

Plaquette 5 - Problèmes de synthèse en probabilités

Dénombrement et probabilités — Licence 1 (L1)

Corrections détaillées

Boîte à outils

Analyse au premier pas : si e_s est l'espérance d'un temps d'attente partant de l'état s , alors $e_s = 1 + \sum_{s'} \mathbb{P}(s \rightarrow s') e_{s'}$.

— **Série géométrique** : $\sum_{k \geq 0} x^k = \frac{1}{1-x}$ ($|x| < 1$) ; **crosse de hockey** :

$$\sum_{k=r}^n \binom{k}{r} = \binom{n+1}{r+1}.$$

— **Approximation de Poisson** : pour n grand et p petit, $B(n, p) \approx P(np)$:

$$\binom{n}{k} p^k (1-p)^{n-k} \approx e^{-\lambda} \frac{\lambda^k}{k!}, \quad \lambda = np$$

— **Linéarité** de l'espérance (indicatrices), **Markov**, **Bienaymé-Tchebychev**.

— **Loi stationnaire** d'une chaîne : $\pi = \pi P$ (la loi ne change pas en un pas).

Correction de l'exercice 1 - Marche aléatoire sur une droite

a) $\mathbb{E}(X_i) = 0$, $V(X_i) = \mathbb{E}(X_i^2) = 1$. Linéarité et indépendance : $\mathbb{E}(S_n) = 0$, $V(S_n) = n$ (l'écart type $\sqrt{n}$ croît moins vite que n).

b) Si B_n pas valent $+1$ et $n - B_n$ valent -1 : $S_n = B_n - (n - B_n) = 2B_n - n$.
 $S_{2n} = 0 \Leftrightarrow B_{2n} = n$:

$$\mathbb{P}(S_{2n} = 0) = \binom{2n}{n} \frac{1}{4^n}$$

c) $\mathbb{P}(S_4 = 0) = \frac{6}{16} = \frac{3}{8}$; $\mathbb{P}(S_6 = 0) = \frac{20}{64} = \frac{5}{16}$ (et $\mathbb{P}(S_n = 0) = 0$ pour n impair : parité).

d) $\mathbb{P}(|S_{100}| \geq 30) \leq \frac{100}{900} = \frac{1}{9} \approx 0,11$.

Réponse. $\mathbb{E}(S_n) = 0$, $V(S_n) = n$; $\mathbb{P}(S_{2n} = 0) = \binom{2n}{n} 4^{-n}$; $\frac{3}{8}$ et $\frac{5}{16}$; borne $\frac{1}{9}$.

Correction de l'exercice 2 - Attendre PP ou PF

a) On attend un premier P (espérance 2), puis, dès lors, le premier F (espérance 2) : un F après un P forme PF , et les P intermédiaires ne font qu'attendre. Espérance 4.

b) Depuis l'état « rien » : un lancer, puis état « pile » (probabilité $\frac{1}{2}$) ou retour à « rien ». Depuis « pile » : un lancer, fin si pile, retour à « rien » si face.

c) $e_0 = 2 + e_1$ et $e_1 = 1 + \frac{1}{2}e_0$ donnent $e_0 = 3 + \frac{1}{2}e_0$:

$$e_0 = 6 \text{ (motif PP)} \quad \text{contre} \quad 4 \text{ (motif PF)}$$

d) Un échec sur PP (on avait P , on obtient F) renvoie **au départ**, tandis qu'un échec sur PF (on avait P , on obtient P) laisse encore un P utilisable. Les motifs PP ont tendance à se chevaucher (PPP en contient deux), donc à arriver « groupés » : à fréquence égale, ils sont plus espacés.

Réponse. Attente moyenne 4 pour PF , 6 pour PP .

Correction de l'exercice 3 - Approximation de Poisson

a) $X \sim B(100; 0,02) : \mathbb{P}(X = 0) = 0,98^{100} \approx 0,1326 ;$

$\mathbb{P}(X = 1) = 100 \times 0,02 \times 0,98^{99} \approx 0,2707.$

b) Poisson : $e^{-2} \approx 0,1353$ et $2e^{-2} \approx 0,2707$: accord à 0,003 près.

c) $\ln\left(1 - \frac{\lambda}{n}\right)^n = n \ln\left(1 - \frac{\lambda}{n}\right) = n\left(-\frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right) \rightarrow -\lambda$; de même $\binom{n}{k} \left(\frac{\lambda}{n}\right)^k \rightarrow \frac{\lambda^k}{k!}$ à k fixé.

d)

$$\mathbb{P}(X \geq 3) \approx 1 - e^{-2} \left(1 + 2 + \frac{2^2}{2}\right) = 1 - 5e^{-2} \approx 0,323$$

Réponse. $B(100; 0,02) : \mathbb{P}(X = 0) \approx 0,133$; Poisson : 0,135 ; $\mathbb{P}(X \geq 3) \approx 0,32$.

Correction de l'exercice 4 - Duel alterné

a) A gagne au k -ième tir si les $k - 1$ tours complets précédents sont ratés par les deux ($\frac{1}{2} \times \frac{2}{3} = \frac{1}{3}$ chacun) puis A touche : $\left(\frac{1}{3}\right)^{k-1} \frac{1}{2}$.

b)

$$\mathbb{P}(A) = \frac{1}{2} \sum_{k \geq 1} \left(\frac{1}{3}\right)^{k-1} = \frac{1}{2} \times \frac{1}{1 - \frac{1}{3}} = \frac{3}{4}$$

c) Si les deux ratent le premier tour, on se retrouve dans la situation initiale : $p = \frac{1}{2} + \frac{1}{3}p$,
 $p = \frac{3}{4}$ ✓

d) Avec B réussissant avec probabilité b : $p = \frac{1/2}{1 - \frac{1}{2}(1-b)} = \frac{1}{1+b}$. Équitable : $\frac{1}{1+b} = \frac{1}{2}$,
 soit $b = 1$! Même infallible, B ne fait que jeu égal : l'avantage de tirer en premier est considérable.

Réponse. $\mathbb{P}(A) = \frac{3}{4}$; en général $\frac{1}{1+b}$: il faudrait $b = 1$ pour l'équité.

Correction de l'exercice 5 - Position du premier as

a) $X > k \Leftrightarrow$ les 4 as sont parmi les $32 - k$ dernières positions ; les positions des as forment une partie uniforme à 4 éléments.

b) $\mathbb{P}(X = 1) = \frac{4}{32} = \frac{1}{8}$; $\mathbb{P}(X > 10) = \frac{\binom{22}{4}}{\binom{32}{4}} = \frac{7315}{35960} \approx 0,203.$

c) $\sum_{k=0}^{28} \binom{32-k}{4} = \sum_{j=4}^{32} \binom{j}{4} = \binom{33}{5}$, donc

$$\mathbb{E}(X) = \frac{\binom{33}{5}}{\binom{32}{4}} = \frac{33}{5} = 6,6$$

d) Même calcul : $\frac{\binom{N+1}{r+1}}{\binom{N}{r}} = \frac{N+1}{r+1}$. Les r as découpent en moyenne le paquet en $r + 1$ tranches égales de $\frac{N-r}{r+1}$ cartes non-as ; le premier as arrive après la première tranche.

Réponse. $\mathbb{P}(X = 1) = \frac{1}{8}$; $\mathbb{P}(X > 10) \approx 0,203$; $\mathbb{E}(X) = \frac{33}{5}$; en général $\frac{N+1}{r+1}$.

Correction de l'exercice 6 - Records d'une permutation aléatoire

a) Parmi les k premières personnes, chacune a la même probabilité d'être la plus grande (symétrie) : la k -ième l'est avec probabilité $\frac{1}{k}$.

b) Indicatrices et linéarité :

$$\mathbb{E}(R_n) = \sum_{k=1}^n \frac{1}{k} \sim \ln n$$

c) Tailles $1 < 2 < 3$: 123 (3 records), 132 (2), 213 (2), 231 (2), 312 (1), 321 (1) :

$$\mathbb{E} = \frac{11}{6} = 1 + \frac{1}{2} + \frac{1}{3} \quad \checkmark$$

d) Variance d'une somme d'indicatrices indépendantes : $V(R_n) = \sum_k \frac{1}{k} \left(1 - \frac{1}{k}\right)$. Pour $n = 3$: $0 + \frac{1}{4} + \frac{2}{9} = \frac{17}{36}$. Contrôle par la loi : $\mathbb{E}(R^2) = \frac{9+12+2}{6} = \frac{23}{6}$ et $\frac{23}{6} - \frac{121}{36} = \frac{138-121}{36} = \frac{17}{36}$ ✓

Réponse. $\mathbb{P}(k \text{ record}) = \frac{1}{k}$; $\mathbb{E}(R_n) = \sum \frac{1}{k}$; $n = 3$: $\frac{11}{6}$ et $V = \frac{17}{36}$.

Correction de l'exercice 7 - Le problème de la secrétaire

a) $k = 0$: on embauche le premier, meilleur avec probabilité $\frac{1}{3}$.

b) Valeurs $1 < 2 < 3$; on refuse le premier, on prend le suivant qui le dépasse :

— 123 → 2 ✗ ; 132 → 3 ✓ ; 213 → 3 ✓ ; 231 → 3 ✓ ; 312 → 2 (défaut) ✗ ; 321 → 1 (défaut) ✗.

— 3 succès sur 6 : $\frac{1}{2} > \frac{1}{3}$.

c) $\frac{11}{24} \approx 0,458 > \frac{5}{12} \approx 0,417$: refuser **un** candidat.

d) On consacre une première phase ($\approx 37\%$ des candidats) à **observer** pour fixer un seuil, puis on agit. Sans information, cette probabilité de succès $\approx 0,37$ ne tend pas vers 0 malgré $n \rightarrow +\infty$, ce qui est remarquable (le hasard seul donnerait $\frac{1}{n}$).

$$\mathbb{P}(\text{succès}) \xrightarrow{n \rightarrow +\infty} \frac{1}{e} \approx 0,368$$

Réponse. $n = 3$: $\frac{1}{3}$ sans refus, $\frac{1}{2}$ en refusant le premier ; $n = 4$: refuser un seul ($\frac{11}{24}$) ; asymptotiquement $\frac{1}{e}$.

Correction de l'exercice 8 - Fixer une prime d'assurance

a) $N \sim B(1000; 0,01)$: $\mathbb{E}(N) = 10$, $V(N) = 9,9$. Prime pure : $\frac{10 \times 10\,000}{1000} = 100$ euros.

b) $\mathbb{P}(N \geq 10 + a) \leq \mathbb{P}(|N - 10| \geq a) \leq \frac{9,9}{a^2} \leq 0,05 \Leftrightarrow a \geq \sqrt{198} \approx 14,1$.

c) Couvrir 24,1 sinistres : $1000p \geq 10\,000 \times 24,1$, soit

$$p \geq 241 \text{ euros}$$

d) $\mathbb{E}(N) = 1000$, $V(N) = 990$, $a \geq \sqrt{19\,800} \approx 141$: couvrir 1141 sinistres, prime $\frac{1141 \times 10\,000}{100\,000} \approx 114$ euros. La marge de sécurité par client est divisée par 10 ($\sqrt{100}$) : plus le portefeuille est grand, plus la prime s'approche de la prime pure.

Réponse. Prime pure 100 € ; avec 1000 clients, ≈ 241 € ; avec 100 000 clients, ≈ 114 €.

Correction de l'exercice 9 - Le modèle d'Ehrenfest à deux boules

a) De 0 ou 2 on va forcément en 1 ; de 1 on va en 0 ou 2 avec probabilité $\frac{1}{2}$:

$$P = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}$$

b) $\pi_0 = \frac{1}{2}\pi_1$, $\pi_2 = \frac{1}{2}\pi_1$, $\pi_1 = \pi_0 + \pi_2$: $\pi = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$, la loi $B(2, \frac{1}{2})$ (chaque boule dans A avec probabilité $\frac{1}{2}$, indépendamment).

c) $X_1 = 1$; puis $X_2 \in \{0, 2\}$ avec probabilité $\frac{1}{2}$ chacun, et ainsi de suite : instants pairs $(\frac{1}{2}, 0, \frac{1}{2})$, impairs $(0, 1, 0)$. La loi **ne converge pas** : la chaîne est **périodique** (période 2).

d) $\frac{1}{2}[(\frac{1}{2}, 0, \frac{1}{2}) + (0, 1, 0)] = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4}) = \pi$ ✓ Les moyennes temporelles convergent vers la loi stationnaire, même quand la loi elle-même oscille.

Réponse. $\pi = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$; lois alternées $(\frac{1}{2}, 0, \frac{1}{2}) / (0, 1, 0)$: pas de convergence (période 2).

Correction de l'exercice 10 - Sondage et taille d'échantillon

a) $nF_n \sim B(n, p)$: $V(F_n) = \frac{np(1-p)}{n^2}$, et $p(1-p) \leq \frac{1}{4}$ (maximum en $p = \frac{1}{2}$).

b) $\mathbb{P}(|F_n - p| \geq 0,1) \leq \frac{1}{4 \times 400 \times 0,01} = \frac{1}{16} \approx 0,063$.

c)

$$\frac{1}{4n \times 0,03^2} \leq 0,05 \Rightarrow n \geq \frac{1}{4 \times 0,0009 \times 0,05} \approx 5556$$

d) Bienaymé-Tchebychev vaut pour toutes les lois et est donc pessimiste ; le **théorème central limite** (loi normale, programme de L2) donne l'intervalle $\pm \frac{1,96}{2\sqrt{n}}$, soit

$$n \approx \left(\frac{1,96}{2 \times 0,03} \right)^2 \approx 1067.$$

Réponse. $V(F_n) \leq \frac{1}{4n}$; borne $\frac{1}{16}$ pour $n = 400$; $n \geq 5556$ par BT (environ 1067 par l'approximation normale).

Correction de l'exercice 11 - Le paradoxe du duc de Toscane

a) 9 : 126, 135, 144, 225, 234, 333. 10 : 136, 145, 226, 235, 244, 334.

b) Trois valeurs distinctes : $3! = 6$ ordres ; deux égales : 3 ; trois égales : 1.

— 9 : $6 + 6 + 3 + 3 + 6 + 1 = 25$. 10 : $6 + 6 + 3 + 6 + 3 + 3 = 27$.

c)

$$\mathbb{P}(S = 9) = \frac{25}{216} \approx 0,116, \quad \mathbb{P}(S = 10) = \frac{27}{216} = \frac{1}{8} = 0,125$$

d) Les décompositions ne sont **pas équiprobables** : 333 n'arrive que d'une façon, 234 de six. L'univers équiprobable est celui des triplets ordonnés (les dés sont discernables). Le duc avait raison d'observer la différence ; Galilée l'a expliquée ainsi.

Réponse. 25 contre 27 triplets : $\mathbb{P}(S = 9) = \frac{25}{216}$, $\mathbb{P}(S = 10) = \frac{1}{8}$.

Correction de l'exercice 12 - Anniversaires : compter les paires

a) $X = \sum_{\{i,j\}} 1_{A_{ij}}$ où A_{ij} : « i et j sont nés le même jour », de probabilité $\frac{1}{365}$:

$$\mathbb{E}(X) = \binom{n}{2} \frac{1}{365}$$

b) $\frac{253}{365} \approx 0,693$ ($\approx \ln 2$, comme dans l'exercice des anniversaires de la plaquette 1).

c) $\mathbb{P}(X \geq 1) \leq \mathbb{E}(X) \approx 0,693$; la vraie valeur 0,507 est plus petite : les paires ne sont pas indépendantes, et une coïncidence entraîne souvent plusieurs paires.

d) $\frac{n(n-1)}{2} \geq 365 \Leftrightarrow n(n-1) \geq 730 : 27 \times 26 = 702 < 730 \leq 28 \times 27 = 756$, donc $n = 28$.

Réponse. $\mathbb{E}(X) = \binom{n}{2} / 365$; 0,693 pour $n = 23$; $\mathbb{E}(X) \geq 1$ dès $n = 28$.

Correction de l'exercice 13 - Le paradoxe de Saint-Petersbourg

a) $\mathbb{P}(\text{premier pile au lancer } k) = \frac{1}{2^k}$:

$$\mathbb{E}(G_m) = \sum_{k=1}^m 2^k \times \frac{1}{2^k} = m$$

b) $\mathbb{E}(G_m) = m \rightarrow +\infty$: le jeu sans limite a une espérance **infinie**, et aucun prix fini ne le rendrait équitable au sens de l'espérance. Pourtant personne ne paierait 1000 euros pour y jouer.

c) $G_m \geq 64 = 2^6 \Leftrightarrow \text{premier pile au lancer } k \geq 6$: probabilité $\sum_{k=6}^m 2^{-k} = 2^{-5} - 2^{-m} \leq \frac{1}{32}$.

d) Avec $m = 20$: $\mathbb{E} = 20$ euros. L'espérance infinie vient de gains **astronomiques** et **improbables** qu'aucune banque ne peut verser ; d'où l'idée (Bernoulli) de mesurer l'utilité d'un gain et non son montant.

Réponse. $\mathbb{E}(G_m) = m$; espérance infinie sans limite ; avec plafond 2^{20} : 20 €.

Correction de l'exercice 14 - Un dé, puis des pièces

a) Sachant $N = n$, $X \sim B(n, \frac{1}{2})$.

b) $\mathbb{P}(X = 0) = \sum_{n=1}^6 \frac{1}{6} \times \frac{1}{2^n} = \frac{1}{6} \left(1 - \frac{1}{64}\right) = \frac{63}{384} = \frac{21}{128} \approx 0,164$.

c)

$$\mathbb{E}(X) = \sum_{n=1}^6 \frac{1}{6} \times \frac{n}{2} = \frac{1}{2} \mathbb{E}(N) = \frac{7}{4} = 1,75$$

d) $X = 6$ impose $N = 6$: probabilité 1.

$\mathbb{P}(N = 1 \mid X = 0) = \frac{\frac{1}{6} \times \frac{1}{2}}{\frac{21}{128}} = \frac{1/12}{21/128} = \frac{128}{252} = \frac{32}{63} \approx 0,508$: n'avoir aucun pile rend « un seul lancer » très probable.

Réponse. $\mathbb{P}(X = 0) = \frac{21}{128}$; $\mathbb{E}(X) = \frac{7}{4}$; $\mathbb{P}(N = 1 \mid X = 0) = \frac{32}{63}$.

Correction de l'exercice 15 - Synthèse : dépistage groupé (méthode de Dorfman)

a) Indépendance : $(1-p)^k = 0,99^k$.

b) Pour un groupe, le nombre de tests T vaut 1 avec probabilité $(1-p)^k$ et $k+1$ sinon :

$$\mathbb{E}(T) = 1 + k(1 - (1-p)^k), \quad f(k) = \frac{\mathbb{E}(T)}{k} = \frac{1}{k} + 1 - (1-p)^k$$

c) $f(5) \approx 0,2 + 0,049 = 0,249$; $f(10) \approx 0,1 + 0,0956 = 0,1956$;

$f(11) \approx 0,0909 + 0,1047 = 0,1956$; $f(15) \approx 0,0667 + 0,1399 = 0,2066$. Optimum vers $k = 10$ ou 11 (valeurs égales à 10^{-4} près).

d) Environ 0,196 test par personne au lieu de 1 : 80 % **d'économie**. Pour $p = 0,3$: $f(2) = 0,5 + 1 - 0,49 = 1,01 > 1$, $f(3) = 0,333 + 1 - 0,343 = 0,99$, $f(4) \approx 0,25 + 0,76 = 1,01$: quasiment aucun gain. La méthode n'est utile que pour une prévalence **faible**.

Réponse. $f(k) = \frac{1}{k} + 1 - 0,99^k$; optimum $k \approx 10-11$, $\approx 0,196$ test par personne (80 % d'économie) ; inutile pour $p = 0,3$.