Introduction à la théorie des extrêmes et des grandes déviations

TD7-MAPI3

2016-2017

1 Statistiques d'ordre et extrêmes

1.1 Étude empirique des lois limites des extrêmes

1.1.1 Étude théorique

On considère un échantillon $X_1, \cdots, X_n$, $(n \in \mathbb{N}^*)$ de variables aléatoires i.i.d. de densité f supportée par un intervalle $\mathbb{R}$. Déterminer la densité de $Z_n := \max_{i=1,\cdots,n} X_i$. Trouver des suites (a_n) et (b_n) telles que $a_n(Z_n - b_n)$ converge en loi dans les cas suivants :

1. f est la loi de Rayleigh, sur $\mathbb{R}^+$ $f(x) := x\exp(-x^2/2)$,
2. f est loi exponentielle de paramètre 1,
3. f est la loi uniforme sur $[0, 1]$,
4. f est la loi de Cauchy,
5. f est la loi de Pareto de densité sur $]1, +\infty[$ $f(x) := \alpha x^{-1-\alpha}$, $(\alpha > 0)$.

1.1.2 Étude Numérique

Simuler N échantillon de taille n de loi f et illustrer les résultats limites mis en évidence dans le paragraphe précédent.

1.2 Médiane

1.2.1 Médiane d'une loi

On se place dans le même cadre que celui de l'exercice précédent. On appelle médiane le plus petit réel μ tel que $\mathbb{P}(X_1 \le \mu) \ge 1/2$. Montrer que μ minimise la fonction $\varphi(x) := \mathbb{E}|X_1 - x|$, $(x \in \mathbb{R})$. Calculer les médianes des lois de l'exercice précédent.

1.2.2 Médiane empirique

soit μ_n l'unique minimiseur de la fonction

$$\varphi_n(x) := \frac{1}{n}\sum_{i=1}^{n} |X_i - x|, \quad (x \in \mathbb{R}).$$

Calculer μ_n. Montrer que μ_n converge en probabilité vers μ et que $\sqrt{n}(\mu_n - \mu)$ converge en loi vers une gaussienne centrée de variance à préciser.

1.2.3 Étude Numérique

Dans le cadre des lois de la section précédente, illustrer par des simulations les deux résultats limites sur la médiane empirique.

2 Grandes déviations et événements rares

2.1 Echantillonnage préférentiel et théorème de Cramér

Pour $n \geq 1$, on considère un échantillon $X_1 \ldots, X_n$ de variables i.i.d. de loi F. On prendra, par exemple, $n = 50, 100, 500$. Comme toujours, on pose $\overline{X_n} := n^{-1} \sum_{i=1}^{n} X_i$. On étudiera les cas suivants :

- F est la loi gaussienne standard,
- F est la loi exponentielle de paramètre 1,
- F est la loi de Poisson de paramètre 1,
- F est la loi de Bernoulli équilibrée,
- F est la loi géométrique de paramètre 1/2.

Pour $c > \mathbb{E}(X_1)$, on pose $p(c) := \mathbb{P}(\overline{X_n} \geq c)$. On souhaite ici comprendre expérimentalement à partir de quel niveau c le théorème de Bahadur-Rao donne une meilleure approximation que celui de la limite centrale.

1. Tracer à l'aide des fonctions `Scilab cdfbin, cdfgam, cdfnbn, cdfnor, cdfpoi` les courbes $p(c)$.
2. Reprendre la question précédente mais estimer les probabilités à l'aide d'une méthode de Monte Carlo *de base*.
3. Reprendre la question précédente mais estimer les probabilités à l'aide d'une méthode de Monte Carlo qui utilise le changement de probabilité des grandes déviations. Conclusions?
4. Tracer les approximations $\widehat{p}(c)$ et $\tilde{p}(c)$ obtenues par utilisation du théorème de Bahadur Rao et de la limite centrale. Conclusions?