High-frequency statistics for Levy processes: a Stein's method perspective

Arturo Jaramillo Gil, Chiara Amorino and Mark Podolskij

Centro de Investigación en Matemáticas (CIMAT)

Motivation

Let S be a random variable of interest that is not accessible, and let $\hat{S}$ be a good approximation of S.

Motivation

Let S be a random variable of interest that is not accessible, and let $\hat{S}$ be a good approximation of S.

How can we describe the fluctuation $F = S - \hat{S}$ of this approximation?

Motivation

Let S be a random variable of interest that is not accessible, and let $\hat{S}$ be a good approximation of S.

How can we describe the fluctuation $F = S - \hat{S}$ of this approximation?

Example:

- $\hat{S} = \mathbb{E}[S] =: \mu$.

Motivation

Let S be a random variable of interest that is not accessible, and let $\hat{S}$ be a good approximation of S.

How can we describe the fluctuation $F = S - \hat{S}$ of this approximation?

Example:

- $\hat{S} = \mathbb{E}[S] =: \mu$.
- Fluctuation is $F = S - \mu$.

Motivation

Let S be a random variable of interest that is not accessible, and let $\hat{S}$ be a good approximation of S.

How can we describe the fluctuation $F = S - \hat{S}$ of this approximation?

Example:

- $\hat{S} = \mathbb{E}[S] =: \mu$.
- Fluctuation is $F = S - \mu$.
- If S is the sum of n i.i.d. random variables with finite moments, then $F \overset{Law}{\approx} \sigma N$, with $N \sim \mathcal{N}(0,1)$ and $\sigma^2 = Var[S]$.

Motivation

Let S be a random variable of interest that is not accessible, and let $\hat{S}$ be a good approximation of S.

How can we describe the fluctuation $F = S - \hat{S}$ of this approximation?

Example:

- $\hat{S} = \mathbb{E}[S] =: \mu$.
- Fluctuation is $F = S - \mu$.
- If S is the sum of n i.i.d. random variables with finite moments, then $F \overset{Law}{\approx} \sigma N$, with $N \sim \mathcal{N}(0, 1)$ and $\sigma^2 = Var[S]$.
- Berry-Esseen theorem: $d_K(\frac{S-\mu}{\sigma}, N) \leq C/\sqrt{n}$.

Context of the problem

General Objective

Understand the cumulative error when approximating a time series with a good predictor.

Context of the problem

General Objective

Understand the cumulative error when approximating a time series with a good predictor.

Ingredients:

1. A non-accessible time series $\boldsymbol{\xi} = \{\xi_k\}_{k \geq 1}$.

Context of the problem

General Objective

Understand the cumulative error when approximating a time series with a good predictor.

Ingredients:

1. A non-accessible time series $\boldsymbol{\xi} = \{\xi_k\}_{k\geq 1}$.
2. A very good estimator $\boldsymbol{\eta} = \{\eta_k\}_{k\geq 1}$ of $\boldsymbol{\xi}$.

Context of the problem

General Objective

Understand the cumulative error when approximating a time series with a good predictor.

Ingredients:

1. A non-accessible time series $\boldsymbol{\xi} = \{\xi_k\}_{k \geq 1}$.
2. A very good estimator $\boldsymbol{\eta} = \{\eta_k\}_{k \geq 1}$ of $\boldsymbol{\xi}$.

Context of the problem

General Objective

Understand the cumulative error when approximating a time series with a good predictor.

Ingredients:

1. A non-accessible time series $\boldsymbol{\xi} = \{\xi_k\}_{k\geq 1}$.
2. A very good estimator $\boldsymbol{\eta} = \{\eta_k\}_{k\geq 1}$ of $\boldsymbol{\xi}$.

Object of interest:

$$\text{Cumulative error} = (\xi_1 - \eta_1) + \cdots + (\xi_n - \eta_n)$$

True context of the problem

I observe a real-valued process $X = \{X_t\}_{t \geq 0}$ in stages:

True context of the problem

I observe a real-valued process $X = \{X_t\}_{t \geq 0}$ in stages: at stage i, I can observe $X_{1/n}, \ldots, X_{(i-1)/n}$.

True context of the problem

I observe a real-valued process $X = \{X_t\}_{t \geq 0}$ in stages: at stage i, I can observe $X_{1/n}, \ldots, X_{(i-1)/n}$.

Quantities of interest

At stage i, I would like to know

$$\xi_i := g(X_{\frac{i-1}{n}}, a_n \Delta_i X),$$

with $\Delta_i X := X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$, $g : \mathbb{R}^2 \to \mathbb{R}$ an appropriate function, and a_n an appropriate scaling.

True context of the problem

I observe a real-valued process $X = \{X_t\}_{t \geq 0}$ in stages: at stage i, I can observe $X_{1/n}, \dots, X_{(i-1)/n}$.

Quantities of interest

At stage i, I would like to know

$$\xi_i := g(X_{\frac{i-1}{n}}, a_n \Delta_i X),$$

with $\Delta_i X := X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$, $g : \mathbb{R}^2 \to \mathbb{R}$ an appropriate function, and a_n an appropriate scaling.

Observation mechanism

At stage i, I only have the information $\mathcal{F}_{i-1} := \sigma(X_{\frac{1}{n}}, \dots, X_{\frac{i-1}{n}})$.

True context of the problem

I observe a real-valued process $X = \{X_t\}_{t \geq 0}$ in stages: at stage i, I can observe $X_{1/n}, \ldots, X_{(i-1)/n}$.

Quantities of interest

At stage i, I would like to know

$$\xi_i := g(X_{\frac{i-1}{n}}, a_n \Delta_i X),$$

with $\Delta_i X := X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$, $g : \mathbb{R}^2 \to \mathbb{R}$ an appropriate function, and a_n an appropriate scaling.

Observation mechanism

At stage i, I only have the information $\mathcal{F}_{i-1} := \sigma(X_{\frac{1}{n}}, \ldots, X_{\frac{i-1}{n}})$.

Estimator:

$$\eta_i := \mathbb{E}[g(X_{\frac{i-1}{n}}, \Delta_i X) \mid \mathcal{F}_{i-1}].$$

Now indeed: the true context of the problem

Suppose that X starts at zero and has independent and stationary increments.

Now indeed: the true context of the problem

Suppose that X starts at zero and has independent and stationary increments.

For $t \geq 0$, we define the cumulative error on $[0, t]$ as

$$Z_n(t) := \frac{1}{\sqrt{n}} \sum_{i=1}^{nt} (g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) - \mathbb{E}[g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) \mid \mathcal{F}_{i-1}]),$$

with $\mathcal{I}_{i,n} := a_n(X_{\frac{i}{n}} - X_{\frac{i-1}{n}}, \ldots, X_{\frac{i+m}{n}} - X_{\frac{i+m-1}{n}})$, and $g : \mathbb{R}^{m+2} \to \mathbb{R}$ an appropriate function.

Now indeed: the true context of the problem

Suppose that X starts at zero and has independent and stationary increments.

For $t \geq 0$, we define the cumulative error on $[0, t]$ as

$$Z_n(t) := \frac{1}{\sqrt{n}} \sum_{i=1}^{nt} (g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) - \mathbb{E}[g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) \mid \mathcal{F}_{i-1}]),$$

with $\mathcal{I}_{i,n} := a_n(X_{\frac{i}{n}} - X_{\frac{i-1}{n}}, \ldots, X_{\frac{i+m}{n}} - X_{\frac{i+m-1}{n}})$, and $g : \mathbb{R}^{m+2} \to \mathbb{R}$ an appropriate function.

Assumptions

- The scaling a_n is such that $a_n X_{1/n}$ converges in law.
- There exists a constant $\alpha > 0$ such that $\mathbb{P}[X \geq s] \leq Cts^{-\alpha}$.

First problem of interest

Our **object of interest** is a process $Z = \{Z_n(t)\}_{t\geq 0}$, defined by

$$Z_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{nt} (g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) - \mathbb{E}[g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) \mid \mathcal{F}_{i-1}]),$$

Key elements:

- Parameter α indicating non-integrability of X.
- Filter function g.

Questions of interest

- What is the limit of Z_n? (if it exists)
- What is the rate of convergence?

Case where S is a r.v. and N is normal

The discrepancy between $F := (S - \mu)/\sigma$ and N is studied using expressions of the form

$$|\mathbb{E}[h(F) - h(N)]|. \tag{1}$$

Case where S is a r.v. and N is normal

The discrepancy between $F := (S - \mu)/\sigma$ and N is studied using expressions of the form

$$|\mathbb{E}[h(F) - h(N)]|. \quad (1)$$

Note that $F \overset{Law}{=} N$ if the following **property** holds: "(1) is zero for sufficiently many test functions h".

Case where S is a r.v. and N is normal

The discrepancy between $F := (S - \mu)/\sigma$ and N is studied using expressions of the form

$$|\mathbb{E}[h(F) - h(N)]|. \tag{1}$$

Note that $F \stackrel{Law}{=} N$ if the following **property** holds: "(1) is zero for sufficiently many test functions h".

Another **property** characterizing the normal distribution is...

Case where S is a r.v. and N is normal

The discrepancy between $F := (S - \mu)/\sigma$ and N is studied using expressions of the form

$$|\mathbb{E}[h(F) - h(N)]|. \tag{1}$$

Note that $F \overset{Law}{=} N$ if the following **property** holds: "(1) is zero for sufficiently many test functions h".

Another **property** characterizing the normal distribution is...

$$|\mathbb{E}[Nf'(N) - f''(N)]| = 0, \tag{2}$$

for sufficiently many functions f.

Case where S is a r.v. and N is normal

The discrepancy between $F := (S-\mu)/\sigma$ and N is studied using expressions of the form

$$|\mathbb{E}[h(F) - h(N)]|. \tag{1}$$

Note that $F \overset{Law}{=} N$ if the following **property** holds: "(1) is zero for sufficiently many test functions h".

Another **property** characterizing the normal distribution is...

$$|\mathbb{E}[Nf'(N) - f''(N)]| = 0, \tag{2}$$

for sufficiently many functions f.

Stein's heuristic

$$|\mathbb{E}[Ff'(F) - f''(F)]| \approx 0 \qquad \Rightarrow \qquad |\mathbb{E}[h(F) - h(N)]| \approx 0.$$

Formalization of Stein's heuristic

For a given function h, consider the equation

$$\boldsymbol{x} \cdot \nabla f(\boldsymbol{x}) - Tr[Hess[f](\boldsymbol{x})\Sigma] = h(\boldsymbol{x}) - \mathbb{E}[f(\boldsymbol{N})].$$

Formalization of Stein's heuristic

For a given function h, consider the equation

$$\boldsymbol{x} \cdot \nabla f(\boldsymbol{x}) - Tr[Hess[f](\boldsymbol{x})\Sigma] = h(\boldsymbol{x}) - \mathbb{E}[f(\boldsymbol{N})].$$

The solution is given by

$$f(\boldsymbol{x}) = \int_0^\infty (\mathbb{E}[h(\boldsymbol{N})] - \mathbb{E}_{\boldsymbol{x}}[h(\boldsymbol{Y}_t)])dt,$$

where $\boldsymbol{Y}_t$ is a Markov process starting at $\boldsymbol{x}$ and converging to $\boldsymbol{N}$.

Formalization of Stein's heuristic

For a given function h, consider the equation

$$\boldsymbol{x} \cdot \nabla f(\boldsymbol{x}) - Tr[Hess[f](\boldsymbol{x})\Sigma] = h(\boldsymbol{x}) - \mathbb{E}[f(\boldsymbol{N})].$$

The solution is given by

$$f(\boldsymbol{x}) = \int_0^\infty (\mathbb{E}[h(\boldsymbol{N})] - \mathbb{E}_{\boldsymbol{x}}[h(\boldsymbol{Y}_t)])dt,$$

where $\boldsymbol{Y}_t$ is a Markov process starting at $\boldsymbol{x}$ and converging to $\boldsymbol{N}$.

We obtain

$$\mathbb{E}[\boldsymbol{F} \cdot \nabla f(\boldsymbol{F}) - Tr[Hess[f](\boldsymbol{F})\Sigma]] = \mathbb{E}[h(\boldsymbol{F}) - f(\boldsymbol{N})].$$

Central issue of Stein's method

If $\boldsymbol{F}$ and $\boldsymbol{N}$ are multivariate and $\boldsymbol{N}$ has covariance Σ, the quantity we need to control is

$$|\mathbb{E}[\boldsymbol{F} \cdot \nabla f_{\Sigma}(\boldsymbol{F}) - Tr[Hess[f_{\Sigma}](\boldsymbol{F})\Sigma]]|.$$

Central issue of Stein's method

If $\boldsymbol{F}$ and $\boldsymbol{N}$ are multivariate and $\boldsymbol{N}$ has covariance Σ, the quantity we need to control is

$$|\mathbb{E}[\boldsymbol{F} \cdot \nabla f_\Sigma(\boldsymbol{F}) - Tr[Hess[f_\Sigma](\boldsymbol{F})\Sigma]]|.$$

Main challenge of the method:

How do we estimate $\mathbb{E}[\boldsymbol{F} \cdot \nabla f_\Sigma(\boldsymbol{F})]$?

Central issue of Stein's method

If $\boldsymbol{F}$ and $\boldsymbol{N}$ are multivariate and $\boldsymbol{N}$ has covariance Σ, the quantity we need to control is

$$|\mathbb{E}[\boldsymbol{F} \cdot \nabla f_\Sigma(\boldsymbol{F}) - Tr[Hess[f_\Sigma](\boldsymbol{F})\Sigma]]|.$$

Main challenge of the method:

How do we estimate $\mathbb{E}[\boldsymbol{F} \cdot \nabla f_\Sigma(\boldsymbol{F})]$?

Typical approach: use the original ideas from Lindeberg's (or Stein's) method.

So what if $\boldsymbol{F} = Z_n$? and Σ is random

The key computation is $\mathbb{E}[\boldsymbol{Z}_n \cdot \nabla f_\Sigma(\boldsymbol{Z}_n)]$.

So what if $\boldsymbol{F} = Z_n$? and Σ is random

The key computation is $\mathbb{E}[\boldsymbol{Z}_n \cdot \nabla f_\Sigma(\boldsymbol{Z}_n)]$.In this case,

$$\begin{aligned}
&\mathbb{E}[\boldsymbol{Z}_n \cdot \nabla f_\Sigma(\boldsymbol{Z}_n)] \\
&\quad = \frac{1}{\sqrt{n}} \sum_{i=1}^{nt} \mathbb{E}[(g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) - \mathbb{E}[g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) \mid \mathcal{F}_{i-1}])) \cdot \nabla f_\Sigma(\boldsymbol{Z}_n)]
\end{aligned}$$

Key observation

So what if $F = Z_n$? and Σ is random

The key computation is $\mathbb{E}[\boldsymbol{Z}_n \cdot \nabla f_\Sigma(\boldsymbol{Z}_n)]$.In this case,

$$\mathbb{E}[\boldsymbol{Z}_n \cdot \nabla f_\Sigma(\boldsymbol{Z}_n)]$$
$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{nt} \mathbb{E}[(g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) - \mathbb{E}[g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) \mid \mathcal{F}_{i-1}])) \cdot \nabla f_\Sigma(\boldsymbol{Z}_n)]$$

Key observation

$$0 = \frac{1}{\sqrt{n}} \sum_{i=1}^{nt} \mathbb{E}[(g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) - \mathbb{E}[g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) \mid \mathcal{F}_{i-1}])) \cdot \nabla f_{\dot{\Sigma}^i}(\dot{\boldsymbol{Z}}_n^i)],$$

where $\dot{\boldsymbol{Z}}_n^i$ and $\dot{\Sigma}^i$ are like $\boldsymbol{Z}_n$ and Σ, but removing the part of X in $[(i-1)/n, i/n]$

So what if $\boldsymbol{F} = Z_n$?

Now we can write

$$\mathbb{E}[\boldsymbol{Z}_n \cdot \nabla f_\Sigma(\boldsymbol{Z}_n)]$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{nt} \mathbb{E}[(g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) - \mathbb{E}[g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) \mid \mathcal{F}_{i-1}]))$$

$$\cdot \nabla (f_\Sigma(\boldsymbol{Z}_n) - f_{\dot{\Sigma}^i}(\dot{\boldsymbol{Z}}_n^i)))]$$

New bottleneck:

So what if $F = Z_n$?

Now we can write

$$\mathbb{E}[\boldsymbol{Z}_n \cdot \nabla f_\Sigma(\boldsymbol{Z}_n)] = \frac{1}{\sqrt{n}} \sum_{i=1}^{nt} \mathbb{E}[(g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) - \mathbb{E}[g(X_{\frac{i-1}{n}}, \mathcal{I}_{i,n}) \mid \mathcal{F}_{i-1}])) \cdot \nabla(f_\Sigma(\boldsymbol{Z}_n) - f_{\dot{\Sigma}^i}(\dot{\boldsymbol{Z}}_n^i)))]$$

New bottleneck: understand

$$f_\Sigma(\boldsymbol{Z}_n) - f_{\dot{\Sigma}^i}(\dot{\boldsymbol{Z}}_n^i)$$

by means of Taylor approximations.

The mixed-Gaussian case

An easy criterion for mixed Gaussian convergence:

The mixed-Gaussian case

An easy criterion for mixed Gaussian convergence:

Theorem (Amorino, Jaramillo, Podolskij)
Consider a sequence $\boldsymbol{F}_n$, which is $\mathcal{G}$-measurable for some σ-algebra. If the convergence

$$\mathbb{E}[Y(\boldsymbol{F}_n \cdot \nabla f(\boldsymbol{F}_n) - Tr[Hess[f](\boldsymbol{F}_n)\Sigma])] \to 0,$$

holds for all bounded $\mathcal{G}$-measurable Y and adequate test functions $h \in C^2(\mathbb{R}^r; \mathbb{R})$ then we obtain

$$S_n \stackrel{Law}{\to} \Sigma^{1/2} N,$$

where $N \sim \mathcal{N}_r(0, id)$ is a standard r-dimensional normal variable defined on an extended space and independent of $\mathcal{G}$.

Main results

Let $X^{(1-m)}, \ldots, X^{(2m)}$ be independent copies of $\boldsymbol{X}$, and define

$$\mathfrak{g}_n(\boldsymbol{x}) := \sum_{j=-m}^{m} Cov[g(\boldsymbol{x}, a_n X_{1/n}^{(1)}, \ldots, a_n X_{1/n}^{(m+1)}), \\ g(\boldsymbol{x}, a_n X_{1/n}^{(j+1)}, \ldots, a_n X_{1/n}^{(j+m+1)})].$$

Main results

Let $X^{(1-m)}, \ldots, X^{(2m)}$ be independent copies of $\boldsymbol{X}$, and define

$$\mathfrak{g}_n(\boldsymbol{x}) := \sum_{j=-m}^{m} Cov[g(\boldsymbol{x}, a_n X_{1/n}^{(1)}, \ldots, a_n X_{1/n}^{(m+1)}), \\ g(\boldsymbol{x}, a_n X_{1/n}^{(j+1)}, \ldots, a_n X_{1/n}^{(j+m+1)})].$$

Theorem (Amorino, Jaramillo, Podolskij, 2023)
If $\mathfrak{g}_n$ converges and $\alpha \in (0,1)$, then

$$Z^{(n)} \overset{Law}{\to} \{\int_0^t \sqrt{\lim_k \mathfrak{g}_k(\boldsymbol{X}_s)} W(ds)\}_{t \geq 0},$$

where W is a Brownian motion independent of $\boldsymbol{X}$.

Main results (part II)

Suppose that X is symmetric α-stable (including $\alpha = 2$), and define

$$d(\mu, \nu) := \sup_h \left| \int h \, d\mu - \int h \, d\nu \right|,$$

where the supremum is taken over all functions h satisfying $\|h^{(i)}\|_\infty \leq 1$ for $i = 0, 1, 2, 3$.

Main results (part II)

Theorem (Amorino, Jaramillo, Podolskij, 2023)
Given a fixed t, there exists a constant $C > 0$ depending only on g, such that:

- *If* $\alpha \in (1, 2)$,

$$d\left(Z_t^{(n)}, \int_0^t \sqrt{\lim_k \mathfrak{g}_k(X_s)}W(ds)\right) \leq C n^{\frac{1}{2}-\frac{1}{\alpha}}.$$

- *If* $\alpha = 1$,

$$d\left(Z_t^{(n)}, \int_0^t \sqrt{\lim_k \mathfrak{g}_k(X_s)}W(ds)\right) \leq C n^{-\frac{1}{2}} \log(n).$$

- *If* $\alpha \in (0, 1)$,

$$d\left(Z_t^{(n)}, \int_0^t \sqrt{\lim_k \mathfrak{g}_k(X_s)}W(ds)\right) \leq C n^{-\frac{1}{2}}.$$

Main results (part II)

Theorem (Amorino, Jaramillo, Podolskij, 2023)
If g is symmetric and t is fixed, then for all $\alpha \in (1,2]$, we have

$$d\left(Z_t^{(n)}, \int_0^t \sqrt{\lim_k \mathfrak{g}_k(X_s)}W(ds)\right) \leq Cn^{-\frac{1}{2}},$$

even for $\alpha = 2$.

Future work

The structure of independent and stationary increments is convenient but not essential.

- Move from high-frequency observations to spaced observations (observe X_k instead of $X_{k/n}$).

Future work

The structure of independent and stationary increments is convenient but not essential.

- Move from high-frequency observations to spaced observations (observe X_k instead of $X_{k/n}$).
- Remove stationarity of increments of X. Do we obtain integrals with time-changed Brownian motion?

Future work

The structure of independent and stationary increments is convenient but not essential.

- Move from high-frequency observations to spaced observations (observe X_k instead of $X_{k/n}$).
- Remove stationarity of increments of X. Do we obtain integrals with time-changed Brownian motion?
- Consider the case when X is a diffusion.

Future work

The structure of independent and stationary increments is convenient but not essential.

- Move from high-frequency observations to spaced observations (observe X_k instead of $X_{k/n}$).
- Remove stationarity of increments of X. Do we obtain integrals with time-changed Brownian motion?
- Consider the case when X is a diffusion.
- Remove independence of increments and consider:
  (i) Gaussianity, replacing independence by weak dependence.
  (ii) Replace independence with exchangeability properties.

Future work

The structure of independent and stationary increments is convenient but not essential.

- Move from high-frequency observations to spaced observations (observe X_k instead of $X_{k/n}$).
- Remove stationarity of increments of X. Do we obtain integrals with time-changed Brownian motion?
- Consider the case when X is a diffusion.
- Remove independence of increments and consider:
  (i) Gaussianity, replacing independence by weak dependence.
  (ii) Replace independence with exchangeability properties.
- Higher-order approximations (Edgeworth expansions).

Future work

The structure of independent and stationary increments is convenient but not essential.

- Move from high-frequency observations to spaced observations (observe X_k instead of $X_{k/n}$).
- Remove stationarity of increments of X. Do we obtain integrals with time-changed Brownian motion?
- Consider the case when X is a diffusion.
- Remove independence of increments and consider:
  (i) Gaussianity, replacing independence by weak dependence.
  (ii) Replace independence with exchangeability properties.
- Higher-order approximations (Edgeworth expansions).
- Allow more flexibility on g (e.g., $g(x) := \delta_0(x)$).

Future work

The structure of independent and stationary increments is convenient but not essential.

- Move from high-frequency observations to spaced observations (observe X_k instead of $X_{k/n}$).
- Remove stationarity of increments of X. Do we obtain integrals with time-changed Brownian motion?
- Consider the case when X is a diffusion.
- Remove independence of increments and consider:
  (i) Gaussianity, replacing independence by weak dependence.
  (ii) Replace independence with exchangeability properties.
- Higher-order approximations (Edgeworth expansions).
- Allow more flexibility on g (e.g., $g(x) := \delta_0(x)$).
- Understand the role of regularity of g (comparison with ItÃť integration).

Gracias!

Contacto
Arturo Jaramillo
jagil@cimat.mx

References

Chiara Amorino, Arturo Jaramillo, Mark Podolskij. Quantitative and stable limits of high-frequency statistics of Levy processes: a Stein's method approach.