BANQUE COMMUNE D'EPREUVES

CODE EPREUVE :

286
ESSECM2_S

Concepteur : ESSEC

OPTION SCIENTIFIQUE

MATHEMATIQUES II

Lundi 16 mai 2005, de 14 h. à 18 h.

La présentation, la lisibilité, l'orthographe, la qualité de la rédaction, la clarté et la précision des raisonnements entreront pour une part importante dans l'appréciation des copies.
*Les candidats sont invités à **encadrer** dans la mesure du possible les résultats de leurs calculs.*
Ils ne doivent faire usage d'aucun document. L'utilisation de toute calculatrice et de tout matériel électronique est interdite. Seule l'utilisation d'une règle graduée est autorisée.
Si au cours de l'épreuve, un candidat repère ce qui lui semble être une erreur d'énoncé, il la signalera sur sa copie et poursuivra sa composition en expliquant les raisons des initiatives qu'il sera amené à prendre.

Dans ce problème, les variables aléatoires sont réelles et toutes définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$. $(X_n)_{n\geqslant 1}$ représente une suite de variables aléatoires et, pour tout $n \geqslant 1$, on note $S_n = \sum_{i=1}^{n} X_i$. Si X est une variable aléatoire réelle, $E(X)$ désigne son espérance.

Préliminaires

1. Soit (X_n) une suite de variables aléatoires réelles de même loi, admettant une espérance m.
Énoncer, *avec précision*, la loi faible des grands nombres pour la suite (X_n).

2. Soit δ un réel strictement positif et A un sous-ensemble de $\mathbb{R}$ tel que l'intervalle $]m-\delta, m+\delta[$ soit inclus dans le complémentaire de A. Déterminer

$$\lim_{n\to+\infty} P\left(\frac{S_n}{n} \in A\right)$$

L'objet du problème est de préciser de manière quantitative les résultats ci–dessus.

I. Un premier exemple. Le cas gaussien

Dans cette partie, $(X_n)_{n\geqslant 1}$ est une suite de variables aléatoires indépendantes suivant toutes une loi normale centrée réduite, $\mathcal{N}(0,1)$.

1. Quelle est la loi de $\dfrac{S_n}{n}$?

2. Soit δ un réel strictement positif. Dans cette partie, on note exp la fonction exponentielle.

a) Montrer que

$$P\left(\left|\frac{S_n}{n}\right| \geqslant \delta\right) = 2P\left(\frac{S_n}{n} \geqslant \delta\right) = \sqrt{\frac{2n}{\pi}} \int_{\delta}^{+\infty} \exp\left(-\frac{nt^2}{2}\right) dt$$

b) En posant $u = n(t-\delta)$, montrer que

$$P\left(\left|\frac{S_n}{n}\right| \geqslant \delta\right) = \sqrt{\frac{2}{n\pi}} \times \exp\left(-\frac{n\delta^2}{2}\right) \int_{0}^{+\infty} \exp\left(-\frac{u^2}{2n} - u\delta\right) du$$

3. a) Montrer que pour tout $x \geqslant 0$, on a $0 \leqslant 1 - \exp(-x) \leqslant x$.

b) Montrer que l'intégrale $\displaystyle\int_0^{+\infty} \exp(-u\delta)\, du$ converge et la calculer.

c) Déterminer

$$\lim_{n \to +\infty} \left(\int_0^{+\infty} \exp(-u\delta)\, du - \int_0^{+\infty} \exp\left(-\frac{u^2}{2n} - u\delta\right) du \right)$$

d) En déduire, lorsque n tend vers l'infini, un équivalent de $P\left(\left|\dfrac{S_n}{n}\right| \geqslant \delta\right)$.

II. Quelques résultats généraux

À l'instar des variables aléatoires discrètes, on admettra que si X, Y sont deux variables aléatoires à densité, indépendantes, admettant une espérance, alors XY admet une espérance et $E(XY) = E(X)E(Y)$.

Soit X une variable aléatoire réelle (discrète ou à densité). Pour tout $s \in \mathbb{R}$ telle que e^{sX} admet une espérance $E(e^{sX})$, on pose

$$\varphi(s) = E(e^{sX})$$

Soit $(X_n)_{n \geqslant 1}$ une suite de variables aléatoires indépendantes suivant toutes la même loi que X.

1. Montrer que pour tout $n \geqslant 1$, pour tout $s \in \mathbb{R}$ tel que $\varphi(s)$ existe, on a $E(e^{s\frac{S_n}{n}}) = (\varphi(s/n))^n$.

2. Soit Y une variable aléatoire réelle et $s > 0$ tel que $E(e^{sY})$ existe.
a) Montrer que pour tout a réel, $1_{(Y \geqslant a)} \leqslant e^{s(Y-a)}$, où $1_{(Y \geqslant a)}$ désigne la fonction indicatrice de l'ensemble $\{\omega \in \Omega \mid Y(\omega) \geqslant a\}$.
b) En déduire que $P(Y \geqslant a) \leqslant e^{-as}E(e^{sY})$,
c) Montrer que $P\left(\dfrac{S_n}{n} \geqslant a\right) \leqslant e^{-as}(\varphi(s/n))^n$.

3. Soit Y une variable aléatoire réelle et $s < 0$ tel que $E(e^{sY})$ existe.
a) Montrer que pour tout a réel, $1_{(Y \leqslant a)} \leqslant e^{s(Y-a)}$.
b) En déduire que $P(Y \leqslant a) \leqslant e^{-as}E(e^{sY})$, puis que $P\left(\dfrac{S_n}{n} \leqslant a\right) \leqslant e^{-as}(\varphi(s/n))^n$.

III. Un second exemple. Le cas binomial

Dans cette partie, $(X_n)_{n \geqslant 1}$ est une suite de variables aléatoires indépendantes suivant toutes une loi de Bernoulli $\mathcal{B}(p)$, avec $0 < p < 1$. On rappelle que $P(X_1 = 1) = p, \quad P(X_1 = 0) = 1 - p = q$.

1. Calculer $\varphi : s \longmapsto \varphi(s)$ et déterminer son domaine de définition.
Soit a un réel fixé de $]0, 1[$.

2. On suppose dans cette question que $a > p$.
a) Étudier sur $\mathbb{R}^+$ les variations de la fonction ℓ_a définie par

$$\ell_a : s \longmapsto as - \ln \varphi(s)$$

b) Montrer que la fonction ℓ_a atteint sur $\mathbb{R}^+$ un maximum strictement positif $h(a, p)$ qu'on exprimera en fonction de a et p.
c) Montrer que

$$P\left(\frac{S_n}{n} \geqslant a\right) \leqslant e^{-n\left(\sup\limits_{t>0}(at - \ln \varphi(t)\right)} = e^{-nh(a,p)}$$

3. On suppose dans cette question que $a < p$, (donc $1 - a > 1 - p$).

a) Déterminer la loi de la variable aléatoire $n - S_n$.

b) Montrer que

$$P\left(\frac{S_n}{n} \leqslant a\right) \leqslant e^{-nh(1-a,1-p)} = e^{-nh(a,p)}$$

4. Soit $\varepsilon > 0$. Déduire des questions précédentes que

$$P\left(\left|\frac{S_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant 2e^{-n\min(h(p+\varepsilon,p),h(p-\varepsilon,p))}$$

5. Soit α un réel de $]0,1[$. Montrer que, pour n assez grand, il est toujours possible de trouver deux réels a_1, a_2 tels que $0 < a_1 < p < a_2 < 1$ qui vérifient les inégalités :

$$\begin{cases} P\left(\dfrac{S_n}{n} \leqslant a_1\right) \leqslant \dfrac{\alpha}{2} \\ P\left(\dfrac{S_n}{n} \geqslant a_2\right) \leqslant \dfrac{\alpha}{2} \end{cases}$$

(on pourra étudier les variations de la fonction $a \mapsto h(a,p)$).

6. Une entreprise souhaite acquérir une machine qui fabrique un certain type d'objets et qui, en fonctionnement normal, produit une proportion p, $(0 < p < 1)$, d'objets défectueux. Le directeur veut connaître la valeur de p. Pour cela il teste la machine et prélève un échantillon de n, $(n \geqslant 1)$, objets qu'il analyse.

Pour tout $i \in [\![1,n]\!]$, soit X_i la variable aléatoire de Bernoulli définie par

$$X_i = \begin{cases} 1 & \text{si le } i\text{-ième objet prélevé est défectueux} \\ 0 & \text{sinon} \end{cases}$$

On suppose que dans les conditions de prélèvement, les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes.

a) Montrer que $F_n = \dfrac{S_n}{n}$ est un estimateur sans biais de p.

b) Calculer le risque quadratique $r_n = E\left((F_n - p)^2\right)$. Déterminer $\lim\limits_{n \to +\infty} r_n$.

7. On souhaite déterminer dans cette question un intervalle de confiance du paramètre p inconnu, au niveau de confiance $1 - \alpha$, à partir de l'échantillon $(X_1, \ldots, X_n)$.

a) Quelle est la limite en loi de la suite $\left(\sqrt{n}\dfrac{F_n - p}{\sqrt{p(1-p)}}\right)_{n \in \mathbb{N}^*}$?

b) Soit f_n la réalisation de F_n sur l'échantillon considéré. Soit t_α le réel défini par $\Phi(t_\alpha) = 1 - \dfrac{\alpha}{2}$, où Φ désigne la fonction de répartition de la loi normale centrée, réduite.

Déterminer en fonction de n, f_n, t_α un intervalle de confiance $[U_n, V_n]$ de p au niveau $1 - \alpha$.

IV. Le cas général

Soit X une variable aléatoire réelle de densité f.

Pour tout réel t tel que $\displaystyle\int_{-\infty}^{+\infty} e^{tu} f(u)du$ converge, on note

$$L_X : t \longmapsto E(e^{tX}) = \int_{-\infty}^{+\infty} e^{tu} f(u)du$$

On supposera que L_X est défini sur un intervalle $]\alpha, \beta[$ contenant 0.

1. Soit $t \in]\alpha, \beta[$, et $\delta > 0$ tel que $[t - \delta, t + \delta] \subset]\alpha, \beta[$.

a) Montrer que pour tout u réel

$$|e^{\delta u} - 1 - \delta u| \leqslant \sum_{n=2}^{\infty} \frac{\delta^n |u^n|}{n!}$$

b) Montrer que, pour tout u réel

$$e^{tu}\left|e^{\delta u}-1-\delta u\right|f(u) \leqslant (e^{(t-\delta)u}+e^{(t+\delta)u})f(u)$$

c) En déduire que l'intégrale $\displaystyle\int_{-\infty}^{+\infty} ue^{tu}f(u)du$ converge, puis que X admet une espérance m.

2. Soit $t \in]\alpha,\beta[$, et $\delta > 0$ tel que $[t-\delta, t+\delta] \subset]\alpha,\beta[$. Soit $h \in \mathbb{R}$ tel que $|h| < \delta$.

a) Montrer que, pour tout u réel,

$$\left|\frac{e^{(t+h)u}-e^{tu}}{h}-ue^{tu}\right|f(u) \leqslant |h|e^{tu}\sum_{n=2}^{\infty}\frac{\delta^{n-2}|u|^n}{n!}f(u)$$

puis que

$$\delta^2\left|\frac{e^{(t+h)u}-e^{tu}}{h}-ue^{tu}\right|f(u) \leqslant |h|e^{tu}e^{\delta|u|}f(u)$$

b) Montrer que L_X est dérivable en t et que

$$L'_X(t) = \int_{-\infty}^{+\infty} ue^{tu}f(u)du$$

On admettra (et on démontrerait de manière analogue) que la fonction L_X est de classe C^2 sur $]\alpha,\beta[$ et que pour tout $t \in]\alpha,\beta[$

$$L''_X(t) = \int_{-\infty}^{+\infty} u^2e^{tu}f(u)du$$

3. On note $\psi(t) = \ln L_X(t)$.

a) Donner le domaine de définition de la fonction ψ.

b) Calculer ψ', ψ'', dérivées première et seconde de ψ.

c) En utilisant l'inégalité de Cauchy Schwarz, montrer que ψ' est strictement croissante sur $]\alpha,\beta[$; en déduire que ψ' admet en α (respectivement β) une limite (finie ou infinie) qu'on notera $L_1(\alpha)$ (respectivement $L_1(\beta)$). Quelle est la valeur de $\psi'(0)$?

d) Montrer que ψ admet en α (respectivement β) une limite (finie ou infinie) qu'on notera $L_0(\alpha)$ (respectivement $L_0(\beta)$).

4. Soit a réel donné. Pour tout $t \in]\alpha,\beta[$, on pose : $\ell_a(t) = at - \psi(t)$.
Dresser le tableau de variations de la fonction ℓ_a (on distinguera trois cas : $a \geqslant L_1(\beta), a \leqslant L_1(\alpha)$ et $L_1(\alpha) < a < L_1(\beta)$).

5. Pour tout a réel, on pose : $h(a) = \sup\limits_{t\in]\alpha,\beta[} \ell_a(t)$. On suppose désormais que $L_1(\alpha) < a < L_1(\beta)$.
Montrer que $h(a) = a(\psi')^{-1}(a) - \psi((\psi')^{-1}(a))$.

6. Montrer que si $a > m$,

$$h(a) = \sup_{t\in]0,\beta[}(at-\psi(t))$$

puis que si $a < m$

$$h(a) = \sup_{t\in]\alpha,0]}(at-\psi(t))$$

7. Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes de même loi que X.
Montrer que si $a > m$, $P\left(\dfrac{S_n}{n} \geqslant a\right) \leqslant e^{-nh(a)}$, puis que si $a < m$, $P\left(\dfrac{S_n}{n} \leqslant a\right) \leqslant e^{-nh(a)}$.

8. Soit $\varepsilon > 0$. Montrer que

$$P\left(\left|\frac{S_n}{n}-m\right| \geqslant \varepsilon\right) \leqslant 2e^{-n\min(h(m-\varepsilon),h(m+\varepsilon))}$$