Feuille de TD n° 1. Corrigé.

Exercice 1.

1. $\Omega = \{1, \dots, 20\}$, $\mathrm{Card}(\Omega) = 20$, $\mathbb{P}(\omega) = 1/20\ \forall \omega \in \Omega$.
2. Le choix le plus simple : $\Omega = \mathbb{N}^2$. On pourrait aussi y ajouter plus d'informations, tel que $\Omega' = \mathbb{N}^2 \times \{R, P, T\}$, où R : "après temps réglementaire", S : "après temps supplémentaire", T : "après tir au but". Dans tous les cas, on a (nécessairement) $\mathrm{Card}(\Omega) = \infty$. Quant à $\mathbb{P}$, on n'en sait rien *a priori.*
3. On a $\Omega = \mathbb{P}(S)$, et $\mathrm{Card}(\Omega) = 2^n$. Quant à $\mathbb{P}$, on n'en sait rien *a priori.*
4. $\Omega = \{0, 1\}^n$, où 0 : "pile", 1 : "face" (par exemple). $\mathrm{Card}(\Omega) = 2^n$, $\mathbb{P}(\omega) = 2^{-n}\ \forall \omega \in \Omega$.
5. Les deux candidats sont Hollande (H) et Sarkozy (S), mais une personne peut aussi bien répondre "je ne sais pas" (N1), "je ne voterai pas" (N2), "je voterai blanc" (N3), où bien ne pas répondre (N0). Un bon sondage prend toutes ces réponses en compte (voir plus...). Un espace de probabilité possible serait $\Omega = \{H, S, N0, N1, N2, N3\}^{1047}$, avec $\mathrm{Card}(\Omega) = 6^{1047}$. Une autre possibilité serait de compter seulement le nombre de répondants pour chaque réponse possible, cela donne $\Omega' = \{0, 1, \dots, 1047\}^6$ avec $\mathrm{Card}(\Omega') = 1048^6 \ll 6^{1047}$. Cependant, le premier espace est plus commode, car il permet d'intégrer facilement l'indépendance entre les répondants (en supposant que ceux-ci ont été tirés avec remise dans la population totale). En effet, si la probabilité d'une réponse $r \in \{H, S, N0, N1, N2, N3\}$ est donnée par p_r, alors on pose $\mathbb{P}(r_1, \dots, r_{1047}) = p_{r_1} \cdots p_{r_{1047}}$.

Exercice 2.

1. Faux. Si B_r désigne la boule ouverte de rayon r autour de l'origine, alors $\cap_n B_{1/n} = \{0\}$, ce qui n'est pas un ouvert.
2. Vrai. On a
   (a) $\{0, 1\}^{\mathbb{N}} = \{0, 1\}^k \times \{0, 1\}^{\mathbb{N}}$, $\emptyset = \emptyset \times \{0, 1\}^{\mathbb{N}}$
   (b) $\{0, 1\}^{\mathbb{N}} \backslash A \times \{0, 1\}^{\mathbb{N}} = (\{0, 1\}^k \backslash A) \times \{0, 1\}^{\mathbb{N}}$
   (c) Pour $A_1, A_2, \dots \subset \{0, 1\}^k$, $\bigcup_n (A_n \times \{0, 1\}^{\mathbb{N}}) = (\bigcup_n A_n) \times \{0, 1\}^{\mathbb{N}}$.
3. Faux (une union dénombrable croissante de tribus n'est pas toujours une tribu) Exemple : $\bigcap_n \{0\}^n \times \{0, 1\}^{\mathbb{N}} = \{0\}^{\mathbb{N}}$ n'est pas de cette forme.
4. *Clarification : « A mesurable » signifie ici que $A \in \mathcal{B}$ pour une tribu quelconque $\mathcal{B}$ sur $\{0, 1\}^{\mathbb{N}}$, par exemple la tribu cylindrique.*

   Vrai (tribu *terminale*) : Comme dans 2., on montre que pour tout n, $\{\{0, 1\}^n \times A : A \subset \{0, 1\}^{\mathbb{N}}$ mesurable$\}$ est une tribu. Une intersection de tribus est concore une tribu.
5. Faux. Par exemple, si A est une partie de $\mathbb{N}$ qui n'admet pas de densité (par exemple, $k \in A$ ssi $\lfloor \log_2 k \rfloor$ est paire), alors pour tout n, $A \cap \{0, \dots, n\}$ admet une densité (nulle), mais $A = \bigcup_n (A \cap \{0, \dots, n\})$ n'en admet pas.

Exercice 3.

1. Par l'hypothèse sur μ, on a pour $e \in E$, $f \in F$,
$$\mu(\{(e, f)\}) = \mu(\{e\} \times \{f\}) = (\mathrm{Card}(E)\mathrm{Card}(F))^{-1} = \mathrm{Card}(E \times F)^{-1}.$$
   Par additivité de la mesure, on a alors pour tout ensemble $C \subset E \times F$,
$$\mu(C) = \mathrm{Card}(C)/\mathrm{Card}(E \times F).$$
   La mesure μ est alors bien la mesure uniforme sur $E \times F$.
2. Ici, les choses ne sont plus aussi simples. Il convient de prendre recours au *lemme des classes monotones.* Notons $\mathcal{B}$ la tribu borélienne sur $[0, 1]^2$ (on rappelle que c'est la tribu engendrée par les ouverts de $[0, 1]^2$). On définit
$$\mathcal{M} = \{A \in \mathcal{B} : \mu(A) = \mathrm{Leb}(A)\}.$$
   On a alors $\mathcal{M} \supset \mathcal{C} = \{I \times J : I, J \subset [0, 1] \text{ intervalles}\}$. On sait/admet que la classe $\mathcal{C}$ engendre la tribu borélienne $\mathcal{B}$. Montrons que $\mathcal{M}$ est une classe monotone :

— Puisque $[0,1]$ est un intervalle, on a $[0,1]^2 = [0,1] \times [0,1] \in \mathcal{C} \subset \mathcal{M}$.
— Soient $A, B \in \mathcal{M}$, $A \subset B$. On a alors

$$\mu(B \backslash A) = \mu(B) - \mu(A) = \mathrm{Leb}(B) - \mathrm{Leb}(A) = \mathrm{Leb}(B \backslash A),$$

donc $B \backslash A \in \mathcal{M}$.
— Soient $A_1, A_2, \ldots \in \mathcal{M}$, avec $A_n \subset A_{n+1}$ pour tout $n \in \mathbb{N}$. Il est une propriété fondamentale de toute mesure ν que $\nu(\bigcup_n A_n) = \lim_{n\to\infty} \nu(A_n)$ pour une telle suite croissante d'evénements. En particulier,

$$\mu\Big(\bigcup_n A_n\Big) = \lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \mathrm{Leb}(A_n) = \mathrm{Leb}\Big(\bigcup_n A_n\Big).$$

Donc $\bigcup_n A_n \in \mathcal{M}$.

Nous avons alors démontré que $\mathcal{M}$ est une classe monotone qui contient la classe $\mathcal{C}$, génèratrice de la tribu borélienne $\mathcal{B}$. De plus, $\mathcal{C}$ est stable par intersection finie car pour tous intervalles $I, J, K, L \subset [0,1]$, on a

$$(I \times J) \cap (K \times L) = (I \cap J) \times (K \cap L) \in \mathcal{C}.$$

Le lemme des classes monotones nous dit alors que $\mathcal{M} \supset \mathcal{B}$ (du coup, on a en fait $\mathcal{M} = \mathcal{B}$). Par conséquent, $\mu(A) = \mathrm{Leb}(A)$ pour tout $A \in \mathcal{B}$ et donc $\mu = \mathrm{Leb}$.

Exercice 4.

1. Puisque μ est une mesure, on a,
   — $\mu_A(\emptyset) = \mu(\emptyset \cap A) = \mu(\emptyset) = 0$, et
   — pour tous $A_1, A_2, \ldots \in \mathcal{A}$ disjoints,

   $$\mu_A\Big(\bigcup_n A_n\Big) = \mu\Big(\Big(\bigcup_n A_n\Big) \cap A\Big) = \mu\Big(\bigcup_n (A_n \cap A)\Big) = \sum_n \mu(A_n \cap A) = \sum_n \mu_A(A_n).$$

   Par conséquent, μ_A est également une mesure.
2. Par le point précédent, l'application $\mathbb{P}(A)\mathbb{P}(\cdot \,|\, A) = \mathbb{P}(\cdot \cap A)$ est une mesure, donc $\mathbb{P}(\cdot \,|\, A)$ l'est aussi car $\mathbb{P}(A) > 0$. De plus, on a $\mathbb{P}(\Omega \,|\, A) = \mathbb{P}(A)/\mathbb{P}(A) = 1$. Par conséquent, $\mathbb{P}$ est une mesure de probabilité.

Exercice 5.

Première solution : Définissons notre espace de probabilité $\Omega = \{FF, FM, MF, MM\}$, où la première lettre donne le sexe de l'ainé(e) et la deuxième celui du cadet/de la cadette. On suppose que ces quatre possibilités sont équiprobables, ce qui est à notre connaissance assez proche de la réalité. L'événement "au moins une fille" est alors $A = \{FF, FM, MF\}$. Sachant cet événement, la probabilité que le deuxième enfant est une fille est alors 1/3, contre 2/3 pour un garçon. Formellement, on calcule

$$\mathbb{P}(\text{les deux enfants sont des filles} \,|\, A) = \mathbb{P}(FF)/\mathbb{P}(A) = 1/3.$$

Deuxième solution : La dernière solution peut ne pas être satisfaisante, car on pourrait se dire que s'il y a deux filles, la chance est plus grande de trouver des chaussures de filles devant la porte, donc nous avons oublié des corrélations importantes. Pour formaliser cela, on élargit l'espace de probabilité : on pose $\Omega = \{F, M\}^2 \times \{0,1\}^2$, où le i-ième chiffre est 1 si le i-ième enfant a laissé ses chaussures devant la porte, et 0 sinon. L'événement A qu'il y a (exactement) une paire de chaussures de filles devant la porte est alors

$$A = \{FM10, MF01, FF10, FF01\}.$$

En ce qui concerne la probabilité sur l'espace Ω, il y a beaucoup de choix possibles. Nous n'allons pas nous restreindre à un choix précis, mais nous allons simplement supposer que la probabilité de laisser ou non ses chaussures devant la porte ne dépend pas du sexe, en particulier, nous supposons que les *éléments de l'événement A sont équiprobables*. Par conséquent, si $B = \{FF\} \times \{0,1\}^2$ désigne l'événement que les deux enfants sont des filles, on a $\mathbb{P}(B \,|\, A) = 1/2$. En conclusion, la probabilité que le deuxième enfant est une fille vaut 1/2, contrairement à 1/3 dans la première solution.

Exercice 6.

1. On a pour tout $k \in \{0, \dots, 2n\}$,

$$\begin{aligned}\mathbb{P}(X+Y=k) &= \sum_{i=0}^{k} \mathbb{P}(X=i,\, Y=k-i) \\ &= \frac{1}{(n+1)^2} \sum_{i=0}^{k} \mathbf{1}_{(i\le n,\, k-i\le n)} \\ &= \frac{1}{(n+1)^2} \begin{cases} \sum_{i=0}^{k} 1, & k \le n \\ \sum_{i=k-n}^{n} 1 & k \ge n, \end{cases} \\ &= \frac{1}{(n+1)^2} \begin{cases} k+1, & k \le n \\ 2n-k+1 & k \ge n, \end{cases}\end{aligned}$$

si bien que

$$\mathbb{P}(X+Y=k) = \frac{\min(k+1, 2n-k+1)}{(n+1)^2} = \frac{n+1-|k-n|}{(n+1)^2}$$

et $\mathbb{P}(X+Y=k)=0$ pour tout $k \notin \{0,\dots,2n\}$. Pour la loi de $X-Y$, on remarque que $(X, n-Y)$ a même loi que (X,Y), si bien que pour $k \in \{-n, \dots, n\}$,

$$\begin{aligned}\mathbb{P}(X-Y=k) &= \mathbb{P}(X+(n-Y)=n+k) \\ &= \mathbb{P}(X+Y=n+k) \\ &= \frac{\min(n+k+1, n-k+1)}{(n+1)^2} = \frac{n+1-|k|}{(n+1)^2}\end{aligned}$$

et $\mathbb{P}(X+Y=k)=0$ pour tout $k \notin \{-n,\dots,n\}$.

2. Première façon. Par linéarité de l'espérance,

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y] = 2\mathbb{E}[X].$$

Or, l'espérance de X vaut

$$\mathbb{E}[X] = \frac{1}{n+1} \sum_{k=0}^{n} k = \frac{n}{2},$$

si bien que $\mathbb{E}[X+Y]=n$ (*un calcul plus astucieux est le suivant : puisque X est de même loi que $n-X$, on a* $2\mathbb{E}[X] = \mathbb{E}[X] + \mathbb{E}[n-X] = \mathbb{E}[X+n-X] = \mathbb{E}[n] = n$).

Deuxième façon. On passe par la loi de $X+Y$: par la première partie de l'exercice,

$$\mathbb{E}[X+Y] = \sum_{k=0}^{2n} k \frac{\min(k+1, 2n-k+1)}{(n+1)^2}.$$

On peut calculer cela de manière un peu fastidieuse, mais on peut aussi remarquer comme ci-dessus que $X+Y$ est égale en loi à $2n-(X+Y)$, si bien que

$$\mathbb{E}[X+Y] = \frac{1}{2}\mathbb{E}[X+Y+2n-(X+Y)] = \frac{1}{2}\mathbb{E}[2n] = n.$$

Exercice 7 (Perte de mémoire). Soit X une v.a. à valeurs dans $\mathbb{N}$. Soit $n, m \in \mathbb{N}$. Alors,

$$\mathbb{P}(X \ge n+m \,|\, X \ge n) = \frac{\mathbb{P}(X \ge n+m)}{\mathbb{P}(X \ge n)},$$

et donc

$$\mathbb{P}(X \ge n+m \,|\, X \ge n) = \mathbb{P}(X \ge m) \iff \mathbb{P}(X \ge n+m) = \mathbb{P}(X \ge m)\mathbb{P}(X \ge n).$$

Posant $m=1$, on voit que ceci implique que $\mathbb{P}(X \ge n+1) = \mathbb{P}(X \ge 1)\mathbb{P}(X \ge n)$ pour tout $n \in \mathbb{N}$, et donc $\mathbb{P}(X \ge n) = q^n$ pour un certain $q \in [0,1]$. Puisque $\lim_{n\to\infty} \mathbb{P}(X \ge n) = 0$, on a en fait $q \in [0,1)$.

Une condition nécessaire est alors que $\mathbb{P}(X \ge n) = q^n$ pour un certain $q \in [0,1)$. Cette condition est également suffisante, car on a alors $\mathbb{P}(X \ge n+m) = q^{n+m} = q^n q^m = \mathbb{P}(X \ge n)\mathbb{P}(X \ge m)$. Les lois cherchées sont alors exactement les lois géométriques $\mathcal{G}(p)$ pour $p \in (0,1]$.

Exercice 8 (Perte de mémoire 2). Soit X une v.a. à valeurs dans $\mathbb{R}_+$ vérifiant les hypothèses dans l'énoncé. Pour $k \in \mathbb{N}$ posons $X_k = \lfloor 2^k X \rfloor$. On a alors pour tout $n, m \in \mathbb{N}$,

$$\mathbb{P}(X_k \geq n + m \mid X_k \geq n) = \mathbb{P}(X \geq 2^{-k}n + 2^{-k}m \mid X \geq 2^{-k}n) = \mathbb{P}(X \geq 2^{-k}m) = \mathbb{P}(X_k \geq m).$$

Par l'Exercice 1, X_k suit alors une loi géométrique de paramètre $p_k = 1 - q_k \in (0, 1]$. On a alors pour tout $n \in \mathbb{N}$,

$$\mathbb{P}(X \geq 2^{-k}n) = \mathbb{P}(X_k \geq n) = q_k^n,$$

si bien que pour tout $k \in \mathbb{N}$,

$$q_k = \mathbb{P}(X \geq 2^{-k}) = \mathbb{P}(X \geq 2^{-(k+1)} \times 2) = q_{k+1}^2,$$

et donc $q_k = q^{2^{-k}}$ pour tout k, avec $q = q_0 \in [0, 1)$. Par conséquent,

$$\forall k, n \in \mathbb{N} : \mathbb{P}(X \geq 2^{-k}n) = q^{2^{-k}n}.$$

Il vient que la fonction $x \mapsto \mathbb{P}(X \geq x) - q^x$ est continue à gauche en $x > 0$ et s'annule en les nombres dyadiques qui forment un ensemble dense dans $\mathbb{R}_+$. On en déduit que $\mathbb{P}(X \geq x) = q^x$ pour tout $x > 0$. De plus, cette égalité est trivialement vraie pour $x = 0$. Par conséquent, X suit la loi exponentielle de paramètre $\lambda = -\log q \in (0, \infty]$ (avec la loi exponentielle de paramètre ∞ la mesure δ_0). Ceci est alors une condition nécessaire. Cette condition est également suffisante car si $\mathbb{P}(X \geq x) = q^x$ pour un $q \in [0, 1)$, alors $\mathbb{P}(X \geq x + y) = \mathbb{P}(X \geq x)\mathbb{P}(X \geq y)$, et donc la loi de X satisfait à l'hypothèse.

Exercice 9. Puisque f est positive et dérivable et $X \geq 0$, on a $f(X) = \int_0^X f'(x)\,dx = \int_0^\infty f'(x)\mathbf{1}_{x \leq X}\,dx$. De plus, puisque f est croissante, sa dérivée $f'(x) \geq 0$ pour tout $x \geq 0$. Par le théorème de Fubini–Tonelli,

$$\mathbb{E}[f(X)] = \mathbb{E}[\int_0^\infty f'(x)\mathbf{1}_{x \leq X}\,dx] = \int_0^\infty \mathbb{E}[f'(x)\mathbf{1}_{x \leq X}]\,dx = \int_0^\infty f'(x)\mathbb{P}(X \geq x)\,dx.$$

Exercice 10 (Formule du crible).

1. Soient $A_1, \ldots, A_n \in \mathcal{A}$. Alors,

$$\mathbf{1}_{A_1 \cup \ldots \cup A_n} = 1 - \mathbf{1}_{A_1^c \cap \ldots \cap A_n^c} = 1 - \prod_{i=1}^n \mathbf{1}_{A_i^c} = 1 - \prod_{i=1}^n (1 - \mathbf{1}_{A_i}).$$

2. Soient $A_1, \ldots, A_n \in \mathcal{A}$. Pour tout $l, n \in \mathbb{N}$, notons $\mathcal{P}_{l,n} = \{I \subset \{1, \ldots, n\} : \mathrm{Card}(I) = l\}$ et pour tout $I \subset \{1, \ldots, n\}$, $A_I = \bigcap_{i \in I} A_i$. Par la formule de la première partie de l'exercice, on a

$$\begin{aligned}
\mathbb{P}(A_1 \cup \ldots \cup A_n) &= \mathbb{E}[\mathbf{1}_{A_1 \cup \ldots \cup A_n}] \\
&= 1 - \mathbb{E}\Big[\prod_{i=1}^n (1 - \mathbf{1}_{A_i})\Big] \\
&= 1 - \mathbb{E}\Big[\sum_{l=0}^n (-1)^l \sum_{I \in \mathcal{P}_{l,n}} \mathbf{1}_{A_I}\Big] \\
&= \mathbb{E}\Big[\sum_{l=1}^n (-1)^{l-1} \sum_{I \in \mathcal{P}_{l,n}} \mathbf{1}_{A_I}\Big] \\
&= \sum_{l=1}^n (-1)^{l-1} \sum_{I \in \mathcal{P}_{l,n}} \mathbb{E}[\mathbf{1}_{A_I}] \\
&= \sum_{l=1}^n (-1)^{l-1} \sum_{I \in \mathcal{P}_{l,n}} \mathbb{P}(A_I).
\end{aligned}$$

Ceci est la formule du crible.

Exercice 11. On calcule d'abord la probabilité que Pascal aurait gagné le jeu (elle est plus facile à calculer que celle que Fermat l'aurait gagné). Au maximum, 4 lancers de pièce sont nécessaires, donc on se met sur l'éspace de probabilité $\Omega = \{P, F\}^4$, ou le i-ième P/F désigne que Pascal/Fermat gagne le i-ième lancer en question. L'événement A que Pascal gagne est alors $A = \{PPPP, PPPF, PPFP, PFPP, FPPP\}$, et sa probabilité est alors $\mathbb{P}(A) = 5/16 = 31.25\% < 46.67\% = 7/15$. On comprend alors que chacun a proposé ce qui l'arrange le plus.

Néanmoins, on peut argumenter que la proposition de Fermat est la plus juste. La continuation du jeu aurait été équivalente à une expérience de Bernoulli qui donne Pascal/Fermat gagnant avec probabilité 31.25%/68.75%. De distribuer ce pourcentage du montant total correspond à donner à chacun *l'espérance* de son gain. Ceci est communément interprété comme une solution “non biaisée” est donc la “plus juste”. Notons aussi que la proposition de Pascal n'a peu à voir avec les mécanismes du jeu et donc avec ce qui a été conclu au préalable.

Exercice 12. Soit (X, Y) un vecteur aléatoire uniforme sur $[0, 1]^2$. On a alors pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ mesurable bornée,

$$\begin{aligned}
\mathbb{E}[f(X+Y)] &= \iint_{[0,1]^2} f(x+y)\,dxdy \\
&= \int_0^1 dx \int_x^{1+x} f(x+y-x)\,dy && \text{chgt de var. } y \mapsto y - x \\
&= \int_0^1 dx \int_0^2 \mathbf{1}_{(x\le y,\,1+x\ge y)} f(y)\,dy \\
&= \int_0^2 dy\, f(y) \int_0^1 \mathbf{1}_{(x\le y,\,x\ge y-1)}\,dx && \text{thm de Fubini} \\
&= \int_0^2 f(y)\min(y, 2-y)\,dy.
\end{aligned}$$

La loi de $X + Y$ est alors celle de densité $f_{X+Y}(y) = \min(y, 2 - y) = 1 - |1 - y|$ sur $[0, 2]$. Puisque $X - Y$ est égale en loi à $X + Y - 1$, cette loi est alors celle de densité $f_{X-Y}(y) = f_{X+Y}(1 + y) = 1 - |y|$ sur $[-1, 1]$.

Exercice 13 (Loi gamma).

1. On a
$$\int_0^\infty x^{\alpha-1}e^{-\beta x}\,dx \overset{\beta x = y}{=} \beta^{-\alpha}\int_0^\infty y^{\alpha-1}e^{-y}\,dy = \frac{\Gamma(\alpha)}{\beta^\alpha}.$$
On a alors $c_{\alpha,\beta} = \frac{\beta^\alpha}{\Gamma(\alpha)}$.

2. La loi exponentielle de paramètre λ est la loi sur $[0, \infty)$ de densité $\lambda e^{-\lambda x}$. C'est donc la loi $\Gamma(1, \lambda)$.

3. Soit $N \sim \mathcal{N}(0, 1)$. On a pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ mesurable bornée,
$$\mathbb{E}[f(N^2)] \propto \int_{\mathbb{R}} f(x^2)e^{-x^2/2}\,dx \overset{x^2=y}{\propto} \int_0^\infty f(y)e^{-y/2}\frac{1}{\sqrt{y}}\,dy,$$
donc la loi de N^2 est celle de densité proportionnelle à $y^{-1/2}e^{-y/2}$ sur $[0, \infty)$. Il s'agit donc de la loi $\Gamma(1/2, 1/2)$.

4. Soit $G \sim \Gamma(\alpha, \beta)$ et $b > 0$. On a alors pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ mesurable bornée,
$$\mathbb{E}[f(bG)] \propto \int_{\mathbb{R}} f(bx)x^{\alpha-1}e^{-\beta x}\,dx \overset{bx=y}{\propto} \int_{\mathbb{R}} f(y)y^{\alpha-1}e^{-(\beta/b)y}\,dy.$$
Ceci montre que $bG \sim \Gamma(\alpha, \beta/b)$.

5. La loi de $G_1 + G_2$ est la loi sur $\mathbb{R}$ de densité la convolué de celles de G_1 et G_2, c'est à dire

$$\begin{aligned}
f_{G_1+G_2}(x) &= \int_{\mathbb{R}} f_{G_1}(y) f_{G_2}(x-y)\, dy \\
&\propto \int_{\mathbb{R}} \mathbf{1}_{y\geq 0} y^{\alpha_1-1} e^{-\beta y} \mathbf{1}_{x-y\geq 0} (x-y)^{\alpha_2-1} e^{-\beta(x-y)}\, dy \\
&= e^{-\beta x} \int_0^x y^{\alpha_1-1}(x-y)^{\alpha_2-1}\, dy \\
&= e^{-\beta x} x^{\alpha_1+\alpha_2-1} \int_0^1 y^{\alpha_1-1}(1-y)^{\alpha_2-1}\, dy \qquad\qquad y \mapsto xy, \\
&\propto e^{-\beta x} x^{\alpha_1+\alpha_2-1}.
\end{aligned}$$

Ceci montre que $G_1 + G_2 \sim \Gamma(\alpha_1+\alpha_2, \beta)$.

6. Par la partie précédente de l'exercice, il vient par récurrence que $\sum_{i=1}^n G_i \sim \Gamma(\sum_{i=1}^n \alpha_i, \beta)$.

7. Soit $G \sim \Gamma(\alpha, \beta)$, $\alpha, \beta > 0$. D'après la partie 4, on a $\beta G \sim \Gamma(\alpha, 1)$, si bien que pour tout $n \in \mathbb{N}$,

$$\mathbb{E}[(\beta G)^n] = \frac{1}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} x^n e^{-x}\, dx = \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)} = \alpha^{(n)},$$

où $\alpha^{(n)} = \alpha \cdots (\alpha+n-1)$ (la fonction de *Pochhammer*). Ceci donne $\mathbb{E}[G^n] = \beta^{-n}\alpha^{(n)}$.
Si $N \sim \mathcal{N}(0,1)$, alors $N^2 \sim \Gamma(1/2, 1/2)$ d'après la partie 3. On a alors pour tout $n \in \mathbb{N}$,

$$\mathbb{E}[N^{2n}] = \mathbb{E}[(N^2)^n] = 2^n (1/2)^{(n)} = 1 \cdots (2n-3)(2n-1).$$

De plus, puisque N et $-N$ ont même loi, on a pour tout n impaire, $\mathbb{E}[N^n] = \mathbb{E}[(-N)^n] = -\mathbb{E}[N^n]$, si bien que $\mathbb{E}[N^n] = 0$.

8. Soit $G \sim \Gamma(\alpha, \beta)$, $\alpha, \beta > 0$. On a alors pour tout $\lambda \in \mathbb{R}$,

$$\varphi_G(\lambda) = \mathbb{E}[e^{i\lambda G}] = c_{\alpha,\beta} \int_0^\infty e^{i\lambda x} x^{\alpha-1} e^{-\beta x}\, dx.$$

Comme pour la loi normale, on dérive en λ et on intègre par parties :

$$\begin{aligned}
\varphi_G'(\lambda) &= c_{\alpha,\beta} i \int_0^\infty e^{i\lambda x} x^{\alpha} e^{-\beta x}\, dx \\
&= c_{\alpha,\beta} i \frac{\alpha}{\beta - i\lambda} \int_0^\infty e^{i\lambda x} x^{\alpha-1} e^{-\beta x}\, dx \\
&= -\frac{\alpha}{\lambda + \beta i} \varphi_G(\lambda).
\end{aligned}$$

Ceci donne pour une constante $C \in \mathbb{C}$,

$$\varphi_G(\lambda) = C(\lambda + \beta i)^{-\alpha},$$

et puisque $\varphi_G(0) = 1$,

$$\varphi_G(\lambda) = \left(\frac{\beta i}{\lambda + \beta i}\right)^\alpha = \left(\frac{\beta}{\beta - i\lambda}\right)^\alpha$$

Montrons les propriétés établies dans les parties 4, 6 et 7 :
— Si $b > 0$, on a $\varphi_{bG}(\lambda) = \varphi_G(b\lambda) = \left(\frac{\beta}{\beta - ib\lambda}\right)^\alpha = \left(\frac{\beta/b}{\beta/b - i\lambda}\right)^\alpha$, et donc $bG \sim \Gamma(\alpha, \beta/b)$.
— Si $G_1, \ldots, G_n$ sont des v.a. indépendantes de lois respectives $\Gamma(\alpha_i, \beta)$, avec $\alpha_1, \ldots, \alpha_n, \beta > 0$, alors

$$\varphi_{G_1+\cdots+G_n}(\lambda) = \varphi_{G_1}(\lambda) \cdots \varphi_{G_n}(\lambda) = \prod_{i=1}^n \left(\frac{\beta}{\beta - i\lambda}\right)^{\alpha_i} = \left(\frac{\beta}{\beta - i\lambda}\right)^{\alpha_1+\cdots+\alpha_n}.$$

Donc, $G_1 + \cdots + G_n \sim \Gamma(\alpha_1 + \cdots + \alpha_n, \beta)$.
— On a pour tout $n \in \mathbb{N}$,

$$\mathbb{E}[G^n] = i^{-n} \frac{d^n}{d\lambda^n} \varphi_G(\lambda)\Big|_{\lambda=0} = \frac{\beta^\alpha \alpha(\alpha+1)\cdots(\alpha+n-1)}{(\beta - i\lambda)^{\alpha+n}}\Big|_{\lambda=0} = \beta^{-n}\alpha^{(n)},$$

avec $\alpha^{(n)} = \alpha(\alpha+1)\cdots(\alpha+n-1)$.

Exercice 14 (Méthode de premier et second moment).

1. Puisque X prend ses valeurs dans $\mathbb{N}$, on a par l'inégalité de Markov,

$$\mathbb{P}(X \neq 0) = \mathbb{P}(X \geq 1) \leq \frac{\mathbb{E}X}{1} = \mathbb{E}X.$$

2. On note que $X = 0$ implique $|X - \mathbb{E}X| \geq \mathbb{E}X$. Par l'inégalité de Chebychev, on a alors

$$\mathbb{P}(X = 0) \leq \mathbb{P}(|X - \mathbb{E}X| \geq \mathbb{E}X) \leq \frac{\mathrm{Var}(X)}{(\mathbb{E}X)^2},$$

si bien que

$$\mathbb{P}(X \neq 0) = 1 - \mathbb{P}(X = 0) \geq 1 - \frac{\mathrm{Var}(X)}{(\mathbb{E}X)^2}.$$

3. Par l'inégalité de Cauchy–Schwarz, on a

$$\mathbb{E}X = \mathbb{E}[X\mathbf{1}_{X\neq 0}] \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[(\mathbf{1}_{X\neq 0})^2]} = \sqrt{\mathbb{E}[X^2]\mathbb{P}(X \neq 0)}.$$

En élevant les deux cotés de l'équation au carré, puis divisant par $\mathbb{E}[X^2]$, on obtient l'inégalité souhaitée :

$$\mathbb{P}(X \neq 0) \geq \frac{(\mathbb{E}X)^2}{\mathbb{E}[X^2]}.$$

4. On observe les équivalences suivantes :

$$\begin{aligned}
& 1 - \frac{\mathrm{Var}(X)}{(\mathbb{E}X)^2} \leq \frac{(\mathbb{E}X)^2}{\mathbb{E}[X^2]} \\
\iff & -\frac{\mathrm{Var}(X)}{(\mathbb{E}X)^2} \leq \frac{(\mathbb{E}X)^2 - \mathbb{E}[X^2]}{\mathbb{E}[X^2]} \\
\iff & \frac{\mathrm{Var}(X)}{(\mathbb{E}X)^2} \geq \frac{\mathrm{Var}(X)}{\mathbb{E}[X^2]}.
\end{aligned}$$

Or, puisque $\mathbb{E}[X^2] = (\mathbb{E}X)^2 + \mathrm{Var}(X) \geq (\mathbb{E}X)^2$, la dernière inégalité est toujours vérifiée. Ceci montre que $1 - \frac{\mathrm{Var}(X)}{(\mathbb{E}X)^2} \leq \frac{(\mathbb{E}X)^2}{\mathbb{E}[X^2]}$, ce qu'il fallait démontrer.

Exercice 15 (Méthode de premier et second moment : application).

1. Si $I_{k,n} \geq 1$, alors il existe un $i \in \{0, \ldots, n-k\}$, tel que $X_{i+1} = \cdots = X_{i+k} = 1$, autrement dit, la suite $X_1, \ldots, X_n$ contient k uns consécutifs. Réciproquement, si la suite $X_1, \ldots, X_n$ contient k uns consécutifs, alors soit i le plus petit indice dans $\{0, \ldots, n-k\}$ tel que $X_{i+1} = \cdots = X_{i+k}$. Si $i > 0$, on a alors $X_i = 0$, sinon $i - 1$ serait également un tel indice, ce qui serait en contradiction avec le fait que i en est le plus petit. De plus, si $i = 0$, alors $X_i = 0$ par définition. On a alors $I_{k,n} \geq 1$.

2. Par linéarité de l'espérance, on a

$$\mathbb{E}[I_{k,n}] = \sum_{i=0}^{n-k} \mathbb{E}[\mathbf{1}_{B_i}] = \sum_{i=0}^{n-k} \mathbb{P}(B_i).$$

Notons que

$$\begin{aligned}
\mathbb{P}(B_0) &= \mathbb{P}(X_1 = \cdots = X_k) = 2^{-k}, \\
\forall i \in \{1, \ldots, n-k\} : \mathbb{P}(B_i) &= \mathbb{P}(X_i = 0,\ X_{i+1} = \cdots = X_{i+k} = 1) = 2^{-(k+1)}.
\end{aligned}$$

Il vient que

$$\begin{aligned}
\mathbb{E}[I_{k,n}] &= 2^{-k} + (n-k)2^{-(k+1)} \\
&= (n-k+2)2^{-k-1}.
\end{aligned}$$

3.

$$\begin{aligned}\mathbb{E}[I_{k,n}^2] &= \mathbb{E}\left[\sum_{0\le i,j\le n-k} \mathbf{1}_{B_i}\mathbf{1}_{B_j}\right]\\ &= \sum_{0\le i,j\le n-k} \mathbb{P}(B_i\cap B_j)\\ &= \sum_{0\le i\le n-k}\mathbb{P}(B_i) + \sum_{0\le i,j\le n-k,\ i\ne j}\mathbb{P}(B_i\cap B_j).\end{aligned}$$

Pour calculer $\mathbb{P}(B_i\cap B_j)$ pour $i,j\in\{0,\dots,n-k\}$, $i\neq j$, on distingue entre deux cas : si $1\le |i-j|\le k$, alors $B_i\cap B_j=\emptyset$, tandis que si $|i-j|>k$, alors B_i et B_j sont indépendants. Dans tous les cas, il vient que

$$\mathbb{P}(B_i\cap B_j)\le \mathbb{P}(B_i)\mathbb{P}(B_j).$$

Il en suit la borne suivante sur $\mathbb{E}[I_{k,n}^2]$:

$$\begin{aligned}\mathbb{E}[I_{k,n}^2] &\le \sum_{0\le i\le n-k}\mathbb{P}(B_i) + \sum_{0\le i,j\le n-k,\ i\ne j}\mathbb{P}(B_i)\mathbb{P}(B_j)\\ &\le \sum_{0\le i\le n-k}\mathbb{P}(B_i) + \left(\sum_{0\le i\le n-k}\mathbb{P}(B_i)\right)^2\\ &= \mathbb{E}[I_{k,n}] + (\mathbb{E}[I_{k,n}])^2.\end{aligned}$$

4. Par l'Exercice 14 de cette feuille et l'égalité des événemnents $A_{k,n}=\{I_{k,n}\neq 0\}=\{I_{k,n}\ge 1\}$, on a

$$\frac{(\mathbb{E}I_{k,n})^2}{\mathbb{E}[I_{k,n}^2]}\le \mathbb{P}(A_{k,n})\le \mathbb{E}I_{k,n}.$$

La borne supérieure découle alors directement de l'expression de $\mathbb{E}I_{k,n}$ obtenue dans la partie 2. De plus, par la partie 3, on a

$$\frac{(\mathbb{E}I_{k,n})^2}{\mathbb{E}[I_{k,n}^2]}\ge \frac{(\mathbb{E}I_{k,n})^2}{\mathbb{E}[I_{k,n}]+(\mathbb{E}[I_{k,n}])^2}=\frac{\mathbb{E}I_{k,n}}{1+\mathbb{E}[I_{k,n}]},$$

si bien que la borne inférieure découle encore de l'expression de $\mathbb{E}I_{k,n}$ de la partie 2.

5.

k	1	2	3	4	5	6	7	8	9	10
borne sup	25.2500	12.5000	6.1875	3.0625	1.5156	0.7500	0.3711	0.1836	0.0908	0.0449
borne inf	0.9619	0.9259	0.8609	0.7538	0.6025	0.4286	0.2707	0.1551	0.0833	0.0430

Exercice 16 (Borne de Chernoff).

1. On a pour tout $\lambda\ge 0$, $\{X\ge x\}=\{e^{\lambda X}\ge e^{\lambda x}\}$, si bien que par l'inégalité de Markov,

$$\mathbb{P}(X\ge x)=\mathbb{P}(e^{\lambda X}\ge e^{\lambda x})\le \mathbb{E}[e^{\lambda X}]e^{-\lambda x}=e^{-(\lambda x-\varphi(\lambda))}.$$

En minimisant sur λ, on a

$$\mathbb{P}(X\ge x)\le \inf_{\lambda\ge 0} e^{-(\lambda x-\varphi(\lambda))}=\exp(-\sup_{\lambda\ge 0}[\lambda x-\varphi(\lambda)])=e^{-I(x)}.$$

2. Soit $\varphi_n(\lambda)=\log\mathbb{E}[e^{\lambda(X_1+\ldots+X_n)}]$ et $I_n(x)=\sup_{\lambda\ge 0}[\lambda x-\varphi_n(\lambda)]$. Par indépendance,

$$\varphi_n(\lambda)=\log\mathbb{E}[e^{\lambda(X_1+\ldots+X_n)}]=\log\mathbb{E}[e^{\lambda X}]^n=n\varphi(\lambda).$$

Par conséquent, on a pour $a\in\mathbb{R}$,

$$I_n(an)=\sup_{\lambda\ge 0}[\lambda an-n\varphi(\lambda)]=nI(a).$$

L'inégalité découle alors de la borne de Chernoff de la deuxième partie.