Náhodný vektor a jeho charakteristiky

V následující kapitole budeme věnovat pozornost pouze dvourozměrnému náhodnému vektoru, i když uvedené pojmy a jejich vlastnosti lze přirozeně zobecnit i na vícerozměrné náhodné vektory.

Představte si, že provádíte náhodný pokus, jehož výsledek ohodnotíte dvojicí čísel, tj. dvourozměrným vektorem. Před provedením pokusu jeho výsledek a tedy ani sledované hodnoty neznáte. Proto je vektor, který připisuje výsledku náhodnému pokusu vámi sledované hodnoty, označován jako náhodný vektor. Náhodný vektor značíme velkým tučným písmenem, např. $\boldsymbol{X}$ a jeho složky značíme X_1 a X_2. Množinu možných hodnot náhodného vektoru nazýváme obor hodnot náhodného vektoru $\boldsymbol{X}$ a značíme jej $\boldsymbol{\mathcal{X}}$. Poté, co je pokus proveden, je naměřená hodnota náhodného vektoru značena malým tučným písmenem, např. $\boldsymbol{x} = [21\text{mm}, 5\text{g}]^{\mathrm{T}}$. Náhodným vektorem může být například

- délka a váha vysoustružené součástky
- počet osob a doba čekání ve frontě
- dnešní teplota a atmosferický tlak
- počet výrobků ze zásilky, které již nelze opravit a které jsou opravitelné

Uveďmě nyní matematicky přesnější popis náhodného vektoru. Z důvodu přístupnosti látky studentům budou pojmy a vlastnosti v tomto textu zavedeny ve stejné podstatě nicméně s mírnými odchylkami od přesných matematických formulací.

1. Pojem **Dvourozměrným náhodným vektorem** (vzhledem k $\mathcal{A}$) rozumíme vektor $\boldsymbol{X} = [X_1, X_2]^{\mathrm{T}}$, jehož složky X_1, X_2 jsou náhodné veličiny na stejném základním prostoru $(\Omega, \mathcal{A})$. Obor hodnot náhodného vektoru $\boldsymbol{X}$ značíme $\boldsymbol{\mathcal{X}}$. Realizaci náhodného vektoru, tj. $\boldsymbol{X}(\omega)$, $\omega \in \Omega$, značíme $\boldsymbol{x}$, popřípadě ve složkách $[x_1, x_2]^{\mathrm{T}}$.

2. Pojem Reálnou funkci $F : (-\infty, \infty) \times (-\infty, \infty) \to \langle 0, 1\rangle$ definovanou předpisem

$$F(x_1, x_2) = \mathrm{P}(X_1 < x_1, X_2 < x_2)$$

nazýváme **(simultánní, sdružená) distribuční funkce** dvourozměrného náhodného vektoru $\boldsymbol{X} = [X_1, X_2]^{\mathrm{T}}$, popřípadě **simultánní (sdružená) distribuční funkce** náhodných veličin X_1 a X_2.

3. Pojem Řekneme, že náhodný vektor je **diskrétní**, resp. má **diskrétní rozdělení pravděpodobnosti**, je-li jeho oborem hodnot nejvýše spočetná množina $\boldsymbol{\mathcal{X}} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots\}$, tj. nabývá nejvýše spočetně mnoha hodnot $\boldsymbol{x}_1 = [x_{11}, x_{12}]^{\mathrm{T}}, \boldsymbol{x}_2 = [x_{21}, x_{22}]^{\mathrm{T}}, \boldsymbol{x}_3 = [x_{31}, x_{32}]^{\mathrm{T}}, \ldots,$ tak, že

$$\sum_{i=1}^{\infty} \mathrm{P}(\boldsymbol{X} = \boldsymbol{x}_i) = \sum_{i=1}^{\infty} \mathrm{P}(X_1 = x_{i1}, X_2 = x_{i2}) = 1.$$

4. Pojem **Simultánní (sdružená) pravděpodobnostní funkce** diskrétního náhodného vektoru $\boldsymbol{X}$ je funkce $p : (-\infty, \infty) \times (-\infty, \infty) \to \langle 0, 1\rangle$ daná předpisem

$$p(x_1, x_2) = \mathrm{P}(X_1 = x_1, X_2 = x_2).$$

5. Příklad Pravděpodobnost, že při přenosu digitální informace dojde k silné, resp. střední, resp. žádné, deformaci bitu je 0.1, 0.3 a 0.6. Předpokládejme, že jsou přeneseny dva bity a rozsah deformace je pro každý bit nezávislý. Náhodný vektor $\boldsymbol{X} = [X_1, X_2]^{\mathrm{T}}$ udává počet bitů se silnou (X_1) a střední (X_2) deformací.

Oborem hodnot náhodného vektoru $\boldsymbol{X}$ je množina $\boldsymbol{\mathcal{X}} = \{[0,0]^{\mathrm{T}}, [0,1]^{\mathrm{T}}, [0,2]^{\mathrm{T}}, [1,0]^{\mathrm{T}}, [1,1]^{\mathrm{T}}, [2,0]^{\mathrm{T}}\}$.

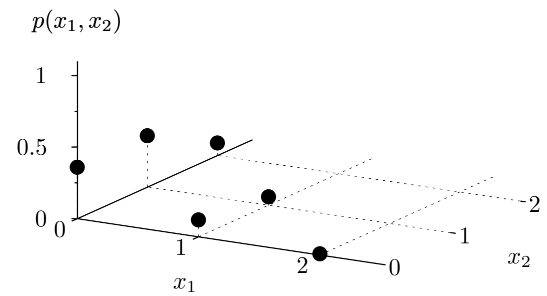
Nejprve vypočteme pravděpodobnostní funkci

$$p(x_1,x_2)=\begin{cases} \mathrm{P}(X_1=0,X_2=0)=0.1^0\cdot 0.3^0\cdot 0.6^2=0.36 & \text{pro } [x_1,x_2]^{\mathrm{T}}=[0,0]^{\mathrm{T}}\\ \mathrm{P}(X_1=0,X_2=1)=2\cdot 0.1^0\cdot 0.3^1\cdot 0.6^1=0.36 & \text{pro } [x_1,x_2]^{\mathrm{T}}=[0,1]^{\mathrm{T}}\\ \mathrm{P}(X_1=0,X_2=2)=0.1^0\cdot 0.3^2\cdot 0.6^0=0.09 & \text{pro } [x_1,x_2]^{\mathrm{T}}=[0,2]^{\mathrm{T}}\\ \mathrm{P}(X_1=1,X_2=0)=2\cdot 0.1^1\cdot 0.3^0\cdot 0.6^1=0.12 & \text{pro } [x_1,x_2]^{\mathrm{T}}=[1,0]^{\mathrm{T}}\\ \mathrm{P}(X_1=1,X_2=1)=2\cdot 0.1^1\cdot 0.3^1\cdot 0.6^0=0.06 & \text{pro } [x_1,x_2]^{\mathrm{T}}=[1,1]^{\mathrm{T}}\\ \mathrm{P}(X_1=2,X_2=0)=0.1^2\cdot 0.3^0\cdot 0.6^0=0.01 & \text{pro } [x_1,x_2]^{\mathrm{T}}=[2,0]^{\mathrm{T}}\\ 0 & \text{jinak}\end{cases}$$

Výsledek můžeme zapsat do pravděpodobnostní tabulky 1 a znázornit graficky, viz. obrázek 1.

$x_2\backslash x_1$	0	1	2
0	0.36	0.12	0.01
1	0.36	0.06	0
2	0.09	0	0

Tabulka 1: Pravděpodobnostní tabulka náhodného vektoru $\boldsymbol{X}=[X_1,X_2]^{\mathrm{T}}$ z příkladu 5



Obrázek 1: Graf nenulových hodnot pravděpodobnostní funkce náhodného vektoru $\boldsymbol{X}=[X_1,X_2]^{\mathrm{T}}$ z příkladu 5 (V ostatních bodech je pravděpodobnostní funkce nulová.)

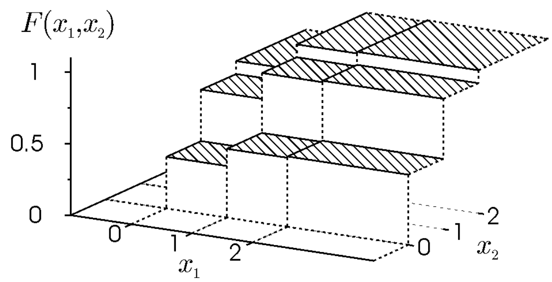
Z pravděpodobnostní funkce odvodíme distribuční funkci, která je pro přehlednost zapsána v Tabulce 2 a znázorněna graficky v Obrázku 2.

$x_2\backslash x_1$	$(-\infty,0\rangle$	$(0,1\rangle$	$(1,2\rangle$	$(2,\infty)$
$(-\infty,0\rangle$	0	0	0	0
$(0,1\rangle$	0	0.36	0.48	0.49
$(1,2\rangle$	0	0.72	0.9	0.91
$(2,\infty)$	0	0.81	0.99	1

Tabulka 2: Distribuční funkce $F(x_1,x_2)$ náhodného vektoru $\boldsymbol{X}=[X_1,X_2]^{\mathrm{T}}$ z příkladu 5

Postup výpočtu hodnot distribuční funkce $F(x_1,x_2)$ z pravděpodobnostní funkce $p(x_1,x_2)$ ukážeme na $F(1.5,0.5)$, tj. na výpočtu hodnoty distribuční funkce na intervalu $(1,2\rangle\times(0,1\rangle$.

$$\begin{aligned} F(1.5,0.5) &= \mathrm{P}(X_1<1.5,X_2<0.5)=\\ &= \mathrm{P}((X_1=0\vee X_1=1)\wedge(X_2=0))=\\ &= \mathrm{P}((X_1=0\wedge X_2=0)\vee(X_1=1\wedge X_2=0))=\\ &= p(0,0)+p(1,0)=0.36+0.12=0.48\end{aligned}$$



Obrázek 2: Graf simultánní distribuční funkce náhodného vektoru $\boldsymbol{X} = [X_1, X_2]^{\mathrm{T}}$ z příkladu 5

6. Pojem Řekneme, že náhodný vektor je **spojitý**, resp. má **spojité rozdělení pravděpodobnosti**, jestliže existuje nezáporná funkce $f(x_1, x_2)$ taková, že

$$F(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f(s,t) \mathrm{d}t \mathrm{d}s.$$

Funkci $f(x_1, x_2)$ nazýváme sdruženou (simultánní) hustotou náhodného vektoru $\boldsymbol{X}$.

7. Vlastnosti Pro simultánní hustotu náhodného vektoru $\boldsymbol{X}$ platí

1. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) \mathrm{d}x_2 \mathrm{d}x_1 = 1$
2. $f(x_1, x_2) = \frac{\partial^2}{\partial x_1 \partial x_2} F(x_1, x_2)$.
3. $\mathrm{P}(\boldsymbol{X} \in M) = \iint\limits_{M} f(x_1, x_2) \mathrm{d}x_2 \mathrm{d}x_1$ pro libovolnou množinu M v $\mathbb{R}^2$

Marginální rozdělení

V souvislosti s rozdělením náhodného vektoru $\boldsymbol{X}$ nazýváme rozdělení jeho složek marginálními rozděleními náhodného vektoru $\boldsymbol{X}$.

8. Pojem **Marginálními distribučními funkcemi** náhodného vektoru $\boldsymbol{X} = [X_1, X_2]^{\mathrm{T}}$ rozumíme distribuční funkce náhodných veličin X_1, X_2, tj.

$$\begin{aligned} F_{X_1}(x_1) &= \mathrm{P}(X_1 < x_1), \\ F_{X_2}(x_2) &= \mathrm{P}(X_2 < x_2). \end{aligned}$$

Pro odlišení distribučních funkcí náhodných veličin X_1 a X_2 byl u předchozího pojmu užit dolní index. Bude-li to nutné, budou stejným způsobem odlišeny i jiné funkční a číselné charakteristiky náhodných veličin X_1 a X_2.

9. Vlastnosti Z definice simultánní distribuční funkce náhodných veličin X_1, X_2 plyne

$$\begin{aligned} F_{X_1}(x_1) &= \mathrm{P}(X_1 < x_1, X_2 < \infty) = \lim_{x_2 \to \infty} F(x_1, x_2), \\ F_{X_2}(x_2) &= \mathrm{P}(X_1 < \infty, X_2 < x_2) = \lim_{x_1 \to \infty} F(x_1, x_2) \end{aligned}$$

10. Pojem Řekneme, že dvě náhodné veličiny jsou **nezávislé**, jestliže

$$F(x_1, x_2) = F_{X_1}(x_1) F_{X_1}(x_2).$$

11. Vlastnosti Diskrétní náhodné veličiny X_1, X_2 jsou nezávislé právě tehdy, když

$$p(x_1, x_2) = p_{X_1}(x_1)p_{X_1}(x_2), \qquad \text{pro všechna } x_1, x_2 \in (-\infty, \infty).$$

Spojité náhodné veličiny X_1, X_2 jsou nezávislé právě tehdy, když

$$f(x_1, x_2) = f_{X_1}(x_1)f_{X_1}(x_2), \qquad \text{pro všechna } x_1, x_2 \in (-\infty, \infty).$$

12. Pojem **Marginálními pravděpodobnostními funkcemi** diskrétního náhodného vektoru $\boldsymbol{X} = [X_1, X_2]^{\mathrm{T}}$ rozumíme pravděpodobnostní funkce náhodných veličin X_1, X_2, tj.

$$\begin{aligned} p_{X_1}(x_1) &= \mathrm{P}(X_1 = x_1), \\ p_{X_2}(x_2) &= \mathrm{P}(X_2 = x_2). \end{aligned}$$

13. Vlastnosti Z definice simultánní pravděpodobnostní funkce náhodných veličin X_1, X_2 plyne

$$\begin{aligned} p_{X_1}(x_1) &= \mathrm{P}(X_1 = x_1, X_2 \in (-\infty, \infty)) = \sum_{x_2=-\infty}^{\infty} p(x_1, x_2), \\ p_{X_2}(x_2) &= \mathrm{P}(X_1 \in (-\infty, \infty), X_2 = x_2) = \sum_{x_1=-\infty}^{\infty} p(x_1, x_2). \end{aligned}$$

14. Pojem **Marginálními hustotami** spojitého náhodného vektoru $\boldsymbol{X} = [X_1, X_2]^{\mathrm{T}}$ rozumíme hustoty náhodných veličin X_1, X_2, které lze vypočíst jako

$$\begin{aligned} f_{X_1}(x_1) &= \int_{-\infty}^{\infty} f(x_1, x_2)\mathrm{d}x_2, \\ f_{X_2}(x_2) &= \int_{-\infty}^{\infty} f(x_1, x_2)\mathrm{d}x_1. \end{aligned}$$

15. Příklad Vypočtěme marginální pravděpodobnostní funkce a marginální distribuční funkce náhodného vektoru $\boldsymbol{X}$ z příkladu 5. Výpočet marginálních pravděpodobnostních funkcí náhodných veličin X_1 a X_2 užitím vlastnosti 13 lze s výhodou provést přímo v pravděpodobnostní tabulce náhodného vektoru $\boldsymbol{X}$, viz. Tabulka 1. Marginální pravděpodobnostní funkce náhodné veličiny X_1, resp. X_2, je součtem hodnot simultánní pravděpodobnostní funkce ve sloupci, resp. v řádku.

$x_2 \backslash x_1$	0	1	2	$p_{X_2}(x_2)$
0	0.36	0.12	0.01	0.49
1	0.36	0.06	0	0.42
2	0.09	0	0	0.09
$p_{X_1}(x_1)$	0.81	0.18	0.01	

Tabulka 3: Výpočet marginálních pravděpodobnostních funkcí náhodného vektoru $\boldsymbol{X} = [X_1, X_2]^{\mathrm{T}}$ z příkladu 5

Marginální distribuční funkce odvodíme z marginálních pravděpodobnostních funkcí.

$$F_{X_1}(x_1) = \begin{cases} 0 & \text{pro } x_1 \leq 0 \\ \mathrm{P}(X_1 = 0) = 0.81 & \text{pro } 0 < x_1 \leq 1 \\ \mathrm{P}(X_1 = 0) + \mathrm{P}(X_1 = 1) = 0.81 + 0.18 = 0.99 & \text{pro } 1 < x_1 \leq 2 \\ \mathrm{P}(X_1 = 1) + \mathrm{P}(X_1 = 1) + \mathrm{P}(X_1 = 2) = 0.81 + 0.18 + 0.01 = 1 & \text{pro } 2 < x_1 \end{cases}$$

$$F_{X_2}(x_2) = \begin{cases} 0 & \text{pro } x_2 \leq 0 \\ \mathrm{P}(X_2 = 0) = 0.49 & \text{pro } 0 < x_2 \leq 1 \\ \mathrm{P}(X_2 = 0) + \mathrm{P}(X_2 = 1) = 0.49 + 0.42 = 0.91 & \text{pro } 1 < x_2 \leq 2 \\ \mathrm{P}(X_2 = 1) + \mathrm{P}(X_2 = 1) + \mathrm{P}(X_2 = 2) = 0.49 + 0.42 + 0.09 = 1 & \text{pro } 2 < x_2 \end{cases}$$

Číselné charakteristiky náhodného vektoru

16. Pojem **Střední hodnotou** náhodného vektoru rozumíme vektor $\mathrm{E}(\boldsymbol{X}) = [\mathrm{E}(X_1), \mathrm{E}(X_2)]^{\mathrm{T}}$.

17. Pojem **Varianční matice** náhodného vektoru je matice

$$\mathrm{var}(\boldsymbol{X}) = \begin{pmatrix} \mathrm{D}(X_1) & \mathrm{C}(X_1, X_2) \\ \mathrm{C}(X_1, X_2) & \mathrm{D}(X_2) \end{pmatrix},$$

kde

$$\mathrm{C}(X_1, X_2) = \mathrm{E}\left([X_1 - \mathrm{E}(X_1)] \cdot [X_2 - \mathrm{E}(X_2)]\right)$$

Reálné číslo $\mathrm{C}(X_1, X_2)$ nazýváme **kovariancí** náhodných veličin X_1 a X_2.

18. Vlastnosti Pro kovarianční matici náhodného vektoru $\boldsymbol{X}$ platí

1. Matice $\mathrm{var}(\boldsymbol{X})$ je symetrická a pozitivně semidefinitní.
2. $\mathrm{C}(X_1, X_2) = \mathrm{E}(X_1 \cdot X_2) - [\mathrm{E}(X_1) \cdot \mathrm{E}(X_2)]$.
3. $\mathrm{var}(\boldsymbol{a} + \boldsymbol{B}\boldsymbol{X}) = \boldsymbol{B}\mathrm{var}(\boldsymbol{X})\boldsymbol{B}^{\mathrm{T}}$, pro libovolný vektor $\boldsymbol{a} \in \mathbb{R}^2$ a matici $\boldsymbol{B}$ typu $n \times 2$.
4. $\mathrm{C}(a_1 + b_1 X_1, a_2 + b_2 X_2) = b_1 b_2 \mathrm{C}(X_1, X_2)$.
5. $\mathrm{C}(X_1, X_1) = \mathrm{D}(X_1)$.
6. $\mathrm{D}(X_1 + X_2) = \mathrm{D}(X_1) + \mathrm{D}(X_2) + 2\mathrm{C}(X_1, X_2)$.
7. Jsou-li náhodné veličiny X_1 a X_2 nezávislé, pak $\mathrm{C}(X_1, X_2) = 0$.

19. Pojem **Korelační matice** náhodného vektoru $\boldsymbol{X}$ je matice

$$\mathrm{cor}(\boldsymbol{X}) = \begin{pmatrix} 1 & \varrho(X_1, X_2) \\ \varrho(X_1, X_2) & 1 \end{pmatrix},$$

kde

$$\varrho(X_1, X_2) = \begin{cases} \frac{\mathrm{C}(X_1, X_2)}{\sqrt{\mathrm{D}(X_1)\mathrm{D}(X_2)}} & \text{je-li } \mathrm{D}(X_1) \neq 0,\ \mathrm{D}(X_2) \neq 0, \\ 0 & \text{jinak.} \end{cases}$$

Reálné číslo $\varrho(X_1, X_2)$ nazýváme **korelace (korelační koeficient)** náhodných veličin X_1 a X_2.

20. Pojem Jestliže $\mathrm{C}(X_1, X_2) = 0$, říkáme, že jsou náhodné veličiny X_1 a X_2 **nekorelované**.

21. Vlastnosti Pro korelační matici náhodného vektoru $\boldsymbol{X}$ platí

1. Matice $\mathrm{cor}(\boldsymbol{X})$ je symetrická a pozitivně semidefinitní.
2. $\varrho(X_1, X_2) \in \langle -1, 1\rangle$.
3. $\varrho(X_1, a_2) = \varrho(a_1, X_2) = \varrho(a_1, a_2) = 0$ pro libovolná $a_1, a_2 \in \mathbb{R}$.
4. $\varrho(X_1, X_1) = 1$.
5. Jestliže $\varrho = 1$, pak existuje $a, b \in \mathbb{R}$, $b > 0$ takové, že $X_2 = a + bX_1$ s pravděpodobností 1.
6. Jestliže $\varrho = -1$, pak existuje $a, b \in \mathbb{R}$, $b < 0$ takové, že $X_2 = a + bX_1$ s pravděpodobností 1.
7. Jsou-li náhodné veličiny X_1 a X_2 nezávislé, pak jsou nekorelované.

22. Poznámka Zdůrazněme, že pojmy nekorelované a nezávislé náhodné veličiny nejsou ekvivalentní. Jsou-li náhodné veličiny nekorelované, nemusí být nutně nezávislé. Bez přidání dalších předpokladů je pouze jisté, že mezi nekorelovanými náhodnými veličinami neexistuje lineární závislost.

23. Příklad Spočtěme číselné charakteristiky náhodného vektoru $\boldsymbol{X}$ z příkladu 5. Střední hodnotu $\mathrm{E}\boldsymbol{X}$ vypočteme z marginálních pravděpodobnostních funkcí v Tabulce 3.

$$\begin{aligned}\mathrm{E}(X_1) &= 0\cdot 0.81 + 1\cdot 0.18 + 2\cdot 0.01 = 0.2\\ \mathrm{E}(X_2) &= 0\cdot 0.49 + 1\cdot 0.42 + 2\cdot 0.09 = 0.6\end{aligned}$$

Pro rozptyl nejdříve vypočteme střední hodnoty druhých mocnin náhodných veličin X_1, X_2

$$\begin{aligned}\mathrm{E}(X_1^2) &= 0^2\cdot 0.81 + 1^2\cdot 0.18 + 2^2\cdot 0.01 = 0.22\\ \mathrm{E}(X_2^2) &= 0^2\cdot 0.49 + 1^2\cdot 0.42 + 2^2\cdot 0.09 = 0.78\end{aligned}$$

odkud dostáváme

$$\begin{aligned}\mathrm{D}(X_1) &= \mathrm{E}(X_1^2) - [\mathrm{E}(X_1)]^2 = 0.22 - 0.^2 = 0.18\\ \mathrm{D}(X_2) &= \mathrm{E}(X_2^2) - [\mathrm{E}(X_2)]^2 = 0.78 - 0.6^2 = 0.42\end{aligned}$$

Abychom mohli vypočíst kovarianci užitím vztahu 2 v odstavci 18 potřebujeme vyčíslit

$$\begin{aligned}\mathrm{E}(X_1\cdot X_2) &= 0\cdot 0\cdot 0.36 + 0\cdot 1\cdot 0.12 + 0\cdot 2\cdot 0.01 + 1\cdot 0\cdot 0.36 +\\ &\quad +1\cdot 1\cdot 0.06 + 1\cdot 2\cdot 0 + 2\cdot 0\cdot 0.09 + 2\cdot 1\cdot 0 + 2\cdot 2\cdot 0 =\\ &= 0.06\end{aligned}$$

odkud

$$\mathrm{C}(X_1, X_2) = 0.06 - 0.2\cdot 0.6 = -0.06$$

Můžeme tedy napsat varianční matici náhodného vektoru $\boldsymbol{X}$

$$\mathrm{var}(\boldsymbol{X}) = \begin{pmatrix} 0.18 & -0.06\\ -0.06 & 0.42\end{pmatrix}$$

Na závěr stanovme korelační koeficient, a tím i korelační matici. Podle definice

$$\varrho(X_1, X_2) = \frac{\mathrm{C}(X_1, X_2)}{\sqrt{\mathrm{D}(X_1)}\sqrt{\mathrm{D}(X_2)}} = \frac{-0.06}{\sqrt{0.18}\sqrt{0.42}} \doteq -0.218$$

a proto

$$\mathrm{cor}(\boldsymbol{X}) = \begin{pmatrix} 1 & -0.218\\ -0.218 & 1\end{pmatrix}$$