BIE-PST – Probability and Statistics

Lecture 7: Random vectors II.

Winter semester 2023/2024

Lecturer:
Francesco Dolce

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

Table of contents

7 Random vectors

7.1 Functions of random vectors

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

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

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

$$F_Z(z) = \mathrm{P}(Z \le z) = \sum_{\{\boldsymbol{x} \in \mathbb{R}^n :\, h(\boldsymbol{x}) \le z\}} \mathrm{P}(\boldsymbol{X} = \boldsymbol{x}).$$

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

$$F_Z(z) = \mathrm{P}(Z \le z) = \underset{\{\boldsymbol{x} \in \mathbb{R}^n :\, h(\boldsymbol{x}) \le z\}}{\int \cdots \int} f_{\boldsymbol{X}}(\boldsymbol{x})\, \mathrm{d}x_1 \dots \mathrm{d}x_n.$$

Expected value of the function of a random vector

The expected value $\mathrm{E}\, h(X, Y)$ 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

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

  if the sum converges absolutely.

- For X and Y continuous random variables it holds that

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

  if the integral converges absolutely.

Now we can prove the linearity of the expectation.

Theorem 7.1 (– linearity of expectation)**.** *For all $a, b \in \mathbb{R}$ and all random variables X and Y it holds that*

$$\mathrm{E}(aX + bY) = a\, \mathrm{E}\, X + b\, \mathrm{E}\, Y.$$

Consequence:

- $\mathrm{E}(aX + b) = a\, \mathrm{E}\, X + b$. This statement was proven before separately.

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

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

For continuous X and Y the proof is analogous:

$$\begin{aligned}
\mathrm{E}(aX+bY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(ax+by)f_{X,Y}(x,y)\ \mathrm{d}x\ \mathrm{d}y\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} axf_{X,Y}(x,y)\ ,\mathrm{d}x\ \mathrm{d}y+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} byf_{X,Y}(x,y)\ \mathrm{d}x\ \mathrm{d}y\\
&= a\int_{-\infty}^{\infty} x\left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)\ \mathrm{d}y\right)\mathrm{d}x\ +\ b\int_{-\infty}^{\infty} y\left(\int_{-\infty}^{\infty} bf_{X,Y}(x,y)\ \mathrm{d}x\right)\mathrm{d}y\\
&= a\int_{-\infty}^{\infty} xf_X(x)\ \mathrm{d}x+b\int_{-\infty}^{\infty} yf_Y(y)\ \mathrm{d}y\\
&= a\,\mathrm{E}\,X+b\,\mathrm{E}\,Y.
\end{aligned}$$

□

7.2 Covariance and correlation

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

Definition 7.2. Let X and Y be random variables with finite second moments. Then we define the *covariance* of the random variables X and Y as

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

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

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

Definition 7.3. Two random variables X and Y are called *non-correlated* if $\operatorname{cov}(X,Y)=0$.

Theorem 7.4. *For the covariance and the correlation coefficient the following properties hold:*

i) $\operatorname{cov}(X, Y) = \mathrm{E}\, XY - \mathrm{E}\, X\, \mathrm{E}\, Y$,

ii) *X and Y are non-correlated if and only if* $\mathrm{E}\, XY = \mathrm{E}\, X\, \mathrm{E}\, Y$,

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)
$$\begin{aligned}\operatorname{cov}(X, Y) &= \mathrm{E}\left((X - \mathrm{E}\, X)(Y - \mathrm{E}\, Y)\right) = \mathrm{E}\left(XY - X\,\mathrm{E}\, Y - Y\,\mathrm{E}\, X + \mathrm{E}\, X\,\mathrm{E}\, Y\right)\\ &= \mathrm{E}\, XY - \mathrm{E}(X\,\mathrm{E}\, Y) - \mathrm{E}(Y\,\mathrm{E}\, X) + \mathrm{E}(\mathrm{E}\, X\,\mathrm{E}\, Y)\\ &= \mathrm{E}\, XY - \mathrm{E}\, X\,\mathrm{E}\, Y - \mathrm{E}\, Y\,\mathrm{E}\, X + \mathrm{E}\, X\,\mathrm{E}\, Y\\ &= \mathrm{E}\, XY - \mathrm{E}\, X\,\mathrm{E}\, Y\end{aligned}$$

ii) Obvious from above. If $\operatorname{cov}(X, Y) = 0$, it means that $\mathrm{E}\, XY - \mathrm{E}\, X\,\mathrm{E}\, Y = 0$, after manipulation we obtain $\mathrm{E}\, XY = \mathrm{E}\, X\,\mathrm{E}\, Y$, which means that the random variables X and Y are non-correlated. Conversely, if X and Y are non-correlated (i.e., $\mathrm{E}\, XY = \mathrm{E}\, X\,\mathrm{E}\, Y$), then $\mathrm{E}\, XY - \mathrm{E}\, X\,\mathrm{E}\, Y = 0$ which means that $\operatorname{cov}(X, Y) = 0$.

iii) From the Schwarz inequality (see bibliography).

iv) Follows straightforwardly by inserting into the definition. Firstly we prepare the quantities $\operatorname{cov}(aX + b, cY + d)$, $\operatorname{var}(aX + b)$ and $\operatorname{var}(cY + d)$:

$$\begin{aligned}\operatorname{cov}(aX + b, cY + d) &= \mathrm{E}[(aX + b - \mathrm{E}(aX + b))(cY + d - \mathrm{E}(cY + d))]\\ &= \mathrm{E}[a(X - \mathrm{E}\, X)c(Y - \mathrm{E}\, Y)] = ac\operatorname{cov}(X, Y),\end{aligned}$$

$$\begin{aligned}\operatorname{var}(aX + b) &= \mathrm{E}(aX + b - \mathrm{E}(aX + b))^2 = \mathrm{E}(a(X - \mathrm{E}\, x))^2 = a^2 \operatorname{var}(X),\\ \operatorname{var}(cY + d) &= c^2 \operatorname{var}(Y).\end{aligned}$$

Inserting them to the definition formula we have

$$\varrho(aX + b, cY + d) = \frac{\operatorname{cov}(aX + b, cY + d)}{\sqrt{\operatorname{var}(aX + b)}\sqrt{\operatorname{var}(cY + d)}} = \frac{ac\operatorname{cov}(X, Y)}{\sqrt{a^2 \operatorname{var}(X)}\sqrt{c^2 \operatorname{var}(Y)}} = \varrho(X, Y).$$

v) Follows from the proof of the Schwarz inequality (see bibliography).

□

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

Definition 7.5. Alternative definition: Two random variables X and Y are called *non-correlated* if

$$\mathrm{E}\, XY = \mathrm{E}\, X\,\mathrm{E}\, Y.$$

Lemma 7.6. *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

$$\mathrm{E}\,XY = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy f_X(x) f_Y(y)\,\mathrm{d}x\,\mathrm{d}y$$
$$= \left(\int_{-\infty}^{+\infty} x f_X(x)\,\mathrm{d}x\right)\left(\int_{-\infty}^{+\infty} y f_Y(y)\,\mathrm{d}y\right) = \mathrm{E}\,X\,\mathrm{E}\,Y.$$

□

It is now possible to obtain the following properties of the variance of sums of two random variables. We recall two formerly mentioned properties of variance and add a theorem about variance of sum of random variables.

Theorem 7.7. *i) For X and Y with finite second moments:*

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

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

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

Proof.

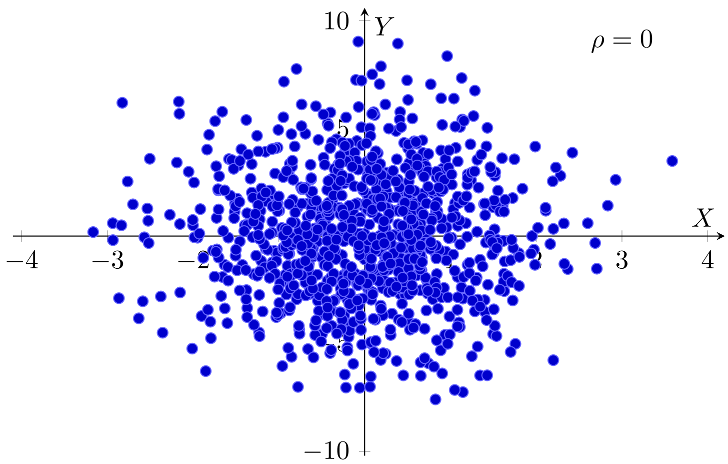
i) Given two random variables X and Y we have:

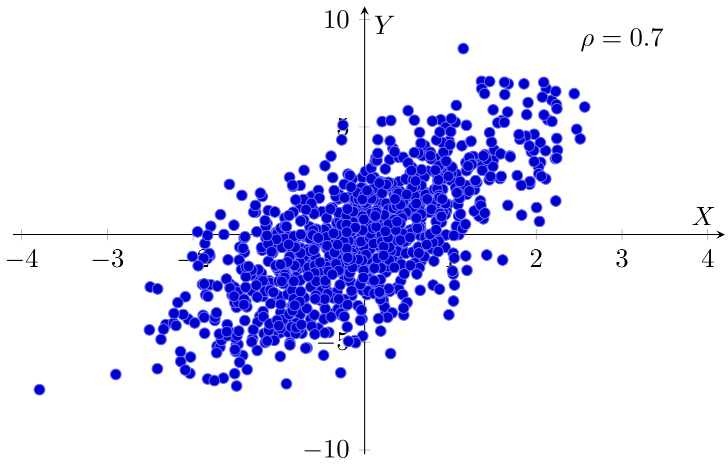
$$\begin{aligned}\operatorname{var}(X \pm Y) &= \mathrm{E}(X \pm Y)^2 - (\mathrm{E}(X \pm Y))^2 = \mathrm{E}(X^2 \pm 2XY + Y^2) - (\mathrm{E}\,X \pm \mathrm{E}\,Y)^2 \\ &= \mathrm{E}\,X^2 \pm 2\,\mathrm{E}\,XY + \mathrm{E}\,Y^2 - (\mathrm{E}\,X)^2 \mp 2\,\mathrm{E}\,X\,\mathrm{E}\,Y - (\mathrm{E}\,Y)^2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

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

□

Correlation – sample of 1000 values





7.3 Sums of random variables – convolution

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

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

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

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

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

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x)\,\mathrm{d}x \quad (\textit{convolution of } \boldsymbol{f_X} \textit{ and } \boldsymbol{f_Y}).$$

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

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

For continuous independent X and Y we have:

$$
\begin{aligned}
F_Z(z) = \mathrm{P}(X+Y \le z) &= \iint\limits_{\{(x,y):\, x+y\le z\}} f_{X,Y}(x,y)\,\mathrm{d}\,(x,y) \\
&= \int\limits_{-\infty}^{\infty} \left(\int\limits_{-\infty}^{z-x} f_{X,Y}(x,y)\,\mathrm{d}y \right) \mathrm{d}x \\
&\overset{y=u-x}{=} \int\limits_{-\infty}^{\infty} \left(\int\limits_{-\infty}^{z} f_{X,Y}(x,u-x)\,\mathrm{d}u \right) \mathrm{d}x \\
&= \int\limits_{-\infty}^{z} \left(\int\limits_{-\infty}^{\infty} f_{X,Y}(x,u-x)\,\mathrm{d}x \right) \mathrm{d}u \\
&= \int\limits_{-\infty}^{z} \left(\int\limits_{-\infty}^{\infty} f_X(x) f_Y(u-x)\,\mathrm{d}x \right) \mathrm{d}u.
\end{aligned}
$$

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

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

Example 7.8 (– sum of two normal distributions)**.** Suppose that X and Y are independent, both having the normal distribution $\mathrm{N}(\mu, 1)$. We want to obtain the distribution of $Z = X+Y$.

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

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

The density of the sum is obtained using convolution:

$$
\begin{aligned}
f_Z(z) = \int\limits_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,\mathrm{d}x &= \int\limits_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-x-\mu)^2}{2}}\,\mathrm{d}x \\
&= \int\limits_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}\left((x-\mu)^2 + (z-x-\mu)^2\right)}\,\mathrm{d}x.
\end{aligned}
$$

The expressions in the exponent can be rewritten as:

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

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:

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

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

Example 7.9. Consider two independent random variables X and Y with the Poisson distribution with parameters λ_1 and λ_2, respectively. Find the distribution of the variable $Z = X + Y$.

$$\mathrm{P}(X = j) = \frac{\lambda_1^j}{j} e^{-\lambda_1} \qquad \mathrm{P}(Y = \ell) = \frac{\lambda_2^\ell}{\ell} e^{-\lambda_2}, \qquad j, \ell = 0, 1, \ldots$$

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

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

✓ *An easier way is to use the 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

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

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

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

Example 7.10. Let $X_1, \ldots, X_n$ be independent *Bernoulli random variables* with parameter p.

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

The random variable $Z = X_1 + \cdots + X_n$ is *binomial* with parameters n and p.
Its generating function is $M_Z(s) = \left(1 - p + pe^s\right)^n$.

Example 7.11. Let X and Y be independent *Poisson random variables* with parameters λ_1 and λ_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 $\lambda_1 + \lambda_2$:

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

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