

Approfondissement Variables aléatoires

Pour aller plus loin : démonstrations de cours, tests groupés, relancer ou s'arrêter, le paradoxe de Saint-Petersbourg, espérance du nombre de succès, inégalité de Bienaymé-Tchebychev

6 exercices · niveau ●●● · durée conseillée : 3 h · calculatrice autorisée

Exercice 1 Démonstrations : linéarité et formule de König-Huygens

●●●

Une variable aléatoire X prend les valeurs $x_1, \dots, x_n$ avec les probabilités $p_1, \dots, p_n$; on note $\mu = E(X)$. On rappelle que $p_1 + \dots + p_n = 1$; a et b sont deux réels, avec $a \neq 0$.

- 1.1 Justifier que $aX + b$ prend les valeurs $ax_i + b$ avec les probabilités p_i , puis démontrer que $E(aX + b) = aE(X) + b$ et que $V(aX + b) = a^2V(X)$.
- 1.2 En développant $(x_i - \mu)^2$ dans la définition de $V(X)$, démontrer la formule de König-Huygens : $V(X) = E(X^2) - \mu^2$. En déduire que $E(X^2) \geq E(X)^2$.
- 1.3 Pour tout réel t , on pose $f(t) = p_1(x_1 - t)^2 + \dots + p_n(x_n - t)^2$ (écart quadratique moyen entre X et t). Démontrer que $f(t) = t^2 - 2\mu t + E(X^2)$, puis que $f(t) = (t - \mu)^2 + V(X)$. Pour quelle valeur de t la fonction f atteint-elle son minimum ? Quel est ce minimum ? Interpréter.
- 1.4 Application. X prend la valeur 1 avec la probabilité p et la valeur 0 avec la probabilité $1 - p$. Démontrer que $V(X) = p(1 - p)$. Pour quelle valeur de p cette variance est-elle maximale ?

Exercice 2 Des tests groupés pour économiser des analyses

●●●

Un laboratoire recherche une infection dans le sang de 4 personnes; chacune est infectée avec la probabilité p , indépendamment des autres, et on pose $q = 1 - p$. Le test est supposé parfaitement fiable, y compris sur un mélange de sangs (il est positif si, et seulement si, au moins un des sangs mélangés est infecté). On compare la **méthode individuelle** (on teste chacun des 4 prélèvements, soit 4 tests) et la **méthode groupée** (on teste le mélange des 4 prélèvements; s'il est négatif, on s'arrête; sinon, on teste en plus chaque prélèvement séparément). On note T le nombre de tests réalisés avec la méthode groupée.

- 2.1 On suppose $p = 0,1$. Donner la loi de T et calculer $E(T)$. Quelle économie moyenne de tests, en pourcentage, la méthode groupée permet-elle ?
- 2.2 Dans le cas général, démontrer que $E(T) = 5 - 4q^4$. Pour quelles valeurs de p la méthode groupée est-elle, en moyenne, plus économe que la méthode individuelle ? (Donner la valeur exacte de la borne, puis un arrondi à 10^{-3} .)
- 2.3 Troisième méthode : on forme deux groupes de deux personnes, et l'on applique la méthode groupée à chaque groupe. On note T' le nombre total de tests. Démontrer que $E(T') = 6 - 4q^2$, puis que $E(T') - E(T) = (2q^2 - 1)^2$. Conclure, puis identifier les trois courbes de la figure 1.

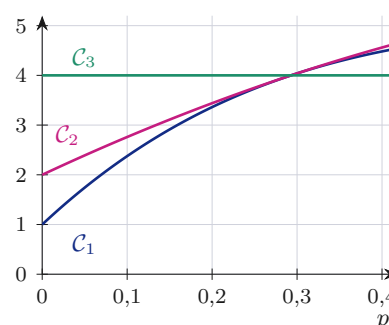


Figure 1 — Nombres moyens de tests en fonction de p

Exercice 3 Relancer ou s'arrêter ?

●●●

On lance un dé cubique équilibré. Le joueur peut garder le résultat, qui est alors son gain en euros, ou relancer le dé une seule fois : le second résultat est alors son gain, même s'il est moins bon. Pour un entier k compris entre 1 et 7, la stratégie $\mathcal{S}_k$ consiste à relancer si, et seulement si, le premier résultat est strictement inférieur à k ; on note G_k le gain obtenu avec cette stratégie.

- 3.1 Que sont les stratégies $\mathcal{S}_1$ et $\mathcal{S}_7$? Justifier que $E(G_1) = E(G_7) = 7/2$.

- 3.2** À l'aide d'un arbre, déterminer la loi de G_4 et démontrer que $E(G_4) = 17/4$.
- 3.3** Démontrer que, pour tout entier k compris entre 1 et 7, $E(G_k) = \frac{-k^2 + 8k + 35}{12}$. En déduire la meilleure stratégie.
- 3.4** On autorise maintenant deux relances (trois lancers au plus). En raisonnant à partir du dernier lancer, justifier qu'au deuxième lancer il faut garder le résultat s'il vaut au moins 4, puis déterminer la stratégie optimale au premier lancer et démontrer que l'espérance du gain vaut alors $14/3$.

Exercice 4 Le paradoxe de Saint-Pétersbourg



Un casino propose le jeu suivant : on lance une pièce équilibrée jusqu'à obtenir « pile », avec au plus n lancers. Si « pile » apparaît pour la première fois au k -ième lancer, le joueur reçoit 2^k euros ; s'il n'apparaît pas au cours des n lancers, le joueur ne reçoit rien. On note G_n la somme reçue.

- 4.1** Pour $n = 4$, construire l'arbre, donner la loi de G_4 , puis calculer $E(G_4)$ et $\sigma(G_4)$.
- 4.2** Démontrer que $E(G_n) = n$ et, à l'aide de la somme des termes d'une suite géométrique, que $V(G_n) = 2^{n+1} - 2 - n^2$.
- 4.3** Quelle mise rend le jeu équitable ? Que se passe-t-il si l'on ne limite plus le nombre de lancers ? Pourtant, peu de joueurs accepteraient de payer 30 euros pour jouer avec $n = 30$: calculer la probabilité de recevoir au moins 30 euros et commenter.

Exercice 5 L'espérance du nombre de succès



On répète quatre fois, de façon indépendante, une épreuve de Bernoulli de paramètre p , et on note X le nombre de succès. On pose $q = 1 - p$.

- 5.1** Justifier, en dénombrant les chemins de l'arbre, que $P(X = 2) = 6p^2q^2$, puis donner la loi de X .
- 5.2** Vérifier que, pour tous réels a et b , $(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$. En déduire que $E(X) = 4p$.
- 5.3** (Hors programme — pour aller plus loin.) Pour i de 1 à 4, on note X_i la variable aléatoire qui vaut 1 si la i -ième épreuve est un succès et 0 sinon, de sorte que $X = X_1 + X_2 + X_3 + X_4$. En admettant que l'espérance d'une somme est la somme des espérances, retrouver sans calcul que $E(X) = 4p$. Vérifier enfin, pour $p = \frac{1}{2}$, que $V(X) = 1$.

Exercice 6 L'inégalité de Bienaymé-Tchebychev



(Hors programme — pour aller plus loin.) X est une variable aléatoire d'espérance μ et d'écart type $\sigma > 0$, qui prend les valeurs $x_1, \dots, x_n$ avec les probabilités $p_1, \dots, p_n$. Soit δ un réel strictement positif.

- 6.1** Dans la somme $V(X) = p_1(x_1 - \mu)^2 + \dots + p_n(x_n - \mu)^2$, ne conserver que les termes pour lesquels $|x_i - \mu| \geq \delta$ et en déduire que $P(|X - \mu| \geq \delta) \leq V(X)/\delta^2$. Que devient cette inégalité pour $\delta = 2\sigma$?
- 6.2** X prend les valeurs $-2, 0$ et 2 avec les probabilités $\frac{1}{8}, \frac{3}{4}$ et $\frac{1}{8}$. Calculer μ et σ , puis $P(|X - \mu| \geq 2\sigma)$. Que montre cet exemple ?
- 6.3** On admet que la moyenne M d'un échantillon de taille n de la loi de X a pour espérance μ et pour variance σ^2/n . Démontrer que la probabilité que M soit en dehors de l'intervalle $]\mu - 2\sigma/\sqrt{n}; \mu + 2\sigma/\sqrt{n}[$ est au plus $1/4$. Les simulations donnent en pratique une proportion voisine de 5 % : commenter.