Plaquette 5 – L'espérance pour décider : jeux, stratégies, algorithmes

Espérance et variance — Prépa MPSI / MP2I (1re année)

Corrections détaillées

Boîte à outils

Décider avec l'espérance.

— **Espérance par conditionnement** : si (A_i) est un système complet,

$$\mathbb{E}(X) = \sum_i \mathbb{P}(A_i)\,\mathbb{E}(X \mid A_i), \qquad \mathbb{E}(X \mid A) = \sum_x x\,\mathbb{P}(X = x \mid A)$$

— **Raisonnement à rebours** : dans un jeu en plusieurs étapes, on détermine d'abord la valeur (espérance optimale) de la dernière étape, puis on remonte.

— **Jeu équitable** : espérance de gain nulle ; un jeu défavorable a une espérance négative, que la loi des grands nombres rend visible à long terme.

— **Coût moyen d'un algorithme** : on décompose le coût en somme d'indicatrices (« la comparaison i a lieu ») et l'on applique la linéarité.

— $\mathbb{E}(X)^2 \leq \mathbb{E}(X^2)$ (car $\mathrm{V}(X) \geq 0$) ; pour $X > 0$: $\mathbb{E}\left(\frac{1}{X}\right) \geq \frac{1}{\mathbb{E}(X)}$.

Correction de l'exercice 1 – Relancer ou s'arrêter ?

a) 3,5.

b) Si l'on relance, l'espérance du second lancer est 3,5. On garde le premier s'il vaut mieux que ce qu'on espère en relançant : on garde 4, 5, 6. Espérance par conditionnement :

$$\mathbb{E}(G) = \mathbb{P}(D_1 \geq 4) \times 5 + \mathbb{P}(D_1 \leq 3) \times 3{,}5 = \frac{1}{2} \times 5 + \frac{1}{2} \times 3{,}5 = 4{,}25$$

— (Pour $s = 5$: $\frac{1}{3} \times 5{,}5 + \frac{2}{3} \times 3{,}5 \approx 4{,}17$; pour $s = 3$: $\frac{2}{3} \times 4{,}5 + \frac{1}{3} \times 3{,}5 \approx 4{,}17$. Le seuil 4 est optimal.)

c) À rebours : avec deux lancers restants, la valeur est 4,25. Au premier lancer, on garde donc un résultat $> 4{,}25$, c'est-à-dire 5 ou 6 :

— $\mathbb{E}(G) = \frac{1}{3} \times 5{,}5 + \frac{2}{3} \times 4{,}25 = \frac{14}{3} \approx 4{,}67$.

Réponse. 3,5 ; seuil 4 : 4,25 ; trois lancers : garder 5, 6 puis seuil 4, espérance $\frac{14}{3}$.

Correction de l'exercice 2 – Le paradoxe de Saint-Pétersbourg

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

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

b) Une mise de n € rend le jeu équitable. L'espérance tend vers $+\infty$: un « jeu infini » aurait une valeur infinie, alors que personne ne paierait une grosse somme pour y jouer (c'est le paradoxe).

c) $2^k > 1\,000 \Leftrightarrow k \geq 10$: $\mathbb{P} = \sum_{k=10}^{20} \frac{1}{2^k} \approx \frac{1}{2^9} = 0{,}002$. L'espérance (20 €) est portée par des gains énormes et très rares ; le gain typique est de 2 ou 4 €. L'espérance seule ne suffit

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

pas à décider : il faut tenir compte du risque.

Réponse. $\mathbb{E}(G) = n$; mise équitable n € ; $\mathbb{P}(G > 1000) \approx 0{,}002$ pour $n = 20$.

Correction de l'exercice 3 – Coût moyen d'une recherche séquentielle

a) Si l'élément est en position k, il faut k comparaisons :

$$\mathbb{E}(C) = \sum_{k=1}^{n} \frac{k}{n} = \frac{n+1}{2}$$

b) Par conditionnement : $\mathbb{E}(C) = q\,n + (1-q)\dfrac{n+1}{2}$.

c) Avec $\mathbb{P}(k) = \dfrac{1/k}{H_n}$ (positions triées par popularité décroissante) : $\mathbb{E}(C) = \sum_k k \dfrac{1/k}{H_n} = \dfrac{n}{H_n}$. Pour $n = 4$: $\dfrac{4}{25/12} = \dfrac{48}{25} = 1{,}92$, contre 2,5 dans le cas uniforme. Placer les éléments fréquents en tête réduit le coût moyen.

Réponse. $\dfrac{n+1}{2}$; $qn + (1-q)\dfrac{n+1}{2}$; Zipf trié : $\dfrac{n}{H_n} = 1{,}92$ pour $n = 4$.

Correction de l'exercice 4 – Comparaisons dans le tri rapide

a) $C_2 = 1$. Pour $n = 3$: 2 comparaisons avec le pivot ; si le pivot est l'élément médian (probabilité $\frac{1}{3}$), il reste deux tableaux d'un élément ; sinon, un tableau de 2 éléments (1 comparaison). $C_3 = 2 + \frac{2}{3} \times 1 = \frac{8}{3}$.

b) Si le pivot est de rang $k+1$ (probabilité $\frac{1}{n}$), les sous-tableaux ont k et $n-1-k$ éléments :

$$C_n = n - 1 + \frac{1}{n}\sum_{k=0}^{n-1} (C_k + C_{n-1-k}) = n - 1 + \frac{2}{n}\sum_{k=0}^{n-1} C_k$$

c) $C_4 = 3 + \frac{2}{4}\left(0 + 0 + 1 + \frac{8}{3}\right) = 3 + \frac{11}{6} = \frac{29}{6}$. Formule : $n = 3$: $8 \times \frac{11}{6} - 12 = \frac{88}{6} - 12 = \frac{8}{3}$; $n = 4$: $10 \times \frac{25}{12} - 16 = \frac{250}{12} - 16 = \frac{29}{6}$. En moyenne, $C_n \sim 2n \ln n$.

Réponse. $C_2 = 1$, $C_3 = \frac{8}{3}$, $C_4 = \frac{29}{6}$; $C_n = 2(n+1)H_n - 4n$.

Correction de l'exercice 5 – Le jeu du Chuck-a-luck

a) $K \sim B\left(3, \frac{1}{6}\right)$: $\mathbb{P}(K=0) = \frac{125}{216}$, $\mathbb{P}(K=1) = \frac{75}{216}$, $\mathbb{P}(K=2) = \frac{15}{216}$, $\mathbb{P}(K=3) = \frac{1}{216}$. $G = K$ si $K \geq 1$, $G = -1$ si $K = 0$.

b) $\mathbb{E}(G) = \dfrac{1 \times 75 + 2 \times 15 + 3 \times 1 - 125}{216} = -\dfrac{17}{216} \approx -0{,}079$ € : le joueur perd en moyenne près de 8 centimes par euro misé. Jeu défavorable.

c) La probabilité de gagner est $\mathbb{P}(K \geq 1) = 1 - \left(\frac{5}{6}\right)^3 = \frac{91}{216} \approx 0{,}42 < \frac{1}{2}$: les événements « le dé i montre le numéro » ne sont pas incompatibles, on ne peut pas additionner leurs probabilités.

Réponse. $K \sim B(3, \frac{1}{6})$; $\mathbb{E}(G) = -\frac{17}{216}$; $\mathbb{P}(\text{gagner}) = \frac{91}{216}$.

Correction de l'exercice 6 – Le problème du marchand de journaux

a) On vend $\min(D, q)$ journaux à 3 € et l'on en a payé q.

b) $\min(D, q)$ est à valeurs dans $[\![0, q]\!]$ et $\min(D, q) \geq k \Leftrightarrow D \geq k$ pour $k \leq q$; on applique $\mathbb{E}(X) = \sum_{k \geq 1} \mathbb{P}(X \geq k)$.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

c) $\mathbb{P}(D \geq k) = \frac{5-k}{5}$: 0,8 ; 0,6 ; 0,4 ; 0,2.

— $\mathbb{E}(B_0) = 0$; $\mathbb{E}(B_1) = 3 \times 0{,}8 - 1 = 1{,}4$; $\mathbb{E}(B_2) = 3 \times 1{,}4 - 2 = 2{,}2$; $\mathbb{E}(B_3) = 3 \times 1{,}8 - 3 = 2{,}4$; $\mathbb{E}(B_4) = 3 \times 2 - 4 = 2$.

— Commande optimale : $q = 3$ journaux.

d) Passer de q à $q+1$ change l'espérance de $3\,\mathbb{P}(D \geq q+1) - 1$: on gagne à commander un journal de plus tant que $\mathbb{P}(D \geq q+1) \geq \frac{1}{3}$, rapport du coût au prix de vente. Ici $\mathbb{P}(D \geq 3) = 0{,}4 \geq \frac{1}{3}$ mais $\mathbb{P}(D \geq 4) = 0{,}2 < \frac{1}{3}$.

Réponse. $\mathbb{E}(B_q)$: $0; 1{,}4; 2{,}2; 2{,}4; 2$; optimum $q = 3$; critère $\mathbb{P}(D \geq q+1) \geq \frac{1}{3}$.

Correction de l'exercice 7 – Collisions dans une table de hachage

a) $C = \sum_{\{i,j\}} \mathbf{1}_{\{h(i) = h(j)\}}$, chaque indicatrice valant 1 avec probabilité $\frac{1}{m}$:

$$\mathbb{E}(C) = \frac{\binom{n}{2}}{m}$$

b) $\frac{499\,500}{m} \leq 5 \Leftrightarrow m \geq 99\,900$ cases (environ $100n$).

c) Pour une clé i donnée, le nombre d'autres clés dans sa case est $\sum_{j \neq i} \mathbf{1}_{\{h(j) = h(i)\}}$, d'espérance $\frac{n-1}{m}$. Avec la clé elle-même : $1 + \frac{n-1}{m}$ comparaisons en moyenne (environ 2 si $m = n$).

Réponse. $\mathbb{E}(C) = \frac{\binom{n}{2}}{m}$; $m \geq 99\,900$; coût $1 + \frac{n-1}{m}$.

Correction de l'exercice 8 – Distance moyenne d'une marche aléatoire

a) $S_n = \sum \varepsilon_i$ avec $\mathbb{E}(\varepsilon_i) = 0$: $\mathbb{E}(S_n) = 0$. $\mathbb{E}(S_n^2) = \mathrm{V}(S_n) = \sum \mathrm{V}(\varepsilon_i) = n$ (indépendance, $\varepsilon_i^2 = 1$).

b) $S_4 \in \{-4, -2, 0, 2, 4\}$ avec probabilités $\frac{1}{16}, \frac{4}{16}, \frac{6}{16}, \frac{4}{16}, \frac{1}{16}$:

$$\mathbb{E}(|S_4|) = \frac{4 + 8 + 0 + 8 + 4}{16} = \frac{24}{16} = 1{,}5$$

c) $\mathbb{E}(|S_n|)^2 \leq \mathbb{E}(S_n^2) = n$ (variance de $|S_n|$ positive). La distance typique au point de départ croît comme $\sqrt{n}$ (et non comme n) : c'est la **diffusion**. Pour $n = 4$: $1{,}5 \leq 2$.

Réponse. $\mathbb{E}(S_n) = 0$, $\mathbb{E}(S_n^2) = n$; $\mathbb{E}(|S_4|) = 1{,}5$; $\mathbb{E}(|S_n|) \leq \sqrt{n}$.

Correction de l'exercice 9 – Espérance par conditionnement : un dé choisi au hasard

a) $\mathbb{E}(X \mid \text{dé 4}) = \frac{5}{2}$ et $\mathbb{E}(X \mid \text{dé 6}) = \frac{7}{2}$:

$$\mathbb{E}(X) = \frac{1}{2} \times \frac{5}{2} + \frac{1}{2} \times \frac{7}{2} = 3$$

b) Pour $k \leq 4$: $\mathbb{P}(X = k) = \frac{1}{2} \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{6} = \frac{5}{24}$; pour $k \in \{5, 6\}$: $\frac{1}{12} = \frac{2}{24}$.
$\mathbb{E}(X) = \frac{5(1+2+3+4) + 2(5+6)}{24} = \frac{50 + 22}{24} = 3$.

c) $\mathbb{E}(X^2) = \frac{5(1+4+9+16) + 2(25+36)}{24} = \frac{150 + 122}{24} = \frac{34}{3}$, $\mathrm{V}(X) = \frac{34}{3} - 9 = \frac{7}{3} \approx 2{,}33$.
Variances conditionnelles : $\frac{15}{12}$ et $\frac{35}{12}$, de moyenne $\frac{25}{12} \approx 2{,}08$. La variance totale est plus grande : elle ajoute la dispersion **entre** les moyennes conditionnelles ($\frac{1}{2}(0{,}5)^2 + \frac{1}{2}(0{,}5)^2 = 0{,}25$), et $\frac{25}{12} + \frac{1}{4} = \frac{7}{3}$.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

Réponse. $\mathbb{E}(X) = 3$; $\mathrm{V}(X) = \frac{7}{3} = \frac{25}{12} + \frac{1}{4}$.

Correction de l’exercice 10 – L’espérance de l’inverse

a) $UV = 1$, donc

$$1 = \mathbb{E}(UV)^2 \le \mathbb{E}(U^2)\,\mathbb{E}(V^2) = \mathbb{E}(X)\,\mathbb{E}\left(\frac{1}{X}\right)$$

b) Durée $T = \frac{10}{V}$: $\mathbb{E}(T) = \frac{1}{2}(1 + \frac{1}{3}) = \frac{2}{3}$ h = 40 min. Mais $\mathbb{E}(V) = 20$ km/h et $\frac{10}{20} = 30$ min. La durée moyenne est plus grande que la durée à vitesse moyenne : c’est a).

c) Égalité dans Cauchy-Schwarz si U et V sont (presque sûrement) proportionnelles, c’est-à-dire X presque sûrement constante.

Réponse. $\mathbb{E}(X)\mathbb{E}(\frac{1}{X}) \ge 1$; 40 min contre 30 min ; égalité ssi X constante.

Correction de l’exercice 11 – Algorithme de Las Vegas

a) $N \ge k$ ($1 \le k \le n$) si et seulement si les $k-1$ premiers essais échouent :

$$\mathbb{E}(N) = \sum_{k=1}^{n} (1-p)^{k-1} = \frac{1-(1-p)^n}{p}$$

b) $(1-p)^n$.

c) $n = 10$: $\mathbb{E}(N) = \frac{1-0{,}8^{10}}{0{,}2} \approx \frac{1-0{,}107}{0{,}2} \approx 4{,}46$ s ; échec $\approx 10{,}7\,\%$. $n = 30$: $\mathbb{E}(N) \approx \frac{1-0{,}00124}{0{,}2} \approx 4{,}99$ s ; échec $\approx 0{,}12\,\%$. Autoriser plus d’essais coûte à peine plus en moyenne mais divise l’échec par 90.

Réponse. $\mathbb{E}(N) = \frac{1-(1-p)^n}{p}$; échec $(1-p)^n$; $n = 10$: 4,46 s, 10,7 % ; $n = 30$: 4,99 s, 0,12 %.

Correction de l’exercice 12 – Une loterie

a) $G = 995$ avec probabilité $\frac{1}{1\,000}$, 195 avec $\frac{5}{1\,000}$, 20 avec $\frac{20}{1\,000}$, -5 avec $\frac{974}{1\,000}$.

b) $\mathbb{E}(G) = \frac{995 + 975 + 400 - 4\,870}{1\,000} = -2{,}5$ € : on perd en moyenne la moitié de sa mise.

$$\mathbb{E}(G^2) = \frac{995^2 + 5 \times 195^2 + 20 \times 20^2 + 974 \times 25}{1\,000} = \frac{990\,025 + 190\,125 + 8\,000 + 24\,350}{1\,000} \approx 1\,212{,}5$$

— $\mathrm{V}(G) \approx 1\,212{,}5 - 6{,}25 = 1\,206{,}25$, $\sigma \approx 34{,}7$ €.

c) Lots : $1\,000 + 1\,000 + 500 = 2\,500$ € sur 5 000 € de recettes, soit 50 %. Les joueurs acceptent une espérance négative pour une petite chance d’un gain important (goût du risque, rêve) : l’espérance ne résume pas leur décision.

Réponse. $\mathbb{E}(G) = -2{,}5$ €, $\sigma \approx 34{,}7$ € ; 50 % redistribués.

Correction de l’exercice 13 – Parier sur une martingale

a) Il a perdu $1 + 2 + \ldots + 2^{k-2} = 2^{k-1} - 1$ puis gagne 2^{k-1} : bilan +1.

b) $1 + 2 + \ldots + 2^{n-1} = 2^n - 1$ €.

c) $q = \frac{19}{37}$; perte totale avec probabilité q^n :

$$\mathbb{E}(G) = (1-q^n) \times 1 - q^n(2^n - 1) = 1 - (2q)^n$$

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s’expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

— $n = 10$: $2q = \frac{38}{37} \approx 1{,}027$, $(2q)^{10} \approx 1{,}306$, donc $\mathbb{E}(G) \approx -0{,}31$ €. L'espérance est négative : on gagne souvent 1 € (99,9 % des cas), mais on perd rarement 1 023 €. Aucune stratégie de mise ne rend positif un jeu défavorable.

Réponse. Gain +1 € si victoire, perte $2^n - 1$ sinon ; $\mathbb{E}(G) = 1 - (2q)^n \approx -0{,}31$ € pour $n = 10$.

Correction de l'exercice 14 – Espérance et médiane

a) $\mathbb{E}(X) = \frac{9 \times 2\,000 + 20\,000}{10} = 3\,800$ € ; $M = 2\,000$ € (90 % des salaires).

b) En $a = 2\,000$: $\mathbb{E}(|X - a|) = \frac{18\,000}{10} = 1\,800$; en $a = 3\,800$: $\frac{9 \times 1\,800 + 16\,200}{10} = 3\,240$. La médiane donne l'écart absolu moyen le plus petit (on peut montrer que c'est un minimum). Pour l'écart quadratique :

$$\mathbb{E}((X - a)^2) = \mathrm{V}(X) + (3\,800 - a)^2$$

— minimal en $a = 3\,800$; en $a = 2\,000$ il vaut $\mathrm{V}(X) + 1\,800^2$.

c) La médiane : l'espérance est tirée vers le haut par une seule valeur extrême, et ne correspond au salaire de personne.

Réponse. $\mathbb{E}(X) = 3\,800$ €, $M = 2\,000$ € ; médiane minimise $\mathbb{E}|X - a|$, espérance minimise $\mathbb{E}(X - a)^2$.

Correction de l'exercice 15 – Synthèse : les tests groupés de Dorfman

a) Le mélange est négatif si les k personnes sont saines, avec probabilité $(1 - p)^k$:

$$\mathbb{E}(T) = 1 \times (1 - p)^k + (1 + k)(1 - (1 - p)^k) = 1 + k - k(1 - p)^k$$

b) Divisé par k : $f(k) = \frac{1}{k} + 1 - (1 - p)^k$.

c) $f(5) = 0{,}2 + 1 - 0{,}99^5 \approx 0{,}249$; $f(10) \approx 0{,}1 + 0{,}0956 = 0{,}1956$; $f(11) \approx 0{,}0909 + 0{,}1047 = 0{,}1956$ (optimum, légèrement meilleur que 10). On économise environ 80 % des tests par rapport au dépistage individuel (1 test par personne).

d) $f(2) < 1 \Leftrightarrow \frac{1}{2} < (1 - p)^2 \Leftrightarrow p < 1 - \frac{1}{\sqrt{2}} \approx 0{,}29$. Au-delà d'environ 30 % de prévalence, grouper n'apporte plus rien (les mélanges sont presque toujours positifs). La méthode, proposée en 1943, a été très utilisée pendant l'épidémie de Covid-19.

Réponse. $\mathbb{E}(T) = 1 + k - k(1 - p)^k$; $f(10) \approx 0{,}196$, optimum $k = 11$; utile si $p < 1 - \frac{1}{\sqrt{2}}$.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).