

# Random vectors II.

(Covariance and correlation, convolution)

Lecturer:  
Francesco Dolce

Department of Applied Mathematics  
Faculty of Information Technology  
Czech Technical University in Prague

© 2011–2023 Rudolf B. Blažek, Francesco Dolce, Roman Kotecký, Jitka Hrabáková, Petr Novák, Daniel Vašata

## Probability and Statistics

BIE-PST, WS 2023/24, Lecture 7



# Content

- **Probability theory:**

- ▶ Events, probability, conditional probability, Bayes' Theorem, independence of events.
- ▶ Random variables, distribution function, functions of random variables, characteristics of random variables: expected value, variance, moments, generating function, quantiles, critical values, important discrete and continuous distributions.
- ▶ **Random vectors**, joint and marginal distributions, functions of random vectors, independence of random variables, conditional distribution, **conditional expected value, covariance and correlation**.
- ▶ Markov's and Chebyshev's inequality, weak law of large numbers, strong law of large numbers, Central limit theorem.

- **Mathematical statistics:**

- ▶ Point estimators, sample mean, sample variance, properties of point estimators, Maximum likelihood method.
- ▶ Interval estimators, hypothesis testing, one-sided vs. two-sided alternatives, linear regression, estimators of regression parameters, testing of linear model.

# Recap

**Joint distribution function of a random vector  $(X, Y)$ :**

$$F_{X,Y}(x, y) = P(X \leq x \cap Y \leq y).$$

**Discrete random variables  $X$  and  $Y$**

Joint probabilities of values:

$$P(X = x \cap Y = y)$$

**Marginal distributions:**

$$P(X = x) = \sum_{\text{all } y} P(X = x \cap Y = y)$$

$$P(Y = y) = \sum_{\text{all } x} P(X = x \cap Y = y)$$

**Continuous random variables  $X$  and  $Y$**

Joint density:

$$f_{X,Y}(x, y)$$

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

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

**Independence of  $X$  and  $Y$ :**

$$P(X = x \cap Y = y) = P(X = x) P(Y = y) \quad | \quad f_{X,Y}(x, y) = f_X(x) f_Y(y)$$

**Conditional probabilities / density of  $X$  given  $Y = y$ :**

$$P(X = x|Y = y) = \frac{P(X = x \cap Y = y)}{P(Y = y)} \quad | \quad f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

**Conditional expectation of  $X$  given  $Y = y$ :**

$$E(X|Y = y) = \sum_x x P(X = x|Y = y) \quad | \quad E(X|Y = y) = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx$$

## Functions of random vectors of variables

Similar formulas as for a function of one random variable also hold for the functions of random vectors. Let

$$Z = h(X_1, \dots, X_n) = h(\mathbf{X}).$$

- When variables  $X_1, \dots, X_n$  have a **joint discrete** distribution with probabilities  $P(\mathbf{X} = \mathbf{x})$ , the following relation holds for the distribution function of  $Z$ :

$$F_Z(z) = P(Z \leq z) = \sum_{\{\mathbf{x} \in \mathbb{R}^n: h(\mathbf{x}) \leq z\}} P(\mathbf{X} = \mathbf{x}).$$

- When variables  $X_1, \dots, X_n$  have a **joint continuous** distribution with density  $f_{\mathbf{X}}(\mathbf{x})$ , the distribution function of  $Z$  is then

$$F_Z(z) = P(Z \leq z) = \int \cdots \int_{\{\mathbf{x} \in \mathbb{R}^n: h(\mathbf{x}) \leq z\}} f_{\mathbf{X}}(\mathbf{x}) \, dx_1 \cdots dx_n.$$

## Expected value of the function of a random vector

The expected value  $\mathbb{E} h(X, Y)$  of a real function  $h$  of random variables  $X$  and  $Y$  can be computed without determining the distribution of the variable  $h(X, Y)$ .

- For  $X$  and  $Y$  discrete random variables it holds that

$$\mathbb{E} h(X, Y) = \sum_{i,j} h(x_i, y_j) \mathbb{P}(X = x_i \cap Y = y_j),$$

if the sum converges absolutely.

- For  $X$  and  $Y$  continuous random variables it holds that

$$\mathbb{E} h(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x, y) f_{X,Y}(x, y) dx dy,$$

if the integral converges absolutely.

# Properties of the expected value

Now we can prove the linearity of the expectation.

## Theorem – linearity of expectation

For all  $a, b \in \mathbb{R}$  and all random variables  $X$  and  $Y$  it holds that

$$E(aX + bY) = a E X + b E Y.$$

## Consequence:

- $E(aX + b) = a E X + b$ . This statement was proven before separately.

# Properties of the expected value

## Proof

From the theory concerning the marginal distributions of discrete random variables  $X$  and  $Y$  we have:

$$\begin{aligned} E(aX + bY) &= \sum_{i,j} (ax_i + by_j) P(X = x_i \cap Y = y_j) \\ &= \sum_{i,j} ax_i P(X = x_i \cap Y = y_j) + \sum_{i,j} by_j P(X = x_i \cap Y = y_j) \\ &= a \sum_i x_i \sum_j P(X = x_i \cap Y = y_j) + b \sum_j y_j \sum_i P(X = x_i \cap Y = y_j) \\ &= a \sum_i x_i P(X = x_i) + b \sum_j y_j P(Y = y_j) = a EX + b EY. \end{aligned}$$

For continuous  $X$  and  $Y$  the proof is analogous:

$$\begin{aligned} E(aX + bY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (ax + by) f_{X,Y}(x, y) dx dy = \dots = \\ &= a \int_{-\infty}^{\infty} x f_X(x) dx + b \int_{-\infty}^{\infty} y f_Y(y) dy = a EX + b EY. \end{aligned}$$

□

## Covariance and correlation coefficient

**Mutual linear dependence** of two random variables  $X$  and  $Y$  can be described in the following way:

### Definition

Let  $X$  and  $Y$  be random variables with finite second moments. Then we define the **covariance** of the random variables  $X$  and  $Y$  as

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

If  $X$  and  $Y$  have positive variances then we define the **correlation coefficient** (or **coefficient of correlation**) as

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var } X} \sqrt{\text{var } Y}}.$$

### Definition

Two random variables  $X$  and  $Y$  are called **non-correlated** if  $\text{cov}(X, Y) = 0$ .



# Covariance and the correlation coefficient – properties

## Theorem

For the covariance and the correlation coefficient the following properties hold:

- i)  $\text{cov}(X, Y) = E XY - E X E Y$ ,
- ii)  $X$  and  $Y$  are non-correlated if and only if  $E XY = E X E Y$ ,
- iii)  $\rho(X, Y) \in [-1, 1]$ ,
- iv)  $\rho(aX + b, cY + d) = \rho(X, Y)$  for all  $a, c > 0$  and  $b, d \in \mathbb{R}$ ,
- v)  $\rho(X, Y) = \pm 1$ , if  $a, b \in \mathbb{R}$ ,  $a > 0$  such that  $Y = \pm aX + b$ .

## Proof

- i) 
$$\begin{aligned} \text{cov}(X, Y) &= E((X - E X)(Y - E Y)) = E(XY - X E Y - Y E X + E X E Y) \\ &= E XY - E(X E Y) - E(Y E X) + E(E X E Y) \\ &= E XY - E X E Y - E Y E X + E X E Y = E XY - E X E Y \end{aligned}$$
- ii) Obvious from above.
- iii) From the Schwarz inequality (see bibliography).
- iv) Follows straightforwardly by inserting into the definition.
- v) Follows from the proof of the Schwarz inequality (see bibliography).



## Non-correlated random variables

Let us study the **expectation of the product**  $XY$  of two random variables  $X$  and  $Y$ .

### Definition

Alternative definition: Two random variables  $X$  and  $Y$  are called **non-correlated** if

$$E XY = E X E Y.$$

### Lemma

*If  $X$  and  $Y$  are independent then they are non-correlated.*

### Proof

Let  $X, Y$  be continuous variables. Independence means that  $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ . Thus we have

$$\begin{aligned} E XY &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f_{X,Y}(x, y) \, dx \, dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f_X(x) f_Y(y) \, dx \, dy \\ &= \left( \int_{-\infty}^{+\infty} x f_X(x) \, dx \right) \left( \int_{-\infty}^{+\infty} y f_Y(y) \, dy \right) = E X E Y. \end{aligned}$$

□

# Properties of the variance

It is now possible to obtain the following properties of the variance of sums of two random variables.

## Theorem

i) For  $X$  and  $Y$  with finite second moments:

$$\text{var}(X \pm Y) = \text{var } X + \text{var } Y \pm 2 \text{cov}(X, Y).$$

ii) For non-correlated (independent) random variables it holds that

$$\text{var}(X \pm Y) = \text{var } X + \text{var } Y.$$

# Properties of variance

## Proof

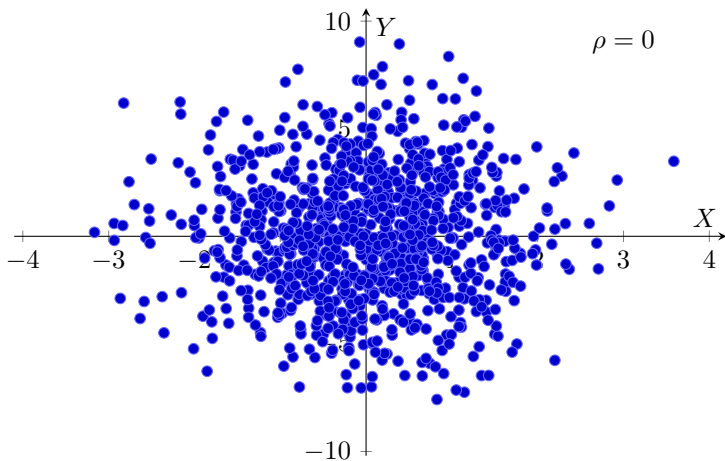
i) Given two random variables  $X$  and  $Y$  we have:

$$\begin{aligned}\text{var}(X \pm Y) &= \text{E}(X \pm Y)^2 - (\text{E}(X \pm Y))^2 = \text{E}(X^2 \pm 2XY + Y^2) - (\text{E} X \pm \text{E} Y)^2 \\ &= \text{E} X^2 \pm 2 \text{E} XY + \text{E} Y^2 - (\text{E} X)^2 \mp 2 \text{E} X \text{E} Y - (\text{E} Y)^2 \\ &= \text{var} X + \text{var} Y \pm (2 \text{E} XY - 2 \text{E} X \text{E} Y) = \text{var} X + \text{var} Y \pm 2 \text{cov}(X, Y).\end{aligned}$$

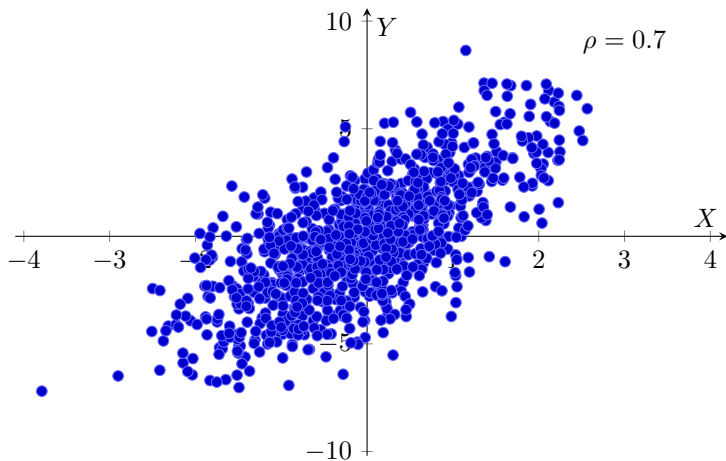
ii) For non-correlated (independent) random variables the covariance is zero.



## Correlation – sample of 1000 values



## Correlation – sample of 1000 values



## Sums of random variables

An important case of a function of multiple random variables is their sum

$$Z = h(\mathbf{X}) = h(X_1, \dots, X_n) = X_1 + \dots + X_n.$$

Consider for simplicity a sum of two random variables:

- If  $X$  and  $Y$  are **discrete** and **independent**, then for  $Z = X + Y$  it holds that

$$P(Z = z) = \sum_x P(X = x) \cdot P(Y = z - x) \quad (\text{discrete convolution}).$$

- If  $X$  and  $Y$  are **continuous** and **independent**, then for  $Z = X + Y$  it holds that

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx \quad (\text{convolution of } f_X \text{ and } f_Y).$$

## Sums of random variables – convolution (discrete case)

The expression for the sum of discrete independent  $X$  and  $Y$  is obtained easily:

$$\begin{aligned} P(Z = z) &= P(X + Y = z) \\ &= \sum_{\{(x_k, y_j): x_k + y_j = z\}} P(X = x_k \cap Y = y_j) \\ &= \sum_{\text{all } x_k} P(X = x_k) P(Y = z - x_k). \end{aligned}$$



## Sums of random variables - convolution (continuous case)

For continuous independent  $X$  and  $Y$  we have:

$$\begin{aligned}
 F_Z(z) = P(X + Y \leq z) &= \iint_{\{(x,y): x+y \leq z\}} f_{X,Y}(x, y) \, d(x, y) \\
 &= \int_{-\infty}^{\infty} \left( \int_{-\infty}^{z-x} f_{X,Y}(x, y) \, dy \right) dx \\
 &\stackrel{y=u-x}{=} \int_{-\infty}^{\infty} \left( \int_{-\infty}^z f_{X,Y}(x, u-x) \, du \right) dx \\
 &= \int_{-\infty}^z \left( \int_{-\infty}^{\infty} f_{X,Y}(x, u-x) \, dx \right) du \\
 &= \int_{-\infty}^z \left( \int_{-\infty}^{\infty} f_X(x) f_Y(u-x) \, dx \right) du.
 \end{aligned}$$

The density  $f_Z$  is any non-negative function, for which  $F_Z(z) = \int_{-\infty}^z f_Z(u) \, du$ .

The expression under the first integral  $f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) \, dx$  is thus the density of  $Z$ .

# Sum of random variables – Normal distribution

## Example – sum of two normal distributions

Suppose that  $X$  and  $Y$  are independent, both having the normal distribution  $N(\mu, 1)$ . We want to obtain the distribution of  $Z = X + Y$ .

The densities of  $X$  and  $Y$  correspond to the normal distribution with variance  $\sigma^2 = 1$ :

$$f_X(x) = \frac{1}{\sqrt{2\pi \cdot 1}} e^{-\frac{(x-\mu)^2}{2 \cdot 1}}, \quad f_Y(y) = \frac{1}{\sqrt{2\pi \cdot 1}} e^{-\frac{(y-\mu)^2}{2 \cdot 1}} \quad x, y \in \mathbb{R}.$$

The density of the sum is obtained using convolution:

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-x-\mu)^2}{2}} dx \\ &= \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}((x-\mu)^2 + (z-x-\mu)^2)} dx. \end{aligned}$$

## Sum of random variables – Normal distribution

### Example – sum of two normal distributions, continuation

The expressions in the exponent can be rewritten as:

$$\begin{aligned}(x - \mu)^2 + (z - x - \mu)^2 &= x^2 - 2\mu x + \mu^2 + z^2 + x^2 + \mu^2 - 2zx - 2\mu z + 2\mu x \\ &= 2 \left(x - \frac{z}{2}\right)^2 + \frac{1}{2} (z - 2\mu)^2.\end{aligned}$$

The expression under the integral can then be split into two multiplicative parts, with one of them not depending on  $x$  and the other one having an integral of 1:

$$\begin{aligned}f_Z(z) &= \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{2(x-z/2)^2}{2}} e^{-\frac{(z-2\mu)^2}{2 \cdot 2}} dx \\ &= \frac{1}{\sqrt{2\pi \cdot 2}} e^{-\frac{(z-2\mu)^2}{2 \cdot 2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(1/2)}} e^{-\frac{(x-z/2)^2}{2 \cdot (1/2)}} dx \\ &= \frac{1}{\sqrt{2\pi \cdot 2}} e^{-\frac{(z-2\mu)^2}{2 \cdot 2}}.\end{aligned}$$

The sum  $Z = X + Y$  has therefore the normal distribution  $N(2\mu, 2)$ . In general, it can be proven that the sum of  $n$  independent normals  $N(\mu, \sigma^2)$  has the distribution  $N(n\mu, n\sigma^2)$ .

## Sum of random variables – Poisson distribution

### Example

Consider two independent random variables  $X$  and  $Y$  with the Poisson distribution with parameters  $\lambda_1$  and  $\lambda_2$ , respectively. Find the distribution of the variable  $Z = X + Y$ .

$$P(X = j) = \frac{\lambda_1^j}{j!} e^{-\lambda_1} \quad P(Y = \ell) = \frac{\lambda_2^\ell}{\ell!} e^{-\lambda_2}, \quad j, \ell = 0, 1, \dots$$

From what we have seen before we know that for  $k = 0, 1, \dots$ :

$$\begin{aligned} P(Z = k) &= \sum_{\{(j, \ell) \in \mathbb{N}_0^2 : j + \ell = k\}} P(X = j) P(Y = \ell) = \sum_{j=0}^k P(X = j) P(Y = k - j) \\ &= \sum_{j=0}^k \frac{\lambda_1^j}{j!} e^{-\lambda_1} \frac{\lambda_2^{k-j}}{(k-j)!} e^{-\lambda_2} = e^{-(\lambda_1 + \lambda_2)} \frac{1}{k!} \sum_{j=0}^k \frac{k!}{j!(k-j)!} \lambda_1^j \lambda_2^{k-j} \\ &= \frac{(\lambda_1 + \lambda_2)^k}{k!} e^{-(\lambda_1 + \lambda_2)}. \quad \sim \text{Poisson}(\lambda_1 + \lambda_2). \end{aligned}$$

✓ An easier way is to use the moment generating function.

## Sums of random variables – moment generating function

The moment generating function can be used to compute moments of random variables.

Taking a sum of independent random variables corresponds to taking a product of their generating functions:

For  $Z = X + Y$  we have

$$\begin{aligned}M_Z(s) &= \mathbb{E}(e^{sZ}) = \mathbb{E}(e^{s(X+Y)}) = \mathbb{E}(e^{sX}e^{sY}) \\ &= \mathbb{E}(e^{sX})\mathbb{E}(e^{sY}) = M_X(s)M_Y(s).\end{aligned}$$

Generally for a vector of independent random variables  $X_1, \dots, X_n$  it holds that:

$$Z = X_1 + \dots + X_n \implies M_Z(s) = M_{X_1}(s) \cdots M_{X_n}(s).$$

### Example

Let  $X_1, \dots, X_n$  be independent **Bernoulli random variables** with parameter  $p$ .

Then  $M_{X_i}(s) = (1-p)e^{0s} + pe^{1s} = 1-p+pe^s$ ,  $i = 1, \dots, n$ .

The random variable  $Z = X_1 + \dots + X_n$  is **binomial** with parameters  $n$  and  $p$ .

Its generating function is  $M_Z(s) = (1-p+pe^s)^n$ .

# Sum of random variables – moment generating function

## Example

Let  $X$  and  $Y$  be independent **Poisson random variables** with parameters  $\lambda_1$  and  $\lambda_2$  respectively. Let  $Z = X + Y$ .

Then

$$M_Z(s) = M_X(s)M_Y(s) = e^{\lambda_1(e^s-1)}e^{\lambda_2(e^s-1)} = e^{(\lambda_1+\lambda_2)(e^s-1)}.$$

$Z$  is again a Poisson random variable, this time with the parameter  $\lambda_1 + \lambda_2$ :

$$P(Z = k) = \frac{(\lambda_1 + \lambda_2)^k}{k!} e^{-(\lambda_1 + \lambda_2)}.$$

Compare with the difficulty of a direct computation of the convolution.

# Summary

**Joint distribution function of a random vector  $(X, Y)$ :**

$$F_{X,Y}(x, y) = P(X \leq x \cap Y \leq y).$$

**Discrete random variables  $X$  and  $Y$**

**Continuous random variables  $X$  and  $Y$**

**Joint probabilities of values / density:**

$$P(X = x \cap Y = y)$$

$$f_{X,Y}(x, y)$$

**Marginal probabilities / density of  $X$ :**

$$P(X = x) = \sum_y P(X = x \cap Y = y)$$

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

$$P(Y = y) = \sum_x P(X = x \cap Y = y)$$

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

**Independence of  $X$  and  $Y$ :**

$$P(X = x \cap Y = y) = P(X = x)P(Y = y)$$

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

**Covariance of  $X$  and  $Y$ :**

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

$X$  and  $Y$  are called non-correlated whenever  $\text{cov}(X, Y) = 0$ .

If  $X$  and  $Y$  are independent, then they are also non-correlated.