COURS

Sommes de variables — Espérance, variance, échantillons

Tle

Programme de terminale générale, enseignement de spécialité — BO spécial n°8 du 25 juillet 2019

Lancer cent dés et prévoir leur somme : sa loi est hors de portée d'un calcul à la main, et pourtant son espérance et sa variance s'obtiennent en une ligne. **L'espérance d'une somme est toujours la somme des espérances ; la variance d'une somme n'est la somme des variances que si les variables sont indépendantes. Pour la moyenne M_n d'un échantillon de taille n, il en découle $E(M_n) = \mu$ et $\sigma(M_n) = \dfrac{\sigma}{\sqrt{n}}$.**

1 Somme de variables aléatoires et linéarité de l'espérance

Rappels de première. Sur un univers fini Ω, une variable aléatoire X est une **fonction** de Ω dans $\mathbb{R}$; $E(X) = \sum p_i\, x_i = \sum_{\omega\in\Omega} P(\{\omega\})\, X(\omega)$ (**moyenne pondérée sur les issues**), $V(X) = E(X^2) - \big(E(X)\big)^2$ et $\sigma(X) = \sqrt{V(X)}$.

DÉFINITION

Soit X et Y deux variables aléatoires définies sur le **même univers** Ω, et a un réel. La variable $X + Y$ associe à chaque issue ω le nombre $X(\omega) + Y(\omega)$; la variable aX lui associe $a\,X(\omega)$. Exemple : pour deux dés, $X_1 + X_2$ est la somme des faces ; un gain en euros G devient $100G$ en centimes.

PROPRIÉTÉ

Linéarité de l'espérance (démontrée). Pour toutes variables aléatoires X et Y définies sur Ω, **indépendantes ou non**, et pour tout réel a : $E(X+Y) = E(X) + E(Y)$ et $E(aX) = aE(X)$. *Preuve* : avec la moyenne pondérée sur les issues, $E(X+Y) = \sum_\omega P(\{\omega\})\big(X(\omega)+Y(\omega)\big)$, que l'on sépare en deux sommes : $E(X) + E(Y)$. De même, $\sum_\omega P(\{\omega\})\, aX(\omega) = aE(X)$. De proche en proche, $E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n)$; et comme une constante b a pour espérance b, $E(aX + b) = aE(X) + b$.

$$E(X+Y) = E(X) + E(Y) \qquad E(aX) = aE(X)$$

CALCULER UNE ESPÉRANCE SANS PASSER PAR LA LOI

1. **Décomposer** : écrire X comme une somme de variables plus simples, souvent une par épreuve ou par objet : le résultat de chaque tirage, ou une variable qui vaut 1 si un événement se produit et 0 sinon.
2. **Calculer l'espérance de chaque terme**, à partir de sa loi, qui est simple.
3. **Additionner**, par linéarité, sans se préoccuper de l'indépendance.
   la loi de X n'est jamais déterminée : c'est tout l'intérêt de la méthode.

EXEMPLE

Une urne contient quatre boules numérotées 1, 2, 3 et 4. On tire successivement deux boules **sans remise** ; X_1 et X_2 sont leurs numéros, et $S = X_1 + X_2$. Calculer $E(S)$.

1. X_1 suit la loi uniforme sur $\{1,2,3,4\}$. Un numéro k sort au second tirage s'il ne sort pas au premier, puis sort parmi les trois boules restantes :
   $\frac{3}{4} \times \frac{1}{3} = \frac{1}{4}$
   X_2 a donc la même loi que X_1, et $E(X_1) = E(X_2) = 2{,}5$.
2. Par linéarité, bien que X_1 et X_2 ne soient pas indépendantes :
   $E(S) = 2{,}5 + 2{,}5 = 5$

$E(S) = 5$, sans avoir déterminé la loi de S ; le tableau des douze tirages le confirme.

1re \ 2e boule	1	2	3	4
1	—	3	4	5
2	3	—	5	6
3	4	5	—	7
4	5	6	7	—

12 issues, chacune de probabilité 1/12
somme moyenne : 60/12 = 5

Les 12 tirages ordonnés sont équiprobables ; la diagonale est impossible sans remise.

PIÈGE

« Les tirages sont sans remise, donc X_1 et X_2 ne sont pas indépendantes : on ne peut pas écrire $E(X_1+X_2)=E(X_1)+E(X_2)$. » — la linéarité de l'espérance ne demande **aucune** hypothèse d'indépendance. Seule l'additivité de la variance en demande une.

Ce que ça coûte : la question, ou de longues minutes passées à chercher une loi qui n'était pas demandée.

Le réflexe : espérance : linéarité, toujours ; variance : additivité, seulement pour des variables indépendantes.

2 Variance d'une somme de variables indépendantes

DÉFINITION

Des variables aléatoires $X_1,\dots,X_n$ sont **indépendantes** si, pour toutes valeurs $x_1,\dots,x_n$, $P(X_1=x_1,\dots,X_n=x_n)=P(X_1=x_1)\times\cdots\times P(X_n=x_n)$, la virgule signifiant « et ». Ici, elles proviennent toujours d'une **succession d'épreuves indépendantes** (lancers, tirages avec remise…), chaque X_i ne dépendant que de la i-ième : l'indépendance est une hypothèse du modèle, qu'il suffit de citer.

PROPRIÉTÉ

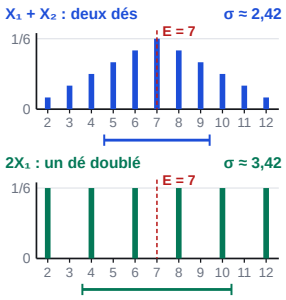
Changement d'échelle (démontré). Pour tous réels a et b, $V(aX+b)=a^2V(X)$. *Preuve* : avec $\mu=E(X)$, la linéarité donne $E(aX+b)=a\mu+b$, donc $V(aX+b)=E\big((aX+b-a\mu-b)^2\big)=E\big(a^2(X-\mu)^2\big)=a^2V(X)$. Ajouter une constante décale les valeurs sans les disperser davantage.

$$V(aX+b)=a^2\,V(X)\qquad \sigma(aX+b)=|a|\,\sigma(X)$$

PROPRIÉTÉ

Additivité de la variance (admise). Plus généralement, si $X_1,\dots,X_n$ sont indépendantes, $V(X_1+\cdots+X_n)=V(X_1)+\cdots+V(X_n)$. Ce sont les **variances** qui s'ajoutent, pas les écarts types. **Sans indépendance, l'additivité tombe** : avec $Y=X$, $V(X+X)=V(2X)=4V(X)$. Pour un dé, $V(X_1)=\dfrac{91}{6}-3{,}5^2=\dfrac{35}{12}$; deux dés ajoutés donnent $V(X_1+X_2)=\dfrac{35}{6}$, un dé doublé $V(2X_1)=\dfrac{35}{3}$, pour la même espérance 7.

$$V(X+Y)=V(X)+V(Y)\quad\text{si } X,Y \text{ indépendantes}$$



Même espérance, dispersions différentes ; sous l'axe, de $E-\sigma$ à $E+\sigma$.

CALCULER UNE VARIANCE PAR DÉCOMPOSITION

1. **Écrire X comme une somme** $X_1+\cdots+X_n$, chaque X_i ne dépendant que d'une épreuve.
2. **Citer le modèle** : les épreuves sont indépendantes, donc les X_i le sont. Sans cette phrase, l'additivité n'est pas justifiée.
3. **Ajouter les variances** (un coefficient sort **au carré**, signe compris), puis **prendre la racine à la fin** pour l'écart type, jamais terme à terme.

EXEMPLE

On lance un dé équilibré, puis une pièce équilibrée. Le score G est le numéro du dé, plus 4 points si la pièce tombe sur pile. Calculer $E(G)$ et $\sigma(G)$.

1. Soit D le numéro du dé et P la variable qui vaut 1 pour pile, 0 sinon : $G=D+4P$. On a $E(D)=3{,}5$, $V(D)=\dfrac{35}{12}$, $E(P)=0{,}5$ et $V(P)=E(P^2)-0{,}5^2=0{,}25$, car $P^2=P$. Par linéarité : $E(G)=3{,}5+4\times0{,}5=5{,}5$
2. D et P dépendent de deux épreuves indépendantes : elles sont indépendantes, et $V(G)=\frac{35}{12}+4^2\times0{,}25=\frac{83}{12}\approx6{,}92$

   d'où $\sigma(G)=\sqrt{\dfrac{83}{12}}\approx2{,}63$.

$E(G)=5{,}5$ points et $\sigma(G)\approx2{,}63$ points.

PIÈGE

Écrire $E(X^2) = \big(E(X)\big)^2$. — la linéarité vaut pour les sommes et la multiplication par un **nombre**, pas pour les carrés : pour un dé, $E(X^2) = \frac{91}{6}$ mais $\big(E(X)\big)^2 = \frac{49}{4}$.

Ce que ça coûte : la variance qui suit, qui vaudrait alors 0.

Le réflexe : calculer $E(X^2) = \sum p_i\, x_i^2$ avec la loi : on élève au carré les valeurs.

PIÈGE

Écrire $V(X-Y) = V(X) - V(Y)$. — $V\big((-1)Y\big) = (-1)^2 V(Y)$: pour X, Y indépendantes, $V(X-Y) = V(X) + V(Y)$. Avec deux dés, la formule fausse donnerait 0.

Ce que ça coûte : le résultat ; une variance nulle ou négative doit alerter.

Le réflexe : un coefficient sort de la variance **au carré**, signe compris.

PIÈGE

Écrire $\sigma(X+Y) = \sigma(X) + \sigma(Y)$, ou $V(2X) = 2V(X)$. — ce sont les variances qui s'ajoutent : avec $\sigma(X) = 3$ et $\sigma(Y) = 4$ indépendantes, $\sigma(X+Y) = 5$, et non 7. Et $V(X+X) = V(2X) = 4V(X)$.

Ce que ça coûte : l'écart type demandé, et tout ce qui en découle.

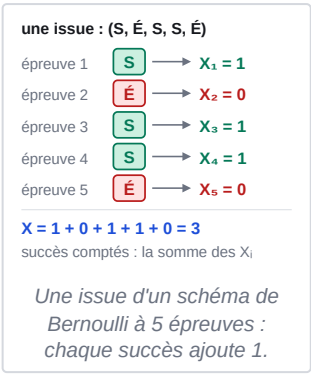
Le réflexe : ajouter les variances, puis prendre une seule racine.

3 La loi binomiale, somme de variables de Bernoulli

PROPRIÉTÉ

Loi de Bernoulli (démontrée). Si Y vaut 1 avec la probabilité p et 0 sinon, $E(Y) = p$; comme $Y^2 = Y$, $E(Y^2) = p$ et $V(Y) = p - p^2 = p(1-p)$. **Loi binomiale (démontrée).** Si X suit $\mathcal{B}(n,p)$, X compte les succès d'un schéma de Bernoulli de n épreuves. Soit X_i la variable qui vaut 1 si la i-ième épreuve est un succès, 0 sinon : chaque X_i suit la loi de Bernoulli de paramètre p, et $X = X_1 + \cdots + X_n$, car chaque succès ajoute 1. Par linéarité, $E(X) = p + \cdots + p = np$. Les épreuves étant indépendantes, les X_i le sont : par additivité, $V(X) = n \times p(1-p)$.

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



Une issue d'un schéma de Bernoulli à 5 épreuves : chaque succès ajoute 1.

EXEMPLE

Un QCM compte 20 questions à quatre réponses, dont une seule juste. Un candidat répond au hasard à chacune, indépendamment ; X est son nombre de bonnes réponses. Une bonne réponse rapporte 1 point, une mauvaise en retire $\frac{1}{3}$. Calculer l'espérance et l'écart type de la note N.

① X suit $\mathcal{B}(20\,;0{,}25)$: $E(X) = 20 \times 0{,}25 = 5$ et $V(X) = 20 \times 0{,}25 \times 0{,}75 = 3{,}75$

② Il y a $20 - X$ mauvaises réponses : $N = X - \dfrac{20-X}{3} = \dfrac{4}{3}X - \dfrac{20}{3}$. Par linéarité, $E(N) = \dfrac{4}{3} \times 5 - \dfrac{20}{3} = 0$, et $V(N) = \left(\dfrac{4}{3}\right)^2 V(X)$:

$\frac{16}{9} \times 3{,}75 = \frac{20}{3} \approx 6{,}67$

d'où $\sigma(N) = \sqrt{\dfrac{20}{3}} \approx 2{,}58$ points.

$E(N) = 0$: ce barème annule en moyenne l'effet du hasard ; $\sigma(N) \approx 2{,}58$ points.

PIÈGE

Utiliser $V(X) = np(1-p)$ pour des tirages sans remise. — les tirages ne sont plus indépendants : $E(X) = np$ reste vrai par linéarité, l'additivité non. Deux boules tirées sans remise parmi 2 rouges et 2 blanches : X vaut 0, 1, 2 avec les probabilités $\frac{1}{6}, \frac{4}{6}, \frac{1}{6}$; $E(X) = 1$ mais $V(X) = \frac{1}{3}$, et non $2 \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{2}$.

Ce que ça coûte : la variance et l'écart type, et la justification de la loi.

Le réflexe : avant d'écrire $\mathcal{B}(n,p)$, vérifier : épreuves identiques, indépendantes, à deux issues.

4 Échantillon d'une loi : somme et moyenne

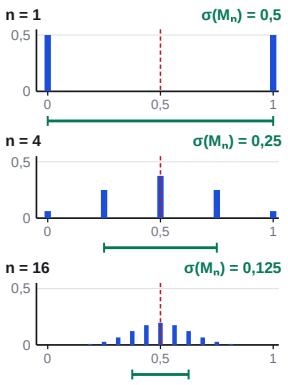
DÉFINITION

Un **échantillon de taille** n d'une loi de probabilité est une liste $(X_1, \ldots, X_n)$ de variables aléatoires **indépendantes** et **identiques**, c'est-à-dire qui suivent toutes cette loi. Il modélise n répétitions indépendantes d'une même expérience : n lancers d'un dé, n mesures, n clients… Sa **somme** est $S_n = X_1 + \cdots + X_n$, sa **moyenne** $M_n = \dfrac{S_n}{n}$.

PROPRIÉTÉ

Somme et moyenne d'un échantillon (démontrées). Si la loi a pour espérance μ, pour variance V et pour écart type $\sigma = \sqrt{V}$, la linéarité donne $E(S_n) = n\mu$ et l'indépendance $V(S_n) = nV$, d'où $\sigma(S_n) = \sqrt{n}\,\sigma$. Puis $M_n = \frac{1}{n}S_n$ donne $E(M_n) = \mu$ et $V(M_n) = \frac{1}{n^2} \times nV$. **Lecture** : la somme se disperse de plus en plus, mais la moyenne se **resserre** autour de μ. Pour diviser $\sigma(M_n)$ par 2, il faut **quadrupler** n. Attention : S_n n'est pas nX_1, qui répète un seul résultat et vérifie $\sigma(nX_1) = n\sigma$. La fiche « Bienaymé-Tchebychev — Concentration, loi des grands nombres » en tire des probabilités.

$$E(M_n) = \mu \qquad V(M_n) = \frac{V}{n} \qquad \sigma(M_n) = \frac{\sigma}{\sqrt{n}}$$



Moyenne de n lancers d'une pièce (1 pour pile) : quadrupler n divise $\sigma(M_n)$ par 2.

ÉTUDIER LA SOMME OU LA MOYENNE D'UN ÉCHANTILLON

1. **Reconnaître l'échantillon** (n répétitions indépendantes d'une même expérience), puis **calculer μ, V et σ pour une seule variable**.
2. **Appliquer les formules** : pour la somme, multiplier μ et V par n ; pour la moyenne, garder μ et diviser V par n. Dans les deux cas, l'écart type varie d'un facteur $\sqrt{n}$.
3. **Pour une précision visée** ε, résoudre $\frac{\sigma}{\sqrt{n}} \leqslant \varepsilon$, soit $n \geqslant \frac{\sigma^2}{\varepsilon^2}$, et prendre le plus petit entier.

EXEMPLE

Une balance donne d'un objet une mesure aléatoire d'espérance μ, sa masse exacte, et d'écart type 0,6 g ; les mesures successives sont indépendantes. On retient la moyenne M_n de n mesures. Calculer $\sigma(M_9)$, puis le nombre de mesures qui garantit $\sigma(M_n) \leqslant 0{,}1$ g.

1. $(X_1, \ldots, X_9)$ est un échantillon : $E(M_9) = \mu$, la moyenne reste centrée sur la masse exacte, et $\sigma(M_9) = \frac{0{,}6}{\sqrt{9}} = 0{,}2$
2. Puis $\frac{0{,}6}{\sqrt{n}} \leqslant 0{,}1 \iff \sqrt{n} \geqslant 6 \iff n \geqslant 36$.

$\sigma(M_9) = 0{,}2$ g ; il faut au moins 36 mesures pour que $\sigma(M_n) \leqslant 0{,}1$ g.

PIÈGE

Écrire $\sigma(M_n) = \frac{\sigma}{n}$, ou $V(M_n) = \frac{V}{n^2}$. — $\frac{1}{n}$ sort de la variance au carré, mais la somme apporte un facteur n : $V(M_n) = \frac{V}{n}$, et la racine donne $\frac{\sigma}{\sqrt{n}}$.

Ce que ça coûte : la valeur, et toute taille d'échantillon qui en découle : 36 mesures deviendraient 6.

Le réflexe : contrôle : quadrupler n divise l'écart type de la moyenne par 2, pas par 4.

À VÉRIFIER

- Ai-je cité l'indépendance des épreuves avant d'ajouter des variances ?
- Un coefficient sort-il de la variance au carré : $V(-Y) = V(Y)$?
- Pour $\mathcal{B}(n, p)$: épreuves identiques, indépendantes, à deux issues ?
- Pour une moyenne : $\sigma(M_n) = \frac{\sigma}{\sqrt{n}}$, et non $\frac{\sigma}{n}$?

À RETENIR

- $E(X+Y) = E(X) + E(Y)$ et $E(aX+b) = aE(X) + b$, **toujours**.
- $V(aX+b) = a^2V(X)$; $V(X+Y) = V(X) + V(Y)$ **si X, Y indépendantes**.
- Bernoulli : $E = p$, $V = p(1-p)$. $\mathcal{B}(n,p)$: $E = np$, $V = np(1-p)$.
- Échantillon : $E(S_n) = n\mu$, $V(S_n) = nV$; $E(M_n) = \mu$, $\sigma(M_n) = \frac{\sigma}{\sqrt{n}}$.
- Les variances s'ajoutent (si indépendance), jamais les écarts types.