On probabilistic approximations and variance estimates

Giovanni Peccati (Luxembourg University)

Journées MAS 2012 — 30/08/2012

Overview

- Survey of a recently developed line of research, studying probabilistic approximations (e.g. Central Limit Theorems or Laws of Small Numbers) using the **Malliavin calculus of variations** and the **Stein** and **Chen-Stein methods**. Keyword: **integration by parts**.

Overview

- Survey of a recently developed line of research, studying probabilistic approximations (e.g. Central Limit Theorems or Laws of Small Numbers) using the **Malliavin calculus of variations** and the **Stein** and **Chen-Stein methods**. Keyword: **integration by parts**.
- Basic message: one can compute Berry-Esseen bounds by means of **variance estimates**, loosely analogous to **second order Poincaré inequalities**. They often rely on moment estimates for chaotic random variables.

Overview

- In a **Gaussian framework**: applications in a number of fields: fractional processes, Gaussian polymers, random fields on homogeneous spaces, random matrices, U-stats. See the monograph: Nourdin-P. 2012.

Overview

- In a **Gaussian framework**: applications in a number of fields: fractional processes, Gaussian polymers, random fields on homogeneous spaces, random matrices, U-stats. See the monograph: Nourdin-P. 2012.
- In a **Poisson framework**: impetus comes since two years from stochastic geometry. Applications to: geometric random graphs, k-flat processes, Poisson-Voronoi, ... (Lachièze-Rey, Last, Penrose, P., Reitzner, Schulte, Thaele).

The basic problem

Fix $n \geq 1$, and let $X = (X_1, ..., X_n) \sim \mathcal{N}_n(0, \mathbb{I}_n)$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be smooth, and $N \sim \mathcal{N}(0, 1)$.

The basic problem

Fix $n \geq 1$, and let $X = (X_1, ..., X_n) \sim \mathcal{N}_n(0, \mathbb{I}_n)$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be smooth, and $N \sim \mathcal{N}(0, 1)$. Define $F = f(X_1, ..., X_n)$ and assume F is centered and has unit variance.

The basic problem

Fix $n \geq 1$, and let $X = (X_1, ..., X_n) \sim \mathcal{N}_n(0, \mathbb{I}_n)$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be smooth, and $N \sim \mathcal{N}(0, 1)$. Define $F = f(X_1, ..., X_n)$ and assume F is centered and has unit variance.

Problem: How distant are the laws of F and N?

The basic problem

Fix $n \geq 1$, and let $X = (X_1, ..., X_n) \sim \mathcal{N}_n(0, \mathbb{I}_n)$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be smooth, and $N \sim \mathcal{N}(0, 1)$. Define $F = f(X_1, ..., X_n)$ and assume F is centered and has unit variance.

Problem: How distant are the laws of F and N?

We shall first tackle this problem by implementing a **smart path method** (used e.g. for proving the Sudakov-Fernique inequality),

The basic problem

Fix $n \geq 1$, and let $X = (X_1, ..., X_n) \sim \mathcal{N}_n(0, \mathbb{I}_n)$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be smooth, and $N \sim \mathcal{N}(0, 1)$. Define $F = f(X_1, ..., X_n)$ and assume F is centered and has unit variance.

Problem: How distant are the laws of F and N?

We shall first tackle this problem by implementing a **smart path method** (used e.g. for proving the Sudakov-Fernique inequality), as well as by using the **Ornstein-Uhlenbeck semigroup** $\{P_t\}_{t \geq 0}$:

$$P_t f(y) = E[f(e^{-t}y + \sqrt{1 - e^{-2t}}X)] \quad \text{(Mehler form)}.$$

Also: L and L^{-1} denote the generator of P_t and its pseudo-inverse.

Approach by smart paths

Assume F and N are independent. Take $\varphi : \mathbb{R} \to \mathbb{R} \in C_b^2$ and define the function

$$\Psi(t) = E[\varphi(\sqrt{t}F + \sqrt{1-t}N)],$$

Assume F and N are independent. Take $\varphi : \mathbb{R} \to \mathbb{R} \in C_b^2$ and define the function

$$\Psi(t) = E[\varphi(\sqrt{t}F + \sqrt{1-t}N)],$$

in such a way that $E[\varphi(F)] - E[\varphi(N)] = \int_0^1 \Psi'(t)dt$, and

$$\begin{aligned}
\Psi'(t) &= \frac{1}{2}\left\{E\left[\frac{F}{\sqrt{t}}\varphi'(\sqrt{t}F + \sqrt{1-t}N)\right]\right. \\
&\quad \left. -E\left[\frac{N}{\sqrt{1-t}}\varphi'(\sqrt{t}F + \sqrt{1-t}N)\right]\right\} \\
&:= \frac{1}{2}(A_t - B_t).
\end{aligned}$$

Approach by smart paths

Of course: $B_t = E\left[\varphi''(\sqrt{t}F + \sqrt{1-t}N)\right]$.

Approach by smart paths

Of course: $B_t = E\left[\varphi''(\sqrt{t}F + \sqrt{1-t}N)\right]$.
Using $F = P_0 f(X) - P_\infty f(X) = -\int_0^\infty (d/dt)P_t f(X)dt = -\int_0^\infty LP_t f(X)dt$, it is now a matter of simple verification that

$$
\begin{aligned}
A_t &= E\left[\frac{F}{\sqrt{t}}\varphi'(\sqrt{t}F + \sqrt{1-t}N)\right] \\
&= E\Big[\varphi''(\sqrt{t}F + \sqrt{1-t}N) \times \langle \nabla f(X), -\nabla L^{-1} f(X)\rangle_{\mathbb{R}^n}\Big],
\end{aligned}
$$

Approach by smart paths

Of course: $B_t = E\left[\varphi''(\sqrt{t}F + \sqrt{1-t}N)\right]$.
Using $F = P_0 f(X) - P_\infty f(X) = -\int_0^\infty (d/dt) P_t f(X) dt = -\int_0^\infty LP_t f(X) dt$, it is now a matter of simple verification that

$$\begin{aligned} A_t &= E\left[\frac{F}{\sqrt{t}}\varphi'(\sqrt{t}F + \sqrt{1-t}N)\right] \\ &= E\left[\varphi''(\sqrt{t}F + \sqrt{1-t}N) \times \langle \nabla f(X), -\nabla L^{-1} f(X)\rangle_{\mathbb{R}^n}\right], \end{aligned}$$

where $-\nabla L^{-1} f(y) = \int_0^\infty e^{-t} P_t \nabla f(y) dt$, yielding that

$$\begin{aligned} \Big|E[\varphi(F)] - E[\varphi(N)]\Big| &\leq \sup_t |\Psi'(t)| \\ &\leq \frac{\|\varphi''\|_\infty}{2} E|1 - G| \leq \frac{\|\varphi''\|_\infty}{2}\sqrt{\mathrm{Var}(G)}. \end{aligned}$$

where $G = \langle \nabla f(X), -\nabla L^{-1} f(X)\rangle_{\mathbb{R}^n}$.

Five questions

- Why is the variance of G relevant to normal approximations?
- Can one consider non-smooth test functions?
- How well does this procedure extend to an infinite-dimensional setting?
- Can we connect these results to Poincaré-type inequalities?
- Are r.v.'s of the form of G uniquely related to normal approximations?

Stein's method in a nutshell (Stein 1972, 1986)

Lemma (Stein's Lemma)

A random variable F has a $\mathcal{N}(0,1)$ distribution if and only if for every smooth function g

$$E\left[g'(F) - Fg(F)\right] = 0.$$

Stein's method in a nutshell (Stein 1972, 1986)

Lemma (Stein's Lemma)

A random variable F has a $\mathcal{N}(0,1)$ distribution if and only if for every smooth function g

$$E\left[g'(F) - Fg(F)\right] = 0.$$

Heuristically, Stein's Lemma suggests that, if F is such that

$$E\left[g'(F) - Fg(F)\right] \simeq 0$$

for a "**sufficiently large**" class of smooth functions g, then the law of F must be **close to Gaussian.**

Stein's method in a nutshell (Stein 1972, 1986)

Formally: fix $h = \mathbf{1}_C$, take $N \sim \mathcal{N}(0,1)$, and introduce the **Stein equation**

$$\boxed{g'(y) - yg(y) = h(y) - E[h(N)], \quad y \in \mathbf{R}} \qquad (\star)$$

(the unknown being g).

Stein's method in a nutshell (Stein 1972, 1986)

Formally: fix $h = \mathbf{1}_C$, take $N \sim \mathcal{N}(0,1)$, and introduce the **Stein equation**

$$\boxed{g'(y) - yg(y) = h(y) - E[h(N)], \quad y \in \mathbf{R}} \qquad (\star)$$

(the unknown being g).

Classic estimates by Stein (1986) yield that there exists a solution to $(\star)$, say g_h, such that

$$\boxed{|g_h| \leq \sqrt{\pi/2} \quad \text{and} \quad |g_h'| \leq 2}.$$

Stein's method in nutshell (Stein 1972, 1986)

The previous results show that, if $N \sim \mathcal{N}(0,1)$, then

$$d_{TV}(F,N) \leq \sup_{\substack{|g'|\leq 2 \\ |g|\leq\sqrt{\pi/2}}} \left| E\left[g'(F) - Fg(F)\right]\right|,$$

which is known as the **Stein's bound** on the total variation distance.

- Let $\mathfrak{H}$ be a real separable Hilbert space. An **isonormal Gaussian process** $X = \{X(h) : h \in \mathfrak{H}\}$ is a centered Gaussian family verifying $E[X(h)X(h')] = \langle h, h' \rangle_{\mathfrak{H}}$. In the finite dimensional case $\mathfrak{H} = \mathbb{R}^n$.

Isonormal processes, chaos and Malliavin

- Let $\mathfrak{H}$ be a real separable Hilbert space. An **isonormal Gaussian process** $X = \{X(h) : h \in \mathfrak{H}\}$ is a centered Gaussian family verifying $E[X(h)X(h')] = \langle h, h' \rangle_{\mathfrak{H}}$. In the finite dimensional case $\mathfrak{H} = \mathbb{R}^n$.
- For every $q \geq 1$ and every $f \in \mathfrak{H}^{\odot q}$, $I_q(f)$ is the multiple Wiener-Itô integral of f with respect to X. Multiple integrals of order q compose the qth **Wiener chaos** of X, noted C_q. In the finite-dimensional case, C_q is just the closed linear space generated by r.v.'s of the type

$$H(X_1, ..., X_n),$$

where X is a n-valued **Hermite polynomial** of exact degree q.

- Recall the **chaotic decomposition**: $L^2(\sigma(X)) = \bigoplus_{q\geq 1} C_q$.

Isonormal processes, chaos and Malliavin

- Recall the **chaotic decomposition**: $L^2(\sigma(X)) = \bigoplus_{q\geq 1} C_q$.
- We use some standard operators of Malliavin calculus: D (= **derivative**); D^2 (= **second derivative**); δ (= **divergence**, adjoint of D), P_t (= **OU semigroup**), L (= **its generator**), L^{-1},... In the finite-dimensional case, $Df(x) = \nabla f(X)$, $D^2 = \mathrm{Hess} f(X)$. P_t is again defined via a Mehler-type formula.

- Recall the **chaotic decomposition**: $L^2(\sigma(X)) = \bigoplus_{q\geq 1} C_q$.
- We use some standard operators of Malliavin calculus: D (= **derivative**); D^2 (= **second derivative**); δ (= **divergence**, adjoint of D), P_t (= **OU semigroup**), L (= **its generator**), L^{-1},... In the finite-dimensional case, $Df(x) = \nabla f(X)$, $D^2 = \mathrm{Hess} f(X)$. P_t is again defined via a Mehler-type formula.
- Recall that C_q are eigenspaces of P_t, L and L^{-1} (with eigenvalues e^{-qt}, $-q$, $-q^{-1}$, respectively).

- Recall the **chaotic decomposition**: $L^2(\sigma(X)) = \bigoplus_{q\geq 1} C_q$.
- We use some standard operators of Malliavin calculus: D (= **derivative**); D^2 (= **second derivative**); δ (= **divergence**, adjoint of D), P_t (= **OU semigroup**), L (= **its generator**), L^{-1},... In the finite-dimensional case, $Df(x) = \nabla f(X)$, $D^2 = \mathrm{Hess} f(X)$. P_t is again defined via a Mehler-type formula.
- Recall that C_q are eigenspaces of P_t, L and L^{-1} (with eigenvalues e^{-qt}, $-q$, $-q^{-1}$, respectively).
- **Chain rule:** if $\varphi : \mathbb{R} \to \mathbb{R}$ is smooth, $D\varphi(F) = \varphi'(F)DF$.

Isonormal processes, chaos and Malliavin

- Recall the **chaotic decomposition**: $L^2(\sigma(X)) = \bigoplus_{q\geq 1} C_q$.
- We use some standard operators of Malliavin calculus: D (= **derivative**); D^2 (= **second derivative**); δ (= **divergence**, adjoint of D), P_t (= **OU semigroup**), L (= **its generator**), L^{-1},... In the finite-dimensional case, $Df(x) = \nabla f(X)$, $D^2 = \mathrm{Hess} f(X)$. P_t is again defined via a Mehler-type formula.
- Recall that C_q are eigenspaces of P_t, L and L^{-1} (with eigenvalues e^{-qt}, $-q$, $-q^{-1}$, respectively).
- **Chain rule:** if $\varphi : \mathbb{R} \to \mathbb{R}$ is smooth, $D\varphi(F) = \varphi'(F)DF$.
- **Important relation:** $-\delta D = L$.

Central formula

Let $F \in \mathrm{dom} D$ be centered and with unit variance. Assume F has a density and g is Lipschitz.

Central formula

Let $F \in \mathrm{dom}D$ be centered and with unit variance. Assume F has a density and g is Lipschitz. Then,

$$E[Fg(F)] = E[LL^{-1}Fg(F)] = E[-\delta(DL^{-1}F)g(F)]$$

Central formula

Let $F \in \mathrm{dom} D$ be centered and with unit variance. Assume F has a density and g is Lipschitz. Then,

$$E[Fg(F)] = E[LL^{-1}Fg(F)] = E[-\delta(DL^{-1}F)g(F)]$$
$$= E[\langle Dg(F), -DL^{-1}F\rangle_{\mathfrak{H}}]$$

Central formula

Let $F \in \mathrm{dom} D$ be centered and with unit variance. Assume F has a density and g is Lipschitz. Then,

$$E[Fg(F)] = E[LL^{-1}Fg(F)] = E[-\delta(DL^{-1}F)g(F)]$$
$$= E[\langle Dg(F), -DL^{-1}F\rangle_{\mathfrak{H}}] = E[g'(F)\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}].$$

Central formula

Let $F \in \mathrm{dom} D$ be centered and with unit variance. Assume F has a density and g is Lipschitz. Then,

$$E[Fg(F)] = E[LL^{-1}Fg(F)] = E[-\delta(DL^{-1}F)g(F)]$$
$$= E[\langle Dg(F), -DL^{-1}F\rangle_{\mathfrak{H}}] = E[g'(F)\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}].$$

Note: if $F \in C_q$, then $\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}} = \frac{1}{q}\|DF\|^2_{\mathfrak{H}}$.

Main bound

Theorem (Nourdin-Peccati, 2009)

Let $F \in \mathrm{dom} D$ be centered and with unit variance, and $N \sim \mathcal{N}(0,1)$.

$$d_{TV}(F,N) \le \sup_{|g'|\le 2} \left| E\left[g'(F) - Fg(F)\right]\right|$$

Main bound

Theorem (Nourdin-Peccati, 2009)

Let $F \in \mathrm{dom} D$ be centered and with unit variance, and $N \sim \mathcal{N}(0,1)$.

$$d_{TV}(F,N) \leq \sup_{|g'|\leq 2} \left| E\left[g'(F) - Fg(F)\right]\right|$$

$$= \sup_{|g'|\leq 2} \left| E\left[g'(F)\left(1 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}\right)\right]\right|$$

$$\leq 2E|1 - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}| \leq 2\mathrm{Var}(\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}})^{1/2}.$$

Focus on Wiener chaos

The most important application of the method is the following. As before, $N \sim \mathcal{N}(0,1)$.

Theorem (Nourdin-Peccati, 2009)

For $q \geq 1$, let $F \in C_q$ have unit variance. Then,

$$\mathrm{Var}\left(\frac{1}{q}\|DF\|_{\mathfrak{H}}^2\right) \leq \frac{q-1}{3q}\left\{E[F^4]-3\right\}.$$

Focus on Wiener chaos

The most important application of the method is the following. As before, $N \sim \mathcal{N}(0,1)$.

Theorem (Nourdin-Peccati, 2009)

For $q \geq 1$, let $F \in C_q$ have unit variance. Then,

$$\mathrm{Var}\left(\frac{1}{q}\|DF\|_{\mathfrak{H}}^2\right) \leq \frac{q-1}{3q}\left\{E[F^4]-3\right\}.$$

In particular,

$$d_{TV}(F,N) \leq 2\mathrm{Var}\left(\frac{1}{q}\|DF\|_{\mathfrak{H}}^2\right)^{1/2} \leq \frac{2}{\sqrt{3}}\sqrt{E[F^4]-3}.$$

Focus on Wiener chaos

The most important application of the method is the following. As before, $N \sim \mathcal{N}(0,1)$.

Theorem (Nourdin-Peccati, 2009)

For $q \geq 1$, let $F \in C_q$ have unit variance. Then,

$$\mathrm{Var}\left(\frac{1}{q}\|DF\|_{\mathfrak{H}}^2\right) \leq \frac{q-1}{3q}\left\{E[F^4]-3\right\}.$$

In particular,

$$d_{TV}(F,N) \leq 2\mathrm{Var}\left(\frac{1}{q}\|DF\|_{\mathfrak{H}}^2\right)^{1/2} \leq \frac{2}{\sqrt{3}}\sqrt{E[F^4]-3}.$$

This result allows one to recover a **fourth moment theorem** on the Wiener chaos, first proved by Nualart and Peccati (2005). Multidimensional version: Peccati and Tudor (2005). See also recent works by Nourdin, Poly and Nualart.

Recall the **Poincaré inequality**: for every $F \in \mathrm{dom} D$,

$$\mathrm{Var}(F) \leq E[\|DF\|^2_{\mathfrak{H}}].$$

Second order Poincaré inequalities

Recall the **Poincaré inequality**: for every $F \in \mathrm{dom} D$,

$$\mathrm{Var}(F) \leq E[\|DF\|^2_{\mathfrak{H}}].$$

Theorem (Nourdin, Peccati, Reinert, 2010)

Let $F \in \mathrm{dom} D$ be centered and with unit variance, and $N \sim \mathcal{N}(0,1)$. Then, one has the second-order estimate

$$d_{TV}(F,N) \leq 2\sqrt{5} E[\|DF\|^4_{\mathfrak{H}}]^{1/4} \times E[\|D^2 F\|^4_{\mathrm{op}}]^{1/4}.$$

Second order Poincaré inequalities

Recall the **Poincaré inequality**: for every $F \in \mathrm{dom} D$,

$$\mathrm{Var}(F) \leq E[\|DF\|_{\mathfrak{H}}^2].$$

Theorem (Nourdin, Peccati, Reinert, 2010)

Let $F \in \mathrm{dom} D$ be centered and with unit variance, and $N \sim \mathcal{N}(0,1)$. Then, one has the second-order estimate

$$d_{TV}(F,N) \leq 2\sqrt{5} E[\|DF\|_{\mathfrak{H}}^4]^{1/4} \times E[\|D^2F\|_{\mathrm{op}}^4]^{1/4}.$$

Note: a second order Poincaré inequality in the finite-dimensional case was introduced by Chatterjee in 2007.

Some extensions

- Criteria for **optimal rates** (Biermé, Bonami, Nourdin, Peccati, 2009, 2011)

Some extensions

- Criteria for **optimal rates** (Biermé, Bonami, Nourdin, Peccati, 2009, 2011)
- Chaos of a general Markov operator (Ledoux, 2011)

Some extensions

- Criteria for **optimal rates** (Biermé, Bonami, Nourdin, Peccati, 2009, 2011)
- Chaos of a general Markov operator (Ledoux, 2011)
- The **Nourdin-Viens formula** (2010). If $g_F(F) = E[\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}} | F] > 0$ a.s., then F has a density:

$$\rho(x) = \frac{E|F|}{2g_F(x)} \exp\left(-\int_0^x \frac{y}{g_F(y)} dy\right).$$

Some extensions

- Criteria for **optimal rates** (Biermé, Bonami, Nourdin, Peccati, 2009, 2011)
- Chaos of a general Markov operator (Ledoux, 2011)
- The **Nourdin-Viens formula** (2010). If $g_F(F) = E[\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}|F] > 0$ a.s., then F has a density:

$$\rho(x) = \frac{E|F|}{2g_F(x)} \exp\left(-\int_0^x \frac{y}{g_F(y)} dy\right).$$

For instance, if $g_F(x) \leq \alpha x + \beta$, then

$$P(F \geq x) \leq \exp\left(-\frac{x^2}{2\alpha x + 2\beta}\right), x \geq 0.$$

- Criteria for **optimal rates** (Biermé, Bonami, Nourdin, Peccati, 2009, 2011)
- Chaos of a general Markov operator (Ledoux, 2011)
- The **Nourdin-Viens formula** (2010). If $g_F(F) = E[\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}|F] > 0$ a.s., then F has a density:

$$\rho(x) = \frac{E|F|}{2g_F(x)} \exp\left(-\int_0^x \frac{y}{g_F(y)} dy\right).$$

For instance, if $g_F(x) \leq \alpha x + \beta$, then

$$P(F \geq x) \leq \exp\left(-\frac{x^2}{2\alpha x + 2\beta}\right), x \geq 0.$$

- Almost sure CLTs (Bercu, Nourdin, Taqqu, 2011)

Some applications

- High-frequency limit theorems for spherical fields (Baldi, Kerkyacharian, Lan, Marinucci, Peccati, Picard, Wigman)
- Asymptotic results for fractional processes (Bandorff-Nielsen, Biermé, Corcuera, Léon, Nourdin, Nualart, Peccati, Podoloskij, Tudor, Viens)
- Gaussian polymers (Viens)
- Universality principles for homogeneous sums (Nourdin, Peccati, Reinert)
- Fluctuations of traces of random matrices (Nourdin, Peccati)

Switching to Poisson

- $(Z, \mathcal{Z})$ is a Polish space.
- Given a σ-finite non atomic measure μ, we denote by η a **Poisson measure with control** μ, and its compensated counterpart is $\widehat{\eta} = \eta(\cdot) - \mu(\cdot)$.
- Recall: for every A, B such that $A \cap B = \emptyset$ and $\mu(A), \mu(B) < \infty$, $\eta(A)$ and $\eta(B)$ are two independent Poisson r.v.'s of parameters $\mu(A), \mu(B)$.

Switching to Poisson

- $(Z, \mathcal{Z})$ is a Polish space.
- Given a σ-finite non atomic measure μ, we denote by η a **Poisson measure with control** μ, and its compensated counterpart is $\widehat{\eta} = \eta(\cdot) - \mu(\cdot)$.
- Recall: for every A, B such that $A \cap B = \emptyset$ and $\mu(A), \mu(B) < \infty$, $\eta(A)$ and $\eta(B)$ are two independent Poisson r.v.'s of parameters $\mu(A), \mu(B)$.
- Recall also the **Chen-Stein Lemma**: a random variable $F \in \mathbb{Z}_+$ has the $\mathrm{Po}(\lambda)$ distribution if and only if, for every g bounded,

$$E[Fg(F)] = \lambda E[g(F+1)].$$

Switching to Poisson

- For every symmetric square-integrable function f in q variables, we define the multiple Wiener-Itô integral

$$I_q(f) = \int_Z \cdots \int_Z f(x_1, ..., x_q) \mathbf{1}_{\{\text{no diagonals}\}} \hat{\eta}(dx_1) \cdots \hat{\eta}(dx_q).$$

- Recall that every $F \in L^2(\sigma(\eta))$ can be written as: $F = E(F) + \sum_{q=1}^{\infty} I_q(f_q)$.

Switching to Poisson

- The **derivative operator** is: $D_z F = \sum_q q I_{q-1}(f_q(z,\cdot))$.
- Nualart and Vives (1990): $D_z F(\eta) = F(\eta + \delta_z) - F(\eta)$ (add-one cost).
- **The O-U generator**: $LF = -\sum_{q\geq 1} q I_q(f_q)$.
- **Pseudo-inverse of the O-U generator**: $L^{-1}F = -\sum_{q\geq 1} q^{-1} I_q(f_q)$.
- **Integration by parts**: for every X derivable and F centered,

$$E[XF] = E[\langle DX, -DL^{-1}F\rangle_\mu].$$

A Gaussian/Poisson alternative

Let $N \sim \mathcal{N}(0,1)$ and $X \sim \mathrm{Po}(\lambda)$, $\lambda > 0$.

Theorem (Peccati-Solé-Taqqu-Utzet, 2010; Peccati 2012)

- *Let $F \in \mathrm{dom} D$ be centered and have unit variance*

$$
\begin{aligned}
d_W(F,N) \quad \le \quad & E|1 - \langle DF, -DL^{-1}F\rangle_\mu| \\
& + E\int_Z (D_z F)^2 |D_z L^{-1} F| \mu(dz).
\end{aligned}
$$

A Gaussian/Poisson alternative

Let $N \sim \mathcal{N}(0,1)$ and $X \sim \mathrm{Po}(\lambda)$, $\lambda > 0$.

Theorem (Peccati-Solé-Taqqu-Utzet, 2010; Peccati 2012)

- *Let $F \in \mathrm{dom} D$ be centered and have unit variance*

$$d_W(F,N) \leq E|1 - \langle DF, -DL^{-1}F\rangle_\mu| + E\int_Z (D_zF)^2 |D_zL^{-1}F| \mu(dz).$$

- *For a $\mathbb{Z}_+$-valued random variable $F \in \mathrm{dom} D$ with mean λ,*

$$d_{TV}(F, \mathrm{Po}(\lambda)) \leq B_\lambda E|\lambda - \langle DF, -DL^{-1}F\rangle_\mu| + C_\lambda E\int_Z |(D_zF)(D_zF-1)D_zL^{-1}F|\mu(dz),$$

where $B_\lambda := \frac{1-e^{-\lambda}}{\lambda} = \lambda C_\lambda$.

Applications

Several new applications in stochastic geometry, starting from a paper by Reitzner and Schulte (2010): geometric random graphs, k-flat processes, Poisson-Voronoi approximations, ... (Lachièze-Rey, Last, Peccati, Penrose, Schulte, Thaele).

Several new applications in stochastic geometry, starting from a paper by Reitzner and Schulte (2010): geometric random graphs, k-flat processes, Poisson-Voronoi approximations, ... (Lachièze-Rey, Last, Peccati, Penrose, Schulte, Thaele).

An important role in these applications is played by **geometric U-statistics**.

A simple example

- Let η be a Poisson measure on $\mathbb{R}^2$, with control equal to the Lebesgue measure. Define

$$W_n = \left[-\frac{1}{2}\sqrt{n}, \frac{1}{2}\sqrt{n}\right]^2, \; n = 1, 2, ..., .$$

- Let $\{r_n\}$ be a non-increasing sequence of positive numbers. For every n, we consider the **disk graph** $G_n = (V_n, E_n)$, where

$$V_n = W_n \cap \eta, \quad E_n = \{(x, y) : 0 < |x - y| < r_n\}.$$

- We are interested in the asymptotic behavior of

$$M_n = \#\{\text{edges of } G_n\}, \quad \widetilde{M}_n = \frac{M_n - E[M_n]}{\sqrt{\mathrm{Var}(M_n)}}.$$

A simple example

One has that:

(i) If $nr_n^2 \to \infty$, then $\widetilde{M}_n \overset{LAW}{\to} \mathcal{N}(0,1)$;

(ii) If $nr_n^2 \to \lambda \in (0,+\infty)$, then $M_n \overset{TV}{\to} \mathrm{Po}(\lambda)$;

(iii) If $nr_n^2 \to 0$, then $M_n, \widetilde{M}_n \overset{L^1}{\to} 0$.

A simple example

One has that:

(i) If $nr_n^2 \to \infty$, then $\widetilde{M}_n \stackrel{LAW}{\to} \mathcal{N}(0,1)$;

(ii) If $nr_n^2 \to \lambda \in (0,+\infty)$, then $M_n \stackrel{TV}{\to} \mathrm{Po}(\lambda)$;

(iii) If $nr_n^2 \to 0$, then $M_n, \widetilde{M}_n \stackrel{L^1}{\to} 0$.

Our bounds then give

$$d_W(\widetilde{M}_n, \mathcal{N}(0,1)) \leq \frac{C_1}{r_n\sqrt{n}}$$

$$d_{TV}(M_n, \mathrm{Po}(\lambda')) \leq |nr_n^2 - \lambda| + C_2 r_n.$$

Thank you! Merci!