DM 3 - sujet A

THÈMES : VARIABLES À DENSITÉ

A. Préliminaires

On note $(c_n)_{n\in\mathbb{N}^*}$ la suite réelle définie pour tout entier n strictement positif par :

$$c_n = \int_0^1 \frac{x^{n-1}}{1+x}\,\mathrm{d}x.$$

1. Montrer que la suite $(c_n)_{n\in\mathbb{N}^*}$ est décroissante.

2. Montrer que pour tout $n\in\mathbb{N}^*$, on a : $\quad c_{n+1}+c_n = \dfrac{1}{n}$.

3. Établir, pour tout entier n supérieur ou égal à 2

$$\frac{1}{n} \leqslant 2c_n \leqslant \frac{1}{n-1}.$$

En déduire un équivalent simple de c_n quand n tend vers l'infini.

4. Calculer c_1 et prouver, pour tout entier n supérieur ou égal à 2 , l'égalité :

$$c_n = (-1)^n\left(\sum_{k=1}^{n-1}\frac{(-1)^{k+1}}{k} - \ln 2\right)$$

5. Écrire un programme Python qui prend en argument $n\in\mathbb{N}^*$ et renvoie c_n.

B. Étude d'une suite de variables aléatoires à densité

Pour tout entier n strictement positif, on note f_n l'application de $\mathbb{R}$ dans $\mathbb{R}$ définie par :

$$f_n(t) = \begin{cases} 0 & \text{si} \quad t<1 \\ \frac{1}{c_n t^n(1+t)} & \text{si} \quad t\geqslant 1 \end{cases}$$

6. À l'aide d'un changement de variable, établir pour tout entier n strictement positif et pour tout réel x supérieur ou égal à 1, l'égalité :

$$\int_1^x \frac{1}{t^n(1+t)}\,\mathrm{d}t = \int_{1/x}^1 \frac{u^{n-1}}{1+u}\,\mathrm{d}u.$$

7. En déduire que, pour tout entier n strictement positif, f_n est une densité de probabilité.

Dans la suite, on suppose que $(X_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires définies sur le même espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$, telle que, pour tout entier n strictement positif, X_n prend ses valeurs dans $[1,+\infty[$ et admet f_n comme densité. On note F_n la fonction de répartition de X_n.

8. Pour quelles valeurs de n la variable aléatoire X_n admet-elle une espérance? Dans le cas où l'espérance de X_n existe, calculer cette espérance en fonction de c_n et de c_{n-1}.

9. Dans cette question, exclusivement, on suppose que $n=1$. Préciser la fonction F_1.

- **a)** En déduire l'ensemble des réels y vérifiant $\mathbf{P}\left([X_1\leqslant y]\right)\geqslant \dfrac{1}{2}$.
- **b)** Déterminer une densité de la variable aléatoire $Z=\ln(X_1)$.

10. Soit x un réel strictement supérieur à 1.

- **a)** Justifier l'encadrement : $\quad 0 \leqslant \displaystyle\int_{1/x}^1 \frac{u^n}{(1+u)^2}\,\mathrm{d}u \leqslant \frac{1}{n+1}$.

b) En déduire la limite suivante : $\lim_{n\to+\infty}\left(\int_{1/x}^{1}\frac{u^n}{(1+u)^2}\,\mathrm{d}u\right)$.

c) Transformer, pour tout entier naturel n non nul, $\mathrm{F}_n(x)$ à l'aide d'une intégration par parties et en déduire l'égalité suivante : $\lim_{n\to+\infty}\mathrm{F}_n(x)=1$.

11. Que vaut $\lim_{n\to+\infty}\mathrm{F}_n(x)$ si x est un réel inférieur ou égal à 1 ?
Nous verrons au chapitre Convergences des variables aléatoires, que la suite de variables aléatoires $(\mathrm{X}_n)_{n\in\mathbb{N}^*}$ *converge en loi.*

– FIN –

DM 3 - sujet *

THÈME : VARIABLES DISCRÈTES ET À DENSITÉ

Distance en variation totale

Les parties I et II sont indépendantes.

Partie I - le cas continu

On dit qu'une fonction $h : \mathbb{R} \to \mathbb{R}$ est lipschitzienne s'il existe une constante $C > 0$ pour laquelle :

$$\forall (x, y) \in \mathbb{R}^2, \qquad |h(x) - h(y)| \leqslant C|x - y|.$$

On note $\mathscr{L}$ l'ensemble de ces fonctions lipschitziennes sur $\mathbb{R}$. Si $h \in \mathscr{L}$, on pose :

$$N(h) = \sup_{\substack{x,y\in\mathbb{R}\\ x\neq y}} \frac{|h(x) - h(y)|}{|x - y|}.$$

Soit $\mathscr{M}$ l'ensemble des variables aléatoires réelles définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ possédant une densité continue sur $\mathbb{R}$ et admettant une espérance.

1. Soit $X \in \mathscr{M}$. Montrer que pour tout $h \in \mathscr{L}$, $h(X)$ admet une espérance.

2. Soit $(X, Y) \in \mathscr{M}^2$.
Montrer qu'il existe une constante $D \geqslant 0$ dépendant de X et Y, telle que, pour tout $h \in \mathscr{L}$:

$$\big|\mathbf{E}\big(h(X)\big) - \mathbf{E}\big(h(Y)\big)\big| \leqslant D\,N(h).$$

On pose, pour tout $(X, Y) \in \mathscr{M}^2$

$$d(X, Y) = \sup_{\substack{h\in\mathscr{L}\\ N(h)\leqslant 1}} \big|\mathbf{E}\big(h(X)\big) - \mathbf{E}\big(h(Y)\big)\big|.$$

3. Montrer que d vérifie :

- **a)** $d(X, X) = 0$;
- **b)** $d(X, Y) = d(Y, X)$;
- **c)** $d(X, Z) \leqslant d(X, Y) + d(Y, Z)$.
- **d)** Soient X et $\tilde{X}$ deux variables de $\mathscr{M}$ ayant même loi. Justifier que $d(X, Y) = d(\tilde{X}, Y)$.

4. Montrer que

$$\big|\mathbf{E}(X - Y)\big| \leqslant d(X, Y) \leqslant \mathbf{E}(|X - Y|).$$

5. On suppose que X suit la loi normale $\mathscr{N}\left(\lambda, \sigma^2\right)$ et Y la loi normale $\mathscr{N}\left(\mu, \sigma^2\right)$. Montrer que

$$d(X, Y) = |\lambda - \mu|.$$

6. On suppose que X suit la loi normale $\mathscr{N}\left(0, \sigma^2\right)$ et Y la loi normale $\mathscr{N}(0, 1)$. Montrer que

$$d(X, Y) \leqslant \sqrt{\frac{2}{\pi}}|1 - \sigma|.$$

Partie II - le cas discret

Dans la suite, on admet ce résultat sur les produits de Cauchy :

- Pour deux suites $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, on définie la suite $(c_n)_{n\in\mathbb{N}}$

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

- Si les séries $\sum a_n$ et $\sum b_n$ sont absolument convergentes alors la séries $\sum c_n$ est aussi absolument convergente et

$$\sum_{n=0}^{+\infty} c_n = \left(\sum_{i=0}^{+\infty} a_i\right)\left(\sum_{j=0}^{+\infty} b_j\right).$$

Toutes les variables aléatoires de ce problème sont définies sur le même espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et à valeurs dans $\mathbb{N}$. De plus, pour X, Y deux variables aléatoires, on pose

$$d(\mathrm{X},\mathrm{Y}) = \sum_{n=0}^{+\infty} \left|\mathbf{P}([\mathrm{X}=n]) - \mathbf{P}([\mathrm{Y}=n])\right|.$$

7. Justifier que la quantité $d(\mathrm{X},\mathrm{Y})$ est bien définie.

- *Exemples*

8. Soient $p\,q, \in]0,1[$, $\mathrm{X} \hookrightarrow \mathscr{B}(p)$ et $\mathrm{Y} \hookrightarrow \mathscr{B}(q)$. Exprimer $d(\mathrm{X},\mathrm{Y})$ en fonction de p et q.

9. Soient $p \in]0,1[$, $\mathrm{X} \hookrightarrow \mathscr{B}(p)$ et $\mathrm{Y} \hookrightarrow \mathscr{P}(p)$.
Établir l'égalité $d(\mathrm{X},\mathrm{Y}) = 2p(1-\mathrm{e}^{-p})$ et en déduire la majoration $d(\mathrm{X},\mathrm{Y}) \leqslant 2p^2$.

- *Distance et sommes*

10. **a)** Soient X, Y, Z, T quatre variables aléatoires mutuellement indépendantes. Montrer que pour tout $n \in \mathbb{N}$

$$|\mathbf{P}(\mathrm{X}+\mathrm{Y}=n) - \mathbf{P}(\mathrm{Z}+\mathrm{T}=n)| \leqslant \sum_{j=0}^{n} \mathbf{P}(\mathrm{Y}=j)\left|\mathbf{P}(\mathrm{X}=n-j) - \mathbf{P}(\mathrm{Z}=n-j)\right| + \sum_{i=0}^{n} \mathbf{P}(\mathrm{Z}=i)\left|\mathbf{P}(\mathrm{Y}=n-i) - \mathbf{P}(\mathrm{T}=n-i)\right|.$$

b) En déduire que $d(\mathrm{X}+\mathrm{Y},\mathrm{Z}+\mathrm{T}) \leqslant d(\mathrm{X},\mathrm{Z}) + d(\mathrm{Y},\mathrm{T})$.

- *Applications*

11. Justifier que si $\mathrm{U} \hookrightarrow \mathscr{B}(n,p)$ et $\mathrm{V} \hookrightarrow \mathscr{B}(n,q)$ alors $d(\mathrm{U},\mathrm{V}) \leqslant 2n|p-q|$.
On pourra supposer les variables U *et* V *indépendantes. On rappelle aussi que si* $\mathrm{A} \hookrightarrow \mathscr{B}(n,p)$, $\mathrm{B} \hookrightarrow \mathscr{B}(m,p)$ *avec* A *et* B *indépendantes, alors* $\mathrm{A}+\mathrm{B} \hookrightarrow \mathscr{B}(n+m,p)$.

12. **a)** Soient $\mathrm{U} \hookrightarrow \mathscr{B}(n,p)$ et $\mathrm{S} \hookrightarrow \mathscr{P}(np)$. Prouver l'inégalité $d(\mathrm{U},\mathrm{S}) \leqslant 2np^2$.
b) Retrouver le théorème de convergence en loi des lois de Poisson.

13. Soient $\lambda, \mu \in \mathbb{R}_*^+$. En utilisant les résultats précédents, montrer que

$$d(\mathrm{X},\mathrm{Y}) \leqslant 2|\lambda-\mu| \quad \text{si} \quad \mathrm{X} \hookrightarrow \mathscr{P}(\lambda), \quad \mathrm{Y} \hookrightarrow \mathscr{P}(\mu).$$

Indication. On pourra remarquer que $d(\mathrm{X},\mathrm{Y}) \leqslant d(\mathrm{X},\mathrm{X}') + d(\mathrm{X}',\mathrm{Y}') + d(\mathrm{Y},\mathrm{Y}')$ *et utiliser la question 12.b).*

14. **a)** Écrire un programme python qui prend en arguments λ, n et renvoie

$$[\mathbf{P}(\mathrm{X}=0), \ldots, \mathbf{P}(\mathrm{X}=n)] \quad \text{où} \quad \mathrm{X} \hookrightarrow \mathscr{P}(\lambda)$$

b) Faire de même pour

$$[\mathbf{P}(\mathrm{Y}=0), \ldots, \mathbf{P}(\mathrm{Y}=n)] \quad \text{où} \quad \mathrm{Y} \hookrightarrow \mathscr{B}(n,p).$$

On peut utiliser la commande `math.comb(n,i)` *pour calculer* $\binom{n}{i}$.

c) Tracer $\left(d(\mathrm{X}_n,\mathrm{Y})\right)_{n\in[[1;20]]}$ où $\mathrm{X}_n \hookrightarrow \mathscr{B}\left(n,\frac{1}{n}\right)$ et $\mathrm{Y} \hookrightarrow \mathscr{P}(1)$.
Pour évaluer numériquement $d(\mathrm{X}_n,\mathrm{Y}_n)$*, on peut se limiter à calculer les 50 premiers termes de la somme.*

– FIN –