M61B, Probabilités et intégration
Licence 3e année
Parcours mathématiques

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 $a < b \in \mathbb{R}$.

Solution : Voir poly. □

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.

Solution :

1) Si X et Y sont des variables discrètes, cela veut dire que leurs espaces d'état $X(\Omega)$ et $Y(\Omega)$ sont (au plus) dénombrables. En particulier,$X(\Omega)\times Y(\Omega) = (X,Y)(\Omega)$ est également dénombrable, donc en particulier
$(X+Y)(\Omega) = \{x+y, (x,y) \in X(\Omega)\times Y(\Omega)\}$ l'est aussi.

2) La convergence en loi ne requiert même pas que les variables aléatoires soient définies sur le même espace de probabilités. En particulier, elle ne peut donc pas impliquer la convergence dans L^1. Par exemple, soit X une variable de Bernoulli de paramètre $1/2$, on pose $X_0 = X$ et pour tout $n > 0$ $X_n = 1 - X_{n-1}$. Tous les X_n ont la même distribution donc en particulier X_n converge en loi vers X, et pourtant $\mathbb{E}(X_n - X) = \mathbf{1}_{\{n \text{ est impair}\}}$, qui ne converge pas vers 0.

3) On peut écrire $F_X(x) = \int_0^x f_X(t)dt$. Or si f_X est continue par morceaux elle est localement bornée, par conséquent F_X est continue.

4) C'est faux, on ne peut pas appliquer le théorème de convergence dominée : on peut prendre le contre exemple du cours, $f_n(t) = \frac{1}{n}\mathbf{1}_{\{0\le t\le n\}}$, qui sont bien bornées d'intégrale 1, et convergent simplement vers $f \equiv 0$ dont l'intégrale est nulle.

5) Si X est intégrable, cela signifie que l'intégrale de $|X|$ est finie. En particulier, comme $|sin(X)| \le 1$ et $|3X^2/(1+X^2)| \le 3$, les deux variables aléatoires Y et Z 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 $\mathbb{E}(X^2) \le 1$, montrer que $V(X) \le 1$.

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 $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = 1/2$.

Solution :

1) (i) on a $|X| \le 1$, et est donc intégrable car la variable aléatoire constant égale à 1 est intégrable, d'intégrale 1 car la mesure $\mathbb{P}$ est une mesure de probabilité. On

peut alors écrire

$$V(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}((X - \mathbb{E}(X))^2).$$

(ii) On a par monotonie de l'espérance $0 \leq \mathbb{E}(X^2) \leq \mathbb{E}(1^2) = 1$, or $|\mathbb{E}(X)| \leq 1$, donc $\mathbb{E}(X)^2 \in [0, 1]$. En utilisant la première expression, on en déduit que

$$\mathbb{E}(X^2) - \mathbb{E}(X)^2 \leq \mathbb{E}(X^2) \leq 1$$

2) (i) Si $V(X) = 1$, on doit avoir que $\mathbb{E}(X^2) = 1$, et $\mathbb{E}(X)^2 = 0$, car sinon on aurait nécessairement $V(X) < 1$. On a donc $\mathbb{E}(X^2) = 1$, et $\mathbb{E}(X) = 0$.

(ii) Pour avoir $\mathbb{E}(X^2) = 1$, on doit avoir $\mathbb{P}(X^2 = 1) = 1$, sinon on aurait $\mathbb{E}(X^2) < 1$. Par conséquent X vaut presque surement 1 ou -1, c'est à dire que $\mathbb{P}(X = 1) + \mathbb{P}(X = -1) = 1$. Comme $\mathbb{E}(X) = 0$, on doit également avoir $\mathbb{P}(X = 1) = \mathbb{P}(X = -1)$, par conséquent $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = 1/2$. □

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.

<u>*Solution*</u> :

1) Pour tout $|t| \leq 1$, la série entière converge absolument, puisque $\sum_{n=0}^{+\infty} \mathbb{P}(X = n) = 1$. Son rayon de convergence en particulier, doit être au moins égal à 1.

2) Par théorème de transfert, on peut écrire que

$$G_X(t) = \mathbb{E}(t^X).$$

3) (i) Une série entière est de classe C^∞ sur son disque ouvert de convergence, donc $G_X(t)$ l'est en particulier sur $]-1,1[\subset]-R,R[$.

(ii) Puisque la série converge normalement sur son disque ouvert de convergence, sa dérivée k-ème est donnée par la k-ème dérivée terme à terme de la série, c'est à dire

$$G_X^{(k)}(t) = \sum_{n=k}^{+\infty} \mathbb{P}(X = n) n(n-1)\dots(n-k+1)t^{n-k},$$

ce qui donne le résultat voulu.

(iii) En $t = 0$, le seul terme qui ne disparait pas est celui pour $n = k$, et on obtient

$$G_X^{(k)}(0) = k!\mathbb{P}(X = k).$$

(iv) En particulier, si G_X=G_Y sur $]-1,1[$, on a pour tout $k \in \mathbb{N}$ que $G_X^{(k)}(0) = G_Y^{(k)}(0)$, et donc $\mathbb{P}(X = k) = \mathbb{P}(Y = k)$ pour tout k.

4) (i) Si X suit une Bernoulli de paramètre p, on a

$$G_X(t) = (1-p) + pt.$$

Si X suit une loi binomiale de paramètres (n, p), on a

$$G_X(t) = \sum_{k=0}^{n} \binom{n}{k} t^k p^k (1-p)^{n-k} = (1-p+pt)^n.$$

(ii) Si X et Y sont indépendantes,

$$\begin{aligned} G_{X+Y}(t) &= \sum_{n=0}^{+\infty} \mathbb{P}(X+Y=n)t^n \\ &= \sum_{n=0}^{+\infty} \sum_{k=0}^{n} \mathbb{P}(X=k, Y=n-k)t^n \\ &= \sum_{n=0}^{+\infty} \sum_{k=0}^{n} \mathbb{P}(X=k)\mathbb{P}(Y=n-k)t^k t^{n-k} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{\infty} \mathbb{P}(X = k)t^k \sum_{n=k}^{\infty} \mathbb{P}(Y = n-k)t^{n-k} \\
&= \sum_{k=0}^{\infty} \mathbb{P}(X = k)t^k G_Y(t) \\
&= G_X(t)G_Y(t)
\end{aligned}
$$

(iii) Par ce qui précède, les fonctions génératrices de X et Y sont données par

$$G_X(t) = (1 - p + pt)^n \quad \text{et} \quad G_Y(t) = (1 - p + pt)^m.$$

Si elles sont indépendantes, la fonction caractéristique de $X + Y$ est alors donnée par

$$G_{X+Y}(t) = G_X(t)G_Y(t) = (1 - p + pt)^{n+m},$$

qui est la fonction génératrice d'une $Bin(n + m, p)$. Comme la fonction génératrice caractérise la loi, $X + Y$ suit une binomiale de paramètres $(n + m, p)$. On retrouve ce résultat en exprimant une loi binomiale comme la somme de n Bernoullis indépendantes.

5) (i) Si X est intégrable, la série $\sum_{n=0}^{\infty} n\mathbb{P}(X = n)$ est absolument convergente, or

$$G'_X(t) = \sum_{n=1}^{+\infty} \mathbb{P}(X = n)nt^{n-1},$$

et donc en particulier $\mathbb{E}(X) = G'_X(1)$.

(ii) De même, Si X est de carré intégrable, la série $\sum_{k=0}^{\infty} n^2\mathbb{P}(X = n)$ est absolument convergente, donc en particulier

$$\mathbb{E}(X^2) = \sum_{k=0}^{\infty} n^2\mathbb{P}(X = n) = \sum_{n=2}^{\infty} n(n-1)\mathbb{P}(X = n) + \sum_{n=0}^{\infty} n\mathbb{P}(X = n) = G''_X(1) + G'_X(1).$$

(iii) On en déduit que

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

□

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 $n \in \mathbb{N}$

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

<u>Solution</u> :

1) Par la loi forte des grands nombres, S_n converge presque surement vers $C := \mathbb{E}(X_1) = 1/2$

2) (i) la variable S_n prend ses valeurs dans $[0, 1]$, et est donc bornée donc de carré intégrable en particulier.

(ii) On écrit, par indépendance des X_k

$$V(S_n) = \frac{1}{n^2}V\left(\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}\sum_{k=1}^{n} V(X_k) = \frac{1}{4n}$$

(iii) L'inégalité est une application directe de l'inégalite de Bienaymé-Tchebychev, puisqu'on a calculé la variance de S_n à la question précédente.

3) (i) $X_n - C$ prend la valeur $1/2$ avec probabilité $1/2$, et vaut $-1/2$ avec probabilité $1/2$. On remarque que $C - X_n$ a exactement la même distribution par conséquent

$$S_n - C = \frac{1}{n}\sum_{k=1}^{n}\{X_k - C\} \quad \text{et} \quad C - S_n = \frac{1}{n}\sum_{k=1}^{n}\{C - X_k\}$$

ont également la même distribution.

(ii) $|S_n - C| > \varepsilon$ implique que $S_n - C > \varepsilon$ ou $C - S_n > \varepsilon$, et donc

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

puisque $S_n - C$ et $C - S_n$ ont la même loi.

(iii) Pour tout $t > 0$, on a

$$S_n - C > \varepsilon \Leftrightarrow tn(S_n - C) > tn\varepsilon \Leftrightarrow \exp(tn(S_n - C)) > \exp(tn\varepsilon).$$

par croissance de la fonction exponentielle. Par définition de S_n, c'est exactement le résultat souhaité.

(iv) Il suffit d'appliquer l'inégalite de Markov à la variable positive $\exp(tn(S_n - C))$ et à la constante $a = \exp(tn\varepsilon)$.

(v) On décompose l'exponentielle, pour obtenir que

$$\exp(tn(S_n - C)) = \exp(-tnC)\exp(tnS_n).$$

Or $\exp(-tnC)$ est une constante, on peut la sortir de l'exponentielle par linéarité de l'espérance, ce qui donne le résultat souhaité grâce à la question ii).
4) (i) Par théorème de transfert
$g(t) = \mathbb{E}(e^{tX}) = e^t\mathbb{P}(X = 1) + e^0\mathbb{P}(X = 0) = (e^t + 1)/2$.
(ii) On fait un développement limité de la fonction $e^t = 1 + t + O(t^2)$, ce qui donne le résultat.
(iii) $Z = \prod_{k=1}^n \exp(tX_k)$ est le produit de n variables indépendantes, par conséquent comme elles sont toutes identiquement distribuées

$$\mathbb{E}(Z) = \prod_{k=1}^{n} \mathbb{E}(\exp(tX_k)) = \mathbb{E}(\exp(tX_1))^n = g(t)^n.$$

(iv) On utilise le développement limité précédent, et on développe à l'ordre 1 $log(1 + u)$ en $u = t/2 + O(t^2)$:

$$\log(g(t)) = \log(1 + t/2 + O(t^2)) = t/2 + O(t^2).$$

En particulier, pour t assez petit, $O(t^2)$ est plus petit que $\varepsilon t/2$, ce qui prouve le résultat.
(v) On écrit que $\mathbb{E}(Z) = \exp(n \log g(t))$ et on utilise la borne précédente pour trouver le résultat.
5) On met ensemble tout ce qu'on a prouvé jusqu'ici :

$$\begin{aligned}\mathbb{P}(|S_n - C| > \varepsilon) \leq \frac{2\mathbb{E}(Z)}{\exp(tn(C + \varepsilon))} &= 2\exp(n\log(g(t)) - tn(C + \varepsilon)) \\ &\leq 2\exp(tn/2 + tn\varepsilon/2 - tn/2 - tn\varepsilon)) \leq 2\exp(-tn\varepsilon/2).\end{aligned}$$

pour tout $t < t_\varepsilon$. Il suffit alors de poser par exemple $K_\varepsilon = t_\varepsilon\varepsilon/2$, qui est positif.

□