Chapitre 8

Variance, corrélation

8.0.24 Variance d'une variable aléatoire

L'espérance mathématique d'une variable aléatoire ne dit rien sur la distribution des valeurs possibles de cette variable autour cette espérance (valeur moyenne). Pour tenir compte de ces variations, on introduit la notion de variance d'une variable aléatoire par la définition suivante:

$$Var(X) = E[(X-\mu)^2] \quad \text{où} \quad \mu = E[X] \tag{8.1}$$

On peut tout de suite vérifier que:

$$Var(X) = E[X^2] - \Big(E[X]\Big)^2 \tag{8.2}$$

Démonstration:

$$\begin{aligned} Var(X) &= E[(X-\mu)^2] = E[X^2 - 2\mu X + \mu^2] = E[X^2] - E[2\mu X] + E[\mu^2] \\ &= E[X^2] - 2\mu E[X] + \mu^2 = E[X^2] - \mu^2 = E[X^2] - \Big(E[X]\Big)^2 \end{aligned} \tag{8.3}$$

Exemples de calcul de variance

A- Variance d'une variable aléatoire normale de paramètres (μ, σ^2)

On rappelle que $E[X] = \mu$. Donc:

$$\begin{aligned} Var(X) &= E[(X-\mu)^2] \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\ (x-\mu)^2 \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dy\, y^2 \exp\left(-\frac{y^2}{2}\right) \quad \text{avec:} \quad y = \frac{x-\mu}{\sigma} \\ &= \frac{-\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y\ \ d\Big(\exp\left(-\frac{y^2}{2}\right)\Big) \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \Big[-y\ \exp\left(-\frac{y^2}{2}\right)|_{-\infty}^{+\infty} + \int_{-\infty}^{\infty} dy \exp\left(-\frac{y^2}{2}\right)\Big] \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dy \exp\left(-\frac{y^2}{2}\right) \\ &= \sigma^2 \end{aligned} \tag{8.4}$$

B- Variance d'une variable aléatoire binomiale de paramètres (n, p)

On montre que:

$$Var(X) = np\,(1-p)$$

Pour voir celà, calculons d'abord:

$$E[X^2] = \sum_{i=0}^{n} i^2\; p(i) = \sum_{i=0}^{n} i^2\; C_i^n p^i (1-p)^{n-i} \tag{8.5}$$

On utilise ensuite l'identité: $i^2 = i(i-1) + i$.

$$\begin{aligned} E[X^2] &= \sum_{i=0}^{n} i(i-1)\frac{n!}{i!\,(n-i)!}p^i(1-p)^{n-i} + \sum_{i=0}^{n} i\,\frac{n!}{i!\,(n-i)!}p^i(1-p)^{n-i} \\ &= \sum_{i=2}^{n} \frac{n!}{(i-2)!\,(n-i)!}\; p^i(1-p)^{n-i} + E[X] \\ &= n(n-1)p^2 \sum_{i=2}^{n} \frac{(n-2)!}{(i-2)!\,(n-i)!}p^{i-2}(1-p)^{n-i} + E[X] \\ &= n(n-1)p^2\;(p + (1-p))^{n-2} + E[X] \\ &= n(n-1)p^2 + np \end{aligned} \tag{8.6}$$

D'où:

$$\begin{aligned} Var(X) &= E[X^2] - (E[X])^2 = n(n-1)p^2 + np - n^2p^2 \\ &= np(1-p) \end{aligned} \tag{8.7}$$

Propriétés de la variance

Si a et b sont deux constantes, on a:

$$Var(aX + b) = a^2 Var(X)$$

Pour celà, on utilise $E[aX + b] = aE[X] + b$. Donc:

$$\begin{aligned} Var(aX + b) &= E[((aX + b) - (aE[X] + b))^2] \\ &= E[(aX + b - aE[X] - b)^2] = E[a^2(X - E[X])^2] \\ &= a^2 E[(X - E[X])^2] = a^2 Var(X) \end{aligned} \tag{8.8}$$

8.0.25 Covariance, corrélation de deux variables aléatoires

Covariance

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

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

Ce qui donne aussi par developpement:

$$Cov(X, Y) = E[XY - E[X]Y - XE[Y] + E[X]E[Y]] = E[XY - E[X]E[Y]]$$

On peut voir que si les deux variables aléatoires X et Y sont indépendantes, ce qui signifie que $E[XY] = E[X]E[Y]$, alors $Cov(X, Y) = 0$. La réciproque est fausse.

Variance de sommes de variables aléatoires

On montre que:

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

Démonstration:

$$Var(X + Y) \;=\; E[(X + Y - E[X + Y])^2] = E[(X + Y - EX - EY])^2]$$

$$
\begin{aligned}
&= E[(X-EX)^2+(Y-EY])^2+2(X-EX)(Y-EY)] \\
&= E[(X-EX)^2]+E[(Y-EY])^2]+2E[(X-EX)(Y-EY)] \\
&= Var(X)+Var(Y)+2Cov(X,Y)
\end{aligned} \tag{8.9}
$$

En général pour plusieurs variables aléatoires, on a:

$$Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i<j} Cov(X_i, Y_j)$$

Si $X_1, X_2, \cdots, X_n$ sont indépendants deux à deux, alors $Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i)$.

Corrélation de deux variables aléatoires

La corrélation de deux variables aléatoires X et Y, notée par $\rho(X,Y)$, est définie par:

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

lorsque $Vat(X)Var(Y)$ est non nul. On montre que $-1 \leq \rho(X,Y) \leq +1$.

Covariance de sommes de variables aléatoires

On montre que:

$$Cov(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j) = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov(X_i, Y_j)$$

<u>Exercice</u>:
Calculer la variance d'une v.a. uniforme sur l'intervalle $[a,b]$.
On a:

$$
\begin{aligned}
E[X] &= \frac{a+b}{2} \\
E[X^2] &= \int_a^b \frac{x^2\ dx}{b-a} = \frac{1}{b-a}\ \frac{x^3}{2}|_a^b = \frac{b^3-a^3}{3(b-a)} \\
Var(X) &= E[X^2]-(E[X])^2 = \frac{(b-a)^2}{12}
\end{aligned}
$$

<u>Exercice</u>:
Calculer la variance d'une v.a. exponentielle de paramètre λ.
On a:

$$
\begin{aligned}
E[X] &= 1/\lambda \\
E[X^2] &= \int_0^\infty dx\ x^2 \lambda e^{-\lambda x} = 2/\lambda^2 \\
Var(X) &= 1/\lambda^2
\end{aligned}
$$

<u>Exercice</u>:
Montrer que:
1- $Cov(X+Y,Z) = Cov(X,Z) + Cov(Y,Z)$
2- $Cov(a+bX, c+dY) = bd\ \ Cov(X,Y)$, a,b,c,d sont des constantes.
<u>Exercice</u>:
La densité de probabilité d'une v.a. X est donnée par:

$$f(x) = \begin{cases} a+bx^2 & (0<x<1) \\ 0 & \text{sinon} \end{cases}$$

Si $E[X] = 3/5$, trouver a et b et calculer $Var(X)$.
Réponse: $a = 3/5$, $b = 6/5$, $Var(X) = 2/25$.
<u>Exercice</u>:

Calculer $E[X]$ et $Var(X)$ lorsque la densité de probabilité de X est donnée par:

$$f(x) = \begin{cases} 5/x^2 & (x > 5) \\ 0 & \text{sinon} \end{cases}$$
$$f(x) = \begin{cases} 3(1-x^2)/4 & (-1 < x < +1) \\ 0 & \text{sinon} \end{cases}$$
$$f(x) = \begin{cases} (xe^{-x/2})/4 & (x > 0) \\ 0 & \text{sinon} \end{cases}$$

Exercice:
La densité de probabilité simultanée de deux v.a. X et Y est donnée par:

$$f(x,y) = c(x^2 - y^2)e^{-y} \quad \text{lorsque} \quad -y \leq x \leq +y, 0 < y < +\infty$$

1- Trouver la valeur de c.
2- Trouver les densités marginales de X et Y.
3- Calculer les espérances mathématiques et les variances de X et Y.
4- Calculer la corrélation de X et Y.