

1. Soit  $p_1, \dots, p_n \in (0, 1)$  et soit  $X_1, \dots, X_n$  des variables aléatoires indépendantes avec  $X_i \sim \text{géom}(p_i)$ .
  - a) Déterminer  $\mathbb{P}(X_i \geq k)$  pour tout  $k \in \mathbb{N}^*$ .
  - b) Soit  $M := \min\{X_1, \dots, X_n\}$ . Déterminer  $\mathbb{P}(M \geq k)$  pour tout  $k \in \mathbb{N}^*$ .
  - c) Dédire du point précédent que  $M \sim \text{géom}(\rho)$  et exprimer  $\rho$  en fonction de  $p_1, \dots, p_n$ .
  - d) En utilisant le fait que  $X_i$  a la même loi que le temps de premier succès d'une suite d'épreuves de Bernoulli de paramètre  $p_i$ , réinterpréter  $M$  comme le temps de premier succès d'une suite d'épreuves de Bernoulli appropriées et en déduire une preuve alternative du résultat du point précédent.

---
2. Soit  $X$  et  $Y$  deux variables aléatoires indépendantes suivant chacune la loi géométrique de paramètre  $p \in (0, 1)$ . Soit  $U := \max\{X, Y\}$  et  $V := \min\{X, Y\}$ .
  - a) Pour  $u, v \in \mathbb{N}^*$ , calculer  $\mathbb{P}(U \leq u, V \geq v)$ . En déduire les lois de  $U$  et de  $V$ .
  - b) Déterminer la loi de  $D := |X - Y|$  et montrer que  $D$  et  $V$  sont indépendantes.

---
3. (Urne de Pólya) On considère une urne contenant initialement  $b$  boules bleues et  $r$  boules rouges. On tire une boule au hasard uniformément et on la replace dans l'urne, ainsi que  $d$  boules supplémentaires de la même couleur que cette dernière. On répète ensuite cette procédure autant de fois que nécessaire. On note  $X_k = 1$  si la  $k$ -ième boule tirée est bleue et  $X_k = 0$  sinon. Soit  $S_n := X_1 + \dots + X_n$  le nombre total de boules bleues extraites lors des  $n$  premiers tirages.
  - a) Calculer la loi de  $X_1$  et la loi de  $X_2$ .
  - b) Calculer  $\mathbb{P}(X_{n+1} = 1 \mid S_n = k)$ .
  - c) En déduire une expression pour  $\mathbb{P}(X_{n+1} = 1)$  en termes de  $\mathbb{E}[S_n]$ .
  - d) Exprimer  $\mathbb{E}[S_n]$  en termes de  $\mathbb{E}[S_{n-1}]$  et  $\mathbb{P}(X_n = 1)$ .
  - e) En déduire que  $\mathbb{P}(X_{n+1} = 1) = \mathbb{P}(X_n = 1)$ .
  - f) Quelle est la loi de  $X_n$  ?

---
4.
  - a) Soit  $X \sim \text{poisson}(\lambda)$  et  $Y \sim \text{poisson}(\mu)$  deux variables aléatoires indépendantes. Montrer que  $X + Y \sim \text{poisson}(\lambda + \mu)$ .
  - b) Soit  $X_1, \dots, X_n$  des variables aléatoires indépendantes suivant chacune une loi de Poisson de paramètre  $\lambda$ . Montrer que  $X_1 + \dots + X_n \sim \text{poisson}(n\lambda)$ .
  - c) Notons  $X_k$ ,  $1 \leq k \leq n$  le nombre de fautes de frappe sur la  $k$ -ième page d'un texte de  $n$  pages. On supposera que les variables aléatoires  $X_k$  sont indépendantes et suivent chacune une loi de Poisson de paramètre  $\lambda$ .
    - i. Calculer la probabilité que la première faute se trouve sur la  $k$ -ième page.
    - ii. Pour  $1 \leq r \leq n$ , notons  $S_r := X_1 + \dots + X_r$  le nombre total de fautes dans les  $r$  premières pages. Soit  $\ell \geq 0$  ; calculer  $\mathbb{P}(S_r \geq \ell)$ .
    - iii. Soit  $\ell \geq 1$ . Calculer la probabilité que la  $\ell$ -ième faute se trouve sur la  $k$ -ième page.

---

---

5. Déterminer la fonction de répartition et la fonction quantile pour les lois suivantes :

1. La loi Béta  $B(\alpha, 1)$ , avec  $\alpha > 0$ , dont la densité est

$$f(x) = \alpha x^{\alpha-1} \mathbf{1}_{[0,1]}(x).$$

2. La loi de Weibull de paramètres  $\alpha > 0$  et  $\lambda > 0$ , dont la densité est

$$f(x) = \alpha \lambda x^{\alpha-1} e^{-\lambda x^\alpha} \mathbf{1}_{\mathbb{R}^+}(x).$$

---

6. Montrer que si  $X$  est une variable aléatoire réelle, les médianes de  $X$  minimisent  $a \mapsto \mathbb{E}|X - a|$ . Commencer par le montrer lorsque  $X$  est une variable aléatoire discrète, uniforme sur  $\{x_1, \dots, x_r\}$  avec  $x_1 < \dots < x_r$ .

---

7. Soit  $X$  une variable aléatoire uniforme sur l'intervalle  $[-1, 3]$ .

Calculer la loi de la variable aléatoire  $Y = \frac{1}{2}(|X - 1| + |X|)$ . Calculer la moyenne et la variance de  $Y$ .

---

8. Soit  $X$  une variable aléatoire de loi uniforme sur  $[1, 5]$ . Déterminer la fonction de répartition et la densité de probabilité de la variable aléatoire  $Y := X/(5 - X)$ .

---

9. Soit  $X$  une variable aléatoire réelle de fonction de répartition  $F_X$ . On définit, pour tout  $u \in ]0, 1[$  l'inverse généralisée continue à gauche de  $F_X$  :

$$F_X^{-1}(u) = \inf\{t : F_X(t) \geq u\}$$

a) Montrer que, pour tout  $t \in \mathbb{R}$ ,  $u \in ]0, 1[$ ,  $(F_X^{-1}(u) \leq t) \Leftrightarrow (u \leq F_X(t))$

b) En déduire que si  $U$  suit la loi uniforme sur  $[0, 1]$  alors  $\tilde{X} = F_X^{-1}(U)$  suit la même loi que  $X$ . Remarquer que cela permet de simuler des lois non uniformes.

c) Expliciter  $F_X^{-1}$  dans les cas d'une loi exponentielle et d'une loi de Cauchy.

d) A l'aide de ce qui précède, montrer que toute fonction croissante  $F$  continue à droite telle que  $\lim_{x \rightarrow -\infty} F(x) = 0$  et  $\lim_{x \rightarrow +\infty} F(x) = 1$  est la fonction de répartition d'une loi de probabilité sur  $\mathbb{R}$ .

---

10. Soit  $X_1, \dots, X_n$  des variables aléatoires réelles indépendantes de même loi de fonction de répartition  $F$ . Soit  $X_{(1)}, \dots, X_{(n)}$  les nombres  $X_1, \dots, X_n$  ordonnés de façon croissante. La variable aléatoire  $X_{(k)}$  est appelée statistique d'ordre de rang  $k$  et fait partie des outils fondamentaux en statistiques.

a) Trouver une expression (sous forme d'une somme appropriée) pour la fonction de répartition de  $X_{(k)}$  pour  $k \in \{1, \dots, n\}$ .

b) Déterminer explicitement les fonctions de répartition de  $X_{(1)}$  et  $X_{(n)}$ .

---

11. Albert collectionne des vignettes, vendues par pochette. Il existe  $n$  vignettes différentes, et chaque pochette contient 5 vignettes distinctes, tirées uniformément parmi toutes les combinaisons possibles. Albert achète une pochette à la fois jusqu'à ce qu'il ait complété sa collection. On note  $N$  le nombre de pochettes qu'il lui a fallu acheter. On numérote les  $N$  pochettes selon leur ordre d'achat.

- a) On note  $T_k$  le numéro de la première pochette contenant une copie de la  $k$ -ième vignette. Montrer que  $N = \max\{T_1, \dots, T_n\}$ .  
 b) On note  $A_k := \{T_k > \ell\}$ . Montrer que

$$\mathbb{P}(N > \ell) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \mathbb{P}(\min\{T_{i_1}, \dots, T_{i_k}\} > \ell).$$

(Indication : appliquer le principe d'inclusion-exclusion à  $\mathbb{P}(A_1 \cup \dots \cup A_n)$ .)

- c) Calculer  $\mathbb{P}(\min(T_{i_1}, \dots, T_{i_k}) > \ell)$  pour  $1 \leq k \leq n$  et  $1 \leq i_1 < i_2 < \dots < i_k \leq n$ .  
 d) En déduire que le nombre moyen de pochettes qu'Albert va devoir acheter pour compléter sa collection est égal à

$$\mathbb{E}[N] = \sum_{k=1}^n (-1)^{k-1} \frac{\binom{n}{k} \binom{n}{5}}{\binom{n}{5} - \binom{n-k}{5}}.$$

- e) Supposons que  $n = 660$ . Calculer explicitement le nombre moyen de pochettes à acheter.

12. Une marque de céréales place une figurine à collectionner dans chacun de ses paquets. Une collection complète comporte  $n$  figurines distinctes. Albert souhaite obtenir la collection complète. Il va donc acheter un paquet à la fois, jusqu'à ce qu'il ait obtenu au moins un exemplaire de chacune des  $n$  figurines, puis il s'arrêtera. Soit  $T_n$  le nombre de paquets qu'il devra acheter. Le but de cet exercice est de démontrer qu'Albert devra acheter approximativement  $n \log n$  paquets (lorsque  $n$  est grand) :

$$\forall \epsilon > 0, \quad \lim_{n \rightarrow \infty} \mathbb{P}\left(\left|\frac{T_n}{n \log n} - 1\right| \geq \epsilon\right) = 0. \quad (\star)$$

On considérera qu'à chaque achat, la probabilité d'obtenir chacune des  $n$  figurines est égale à  $1/n$ .

- a) Soit  $\tau_n^k$  le nombre minimum de paquets qu'Albert doit acheter afin d'avoir en sa possession  $k$  ( $k \leq n$ ) figurines distinctes. Évidemment,  $\tau_n^0 = 0$  et  $\tau_n^n = T_n$ . Exprimer  $T_n$  en termes des variables aléatoires  $\xi_k := \tau_n^k - \tau_n^{k-1}$  (le nombre de paquets à acheter pour faire passer sa collection de  $k-1$  figurines à  $k$  figurines).  
 b) Expliquer pourquoi les variables aléatoires  $\xi_1, \xi_2, \dots, \xi_n$  sont indépendantes.  
 c) Quelle est la loi de  $\xi_k$  ? Déterminer son espérance et sa variance.  
 d) En déduire que

$$\mathbb{E}[T_n] = n \sum_{k=1}^n k^{-1}, \quad \text{Var}[T_n] \leq n^2 \sum_{k=1}^n k^{-2},$$

et donc, en particulier, que  $\lim_{n \rightarrow \infty} \mathbb{E}[T_n]/(n \log n) = 1$ .

- e) Soient  $(S_k)_{k \geq 1}$  des variables aléatoires telles que  $\mathbb{E}[S_n] = \mu_n$  et  $\text{Var}[S_n] = \sigma_n^2$ . Soit  $(b_n)_{n \geq 1}$  une suite de nombres réels tels que  $\lim_{n \rightarrow \infty} \sigma_n^2/b_n^2 = 0$ . Démontrer, à l'aide de l'inégalité de Bienaymé-Tchebychev, que

$$\forall \epsilon > 0, \quad \lim_{n \rightarrow \infty} \mathbb{P}(|S_n - \mu_n| > \epsilon b_n) = 0.$$

- f) Vérifier qu'on peut appliquer le point précédent avec  $S_n = T_n$  et  $b_n = n \log n$ . En déduire  $(\star)$ .

Tournez la page.

- 
13. Soit  $r : \mathbb{N} \rightarrow \mathbb{R}$  telle que  $\lim_{\ell \rightarrow \infty} r(\ell) = 0$ . Soit  $(X_n)_{n \geq 1}$  des variables aléatoires réelles satisfaisant, pour tout  $n \geq 1$ ,  $\mathbb{E}[X_n] = \mu$  et  $\text{Cov}[X_n, X_{n+\ell}] \leq r(\ell)$  pour tout  $\ell \geq 0$ . Soit  $\bar{X}_n$  la moyenne empirique de  $X_1, \dots, X_n$ . Montrer que

$$\forall \epsilon > 0, \quad \lim_{n \rightarrow \infty} \mathbb{P}(|\bar{X}_n - \mu| \geq \epsilon) = 0.$$

*(Remarque : Cet exercice montre que la loi faible des grands nombres s'étend à certaines suites de variables aléatoires corrélées.)*

---

14. Soit  $(X_n)_n$  une suite de variables aléatoires telle que chaque  $X_n$  suit une loi uniforme sur l'ensemble  $\{\frac{1}{n}, \frac{2}{n}, \dots, 1\}$ . Montrer que cette suite converge en loi vers une variable aléatoire  $X$  de loi uniforme sur l'intervalle  $[0, 1]$ .
- 

15. Soit  $\theta > 0$  et soit  $(X_n)_n$  une suite de variables aléatoires telle que chaque  $X_n$  suit une loi géométrique de paramètre  $p_n = \frac{\theta}{n}$ . Montrer que la suite  $T_n/n$  converge en loi et déterminer sa limite.
- 

16. Soit  $(X_i)_{i \in \mathbb{N}}$  une suite de variables aléatoires de Bernoulli de paramètre  $p_i$ , indépendantes.
1. Si  $p_i \rightarrow p$ , montrer que la suite  $X_i$  converge en loi vers une variable aléatoire de Bernoulli de paramètre  $p$ .
  2. Si  $\sum_i p_i < \infty$ , montrer que  $X_i$  converge presque sûrement vers 0.
  3. Si  $p_i \rightarrow 0$  et  $\sum_i p_i = \infty$ , montrer que  $X_i$  diverge p.s. mais converge en loi vers une variable constante égale à 0.