M61B, Probabilités et intégration
Licence 3e année
Parcours mathématiques

Université de Lille — FACULTÉ DES SCIENCES ET TECHNOLOGIES, Département Mécanique

DS2-3h

9 Mai 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.*
- *Attention aux justifications : une bonne réponse mal justifiée ne vaut pas grand chose...*

Exercice 1 : Questions de cours, 4 pts

1) Énoncer le théorème de convergence dominée.
2) Donner la définition de la tribu borélienne sur $\mathbb{R}$ (Il n'est pas nécessaire de redonner la définition d'une tribu).
3) Après avoir donné ses conditions d'applications, énoncer l'inégalité de Markov.
4) Qu'est ce qu'un espace mesuré ?
5) Donner la définition de la loi uniforme sur le segment $[a, b]$, pour tous réels $a < b \in \mathbb{R}$.

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) Si X et Y sont des variables aléatoires discrètes, alors la variable aléatoire $Z = X + Y$ est également discrète.
2) Si une suite de variables aléatoires réelles $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X, alors elle converge aussi dans L^1.
3) La fonction de répartition F_X d'une variable aléatoire X à densité continue par morceaux est une fonction continue.
4) Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions intégrables et bornées sur $\mathbb{R}$, on suppose

que (f_n) converge simplement vers une fonction f. Alors f est intégrable, et

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n(t)dt = \int_{\mathbb{R}} f(t)dt.$$

5) Si la variable aléatoire réelle X est intégrable, alors les variables

$$Y := Xsin(X) \quad \text{et} \quad Z := 3X^3/(1+X^2)$$

sont également intégrables.

Exercice 3 : Variables presque-sûres, 5 pts

1) Soit X une variable aléatoire à valeurs dans $[-1, 1]$.
(i) Justifier que X est de carré intégrable, et donner deux expressions pour sa variance $V(X)$.
(ii) En justifiant que $\mathbb{E}(X^2) \leq 1$, montrer que $V(X) \leq 1$.
2) On suppose maintenant que $V(X) = 1$.
(i) Quelles sont les valeurs possibles de $\mathbb{E}(X^2)$ et de $\mathbb{E}(X)$?
(ii) En déduire que $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = 1/2$.

Exercice 4 : Fonctions génératrices, 10 pts

Soient X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$. On appelle *fonction génératrice* de X la série entière

$$G_X(t) = \sum_{n=0}^{+\infty} \mathbb{P}(X = n)t^n.$$

1) Soit R le rayon de convergence de cette série, montrer que $R \geq 1$.
2) À l'aide du théorème de transfert, exprimer $G_X(t)$ comme l'espérance d'une fonction de X pour $t < R$.
3) (i) Justifier que G_X est de classe C^∞ sur $]-1, 1[$.
(ii) En calculant ses premières dérivées, justifier que sa dérivée k-ème $G_X^{(k)}$ est donnée par

$$G_X^{(k)}(t) = \sum_{n=k}^{+\infty} \mathbb{P}(X = n)\frac{n!}{(n-k)!}t^{n-k}.$$

(iii) Exprimer $G_X^{(k)}(0)$ en fonction de la distribution de X.
(iv) En déduire que si que si $G_X = G_Y$ sur $]-1, 1[$, alors X et Y ont même loi.
4) (i) Calculer G_X lorsque X suit une loi de Bernoulli de paramètre p, puis lorsque X suit une loi binomiale de paramètres (n, p).
(ii) On suppose que X et Y sont indépendantes. Démontrer que pour tout $t \in]-1, 1[$

$$G_{X+Y}(t) = G_X(t)G_Y(t).$$

(iii) Soit $X \sim Bin(n,p)$ et $Y \sim Bin(m,p)$ deux variables aléatoires indépendantes. En utilisant les fonctions génératrices, déterminer la loi de $X+Y$. Retrouver ce résultat sans les fonctions génératrices.

5) On suppose maintenant que $R > 1$.

(i) On suppose que X est intégrable. Calculer la dérivée $G'_X = G_X^{(1)}$, et exprimer $\mathbb{E}(X)$ en fonction de G'_X.

(ii) On suppose maintenant X de carré intégrable. Calculer la dérivée seconde $G''_X = G_X^{(2)}$, et montrer que

$$\mathbb{E}(X^2) = G''_X(1) + G'_X(1).$$

(iii) En déduire en fonction de G_X l'expression de la variance de X.

Problème : Un principe de grandes déviations, 14 pts

Soit $(X_n)_{n\in\mathbb{N}}$ une suite i.i.d. de variables de bernoulli de paramètre $1/2$. On note

$$S_n = \frac{1}{n}\sum_{k=1}^{n} X_k$$

la moyenne empirique de (X_n). On souhaite étudier le comportement, quand $n \to \infty$, de S_n.

1) Justifier qu'il existe une constante C, que l'on précisera, telle que S_n converge presque surement vers C lorsque n tend vers l'infini.

Pour $\varepsilon > 0$, on va maintenant estimer la probabilité $\mathbb{P}(|S_n - C| > \varepsilon)$.

Partie I : Décroissance polynomiale.

2) (i) Justifier que S_n est de carré intégrable.

(ii) Calculer sa variance $V(S_n)$.

(iii) Montrer que pour tout entier ≥ 1, et pour tout $\varepsilon > 0$,

$$\mathbb{P}(|S_n - C| > \varepsilon) \leq \frac{1}{4n\varepsilon^2}.$$

Partie II : Décroissance exponentielle.

La question précédente montre que la probabilité que $|S_n - C|$ soit plus grande que $\varepsilon > 0$ décroit polynomialement, c'est à dire qu'elle est de l'ordre au plus $O(n^{-a})$ (avec ici $a = 1$). On souhaite montrer un résultat plus fort : en réalité, cette probabilité décroit comme $\exp(-nK_\varepsilon)$, pour une certaine constante $K_\varepsilon > 0$.

3) (i) En comparant les lois des variables $X_n - C$ et $C - X_n$, justifier que les variables aléatoires $S_n - C$ et $C - S_n$ ont la même distribution.

(ii) En déduire que

$$\mathbb{P}(|S_n - C| > \varepsilon) \leq 2\mathbb{P}(S_n - C > \varepsilon).$$

(iii) Soit $\varepsilon > 0$ et $t > 0$, montrer que

$$S_n - C \geq \varepsilon \quad \Leftrightarrow \quad \exp\left(t\left[\sum_{k=1}^{n} X_k - nC\right]\right) \geq \exp(tn\varepsilon).$$

(iv) En déduire que

$$\mathbb{P}(S_n - C > \varepsilon) \leq \frac{\mathbb{E}\left(\exp\left(t\sum_{k=1}^{n} X_k - tnC\right)\right)}{\exp(tn\varepsilon)}.$$

(v) À l'aide des questions précédentes, montrer que

$$\mathbb{P}(|S_n - C| > \varepsilon) \leq \frac{2\mathbb{E}(Z)}{\exp(tn(C+\varepsilon))}.$$

où on a défini $Z := \exp\left(t\sum_{k=1}^{n} X_k\right)$.

4) On souhaite maintenant calculer $\mathbb{E}(Z)$.

(i) Soit X une variable de Bernoulli de paramètre $1/2$ et $t > 0$. À l'aide du théorème de transfert, calculer $g(t) := \mathbb{E}(\exp(tX))$.

(ii) À l'aide d'un développement limité, en déduire que pour t petit,

$$g(t) = 1 + \frac{t}{2} + O(t^2).$$

(iii) Montrer que $\mathbb{E}(Z) = g(t)^n$.

(iv) Montrer que pour tout $\varepsilon > 0$, il existe $t_\varepsilon > 0$ tel que pour tout $t \leq t_\varepsilon$,

$$\log(g(t)) \leq \frac{t}{2} + \frac{t\varepsilon}{2}.$$

(v) Déduire de ce qui précède que pour tout $t > t_\varepsilon$

$$\mathbb{E}(Z) \leq \exp\left(\frac{nt}{2} + \frac{nt\varepsilon}{2}\right).$$

5) Déduire de tout ce qui précède le résultats souhaité, c'est à dire qu'il existe une constant $K_\varepsilon > 0$ telle que pour tout $n \in \mathbb{N}$

$$\mathbb{P}(|S_n - C| > \varepsilon) \leq 2\exp(-nK_\varepsilon).$$