Plaquette 3 – Variance, covariance et corrélation

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

Exercices

Méthode

La variance mesure la dispersion autour de la moyenne ; on la calcule le plus souvent par **König-Huygens** $\mathrm{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$. Pour une **somme**, on développe : $\mathrm{V}\left(\sum X_i\right) = \sum \mathrm{V}(X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j)$, et les covariances disparaissent si les variables sont indépendantes (deux à deux suffit). Pour une variable de comptage $\sum \mathbf{1}_{A_i}$, les covariances se calculent avec $\mathbb{P}(A_i \cap A_j)$. La covariance est une forme bilinéaire symétrique positive : on en tire Cauchy-Schwarz ($|\rho| \leq 1$) et la meilleure approximation affine (régression).

Exercice 1 – Variance d'une loi uniforme

Soit X uniforme sur $[\![1, n]\!]$.

a) Calculer $\mathbb{E}(X^2)$ à l'aide de $\sum_{k=1}^{n} k^2 = \dfrac{n(n+1)(2n+1)}{6}$.

b) En déduire $\mathrm{V}(X) = \dfrac{n^2 - 1}{12}$.

c) Application : écart type d'un dé ; d'un nombre tiré dans $[\![1, 100]\!]$.

Exercice 2 – Variance de la loi binomiale

Soit $X \sim B(n, p)$.

a) Écrire X comme somme de n variables de Bernoulli indépendantes et en déduire $\mathrm{V}(X) = np(1-p)$.

b) Retrouver ce résultat par la fonction génératrice : $\mathrm{V}(X) = G''(1) + G'(1) - G'(1)^2$.

c) Application : 1 000 lancers d'une pièce équilibrée ; espérance et écart type du nombre de Piles.

Exercice 3 – Variance de la loi hypergéométrique

Une urne contient N boules dont m blanches ; on en tire n sans remise. $X = \sum_{i=1}^{m} \mathbf{1}_{A_i}$ où A_i : « la blanche i est tirée ».

a) Calculer $\mathbb{P}(A_i) = \dfrac{n}{N}$ et $\mathbb{P}(A_i \cap A_j) = \dfrac{n(n-1)}{N(N-1)}$ pour $i \neq j$.

b) En déduire $\mathrm{Cov}(\mathbf{1}_{A_i}, \mathbf{1}_{A_j})$, puis $\mathrm{V}(X) = n\dfrac{m}{N}\left(1 - \dfrac{m}{N}\right)\dfrac{N-n}{N-1}$.

c) Comparer avec la variance de la binomiale $B\left(n, \dfrac{m}{N}\right)$ et commenter le cas $n = N$.

Exercice 4 – Variance du nombre de points fixes

Soit F le nombre de points fixes d'une permutation uniforme de $[\![1, n]\!]$ ($n \geq 2$), $F = \sum_i \mathbf{1}_{A_i}$ avec $A_i = \{\sigma(i) = i\}$.

a) Calculer $\mathrm{Cov}(\mathbf{1}_{A_i}, \mathbf{1}_{A_j})$ pour $i \neq j$.

b) En déduire $\mathrm{V}(F) = 1$.

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) Vérifier pour $n = 4$ (loi $\frac{9}{24}, \frac{8}{24}, \frac{6}{24}, 0, \frac{1}{24}$).

Exercice 5 – Covariance à partir d'un tableau

On reprend la loi conjointe :

	$Y=0$	$Y=1$
$X=0$	0,1	0,2
$X=1$	0,2	0,1
$X=2$	0,3	0,1

a) Calculer $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\mathbb{E}(XY)$ et $\mathrm{Cov}(X,Y)$.
b) Calculer $\mathrm{V}(X)$, $\mathrm{V}(Y)$ et le coefficient de corrélation.
c) Calculer $\mathrm{V}(X+Y)$ de deux façons.

Exercice 6 – Le coefficient de corrélation est entre −1 et 1

Soient X, Y de variances non nulles.

a) Montrer que $t \mapsto \mathrm{V}(X+tY)$ est un polynôme de degré 2, positif sur $\mathbb{R}$.
b) En déduire $\mathrm{Cov}(X,Y)^2 \leq \mathrm{V}(X)\mathrm{V}(Y)$, puis $|\rho| \leq 1$.
c) Montrer que $|\rho| = 1$ si et seulement si $Y = aX + b$ presque sûrement (avec $a \neq 0$).
d) Exemple : X résultat d'un dé, $Y = 7 - X$. Calculer ρ.

Exercice 7 – Covariance nulle et indépendance

Soit X uniforme sur $\{-1, 0, 1\}$ et $Y = X^2$.

a) Calculer $\mathrm{Cov}(X,Y)$.
b) X et Y sont-elles indépendantes ?
c) Montrer en revanche que pour des variables de **Bernoulli**, covariance nulle entraîne indépendance.

Exercice 8 – La moyenne minimise l'écart quadratique

Soit X une variable finie.

a) Montrer que $\mathbb{E}((X-a)^2) = \mathrm{V}(X) + (\mathbb{E}(X) - a)^2$ pour tout réel a.
b) En déduire que $a \mapsto \mathbb{E}((X-a)^2)$ est minimale en $a = \mathbb{E}(X)$, de minimum $\mathrm{V}(X)$.
c) Application : un dé ; quelle constante approche le mieux le résultat au sens quadratique ? Quelle erreur moyenne ?

Exercice 9 – Diversifier un portefeuille

Deux placements ont des rendements aléatoires R_1, R_2 de variances $\sigma_1^2 = 0{,}04$ et $\sigma_2^2 = 0{,}09$, et de covariance $c = 0{,}012$. On place une proportion $t \in [0,1]$ sur le premier.

a) Exprimer $v(t) = \mathrm{V}(tR_1 + (1-t)R_2)$.
b) Déterminer t^* qui minimise v, et $v(t^*)$.

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) Comparer l'écart type obtenu avec celui du placement le moins risqué. Commenter.

Exercice 10 – Droite de régression

Pour le couple (X, Y) de l'exercice « covariance à partir d'un tableau » ($\mathbb{E}(X) = 1{,}1$, $\mathbb{E}(Y) = 0{,}4$, $\mathrm{V}(X) = 0{,}69$, $\mathrm{Cov}(X, Y) = -0{,}14$) :

a) Montrer que $\mathbb{E}((Y - aX - b)^2)$ est minimal pour $a = \dfrac{\mathrm{Cov}(X, Y)}{\mathrm{V}(X)}$ et $b = \mathbb{E}(Y) - a\mathbb{E}(X)$.

b) Calculer a et b.

c) Montrer que l'erreur minimale vaut $\mathrm{V}(Y)(1 - \rho^2)$ et la calculer.

Exercice 11 – Variance nulle

a) Montrer que $\mathrm{V}(X) = 0$ si et seulement si X est presque sûrement constante ($\mathbb{P}(X = c) = 1$).

b) En déduire que si $\mathrm{V}(X + Y) = 0$ avec X, Y indépendantes, alors X et Y sont presque sûrement constantes.

c) Construire X et Y non constantes (non indépendantes) telles que $\mathrm{V}(X + Y) = 0$.

Exercice 12 – Effectifs de deux faces : covariance négative

On lance n dés ; N_1 est le nombre de 1, N_6 le nombre de 6.

a) Écrire $N_1 = \sum_i \mathbf{1}_{\{D_i = 1\}}$, $N_6 = \sum_i \mathbf{1}_{\{D_i = 6\}}$ et calculer $\mathrm{Cov}(\mathbf{1}_{\{D_i = 1\}}, \mathbf{1}_{\{D_j = 6\}})$ selon que $i = j$ ou $i \neq j$.

b) En déduire $\mathrm{Cov}(N_1, N_6) = -\dfrac{n}{36}$ et le coefficient de corrélation.

c) Interpréter le signe.

Exercice 13 – Variance du nombre de boîtes vides

On lance 3 boules dans 3 boîtes ; $V = \sum_{j=1}^{3} \mathbf{1}_{A_j}$ est le nombre de boîtes vides, avec $\mathbb{P}(A_j) = \dfrac{8}{27}$.

a) Calculer $\mathbb{P}(A_i \cap A_j)$ pour $i \neq j$ et la covariance correspondante.

b) En déduire $\mathrm{V}(V)$.

c) Vérifier avec la loi $\dfrac{6}{27}, \dfrac{18}{27}, \dfrac{3}{27}$.

Exercice 14 – Variable centrée réduite

Soit X de variance non nulle, $m = \mathbb{E}(X)$, $\sigma = \sigma(X)$, et $Z = \dfrac{X - m}{\sigma}$.

a) Montrer que $\mathbb{E}(Z) = 0$ et $\mathrm{V}(Z) = 1$.

b) Pour $X \sim B(100, 0{,}5)$, calculer Z lorsque $X = 60$.

c) Deux élèves ont eu 14 dans une classe où la moyenne est 11 et l'écart type 2, et 15 dans une classe de moyenne 12 et d'écart type 4. Lequel est le « meilleur » relativement à sa classe ?

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).

Exercice 15 – Synthèse : variance du nombre de records

Dans une permutation uniforme de n valeurs, les indicatrices de records $I_i \sim B\left(\frac{1}{i}\right)$ sont (on l'admet) indépendantes. $R = \sum_i I_i$.

a) Montrer que $\mathrm{V}(R) = \sum_{i=1}^{n} \frac{1}{i}\left(1 - \frac{1}{i}\right) = H_n - \sum_{i=1}^{n} \frac{1}{i^2}$.

b) Calculer pour $n = 4$ et vérifier avec la loi $\frac{6}{24}, \frac{11}{24}, \frac{6}{24}, \frac{1}{24}$.

c) Montrer que $\mathrm{V}(R) \sim \ln n$ quand $n \to +\infty$ (admettre $H_n \sim \ln n$ et que $\sum \frac{1}{i^2}$ est bornée).

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).