

Statistiques et probabilités

Cours n°5

Guillaume Postic

Université Paris-Saclay, Univ. Evry
Département informatique

Master 1 MIAE - 2022/2023

Loi de probabilité jointe

Loi de probabilité à **plusieurs variables**

- Cas discret : fonction de masse $p(x_i, y_j)$
- Cas continu : fonction de densité $f(x, y)$
- Fonction de répartition $F(x, y) = P(X \leq x, Y \leq y)$

Loi de probabilité jointe : cas discret

Exemple 1

Jet de deux dés à 6 faces : premier X , second Y

Table des probabilités jointes

$X \backslash Y$	1	2	3	4	5	6
1	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
6	1/36	1/36	1/36	1/36	1/36	1/36

Loi de probabilité jointe : cas discret

Exemple 2

Jet de deux dés à 6 faces : premier X , total (somme) T

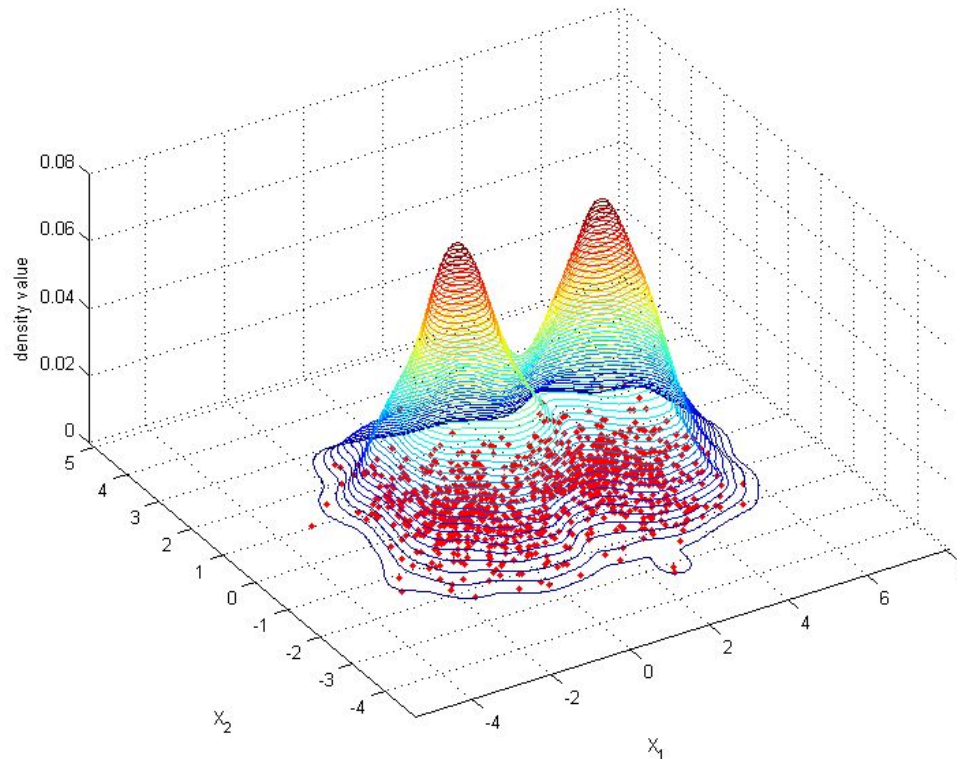
Table des probabilités jointes

$X \backslash T$	2	3	4	5	6	7	8	9	10	11	12
1	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0	0
2	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0	0
3	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0	0
4	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0	0
5	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36	0
6	0	0	0	0	0	1/36	1/36	1/36	1/36	1/36	1/36

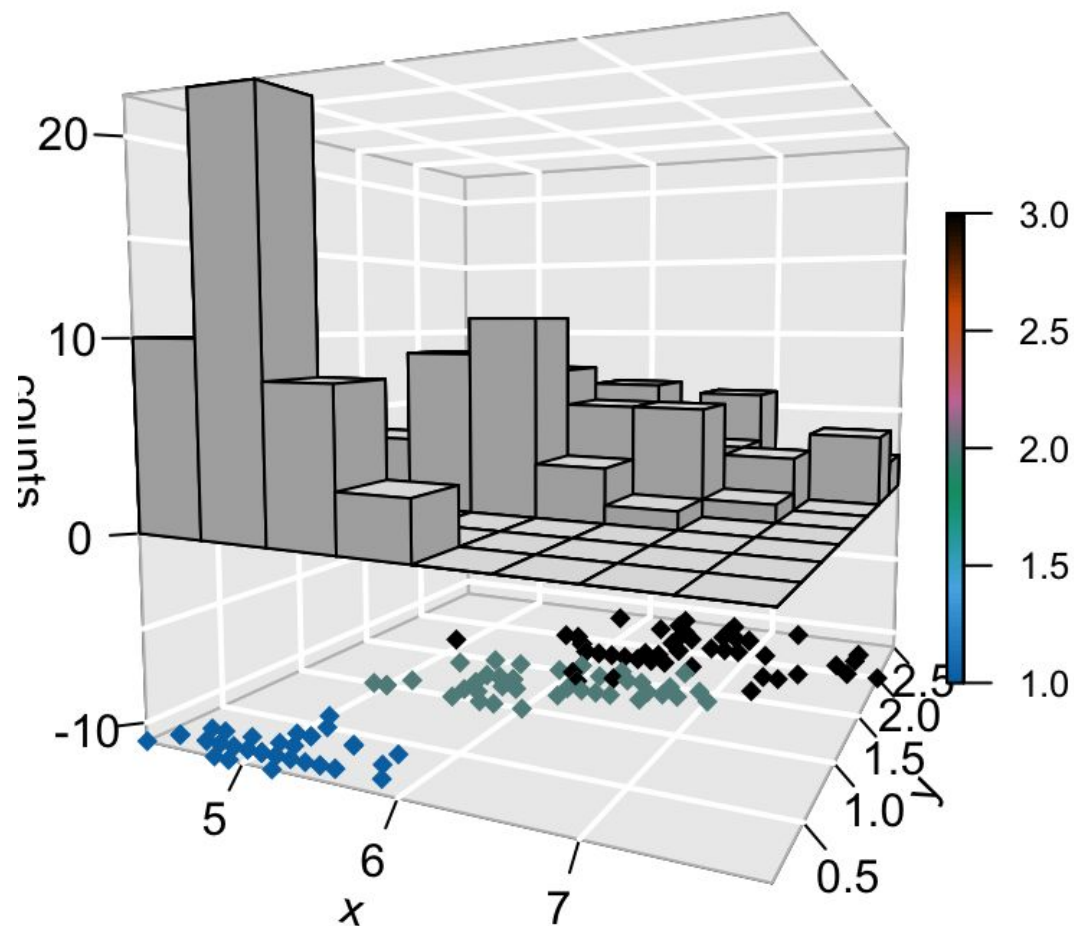
Loi de probabilité jointe : cas continu

Exemple

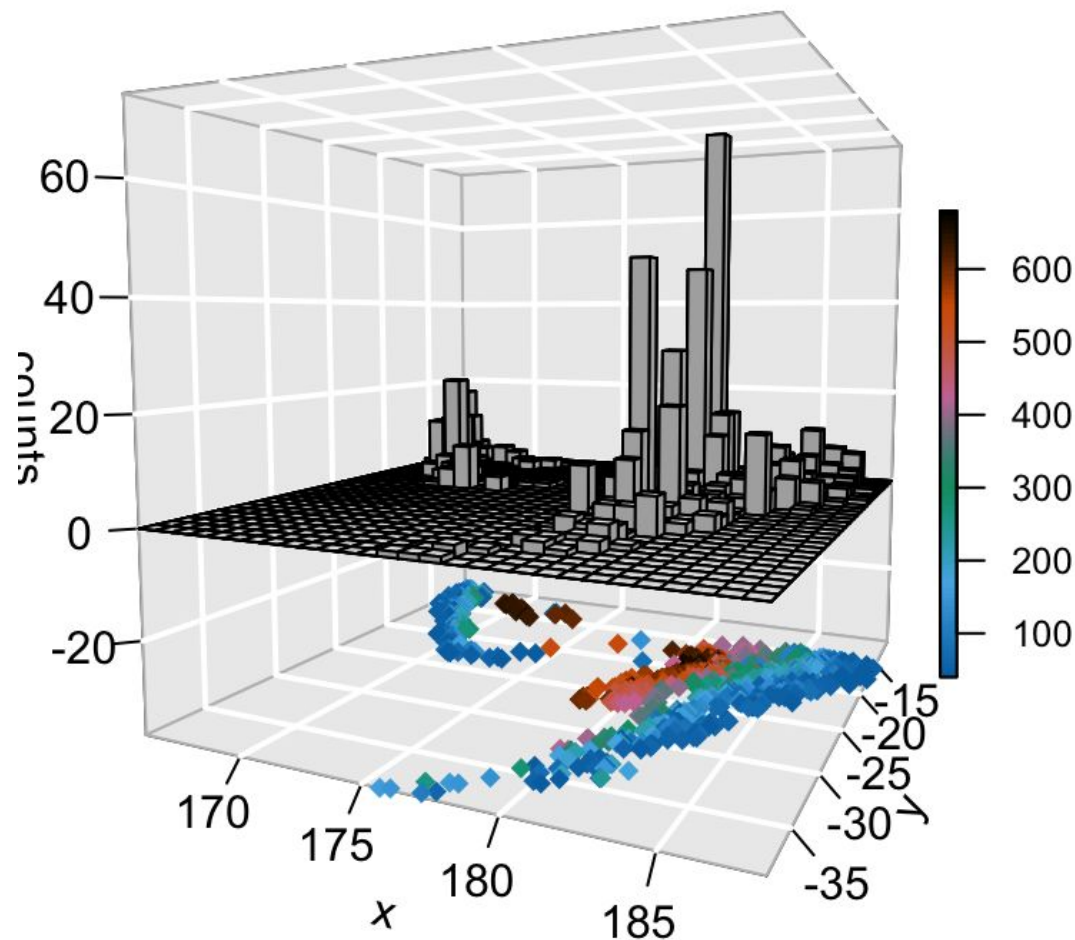
- X prend des valeurs dans l'intervalle $[a, b]$ et Y dans $[c, d]$
- (X, Y) prend des valeurs dans $[a, b] \times [c, d]$



Loi de probabilité jointe : cas continu



Loi de probabilité jointe : cas continu



Loi de probabilité jointe

Variables discrète et nominale : **table de contingence**

	Alice	Bob	Charlie	David	Classe
A	4	0	1	2	7
B	1	1	1	2	5
C	0	1	2	0	3
D	0	2	1	0	3
E	0	0	0	1	1
F	0	1	0	0	1
	5	5	5	5	20

- Probabilité **jointe** : $p(A \cap \text{Alice}) = 4/20 = 1/5$
- Probabilité **conditionnelle** :
 - $p(A | \text{Alice}) = p(A \cap \text{Alice}) / p(\text{Alice}) = (4/20) / (5/20) = 4/5$
- Probabilités **marginales** :
 - $p(A) = 7/20$ et $p(\text{Alice}) = 5/20 = 1/4$

Loi de probabilité jointe

Propriétés des fonctions de masse et densité jointes

Cas discret

1. $0 \leq p(x_i, y_j) \leq 1$

2. Probabilité totale vaut 1

$$\sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) = 1$$

Cas continu

1. $0 \leq f(x, y)$

2. Probabilité totale vaut 1

$$\int_c^d \int_a^b f(x, y) dx dy = 1$$

Note : $f(x, y)$ peut être plus grand que 1,

car **c'est une densité, pas une probabilité**

Covariance

Mesure du degré avec lequel deux variables aléatoires varient conjointement.

Par exemple : le poids et la taille d'individus

Pour X et Y variables aléatoires de moyennes μ_X et μ_Y

$$\text{Cov}(X, Y) = E((X - \mu_X) (Y - \mu_Y))$$

Covariance

Propriétés

1. $\text{Cov}(aX + b, cY + d) = ac\text{Cov}(X, Y)$, pour des constantes a, b, c, d
2. $\text{Cov}(X_1 + X_2, Y) = \text{Cov}(X_1, Y) + \text{Cov}(X_2, Y)$
3. $\text{Cov}(X, X) = \text{Var}(X)$
4. $\text{Cov}(X, Y) = E(XY) - \mu_X \mu_Y$
5. Si X et Y sont indépendants, alors $\text{Cov}(X, Y) = 0$
6. ⚠ La réciproque n'est pas vraie : même si la covariance est nulle, les variables peuvent ne pas être indépendantes

Corrélation

Le coefficient de corrélation (de Pearson) entre X et Y est défini par :

$$\text{Cor}(X, Y) = \rho = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

Propriétés :

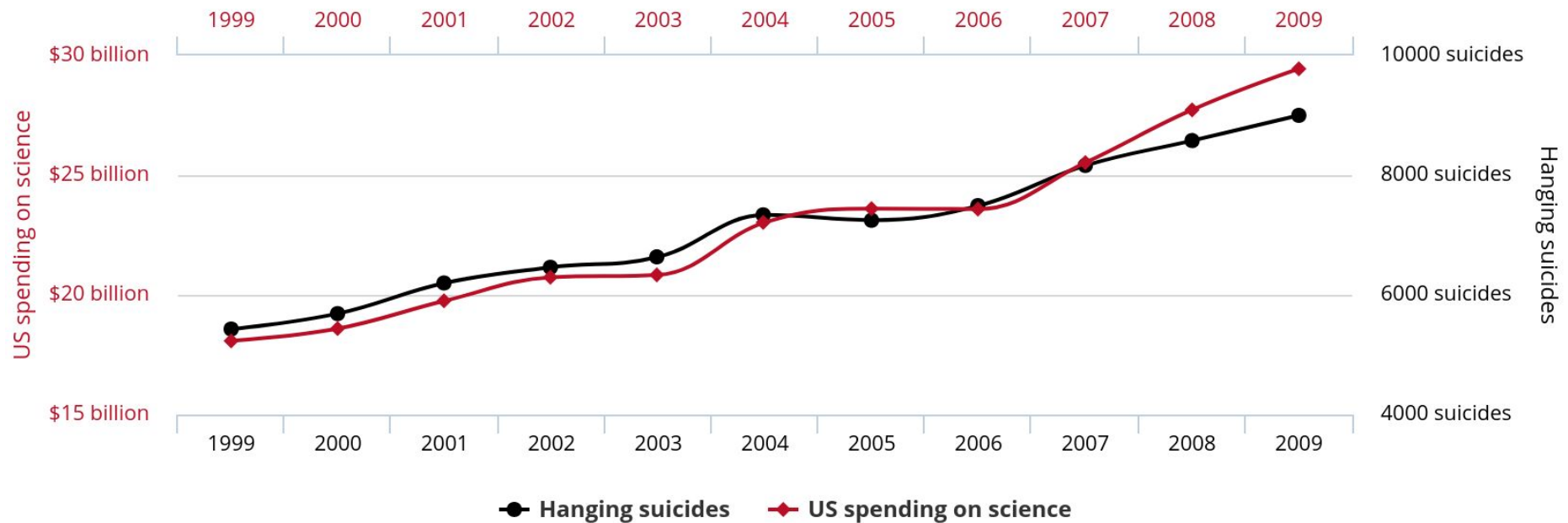
1. ρ est la covariance des versions centrées réduites de X et Y
2. ρ est adimensionnelle
3. $-1 \leq \rho \leq 1$
 $\rho = 1$ ssi $Y = aX + b$ avec $a > 0$
et $\rho = -1$ ssi $Y = aX + b$ avec $a < 0$

Corrélations fallacieuses (*spurious*)

US spending on science, space, and technology
correlates with

Suicides by hanging, strangulation and suffocation

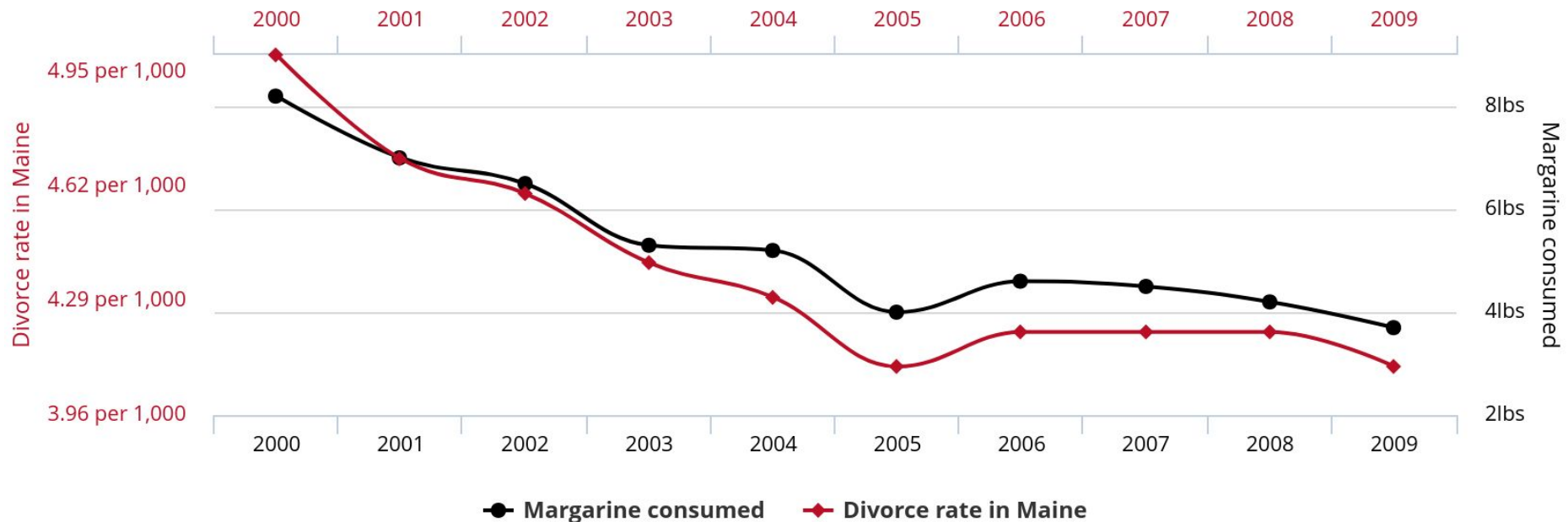
Correlation: 99.79% ($r=0.99789126$)



Corrélations fallacieuses (*spurious*)

Divorce rate in Maine correlates with Per capita consumption of margarine

Correlation: 99.26% ($r=0.992558$)



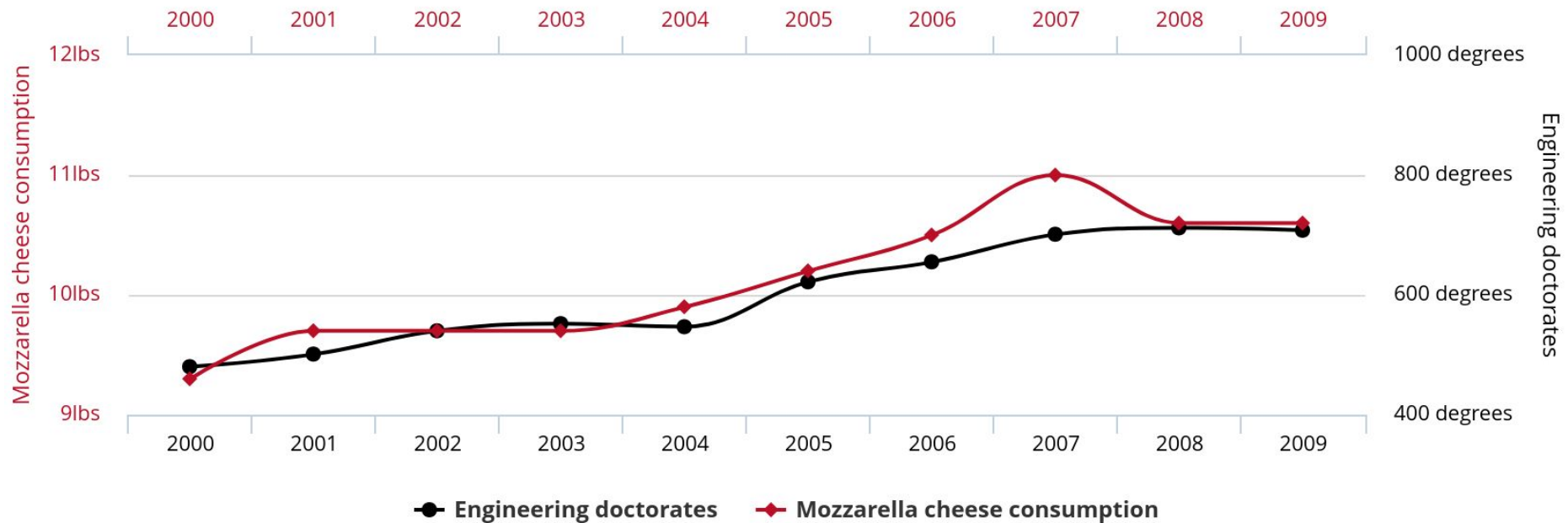
Corrélations fallacieuses (*spurious*)

Per capita consumption of mozzarella cheese

correlates with

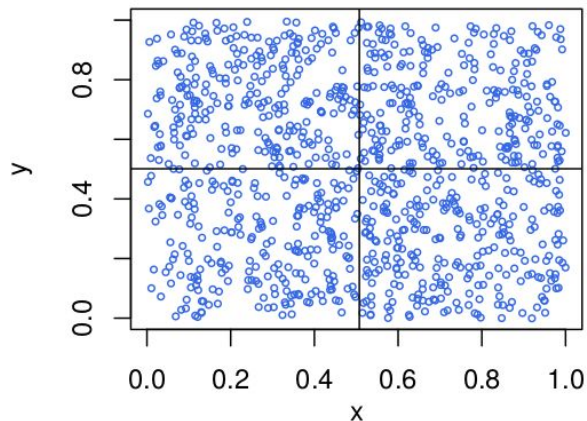
Civil engineering doctorates awarded

Correlation: 95.86% ($r=0.958648$)

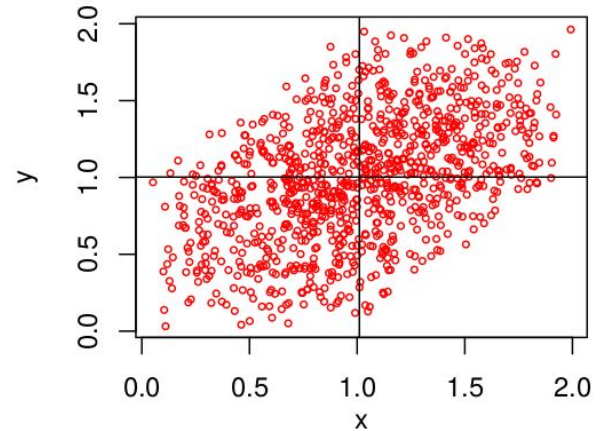


Nuage de points (*scatter plot*)

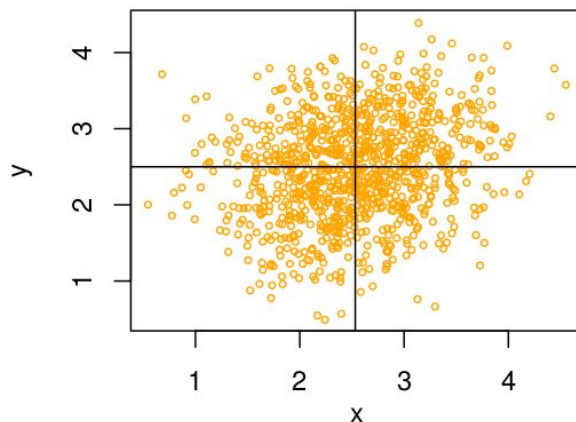
(1, 0) $\text{cor}=0.00$, $\text{sample_cor}=-0.07$



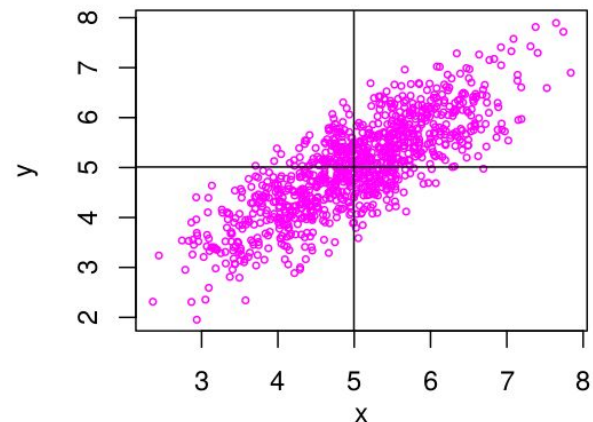
(2, 1) $\text{cor}=0.50$, $\text{sample_cor}=0.48$



(5, 1) $\text{cor}=0.20$, $\text{sample_cor}=0.21$



(10, 8) $\text{cor}=0.80$, $\text{sample_cor}=0.81$



Transformation des données

