M61B, Probabilités et intégration
Licence 3e année
Parcours mathématiques

DS1-2h

19 Mars 2022

- *La note maximale est 20, il n'y a pas besoin de traiter tout le DS pour l'obtenir. Parcourez rapidement tous les exercices avant de vous lancer, afin de vous concentrer en priorité sur ceux que vous maîtrisez le mieux.*
- *Les exercices sont indépendants entre eux, mais dans un même exercice, les questions ne le sont pas : pensez à utiliser les questions précédentes.*
- *Attention, les questions faciles valent moins de points, attaquez vous aussi à des questions difficiles !*
- *Il est possible d'admettre certaines questions pour traiter les suivantes.*

Exercice 1 : Questions de cours, 4 pts

1) Soit E un ensemble, qu'est ce qu'une tribu sur E ?
2) Qu'est ce qu'un espace de probabilité ?
3) Donner la définition d'une *fonction mesurable* de $(E, \mathscr{T})$ dans $(F, \mathscr{T}')$.
4) Donner la définition et les conditions d'existence de l'espérance, pour une variable aléatoire discrète X à valeurs dans $\mathbb{N}$ et pour une variable aléatoire Y, réelle à densité f_Y.
5) Donner la définition et la fonction de répartition d'une variable de loi exponentielle de paramètre $\lambda > 0$.

Exercice 2 : Quizz, 5 pts

Pour chaque affirmation, indiquer si elle est vraie ou fausse. Si elle est vraie, la justifier, si elle est fausse, l'infirmer, par exemple à l'aide d'un contre exemple.

1) Les ensembles

$$\mathscr{T} := \{A \subset \mathbb{N},\ A \text{ ne contient que des entiers pairs}\}$$

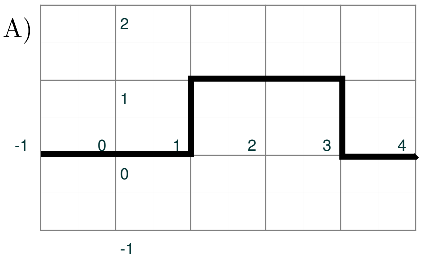
et

$$\mathscr{T}' := \{A \subset \mathbb{N},\ card(A) < \infty\}$$

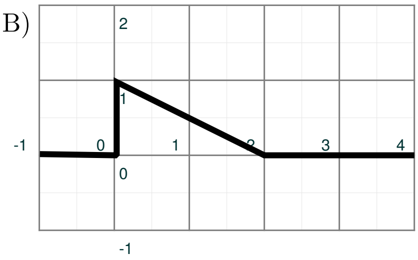
sont ils des tribus sur $\mathbb{N}$?

2) Parmi les fonctions représentées ci-dessous, indiquer lesquelles sont des densités de probabilité.

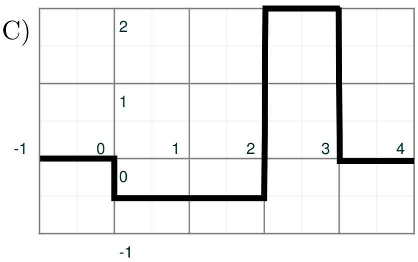
A)



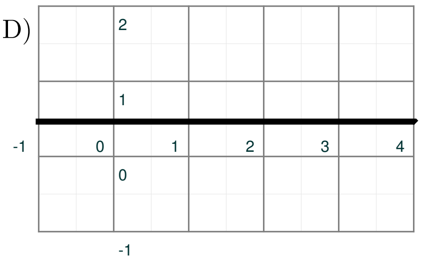
B)



C)



D)



3) Soient X, Y et Z trois variables aléatoires à valeurs dans $\mathbb{N}$. On suppose que X est indépendante de Y et de Z, et que Y et Z sont indépendantes. Alors, pour tous entiers x, y, $z \in \mathbb{N}$, on a

$$\mathbb{P}(X = x, Y = y, Z = z) = \mathbb{P}(X = x)\mathbb{P}(Y = y)\mathbb{P}(Z = z).$$

Exercice 3 : Grippe bovine, 8 pts

On souhaite dépister, grâce à un contrôle sanguin, quels sont les bovins atteints par un certain virus dans un troupeau de n bêtes. On sait que chaque bête a, indépendamment des autres, une probabilité $p \in]0,1[$ d'être atteinte. Comme chaque test coûte cher, on considère la stratégie suivante : on forme des groupes de m bêtes (où m divise n) ; dans chaque groupe, on mélange les prélèvements sanguins des bovins en un seul « superéchantillon », dans lequel on teste la présence du virus (il suffit qu'une seule bête du groupe soit infectée pour que le test soit positif). Si on décèle la présence du virus dans un super-échantillon, on teste toutes les bêtes du groupe séparément. On appelle cette stratégie “test de super-échantillon de taille m”, et on appelle “test individuel” la stratégie consistant à tester automatiquement chaque bête individuellement.

1) Quelle est la loi du nombre de bêtes atteintes dans le troupeau ? Rappeler (sans calcul) son espérance et sa variance.

2) On veut calculer l'espérance $\mathbb{E}(N)$ du nombre N de tests effectués pour la stratégies des super-échantillons.

(i) Quelle est la probabilité qu'un groupe donné soit testé positif ? Quel est alors le nombre de tests effectués sur le groupe ?

(ii) En déduire l'espérance de N en fonction de n, m et p.

3) On fixe un $m \geq 2$ quelconque, on veut comparer les deux stratégies.

(i) Que vaudrait $\mathbb{E}(N)$ pour la stratégie des tests individuels ?

(ii) Si la probabilité d'infection est très grande (p très proche de 1), est-il plus intéressant de tester des super-échantillons de taille m, plutot que d'effectuer des tests individuels ? Justifier à l'aide de l'expression de $\mathbb{E}(N)$ obtenues aux questions 2) - (ii) et 3) - (i).

(iii) Même question si p est très petit (très peu d'infections).

4) On souhaite maintenant étudier l'efficacité des tests de super-échantilons de taille n par rapport aux test individuels.

(i) On pose $q = (1-p)^m$. Que représente cette quantité ?

(ii) On note à nouveau N le nombre de tests effectués pour la stratégie des super-échantillons. Exprimer $n - \mathbb{E}(N)$ en fonction de q.

(iii) On pose $p_0(m) = 1 - \frac{1}{m^{1/m}}$. Pour quelles valeurs de p l'utilisation de super-échantillons est-elle plus intéressante que l'utilisation de tests individuels ?

Problème : Théorème de Scheffé, 14 pts

Dans tout le problème, on fixe un espace de probabilités $(\Omega, \mathscr{F}, \mathbb{P})$ sur lequel seront définies toutes les variables aléatoires considérées.

Partie I : Théorème de Scheffé

1) Étant donnée une fonction $g : \mathbb{R} \to \mathbb{R}$, on note

$$g^+ : x \mapsto g(x)\mathbf{1}_{\{g(x)>0\}}$$

la partie positive de g, et

$$g^- : x \mapsto -g(x)\mathbf{1}_{\{g(x)<0\}}$$

sa partie négative. Exprimer les fonctions g et $|g|$ en fonction de g^+ et g^-.

2) Soit $(f_n)_{n\in\mathbb{N}}$ une suite de densités de probabilités sur $\mathbb{R}$. Soit f une autre densité de probabilités sur $\mathbb{R}$, on suppose que f_n converge simplement vers f.

(i) Justifier les inégalités

$$(f - f_n)^+ \leq |f - f_n| \quad \text{et} \quad (f - f_n)^- \leq |f - f_n|,$$

et montrer que $(f - f_n)^+$ et $(f - f_n)^-$ convergent simplement vers 0.

(ii) Montrer que

$$0 \leq (f - f_n)^+(x) \leq f(x).$$

(iii) Justifier que $(f - f_n)^+$ et $(f - f_n)^-$ sont toutes les deux intégrables. Montrer que

$$\int_{\mathbb{R}} (f - f_n)^+(x)dx = \int_{\mathbb{R}} (f - f_n)^-(x)dx = \frac{1}{2}\int_{\mathbb{R}} |f - f_n|(x)dx.$$

(iv) Déduire de ce qui précède la valeur de $\lim_{n\to\infty} \int_{\mathbb{R}} (f - f_n)^+(x)dx$, puis déterminer celle de $\lim_{n\to\infty} \int_{\mathbb{R}} |f - f_n|(x)dx$.

3) On note F_n et F les fonctions de répartition associées à f_n et f respectivement.

(i) Rappeller l'expression de F_n et F en fonction de f_n et f. Soit $x \in \mathbb{R}$, montrer que

$$|F_n(x) - F(x)| \leq \int_{\mathbb{R}} |f_n - f|\,(y)dy.$$

(ii) Déduire de ce qui précède le Théorème de Scheffé, qui affirme que si une suite de densités $(f_n)_{n\in\mathbb{N}}$ converge simplement vers une densité f, alors les fonctions de répartitions correspondantes (F_n) convergent simplement vers la fonction de répartition F associée à f.

Partie II : Réciproque du théorème de Scheffé

4) Pour tout $n \in \mathbb{N}$, on définit

$$f_n(x) = (1 + \cos(2\pi n x))\,\mathbf{1}_{[0,1]}(x)$$

(i) Montrer que f_n est une densité de probabilités.

(ii) Pour $n \in \mathbb{N}$ fixé, et pour $x \in \mathbb{R}$, calculer $F_n(x)$.

(iii) Montrer que $F_n(x)$ converge simplement vers une fonction F que l'on déterminera. La fonction F est-elle une fonction de répartition ?

(iv) Soit $x \in \mathbb{R}$ fixé. La suite $(f_n(x))_{n\in\mathbb{N}}$ converge t-elle lorsque $n \to \infty$? La réciproque du théorème de Scheffé est-elle vraie ?