TP de Modélisation Stochastique n^0 10 : Principe de Grandes Déviations

Rappels

1. On considère une mesure de probabilité μ sur $\mathbb{R}$, intégrable. On suppose que X est une v.a. de loi μ. On appelle:
   - Transformée de Laplace de μ : $0 \leq L_\mu(\theta) = \int e^{\theta x}\mu(dx) = \mathbb{E}(e^{\theta X})$, pour θ dans un intervalle I_μ de $\mathbb{R}$ où cette transformée est finie.
   - Fonction des cumulants de μ : $\psi_\mu(\theta) = \log(L_\mu(\theta))$, pour $\theta \in I_\mu$.
   - Transformée de Cramer (Legendre) de μ : $h_\mu(x) = \sup_{\theta \in I_\mu} (\theta.x - \psi_\mu(\theta))$ pour $x \in \mathbb{R}$, fonction convexe, telle que si μ a pour moyenne m, alors h_μ décroît sur $]-\infty, m[$ et croît sur $]m, +\infty[$. Ainsi, pour:

     (a) X de loi de Bernoulli de paramètre p:

     $$h_\mu(x) = x\log(\frac{x}{p}) + (1-x)\log(\frac{1-x}{1-p}) \ \text{ si } x \in [0,1], \quad +\infty \text{ sinon.}$$

     (b) X de loi de Poisson de moyenne λ:

     $$h_\mu(x) = x\log(\frac{x}{\lambda}) + \lambda(1 - \frac{x}{\lambda}) \ \text{ si } x > 0, \quad +\infty \text{ sinon.}$$

     (c) X de loi gaussienne $\mathcal{N}(m, \sigma^2)$:

     $$h_\mu(x) = \frac{(x-m)^2}{2\sigma^2} \ \text{ pour } x \in \mathbb{R}.$$

     (d) X de loi exponentielle de moyenne λ:

     $$h_\mu(x) = \frac{x}{\lambda} - 1 - \log(\frac{x}{\lambda}) \ \text{ si } x > 0, \quad +\infty \text{ sinon.}$$

2. On suppose que $(X_1, X_2, .., X_n)$ est un n-échantillon de loi μ, de moyenne m finie. Soit la moyenne empirique, $S_n = \frac{1}{n}(X_1 + X_2 + .. + X_n)$. Alors,
   - Pour $a > m$, on a $\lim_{n\to+\infty} \frac{1}{n}\log(P(\frac{S_n}{n} > a)) = -h_\mu(a)$.
   - Pour $a < m$, on a $\lim_{n\to+\infty} \frac{1}{n}\log(P(\frac{S_n}{n} < a)) = -h_\mu(a)$.

Exercices sur le principe de grandes déviations

1. Retrouver la transformée de Cramer pour la loi de Bernoulli et la loi gaussienne en utilisant une méthode numérique (pour cela, on simule un échantillonde grande taille, on approche la transformée de Laplace à partir de moyenne, et on calcule le sup (utilser la fonction Matlab $fmin$) pour des x suffisament nombreux et bien répartis). Quelle est l'erreur commise sur ces calculs? Tracer ces transformées. Faire la même chose pour une loi binomiale, une loi de Weibull et un Khi-deux.

2. Comparer le logarithme de la probabilité d'un événement rare avec ses approximations par le Théorème Central Limite et par le principe de grandes déviations, dans le cas de la moyenne empirique d'un échantillon de v.a. de Bernoulli. Faire la même chose avec une loi de Poisson, puis un Khi-deux. Et avec une loi normale?

3. Comparer les intervalles de confiance de différents niveaux de confiance $(1-\alpha)$ obtenus par un calcul exact, puis approchés par un Théorème Central Limite et par un Principe de Grandes Déviations. On considère un n-échantillon de v.a.i.i.d. suivant une loi de Bernoulli de paramètre p (à choisir) et on estime p à l'aide de la moyenne empirique. Pour déterminer l'intervalle de confiance obtenu par PGD, il faut résoudre une équation dont les deux solutions seront les "quantiles" cherchés (pour chercher les zéros de cette fonction, on pourra utiliser la commande Matlab $fzero$). On tracera les différents intervalles de confiance. En jouant sur p et α, on pourra chercher les cas où le principe de grandes déviations est plus intéressant que le TLC. Faire la même chose avec une loi de Poisson, puis un Khi-deux.