Moments d'une variable aléatoire

Prof. Mohamed El Merouani

Université Abdelmalek Essaâdi
Faculté des Sciences de Tétouan
Département de Mathématiques

2019/2020

- Une v.a. peut être caractérisée par certaines valeurs typiques associées aux notions de valeur centrale, de dispersion et de forme de la distribution.

Définition :

Soit X une v.a. discrète de loi de probabilité $P(X = x_i) = p_i$, $i = 1, 2, \cdots$
Si $\sum_{i=1}^{\infty} |x_i| p_i < \infty$, on définit l'espérance mathématique de X par :

$$E(X) = \sum_{i=1}^{\infty} x_i P(X = x_i)$$

Remarque :

- Lorsque $X(\Omega) = \{x_1, x_2, \cdots, x_n\}$ est fini, cette somme est finie et on a :

$$E(X) = \sum_{i=1}^{n} x_i P(X = x_i)$$

- Si $X(\Omega)$ est infini, on a la somme d'une série qui peut ne pas exister.

Exemple :

Soit X une v.a. discrète de loi de probabilité définie par :
$p_i = P(X = x_i = (-1)^{i+1}\frac{3^i}{i}) = \frac{2}{3^i}$, pour $i = 1, 2, 3, \cdots$
Puisque $\sum_{i=1}^{\infty} |x_i| p_i = \sum_{i=1}^{\infty} \frac{3^i}{i} . \frac{2}{3^i} = \sum_{i=1}^{\infty} \frac{2}{i} = \infty$, l'espérance mathématique de X n'existe pas.

Définition :

Soit X une v.a. continue de densité f telle que $\int_{-\infty}^{+\infty} |x| f(x)dx < \infty$. On définit l'espérance mathématique de X par :

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_{-\infty}^{+\infty} xdF(x)$$

Exemple :

Soit X une v.a. continue de densité f définie par : $f(x) = \frac{1}{\pi}\frac{1}{1+x^2}$; $x \in \mathbb{R}$ (Loi de Cauchy).
L'intégrale $\frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{|x|}{1+x^2}dx$ n'existe pas ; cet intégrale diverge. En effet, en posant $t = 1 + x^2$, on obtient $\int_0^{\infty} \frac{x}{1+x^2}dx = \frac{1}{2}\int_1^{\infty} \frac{dt}{t} = \frac{1}{2}\ln t]_1^{\infty}$.
Comme ce nombre est infini, l'espérance mathématique de X n'existe pas.

Espérance d'une fonction d'une v.a. :

Théorème de transfert :

Soit X une v.a. et g une fonction Borel-mesurable, et soit Y=g(X). Alors $E(g(X)) = E(Y)$.

- $E(Y) = \sum_{i=1}^{\infty} g(x_i) P(X = x_i) = \sum_{j=1}^{\infty} y_j P(Y = y_j)$
- $E(Y) = \int_{-\infty}^{+\infty} g(x) f(x) dx = \int_{-\infty}^{+\infty} y f(g^{-1}(y))(g^{-1}(y))' dy$

Exemples :

1) Soit X une v.a. discrète de loi :

x_i	-3	0	3
$P(X = x_i)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

En posant $Y = X^2$, d'après le théorème précédent, on a :

Exemples :

$E(Y) = \sum_i g(x_i)P(X = x_i) = (-3)^2.\frac{1}{4} + 0^2.\frac{1}{2} + (3)^2\frac{1}{4} = \frac{9}{2}$. Mais, on peut aussi écrire la loi de Y :

y_j	0	9
$P(X = x_i)$	$\frac{1}{2}$	$\frac{1}{2}$

d'où $E(Y) = 0.\frac{1}{2} + 9.\frac{1}{2} = \frac{9}{2}$. Dans cet exemple, on a $g^{-1}(0) = 0; g^{-1}(9) = \{-3, 3\}$ et $E(Y) = 0.\frac{1}{2} + 9(\frac{1}{4} + \frac{1}{4}) = \frac{9}{2}$

2) Soit X une v.a. de densité de probabilité :
$f(x) = \begin{cases} \frac{3}{2}(1 - x^2) & \text{si } 0 \leq x \leq 1 \\ 0 & \text{sinon} \end{cases}$. Cherchons $E(X^2)$. L'application directe du théorème précédent donne :
$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)dx = \frac{3}{2}\int_0^1 x^2(1 - x^2)dx = \frac{1}{5}$

Appelation :

Si $E(X) = 0$, la v.a. X est dite centrée.

Propriétés :

Soient X et Y deux v.a. discrètes (ou continues) et a une constante, on a :

1. $E(a) = a$
2. $E(X + a) = E(X) + a$
3. $E(aX) = aE(X)$
4. $E[X - E(X)] = 0$
5. $E(X + Y) = E(X) + E(Y)$

Preuve :

1. Soit $X = a$ et $P(X = a) = 1$, alors $E(X) = a.P(X = a) = a.1 = a$

Preuve :

2. $$E(X+a)=\sum_i (x_i+a)P(X=x_i)$$
$$=\sum_i x_iP(X=x_i)+a\sum_i P(X=x_i)$$
$$=E(X)+a$$
3. $$E(aX)=\sum_i ax_iP(X=x_i)$$
$$=a\sum_i x_iP(X=x_i)$$
$$=aE(X)$$
4. $$E[X-E(X)]=\sum_i (x_i-E(X))P(X=x_i)$$
$$=\sum_i x_iP(X=x_i)-\sum_i E(X)P(X=x_i)$$
$$=E(X)-E(X)\sum_i P(X=x_i)$$
$$=E(X)-E(X)=0$$

Preuve :

5. $E(X+Y) = \sum_i \sum_j (x_i + y_j) P(X = x_i, Y = y_j)$
$= \sum_i x_i \left(\sum_j P(X = x_i, Y = y_j) \right) + \sum_j y_j \left(\sum_i P(X = x_i, Y = y_j) \right)$
$= \sum_i x_i P(X = x_i) + \sum_j y_j P(Y = y_j)$
$= E(X) + E(Y)$

Remarque :

Des démonstrations analogues peuvent se faire dans le cas de v.a. continues. Ces propriétés sont valables aussi pour des v.a. continues.

Soient X et Y deux v.a. indépendantes, alors, on a :

$$E(XY) = E(X).E(Y)$$

En effet :

$$E(XY) = \sum_i \sum_j x_i y_j P(X = x_i, Y = y_j)$$

$$= \sum_i \sum_j x_i y_j P(X = x_i).P(Y = y_j)$$

$$= \sum_i \left(x_i P(X = x_i) \sum_j y_j P(Y = y_j) \right)$$

$$= \sum_i x_i P(X = x_i) E(Y)$$

$$= E(X).E(Y)$$

Cette propriété est vraie aussi dans le cas continue.

Espérance d'un produit et indépendance de deux v.a. :

La réciproque est fausse. $E(XY) = E(X).E(Y)$ n'implique pas en général l'indépendance de X et Y.

Contre-exemple :

Soit la distribution suivante :

$Y \backslash X$	-3	0	3	$P(Y=j)$
-1	0	$\frac{1}{4}$	0	$\frac{1}{4}$
0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
1	0	$\frac{1}{4}$	0	$\frac{1}{4}$
$P(X=i)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

On a $E(X) = 0\,;\ E(Y) = 0$ et $E(XY) = 0$. Mais l'égalité $E(XY) = E(X)E(Y)$ n'entraîne pas l'indépendance de X et Y.
Il suffit de vérifier, par exemple, que
$P(X = 0, Y = 1) = \frac{1}{4} \neq P(X = 0)P(Y = 1) = \frac{1}{2}.\frac{1}{4} = \frac{1}{8}$.

Variance et écart-type :

On définit une mesure de dispersion de X autour de son espérance mathématique dite variance de X.

Définition :

Si $E(X^2)$ existe, on appelle variance d'une v.a. X, le nombre $Var(X) = E\left((X - E(X))^2\right)$. On appelle écart-type de X, le nombre $\sigma(X) = \sqrt{Var(X)}$.

Formule réduite :

$$Var(X) = E(X^2) - E(X)^2$$

En effet,

$$\begin{aligned} Var(X) &= E\left(X^2 - 2XE(X) + E(X)^2\right) \\ &= E(X^2) - 2E(X)^2 + E(X)^2 \\ &= E(X^2) - E(X)^2 \end{aligned}$$

Variance et écart-type :

Propriétés :

$\forall a, b \in \mathbb{R}$, on a :

1. $Var(aX + b) = a^2 Var(X)$
2. $Var(X) = 0 \Leftrightarrow X$ est dégénerée,
   i.e. $X = E(X)$ avec $P(X = E(X)) = 1$
3. $Var(X) < E(X - C)^2; \forall C \neq E(X)$
   car, $E(X - C)^2 = E(X - E(X))^2 + (C - E(X))^2$
   puisque $2(E(X) - C)E(X - E(X)) = 0$
4. Si $E(|X|^2) < \infty$, la v.a. $Z = \frac{X - E(X)}{\sigma(X)}$ est telle que $E(Z) = 0$ et $\sigma(Z) = 1$.
   Z est dite v.a. centrée et réduite ($Var(Z) = \frac{Var(X)}{\sigma^2(X)} = 1$) associée à la v.a. X.

Inégalité de Bienaymé-Tchebychev :

Faute de connaître une probabilité exacte, il suffit parfois de trouver une borne supérieure ou inférieure à cette probabilité.
Le théorème suivant qui lie l'espérance mathématique et l'écart-type répond à ce genre de questions.

Théorème :

Soit X une v.a. telle que $E(X)$ et $Var(X)$ existent. Pour tout ε réel $(\varepsilon > 0)$ on a :

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon}$$

Remarque :

En posant $\varepsilon = t\sigma$, on obtient une autre variante du théorème :

$$P(|X - E(X)| \geq t\varepsilon) \leq \frac{1}{t^2}$$

et $$P(|X - E(X)| < t\varepsilon) \geq 1 - \frac{1}{t^2}$$

Preuve :

• **Si X v.a. discrète :**

$Var(X) = E(X - E(X))^2 = \sum_i (x_i - E(X))^2 P(X = x_i)$

$= \sum_{i \in I} (x_i - E(X))^2 P(X = x_i) + \sum_{i \in J} (x_i - E(X))^2 P(X = x_i)$

où $I = \{i \in \mathbb{N} / |x_i - E(X)| \geq \varepsilon\}$ et $J = I^c$

On obtient $Var(X) \geq \sum_{i \in I} (x_i - E(X))^2 P(X = x_i)$

Puisque, on a $|x_i - E(X)| \geq \varepsilon$, on peut écrire :

$Var(X) \geq \sum_{i \in I} (x_i - E(X))^2 P(X = x_i) \geq \varepsilon^2 \sum_{i \in I} P(X = x_i)$

Mais, comme $\sum_{i \in I} P(X = x_i) = P\left(\bigcup_{i \in I} (X = x_i)\right) = P(|X - E(X)| \geq \varepsilon)$

On obtient finalement $Var(X) \geq \varepsilon^2 P(|X - E(X)| \geq \varepsilon)$

d'où le résultat.

Preuve :

• **Si X v.a. absolument continue de densité f :**

Soit $\varepsilon > 0$, on a : $Var(X) = \int_{-\infty}^{+\infty}(x - E(X))^2 f(x)dx$
$= \int_{-\infty}^{E(x)-\varepsilon}(x - E(X))^2 f(x)dx + \int_{E(x)-\varepsilon}^{E(x)+\varepsilon}(x - E(X))^2 f(x)dx +$
$\int_{E(x)+\varepsilon}^{+\infty}(x - E(X))^2 f(x)dx$

Ces trois intégrales sont positives ou nulles, d'où :
$Var(X) \geq \int_{-\infty}^{E(x)-\varepsilon}(x - E(X))^2 f(x)dx + \int_{E(x)+\varepsilon}^{+\infty}(x - E(X))^2 f(x)dx$
sur l'intervalle $]-\infty, E(X) - \varepsilon]$, on a $x \leq E(X) - \varepsilon$ et sur l'intervalle $[E(X) + \varepsilon, +\infty[$, on a $E(X) + \varepsilon \leq x$. On a donc $x - E(X) \leq -\varepsilon$ et $x - E(X) \geq \varepsilon$ c'est-à-dire $|x - E(X)| \leq \varepsilon$ ou $(x - E(X))^2 \leq \varepsilon^2$

On obtient, alors, $Var(X) \geq \int_{-\infty}^{E(x)-\varepsilon} \varepsilon^2 f(x)dx + \int_{E(x)+\varepsilon}^{+\infty} \varepsilon^2 f(x)dx$
ou $Var(X) \geq \varepsilon^2 \left(\int_{-\infty}^{E(x)-\varepsilon} f(x)dx + \int_{E(x)+\varepsilon}^{+\infty} f(x)dx\right)$
c'est-à-dire $Var(X) \geq \varepsilon^2 \left(P(X \leq E(X) - \varepsilon) + P(X \geq E(X) + \varepsilon)\right)$
où $Var(X) \geq \varepsilon^2 P(|X - E(X)| \geq \varepsilon)$
ou encore $P(|X - E(X)| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$ ■

Moments d'ordre supérieur :

Définition :

On appelle moment d'ordre k d'une v.a. X, le nombre m_k définie par :

$$m_k = E(X^k) = \begin{cases} \sum_i x_i^k P(X = x_i) & \text{si } X \text{ est discrète} \\ \int_{-\infty}^{+\infty} x^k f(x)dx & \text{si } X \text{ est continue} \end{cases}$$

Définition :

On appelle moment centré d'ordre k d'une v.a. X, le nombre μ_k définie par :

$$\mu_k = E(X - E(X))^k = \begin{cases} \sum_i (x_i - E(X))^k P(X = x_i) & \text{si } X \text{ est discrète} \\ \int_{-\infty}^{+\infty} (x - E(X))^k f(x)dx & \text{si } X \text{ est continue} \end{cases}$$

Remarques :

1. Le moment d'ordre 1, m_1 ou m est l'espérance mathématique $E(X) = m$.
2. Le moment centré d'ordre 2, μ_2 est la variance $\mu_2 = Var(X)$.
3. Comme pour l'espérance et la variance, les moments peuvent parfois ne pas exister (la série ou l'intégrale correspondante diverge).

Définition :

La covariance de deux variables aléatoires X et Y, notée $Cov(X,Y)$, est définie par :

$$Cov(X,Y) = E\left[(X - E(X))(Y - E(Y))\right]$$

ou encore : $$Cov(X,Y) = E(XY) - E(X)E(Y)$$

Cas discrèt :

$$Cov(X,Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j p_{ij} - E(X)E(Y)$$

Cas continue :

$$Cov(X,Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)dxdy - E(X)E(Y)$$

Propriétés :

Soient X, Y et Z trois v.a.

La covariance est une forme bilinéaire symétrique :

1. $Cov(X,Y) = Cov(Y,X)$
2. $Cov(X+Y,Z) = Cov(X,Z) + Cov(Y,Z)$
3. $Cov(\lambda X,Y) = \lambda Cov(X,Y)$

Preuve :

1. $Cov(X,Y) = E\left[(X - E(X))(Y - E(Y))\right] =$
   $= E\left[(Y - E(Y))(X - E(X))\right] = Cov(Y,X)$
2. $Cov(X+Y,Z) = E\left[((X+Y) - E(X+Y))(Z - E(Z))\right]$
   $= E[((X - E(X)) + (Y - E(Y)))(Z - E(Z))]$
   $= E[(X - E(X))(Z - E(Z))] + E[(Y - E(Y))(Z - E(Z))]$
   $= Cov(X,Z) + Cov(Y,Z)$

Preuve :

3. $Cov(\lambda X, Y) = E[(\lambda X - E(\lambda X)(Y - E(Y))]$
$= E[\lambda(X - E(X))(Y - E(Y))] = \lambda E[(X - E(X))(Y - E(Y))]$
$= \lambda Cov(X, Y)$

Remarque :

Pour $X = Y$, on retrouve la variance de X comme covariance de (X, X) :

$$Cov(X, X) = E[(X - E(X))(X - E(X))]$$
$$= E[(X - E(X))^2] = Var(X)$$

Propriété :

$$Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$$

Covariance et coefficient de corrélation :

Propriété :

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

En effet :

$$Var(X+Y) = E[(X+Y) - E(X+Y)]^2$$

$$= E[X + Y - E(X) - E(Y)]^2$$

$$= E[(X - E(X)) + (Y - E(Y))]^2$$

$$= E[(X - E(X))^2 + 2(X - E(X))(Y - E(Y)) + (Y - E(Y))^2]$$

$$= E(X - E(X))^2 + E(Y - E(Y))^2 + 2E((X - E(X))(Y - E(Y)))$$

$$= Var(X) + Var(Y) + 2Cov(X,Y)$$

Coefficient de corrélation :

Définition :

Pour deux variables aléatoires X et Y telles que $Var(X) \neq 0$ et $Var(Y) \neq 0$, le coefficient de corrélation entre X et Y, noté $\rho(X,Y)$, est définie par :

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma(X)\sigma(Y)}$$

Propriétés :

1. On a : $-1 \leq \rho \leq 1$
2. **Inégalité de Schwarz :** On a : $|E(XY)| \leq \sqrt{E(X^2)E(Y^2)}$
   Si X et Y sont deux v.a.r. admettant un moment d'odre 2, la v.a. XY admet une moyenne.

Preuve :

1. En considérant la variable aléatoire $aX+Y$, on a :
$$Var(aX+Y) = Var(aX) + Var(Y) + 2Cov(aX+Y)$$
$$= a^2Var(X) + Var(Y) + 2aCov(X,Y)$$
où $\forall a \in \mathbb{R}, Var(aX+Y) \geq 0$.
La quantité positive $Var(aX+Y)$ est considérée comme un trinôme en a de signe constant. Son discriminant Δ' est négatif ou nul :
$$\Delta' = [Cov(X,Y)]^2 - Var(X)Var(Y) \leq 0$$
On en déduit : $\left|\frac{Cov(X,Y)}{\sigma(X)\sigma(Y)}\right| = |\rho| \leq 1$
2. Soit $a \in \mathbb{R}$ et on considère la v.a. $(aX+Y)$. On a $E(aX+Y)^2 \geq 0$. Comme $E(aX+Y)^2 = E(a^2X^2 + 2aXY + Y^2)$
$$= a^2E(X^2) + 2aE(XY) + E(Y^2)$$

Preuve : (suite)

On considère le trinôme en a de signe constant. Son discriminant Δ' est négatif ou nul :

$$\Delta' = E(XY)^2 - E(X^2)E(Y^2) \leq 0$$

On en déduit $|E(XY)| \leq \sqrt{E(X^2)E(Y^2)}$.