

Plaquette 3 - Espérance et variance

Loi de Bernoulli et loi binomiale — Première spé maths

Corrections détaillées

Boîte à outils

Pour $X \sim B(n; p)$:

$$E(X) = np \quad V(X) = np(1 - p) \quad \sigma(X) = \sqrt{np(1 - p)}$$

Pourquoi. X est la somme de n variables de Bernoulli indépendantes, chacune d'espérance p et de variance $p(1 - p)$: les espérances s'ajoutent, et les variances aussi grâce à l'indépendance.

Interprétation.

- $E(X) = np$: le nombre **moyen** de succès sur n répétitions ;
- $\sigma(X)$: l'ampleur typique des **fluctuations** autour de cette moyenne.

Ordre de grandeur des fluctuations. L'écart type croît comme $\sqrt{n}$, alors que l'espérance croît comme n : en proportion, les fluctuations **diminuent** quand l'échantillon grandit.

$$\frac{\sigma(X)}{E(X)} = \frac{\sqrt{np(1 - p)}}{np} = \sqrt{\frac{1 - p}{np}}$$

La fréquence de succès $F = \frac{X}{n}$ vérifie :

$$E(F) = p \quad \sigma(F) = \frac{\sqrt{p(1 - p)}}{\sqrt{n}}$$

Règle pratique de décision. Un résultat situé à plus de **deux écarts types** de l'espérance est considéré comme inhabituel : il conduit à interroger le modèle.

Cas particulier. $V(X)$ est maximale pour $p = 0,5$: l'incertitude est la plus grande quand les deux issues sont équiprobables.

Correction de l'exercice 1 - Appliquer les deux formules

On applique directement les formules du cours.

$$E(X) = np = 50 \times 0,2 = 10$$

$$V(X) = np(1 - p) = 50 \times 0,2 \times 0,8 = 8$$

$$\sigma(X) = \sqrt{8} = 2\sqrt{2} \approx 2,83$$

Interprétation. Sur 50 répétitions, on obtient en moyenne 10 succès, avec des fluctuations typiques de l'ordre de 3 : observer entre 7 et 13 succès est banal.

Contrôles.

- $E(X) = 10$ est bien compris entre 0 et $n = 50$ ✓
- $V(X) \leq E(X)$ toujours, puisque $1 - p \leq 1$ ✓
- $\sigma(X) \approx 2,83$ est bien inférieur à $\sqrt{np} \approx 3,16$ ✓

L'erreur à éviter. Écrire $V(X) = np^2$ ou $\sigma(X) = np(1 - p)$: la variance contient **un seul** p et **un seul** $(1 - p)$, et l'écart type est sa **racine carrée**.

Réponse. $E(X) = 10$, $V(X) = 8$, $\sigma(X) \approx 2,83$.

Correction de l'exercice 2 - Combien de six en moyenne ?

X suit $B\left(60; \frac{1}{6}\right)$.

a) **Espérance** :

$$E(X) = 60 \times \frac{1}{6} = 10$$

— **Variance** :

$$V(X) = 60 \times \frac{1}{6} \times \frac{5}{6} = \frac{300}{36} \approx 8,33$$

— **Écart type** : $\sigma(X) = \sqrt{8,33} \approx 2,89$.

— b) **Intervalle à deux écarts types** :

$$[E(X) - 2\sigma; E(X) + 2\sigma] \approx [10 - 5,77; 10 + 5,77] = [4,23; 15,77]$$

— Le nombre de 6 étant entier, on retient l'intervalle $[5; 15]$.

Interprétation. Obtenir entre 5 et 15 fois le 6 sur 60 lancers est tout à fait ordinaire. En revanche, obtenir 25 fois le 6 serait **très** inhabituel : cela ferait plus de cinq écarts types, et conduirait à soupçonner un dé truqué.

Contrôle de bon sens. Un sixième de 60 vaut bien 10 ✓

Réponse. a) $E(X) = 10$ et $\sigma(X) \approx 2,89$ — b) environ $[5; 15]$.

Correction de l'exercice 3 - Retrouver le nombre de répétitions

On part de la formule de l'espérance :

$$E(X) = np \Leftrightarrow 12 = n \times 0,25$$

— **Résolution** :

$$n = \frac{12}{0,25} = 48$$

— **Écart type** :

$$V(X) = 48 \times 0,25 \times 0,75 = 9 \quad \text{donc} \quad \sigma(X) = 3$$

Vérification : $48 \times 0,25 = 12$ ✓

Interprétation. Sur 48 répétitions, on attend 12 succès avec une fluctuation typique de 3 : cela signifie qu'observer 9 ou 15 succès n'a rien d'anormal.

Méthode générale. Les deux formules forment un système : connaître $E(X)$ et p donne n ; connaître $E(X)$ et $V(X)$ donne les deux paramètres, puisque

$$\frac{V(X)}{E(X)} = \frac{np(1-p)}{np} = 1 - p$$

Réponse. $n = 48$ et $\sigma(X) = 3$.

Correction de l'exercice 4 - Retrouver les deux paramètres

Le rapport élimine n :

$$\frac{V(X)}{E(X)} = \frac{np(1-p)}{np} = 1 - p$$

— **Calcul** :

$$1 - p = \frac{14,4}{24} = 0,6$$

- Donc $p = 0,4$.
- **Retour à l'espérance** : $np = 24$ avec $p = 0,4$ donne

$$n = \frac{24}{0,4} = 60$$

- **Vérification** : $E(X) = 60 \times 0,4 = 24 \checkmark$ et $V(X) = 60 \times 0,4 \times 0,6 = 14,4 \checkmark$
 $X \sim B(60; 0,4)$

La technique à retenir. Diviser la variance par l'espérance fait disparaître n et donne directement $1 - p$: c'est le raccourci attendu dans ce type de question, bien plus rapide qu'un système à deux inconnues résolu par substitution.

Réponse. $n = 60$ et $p = 0,4$.

Correction de l'exercice 5 - La fréquence de succès

- a) **Pour le nombre** :

$$E(X) = 400 \times 0,35 = 140 \quad V(X) = 400 \times 0,35 \times 0,65 = 91$$

- $\sigma(X) = \sqrt{91} \approx 9,54$.
- b) **Pour la fréquence**, on applique les formules de transformation avec $a = \frac{1}{400}$:

$$E(F) = \frac{E(X)}{400} = \frac{140}{400} = 0,35 = p$$

$$\sigma(F) = \frac{\sigma(X)}{400} = \frac{9,54}{400} \approx 0,0239$$

Interprétation. La fréquence observée vaut en moyenne exactement p — le sondage est **sans biais** —, avec une fluctuation typique de 2,4 points de pourcentage.

Ce que dit la formule. Comme $\sigma(F) = \sqrt{\frac{p(1-p)}{n}}$, il faut **quadrupler** la taille de l'échantillon pour diviser la marge d'erreur par deux. C'est la raison économique pour laquelle les sondages s'arrêtent souvent à 1000 personnes.

Réponse. a) 140 et $\approx 9,54$ — b) $E(F) = 0,35$ et $\sigma(F) \approx 0,024$.

Correction de l'exercice 6 - Quand la variance est-elle maximale ?

- a) **Formule** :

$$V(X) = 100 p(1 - p) = -100p^2 + 100p$$

- b) C'est un **trinôme en p** , de coefficient dominant $-100 < 0$: il admet un maximum au sommet.

$$\alpha = -\frac{100}{2 \times (-100)} = 0,5$$

- **Valeur maximale** :

$$V(X) = 100 \times 0,5 \times 0,5 = 25$$

- L'écart type vaut alors $\sigma(X) = 5$.

Interprétation. L'incertitude est **maximale** quand les deux issues sont équiprobables : c'est le cas d'une pièce équilibrée, où l'on ne peut rien prévoir.

À l'inverse, pour $p = 0,01$, on obtient $V(X) = 0,99$ et $\sigma(X) \approx 1$: le résultat est presque certain (très peu de succès), donc peu dispersé.

Une conséquence pratique en sondage. Comme la variance est maximale en $p = 0,5$, les instituts calculent leur marge d'erreur avec cette valeur : elle donne le **pire cas**, donc

une marge garantie.

Réponse. a) $V(X) = 100p(1 - p)$ — b) maximale pour $p = 0,5$, avec $V(X) = 25$.

Correction de l'exercice 7 - Le stand de tir revisité

Le nombre X de cibles touchées suit $B(40; 0,75)$.

a) **Espérance :**

$$E(X) = 40 \times 0,75 = 30$$

b) **Variance et écart type :**

$$V(X) = 40 \times 0,75 \times 0,25 = 7,5 \quad \sigma(X) = \sqrt{7,5} \approx 2,74$$

c) **Écart à la moyenne :** $30 - 22 = 8$, soit

$$\frac{8}{2,74} \approx 2,9 \text{ écarts types}$$

Conclusion. Un résultat situé à près de **trois écarts types** en dessous de la moyenne est **surprenant** : il arrive moins d'une fois sur cent.

Que peut-on en déduire ? Soit le tireur a eu une très mauvaise série (le hasard le permet), soit son taux de réussite réel n'est pas de 75 % — fatigue, matériel, conditions. Une seule observation ne permet pas de trancher, mais elle justifie d'y regarder de plus près.

La règle des deux écarts types. Au-delà de 2σ , on considère l'écart comme significatif ; c'est le principe des intervalles de fluctuation.

Réponse. a) 30 — b) $\approx 2,74$ — c) oui, environ 2,9 écarts types en dessous.

Correction de l'exercice 8 - Comparer deux tailles d'échantillon

a) **Pour** $n = 100$:

$$E(X) = 50 \quad \sigma(X) = \sqrt{100 \times 0,25} = 5$$

— **Pour** $n = 10\,000$:

$$E(X) = 5000 \quad \sigma(X) = \sqrt{10\,000 \times 0,25} = 50$$

b) **Fluctuation relative :**

$$\frac{\sigma}{E} = \frac{5}{50} = 0,1 \quad \text{contre} \quad \frac{50}{5000} = 0,01$$

Le résultat essentiel. En multipliant l'échantillon par 100, l'écart type est multiplié par 10 seulement ($\sqrt{100}$), donc la fluctuation **relative** est divisée par 10.

Traduction en fréquences. L'écart type de la fréquence passe de $\frac{5}{100} = 0,05$ à $\frac{50}{10\,000} = 0,005$: de 5 points à un demi-point.

C'est la loi des grands nombres. La fréquence observée se rapproche de p quand n grandit — mais lentement, en $\frac{1}{\sqrt{n}}$: c'est pourquoi les sondages gardent toujours une marge d'erreur de quelques points.

Réponse. a) 50 et 5 ; 5000 et 50 — b) fluctuation relative divisée par 10.

Correction de l'exercice 9 - La démonstration des formules

a) Avec les coefficients 1, 3, 3, 1 et $q = 1 - p$:

k	0	1	2	3
$P(X = k)$	q^3	$3pq^2$	$3p^2q$	p^3

— **Contrôle** : la somme vaut $q^3 + 3pq^2 + 3p^2q + p^3 = (p + q)^3 = 1$ ✓ — on reconnaît le développement du cube d'une somme.

— b) **Espérance** :

$$E(X) = 0 \times q^3 + 1 \times 3pq^2 + 2 \times 3p^2q + 3 \times p^3$$

— On factorise par $3p$:

$$E(X) = 3p(q^2 + 2pq + p^2) = 3p(p + q)^2$$

— Or $p + q = 1$, donc

$$E(X) = 3p$$

Ce que la démonstration montre. La formule $E(X) = np$ n'est pas une recette : elle découle de l'identité remarquable $(p + q)^n = 1$, qui organise toute la loi binomiale.

Vérification numérique avec $p = 0,4$: $E(X) = 1,2$, et le calcul direct donne $3 \times 0,288 + 2 \times 0,288 + 3 \times 0,064 \dots$ soit $0,432 + 0,576 + 0,192 = 1,2$ ✓

Réponse. b) $E(X) = 3p(p + q)^2 = 3p$.

Correction de l'exercice 10 - Le budget de l'assureur

X suit $B(2000; 0,03)$.

a) **Espérance** :

$$E(X) = 2000 \times 0,03 = 60 \text{ sinistres}$$

— **Variance** : $V(X) = 2000 \times 0,03 \times 0,97 = 58,2$, donc

$$\sigma(X) = \sqrt{58,2} \approx 7,63$$

b) **Seuil à deux écarts types** :

$$60 + 2 \times 7,63 \approx 75,3$$

— On retient 76 sinistres, soit un budget de

$$76 \times 1500 = 114\,000 \text{ €}$$

Interprétation. Le coût moyen n'est que de $60 \times 1500 = 90\,000$ €, mais provisionner cette somme exposerait l'assureur à un dépassement une année sur deux. La marge de 24 000 € couvre les fluctuations ordinaires.

Le principe de la mutualisation. Avec 2000 clients, la fluctuation relative vaut $\frac{7,63}{60} \approx 13\%$; avec 200 000 clients, elle tomberait à 1,3 %. Plus l'assureur a de clients, plus il peut prévoir précisément — et réduire ses marges.

Réponse. a) $E(X) = 60$ et $\sigma(X) \approx 7,63$ — b) environ 114 000 €.

Correction de l'exercice 11 - Un résultat est-il plausible ?

Sous l'hypothèse du fabricant, le nombre X de pièces conformes suit $B(200; 0,9)$.

— **Espérance** :

$$E(X) = 200 \times 0,9 = 180$$

— **Écart type** :

$$\sigma(X) = \sqrt{200 \times 0,9 \times 0,1} = \sqrt{18} \approx 4,24$$

— **Écart observé** : $180 - 169 = 11$, soit

$$\frac{11}{4,24} \approx 2,6 \text{ écarts types}$$

Conclusion. L'observation se situe à plus de deux écarts types **en dessous** de ce que prévoit le modèle : elle est **peu compatible** avec l'affirmation du fabricant. Le client est fondé à demander une vérification.

Prudence de rédaction. On ne « démontre » pas que le fabricant ment : on constate que, si son taux était vraiment de 90 %, un tel échantillon serait rare (moins de 1 % des cas). Cela **fragilise** l'affirmation sans la réfuter avec certitude.

Le seuil usuel. Dans l'intervalle $[E - 2\sigma; E + 2\sigma] \approx [171,5; 188,5]$, on accepte l'hypothèse ; 169 en sort ✓

Réponse. Non : 169 est à environ 2,6 écarts types sous l'espérance 180.

Correction de l'exercice 12 - Espérance d'un gain binomial

Le nombre X de piles suit $B(5; 0,5)$, et le gain algébrique vaut

$$G = 3X - 8$$

— **Espérance de X** : $E(X) = 5 \times 0,5 = 2,5$.

— **Par linéarité** :

$$E(G) = 3E(X) - 8 = 3 \times 2,5 - 8 = 7,5 - 8 = -0,5$$

Le joueur perd en moyenne 0,50 € par partie : le jeu est **défavorable**.

Écart type du gain. $V(X) = 5 \times 0,25 = 1,25$, donc $\sigma(X) \approx 1,12$ et

$$\sigma(G) = 3 \times \sigma(X) \approx 3,35 \text{ €}$$

Lecture. La perte moyenne (0,50 €) est **six fois plus petite** que la fluctuation typique (3,35 €) : sur quelques parties, le joueur peut très bien gagner. C'est sur le long terme que le désavantage se manifeste.

Quelle mise rendrait le jeu équitable ? Il faudrait $3 \times 2,5 = 7,50$ € : cinquante centimes de moins.

Réponse. $E(G) = -0,5$ € (et $\sigma(G) \approx 3,35$ €).

Correction de l'exercice 13 - Deux productions comparées

a) **Atelier A** : $X_A \sim B(500; 0,02)$.

$$E(X_A) = 10 \quad \sigma(X_A) = \sqrt{500 \times 0,02 \times 0,98} = \sqrt{9,8} \approx 3,13$$

— **Atelier B** : $X_B \sim B(2000; 0,01)$.

$$E(X_B) = 20 \quad \sigma(X_B) = \sqrt{2000 \times 0,01 \times 0,99} = \sqrt{19,8} \approx 4,45$$

— b) **En nombre absolu**, l'atelier B produit **deux fois plus** de pièces défectueuses (20 contre 10) — mais il produit aussi quatre fois plus de pièces.

— **Fluctuation relative** :

$$\frac{3,13}{10} \approx 0,31 \quad \frac{4,45}{20} \approx 0,22$$

— L'atelier B est donc **plus régulier** : son nombre de défauts varie proportionnellement moins.

La leçon. Comparer deux productions de tailles différentes exige de raisonner en **taux** et

en fluctuation **relative**, jamais en effectifs bruts.

Contrôle des taux : $\frac{10}{500} = 2\%$ et $\frac{20}{2000} = 1\%$ ✓ conformes aux données.

Réponse. a) 10 et $\approx 3,13$; 20 et $\approx 4,45$ — b) B en nombre, mais B est aussi le plus régulier.

Correction de l'exercice 14 - Choisir la taille de l'échantillon

Écart type de la fréquence :

$$\sigma(F) = \sqrt{\frac{p(1-p)}{n}}$$

— Avec $p = 0,5$, le numérateur vaut 0,25 :

$$\sigma(F) = \frac{0,5}{\sqrt{n}}$$

— **Condition** : $\frac{0,5}{\sqrt{n}} < 0,01$, soit $\sqrt{n} > 50$.

$$n > 2500$$

— Il faut donc interroger **au moins 2501 personnes**, en pratique 2500 suffit largement puisque l'inégalité est presque atteinte.

Vérification : pour $n = 2500$, $\sigma(F) = \frac{0,5}{50} = 0,01$ exactement ✓

Pourquoi $p = 0,5$? C'est la valeur qui **maximise** $p(1-p)$: en l'utilisant, on garantit la précision quelle que soit la vraie valeur de p . C'est le calcul de « pire cas » des instituts de sondage.

Ordre de grandeur à retenir. Pour une marge de deux écarts types (soit ± 2 points), $n = 2500$; pour ± 3 points, $n \approx 1100$ — la taille habituelle des sondages d'opinion.

Réponse. $n > 2500$ personnes.

Correction de l'exercice 15 - Synthèse : la chaîne d'embouteillage

a) Chaque bouteille est mal bouchée ou non, avec la même probabilité, indépendamment des autres :

$$X \sim B(1200; 0,015)$$

— b) **Espérance** : $E(X) = 1200 \times 0,015 = 18$.

— **Variance** : $V(X) = 1200 \times 0,015 \times 0,985 = 17,73$, donc

$$\sigma(X) = \sqrt{17,73} \approx 4,21$$

c) **Écart** : $30 - 18 = 12$, soit

$$\frac{12}{4,21} \approx 2,85 \text{ écarts types}$$

— Au-delà de deux écarts types : l'écart est **significatif**, l'inquiétude est justifiée. Un réglage a probablement dérivé.

— d) **Avec $p = 0,01$** : $E(X) = 12$ et $\sigma(X) = \sqrt{1200 \times 0,01 \times 0,99} \approx 3,45$.

— On gagnerait 6 bouteilles mal bouchées par heure, soit environ 48 par journée de 8 heures.

Le seuil d'alerte adapté. Avec le nouveau taux, le seuil à deux écarts types tomberait à $12 + 2 \times 3,45 \approx 19$: l'alerte se déclencherait bien plus tôt, ce qui est le but d'un contrôle plus exigeant.

Réponse. a) $B(1200 ; 0,015)$ — b) 18 et $\approx 4,21$ — c) oui, $\approx 2,85\sigma$ — d) 6 bouteilles de moins par heure.