ULC 235 **J. 2043**

SESSION 2002

Filière BCPST

(Epreuve commune aux ENS de Paris, Lyon et Cachan)

MATHEMATIQUES

Durée : 4 heures

L'usage de calculatrices électroniques de poche à alimentation autonome, non imprimantes et sans document d'accompagnement, est autorisé. Cependant, une seule calculatrice à la fois est admise sur la table ou le poste de travail, et aucun échange n'est autorisé entre les candidats.

Ce problème aborde l'étude d'une classe d'objets probabilistes connus sous le nom de "marches aléatoires". Dans tout le problème, N (respectivement N^*) désigne l'ensemble des nombres entiers naturels (respectivement, des entiers naturels strictement positifs); Z, l'ensemble des nombres entiers relatifs; R (respectivement R_+, R_+^*), l'ensemble des nombres réels (respectivement, des réels positifs ou nuls, des réels strictement positifs). On note e la base du logarithme néperien (ln), et la fonction exponentielle de x est notée indifféremment $x \mapsto e^x$ ou $x \mapsto \exp x$. L'espérance et la variance d'une variable aléatoire réelle X seront notées respectivement $\mathrm{E}[X]$ et $\mathrm{var}\, X$.

Soit $(X_i)_{i \in N^*}$ une suite de variables aléatoires réelles definies sur un même espace de probabilité $(\Omega, \mathcal{A}, \mathrm{P})$, indépendantes, identiquement distribuées (c'est-à-dire, pour tout entier $k > 0$, les variables $X_1, X_2, \ldots, X_k$ sont indépendantes et identiquement distribuées). Pour tout entier $n > 0$, on pose:

$$S_n = \sum_{i=1}^{n} X_i .$$

Sauf mention contraire (questions II.2 et II.3), on pose $S_0 = 0$. La suite $(S_n)_{n \in N}$ est appelée "marche aléatoire". Dans tout le problème, on suppose $\mathrm{P}(X_1 = 0) < 1$, $\mathrm{E}[X_1] < \infty$ et $0 < \mathrm{var}\, X_1 < \infty$.

Tournez la page S.V.P.

Pour tout événement A de l'espace de probabilité considéré, on note $\mathbf{1}_A$ la fonction indicatrice de A, c'est-à-dire la variable aléatoire qui vaut 1 si A est réalisé, 0 sinon.

Soient a et b deux nombres réels tels que $a<0<b$. On note T la variable aléatoire définie par:

$$T=\min\{n\in \mathrm{N}^* \mid S_n \leq a \text{ ou } S_n \geq b\}.$$

Le problème vise à décrire la loi de T, dont l'étude connaît d'importantes applications en biologie des populations – évaluation du risque d'extinction d'une espèce, planification des stratégies de contrôle des épidémies, etc. La partie II utilise les résultats de la partie I. Les parties III et IV sont indépendantes et peuvent être traitées indépendamment des parties I et II.

PARTIE I

I.1. Montrer que la variable aléatoire T est finie presque surement.

On suppose dans la suite de cette partie I que $\mathrm{E}[T]$ est fini (la démonstration de ce résultat fera l'objet de la question III.4.e).

I.2. Montrer les égalités:

(a) $\mathrm{E}[S_T]=\mathrm{E}\left[\sum_{j=1}^{+\infty} X_j\left(1-\mathbf{1}_{\{T<j\}}\right)\right]$,

(b) $\mathrm{E}[S_T^2]=\mathrm{E}\left[\sum_{j=1}^{+\infty}\sum_{k=1}^{+\infty} X_j\left(1-\mathbf{1}_{\{T<j\}}\right)X_k\left(1-\mathbf{1}_{\{T<k\}}\right)\right]$.

I.3. Montrer, pour tout $j\in \mathrm{N}^*$, que les variables aléatoires X_j et $\left(1-\mathbf{1}_{\{T<j\}}\right)$ sont indépendantes, de même que les variables aléatoires X_k et $X_j\left(1-\mathbf{1}_{\{T<j\}}\right)\left(1-\mathbf{1}_{\{T<k\}}\right)$ pour tout entier k tel que $k>j$.

I.4. On admet que l'espérance et la somme infinie qui apparaissent dans l'expression (a) de la question I.2 peuvent être permutées. Démontrer l'égalité:

$$\mathrm{E}[S_T]=\mathrm{E}[X_1]\cdot \mathrm{E}[T].$$

(C'est la "première identité de Wald".)

I.5. On suppose $\mathrm{E}[X_1]=0$. En admettant que l'espérance et la somme infinie qui apparaissent dans l'expression (b) de la question I.2 peuvent être permutées, montrer l'égalité:

$$\mathrm{E}[S_T^2]=\operatorname{var} X_1\cdot \mathrm{E}[T],$$

et en déduire:

$$\operatorname{var} S_T=\operatorname{var} X_1\cdot \mathrm{E}[T].$$

(C'est la "deuxième identité de Wald".)

PARTIE II

On suppose *dans cette partie* II *seulement* que a et b sont des entiers relatifs tels que $a<0<b$, et que l'on a $\mathrm{P}(X_1=-1 \text{ ou } 0 \text{ ou } 1)=1$. Plus précisément, on pose:

$$\mathrm{P}(X_1=1)=\mathrm{P}(X_1=-1)=p$$

et

$$\mathrm{P}(X_1=0)=q,$$

où p et q sont des réels positifs ou nuls tels que $2p+q=1$ et $q<1$.

On admettra, comme à la partie I, que $\mathrm{E}[T]$ est fini.

II.1. Utiliser la première identité de Wald (*cf.* question I.4) pour calculer $\mathrm{E}[S_T]$; en déduire les expressions de $\mathrm{P}(S_T=a)$ et $\mathrm{P}(S_T=b)$, puis celle de $\mathrm{E}[S_T^2]$. En utilisant la deuxième identité de Wald (*cf.* question I.5), établir l'égalité:

$$\mathrm{E}[T]=\frac{|ab|}{1-q}.$$

Soient Δ une partie de Z, et x et y, deux entiers naturels tels que $x\notin\Delta$ et $y\notin\Delta$. On note $\mathrm{P}_\Delta(x,y)$ la probabilité que la marche aléatoire $(S_n)_{n\in\mathrm{N}}$ partant de x (ce qui signifie: $S_0=x$) atteigne la valeur y à un instant $n\in\mathrm{N}^*$ inférieur à l'instant auquel elle atteint la partie Δ pour la première fois.

II.2. On suppose dans cette question II.2 seulement que $\Delta=\{a,b\}$. Soit y un entier relatif tel que $a<y<b$. Montrer l'égalité:

$$\mathrm{P}_\Delta(y,y)=1-\frac{p(b-a)}{(y-a)(b-y)}.$$

II.3. On note $\mathrm{E}_\Delta(x,y)$ l'espérance de la variable aléatoire $W_{(x,y)}$ égale au nombre de passages par la valeur y que la marche aléatoire $(S_n)_{n\in\mathrm{N}}$ partant de x effectue avant d'atteindre Δ pour la première fois. On se propose de montrer l'égalité:

$$\mathrm{E}_\Delta(x,y)=\frac{\mathrm{P}_\Delta(x,y)}{1-\mathrm{P}_\Delta(y,y)}.$$

II.3.a. On se place d'abord dans le cas particulier $x=y$. Montrer que $W_{(y,y)}$ suit une loi géométrique de paramètre $1-\mathrm{P}_\Delta(y,y)$. En déduire l'égalité annoncée.

II.3.b. Dans le cas $x\neq y$, montrer l'égalité:

$$\mathrm{E}_\Delta(x,y)=\mathrm{P}_\Delta(x,y)[1+\mathrm{E}_\Delta(y,y)],$$

et conclure.

PARTIE III

Cette partie III vise à établir une généralisation des identités de Wald démontrées dans la partie I. On suppose ici que la fonction génératrice $G(t)=\mathrm{E}[e^{tX_1}]$ de la variable aléatoire X_1 est définie pour tout réel t dans un certain intervalle (fini ou infini). Dans les questions III.1 à III.3, on supposera t quelconque dans cet intervalle.

Pour toute variable aléatoire X définie sur l'espace de probabilité $(\Omega,\mathcal{A},\mathrm{P})$ et tout événement A on pose: $\mathrm{E}_A[X]=\mathrm{E}[\mathbf{1}_A X]$.

III.1. Montrer que, quel que soit $n\in\mathrm{N}^*$, $\mathrm{E}[e^{tS_n}]=G(t)^n$.

III.2. Soit $n\in\mathrm{N}^*$. Conditionnellement à l'événement $\{T\le n\}$, les variables aléatoires S_n-S_T et S_T sont-elles indépendantes ?

III.3. Démontrer l'égalité:

$$\mathrm{E}_{\{T\le n\}}[e^{tS_n}]=\mathrm{E}_{\{T\le n\}}[e^{tS_T}G(t)^{n-T}].$$

On pourra écrire: $S_n=S_T+(S_n-S_T)$.

III.4. On établit dans cette question un lemme (III.4.e) qui sera utilisé à la question suivante pour établir une généralisation des identités de Wald.

III.4.a. En appliquant l'inégalité de Tchebychev, montrer qu'il existe un entier $n_0\in\mathrm{N}^*$ tel que pour tout $n\ge n_0$, on a:

$$\mathrm{P}\left(|S_n|\le 2|a|+2b\right)<\frac{1}{2}.$$

III.4.b. Montrer que pour tout entier $k>1$, les variables aléatoires S_{n_0}, $S_{2n_0}-S_{n_0}$, $S_{3n_0}-S_{2n_0}$, ..., $S_{kn_0}-S_{(k-1)n_0}$ sont indépendantes.

III.4.c. Montrer que l'on a:

$$\mathrm{P}\left(\forall k\in\mathrm{N}^*, |S_{kn_0}-S_{(k-1)n_0}|\le 2|a|+2b\right)=0.$$

III.4.d. Montrer qu'il existe des constantes réelles $M>0$ et $c>0$ telles que, quel que soit $n\in\mathrm{N}^*$,

$$\mathrm{P}\left(\forall k\in\{1,2,\ldots,n\}, |S_k|\le |a|+b\right)<M\,e^{-cn}.$$

III.4.e. En déduire que, quel que soit $n\in\mathrm{N}^*$, $\mathrm{P}(T>n)<M\,e^{-cn}$, puis que les moments de T de tous ordres sont finis.

III.5. En écrivant $\mathrm{E}[e^{tS_n}]$ sous la forme:

$$\mathrm{E}[e^{tS_n}]=\mathrm{E}_{\{T\le n\}}[e^{tS_n}]+\mathrm{E}_{\{T>n\}}[e^{tS_n}]$$

(expression que l'on justifiera), montrer que pour tout réel t tel que $1 \leq G(t) < +\infty$, on a l'égalité:

$$\mathrm{E}[e^{tS_T} G(t)^{-T}] = 1$$

("identité de Wald généralisée"). Pour cela, on utilisera les résultats des questions précédentes (III.1, III.3 et III.4.e), et l'on remarquera que la condition $T > n$ implique $a < S_n < b$.

PARTIE IV

On suppose ici que les variables aléatoires (indépendantes et identiquement distribuées) $X_1, X_2, \ldots$ sont entières, à valeurs dans l'ensemble

$$\{-r, -(r-1), \ldots, -1, 0, 1, 2, \ldots, s\}$$

où r et s sont fixés dans $\mathbb{N}^*$. On définit une variable aléatoire U_n en posant:

$$U_n = \max_{0 \leq k \leq n} S_k .$$

Etant donnés $j \in \mathbb{N}^*$, $i_1, i_2, \ldots, i_j$ des entiers positifs distincts, et φ une fonction réelle quelconque définie sur $\{-r, \ldots, s\}^j$, on admettra que l'espérance de la variable aléatoire $\varphi(X_{i_1}, X_{i_2}, \ldots, X_{i_j})$ est donnée par:

$$\mathrm{E}[\varphi(X_{i_1}, X_{i_2}, \ldots, X_{i_j})] = \sum_{\substack{(n_1, n_2, \ldots, n_j) \\ \in \{-r, \ldots, s\}^j}} \varphi(n_1, n_2, \ldots, n_j) \mathrm{P}(X_{i_1} = n_1) \mathrm{P}(X_{i_2} = n_2) \ldots \mathrm{P}(X_{i_j} = n_j).$$

On pose $\mathrm{E}_A[\varphi(X_{i_1}, X_{i_2}, \ldots, X_{i_j})] = \mathrm{E}[\mathbf{1}_A \varphi(X_{i_1}, X_{i_2}, \ldots, X_{i_j})]$, où $\mathbf{1}_A$ désigne la fonction indicatrice d'un événement A de l'espace de probabilité considéré.

IV.1. Donner (sans démonstration) le développement en série entière des fonctions $z \mapsto e^z$, $z \mapsto \ln(1-z)$ et $z \mapsto 1/(1-z)$, en précisant leur rayon de convergence.

IV.2. On suppose la variable aléatoire T définie pour $a = -\infty$. Quelle relation existe-t-il entre $\mathrm{P}(T > n)$ et $\mathrm{P}(U_n < b)$?

IV.3. On définit les fonctions réelles de deux variables réelles:

$$f_+(z, w) = \exp \sum_{n=1}^{+\infty} \frac{z^n}{n} \mathrm{E}_{\{S_n > 0\}}\left[w^{S_n}\right] \text{ et } f_-(z, w) = \exp \sum_{n=1}^{+\infty} \frac{z^n}{n} \mathrm{E}_{\{S_n \leq 0\}}\left[w^{S_n}\right],$$

et on note $\pi(w)$ la fonction génératrice de X_1: $\pi(w) = \mathrm{E}[w^{X_1}]$.

IV.3.a. Montrer que les fonctions f_+ et f_- sont bien définies pour $|z| < 1$ et $|w| \leq 1$.

IV.3.b. Démontrer l'égalité:

$$f_+(z,w)\, f_-(z,w) = \frac{1}{1 - z\pi(w)}$$

pout tout $(z,w) \in]-1,1[\times [-1,1]$.

On admettra que l'on peut écrire $f_+(z,w)$ et $f_-(z,w)$ sous la forme $\sum_{n=0}^{+\infty} \alpha_n(w)\, z^n$ et $\sum_{n=0}^{+\infty} \beta_n(w)\, z^n$, où $\alpha_n(w)$ et $\beta_n(w)$ sont des produits de puissances de termes de la forme $\mathrm{E}_{\{S_n>0\}}[w^{S_n}]$ et $\mathrm{E}_{\{S_n\le 0\}}[w^{S_n}]$, respectivement.

IV.3.c. Montrer l'identité:

$$\pi(w)^n = \sum_{k=0}^{n} \alpha_k(w)\, \beta_{n-k}(w)$$

pour tout $n \in \mathrm{N}$ et tout $w \in [-1,1]$.

IV.4. Pour tout entier n, on définit la variable aléatoire L_n égale au premier entier $k \in \{0,1,\ldots,n\}$ tel que $S_k = U_n$.

IV.4.a. Montrer que $L_n = 0$ si et seulement si, quel que soit l'entier $i \in \{1,2,\ldots,n\}$, on a $S_i \le 0$.

IV.4.b. Montrer que pour tout entier $k \in \{1,2,\ldots,n\}$, $L_n = k$ si et seulement si les conditions $S_k > 0$, $S_k > S_1$, ..., $S_k > S_{k-1}$ et $S_k \ge S_{k+1}$, $S_k \ge S_{k+2}$, ..., $S_k \ge S_n$ sont toutes simultanément satisfaites.

IV.5. Pour tout entier n et tout réel w tel que $|w| \le 1$, on pose $p_n(w) = \mathrm{E}_{\{L_n=n\}}[w^{S_n}]$ et $q_n(w) = \mathrm{E}_{\{L_n=0\}}[w^{S_n}]$ où $w \in [-1,1]$.

IV.5.a. Montrer, pour tout entier $k \in \{0,1,\ldots,n\}$, l'égalité:

$$\mathrm{E}_{\{L_n=k\}}[w^{S_n}] = \mathrm{E}_{\{L_k=k\}}[w^{S_k}]\, \mathrm{E}_{\{L_{n-k}=0\}}[w^{S_{n-k}}].$$

IV.5.b. En déduire la relation:

$$\pi(w)^n = \sum_{k=0}^{n} p_k(w)\, q_{n-k}(w)$$

pour tout $n \in \mathrm{N}$ et tout $w \in [-1,1]$.

IV.6. A partir des résultats des questions IV.3.c et IV.5.b, établir les identités $\alpha_n(w) = p_n(w)$ et $\beta_n(w) = q_n(w)$ pour tout $n \in \mathrm{N}$ et tout $w \in [-1,1]$.

IV.7. Démontrer l'égalité:

$$\mathrm{E}[w^{U_n}] = \sum_{k=0}^{n} p_k(w)\, q_{n-k}(1)$$

pour tout $n \in \mathrm{N}$ et tout $w \in [-1,1]$. En déduire que pour tous réels z tel que $|z|<1$ et w tel que $|w| \leq 1$, on a la relation:

$$\sum_{n=0}^{+\infty} \mathrm{E}[w^{U_n}]\, z^n = f_+(z,w)\, f_-(z,1)\,.$$

IV.8. On note $\phi_n(t)$ et $\psi_n(t)$ les fonctions caractéristiques de U_n et de la variable aléatoire $S_n^+ = \max(0, S_n)$, respectivement. Démontrer l'égalité:

$$\sum_{n=0}^{+\infty} \phi_n(t)\, z^n = \exp \sum_{n=1}^{+\infty} \psi_n(t) \frac{z^n}{n}\,.$$

IV.9. Montrer que l'on a, pour tout entier $n \in \mathrm{N}^*$:

$$\mathrm{E}[U_n] = \sum_{k=1}^{n} \frac{1}{k} \mathrm{E}[S_k^+].$$

Examiner le cas particulier où la loi de X_1 est gaussienne de moyenne nulle et de variance unité.