

MVE055 2017 Lecture 4

Marco Longfils

Department of Mathematical Sciences, Chalmers University of Technology and
University of Gothenburg

Wednesday 6th September, 2017

- Up to now: univariate distribution \rightarrow a single random value.
- Typically we need to consider multivariate distribution, used to model many uncertain values. In this way, we can take into account the dependencies between these quantities.
- We will now focus on bivariate distribution, but generalization to more variables is straightforward.

Definition (discrete joint density)

Let X, Y be two discrete random variables. The vector (X, Y) is a bivariate discrete random variable and a function f_{XY} which satisfies

$$f_{XY}(x, y) = \Pr[X = x, Y = y], \text{ for all } (x, y) \in \mathbb{R}^2$$

is called joint density for the vector (X, Y) .

Theorem

A function $f(x, y)$ is a discrete joint density if and only if

- $f(x, y) \geq 0$
- $\sum_{\text{all } (x, y)} f(x, y) = 1$

Definition (discrete marginal density)

Let (X, Y) be a bivariate discrete random vector with joint density f_{XY} . The marginal density f_X for X is given by

$$f_X(x) = \sum_{\text{all } y} f_{XY}(x, y)$$

and similarly the marginal density for Y is

$$f_Y(y) = \sum_{\text{all } x} f_{XY}(x, y)$$

Joint density function/2

Definition (continuous joint density)

Let X, Y be two continuous random variables. The vector (X, Y) is a bivariate continuous random variable and a function f_{XY} which satisfies

- $f(x, y) \geq 0$
- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dy dx = 1$
- $\Pr[X \in [a, b], Y \in [c, d]] = \int_a^b \int_c^d f(x, y) dy dx$, for all $a, b, c, d \in \mathbb{R}$ is called joint density for the vector (X, Y) .

Furthermore, the marginal densities f_X and f_Y for, respectively, X and Y are given by

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx$$

Independence of random variable

- Recall: two events A, B are said to be independent if $\Pr[A \cap B] = \Pr[A] \Pr[B]$.

Definition (independence for random variables)

Two random variables X and Y with joint density f_{XY} and marginal densities f_X, f_Y are independent if and only if

$$f_{XY}(x, y) = f_X(x)f_Y(y)$$

for all x and y

Expected value

- In general, the expected value of a function of $H(X, Y)$ is given by

$$\mathbb{E}[H(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(x, y) f_{XY}(x, y) dx dy$$

if X, Y are discrete and by

$$\mathbb{E}[H(X, Y)] = \sum_{\text{all } (x, y)} H(x, y) f_{XY}(x, y)$$

- Same properties as in the discrete case.
- If X, Y are independent then

$$\mathbb{E}[XY] = \mathbb{E}[X] \mathbb{E}[Y]$$

(the viceversa is not true in general).

Definition (Covariance)

Let X and Y be two random variables. The covariance between X and Y is defined as

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

It holds that

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

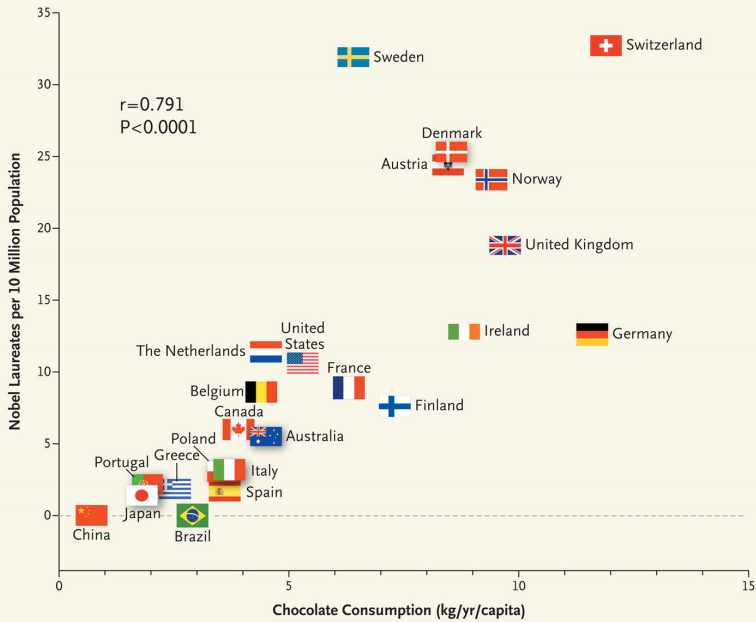
- If X, Y are independent $\rightarrow \mathbb{E}[XY] = \mathbb{E}[X] \mathbb{E}[Y] \rightarrow \text{Cov}(X, Y) = 0$ (the viceversa is not true in general).
- $\text{Cov}(X, Y)$ gives an indication of association between X and Y
- $\text{Cov}(X, Y)$ can be any real value \rightarrow no information about the strength of the dependence.

Definition (Correlation)

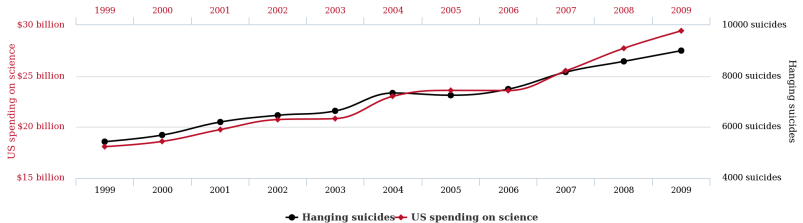
Let X and Y be two random variables. The correlation between X and Y is defined as

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}[X] \text{Var}[Y]}}.$$

- ρ_{XY} measures linear dependence between X and Y .
- ρ_{XY} can be any real value between -1 and 1 .
- $|\rho_{XY}| = 1$ if and only if $Y = \beta_0 + \beta_1 X$ for some β_0 and $\beta_1 \neq 0$.



US spending on science, space, and technology correlates with Suicides by hanging, strangulation and suffocation



tylervigen.com

- Recall: given two events A, B (if $\Pr[B] > 0$) we have
$$\Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]}.$$

Definition (conditional density)

Given two random variables X and Y with joint density f_{XY} and marginal densities f_X, f_Y we define the conditional density for X given $Y = y$ as

$$f_{X|y} = \frac{f_{XY}(x, y)}{f_Y(y)}, \text{ if } f_Y(y) > 0$$

Transformation of variables

Theorem

Let (X, Y) be a continuous bivariate vector with density f_{XY} . Moreover, let (U, V) be a continuous bivariate vector with density f_{UV} and

$$(X, Y) = (h_1(U, V), h_2(U, V))$$

where h_1 and h_2 define a one-to-one transformation and have continuous partial derivatives. Then

$$f_{UV}(u, v) = f_{XY}(h_1(u, v), h_2(u, v)) |J|$$

where J is the given by

$$J = \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}.$$