Agrégation Interne de Mathématiques
Thierry Champion
séance du 17 novembre 2017

Maths 2 HEC 2017 : commentaires et compléments

1 Commentaires sur le sujet

Ce sujet est centré sur l'étude d'une loi de probabilité continue un peu particulière, la loi de Cauchy.

La partie I propose une démonstration probabiliste de la formule de Stirling, afin de se concentrer sur l'aspect "probas" je vous suggère en première lecture de sauter les questions 2 et 3.

La partie II est consacrée à la loi de Cauchy proprement dite, c'est un bon exercice pour se remettre dans le bain des variables aléatoires continues, à nouveau on peut sauter la question 6 (très classique pour vous).

La partie III aborde la loi de moyenne empirique. Dans la quesion 9 des résultats sont admis qu'en fait vous devriez savoir démontrer (développement en éléments simples), je vous suggère cependant de sauter cette question en première lecture, ainsi que la question 11, que nous commenterons durant la séance.

Dans la partie IV on étudie la loi de la médiane empirique, et on en vient aux estimateurs.

2 Compléments pour ce sujet

Avant d'aborder le sujet, il convient de vous replonger dans vos cours ou manuels de référence pour réviser les éléments de base concernant les lois continues, les définitions des notions de convergence (presque sûre, en probabilité, en loi), la loi des grands nombres et le Théorème Central Limite.

2.1 Un résultat nécessaire

Ce problème utilise un résultat classique, le théorème de Slutsky, rappelé ci-dessous, avec quelques étapes pour la preuve (ce théorème était traité l'an dernier sur la fiche concernant le TCL, mais je ne sais plus si on avait corrigé cette partie).

Théorème (Théorème de Slutsky)**.** *Si $(Q_n)_n$ est une suite de variables aléatoires qui converge en loi vers une variable aléatoire à densité Q et si $(R_n)_n$ est une suite de variables aléatoires qui converge presque sûrement vers une constante $c > 0$ alors $(R_n\, Q_n)_n$ converge en loi vers $c\,Q$, c'est-à-dire :*

$$\forall t \in \mathbb{R}, \qquad \lim_{n\to+\infty} \mathbb{P}(R_n\, Q_n \leq t) = \mathbb{P}(c\,Q \leq t) \tag{R Q}$$

La preuve repose sur les étapes suivantes :

1. Soit $\varepsilon > 0$, montrer que

$$\bigcap_{N\geq 1} \bigcup_{n\geq N} (|R_n - c| \geq \varepsilon) \;\subset\; \Omega \setminus (\lim_{n\to+\infty} R_n = c)$$

2. En déduire que

$$\forall\varepsilon > 0,\ \forall\delta \in]0,1[\,,\ \exists N,\ \forall n \geq N, \qquad \mathbb{P}(|R_n - c| \leq \varepsilon) > 1 - \delta$$

3. Démontrer (R Q).

2.2 Sur les estimateurs

Concernant les estimateurs statistiques, nous aurons besoin des définitions suivants

— Un *n-échantillon i.i.d.* de la variable aléatoire X est une famille $(X_1, \ldots, X_n)$ de variables aléatoires indépendantes suivant la même loi que X.

— Une variable aléatoire Y est un *estimateur* du paramètre θ de la variable aléatoire X (par exemple ce paramètre peut être son espérance, sa variance,. . .) si $Y = g(X_1, \ldots, X_n)$ est une fonction mesurable d'un échantillon i.i.d. de la v.a. X, qui ne dépend pas de θ mais qu'on peut envisager comme une estimation de ce paramètre.

— Une v.a. $Y = g(X_1, \ldots, X_n)$ est un *estimateur sans biais* du paramètre θ si $E(Y) = E(g(X_1, \ldots, X_n)) = \theta$.

— L'estimateur $g(X_1, \ldots, X_n)$ du paramètre θ est *convergent* si la suite de variables aléatoires $(g_n(X_1, \ldots, X_n))_n$ vérifie les propriétés :

$$\begin{cases} \mathbb{E}(g_n(X_1, \ldots, X_n)) \to \theta \\ Var(g_n(X_1, \ldots, X_n)) \to 0 \end{cases} \qquad \text{quand } n \to +\infty$$