

ESSEC II 2014

I. Autour de la loi de Benford

Soit $x \in \mathbb{R}$. On note $\lfloor x \rfloor$ sa partie entière, c'est à dire le plus grand entier relatif inférieur ou égal à x , et $\{x\}$ sa partie fractionnaire : $\{x\} = x - \lfloor x \rfloor$. On note $\log(z)$ le logarithme en base 10 du réel $z > 0$. On a donc $\log(z) = \frac{\ln(z)}{\ln(10)}$. On rappelle en particulier les propriétés suivantes, qu'on pourra utiliser sans démonstration

$$\forall z > 0, 10^{\log(z)} = z$$

$$\forall z > 0, \forall z' > 0, \log(z \times z') = \log(z) + \log(z')$$

$$\forall a \in \mathbb{R}, \log(10^a) = a$$

On a par exemple $\log(100) = 2$ et $\log(\sqrt{10}) = \frac{1}{2}$.

1. a) Montrer que pour tout réel x positif et non nul, on a

$$x = 10^{\{\log(x)\}} \times 10^{\lfloor \log(x) \rfloor}$$

Cette décomposition est dite *notation scientifique* de x .

b) Montrer que pour tout $x > 0$, le couple $(10^{\{\log(x)\}}, \lfloor \log(x) \rfloor)$ est l'unique couple (α, n) dans $[1, 10[\times \mathbb{Z}$ tel que $x = \alpha \times 10^n$.

c) Soit $x > 0$. On pose $\gamma = \lfloor 10^{\{\log(x)\}} \rfloor$. Montrer que $\gamma \in \{1, 2, \dots, 9\}$.

2. Pour tout entier naturel k tel que $1 \leq k \leq 9$, on pose $p_k = \log(1 + \frac{1}{k})$. Montrer que $\sum_{k=1}^9 p_k = 1$.
 $(p_k)_{1 \leq k \leq 9}$ définit donc une loi de probabilité sur $\{1, 2, \dots, 9\}$ dite *loi de Benford*.

3. Soit X une variable aléatoire réelle strictement positive. On suppose que la variable aléatoire réelle $Y = \{\log(X)\}$ suit une loi uniforme sur $[0, 1[$.

a) Soit $k \in \{1, 2, \dots, 9\}$. Montrer que $\lfloor 10^Y \rfloor = k \Leftrightarrow k \leq 10^Y < k + 1$.

b) On considère la variable aléatoire $\Gamma = \lfloor 10^{\{\log(X)\}} \rfloor$ égale au premier chiffre significatif de X . Déterminer la loi de la variable aléatoire Γ .

4. Soit Y une variable aléatoire réelle admettant une densité g continue sur \mathbb{R} . On suppose que

(h₁) g atteint son maximum M en un unique point $a_0 \in \mathbb{R}$.

(h₂) g est croissante sur $] - \infty, a_0]$ et décroissante sur $[a_0, +\infty[$.

a) i. Montrer que pour tout $y \in \mathbb{R}$ et tout $n \in \mathbb{Z}$, $\{y\} = \{y - n\}$.

ii. Dédire que pour tout $n \in \mathbb{Z}$, la loi de $\{Y\}$ est identique à celle de $\{Y - n\}$.

iii. Déterminer une fonction de densité \tilde{g} continue de la variable aléatoire $Y - \lfloor a_0 \rfloor$.

iv. Montrer que \tilde{g} admet un unique maximum en un point $\tilde{a}_0 \in [0, 1[$.

v. Montrer que \tilde{g} vérifie les conditions (h₁) et (h₂) ci-dessus avec \tilde{a}_0 remplaçant a_0 .

On supposera donc désormais que $a_0 \in [0, 1[$. On fixe $x \in]0, 1[$ et on note $I_{n,x} = [n, n + x[$ pour tout $n \in \mathbb{Z}$.

b) i. Soit φ une fonction positive continue et croissante sur $[0, 1]$. Montrer, en utilisant un changement de variable, que $\int_0^x \varphi(t) dt \leq x \times \int_0^1 \varphi(u) du$.

ii. Dédurre que pour tout $n \in \mathbb{Z}$ tel que $n \leq -1$, on a $\frac{1}{x} \int_{I_{n,x}} g(t) dt \leq \int_n^{n+1} g(t) dt$.

On **admettra** qu'on montrerait de même que pour $n \geq 2$, $\frac{1}{x} \int_{I_{n,x}} g(t) dt \leq \int_{n-1+x}^{n+x} g(t) dt$

iii. Montrer que $\frac{1}{x} \sum_{n \geq 1} \int_{I_{-n,x}} g(t) dt \leq \int_{-\infty}^0 g(t) dt$ et que $\frac{1}{x} \sum_{n \geq 2} \int_{I_{n,x}} g(t) dt \leq \int_{1+x}^{+\infty} g(t) dt$

iv. Montrer que $\int_{I_{0,x}} g(u) du \leq xM$ et que $\int_{I_{1,x}} g(u) du \leq xM$

v. Conclure que $\frac{1}{x} \sum_{n \in \mathbb{Z}} \int_{I_{n,x}} g(u) du = \frac{1}{x} \sum_{n \geq 1} \int_{I_{-n,x}} g(u) du + \frac{1}{x} \sum_{n \geq 0} \int_{I_{n,x}} g(u) du \leq 1 + 2M$.

On montrerait de même que $\frac{1}{x} \sum_{n \in \mathbb{Z}} \int_{I_{n,x}} g(u) du \geq 1 - 2M$, inégalité qu'on **admettra**.

vi. Montrer que l'événement $\{Y\} < x$ est égal à $\bigcup_{n \in \mathbb{Z}} [Y \in I_{n,x}]$.

vii. Dédurre que $|\mathbb{P}(\{Y\} < x) - x| \leq 2M$.

5. Soit $(Z_n)_{n \geq 1}$ une suite de variables aléatoires, où Z_n suit une loi exponentielle de paramètre $\frac{1}{n}$. On pose $X_n = 10^{\sqrt{Z_n}}$ et $Y_n = \log(X_n) = \sqrt{Z_n}$.

a) Déterminer une densité g_n de la loi Y_n , continue sur \mathbb{R} .

b) Étudier les variations de g_n sur \mathbb{R}_+ et déterminer son maximum.

c) Montrer que pour tout $x \in]0, 1[$, $|\mathbb{P}(\{Y_n\} < x) - x| \leq 2\sqrt{\frac{2}{n}}e^{-\frac{1}{2}}$.

d) Seulement pour les cubes :

Montrer que la suite $(\{Y_n\})_{n \geq 1}$ converge en loi vers la loi uniforme sur $[0, 1[$.

II. Répartition des valeurs dans une table numérique

Henri Poincaré (1854-1912) a proposé au début du 20^{ème} siècle une façon originale d'étudier la répartition des valeurs d'une table numérique en montrant que pour un bon choix d'une fonction F de période assez grande par rapport à l'incrément des valeurs de la table, la moyenne des valeurs prises par F sur la table sera petite, ce qui indique une certaine forme d'équilibre dans la répartition de ces valeurs.

Poincaré considère l'exemple des valeurs d'une table de logarithmes :

$$z_n = \ln \left(1 + \frac{n}{100\,000} \right)$$

pour $n = 1, 2, 3, \dots, 10\,000$ et pose $F(y) = \sin(1000 \times \pi \times y)$, fonction de période $\frac{1}{500}$, grande par rapport à l'incrément $\frac{1}{100\,000}$ dans la table. Il s'intéresse à la moyenne des valeurs de F sur la table, c'est à dire à :

$$S = \frac{1}{10\,000} \sum_{k=1}^{10\,000} F(z_k)$$

et désire montrer que cette valeur est petite.

Posons

$$J = \frac{1}{10\,000} \int_{\frac{1}{2}}^{10\,000 + \frac{1}{2}} F \left(\ln \left(1 + \frac{x}{100\,000} \right) \right) dx$$

6. a) Soit φ une fonction de classe \mathcal{C}^2 sur $[\frac{1}{2}, +\infty[$. On suppose qu'il existe un réel $M > 0$ tel que, pour tout $u \in [\frac{1}{2}, +\infty[$, $|\varphi''(u)| \leq M$.
Montrer que pour $n \in \mathbb{N}^*$ et $|h| \leq \frac{1}{2}$, on a

$$-\frac{M}{8} \leq \varphi(n+h) - \varphi(n) - h\varphi'(n) \leq \frac{M}{8}$$

- b) Dédurre que pour tout $n \in \mathbb{N}^*$,

$$-\frac{M}{8} \leq \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} \varphi(u) du - \varphi(n) \leq \frac{M}{8}$$

- c) À partir de cette question, dans la partie II, on utilise des fonctions trigonométriques qui ne sont pas connues en voie E. Je les ai donc supprimées mais vous pouvez aller les consulter sur internet.

III. Sur les nombres normaux

Dans cette partie, on se donne un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et on notera comme d'habitude, sous réserve d'existence, $\mathbb{E}(X)$ et $\mathbb{V}(X)$ l'espérance et la variance d'une variable aléatoire réelle X .

On commence par rappeler les deux points de théorie suivants.

- (i) Pour toute suite d'événements $(C_k)_{k \in \mathbb{N}}$ dans Ω , on a $\mathbb{P}(\bigcup_{k=0}^{+\infty} C_k) \leq \sum_{k=0}^{\infty} \mathbb{P}(C_k)$ avec la convention que cette série vaut $+\infty$ si elle diverge.
- (ii) Si $(C_k)_{k \in \mathbb{N}}$ est une suite décroissante d'événements dans Ω , au sens où $\forall k \geq 0, C_k \supset C_{k+1}$, on a $\mathbb{P}(\bigcap_{k \geq 0} C_k) = \lim_{k \rightarrow +\infty} \mathbb{P}(C_k)$

On rappelle aussi l'*inégalité de Markov* : si Z est une variable aléatoire positive admettant une espérance $\mathbb{E}(Z)$, pour tout $\alpha > 0$, on a $\mathbb{P}([Z > \alpha]) \leq \frac{\mathbb{E}(Z)}{\alpha}$.

On considère ici le tirage au sort d'un nombre réel entre 0 et 1 qu'on modélise de la façon suivante : $(X_n)_{n \geq 1}$ est une suite de variables aléatoires indépendantes, de même loi uniforme à valeurs dans $\{0, 1, 2, \dots, 9\}$. Les $(X_n)_{n \geq 1}$ représentent les décimales du nombre tiré au hasard c'est à dire que ce nombre est $\sum_{k=1}^{+\infty} \frac{X_k}{10^k}$.

On définit enfin pour tout $k \geq 1$ une variable aléatoire Y_k à valeurs 0 ou 1 par $Y_k = 1$ si $X_k = 1$ et $Y_k = 0$ si $X_k \neq 1$.

7. a) Montrer que les variables Y_k sont indépendantes et de même loi que l'on précisera.

- b) Déterminer $\mathbb{E}(Y_k)$ et $\mathbb{V}(Y_k)$.

On pose $S_n = \sum_{k=1}^n Y_k$. par conséquent, $\frac{S_n}{n}$ représente la fréquence des 1 dans la suite des décimales du nombre tiré.

- c) Calculer $\mathbb{V}(\frac{S_n}{n})$ en fonction de n .

- d) Soit $\varepsilon > 0$ fixé. Montrer $\mathbb{P}(|\frac{S_n}{n} - \frac{1}{10}| > \varepsilon) \leq \frac{\mathbb{V}(\frac{S_n}{n})}{\varepsilon^2}$.

- e) En déduire que

$$\lim_{n \rightarrow +\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - \frac{1}{10}\right| > \varepsilon\right) = 0$$

On va dans la suite améliorer ce résultat en montrant qu'en fait pour la plupart des nombres réels, la fréquence des 1 dans leurs décimales vaut $\frac{1}{10}$.

8. *a)* On pose $A = \bigcap_{N \geq 1} \bigcup_{k=N}^{\infty} A_k$. Montrer que A est l'ensemble des ω qui appartiennent à une infinité d'événements A_k .
- b)* On pose, pour tout $N \geq 1$, $B_N = \bigcup_{k=N}^{\infty} A_k$. Montrer que $\forall N \geq 0, B_N \supset B_{N+1}$.
- c)* Dédurre que $\lim_{N \rightarrow +\infty} \mathbb{P}(B_N) = \mathbb{P}(A)$.
- d)* On suppose que $\sum_{k=1}^{\infty} \mathbb{P}(A_k) < +\infty$.
- i.* Que vaut $\lim_{N \rightarrow +\infty} \sum_{k=N}^{+\infty} \mathbb{P}(A_k)$?
- ii.* Conclure que $\mathbb{P}(A) = 0$.
9. On pose, pour tout $k \geq 1$, $Y'_k = Y_k - \frac{1}{10}$.
- a)* Montrer que $\frac{S_n}{n} - \frac{1}{10} = \frac{1}{n} \sum_{k=1}^n Y'_k$.
- b)* Montrer que les variables Y'_k sont indépendantes, d'espérance nulle et telles que $|Y'_k| \leq 1$.
- c)* Montrer que $\mathbb{E} \left(\left(\sum_{k=1}^n Y'_k \right)^4 \right) \leq n + 3n(n-1)$.
- d)* Dédurre que
- $$\mathbb{E} \left(\left(\frac{S_n}{n} - \frac{1}{10} \right)^4 \right) \leq \frac{3}{n^2}$$
- e)* On pose, pour $k \geq 1$, $A_k = \left[\left(\frac{S_k}{k} - \frac{1}{10} \right)^4 > \frac{1}{\sqrt{k}} \right]$. Montrer que $\mathbb{P}(A_k) \leq \frac{3}{k^{\frac{3}{2}}}$.
- f)* Dédurre que $\sum_{k \geq 1} \mathbb{P}(A_k)$ est une série convergente.
- g)* On considère l'événement $A = \{ \omega \in \Omega, \left(\frac{S_k(\omega)}{k} - \frac{1}{10} \right)^4 > \frac{1}{\sqrt{k}} \text{ pour une infinité de } k \}$. Montrer que $\mathbb{P}(A) = 0$.
- h)* Dédurre qu'avec probabilité 1, on peut trouver N tel que pour tout $k \geq N$, $\left| \frac{S_k}{k} - \frac{1}{10} \right| \leq \frac{1}{\sqrt[8]{k}}$.
- i)* Conclure qu'avec probabilité 1, on a, $\lim_{k \rightarrow +\infty} \frac{S_k}{k} = \frac{1}{10}$.