

Probabilités et Statistiques

Année 2009/2010

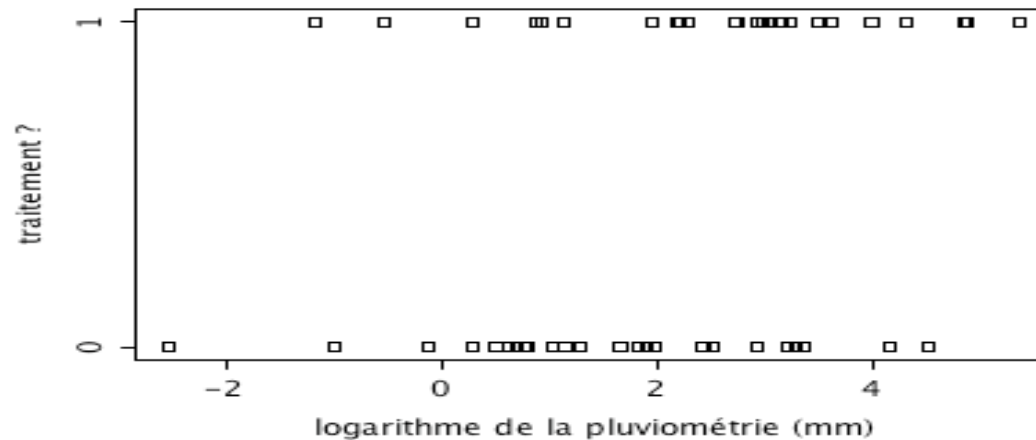
laurent.carraro@telecom-st-etienne.fr

olivier.roustant@emse.fr

Cours n°7

Variance – Covariance
Loi des grands nombres

Des indicateurs de dispersion



	Sans traitement	Avec traitement
DISPERSION		
Ecart-type	1.64	1.60
q(75%) - q(25%)	1.86	1.42
q(5%)	- 0.99	- 0.32
q(95%)	3.95	4.86

Version probabiliste

- quantile d'ordre α : q tel que $P(X < q) = \alpha$
 - Ou encore : $q = F_X^{-1}(\alpha)$
- intervalle interquartile : $F_X^{-1}(0.75) - F_X^{-1}(0.25)$
- variance : $\text{var}(X) = E[(X - E[X])^2]$
- écart-type : $\sigma(X) = (\text{var}(X))^{1/2}$

QUESTION :

Quel indicateur est le plus facile à manipuler ?

Variance et covariance

- La variance est une **forme quadratique**, de forme bilinéaire associée **la covariance** :

$$\text{cov}(X,Y) = E[(X-E[X])(Y-E[Y])]$$

- $\text{var}(X) = \text{cov}(X,X)$
- **cov** est une forme bilinéaire symétrique positive
 - $\text{cov}(X,Y) = \text{cov}(Y,X)$; $\text{cov}(X,X) \geq 0$
 - $\text{cov}(aX + bY, Z) = a \text{cov}(X,Z) + b \text{cov}(Y,Z)$
 - $\text{var}(aX) = \text{cov}(aX, aX) = a^2 \text{var}(X)$
- $\text{cov}(X,X) = 0 \rightarrow X = E[X]$ avec proba. 1

Des propriétés bien utiles

➤ Propriétés

$$\text{var}(X) = E[X^2] - E[X]^2$$

$$\text{cov}(X,Y) = E[XY] - E[X]E[Y]$$

Suggestion pour la démonstration : noter

$$m_X = E[X] \text{ et } m_Y = E[Y]$$

et remarquer que ce sont des nombres fixés.

Corrélation

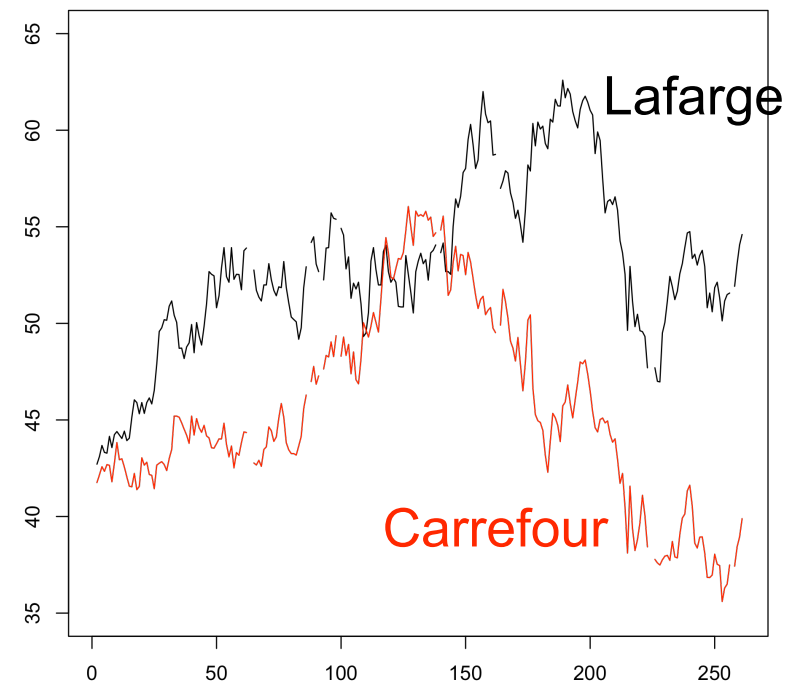
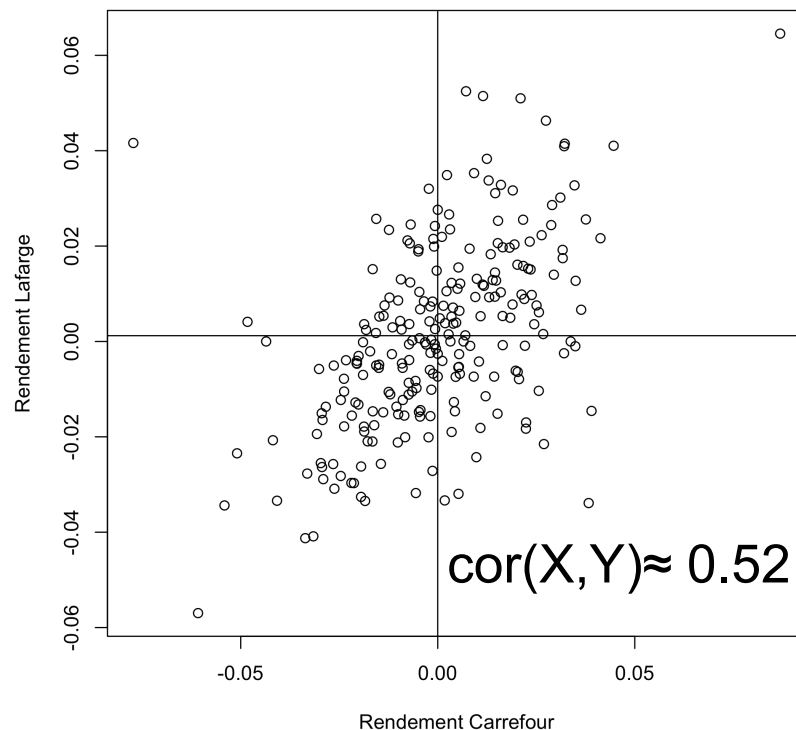
- La **corrélation** est la covariance des variables **centrées et réduites** :

$$\begin{aligned}\text{cor}(X,Y) &= \text{cov}[(X-E[X])/\sigma(X), (Y-E[Y])/\sigma(Y)] \\ &= \text{cov}(X,Y)/(\sigma(X)\sigma(Y))\end{aligned}$$

- Ainsi la corrélation est indépendante de changements d'unité. Comparer :
- $\text{cov}(10X, 100Y) = 1.000 \text{ cov}(X,Y)$
 - $\text{cor}(10X, 1000Y) = \text{cor}(X,Y)$

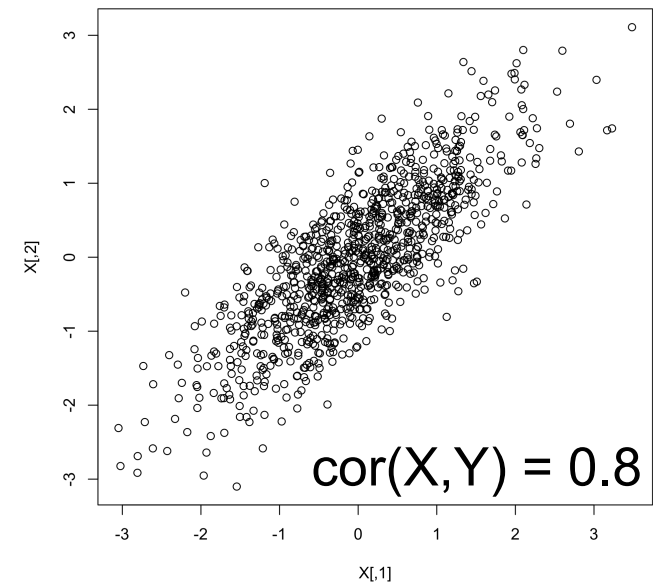
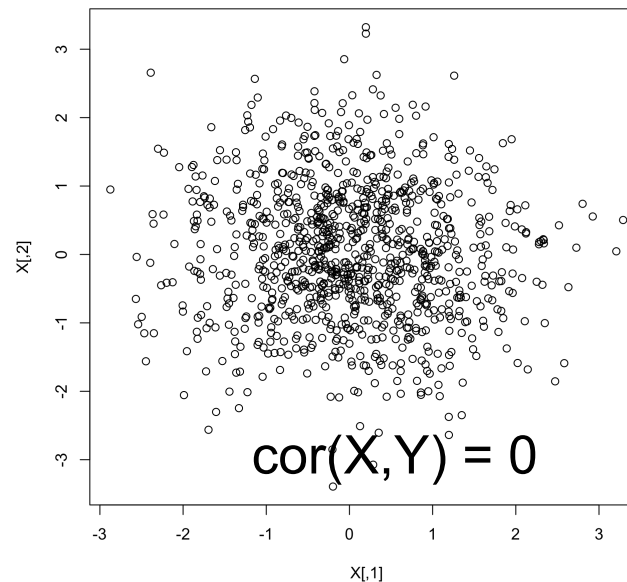
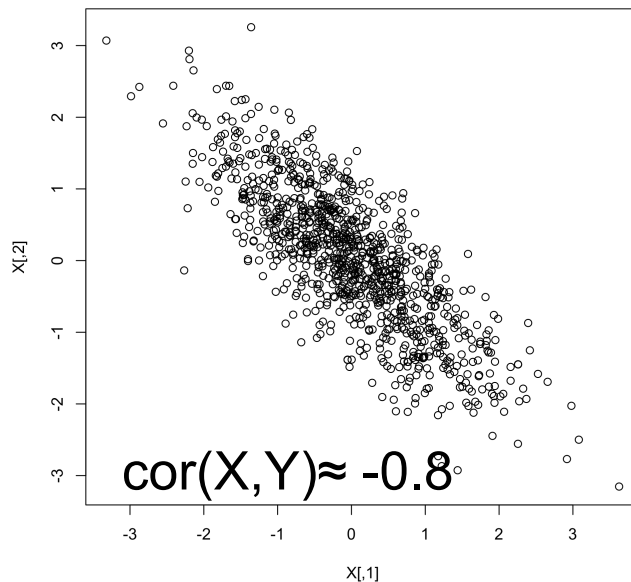
Interprétation de la covariance

- Cours journalier de 2 actions du CAC40 (droite) et leurs taux de rendements (gauche)



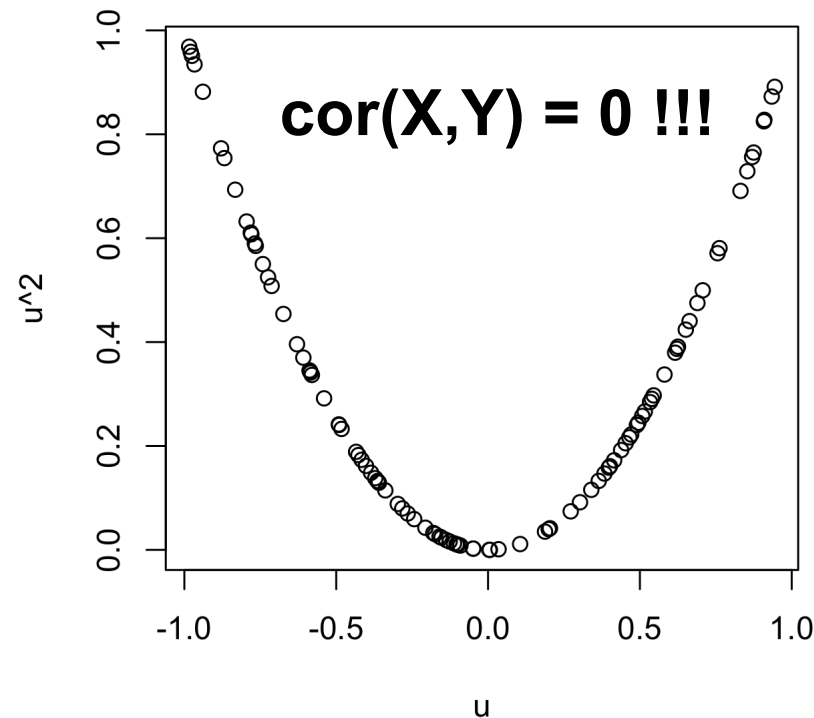
Interprétation de la covariance

- 500 réalisations d'un vecteur [gaussien] (X,Y) , avec $\text{cor}(X,Y)$ fixé.



Non corrélation $\neq$ indépendance !

- Soit $X = U$ de loi uniforme sur $[-1,1]$, et $Y = U^2$



Covariance et indépendance

➤ Théorème

X et Y sont **indépendantes** ssi pour toutes fonctions f et g ("mesurables"), on a:

$$\text{cov}(f(X), g(Y)) = 0$$

➤ Corollaire

Si X et Y sont indépendantes, alors $\text{cov}(X, Y) = 0$

On dit alors que X et Y sont **non corrélées**

La réciproque est FAUSSE (cf ex. précédent)

Variançe d'une somme

Proposition

- Si X et Y sont non corrélées, alors :
$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$$
- Généralisation. Si $X_1, \dots, X_n$ sont non corrélées,
$$\text{var}(X_1 + \dots + X_n) = \text{var}(X_1) + \dots + \text{var}(X_n)$$

Un problème

On veut calculer $\mu = E(g(U))$ avec U uniforme sur $[0,1]$

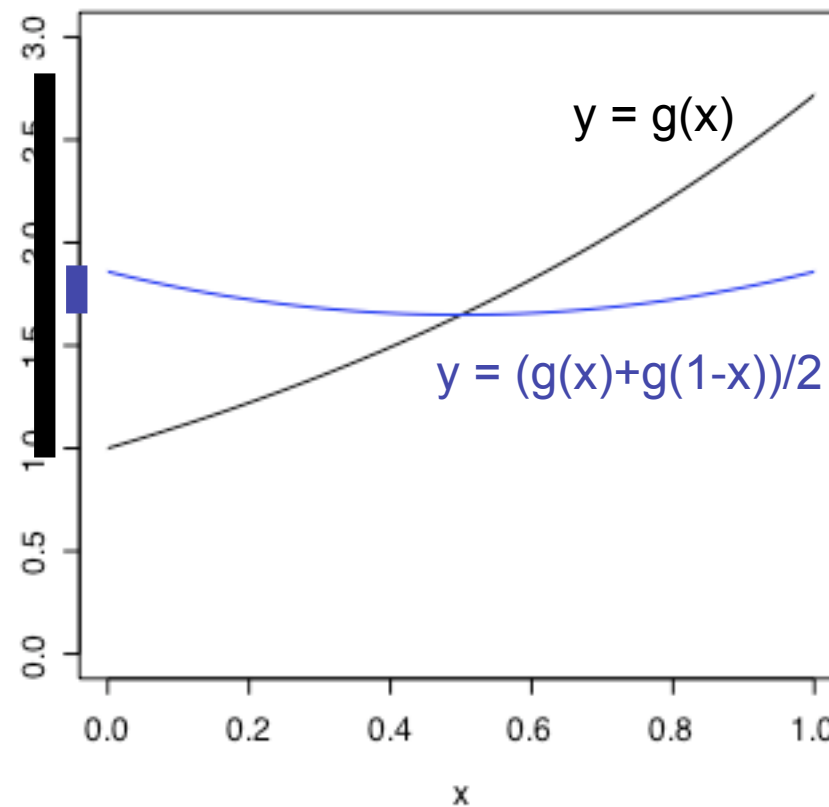
Un estimateur naturel M_n est la moyenne des valeurs obtenues sur n simulations.

- Que se passe-t-il quand $n \rightarrow +\infty$
- Calculer l'erreur quadratique moyenne :
$$E[(M_n - \mu)^2]$$
- Comment faire pour accélérer la convergence ?

Exemple de technique de réduction de variance

novembre 09

Réduction de variance par
variables antithétiques



Loi forte des grands nombres

Théorème

Soit $X_1, \dots, X_n$ des v.a. indépendantes et de même loi (i.i.d.) telles que $\mu := E(X_i)$ existe.

Alors $(X_1 + \dots + X_n)/n \rightarrow \mu$ p.s.

p.s. = "presque sûrement" = avec probabilité 1

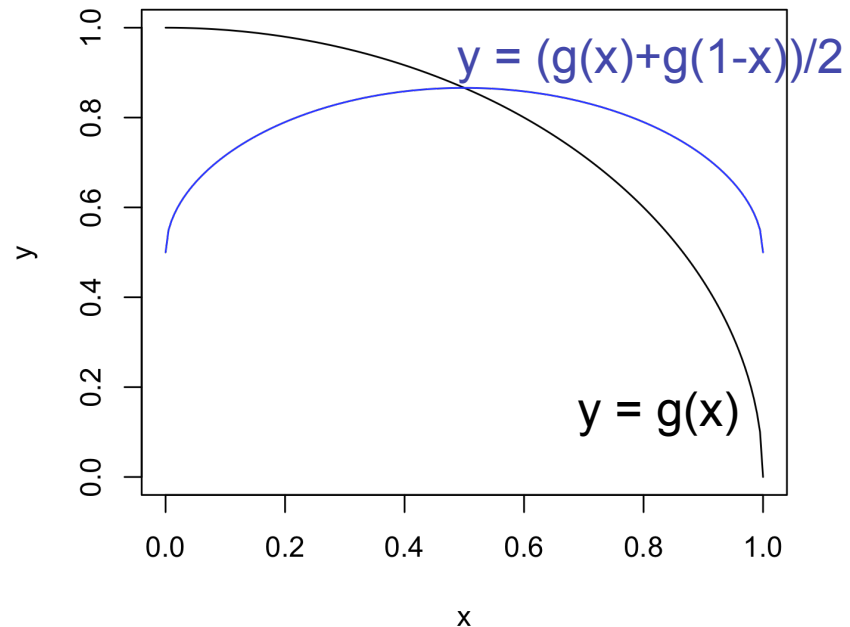
Exercice

On veut calculer π par simulation, en remarquant que $\pi/4 = E(g(U))$, avec U unif. sur $[0,1]$ et $g(x)=(1-x^2)^{1/2}$ (arc de cercle...)

Que vaut l'erreur quadratique moyenne dans ce cas ?

On propose d'accélérer la convergence en notant que $\pi/4 = E(h(U))$, avec $h(x)=(g(x)+g(1-x))/2$. Faire un dessin. Montrer qu'on divise par un facteur 7 l'erreur quad. moyenne.

Exercice (illustration)



Calcul de pi avec 1000 simulations

