A generalization of Itô calculus and large deviations theory

Sudip Sinha

2019-04-05

Advisors
Prof. Hui-Hsiung Kuo
Prof. Padmanabhan Sundar

Outline

Section 1

Introduction and motivation

Quick revision and notations

- Let $T \in (0, \infty)$, and denote $t \in [0, T]$.
- Let $(\Omega, \mathcal{F}, \mathcal{F}_\bullet, \mathbb{P})$ be a filtered probability space.
- $B_\bullet$ is a Brownian motion on $(\Omega, \mathcal{F}, \mathcal{F}_\bullet, \mathbb{P})$.
- Properties of $B_\bullet$
  - starts at 0
  - has independent increments
  - $B_t - B_s \sim \mathcal{N}(0, t - s)$
  - continuous paths (a.s.)
  - is a.s. nowhere differentiable
  - has unbounded linear variation ☹
  - has bounded quadratic variation ☺
  - $\mathbb{E}(B_t B_s) = s \wedge t$
  - martingale
- Naive integration w.r.t. B_t: not possible.
- A stochastic process $X_\bullet$ is called adapted to $\mathcal{F}_\bullet$ if X_t is measurable w.r.t. $\mathcal{F}_t$ $\forall t$.

Martingales and Markov processes

Definition Let $X_\bullet$ is a L^1-bounded $\mathcal{F}_\bullet$-adapted stochastic process, and let $0 \le s \le t \le T$. Then $X_\bullet$ is called a

- martingale if $\mathbb{E}(X_t \mid \mathcal{F}_s) = X_s$,
- submartingale if $\mathbb{E}(X_t \mid \mathcal{F}_s) \ge X_s$, and
- supermartingale if $\mathbb{E}(X_t \mid \mathcal{F}_s) \le X_s$.

Remark If (X_n) is a discrete-time martingale and (H_n) is an adapted process, then the process $Y_n = \sum_{j=0}^{n-1} H_j(X_{j+1} - X_j) =: (H \bullet X)_n$ is itself a martingale and called a martingale transform of (X_n).

Definition A stochastic process $X_\bullet$ is called Markov if for any $0 \le s \le t \le T$, we have

$$\mathbb{P}(X_t \in E \mid \mathcal{F}_s) = \mathbb{P}(X_t \in E \mid X_s).$$

Wiener integral for $f \in L^2[0,T]$

- Definition of the integral:

  1. Step functions $f = \sum_{j=0}^{n-1} c_j \mathbb{1}_{[t_j, t_{j+1})}(t)$: Define $\int_0^T f(t)\,\mathrm{d}B_t = \sum_{j=0}^{n-1} c_j \Delta B_j$, where $\Delta B_j = B_{t_{j+1}} - B_{t_j}$.
  2. $f \in L^2[0,T]$: Use step functions approximating f to extend the integral a.s.

- Properties of the integral:

  - Linear.
  - Gaussian distribution with mean 0 and variance $\|f\|^2_{L^2[0,T]}$ (Itô isometry).
  - Agrees with the Riemann–Stieltjes integral for continuous functions of bounded variation.

- Properties of the associated process $I_\bullet = \int_0^\bullet f(t)\,\mathrm{d}B_t$:

  - continuity.
  - martingale.

- Problem: Cannot integrate stochastic processes.

Trying to integrate stochastic processes

- $\int_0^T B_t \, \mathrm{d}B_t \stackrel{?}{=}$
  Since B_t is continuous, let us try Riemann–Stieltjes integral. Consider a sequence of partitions Δ_n such that $\|\Delta_n\| \to 0$. Then

$$\int_0^T B_t \, \mathrm{d}B_t = \lim \sum_{j=0}^{n-1} B_{t_j^*} \Delta B_j.$$

- Choosing different endpoints for t_j^* gives us different results.

t_j^*	$\int_0^t B_s \, \mathrm{d}B_s$	Intuitive?	$\mathbb{E}$	Martingale?	Theory
left	$\frac{1}{2}\left(B_t^2 - t\right)$	☹	0	☺	Itô
mid	$\frac{1}{2}\left(B_t^2\right)$	☺	$\frac{1}{2}t$	☹	Stratonovich
right	$\frac{1}{2}\left(B_t^2 + t\right)$	☹	t	☹	

- Which one do we choose?

Itô integral [Itô44] for $X_\bullet \in L^2_{\text{ad}}([0,T]\times\Omega)$

▷ Definition of the integral:

1. Adapted step processes $X_t(\omega) = \sum_{j=0}^{n-1} \xi_j(\omega)\mathbb{1}_{[t_j,t_{j+1})}(t)$: define $\int_0^T X_t\,\mathrm{d}B_t = \sum_{j=0}^{n-1}\xi_j \Delta B_j$.
2. $X \in L^2_{\text{ad}}([0,T]\times\Omega)$: use step processes approximating X to extend the integral in $L^2(\Omega)$.

▷ Properties of the integral:

- Linear.
- Mean 0 and variance $\|f\|^2_{L^2[0,T]}$ (Itô isometry).
- For $X_\bullet$ continuous, $\int_0^T X_t\,\mathrm{d}B_t = \lim \int_0^T X_{\lfloor \frac{tn}{n} \rfloor}\,\mathrm{d}B_t = \lim \sum_{j=0}^{n-1} X_{t_j}\Delta B_j$.

▷ Properties of the associated process $I_\bullet = \int_0^\bullet X_t\,\mathrm{d}B_t$:

- continuity.
- martingale.

▷ Example: $\int_0^t B_u\,\mathrm{d}B_u = \frac{1}{2}(B_t^2 - t) \quad \forall t$.

Itô integral for $X_\cdot$ such that $\int_0^T X_t^2 \,\mathrm{d}t < \infty$ a.s.

- Definition: Use sequences of processes in $L^2_{\mathrm{ad}}([0,T] \times \Omega)$ approximating X in probability to extend the integral in probability.

- Properties of the integral:

  - Linear.
  - Mean and variance? ☹

- Properties of the associated process $I_\bullet = \int_0^\bullet X_t \,\mathrm{d}B_t$:

  - continuity.
  - local martingale.

- Example: $\int_0^T e^{B_t^2} \,\mathrm{d}B_t = \int_0^{B_T} e^{t^2} \,\mathrm{d}t - \int_0^T B_t e^{B_t^2} \,\mathrm{d}t$.

The Itô formula

- An Itô process is a process of the form $X_\bullet = X_0 + \int_0^\bullet m_t \, \mathrm{d}t + \int_0^\bullet \sigma_t \, \mathrm{d}B_t$.
  Equivalently expressed as $\mathrm{d}X_t = m_t \, \mathrm{d}t + \sigma_t \, \mathrm{d}B_t$.

Theorem ([Itô44]) Let X_t be a d-dimensional Itô process, and assume $f \in C^{1,2}(\mathbb{R} \times \mathbb{R})$. Then $f(t, X_t)$ is also a d-dimensional Itô process given by

$$\mathrm{d}f(t, X_t) = \frac{\partial f}{\partial t}(t, X_t) \, \mathrm{d}t + \langle (\mathrm{D}f)(t, X_t), \mathrm{d}X_t \rangle + \frac{1}{2} \Big\langle \mathrm{d}X_t, (D^2 f)(t, X_t) \, \mathrm{d}X_t \Big\rangle ,$$

where we use the rule $\mathrm{d}B_t \otimes \mathrm{d}B_t = I_d \, \mathrm{d}t$.

- Example: For σ constant, $\mathcal{E}_t = \exp\left(\sigma B_t - \frac{1}{2}\sigma^2 t\right)$, $\mathrm{d}\mathcal{E}_t = -\frac{1}{2}\sigma^2 \mathcal{E}_t \, \mathrm{d}t + \sigma \mathcal{E}_t \, \mathrm{d}B_t + \frac{1}{2}\sigma^2 \mathcal{E}_t (\mathrm{d}B_t)^2$.

Exponential processes and the Girsanov theorem

- Let $h_\bullet$ be a stochastic process. The associated exponential process is defined as

$$\mathcal{E}_\bullet^{(h)} = \exp\left(\int_0^\bullet h_t \, \mathrm{d}B_t - \frac{1}{2} \int_0^\bullet h_t^2 \, \mathrm{d}t \right).$$

- The exponential process is a martingale if and only if $\mathbb{E}\mathcal{E}_t^{(h)} = 1 \ \forall t$.

- The Novikov condition: The exponential process is a martingale if $\mathbb{E} \exp\left(\frac{1}{2}\int_0^T h_t^2 \, \mathrm{d}t\right) < \infty$.

- The Girsanov theorem [Gir60]: The translated stochastic process $W_\bullet = B_\bullet + \int_0^\bullet h_t \, \mathrm{d}t$ is a Brownian motion under the probability measure $\widetilde{\mathbb{P}}$ defined by the Radon-Nikodym derivative $\frac{\mathrm{d}\widetilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}} = \mathcal{E}_T^{(h)}$.

Stochastic differential equations

- Let $\xi \in L^2(\Omega)$ be independent of $B_\bullet$, and $m, \sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$ have 'nice' measurability. Then a $\mathcal{F}_t$-adapted stochastic process X_t is called a solution of the stochastic *integral* equation $X_t = \xi + \int_0^t m(s, X_s)\, \mathrm{d}s + \int_0^t \sigma(s, X_s)\, \mathrm{d}B_s$ if for each t, the X_t satisfies the integral equation a.s.

- Stochastic *differential* equation $\mathrm{d}X_t = m(t, X_t)\, \mathrm{d}t + \sigma(t, X_t)\, \mathrm{d}B_t,\ X_0 = \xi$ is a *formal representation*.

Theorem (Existence and uniqueness, Markov property) The SDE above has a unique solution if there exists an $M > 0$ such that the following two conditions are satisfied:

- (Lipschitz condition) $|m(t,x) - m(t,y)|^2 + |\sigma(t,x) - \sigma(t,y)|^2 \leq M\,|x-y|^2$ a.s.
- (growth condition) $|m(t,x)|^2 + |\sigma(t,y)|^2 \leq M\left(1 + |x|^2\right)$ a.s.

The solution is a Markov process.

Moreover if $\xi \in \mathbb{R}$ and m, σ are function of only x, then the solution is also stationary.

- Example: For σ constant, $\mathrm{d}\mathcal{E}_t = \sigma\mathcal{E}_t\, \mathrm{d}B_t, \mathcal{E}_0 = 1$ is solved by $\mathcal{E}_t = \exp\left(\sigma B_t - \frac{1}{2}\sigma^2 t\right)$.

Multiple Wiener–Itô integrals

- How do we define the double integral?

- Naive idea: $\int_0^t \int_0^t \mathrm{d}B_u \, \mathrm{d}B_v = \int_0^t \mathrm{d}B_u \int_0^t \mathrm{d}B_v = B_t^2$.
  But $\mathbb{E}B_t^2 = t \neq 0$, so no martingale property. ☹

- Itô's idea: remove the diagonal to get

$$\int_0^t \int_0^t \mathrm{d}B_u \, \mathrm{d}B_v = 2 \int_0^t \int_0^v \mathrm{d}B_u \, \mathrm{d}B_v = 2 \int_0^t B_v \, \mathrm{d}B_v = B_t^2 - t.$$

Theorem ([Itô51]) Let $f \in L^2\left([0,T]^n\right)$ and $\hat{f}$ be its symmetrization. Then

$$\int_{[0,T]^n} f(t_1, \ldots, t_n) \, \mathrm{d}B_{t_1} \cdots \mathrm{d}B_{t_n} = n! \int_0^T \cdots \int_0^{t_{n-2}} \left(\int_0^{t_{n-1}} \hat{f}(t_1, \ldots, t_n) \, \mathrm{d}B_{t_n} \right) \mathrm{d}B_{t_{n-1}} \cdots \mathrm{d}B_{t_1}.$$

Section 2

A Generalization of Itô calculus

Motivation

- Iterated integrals: Consider the iterated integral $\int_0^T \int_0^T \mathrm{d}B_s\,\mathrm{d}B_t = \int_0^T B_T\,\mathrm{d}B_t \overset{?}{=} B_T^2$.
- Note that $\mathbb{E}(B_T B_t) = T \wedge t = t \neq 0$, so no martingale property ☹.
- Stochastic differential equations with anticipation:

$$\begin{aligned} \mathrm{d}X_t &= X_t\,\mathrm{d}B_t & \qquad \mathrm{d}Y_t &= B_T\,\mathrm{d}B_t \\ X_0 &= B_T & \qquad Y_0 &= 1 \end{aligned}$$

- Problem: We want to define $\int_0^T Z(\bullet)\,\mathrm{d}B_t$, where $Z(\bullet)$ is not (necessarily) adapted.
- Some approaches:
  - Enlargement of filtration $\mathcal{G}_\bullet = \mathcal{F}_\bullet \vee B_T$, with Itô's decomposition of integrand [Itô78] $B_t = \left(B_t - \int_0^t \frac{B_T - B_s}{T-s}\,\mathrm{d}s\right) + \int_0^t \frac{B_T - B_s}{T-s}\,\mathrm{d}s$.
  - White noise theory
  - Malliavin calculus

The new integral [AK08; AK10]: Idea

- A process $Y^{\bullet}$ and filtration $\mathcal{F}_{\bullet}$ are called instantly independent if Y^t and $\mathcal{F}_t$ are independent $\forall t$. Example: The process $(B_T - B_{\bullet})$ is instantly independent of the filtration generated by $B_{\bullet}$.
- Ideas
  1. Decompose the integrand into adapted and instantly independent parts.
  2. Evaluate the adapted and the instantly independent parts at the left and right endpoints.
- Consider two continuous stochastic processes, X_t adapted and Y^t instantly independent w.r.t. $\mathcal{F}_{\bullet}$. Then the integral $\int_0^T X_t Y^t \, \mathrm{d}B_t$ is defined as

$$\int_0^T X_t Y^t \, \mathrm{d}B_t \triangleq \lim_{\|\Delta_n\| \to 0} \sum_{j=0}^{n-1} X_{t_j} Y^{t_{j+1}} \Delta B_j,$$

  provided that the limit exists in probability.
- Now, for any stochastic process $Z(t) = \sum_{k=1}^{n} X_t^{(k)} Y_{(k)}^t$ we extend the definition by linearity.
- This is well-defined [HKS+16].

A simple example

- In the following, lim is the limit in L^2.

$$
\begin{aligned}
\int_0^t B_T \,\mathrm{d}B_t &= \int_0^t (B_t + (B_T - B_t))\,\mathrm{d}B_t = \int_0^t B_t \,\mathrm{d}B_t + \int_0^t (B_T - B_t)\,\mathrm{d}B_t \\
&= \lim \sum_{j=0}^{n-1} B_{t_j} \Delta B_j + \lim \sum_{j=0}^{n-1} (B_T - B_{t_{j+1}}) \Delta B_j \\
&= \lim \sum_{j=0}^{n-1} \left(B_T - \Delta B_j\right) \Delta B_j \\
&= B_T \lim \sum_{j=0}^{n-1} \Delta B_j - \lim \sum_{j=0}^{n-1} (\Delta B_j)^2 = B_T B_t - t
\end{aligned}
$$

- Note that $\mathbb{E}(B_T B_t - t) = 0$.
- In general, $\mathbb{E} \int_0^t Z(s)\,\mathrm{d}B_s = 0$. ☺

The near-martingale property

- Question: What are the analogues of the martingale property and the Markov property?
- Example: $\mathbb{E}(B_T B_t - t \mid \mathcal{F}_s) = B_s^2 - s \neq B_T B_s - s$ ☹
  But $\mathbb{E}(B_T B_s - s \mid \mathcal{F}_s) = B_s^2 - s$. ☺
- Let $Z(t)$ be a process such that $\mathbb{E}\,|Z(t)| < \infty\ \forall t$, and $0 \leq s \leq t \leq T$. Then $Z(t)$ is called a
  - near-martingale if $\mathbb{E}(Z(t) \mid \mathcal{F}_s) = \mathbb{E}(Z(s) \mid \mathcal{F}_s)$,
  - near-submartingale if $\mathbb{E}(Z(t) \mid \mathcal{F}_s) \geq \mathbb{E}(Z(s) \mid \mathcal{F}_s)$, and
  - near-supermartingale if $\mathbb{E}(Z(t) \mid \mathcal{F}_s) \leq \mathbb{E}(Z(s) \mid \mathcal{F}_s)$.

Theorem ([KSS12b]) Let f, ϕ be continuous functions on $\mathbb{R}$. Under integrability conditions, the processes $X_\bullet = \int_0^\bullet f(B_t)\phi(B_T - B_t)\,\mathrm{d}B_t$ and $Y^\bullet = \int_\bullet^T f(B_t)\phi(B_T - B_t)\,\mathrm{d}B_t$ are near-martingales.

Theorem ([HKS+17]) Let $Z(\cdot)$ be a stochastic process bounded in L^1, and $X_\bullet = \mathbb{E}(Z(\cdot) \mid \mathcal{F}_\bullet)$. Then $X_\bullet$ is a (sub/super)martingale if and only if $Z(\cdot)$ is a near-(sub/super)martingale.

A generalized Itô formula [HKS+16]

Process	Definition	Representation
Itô	$X_\bullet = X_0 + \int_0^\bullet m_t \, \mathrm{d}t + \int_0^\bullet \sigma_t \, \mathrm{d}B_t$	$\mathrm{d}X_t = m_t \, \mathrm{d}t + \sigma_t \, \mathrm{d}B_t$
instantly independent	$Y^\bullet = Y^0 + \int_\bullet^T \eta^t \, \mathrm{d}t + \int_\bullet^T \varsigma^t \, \mathrm{d}B_t$	$\mathrm{d}Y^t = -\eta^t \, \mathrm{d}t - \varsigma^t \, \mathrm{d}B_t$

Here η^t and ς^t are instantly independent such that Y^t is also instantly independent.

Theorem ([HKS+16]) Let $\mathrm{d}X_t = m_t \, \mathrm{d}t + \sigma_t \, \mathrm{d}B_t$ be an d-dimensional Itô process, and $\mathrm{d}Y^t = -\eta^t \, \mathrm{d}t - \varsigma^t \, \mathrm{d}B_t$ be a $\tilde{d}$-dimensional instantly independent process. If $f(t,x,y) \in C^{1,2,2}(\mathbb{R} \times \mathbb{R} \times \mathbb{R})$, then

$$\mathrm{d}f(t, X_t, Y^t) = \frac{\partial f}{\partial t}(t, X_t, Y^t) \, \mathrm{d}t + \langle (\mathrm{D}_x f)(t, X_t, Y^t), \mathrm{d}X_t \rangle + \frac{1}{2} \left\langle \mathrm{d}X_t, (D_x^2 f)(t, X_t, Y^t) \, \mathrm{d}X_t \right\rangle$$
$$+ \left\langle (\mathrm{D}_y f)(t, X_t, Y^t), \mathrm{d}Y^t \right\rangle - \frac{1}{2} \left\langle \mathrm{d}Y^t, (D_y^2 f)(t, X_t, Y^t) \, \mathrm{d}Y^t \right\rangle,$$

where we use the rules $\mathrm{d}B_t \otimes \mathrm{d}B_t = I_d \, \mathrm{d}t$.

Iterated integrals

Theorem ([Itô51]) Let $f \in L^2([0,T]^n)$ and $\hat{f}$ be its symmetrization. Then

$$\int_{[0,T]^n} f(t_1,\ldots,t_n)\,\mathrm{d}B_{t_1}\ldots\mathrm{d}B_{t_n} = n!\int_0^T \cdots \int_0^{t_{n-2}} \left(\int_0^{t_{n-1}} \hat{f}(t_1,\ldots,t_n)\,\mathrm{d}B_{t_n} \right) \mathrm{d}B_{t_{n-1}}\ldots\mathrm{d}B_{t_1}.$$

Theorem ([AK10]) Let $f \in L^2([0,T]^n)$. Then

$$\int_{[0,T]^n} f(t_1,\ldots,t_n)\,\mathrm{d}B_{t_1}\ldots\mathrm{d}B_{t_n} = \int_0^T \cdots \int_0^T f(t_1,\ldots,t_n)\,\mathrm{d}B_{t_n}\ldots\mathrm{d}B_{t_1}.$$

Example[HKS+16]: For the new integral, $\int_0^T \left(\int_0^T B_u\,\mathrm{d}u \right) \mathrm{d}B_v = \int_0^T \left(\int_0^T B_u\,\mathrm{d}B_v \right) \mathrm{d}u$.

A generalization of Itô isometry

Theorem ([KSS12b]) Let ϕ be an analytic function on $\mathbb{R}$. Then under integrability conditions and for each t,

$$\mathbb{E}\left[\left(\int_0^t \phi(B_T - B_s)\,\mathrm{d}B_s\right)^2\right] = \int_0^t \mathbb{E}\left[(\phi(B_T - B_s))^2\right]\mathrm{d}s$$

Theorem ([KSS13]) Let f, ϕ be C^1 function on $\mathbb{R}$. Then

$$\mathbb{E}\left[\left(\int_0^T f(B_t)\phi(B_T - B_t)\,\mathrm{d}B_t\right)^2\right] = \int_0^T \mathbb{E}\left[(f(B_t)\phi(B_T - B_t))^2\right]\mathrm{d}t$$

$$+ 2\int_0^T \int_0^t \mathbb{E}\left[f(B_s)\phi'(B_T - B_s)f'(B_s)\phi(B_T - B_s)\right]\mathrm{d}s\,\mathrm{d}t.$$

A generalization of Girsanov theorem

Theorem ([KPS13]) Let $X_\bullet$ and $Y^\bullet$ be continuous square-integrable stochastic processes such that $X_\bullet$ is adapted and $Y^\bullet$ is instantly independent.

Then translated stochastic process $W_\bullet = B_\bullet + \int_0^\bullet (X_t + Y^t)\,\mathrm{d}t$ is a near-martingale under the probability measure $\widetilde{\mathbb{P}}$ defined by the Radon-Nikodym derivative $\frac{\mathrm{d}\widetilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}} = \mathcal{E}_T^{(X+Y)}$.

Section 3
Large deviations theory

What is it about?

- A theory to find probabilities of rare events that decay exponentially.
- Started by Swedish actuarials Fredrik Esscher, Harald Cramér, Filip Lundberg.
- Unified by Varadhan in his 1966 paper [Var66].
- Example: A problem faced by the insurance industry.
  - Value of claims received on the nth day: X_n \$.
  - Steady income from premium: x \$/day.
  - Planning period: n days.
  - Average expenditure: $\overline{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$ \$/day.
  - *Question*: How should the company decide on the premium?
  - *Idea*: Determine x such that $\mathbb{P}\left\{\overline{X}_n > x\right\} < \varepsilon$ (specified).

Insurance problem: setup

1. Let the following hold:

    - $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space.
    - (X_n) is a sequence of i.i.d. random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with finite moment generating function M.
    - $\mathbb{E}X_1 = m$, $\mathbb{V}X_1 = \sigma^2$, and $X_1 \sim \mu$.
    - $\overline{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$.

2. Asymptotic behavior of $\overline{X}_n$:

    - Weak law of large numbers: $\overline{X}_n \xrightarrow{\mathbb{P}} m$.
    - Central limit theorem: $\sqrt{n}(\overline{X}_n - m) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2)$.

3. What is the rate for LLN?

Insurance problem: large deviation bounds

1. For $x > m$ and an arbitrary $\theta > 0$, we get

$$\mathbb{P}\left\{\overline{X}_n \geq x\right\} = \mathbb{P}\left\{e^{\theta n \overline{X}_n} \geq e^{\theta n x}\right\} \leq e^{-\theta n x}\mathbb{E}\left(e^{\theta n \overline{X}_n}\right) = e^{-\theta n x} M_X(\theta)^n = e^{-n(\theta x - \log M_X(\theta))}.$$

2. Since θ was arbitrary, we have

$$\mathbb{P}\left\{\overline{X}_n \geq x\right\} \leq \inf_\theta e^{-n(\theta x - \log M_X(\theta))} = e^{-n \sup_\theta (\theta x - \log M_X(\theta))} =: e^{-nI(x)}.$$

3. Generalizing, we get the large deviation upper bound

$$\overline{\lim} \frac{1}{n} \log \mathbb{P}\left\{\overline{X}_n \in F\right\} \leq -\inf_F I \qquad \forall F \text{ closed.}$$

4. We can also obtain a large deviation lower bound using an exponential change of measure

$$\underline{\lim} \frac{1}{n} \log \mathbb{P}\left\{\overline{X}_n \in G\right\} \geq -\inf_G I \qquad \forall G \text{ open.}$$

5. We informally write $\mathbb{P}\left\{\overline{X}_n \in \, \mathrm{d}x\right\} \asymp e^{-nI(x)}\,\mathrm{d}x$ for $x \in \mathbb{R}$.

Definition of large deviation principle

- The setup: (X_n) is a stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in a Polish space $(\mathcal{X}, d)$.
- A function $I : \mathcal{X} \to [0, \infty]$ is called a rate function if it has compact lower level sets.

Definition (X_n) is said to satisfy the large deviation principle on $\mathcal{X}$ with rate function I if the large deviation upper and lower bounds hold.

Example

Theorem ([Cra38]) Let (X_n) be a sequence of i.i.d. real random variables with finite moment generating function M. Then (X_n) follows large deviation principle with rate function $I(x) = \sup_\theta (\theta x - \log M(\theta))$.

Applications of the Cramér theorem

Rate functions for some common distributions

Distribution	$M(\theta)$	$I(x)$
Bernoulli(p)	$1-p+pe^{\theta}$	$\left(x\log x+(1-x)\log(1-x)-\left(x\log\frac{1-p}{p}+\log p\right)\right)\mathbb{1}_{[0,1]}(x)+\infty\mathbb{1}_{[0,1]^{\complement}}(x)$
Poisson(λ)	$e^{\lambda(e^{\theta}-1)}$	$\left(\lambda-x+x\log\frac{x}{\lambda}\right)\mathbb{1}_{[0,\infty)}(x)+\infty\mathbb{1}_{(-\infty,0)}(x)$
Exp(λ)	$\left(1-\frac{\theta}{\lambda}\right)^{-1}$	$(\lambda x-1+x\log(\lambda x))\,\mathbb{1}_{[0,\infty)}(x)+\infty\mathbb{1}_{(-\infty,0)}(x)$
$\mathcal{N}(m,\sigma^2)$	$e^{m\theta+\frac{1}{2}\sigma^2\theta^2}$	$\frac{(x-m)^2}{2\sigma^2}$
$\chi^2(k)$	$(1-2\theta)^{-\frac{k}{2}}$	$\frac{1}{2}\left(x-k+k\log\frac{k}{x}\right)$

The Schilder theorem

- Aim: Estimate the probability that a scaled-down sample path of a Brownian motion will stray far from the mean path.
- Let C_x denote the set of continuous functions from $[0, T]$ to $\mathbb{R}^d$ starting at x, and let $\mathrm{CM}_x = \{\omega \in C_x : \omega \text{ is absolutely continuous and } \omega_t' \in L^2[0, T]\}$.
- Let $B_\bullet$ be a d-dimensional Brownian motion, so $B_\bullet \in C_0 = C_0([0, T]; \mathbb{R}^d)$
- Let $\frac{1}{\sqrt{n}} B_t \sim W^{(n)}$. Then $W^{(n)} = \mathcal{N}\left(0, \frac{t}{n}\right) \xrightarrow{\mathcal{D}} \delta_0$ as $n \to \infty$.

Theorem ([Sch66]) On the Banach space $(C_0, \|\cdot\|_\infty)$, the sequence of probability measures $\left(W^{(n)}\right)$ satisfies LDP with the rate function $I : C_0 \to \overline{\mathbb{R}}$ given by

$$I(\omega) = \left(\frac{1}{2} \int_0^T |\omega_t'|^2 \, \mathrm{d}t\right) \mathbb{1}_{\mathrm{CM}_0}(\omega) + \infty \mathbb{1}_{\mathrm{CM}_0^\complement}(\omega)$$

The Freidlin–Wentzell theorem

- Aim: Estimate the probability that a scaled-down sample path of an Itô diffusion will stray far from the mean path.
- Let $X^{(n)}_\bullet$ be the solution of the d-dimensional stochastic differential equation $\mathrm{d}X^{(n)}_t = m(X^{(n)}_t)\,\mathrm{d}t + \frac{1}{\sqrt{n}}\sigma(X^{(n)}_t)\,\mathrm{d}B_t,\ X^{(n)}_0 = x$, where m, σ are sufficiently nice.
- Let $W^{(n)}_x$ denote the law of $X^{(n)}_\bullet$ starting at x.
- As $n \to \infty$, $W^{(n)}_x \xrightarrow{\mathcal{D}} \delta_\xi$, where ξ solves the ODE $\dot{\xi}(t) = m(\xi(t)),\ \xi(0) = x$.

Theorem ([FW12]) For any fixed x, the sequence of probability measures $\left(W^{(n)}_x\right)$ satisfies LDP with the rate function $I_x : C_0 \to \overline{\mathbb{R}}$ given by

$$I_x(\omega) = \left(\frac{1}{2}\int_0^T \left\langle \omega'_t - b(\omega_t), A^{-1}(\omega_t)\left(\omega'_t - b(\omega_t)\right)\right\rangle \mathrm{d}t\right) \mathbb{1}_{\mathrm{CM}_x}(\omega) + \infty \mathbb{1}_{\mathrm{CM}_x^\complement}(\omega),$$

where $A = \sigma\sigma^*$.

Section 4

The way forward

Possible research directions

- ▷ Identify the class of integrable processes under the new integral.
- ▷ Give a broader generalization of the Itô isometry for the new integral.
- ▷ Provide a broader generalization of the Girsanov theorem.
- ▷ Formulate the extension to stochastic differential equations with anticipating coefficients.
- ▷ Develop the near-Markov property for the new integral.
- ▷ Prove Freidlin–Wentzell type results for stochastic differential equations with anticipating initial conditions.
- ▷ Study LDP results for SDEs with anticipating coefficients.
- ▷ Analyze LDP for linear SPDEs with anticipating initial conditions.

Thank you!

Appendix

Laplace principle and equivalence to LDP

Definition (Laplace principle) (X_n) is said to satisfy the Laplace principle on $\mathcal{X}$ with rate function I if for all bounded continuous functions h, we have

$$\lim \frac{1}{n} \log \mathbb{E} \exp(-nh(X_n)) = \inf_{\mathcal{X}} (h + I)$$

Theorem (X_n) satisfies LP on $\mathcal{X}$ with rate function I if and only if (X_n) satisfies LDP on $\mathcal{X}$ with the same rate function I.

Some important results

- Uniqueness of the rate function.
- Continuity principle.
- Superexponential approximation preserves Laplace principle.

Thank you!

Bibliography

Ayed, W. & Kuo, H. H. (2008). An extension of the Itô integral. *Communications on Stochastic Analysis*, *2*(3). doi:10.31390/cosa.2.3.05

——— (2010). An extension of the Itô integral: Toward a general theory of stochastic integration. *Theory of Stochastic Processes*, *16*(32), 17–28. Retrieved from http://mi.mathnet.ru/thsp56

Cramér, H. (1938). Sur un nouveau théorème-limite de la théorie des probabilités. *Actualités Scientifiques et Industrielles*, *736*, 5–23.

Freidlin, M. I. & Wentzell, A. D. (2012). *Random Perturbations of Dynamical Systems*. (3 ed.). Springer-Verlag. doi:10.1007/978-3-642-25847-3

Girsanov, I. V. (1960). On Transforming a Certain Class of Stochastic Processes by Absolutely Continuous Substitution of Measures. *Theory of Probability & Its Applications*, *5*, 285–301. doi:10.1137/1105027

Hwang, C. R., Kuo, H. H., Saitô, K., & Zhai, J. (2016). A general Itô formula for adapted and instantly independent stochastic processes. *Communications on Stochastic Analysis*, *10*(3). doi:10.31390/cosa.10.3.05

——— (2017). Near-martingale Property of Anticipating Stochastic Integration. *Communications on Stochastic Analysis*, *11*(4). doi:10.31390/cosa.11.4.06

Itô, K. (1944). Stochastic integral. *Proc. Imp. Acad.*, *20*(8), 519–524. doi:10.3792/pia/1195572786

——— (1951). Multiple Wiener Integral. *J. Math. Soc. Japan*, *3*(1), 157–169. doi:10.2969/jmsj/00310157

——— (1978). Extension of stochastic integrals. In Extension of stochastic integrals., *Proceedings of the International Symposium on Stochastic Differential Equations*. Kinokuniya.

Kuo, H. H., Peng, Y., & Szozda, B. (2013). Generalization of the anticipative Girsanov theorem. *Communications on Stochastic Analysis*, *7*(4). doi:10.31390/cosa.7.4.06

Kuo, H. H., Sae-Tang, A., & Szozda, B. (2012b). A stochastic integral for adapted and instantly independent stochastic processes. In A stochastic integral for adapted and instantly independent stochastic processes., *Stochastic Processes, Finance and Control*. (Vol. 29, pp. 53–71). Author. doi:10.1142/9789814383318_0003

——— (2013). And isometry formula for a new stochastic integral. In And isometry formula for a new stochastic integral., *Quantum Probability and Related Topics*. (Vol. 29, pp. 222–232). Author. doi:10.1142/9789814447546_0014

Schilder, M. (1966). Some Asymptotic Formulas for Wiener Integrals. *Transactions of the American Mathematical Society*, *125*(1), 63–85. doi:10.2307/1994588

Varadhan, S. R. S. (1966). Asymptotic probabilities and differential equations. *Communications on Pure and Applied Mathematics*, *19*, 261–286. doi:https://doi.org/10.1002/cpa.3160190303