PS
(seminar 11)

1. Fie X o variabilă cu valorile $-2, -1, 0, 1, 2$, fiecare valoare fiind luată cu probabilitate $p = \frac{1}{5}$. Se consideră $Y = X^2$. Să se arate că $cov(X, Y) = 0$, dar variabilele X şi Y nu sunt independente.

Rezolvare: Distribuţia de probabilitate a vectorului (X,Y) este:

		Y			$\sum p_i$
		0	1	4	
	-2	0	0	$\frac{1}{5}$	$\frac{1}{5}$
X	-1	0	$\frac{1}{5}$	0	$\frac{1}{5}$
	0	$\frac{1}{5}$	0	0	$\frac{1}{5}$
	1	0	$\frac{1}{5}$	0	$\frac{1}{5}$
	2	0	0	$\frac{1}{5}$	$\frac{1}{5}$
$\sum q_i$		$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$	

$$X = \begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix} \quad M(X) = -\frac{2}{5} - \frac{1}{5} + 0 + \frac{1}{5} + \frac{2}{5} = 0$$

$$Y = \begin{pmatrix} 0 & 1 & 4 \\ \frac{1}{5} & \frac{2}{5} & \frac{2}{5} \end{pmatrix} \quad M(Y) = 0 + \frac{2}{5} + \frac{8}{5} = \frac{10}{5} = 2$$

$$P(X = -2, Y = 0) = P(X = -2) \cdot P(Y = 0)$$

$$0 = \frac{1}{5} \cdot \frac{1}{5} = \frac{1}{25} \text{ fals}$$

$\Rightarrow X, Y$ – dependente

$$\Rightarrow cov(X, Y) = M(XY) - M(X) \cdot M(Y)$$

$$= 0 - 0 = 0 \checkmark$$

$$XY = \begin{pmatrix} -8 & -4 & -2 & -1 & 0 & 1 & 2 & 4 & 8 \\ \frac{1}{5} & 0 & 0 & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & 0 & 0 & \frac{1}{5} \end{pmatrix}$$

$$P(X = -2, Y = 4) = \frac{1}{5}$$

$$M(X \cdot Y) = -\frac{8}{5} - \frac{1}{5} + \frac{1}{5} + \frac{8}{5} = 0$$

2. Fie X o variabilă aleatoare ce are media $M(X) = 3$ şi dispersia $\sigma^2(X) = 1$, iar $Y = -2X + 5$. Să se calculeze covarianţa şi coeficientul de corelaţie pentru variabilele X, Y. Să se determine matricea de covarianţă asociată vectorului (X, Y).

- Dacă între variabilele aleatoare X şi Y există o relaţie liniară de forma

$$Y = aX + b, \ a, b \in \mathbb{R}, \ a \neq 0,$$

atunci

$$\rho(X,Y) = \begin{cases} -1, & \text{dacă } a < 0, \\ 1, & \text{dacă } a > 0. \end{cases}$$

$a < 0 \Rightarrow \rho(X,Y) = -1$

coef de corelatie

- Reciproc, dacă $|\rho(X,Y)| = 1$, atunci între ele există o relaţie liniară,

$$Y = aX + b, \quad a \neq 0$$

Legătura dintre două variabile X şi Y poate fi determinată folosind coeficientul de corelaţie astfel:

- $\rho(X,Y) = 0$, atunci X şi Y sunt **necorelate**;
- $\rho(X,Y)$ este apropiat de zero, atunci X şi Y sunt **slab corelate** (intensitatea legăturii dintre ele este redusă);
- $\rho(X,Y) = 1$, atunci $Y = aX + b, \quad a > 0$, X şi Y sunt **pozitiv corelate**;
- $\rho(X,Y) = -1$, atunci $Y = aX + b, \quad a < 0$, X şi Y sunt **negativ corelate**;
- $|\rho(X,Y)|$ are o valoare apropiată de 1, **relaţia dintre variabilele aleatoare este "aproape liniară"**, adică valorile (x, y) ale vectorului aleator (X, Y) sunt uşor dispersate în jurul unei drepte de ecuaţie $y = ax + b$.

$$\rho(X,Y) = \frac{cov(X,Y)}{\sigma(X) \cdot \sigma(Y)} = -1$$

$$\Rightarrow cov(X,Y) = -\sigma(X) \cdot \sigma(Y) = -(1)(2) = -2$$

$$\underbrace{\sigma^2(X)}_{\text{dispersie}} = 1 \qquad M(X) = 3 \qquad \boxed{M(X^2) - 9}$$

$$\sigma^2(Y) = \sigma^2(-2X + 5) = (-2)^2 \sigma^2(X) = 4 \cdot 1 = 4$$

Matricea de covarianţă a vectorului aleator $\mathbf{X} = (X_1, X_2, \ldots, X_n)^T$ este matricea notată cu Σ, ale cărei elemente sunt $\sigma_{ij} = \text{cov}(X_i, X_j)$, $i, j \in \{1, \ldots, n\}$. **Observaţii:**

- $\sigma_{ii} = cov(X_i, X_i) = \sigma^2(X_i)$
- Σ este simetrică şi semipozitiv definită
- $\Sigma = M(\mathbf{YY}^T)$, unde $\mathbf{Y} = \mathbf{X} - \mathbf{m} = (X_1 - m_1, X_2 - m_2, \ldots, X_n - m_n)^T$ iar $M(\mathbf{YY^T})$ notează matricea mediilor elementelor matricii $\mathbf{YY^T}$.

$$\Sigma = \begin{pmatrix} \sigma^2(X) & cov(X,Y) \\ cov(X,Y) & \sigma^2(Y) \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}$$

3. Fie X şi Y două variabile aleatoare independente, $X, Y \sim N(0,1)$. Se consideră $Z = 1 + X + XY^2$ şi $W = 1 + X$. Să se calculeze $cov(Z, W)$.

$M = 0 \quad \sigma^2 = 1$

$$cov(Z,W) = cov(1+X+XY^2, \ 1+X) =$$

$$= cov(X+XY^2, X) =$$

$$= cov(X,X) + cov(XY^2, X) =$$

$$= \sigma^2(X) + M(X^2Y^2) - M(XY^2)M(X)$$

$$= \sigma^2(X) + M(X^2)M(Y^2) - M(X)^2 M(Y^2) =$$

$$= 1 + 1 \cdot 1 - 0 \cdot 1 =$$

$$= 2$$

Covarianţa unui vector (X, Y) se defineşte ca fiind

$$cov(X,Y) := M[(X - M(X))(Y - M(Y))]$$

Covarianţa are următoarele proprietăţi:

- $cov(X,Y) = M(XY) - M(X)M(Y)$
- $cov(X,X) = \sigma^2(X)$;
- $cov(X,a) = 0$
- $cov(X,Y) = cov(Y,X)$;
- $cov(aX,Y) = a\,cov(X,Y)$, $a \in \mathbb{R}$;
- $cov(X+Y,Z) = cov(X,Z) + cov(Y,Z)$;
- $\sigma^2(X+Y) = \sigma^2(X) + \sigma^2(Y) + 2cov(X,Y)$;
- Dacă în plus X şi Y sunt variabile aleatoare independente, atunci
$$\sigma^2(X+Y) = \sigma^2(X) + \sigma^2(Y).$$
- X, Y sunt independente, atunci $cov(X,Y) = 0$ (**Reciproca nu este adevărată!**)

$M(X^2) - M(X)^2$

$\boxed{\sigma^2(X) = M(X^2)}$ 0

$= 1$

Analog $M(Y^2) = 1$

itemize **Coeficientul de corelaţie** a două variabile aleatoare X şi Y, de abateri standard nenule σ_X, σ_Y, este

$$\rho(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{cov(X,Y)}{\sigma(X)\sigma(Y)} \in [-1, 1]$$

4. Se consideră vectorul aleator continuu (X, Y) cu densitatea de probabilitate

$$f(x,y) = \begin{cases} 2 & \text{dacă } 0 \le y \le x \le 1 \\ 0 & \text{dacă altfel} \end{cases}$$

$cov(X,Y) = ?$

$$cov(X,Y) = M(XY) - M(X) \cdot M(Y)$$

$$f_X(x) = \int_{-\infty \Rightarrow 0}^{\infty \Rightarrow x} 2\,dy = 2y\Big|_0^x = 2x \Rightarrow M(X) = \int_0^1 2x^2\,dx = \frac{2}{3}$$

$$f_Y(y) = \int_y^1 2\,dx = 2x\Big|_y^1 = 2 - 2y \Rightarrow M(Y) = \int_0^1 2y - 2y^2 =$$

$$= \frac{^{3)}2}{2} - \frac{2}{3} = \frac{1}{3}$$

$$M(XY) = \int_0^1\int_0^x xy\,2\,dy\,dx = \int_0^1 x^3\,dx = \frac{1}{4}$$

$$\Rightarrow cov(X,Y) = \frac{^{9)}1}{4} - \frac{^{4)}2}{3} \cdot \frac{1}{3} = \frac{1}{36}$$