

Probabilités et Statistiques

Année 2010/2011

olivier.roustant@emse.fr

laurent.carraro@telecom-st-etienne.fr

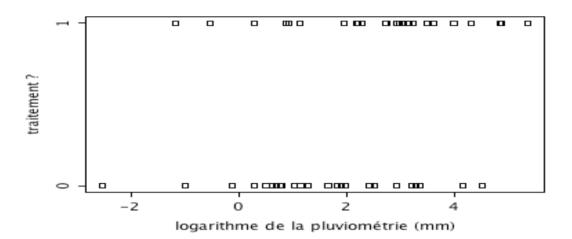


Cours n°8

Variance - Covariance Variance d'une somme Vers la loi d'une somme



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) = α

 Ou encore : q = $F_x^{-1}(\alpha)$
- \rightarrow intervalle interquartile: $F_X^{-1}(0.75) F_X^{-1}(0.25)$
- \triangleright variance: var(X) = E[(X E[X])²)
- \triangleright écart-type : $\sigma(X) = (var(X))^{1/2}$

QUESTION:

Quel indicateur est le plus facile à manipuler?

Variance et covariance

- La variance est une forme quadratique, de forme bilinéaire associée la covariance : cov(X,Y) = E[(X-E[X])(Y-E[Y])]
- \triangleright var(X) = cov(X,X)
- > cov est une forme bilinéaire symétrique positive
 - cov(X,Y) = cov(Y,X); $cov(X,X) \ge 0$
 - = cov(aX + bY, Z) = a cov(X,Z) + b cov(Y,Z)
 - $var(aX) = cov(aX, aX) = a^2 var(X)$
- > cov(X,X) = 0 -> X = E[X] avec proba. 1



Corrélation

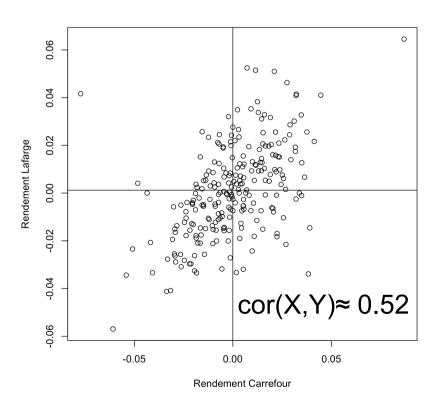
La corrélation est la covariance des variables centrées et réduites : cor(X,Y) = cov[(X-E[X])/σ(X), (Y-E[Y])/σ(Y)] = cov(X,Y)/(σ(X)σ(Y))

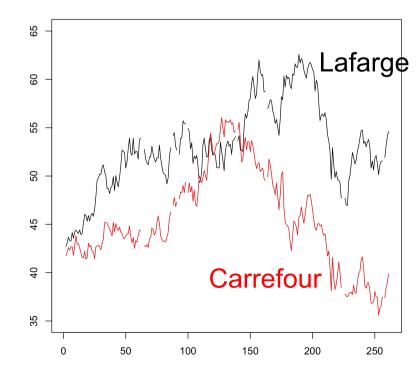
- > Ainsi la corrélation est indépendante de changements d'unité. Comparer :
 - cov(10X, 100Y) = 1.000 cov(X,Y)
 - cor(10X, 1000Y) = cor(X,Y)



Interprétation de la covariance

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

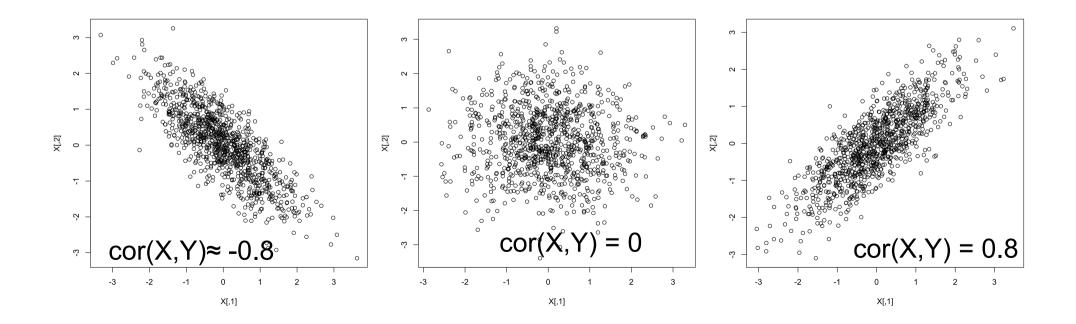






Interprétation de la covariance

> 500 réalisations d'un vecteur [gaussien] (X,Y), avec cor(X,Y) fixé.





Des propriétés bien utiles

Propriétés

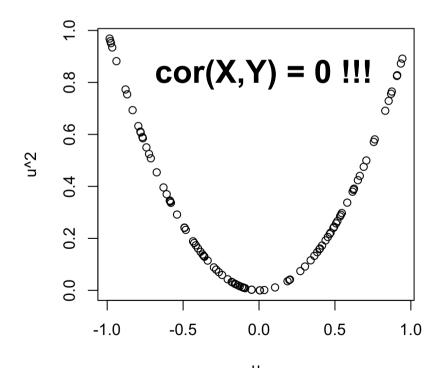
$$var(X) = E[X^2] - E[X]^2$$
$$cov(X,Y) = E[XY] - E[X]E[Y]$$

Suggestion pour la démonstration : noter $m_X = E[X]$ et $m_Y = E[Y]$ et remarquer que ce sont des nombres fixés.



Non corrélation ≠ 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: cov(f(X), g(Y)) = 0

> Corollaire

Si X et Y sont indépendantes, alors cov(X,Y)=0 La réciproque est FAUSSE (cf ex. précédent)

Variance d'une somme

Proposition

- Si X et Y sont non corrélées, alors : var(X + Y) = var(X) + var(Y)
- Figure 6 Généralisation. Si X_1 , ..., X_n sont non corrélées, $var(X_1 + ... + X_n) = var(X_1) + ... + var(X_n)$

Remarque

 $Cov(X,Y) \le 0 \rightarrow var(X+Y) \le var(X) + var(Y)$ (diversification -> réduction du risque)



Application

On veut estimer la proportion p de pièces défectueuses dans une chaîne de fabrication.

Combien de pièces faut-il tester pour que p soit estimée avec une précision (absolue) de 1% ?



Application

On veut estimer la proportion p de pièces défectueuses dans une chaîne de fabrication.

Combien de pièces faut-il tester pour que p soit estimée avec une précision (absolue) de 1% ?

Indications de solution :

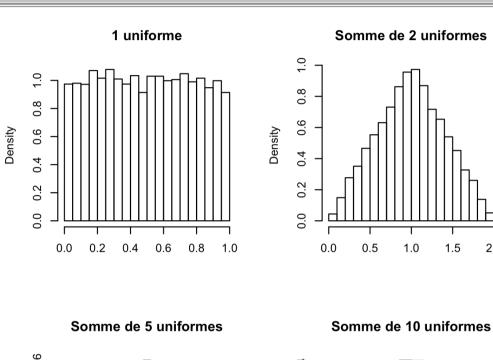
Estimation de $p: p_n = (X_1 + ... + X_n)/n$, avec X_i iid de loi B(p) (modèle). Idéalement, on voudrait $|p_n - p| \le 1\%$ avec une bonne probabilité. Aujourd'hui, on va seulement demander $\sigma(p_n) \le 1\%$. On calcule $var(p_n)=p(1-p)/n$, d'où $n \ge 10^4p(1-p)$ (ex : p=5% -> $n \ge 475$)

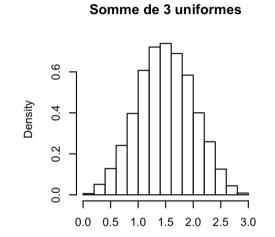


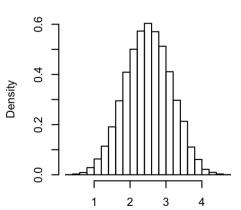
Vers la loi d'une somme de v.a. i.i.d ...

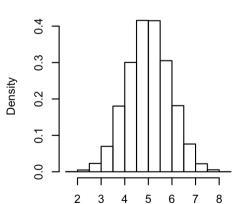


Loi d'une somme de v.a. i.i.d?

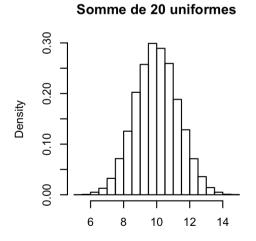








2.0





Loi d'une somme de v.a. i.i.d?

