mathbf{EFf(F)} = \mathbf{E}\delta\mathbf{D}(-\mathbf{L})^{-1}\mathbf{Ff(F)} = \mathbf{Ef'(F)}\langle\mathbf{D}(-\mathbf{L})^{-1}\mathbf{F}, \mathbf{DF}\rangle$$

so

$$\mathbf{Ef'(F)} - \mathbf{EFf(F)} = \mathbf{E}(\mathbf{f'(F)}(1 - \langle\mathbf{D}(-\mathbf{L})^{-1}\mathbf{F}, \mathbf{DF}\rangle)$$

Use Chauchy-Schwarz and remember that the derivative of the solution to the Stein equation is bounded.

we obtain

$$\sup_{z\in\mathbb{R}} |P(F < z) - P(Z < z)| \leq C \left(E \left(1 - \langle DF, D(-L)^{-1} F \rangle\right)^2 \right)^{\frac{1}{2}}$$

C is a constant that majorize the norm infinity of f ' (recall : f ' is bounded)

Note : the right-hand side does not depends on z !

Conclusion

$$\mathbf{d}(\mathbf{F},\mathbf{Z}) \leq \mathbf{C}\left(\mathbf{E}\left(1-\langle \mathbf{DF}, \mathbf{D}(-\mathbf{L})^{-1}\mathbf{F}\rangle\right)^{2}\right)^{\frac{1}{2}}$$

If we are able to compute

$$\left(\mathbf{E}\left(1-\langle \mathbf{DF}, \mathbf{D}(-\mathbf{L})^{-1}\mathbf{F}\rangle\right)^{2}\right)^{\frac{1}{2}}$$

one obtains and estimation for the distance netween the law of **F** and the standard normal law

How good is this bound ?
Suppose $\mathbf{F} = \mathbf{I_n}(\mathbf{f})$ is a multiple stochastic integral with respect to a Gaussian process. Then the right hand side of the above inequality gives

$$\mathbf{C}\sqrt{|3 - \mathbf{E}\mathbf{F}^4|}$$

The Fourth Moment Theorem states as follows.

Theorem

Fix $\mathbf{n} \geq 1$. *Consider a sequence* $\{\mathbf{F_k} = \mathbf{I_n}(\mathbf{f_k}), \mathbf{k} \geq 1\}$ *of square integrable random variables in the* $\mathbf{n}$*-th Wiener chaos. Assume that*

$$\lim_{\mathbf{k}\to\infty} \mathbf{E}[\mathbf{F}_\mathbf{k}^2] = \lim_{\mathbf{k}\to\infty} \|\mathbf{f_k}\|^2_{\mathbf{H}^{\odot \mathbf{n}}} = 1. \qquad (1)$$

Then, the following statements are equivalent.

1. *The sequence of random variables* $\{\mathbf{F_k} = \mathbf{I_n}(\mathbf{f_k}), \mathbf{k} \geq 1\}$ *converges to* $\mathbf{N}(0,1)$ *in distribution as* $\mathbf{k} \to \infty$.
2. $\lim_{\mathbf{k}\to\infty} \mathbf{E}[\mathbf{F}_\mathbf{k}^4] = 3$.
3. $\lim_{\mathbf{k}\to\infty} \|\mathbf{f_k} \otimes_\mathbf{l} \mathbf{f_k}\|_{\mathbf{H}^{\otimes 2(\mathbf{n}-\mathbf{l})}} = 0$ *for* $\mathbf{l} = 1, 2, \ldots, \mathbf{n}-1$.
4. $\|\mathbf{DF_k}\|^2_\mathbf{H}$ *converges to* $\mathbf{n}$ *in* $\mathbb{L}^2$ *as* $\mathbf{k} \to \infty$.

- the Stein bound is sharp for multiple integrals
- To check that a sequence of multiple integrals goes to $\mathbf{N}(0,1)$, it suffices to check that the second moment goes to 1 and the fourth moment goes to 3 !
- many applications to statistics, limit theorems etc.

Quadratic variations of a Gaussian process

Define the centered quadratic variation of the Gaussian process (U_t)

$$V_N := \sum_{i=0}^{N-1} \left[\left(U_{t_{i+1}} - U_{t_i} \right)^2 - E \left(U_{t_{i+1}} - U_{t_i} \right)^2 \right].$$

Purpose : find the limit in distribution, as $N \to \infty$, of the sequence V_N

Let $\mathbf{I_n}$ denote the multiple integral with respect to the Gaussian process $(\mathbf{U_t})$.

-multiple integrals can be defined with respect to any Gaussian process ; a multiple integral of order $\mathbf{n}$ is an isometry between $\mathcal{H}^{\odot \mathbf{n}}$ and $\mathbf{L}^2(\Omega)$ ($\mathcal{H}$ is the canonical Hilbert space associated to the Gaussian process)

Back to the process $\mathbf{U}$: we write $\mathbf{V_N}$ as a multiple integral. Since

$$\mathbf{U}_{\mathbf{t}_{\mathbf{i}+1}} - \mathbf{U}_{\mathbf{t}_{\mathbf{i}}} = \mathbf{I}_1\left(1_{(\mathbf{t}_{\mathbf{i}},\mathbf{t}_{\mathbf{i}+1})}\right)$$

and thanks to the product formula we can express the sequence $\mathbf{V_N}$ as a multiple integral of order 2 :

$$\mathbf{V_N} = \mathbf{I}_2\left(\sum_{\mathbf{i}=0}^{\mathbf{N}-1} 1_{(\mathbf{t}_{\mathbf{i}},\mathbf{t}_{\mathbf{i}+1})}^{\otimes 2}\right).$$

Steps : Find $\mathbf{a_N}$ such that

$$\mathbf{E}(\mathbf{a_N V_N})^2 \to_{\mathbf{N}} 1$$

(in the Fourth Moment Theorem we need the second moment to converge to 1)
Then : Let

$$\mathbf{F_N} = \mathbf{a_N V_N}$$

Prove that

$$\mathbf{E F_N^4} \to_{\mathbf{N}} 3$$

We obtain the convergence of $\mathbf{F_N}$ to the standard normal law.

Fractional Brownian motion (fBm)

Important particular case : the fractional Brownian motion (fBm)

The fBm is a centered Gaussian process $(\mathbf{B}^{\mathbf{H}}_{\mathbf{t}})_{\mathbf{t}\in[0,1]}$ with covariance

$$\mathbf{R}_{\mathbf{H}}(\mathbf{t},\mathbf{s}) = \mathbf{E}\mathbf{B}^{\mathbf{H}}_{\mathbf{t}}\mathbf{B}^{\mathbf{H}}_{\mathbf{s}} = \frac{1}{2}(\mathbf{t}^{2\mathbf{H}} + \mathbf{s}^{2\mathbf{H}} - |\mathbf{t}-\mathbf{s}|^{2\mathbf{H}}). \qquad (2)$$

It can be also defined as the only Gaussian process which is self-similar and has stationary increments.

Let $\mathbf{B}$ be a fBm with $\mathbf{H} \in (0,1)$. Define

$$\mathbf{V_N} := \sum_{\mathbf{i}=0}^{\mathbf{N}-1} \left[\left(\mathbf{B_{t_{i+1}}} - \mathbf{B_{t_i}} \right)^2 - \mathbf{E} \left(\mathbf{B_{t_{i+1}}} - \mathbf{B_{t_i}} \right)^2 \right]. \tag{3}$$

Then, if $\mathbf{H} < \frac{3}{4}$,

$$\mathbf{c} \mathbf{N}^{2\mathbf{H} - \frac{1}{2}} \mathbf{V_N} \to_{\mathbf{N}}^{\mathbf{d}} \mathbf{N}(0,1)$$

and if $\mathbf{H} > \frac{3}{4}$ then

$$\mathbf{N} \mathbf{V_N} \to_{\mathbf{N}} \textbf{Rosenblatt}.$$

The last convergence holds also in $\mathbf{L}^2$.

Purpose : generalize the Stein's bound to invariant measure of diffusions

Let $\mathbf{S}$ be the interval $(\mathbf{l}, \mathbf{u})$ $(-\infty \leq \mathbf{l} < \mathbf{u} \leq \infty)$ and μ be a probability measure on $\mathbf{S}$ with a density function $\mathbf{p}$ which is

- continuous
- bounded
- strictly positive on $\mathbf{S}$
- admits finite variance.

We want to find a diffusion process

$$\mathbf{dX_t} = \mathbf{b(X_t)dt} + \sqrt{\mathbf{a(X_t)}}\mathbf{dW_t}, \qquad \mathbf{t} \geq 0$$

which admits μ as invariant measure

-how to find the coefficients **a** and **b** ?

-the construction is not unique

The drift coefficient

Consider a continuous function $\mathbf{b}$ on $\mathbf{S}$ such that there exists $\mathbf{k} \in (\mathbf{l}, \mathbf{u})$ such that :

- $\mathbf{b}(\mathbf{x}) > 0$ for $\mathbf{x} \in (\mathbf{l}, \mathbf{k})$ and $\mathbf{b}(\mathbf{x}) < 0$ for $\mathbf{x} \in (\mathbf{k}, \mathbf{u})$
- $\mathbf{b} \in \mathbf{L}^1(\mu)$, $\mathbf{bp}$ is bounded on $\mathbf{S}$
- $$\int_{\mathbf{l}}^{\mathbf{u}} \mathbf{b}(\mathbf{x})\mathbf{p}(\mathbf{x})\mathbf{dx} = 0.$$

Basic example :

$$\mathbf{b}(\mathbf{x}) = -(\mathbf{x} - \mathrm{E}\mu)$$

The diffusion coefficient

Define

$$\mathbf{a}(\mathbf{x}) := \frac{2\int_{\mathbf{l}}^{\mathbf{x}} \mathbf{b}(\mathbf{y})\mathbf{p}(\mathbf{y})\mathbf{dy}}{\mathbf{p}(\mathbf{x})}, \quad \mathbf{x} \in \mathbf{S}. \tag{4}$$

Then, the stochastic differential equation :

$$\mathbf{dX_t} = \mathbf{b}(\mathbf{X_t})\mathbf{dt} + \sqrt{\mathbf{a}(\mathbf{X_t})}\mathbf{dW_t}, \quad \mathbf{t} \geq 0$$

has a unique Markovian weak solution, ergodic with invariant density $\mathbf{p}$.
Based on this fact, it is possible to define a so-called Stein's equation for a given function $\mathbf{f} \in \mathbf{L}^1(\mu)$.

For $\mathbf{f} \in \mathbf{L}^1(\mu)$, let $\mathbf{m_f} := \int_{\mathbf{l}}^{\mathbf{u}} \mathbf{f}(\mathbf{x})\mu(\mathbf{dx})$ and define $\tilde{\mathbf{g}}_{\mathbf{f}}$ by, for every $\mathbf{x} \in \mathbf{S}$,

$$\tilde{\mathbf{g}}_{\mathbf{f}}(\mathbf{x}) := \frac{2}{\mathbf{a}(\mathbf{x})\mathbf{p}(\mathbf{x})} \int_{\mathbf{l}}^{\mathbf{x}} (\mathbf{f}(\mathbf{y}) - \mathbf{m_f})\mathbf{p}(\mathbf{y})\mathbf{dy}.$$

We have

$$\tilde{\mathbf{g}}_{\mathbf{f}}(\mathbf{x}) = \int_{\mathbf{l}}^{\mathbf{x}} \frac{2(\mathbf{f}(\mathbf{y}) - \mathbf{m_f})}{\mathbf{a}(\mathbf{y})} \exp\left(- \int_{\mathbf{y}}^{\mathbf{x}} \frac{2\mathbf{b}(\mathbf{z})}{\mathbf{a}(\mathbf{z})}\mathbf{dz} \right) \mathbf{dy}, \qquad \mathbf{x} \in \mathbf{S}.$$

Then, $\mathbf{g_f(x)} := \int_0^{\mathbf{x}} \tilde{\mathbf{g}}_{\mathbf{f}}(\mathbf{y})\mathbf{dy}$ satisfies that $\mathbf{f} - \mathbf{m_f} = \mathbf{A}\mathbf{g_f}$ ($\mathbf{A}$ is the infinitesimal generator of the diffusion $(\mathbf{X_t})_{\mathbf{t}\geq 0},$) μ-almost everywhere and

$$\mathbf{f(x)} - \mathbf{E[f(X)]} = \frac{1}{2}\mathbf{a(x)}\tilde{\mathbf{g}}_{\mathbf{f}}'(\mathbf{x}) + \mathbf{b(x)}\tilde{\mathbf{g}}_{\mathbf{f}}(\mathbf{x}), \quad \mu\text{-a.e. } \mathbf{x} \qquad (5)$$

where $\mathbf{X}$ is a random variable with its law μ.

The equation (5) is a **generalized version of Stein's equation.**

Note : we constructed the equation **and** its solution !

Characterization of the law

Assume that $\int_{\mathbf{S}} \mathbf{a}(\mathbf{x})\mu(\mathbf{dx}) < \infty$. Let $\mathbf{Y}$ be a random variable on $\mathbf{S}$. Then, the distribution of $\mathbf{Y}$ coincides with μ if and only if

$$\mathbf{E}\left[\frac{1}{2}\mathbf{a}(\mathbf{Y})\mathbf{h}'(\mathbf{Y}) + \mathbf{b}(\mathbf{Y})\mathbf{h}(\mathbf{Y})\right] = 0$$

for $\mathbf{h} \in \mathbf{C}^1(\mathbf{S})$ such that $\mathbf{E}[|\mathbf{b}(\mathbf{Y})\mathbf{h}(\mathbf{Y})|] < \infty$ and $\mathbf{E}[|\mathbf{a}(\mathbf{Y})\mathbf{h}'(\mathbf{Y})|] < \infty$.

An alternative characterization of the random variables $\mathbf{Y}$ with distribution μ :

Consider a random variable $\mathbf{Y} \in \mathbb{D}^{1,2}$ with its values on $\mathbf{S}$ which satisfies that $\mathbf{b}(\mathbf{Y}) \in \mathbf{L}^2(\Omega)$. Then, $\mathbf{Y}$ has probability distribution μ if and only if $\mathbf{E}[\mathbf{b}(\mathbf{Y})] = 0$ and

$$\mathbf{E}\left[\frac{1}{2}\mathbf{a}(\mathbf{Y}) + \langle \mathbf{D}(-\mathbf{L})^{-1}\mathbf{b}(\mathbf{Y}), \mathbf{D}\mathbf{Y}\rangle_{\mathbf{H}} \middle| \mathbf{Y}\right] = 0.$$

The Stein bound obtained in previous work : if $\mathbf{Y}$ is a r.v. regular in the Malliavin calculus sense,

$$\mathbf{d}(\mathcal{L}(\mathbf{Y}),\mu) \leq \mathbf{CE}\left[\left|\frac{1}{2}\mathbf{a}(\mathbf{Y}) + \langle \mathbf{D}(-\mathbf{L})^{-1}\{\mathbf{b}(\mathbf{Y}) - \mathbf{E}[\mathbf{b}(\mathbf{Y})]\}, \mathbf{DY}\rangle_{\mathbf{H}}\right|\right] + \mathbf{C}|\mathbf{E}[\mathbf{b}(\mathbf{Y})]|, \quad \mathbf{Y} \in \mathbb{D}^{1,2}$$

where $\mathbf{C}$ is a positive constant and $\mathcal{L}(\mathbf{Y})$ is the law of $\mathbf{Y}$.

If $\mathbf{Y}$ is centered and $\mathbf{b}(\mathbf{x}) = -\mathbf{x}$ then

$$\mathbf{d}(\mathcal{L}(\mathbf{Y}), \mu) \leq \mathbf{CE}\left[\left|\frac{1}{2}\mathbf{a}(\mathbf{Y}) - \langle \mathbf{D}(-\mathbf{L})^{-1}\mathbf{Y}, \mathbf{DY}\rangle_{\mathbf{H}}\right|\right]$$

we just need to know what is $\mathbf{a}$

Using the conditional expectation :

Consider a random variable $\mathbf{Y} \in \mathbb{D}^{1,2}$ with its values on $\mathbf{S}$ which satisfies that $\mathbf{b}(\mathbf{Y}) \in \mathbf{L}^2(\Omega)$. Then, $\mathbf{Y}$ has probability distribution μ if and only if $\mathbf{E}[\mathbf{b}(\mathbf{Y})] = 0$ and

$$\mathbf{E}\left[\frac{1}{2}\mathbf{a}(\mathbf{Y}) + \langle \mathbf{D}(-\mathbf{L})^{-1}\mathbf{b}(\mathbf{Y}), \mathbf{D}\mathbf{Y}\rangle_{\mathbf{H}} \middle| \mathbf{Y}\right] = 0.$$

Example

The normal distribution $\mathbf{N}(0,\gamma)$, $\gamma > 0$. In this case $\mathbf{a}(\mathbf{x}) = 2\gamma$.

In this case

$$\mathbf{d}(\mathcal{L}(\mathbf{Y}),\mu) \leq \mathbf{CE}\left[\left|1 - \langle \mathbf{D}(-\mathbf{L})^{-1}\mathbf{Y}, \mathbf{DY}\rangle_{\mathbf{H}}\right|\right]$$

Example

The Gamma $\Gamma(a,\lambda)$, $a,\lambda>0$ law. Here the density is $f(x)=\frac{\lambda^a}{\Gamma(a)}x^{a-1}e^{-\lambda x}$ for $x>0$ and $f(x)=0$ otherwise. Also $EX=\frac{a}{\lambda}$ and the centered Gamma law has

$$a(x)=\frac{2}{\lambda}(x+\frac{a}{\lambda})$$

In this case

$$d(\mathcal{L}(Y),\mu)\leq CE\left[\left|\frac{1}{\lambda}(F+\frac{a}{\lambda})-\langle D(-L)^{-1}Y,DY\rangle_H\right|\right]$$

Example

The uniform U$(0,1)$ distribution. Here the density is $\mathbf{f}(\mathbf{x}) = 1_{[0,1]}(\mathbf{x})$, the mean is $\mathbf{EX} = \frac{1}{2}$ and $\mathbf{U}[0,1] - \mathbf{EU}[0,1]$ has quared diffusion coefficient

$$\mathbf{a}(\mathbf{x}) = (\mathbf{x} + \frac{1}{2})(\frac{1}{2} - \mathbf{x}) = \frac{1}{4} - \mathbf{x}^2.$$

Here

$$\mathbf{d}(\mathcal{L}(\mathbf{Y}), \mu) \leq \mathbf{CE}\left[\left|\frac{1}{2} - \frac{1}{2}\mathbf{Y}^2 - \langle \mathbf{D}(-\mathbf{L})^{-1}\mathbf{Y}, \mathbf{DY}\rangle_{\mathbf{H}}\right|\right]$$

Example

The Pareto Pareto(ν) distribution, $\nu > 1$. The density of this law is $\mathbf{f}(\mathbf{x}) = \nu(1+\mathbf{x})^{-\nu-1}$ for $\mathbf{x} \in \mathbb{R}$. This implies **EX** $\sim \frac{1}{\nu-1}$ and

$$\mathbf{a}(\mathbf{x}) = \frac{2}{\nu-1}(\mathbf{x} + \frac{1}{\nu-1})(1 + \mathbf{x} + \frac{1}{\nu-1}).$$

Example

The inverse Gamma distribution with parameters $\delta > 0, \lambda > 1$. The density function is given by

$$f(x) = \frac{\delta^{\lambda}}{\Gamma(\lambda)} x^{-\lambda-1} e^{-\frac{\delta}{x}} 1_{(0,\infty)}(x).$$

The expectation of this law $EX = \frac{\delta}{\lambda-1}$ and the centered inverse Gamma distribution is associated with

$$a(x) = \frac{2}{\lambda-1}(x + \frac{\delta}{\lambda-1})^2$$

Example

The F distribution with parameters $a \geq 2, b > 2$. Here

$$f(x) = \frac{a^{\frac{a}{2}} b^{\frac{b}{2}}}{\beta(\frac{a}{2}, \frac{b}{2})} \frac{x^{\frac{a}{2}-1}}{(b+ax)^{\frac{a+b}{2}}} 1_{(0,\infty)}(x).$$

Moreover $\mathbf{E}X = \frac{b}{b-2}$ and the centered **F** law has

$$a(x) = \frac{4}{a(b-2)}(x + \frac{b}{b-2}) \left[b + a(x + \frac{b}{b-2}) \right].$$

Example

The Beta $\beta(\mathbf{a}, \mathbf{b})$ law, $\mathbf{a}, \mathbf{b} > 0$. In this case the probability density function is

$$\mathbf{f}(\mathbf{x}) = \frac{\Gamma(\mathbf{a}+\mathbf{b})}{\Gamma(\mathbf{a})\Gamma(\mathbf{b})}\mathbf{x}^{\mathbf{a}-1}(1-\mathbf{x})^{\mathbf{b}-1}1_{(0,1)}(\mathbf{x}),$$

$\mathbf{EX} = \frac{\mathbf{a}}{\mathbf{a}+\mathbf{b}}$ and the centered beta law has

$$\mathbf{a}(\mathbf{x}) = \frac{2}{\mathbf{a}+\mathbf{b}}(\mathbf{x} + \frac{\mathbf{a}}{\mathbf{a}+\mathbf{b}})(\frac{\mathbf{b}}{\mathbf{a}+\mathbf{b}} - \mathbf{x}).$$

Therefore $\alpha = -\frac{2}{\mathbf{a}+\mathbf{b}}, \beta = \frac{2}{\mathbf{a}+\mathbf{b}}\frac{\mathbf{b}-\mathbf{a}}{\mathbf{a}+\mathbf{b}}, \gamma = \frac{2}{\mathbf{a}+\mathbf{b}}\frac{\mathbf{a}}{\mathbf{a}+\mathbf{b}}\frac{\mathbf{b}}{\mathbf{a}+\mathbf{b}}$.

By "Significance of the bound" we mean the following : given a random variable whose probability law is the invariant measure $\mathbf{X}$, then the distance between its law and $\mathbf{X}$ is zero.

we need to calculate the random variable $\langle \mathbf{DY}, \mathbf{D}(-\mathbf{L})^{-1}\mathbf{b}(\mathbf{Y})\rangle$.

This random variable (and its conditional expectation given $\mathbf{Y}$) appears in several works related to Malliavin calculus and Stein's method

In general, it is difficult to find an explicit expression for it for general $\mathbf{Y}$.

But in the case when $\mathbf{Y}$ is a function of a Gaussian vector we have a very useful formula : if $\mathbf{Y} = \mathbf{h}(\mathbf{N}) - \mathbf{E}\mathbf{h}(\mathbf{N})$ where $\mathbf{h} : \mathbb{R}^{\mathbf{n}} \to \mathbb{R}$ is a function of class $\mathbf{C}^1$ with bounded derivatives and $\mathbf{N} = (\mathbf{N}_1, ..., \mathbf{N}_{\mathbf{n}})$ is a Gaussian vector with zero mean and covariance matrix $\mathbf{K} = (\mathbf{K}_{\mathbf{i},\mathbf{j}})_{\mathbf{i},\mathbf{j}=1,..,\mathbf{n}}$ then (we will omit in the sequel the index $\mathbf{H}$ for the scalar product)

$$\langle \mathbf{D}(-\mathbf{L})^{-1}(\mathbf{Y} - \mathbf{E}\mathbf{Y}), \mathbf{D}\mathbf{Y} \rangle$$
$$= \int_0^\infty \mathbf{e}^{-\mathbf{u}} \mathbf{d}\mathbf{u} \mathbf{E}' \sum_{\mathbf{i},\mathbf{j}=1}^{\mathbf{n}} \mathbf{K}_{\mathbf{i},\mathbf{j}} \frac{\partial \mathbf{h}}{\partial \mathbf{x}_{\mathbf{i}}}(\mathbf{N}) \frac{\partial \mathbf{h}}{\partial \mathbf{x}_{\mathbf{j}}}(\mathbf{e}^{-\mathbf{u}}\mathbf{N} + \sqrt{1 - \mathbf{e}^{-2\mathbf{u}}}\mathbf{N}').$$

Here $\mathbf{N}'$ denotes and independent copy of $\mathbf{N}$ a

In the case of the uniform distribution :

Let $\mathbf{f}, \mathbf{g} \in \mathbf{L}^2([0, \mathbf{T}])$ such that

$$\|\mathbf{f}\|_{\mathbf{L}^2([0,\mathbf{T}])} = \|\mathbf{g}\|_{\mathbf{L}^2([0,\mathbf{T}])} = 1 \text{ and } \langle \mathbf{f}, \mathbf{g} \rangle_{\mathbf{L}^2([0,\mathbf{T}])} = 0.$$

Then $\mathbf{W}(\mathbf{f})$ and $\mathbf{W}(\mathbf{g})$ are independent standard normal random variables. Define the random variable $\mathbf{Y}$ by

$$\mathbf{Y} = \mathbf{e}^{-\frac{1}{2}\left(\mathbf{W}(\mathbf{f})^2 + \mathbf{W}(\mathbf{g})^2\right)}.$$

Then it is well-known that $\mathbf{Y}$ has uniform distribution $\mathbf{U}([0, 1])$ since the random variable $-\frac{1}{2}\left(\mathbf{W}(\mathbf{f})^2 + \mathbf{W}(\mathbf{g})^2\right)$ has exponential distribution with parameter 1.

One can make all the compatations and we find

$$\langle \mathbf{D}(-\mathbf{L})^{-1}(\mathbf{Y}-\frac{1}{2}), \mathbf{DY}\rangle = \mathbf{a}(\mathbf{Y}) = \mathbf{Y}(1-\mathbf{Y}).$$

Assume that there exists a random variable $\mathbf{G} \in \mathbb{D}^{1,4}$ such that :

- the distribution of $\mathbf{G}$ is equal to μ
- $\langle \mathbf{DG}, \mathbf{D}(-\mathbf{L})^{-1}\mathbf{G} \rangle_{\mathbf{H}}$ is measurable with respect to the σ-field generated by $\mathbf{G}$.

Note : The second assumption is satified by several common distributions (Gaussian, gamma, uniform, Pareto)

Then, the following statements are equivalent.

1. The vector valued random variable $(\mathbf{F_m}, \frac{1}{\mathbf{n}}\|\mathbf{DF_m}\|^2_{\mathbf{H}})$ converge to
$(\mathbf{G}, \langle \mathbf{DG}, \mathbf{D}(-\mathbf{L})^{-1}\mathbf{G}\rangle_{\mathbf{H}})$ in distribution as $\mathbf{m} \to \infty$.
2. $\frac{1}{2}\mathbf{a}(\mathbf{F_m}) - \frac{1}{\mathbf{n}}\|\mathbf{DF_m}\|^2_{\mathbf{H}}$ converges to 0 in $\mathbb{L}^2$ as $\mathbf{m} \to \infty$.

Comments

- For the standard normal law, $\mathbf{a}(\mathbf{x}) = 2$. So, the condition that $\frac{1}{2}\mathbf{a}(\mathbf{F_m}) - \frac{1}{\mathbf{n}}\|\mathbf{DF_m}\|^2_{\mathbf{H}}$ converges to 0 in $\mathbb{L}^2$ as $\mathbf{m} \to \infty$ means $1 - \frac{1}{\mathbf{n}}\|\mathbf{DF_m}\|^2_{\mathbf{H}}$ converges to 0 in $\mathbb{L}^2$
- It fits also for the Gamma case
- the condition that $\frac{1}{2}\mathbf{a}(\mathbf{F_m}) - \frac{1}{\mathbf{n}}\|\mathbf{DF_m}\|^2_{\mathbf{H}}$ converges to 0 in $\mathbb{L}^2$ is equivalent to the condition that $\frac{1}{4}\mathbf{E}[\mathbf{a}(\mathbf{F_m})^2] - \frac{1}{\mathbf{n}^2}\mathbf{E}[\|\mathbf{DF_m}\|^4_{\mathbf{H}}]$ converges to 0 as $\mathbf{m} \to \infty$.

Sketsch of the proof :
Assume (2) : By the Stein bound, $\mathbf{F_m}$ converges to $\mathbf{G}$ in distribution. Since $\langle \mathbf{DG}, \mathbf{D}(-\mathbf{L})^{-1}\mathbf{G}\rangle$ is measurable with respect to the σ-field generated by $\mathbf{G}$, we obtain

$$\frac{1}{2}\mathbf{a}(\mathbf{G}) = \langle \mathbf{DG}, \mathbf{D}(-\mathbf{L})^{-1}\mathbf{G}\rangle_{\mathbf{H}}$$

almost surely.

Hence, by the convergence of $\mathbf{F_m}$ to $\mathbf{G}$ in distribution and 2, we have for $\mathbf{h}_1, \mathbf{h}_2 \in \mathbf{C_b}(\mathbb{R})$

$$\limsup_{\mathbf{m}\to\infty} \left| \mathbf{E}\left[\mathbf{h}_1(\mathbf{F_m})\mathbf{h}_2\left(\frac{1}{\mathbf{n}}\|\mathbf{DF_m}\|^2_{\mathbf{H}}\right)\right] - \mathbf{E}\left[\mathbf{h}_1(\mathbf{G})\mathbf{h}_2\left(\langle \mathbf{DG}, \mathbf{D}(-\mathbf{L})^{-1}\mathbf{G}\rangle_{\mathbf{H}}\right)\right.\right.$$
$$\leq \limsup_{\mathbf{m}\to\infty} \left| \mathbf{E}\left[\mathbf{h}_1(\mathbf{F_m})\mathbf{h}_2\left(\frac{1}{2}\mathbf{a}(\mathbf{F_m})\right)\right] - \mathbf{E}\left[\mathbf{h}_1(\mathbf{G})\mathbf{h}_2\left(\frac{1}{2}\mathbf{a}(\mathbf{G})\right)\right]\right|$$
$$+ \limsup_{\mathbf{m}\to\infty} \left| \mathbf{E}\left[\mathbf{h}_1(\mathbf{F_m})\left\{\mathbf{h}_2\left(\frac{1}{2}\mathbf{a}(\mathbf{F_m})\right) - \mathbf{h}_2\left(\frac{1}{\mathbf{n}}\|\mathbf{DF_m}\|^2_{\mathbf{H}}\right)\right\}\right]\right|$$
$$= 0.$$

Assume (1). Since $\langle \mathbf{DG}, \mathbf{D}(-\mathbf{L})^{-1}\mathbf{G}\rangle$ is measurable with respect to the σ-field generated by $\mathbf{G}$, again

$$\frac{1}{2}\mathbf{a}(\mathbf{G}) = \langle \mathbf{DG}, \mathbf{D}(-\mathbf{L})^{-1}\mathbf{G}\rangle_{\mathbf{H}}$$

almost surely. Therefore, by (1)

$$\lim_{\mathbf{m}\to\infty}\left(\frac{1}{4}\mathbf{E}[\mathbf{a}(\mathbf{F_m})^2] - \frac{1}{\mathbf{n}^2}\mathbf{E}[\|\mathbf{DF_m}\|_{\mathbf{H}}^4]\right)$$
$$= \frac{1}{4}\mathbf{E}[\mathbf{a}(\mathbf{G})^2] - \mathbf{E}\left[\langle \mathbf{DG}, \mathbf{D}(-\mathbf{L})^{-1}\mathbf{G}\rangle_{\mathbf{H}}^2\right] = 0.$$

The measurablility of $\langle \mathbf{DG}, \mathbf{D}(-\mathbf{L})^{-1}\mathbf{G}\rangle_{\mathbf{H}}$ with respect to the σ-field generated by $\mathbf{G}$ is assumed. In the special cases this condition immediately follows.
If :

- $\mathbf{F} = \mathbf{cW}(\mathbf{h})$ where $\mathbf{c} \in \mathbb{R}$ and $\mathbf{h} \in \mathcal{H}$ such that $\|\mathbf{h}\|_{\mathbf{H}} = 1$.

  Note : Case (i) includes the centered Gaussian distribution

- $\mathbf{F} = \mathbf{c}(\mathbf{W}(\mathbf{h})^2 - 1)$ where $\mathbf{c} \in \mathbb{R}$ and $\mathbf{h} \in \mathcal{H}$ such that $\|\mathbf{h}\|_{\mathbf{H}} = 1$.

  Note : Case (ii) includes the centered Gamma distribution.

- $\mathbf{F} = \mathbf{e}^{\mathbf{cW}(\mathbf{h})}$ where $\mathbf{c} \in \mathbb{R}$ and $\mathbf{h} \in \mathcal{H}$ such that $\|\mathbf{h}\|_{\mathbf{H}} = 1$.

  Note : Case (iii) contains the random variables with the centered log-normal law.

- $\mathbf{F} = \mathbf{e}^{\mathbf{c}\sum_{\mathbf{k}=1}^{\mathbf{n}} \mathbf{W}(\mathbf{h_k})^2}$ where $\mathbf{n} \in \mathbb{N}$, $\mathbf{c} \in (-\infty, 1/2)$ and $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h_n} \in \mathcal{H}$ such that $\|\mathbf{h_k}\|_{\mathbf{H}} = 1$ for $\mathbf{k} = 1, 2, \ldots, \mathbf{n}$, and $(\mathbf{h_k}, \mathbf{h_l})_{\mathbf{H}} = 0$ for $\mathbf{k}, \mathbf{l} = 1, 2, \ldots, \mathbf{n}$.

Note : Case (iv) includes :

- the uniform distribution $\mathbf{U}[0, 1]$ (by taking for example $\mathbf{n} = 2$ and $\mathbf{c} = -\frac{1}{2}$)
- the centered Pareto distribution (if we consider $\mathbf{n} = 2$ and $\mathbf{c} = \frac{1}{4}$ we obtain a centered Pareto distribution with parameter 2)
- the centered beta distribution (by taking $\mathbf{c} = -1$, $\mathbf{n} = 2$ with have a random variable with centered beta law with parameters $\frac{1}{2}$ and 1).

we can show one-way implication as follows.

Proposition

Let $\mathbf{F_m}$ *as before and assume that*

$$\sup_{\mathbf{m}} \mathbf{E}[\mathbf{F_m^2}] = \sup_{\mathbf{m}} \|\mathbf{f_m}\|^2_{\mathbf{H}^{\odot n}} < \infty. \tag{6}$$

If the distribution of $\mathbf{F_m}$ *converges to* μ,

$$\lim_{\mathbf{m}\to\infty} \mathbf{E}\left[\mathbf{F_m^4} - \frac{3}{2}\mathbf{F_m^2}\mathbf{a}(\mathbf{F_m})\right] = 0.$$

Since we are studying the convergence of a sequence of multiple stochastic integrals, whose expectation is zero, we will assume that the measure μ is centered and the drift coefficient is $\mathbf{b}(\mathbf{x}) = -\mathbf{x}$.

We will also assume that the diffusion coefficient is a polynomial of second degree expressed as

$$\mathbf{a}(\mathbf{x}) = \alpha \mathbf{x}^2 + \beta \mathbf{x} + \gamma, \quad \mathbf{x} \in \mathbf{S}, \; \alpha, \beta, \gamma \in \mathbb{R} \tag{7}$$

such that $\mathbf{a}(\mathbf{x}) > 0$ for every $\mathbf{x} \in \mathbf{S}$.

We study when the necessary and sufficient condition for the weak convergence of a sequence of multiple integrals toward the law μ is satisfied.

This class contains the known continuous probability distributions. Let us list below several examples.

Example

The normal distribution $\mathbf{N}(0,\gamma)$, $\gamma > 0$. In this case $\mathbf{a}(\mathbf{x}) = 2\gamma$.

Example

The Student $\mathbf{t}(\nu)$ distribution, $\nu > 1$. In this case, if $\mathbf{X} \sim \mathbf{t}(\nu)$ then the probability density of $\mathbf{X}$ is

$$\mathbf{f}(\mathbf{x}) = \frac{\Gamma(\frac{\nu+1}{2})\nu^{\frac{\nu}{2}}}{\sqrt{\pi}\Gamma(\frac{\nu}{2})}(\nu + \mathbf{x}^2)^{-\frac{\nu+1}{2}}$$

for $\mathbf{x} \in \mathbb{R}$. In particular $\mathbf{EX} = 0$. The squared diffusion coefficient is

Example

The Pareto Pareto(ν) distribution, $\nu > 1$**.** The density of this law is $\mathbf{f}(\mathbf{x}) = \nu(1+\mathbf{x})^{-\nu-1}$ for $\mathbf{x} \in \mathbb{R}$. This implies $\mathbf{EX} \sim \frac{1}{\nu-1}$ and

$$\mathbf{a}(\mathbf{x}) = \frac{2}{\nu-1}(\mathbf{x} + \frac{1}{\nu-1})(1+\mathbf{x}+\frac{1}{\nu-1}).$$

Thus $\alpha = \frac{2}{\nu-1}, \beta = \frac{2}{\nu-1}(1+\frac{2}{\nu-1}), \gamma = \frac{2}{(\nu-1)^2}(1+\frac{1}{\nu-1})$.

Example

The Gamma $\Gamma(a, \lambda)$, $a, \lambda > 0$ law. Here the density is $f(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x}$ for $x > 0$ and $f(x) = 0$ otherwise. Also $EX = \frac{a}{\lambda}$ and the centered Gamma law has

$$a(x) = \frac{2}{\lambda}(x + \frac{a}{\lambda})$$

meaning that $\alpha = 0, \beta = \frac{2}{\lambda}, \gamma = \frac{2a}{\lambda^2}$.

Example

The inverse Gamma distribution with parameters $\delta > 0, \lambda > 1$. The density function is given by

$$\mathbf{f}(\mathbf{x}) = \frac{\delta^{\lambda}}{\Gamma(\lambda)}\mathbf{x}^{-\lambda-1}\mathbf{e}^{-\frac{\delta}{\mathbf{x}}}1_{(0,\infty)}(\mathbf{x}).$$

The expectation of this law $\mathbf{EX} = \frac{\delta}{\lambda-1}$ and the centered inverse Gamma distribution is associated with

$$\mathbf{a}(\mathbf{x}) = \frac{2}{\lambda-1}(\mathbf{x} + \frac{\delta}{\lambda-1})^2$$

which gives $\alpha = \frac{2}{\lambda-1}, \beta = \frac{4\delta}{(\lambda-1)^2}, \gamma = \frac{2\delta^2}{(\lambda-1)^3}$.

Example

The F distribution with parameters $a \geq 2, b > 2$. Here

$$f(x) = \frac{a^{\frac{a}{2}} b^{\frac{b}{2}}}{\beta(\frac{a}{2}, \frac{b}{2})} \frac{x^{\frac{a}{2}-1}}{(b+ax)^{\frac{a+b}{2}}} 1_{(0,\infty)}(x).$$

Moreover $\mathbf{EX} = \frac{b}{b-2}$ and the centered **F** law has

$$a(x) = \frac{4}{a(b-2)}(x + \frac{b}{b-2}) \left[b + a(x + \frac{b}{b-2}) \right].$$

Example

The uniform $\mathbf{U}(0,1)$ distribution. Here the density is $\mathbf{f}(\mathbf{x}) = 1_{[0,1]}(\mathbf{x})$, the mean is $\mathbf{EX} = \frac{1}{2}$ and $\mathbf{U}[0,1] - \mathbf{EU}[0,1]$ has quared diffusion coefficient

$$\mathbf{a}(\mathbf{x}) = (\mathbf{x} + \frac{1}{2})(\frac{1}{2} - \mathbf{x}) = \frac{1}{4} - \mathbf{x}^2.$$

So $\alpha = -1, \beta = 0, \gamma = \frac{1}{4}$.

Example

The Beta $\beta(\mathbf{a},\mathbf{b})$ law, $\mathbf{a},\mathbf{b} > 0$. In this case the probability density function is

$$\mathbf{f}(\mathbf{x}) = \Gamma(\mathbf{a}+\mathbf{b}) \ldots \mathbf{x}^{\mathbf{a}-1}(1-\mathbf{x})^{\mathbf{b}-1} 1 \ldots (\mathbf{x})$$

Idea : find some relation on the moments of $\mathbf{X}$ with law μ¿

Actually, using the charcacterization of the r.v with law μ, we find : for every $\mathbf{k} \in \mathbb{R}$, $\mathbf{k} \geq 1$ such that $\mathbf{E}\mathbf{X}^{2\mathbf{k}} < \infty$ one has

$$\left(1 - \frac{2\mathbf{k}-1}{2}\alpha\right)\mathbf{E}\mathbf{X}^{2\mathbf{k}} = \frac{2\mathbf{k}-1}{2}\beta\mathbf{E}\mathbf{X}^{2\mathbf{k}-1} + \frac{2\mathbf{k}-1}{2}\gamma\mathbf{E}\mathbf{X}^{2\mathbf{k}-2}. \quad (8)$$

In particular, if $\alpha \neq 2$,

$$\mathbf{E}X^2 = \frac{\gamma}{2-\alpha}, \tag{9}$$

if $\alpha \neq 1, 2$,

$$\mathbf{E}X^3 = \frac{\beta}{1-\alpha}\mathbf{E}X^2 = \frac{\beta\gamma}{(1-\alpha)(2-\alpha)} \tag{10}$$

and if $\alpha \neq 2, \frac{2}{3}$, then

$$\mathbf{E}X^4 = \frac{3(\frac{\beta^2}{1-\alpha}+\gamma)}{2-3\alpha}\mathbf{E}X^2 = \frac{3\gamma(\frac{\beta^2}{1-\alpha}+\gamma)}{(2-\alpha)(2-3\alpha)}. \tag{11}$$

We can compute the third and fourth moment of a random variable in a fixed Wiener chaos.

Lemma

Let $\mathbf{F} = \mathbf{I_n}(\mathbf{f})$ *with* $\mathbf{n} \geq 1$ *and* $\mathbf{f} \in \mathbf{H}^{\odot \mathbf{n}}$. *Then*

$$\mathbf{EF}^3 = \frac{\mathbf{n}!^3}{\left(\frac{\mathbf{n}}{2}\right)!^3} \langle \mathbf{f}, \mathbf{f} \tilde{\otimes}_{\frac{\mathbf{n}}{2}} \mathbf{f} \rangle 1_{\{n \text{ is even}\}}$$

and

$$\mathbf{EF}^4 = 3(\mathbf{EF}^2)^2$$

$$+3\mathbf{n} \sum_{\mathbf{p}=1}^{\mathbf{n}-1} (\mathbf{p}-1)! \begin{pmatrix} \mathbf{n}-1 \\ \mathbf{p}-1 \end{pmatrix}^2 \mathbf{p}! \begin{pmatrix} \mathbf{n} \\ \mathbf{p} \end{pmatrix}^2 (2\mathbf{n}-2\mathbf{p})! \|\mathbf{f_m} \tilde{\otimes}_{\mathbf{p}} \mathbf{f_m}\|^2.$$

The case $\beta = 0$. This is the case of the Student, uniform and beta (with parameters $\mathbf{a} = \mathbf{b}$) distributions. In this situation, since $\mathbf{EX}^3 = 0$, the order of the chaos $\mathbf{n}$ can be even or odd in principle.

Theorem

Assume $\alpha \neq 2, \frac{2}{3}$ and $\beta = 0$. Fix $\mathbf{n} \geq 1$ and let $\{\mathbf{F_m} = \mathbf{I_n}(\mathbf{f_m}, \mathbf{m} \geq 1\}$ satisfying

$$\mathbf{EF_m^2} \to_{\mathbf{m}\to\infty} \mathbf{EX}^2, \mathbf{EF_m^4} \to_{\mathbf{m}\to\infty} \mathbf{EX}^4 \text{ and } \mathbf{EF_m^3} \to_{\mathbf{m}\to\infty} \mathbf{EX}^3.$$

Then $\alpha = 0, \gamma > 0$ and $\mathbf{X}$ follows a centered normal distribution with variance γ.

Use the formula for $\mathbf{EX}^2$ and $\mathbf{EX}^4$ (here $\mathbf{EX}^3 = 0$).

As a consequence, we notice that several probability distributions cannot be limits in distribution of sequences of multiple stochastic integrals.

Corollary

A sequence a random variables in a fixed Wiener chaos cannot converge to the uniform, Student or to the beta distribution $\beta(\mathbf{a}, \mathbf{b})$ *with* $\mathbf{a} = \mathbf{b}$.

The case $\beta \neq 0$. This is the case of the Pareto, Gamma, inverse Gamma and **F** distributions.
Fix $\mathbf{n} \geq 1$. Consider throughout this section that $\{\mathbf{F_m}, \mathbf{m} \geq 1\}$ a sequence of random variables expressed as $\mathbf{F_m} = \mathbf{I_n}(\mathbf{f_m})$ with $\mathbf{f_m} \in \mathbf{H}^{\odot \mathbf{n}}$. Since the third moment of a multiple Wiener -Itô integral of odd order is zero, from (10) we may assume in this paragraph that **n is even.**

Theorem

Assume $\alpha \neq 1, 2, \frac{2}{3}$. *Fix* $\mathbf{n} \geq 1$ *and let* $\{\mathbf{F_m} = \mathbf{I_n}(\mathbf{f_m}), \mathbf{m} \geq 1\}$ *satisfying*

$$\mathbf{E}\mathbf{F}_\mathbf{m}^2 \rightarrow_{\mathbf{m}\rightarrow\infty} \mathbf{E}\mathbf{X}^2, \mathbf{E}\mathbf{F}_\mathbf{m}^4 \rightarrow_{\mathbf{m}\rightarrow\infty} \mathbf{E}\mathbf{X}^4 \text{ and } \mathbf{E}\mathbf{F}_\mathbf{m}^3 \rightarrow_{\mathbf{m}\rightarrow\infty} \mathbf{E}\mathbf{X}^3.$$

Moreover, let us assume $\frac{\alpha}{2-3\alpha} \leq 0$ *(that is,* $\alpha \in \mathbb{R} \setminus (0, \frac{2}{3}]$*). Then* $\alpha = 0$ *and* $\mathbf{X}$ *follows a centered Gamma law* $\Gamma(\mathbf{a}, \lambda) - \mathbf{E}\Gamma(\mathbf{a}, \lambda)$ *where* $\beta = \frac{2}{\lambda}, \gamma = \frac{2\mathbf{a}}{\lambda^2}$.

As a consequence, a sequence of multiple stochastic integrals in a fixed Wiener chaos cannot converge to a Pareto distribution with parameter $\mu < 4$, to a inverse Gamma distribution with parameter $\lambda < 4$ or to a **F** distribution with parameter $\mathbf{b} < 10$.

In the case of the centered Gamma distribution, the result reads as follows : a sequence $\mathbf{F_m} = \mathbf{I_n}(\mathbf{f_m})$ such that $\mathbf{E}\mathbf{F_m^2} \to_{\mathbf{m}\to\infty} \frac{\mathbf{a}}{\lambda^2}$ converges to the centered Gamma law $\Gamma(\mathbf{a}, \lambda) - \mathbf{E}\Gamma(\mathbf{a}, \lambda)$ if and only if the following assertions are satisfied :

- $\mathbf{E}\mathbf{F_m^3} \to_{\mathbf{m}} \frac{2\mathbf{a}}{\lambda^3}$ and $\mathbf{E}\mathbf{F_m^4} \to_{\mathbf{m}} \frac{3\mathbf{a}(\mathbf{a}+2)}{\lambda^4}$
- $\|\frac{2}{\lambda}\mathbf{c_n}\mathbf{f_m} - \mathbf{f_m}\tilde{\otimes}_{\mathbf{n}/2}\mathbf{f_m}\| \to_{\mathbf{m}} 0$ ($\mathbf{c_n}$ is a constant)
- $\frac{1}{\lambda}\mathbf{F_m^2} + \frac{\mathbf{a}}{\lambda^2} - \frac{1}{2}\|\mathbf{D}\mathbf{F_m}\|^2_{\mathbf{H}}$ converges to zero in $\mathbf{L}^2(\Omega)$.

When $\lambda = \frac{1}{2}$ and $\mathbf{a} = \frac{\nu}{2}$ we retrieve known results.