



Concours Blanc n°2 - Math II



Samedi 3 Juin
Durée : 4 heures

Dans tout le sujet, concernant les questions d'informatique sous Python, on suppose que les bibliothèques d'usage sont déjà importées sous leurs alias habituels :

```
import numpy as np
import numpy.linalg as al
import numpy.random as rd
import matplotlib.pyplot as plt
```

Toutes les variables aléatoires intervenant dans le problème sont définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$. Pour une variable aléatoire X , on notera $E(X)$ son espérance et $V(X)$ sa variance, lorsqu'elles existent.

Partie 1 : Contexte et préliminaires

Un joueur joue à *Pile* ou *Face* avec une pièce équilibrée : pile lui rapporte 1 euro, face lui coûte 1 euro.

Notons, pour $n \in \mathbb{N}^*$, S_n la variable aléatoire égale à son gain total (éventuellement négatif) après n lancers de la pièce.

- (1) Écrire, en langage Python, une fonction d'en-tête `def simul_S(n):` qui renvoie une simulation de S_n .
- (2) On note, pour $n \in \mathbb{N}^*$, T_n la variable aléatoire correspondant au nombre de *Pile* obtenus lors des n lancers de la pièce.
 - (a) Quelle est la loi de T_n ? Rappeler la valeur de son espérance et de sa variance.
 - (b) Exprimer S_n en fonction de T_n . En déduire que $E(S_n) = 0$.

Le joueur est intéressé à estimer sa probabilité de gagner en n coups plus de K euros, ou au moins d'obtenir une minoration de $P(S_n \geq K)$.

- (3) Donner l'expression, pour $K \in [-n, n]$, de $P(S_n \geq K)$ sous forme d'une somme qu'on ne cherchera pas à simplifier.

L'objet de toute la suite du problème est d'obtenir une telle minoration sous la forme d'une inégalité dite *de grandes déviations*.

Partie 2 : Obtention de quelques résultats du cours de deuxième année

(4) Soient X et Y deux variables aléatoires finies telles que $X(\Omega) = \{x_1, \dots, x_\ell\}$ et $Y(\Omega) = \{y_1, \dots, y_m\}$.

(a) Justifier que XY admet une espérance. On **admet** alors que

$$E(XY) = \sum_{i=1}^{\ell} \sum_{j=1}^m x_i y_j P([X = x_i] \cap [Y = y_j]).$$

(b) On cherche à écrire une fonction Python permettant le calcul de $E(XY)$. On suppose que les valeurs prises par X et Y sont stockées respectivement dans des variables X et Y et que la loi conjointe de X et Y est stockée dans une matrice M (c'est à dire que si $M = (M_{i,j})_{\substack{1 \leq i \leq \ell, \\ 1 \leq j \leq m}}$, on a $M_{i,j} = P([X = x_i] \cap [Y = y_j])$). Compléter la fonction ci-dessous afin qu'elle renvoie la valeur souhaitée.

```
def esp_prod(X,Y,M) :
    l=.....
    m=.....
    s=0
    for i in range(.....) :
        for j in range(.....) :
            s = .....
    return s
```

(c) On rappelle qu'on dit que X et Y sont indépendantes si, pour tous $(i, j) \in \llbracket 1, \ell \rrbracket \times \llbracket 1, m \rrbracket$, on a $P([X = x_i] \cap [Y = y_j]) = P(X = x_i)P(Y = y_j)$.
Montrer que, si X et Y sont indépendantes, alors

$$E(XY) = E(X)E(Y).$$

(d) Écrire alors une fonction d'en-tête `def test_nec(X,Y,M):` qui prend en argument les valeurs prises par X et Y , la matrice de la loi conjointe et renvoie 0 si les variables ne sont pas indépendantes et -1 si on ne peut pas conclure.

(5) Soit Y une variable aléatoire finie à valeurs positives d'univers image $Y(\Omega) = \{y_1, \dots, y_\ell\} \subset \mathbb{R}_+$ avec $0 \leq y_1 < y_2 < \dots < y_\ell$. Soit $t > 0$. En remarquant que

$$E(Y) = \sum_{i: y_i < t} y_i P(Y = y_i) + \sum_{i: y_i \geq t} y_i P(Y = y_i),$$

montrer que

$$P(Y \geq t) \leq \frac{E(Y)}{t}.$$

Cette inégalité est un résultat du cours de deuxième année connu sous le nom d'*Inégalité de Markov*.

Dans toute la suite du problème, on **admet** le résultat suivant (connu sous le nom de *Lemme des coalitions*) :

Si $(X_n)_{n \geq 0}$ est une suite de variables aléatoires mutuellement indépendantes, alors, pour tout $n \in \mathbb{N}$, toute variable aléatoire obtenue comme fonction de X_{n+1} est indépendante de toute variable aléatoire obtenue comme fonction des variables $X_0, X_1, \dots, X_n$.

Partie 3 : Fonction génératrice des moments

- (6) On considère une variable aléatoire finie X .
Montrer que pour tout entier $k \geq 1$, $E(X^k)$ existe.

On introduit la *fonction génératrice des moments* de X , définie pour tout $t \in \mathbb{R}$ par :

$$M_X(t) = E(e^{tX}).$$

- (7) Soit X une variable aléatoire finie d'univers image $X(\Omega) = \{x_1, x_2, \dots, x_\ell\}$.
On note $p_i = P(X = x_i)$ avec $p_i > 0$ pour $i = 1, \dots, \ell$.

(a) (i) Montrer que :
$$M_X(t) = \sum_{i=1}^{\ell} p_i e^{tx_i}.$$

(ii) En déduire que, pour tout $t \in \mathbb{R}$, on a : $M_X(t) > 0$.

(b) Pour $k \in \mathbb{N}$, on note $M_X^{(k)}$ la dérivée k -ième de M_X .

(i) Montrer par récurrence sur $k \in \mathbb{N}$ que, pour tout $t \in \mathbb{R}$, : $M_X^{(k)}(t) = E(X^k e^{tX})$.

(ii) Soit $k \in \mathbb{N}$. Montrer que : $M_X^{(k)}(0) = E(X^k)$.

(iii) Exprimer $V(X)$ en fonction de $M_X'(0)$ et $M_X''(0)$.

(c) Montrer que la fonction M_X est convexe.

- (8) Soient $X_1, \dots, X_n$ des variables aléatoires finies mutuellement indépendantes.
On pose $S_n = X_1 + \dots + X_n$.

Montrer que pour tout $t \in \mathbb{R}$, on a :

$$M_{S_n}(t) = \prod_{i=1}^n M_{X_i}(t)$$

(On pourra utiliser une récurrence et des éléments de la Partie 2).

- (9) *Un exemple.* On suppose ici que X suit la loi de Bernoulli de paramètre p où p est un réel donné dans $]0, 1[$.

(a) Calculer $M_X(t)$.

(b) Calculer $M_X'(0)$ et $M_X''(0)$.

(c) A l'aide de la formule trouvée dans la Question (7(b)iii), retrouver l'expression de $V(X)$.

- (10) Soit Y une variable aléatoire d'univers image $Y(\Omega) = \{y_1, \dots, y_\ell\}$.

On pose $P(Y = y_i) = p_i > 0$.

On suppose que $E(Y) < 0$ et que $P(Y > 0) > 0$.

(a) Montrer que : $M_Y'(0) < 0$.

(b) Montrer que : $\lim_{t \rightarrow +\infty} M_Y(t) = +\infty$.

(c) Déduire qu'il existe $\tau \in]0, +\infty[$ tel que pour tout $t > 0$ on ait : $M_Y(t) \geq M_Y(\tau)$.

- (11) Soit $a < 0 < b$ deux réels.

Soit X une variable aléatoire finie d'espérance $E(X) = 0$, telle que $P(a \leq X \leq b) = 1$.

(a) On rappelle une propriété de convexité :

Pour toute fonction convexe f sur $[a, b]$, pour tous réels λ et μ tels que $\lambda + \mu = 1$, on a :

$$f(\lambda a + \mu b) \leq \lambda f(a) + \mu f(b).$$

Soit $t > 0$ un réel strictement positif. Montrer que pour tout $u \in [a, b]$, on a :

$$e^{tu} \leq \frac{b-u}{b-a} e^{ta} + \frac{u-a}{b-a} e^{tb}.$$

(b) En déduire que pour tout $t > 0$, on a

$$M_X(t) \leq \frac{b}{b-a} e^{ta} - \frac{a}{b-a} e^{tb}.$$

On pose $p = \frac{-a}{b-a} \in [0, 1]$ et on définit la fonction ψ sur $\mathbb{R}_+$ par $\psi(s) = \ln(1 - p + pe^s) - ps$.

(a) Montrer que pour tout $t > 0$, on a : $M_X(t) \leq \exp(\psi(t(b-a)))$.

(b) (i) Montrer que $\psi(0) = \psi'(0) = 0$.

(ii) Montrer que pour tous réels positifs $u, v > 0$, on a : $uv \leq \frac{1}{4}(u+v)^2$.

(iii) Montrer que pour tout réel $s \in \mathbb{R}_+^*$, on a : $\psi''(s) \leq \frac{1}{4}$.

(iv) Conclure que pour tout réel $s \in \mathbb{R}_+^*$, on a : $\psi(s) \leq \frac{1}{8}s^2$.

(c) Conclure que pour tout $t > 0$, on a

$$M_X(t) \leq \exp\left(\frac{1}{8}t^2(b-a)^2\right).$$

Partie 4 : Quelques inégalités

On rappelle que la fonction logarithme est concave, ce qui se traduit par la propriété suivante :

Si $p_1, \dots, p_\ell$ sont des nombres réels positifs tels que $\sum_{i=1}^{\ell} p_i = 1$ et $x_1, \dots, x_\ell$ sont des réels strictement positifs, on a :

$$\ln(p_1 x_1 + \dots + p_\ell x_\ell) \geq p_1 \ln(x_1) + \dots + p_\ell \ln(x_\ell).$$

(12) Soit X une variable aléatoire finie à valeurs strictement positives.

Montrer l'inégalité de Jensen :

$$\ln(E(X)) \geq E(\ln(X)).$$

(13) Soient p et q deux nombres réels strictement supérieurs à 1 tels que $\frac{1}{p} + \frac{1}{q} = 1$.

(a) Soient a et b deux réels positifs. Montrer que :

$$\frac{1}{p}a^p + \frac{1}{q}b^q \geq ab.$$

(b) On considère, pour les Questions (13b) à (13f), X et Y deux variables aléatoires finies, à valeurs **strictement positives**.

Montrer que :

$$E(XY) \leq \frac{1}{p}E(X^p) + \frac{1}{q}E(Y^q).$$

(c) Soit $\lambda > 0$. Montrer que :

$$E(XY) \leq \frac{1}{p}\lambda^{p-1}E(X^p) + \frac{1}{q}\lambda^{-1}E(Y^q).$$

(d) Soient α et β deux réels strictement positifs. On considère la fonction φ définie sur $\mathbb{R}_+^*$ par :

$$\varphi(\lambda) = \frac{1}{p}\lambda^{p-1}\alpha + \frac{1}{q}\lambda^{-1}\beta.$$

- (i) Dresser le tableau de variations de φ .
- (ii) Trouver le réel λ_0 où φ atteint son minimum.
- (iii) Déterminer la valeur minimum de φ .

(e) Montrer l'inégalité de Hölder :

$$E(XY) \leq E(X^p)^{1/p} E(Y^q)^{1/q}$$

(f) Expliciter l'inégalité dans le cas $p = q = 2$.

(14) Soit X une variable aléatoire finie, telle que $E(X) = 0$.

On note σ le réel positif tel que $\sigma^2 = E(X^2) > 0$ et ξ le réel positif tel que $\xi^4 = E(X^4) > 0$.
On considère les deux variables aléatoires positives $X^+ = \max(X, 0)$ et $X^- = \max(-X, 0)$.
On pose $\delta = P(X \geq 0)$ et on cherche à estimer δ .

(a) Montrer que $X = X^+ - X^-$. Observer que ceci implique $E(X^+) = E(X^-)$.

(b) Montrer que $X^2 = (X^+)^2 + (X^-)^2$. Observer que ceci implique

$$\sigma^2 = E[(X^+)^2] + E[(X^-)^2].$$

(c) Montrer que : $E[(X^+)^4] \leq \xi^4$ et $E[(X^-)^4] \leq \xi^4$.

(d) Soit W la variable aléatoire définie par $W(\omega) = 1$ si $X(\omega) \geq 0$ et $W(\omega) = 0$ si $X(\omega) < 0$.

(i) Montrer que : $X^+ = WX$.

(ii) Montrer que : $E((X^+)^2) \leq \sqrt{\delta}\xi^2$.

(e) En appliquant l'inégalité de Hölder avec les coefficients $p = 3/2$ et $q = 3$, montrer que :

$$E((X^-)^2) \leq (E(X^-))^{2/3} (E(X^-)^4)^{1/3}.$$

(f) En appliquant l'inégalité de Hölder avec les coefficients $p = 4$ et $q = 4/3$, montrer que :

$$E(X^-) = E(X^+) \leq \xi\delta^{3/4}.$$

(g) Dédire que : $\sigma^2 \leq 2\sqrt{\delta}\xi^2$.

(h) Conclure que : $P(X \geq 0) \geq \frac{\sigma^4}{4\xi^4}$.

Partie 5 : Grandes déviations.

On considère une variable aléatoire finie Y d'univers image $\{y_1, \dots, y_\ell\}$.

On note $P(Y = y_i) = p_i > 0$.

Comme dans la Question 10, on suppose que $E(Y) < 0$ et $P(Y > 0) > 0$.

Comme $P(Y > 0) > 0$, il existe des valeurs strictement positives parmi les y_i . Quitte à ré-indexer les y_i , on suppose qu'il existe $k \in \{1, \dots, \ell - 1\}$ tel que $y_1, \dots, y_k$ sont positifs et $y_{k+1}, \dots, y_\ell$ sont strictement négatifs.

Au vu de la Question 10c, il existe $\tau > 0$ tel que $M_Y(\tau) = \inf_{t \in [0, +\infty[} M_Y(t)$. On pose $\rho = M_Y(\tau)$.

(15) (a) Montrer que : $P(Y \geq 0) = P(e^{\tau Y} \geq 1)$.

(b) En déduire, à l'aide d'un résultat de la Partie 2, que : $P(Y \geq 0) \leq \rho$.

(16) Montrer que

$$\sum_{i=1}^{\ell} \frac{e^{\tau y_i}}{\rho} P(Y = y_i) = 1.$$

Soit Z une variable aléatoire d'univers image $\{y_1, \dots, y_\ell\}$ telle que : $\forall 1 \leq i \leq \ell$, on a

$$P(Z = y_i) = \frac{e^{\tau y_i}}{\rho} P(Y = y_i).$$

- (16) (a) Montrer que : $M_Z(t) = \rho^{-1} M_Y(\tau + t)$.
 (b) Montrer que : $E(Z) = 0$.
 (c) Montrer que : $E(Z^2) > 0$.

(17) On rappelle que les valeurs positives parmi les y_i sont : $y_1, \dots, y_k$.

(a) Montrer que :

$$P(Y \geq 0) = \rho \sum_{i=1}^k e^{-\tau y_i} P(Z = y_i).$$

On pose $\delta = P(Z \geq 0)$.

(a) Montrer que : $\sum_{i=1}^k \delta^{-1} P(Z = y_i) = 1$.

(b) Montrer que :

$$\ln \left(\sum_{i=1}^k e^{-\tau y_i} \delta^{-1} P(Z = y_i) \right) \geq -\tau \delta^{-1} \sum_{i=1}^k y_i P(Z = y_i).$$

(c) On considère $\sigma > 0$ le réel tel que $\sigma^2 = E(Z^2)$.

(i) Montrer que : $\sum_{i=1}^k y_i P(Z = y_i) \leq E(|Z|)$.

(ii) En déduire que : $\sum_{i=1}^k y_i P(Z = y_i) \leq \sigma$.

(d) Montrer que :

$$\ln \left(\sum_{j=1}^k e^{-\tau y_j} P(Z = y_j) \right) \geq \ln \delta - \frac{\tau \sigma}{\delta}.$$

(e) Conclure que :

$$P(Y \geq 0) \geq \rho \exp \left(- \left[\frac{\tau \sigma}{\delta} - \ln \delta \right] \right).$$

(f) On considère le réel positif ξ tel que $\xi^4 = E(Z^4) > 0$. Montrer que :

$$P(Y \geq 0) \geq \rho \exp \left(- \left[\frac{4\tau \xi^4}{\sigma^3} - \ln \frac{\sigma^4}{4\xi^4} \right] \right).$$

(18) On suppose dans cette question qu'il existe deux réels $c < 0 < d$ tels que $P(c \leq Y \leq d) = 1$. On rappelle que l'on a supposé que $E(Y) < 0$ et que $P(Y > 0) > 0$.

(a) On pose $X = Y - E(Y)$.

(i) Montrer que X vérifie les hypothèses de la Question 11, avec $a = c - E(Y)$ et $b = d - E(Y)$.

On vérifiera en particulier que $a < 0 < b$.

(ii) En déduire que pour tout $t > 0$

$$M_X(t) \leq \exp\left(\frac{1}{8}t^2(d-c)^2\right).$$

Soit $n \in \mathbb{N}^*$ et soient $Y_1, \dots, Y_n$ des variables aléatoires indépendantes et de même loi que Y . Pour tout $1 \leq i \leq n$, on pose $X_i = Y_i - E(Y)$.

(a) Montrer que pour tout $t > 0$, on a

$$P\left(\sum_{i=1}^n Y_i \geq 0\right) = P\left(e^{t(X_1 + \dots + X_n)} \geq e^{-ntE(Y)}\right).$$

(b) Montrer que pour tout $t > 0$

$$P\left(\sum_{i=1}^n Y_i \geq 0\right) \leq e^{ntE(Y)} M_{X_1 + \dots + X_n}(t).$$

(On pourra utiliser à nouveau un résultat de la Partie 2.)

(c) En déduire que pour tout réel $t > 0$, on a

$$P\left(\sum_{i=1}^n Y_i \geq 0\right) \leq \exp\left(ntE(Y) + \frac{1}{8}nt^2(d-c)^2\right).$$

(d) Conclure que pour tout $n \in \mathbb{N}^*$, on a

$$P\left(\sum_{i=1}^n Y_i \geq 0\right) \leq \exp\left(-\frac{2E(Y)^2}{(d-c)^2}n\right).$$

