Grandes déviations pour une moyenne empirique pondérée et applications

J. Najim

Laboratoire Modal'X, Université Paris 10

200, av de la République, 92001 Nanterre Cedex

Objet de l'exposé

Principe de grandes déviations

Établir un principe de grandes déviations pour la moyenne empirique pondérée:

$$\langle L_n, \mathbf{f} \rangle = \frac{1}{n} \sum_1^n \mathbf{f}(x_i^n) \cdot Z_i,$$

où

- $\mathbf{f}$ continue bornée,
- Les v. a. $(Z_i)_{i \in \mathbb{N}}$ sont iid à valeurs dans $\mathbb{R}^d$ et vérifient:

$$E e^{\alpha |Z|} < +\infty \quad \text{pour un } \alpha > 0.$$

- Les x_i^n sont déterministes à valeurs dans $\mathcal{X}$ et satisfont:

$$\frac{1}{n} \sum_1^n \delta_{x_i^n} \xrightarrow[n \to \infty]{conv.etr.} R.$$

Applications

Applications à:

- la mesure empirique MEM $L_n = \frac{1}{n} \sum_1^n Z_i \delta_{x_i^n}$.
- certains processus empiriques

$$t \mapsto \bar{Z}_n(t) = \frac{1}{n} \sum_{i=1}^{[nt]} Z_i.$$

Plan

1. Rappels de grandes déviations

2. Grandes déviations pour la moyenne empirique

3. Applications

Rappels sur les grandes déviations

Définition

Y_n satisfait un PGD de fonction de taux I si

$$P\{Y_n \in A\} \asymp e^{-nI(A)} \quad \text{pour } n \to \infty$$

où $I(A) = \inf\{I(y),\ y \in A\}$.

Exemple

Moyenne empirique $\bar{X}_n = \frac{1}{n}\sum_1^n X_i$ de variables iid.
Si:

$$Ee^{\alpha|X|} < +\infty \quad \text{pour un } \alpha > 0,$$

alors

$$P\left\{\left|\frac{1}{n}\sum_1^n X_i - EX\right| > \epsilon\right\} \asymp e^{-n\Lambda^*(A)}$$

où

- $A = \{x,\ |x - EX| > \epsilon\}$,
- $\Lambda(\lambda) = \log Ee^{\lambda X}$,
- $\Lambda^*(x) = \sup_\lambda\{\lambda x - \Lambda(\lambda)\}$.

Vraie définition

Y_n satisfait un PGD de fonction de taux I si

$$\begin{aligned}\limsup_n \frac{1}{n}\log P\{Y_n \in A\} &\leq -I(\bar{A}) \\ \liminf_n \frac{1}{n}\log P\{Y_n \in A\} &\geq -I(\mathring{A})\end{aligned}$$

où

$$I(B) = \inf\{I(y),\ y \in B\}.$$

Remarque: si $I(\bar{A}) = I(\mathring{A}) = I_A$, alors

$$\lim_{n\to\infty} \frac{1}{n}\log P\{Y_n \in A\} = -I_A$$

Soit

$$P\{Y_n \in A\} \asymp_n e^{-nI_A}$$

Techniques d'établissement de la borne inférieure

on remarque qu'il suffit de montrer:

$$\liminf_n \frac{1}{n} \log P\{Y_n \in V_y\} \geq -I(y),$$

pour tout voisinage V_y de y.

1. Changement de probabilité exponentielle.

- Cramér '38
- Gärtner '77
- Ellis '84

Illustration dans le cas de la moyenne empirique iid $\frac{1}{n}\sum_1^n X_i$ où $X_i \sim P$. On veut minorer

$$P\{\bar{X}_n \in]x-\delta, x+\delta[\}.$$

Introduisons la **log-Laplace**

$$\Lambda(\lambda) = \log E e^{\lambda X_1},$$

et sa dérivée

$$\Lambda'(\lambda) = \frac{EX_1 e^{\lambda X_1}}{E e^{\lambda X_1}} = EX_1 e^{\lambda X_1 - \Lambda(\lambda)}.$$

On détermine λ_x solution de:

$$\Lambda'(\lambda) = x.$$

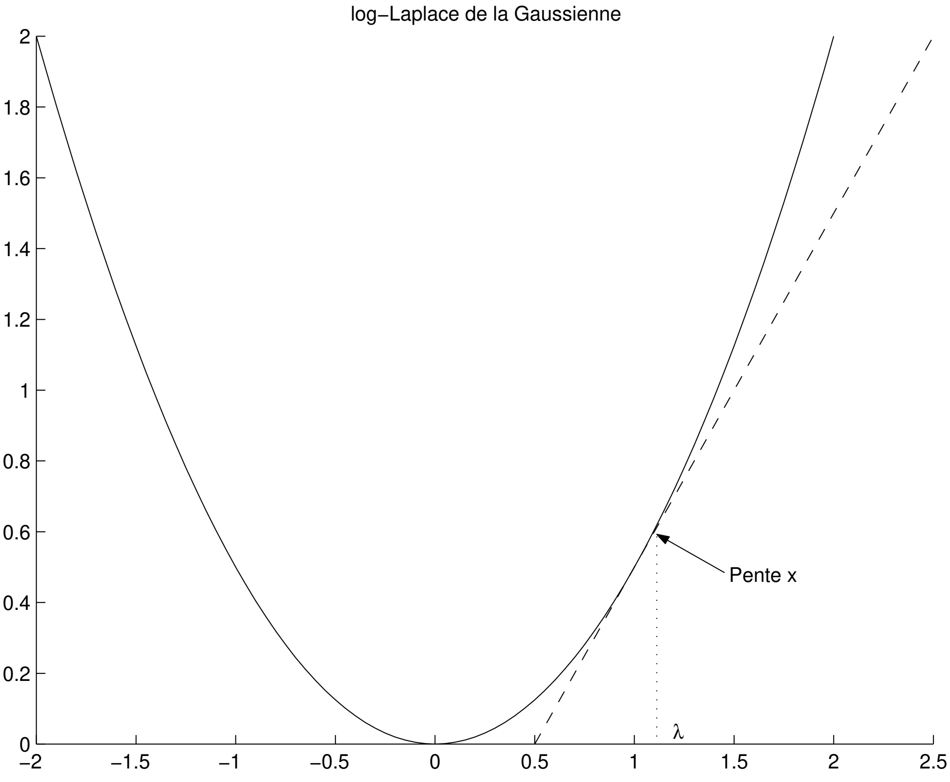

Figure 1: Log-Laplace de la Gaussienne

On introduit des variables auxiliaires Y_i telles que:

$$\frac{dP_Y}{dP}(u) = e^{\lambda_x u - \Lambda(\lambda_x)}.$$

Alors les Y_i sont **centrées en** x:

$$EY_1 = EX_1 e^{\lambda_x X_1 - \Lambda(\lambda_x)} = x$$

et

$$P\{X_1 \in]x-\delta, x+\delta[\}$$

$$\begin{aligned}
&= \int_{]x-\delta,x+\delta[} e^{\lambda_x u - \Lambda(\lambda_x) - (\lambda_x u - \Lambda(\lambda_x))} P(du) \\
&\geq e^{-(\lambda_x(x+\delta) - \Lambda(\lambda_x))} \int_{]x-\delta,x+\delta[} e^{\lambda_x u - \Lambda(\lambda_x)} P(du) \\
&\geq e^{-(\lambda_x(x+\delta) - \Lambda(\lambda_x))} P\{Y_1 \in]x-\delta, x+\delta[\}
\end{aligned}$$

soit, pour la moyenne empirique

$$P\{\bar{X}_n \in]x-\delta, x+\delta[\}$$

$$\geq e^{-n(\lambda_x(x+\delta) - \Lambda(\lambda_x))} P\{\bar{Y}_n \in]x-\delta, x+\delta[\}$$

en passant au logarithme,

$$\frac{1}{n}\log P\{\bar{X}_n \in]x-\delta, x+\delta[\}$$

$$\geq \quad -(\lambda_x(x+\delta) - \Lambda(\lambda_x))$$

$$+\frac{1}{n}\log P\{\bar{Y}_n \in]x-\delta, x+\delta[\}$$

Soit enfin,

$$\liminf_n \frac{1}{n}\log P\{\bar{X}_n \in]x-\delta, x+\delta[\}$$

$$\geq \quad -(\lambda_x(x-\delta) - \Lambda(\lambda_x))$$

Remarque: méthode très robuste, qui peut marcher dans le cas d'une **butée** si la log-Laplace est **escarpée**:

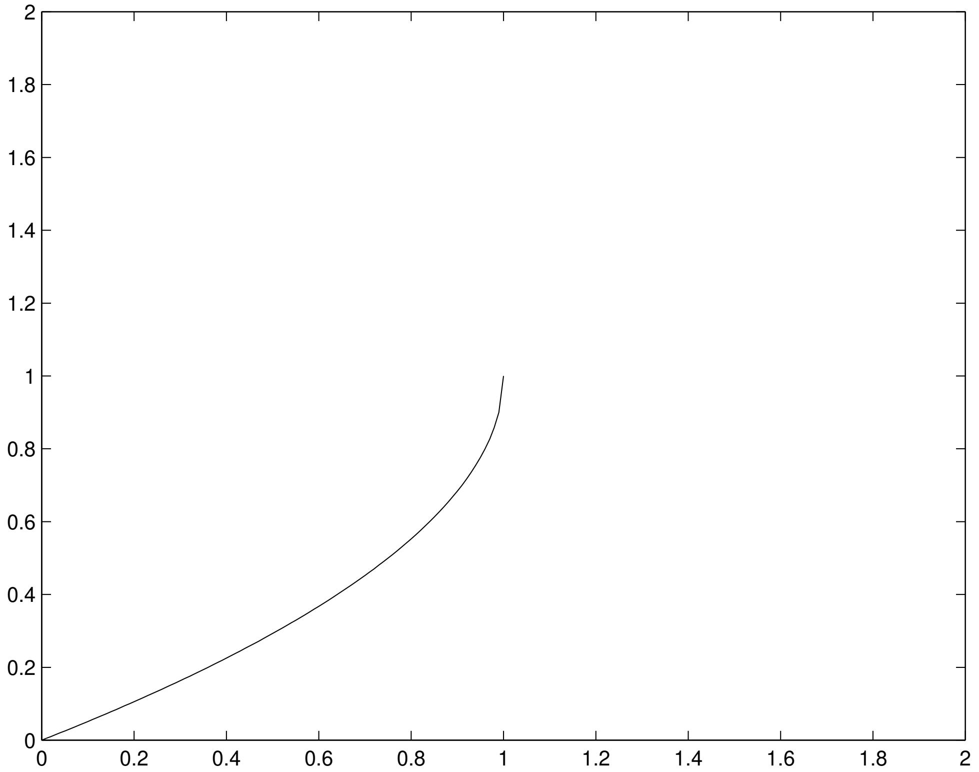

Figure 2: Log-Laplace escarpée

Correspond au cas où

$$Ee^{\alpha|Z_i|} < \infty \quad \text{pour un } \alpha > 0.$$

Il se peut que la butée ne soit pas escarpée:

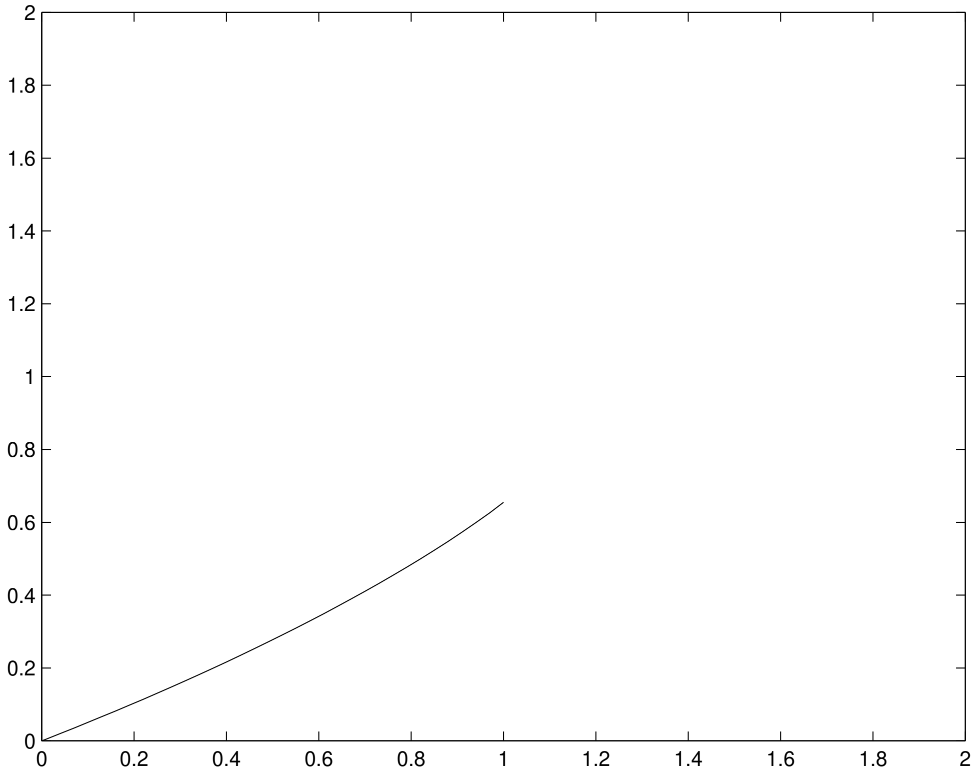

Figure 3: Log-Laplace non escarpée

- la technique de changement de proba exponentielle ne marche pas.
- le théorème de Cramér reste valide.
- sa démonstration est basée sur l'idée de **sous-additivité**.

2. Sous-additivité.

- Ruelle, Lanford '65, '67, '73,
- Bahadur, Zabell '79.

Remarquons que

$$\begin{aligned}\bar{X}_{n+m} &= \frac{n}{n+m}\left(\frac{1}{n}\sum_{1}^{n} X_i\right) \\ &\quad + \frac{m}{n+m}\left(\frac{1}{m}\sum_{1}^{m} X_{n+i}\right)\end{aligned}$$

et $]x-\delta, x+\delta[$ est convexe donc:

$$\begin{aligned}&\{\bar{X}_n \in]x-\delta, x+\delta[\} \\ &\cap \quad \{1/m\sum_{1}^{m} X_{n+i} \in]x-\delta, x+\delta[\} \\ &\qquad \subset \{\bar{X}_{n+m} \in]x-\delta, x+\delta[\}\end{aligned}$$

Par suite (indépendance et equidistribution):

$$\begin{aligned} P\{\bar{X}_n \in]x-\delta, x+\delta[\} \\ \times \quad P\{\bar{X}_m \in]x-\delta, x+\delta[\} \\ \leq P\{\bar{X}_{n+m} \in]x-\delta, x+\delta[\} \end{aligned}$$

et

$$f(n) = -\log P\{\bar{X}_n \in]x-\delta, x+\delta[\}$$

est une fonction **sous-additive**:

$$f(n) + f(m) \geq f(n+m),$$

ce qui assure la convergence de:

$$\frac{f(n)}{n} = -\frac{1}{n} \log P\{\bar{X}_n \in]x-\delta, x+\delta[\}$$

La convergence implique la minoration.

Remarque: La méthode nécessite seulement que:

$$Ee^{\alpha|Z|} < \infty \text{ pour un } \alpha > 0$$

Bilan: Méthode plus performante **mais** domaine d'application moindre.

Grandes déviations pour la moyenne empirique

Le modèle

$$\langle L_n, \mathbf{f} \rangle = \frac{1}{n} \sum_1^n \mathbf{f}(x_i^n) \cdot Z_i, = \begin{pmatrix} f_1(x_i^n) \cdot Z_i \\ \vdots \\ f_m(x_i^n) \cdot Z_i \end{pmatrix}$$

où

- $\mathbf{f}$ continue bornée,
- Les v. a. $(Z_i)_{i \in \mathbb{N}}$ sont iid à valeurs $\mathbb{R}^d$ et:

$$E e^{\alpha |Z|} < +\infty \quad \text{pour un } \alpha > 0.$$

- Les x_i^n sont déterministes à valeurs dans $\mathcal{X}$ et satisfont:

$$\frac{1}{n} \sum_1^n \delta_{x_i^n} \xrightarrow[n \to \infty]{conv.etr.} R.$$

Théorème

$\langle L_n, \mathbf{f} \rangle$ satisfait un PGD de fonction de taux:

$$I_{\mathbf{f}}(z) = \sup_{\lambda \in \mathbb{R}^m} \left\{ \lambda \cdot z - \int_{\mathcal{X}} \Lambda \Big[\sum_{j=1}^{m} \lambda_j f_j(x) \Big] \, R(dx) \right\}$$

pour $z \in \mathbb{R}^m$.

Éléments de preuve

L'idée consiste à approximer $\mathbf{f}$ par une fonction en escalier qui converge uniformément vers $\mathbf{f}$.

1. Théorème de Cramér 'rescalé'

Si

$$\frac{N_A(n)}{n} \to R_A > 0,$$

alors $\frac{1}{n} \sum_1^{N_A(n)} Z_i$ satisfait un PGD.

Preuve -suite-

2. PGD pour une fonction en escalier

Si

$$f(x) = a\,1_A(x) + b\,1_B(x)$$

avec:

$$R(A) > 0, \quad R(\partial A) = 0,$$

de même pour B. Alors $\langle f, L_n \rangle$ satisfait un PGD:

$$\langle f, L_n \rangle = \frac{a}{n} \sum_{x_i \in A} Z_i + \frac{b}{n} \sum_{x_i \in B} Z_i$$

et

$$\sum_{x_i \in A} Z_i = \sum_{1}^{N_A(n)} Z_{\sigma(i)}$$

où

$$\frac{N_A(n)}{n} = \frac{1}{n} \sum_{1}^{n} 1_A(x_i) \to R(A) > 0$$

on applique Cramér rescalé puis le principe de contraction.

Preuve -suite-

3. Approximation exponentielle

Soit des fonctions étagées f^ϵ telles que:

$$\|f^\epsilon - f\| \le \epsilon$$

Alors

$$\begin{aligned} &\limsup_n \frac{1}{n} \log P\{|\langle f, L_n\rangle - \langle f^\epsilon, L_n\rangle| > \delta\} \\ \le \quad & \limsup_n \frac{1}{n} \log P\{\frac{1}{n} \sum |Z_i| > \frac{\delta}{\epsilon}\} \\ = \quad & \alpha(\epsilon) \xrightarrow[\epsilon \to 0]{} -\infty \end{aligned}$$

4. Conclusion

$\langle f, L_n\rangle$ satisfait un PGD.

Contrexemples

Cas où l'hypothèse:

$$R(U) > 0 \quad \text{pour} \quad U \,\text{ouvert non vide}$$

n'est pas vérifiée. On considère

- la mesure deterministe

  $$R_n = \frac{1}{n} \sum_1^{n-1} \delta_{\frac{i}{n}} + \frac{\delta_2}{n} \in \mathcal{P}([0,2])$$

  alors

  $$R_n \to leb_{[0,1]} \in \mathcal{P}([0,2])$$

  qui ne satisfait pas l'hypothèse.

- les variables Z_i de loi:

  $$P\{Z = a_n\} = c\frac{e^{-a_n}}{1+a_n^3}, \quad a_n = 16^n.$$

Contrexemple 1:

Soit

$$f = \begin{cases} 0 & \text{sur } [0,1] \\ 1 & \text{en } 2. \end{cases}$$

Alors

$$\langle f, L_n \rangle = \frac{1}{n} \sum f(x_i) Z_i = \frac{Z_n}{n}$$

ne satisfait pas de PGD.

Contrexemple 2:

Soit

$$f = \begin{cases} 1 & \text{sur } [0,1] \\ -1 & \text{en } 2. \end{cases}$$

Alors

$$\langle f, L_n \rangle = \frac{1}{n} \sum f(x_i) Z_i = \frac{1}{n} \sum_{1}^{n-1} Z_i - \frac{Z_n}{n}$$

ne satisfait pas de PGD.

Applications

La mesure empirique MEM

On étudie les GD de

$$L_n = \frac{1}{n} \sum_{i=1}^{n} Z_i \delta_{x_i}$$

Intérêt:

- Théorie des gaz, Ellis, Gough, Pulé '92,
- Maximum d'Entropie en Moyenne, Gamboa & Gassiat '97, Csiszàr, G & G '99.

On suppose $\mathcal{X}$ compact.

Théorème

La mesure empirique L_n satisfait un PGD de fonction de taux:

$$I(\mu) = \int_{\mathcal{X}} \Lambda^* \left[\frac{d\mu_a}{dR}(x) \right] R(dx) + \int_{\mathcal{X}} \rho \left[\frac{d\mu_s}{d\theta}(x) \right] \theta(dx),$$

dans $M^d(\mathcal{X})$.

PGD pour certains processus empiriques

On s'intéresse à la ligne discontinue

$$t \mapsto \bar{Z}_n(t) = \frac{1}{n} \sum_{i=1}^{[nt]} Z_i$$

ainsi qu'à la ligne brisée

$$t \mapsto \tilde{Z}_n(t) = \bar{Z}_n(t) + \left(t - \frac{[nt]}{n}\right) Z_{[nt]+1}.$$

Grandes déviations étudiées par:

- Mogul'skii, '76.
- De Acosta, '94.

Théorème

$\bar{Z}_n$ et $\tilde{Z}_n$ satisfont un PGD dans l'ensemble des fonctions à variation bornée, de fonction de taux:

$$\Phi(f) = \int_{[0,1]} \Lambda^*(f_a'(t))\, dt + \int_{[0,1]} \rho(f_s'(t))\, d\theta(t)$$

Perspectives

Deux directions d'études se dégagent:

- étudier $\frac{1}{n}\sum_{i=1}^{n} Z_i f(x_i)$ lorsque les variables Z_i sont à valeurs dans un espace de dimension infinie.
- étudier $\frac{1}{n}\sum_{i=1}^{n} Z_i f(x_i)$ lorsque les Z_i demeurent des variables aléatoires indépendantes mais ne sont plus identiquement distribuées: la loi de Z_i dépend de x_i.