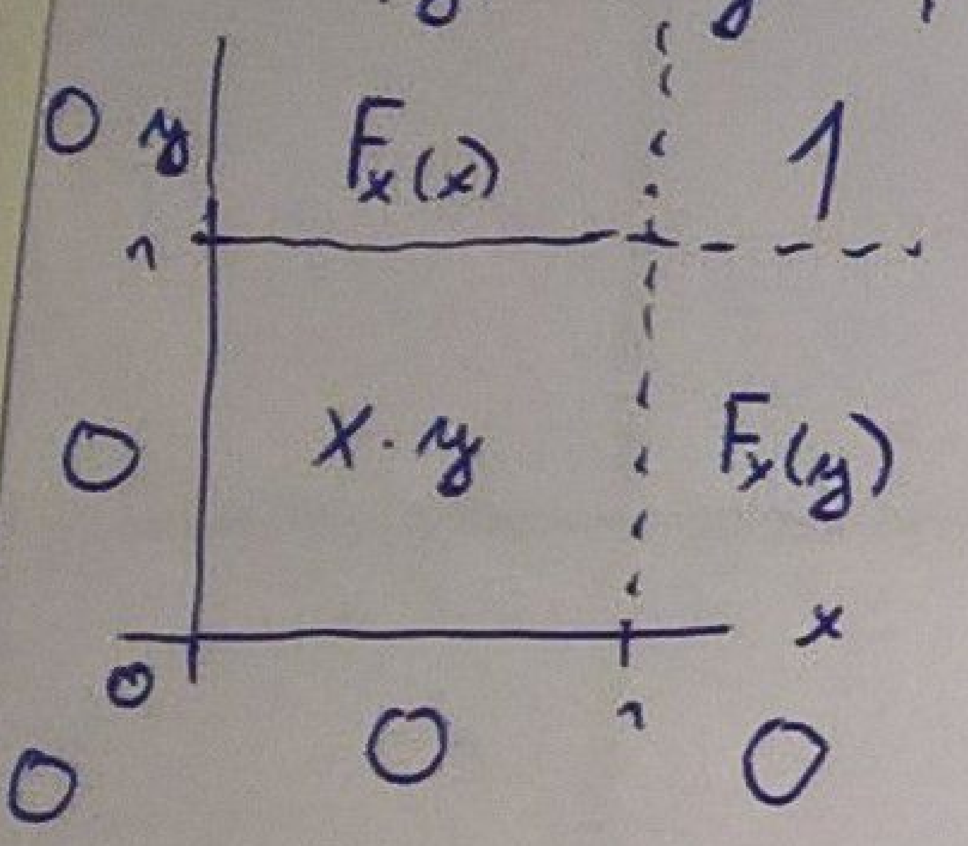
Kopuly

Příklad (motivační)

Nechť náhodný vektor $(X,Y)^T$ má sdruženou (simultánní) distribuční funkci $F(x,y)$. Určete marginální rozdělení jeho složek X a Y.

a) $F(x,y) = x\cdot y \quad , 0\le x\le 1,\ 0\le y\le 1$

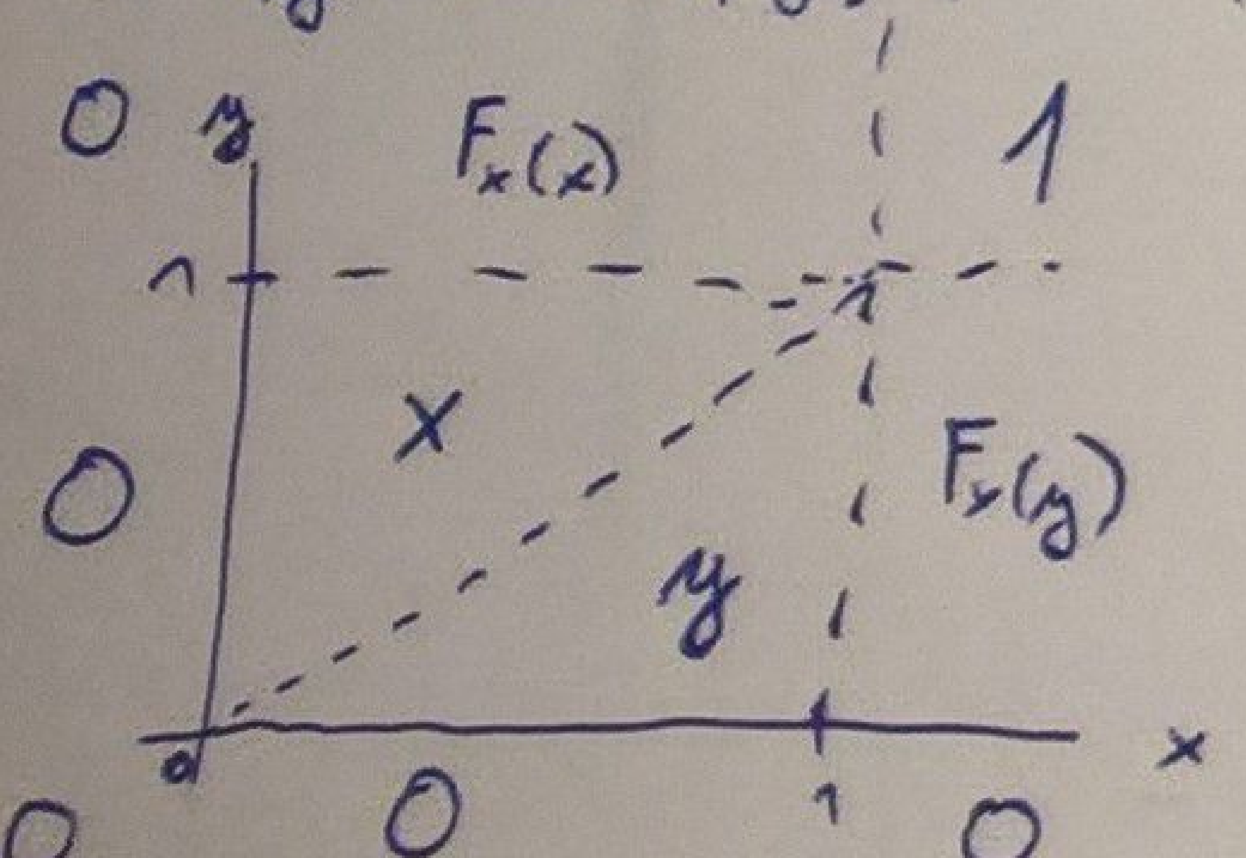


$$F_X(x) = \lim_{y\to\infty} F(x,y) = \lim_{y\to 1} x\cdot y = x \quad , 0\le x\le 1$$
$$= 0 \quad , x<0$$
$$= 1 \quad , x\ge 1$$

$$F_X(x) = \begin{cases} 0 & , x<0 \\ x & , 0\le x\le 1 \\ 1 & , x>1 \end{cases} \qquad X\sim Rs(0,1)$$

a analogicky $F_Y(y) = \begin{cases} 0 & , y<0 \\ y & , 0\le y\le 1 \\ 1 & , y>1 \end{cases}$ $\quad Y\sim Rs(0,1)$. Navíc X a Y jsou nezávislé.

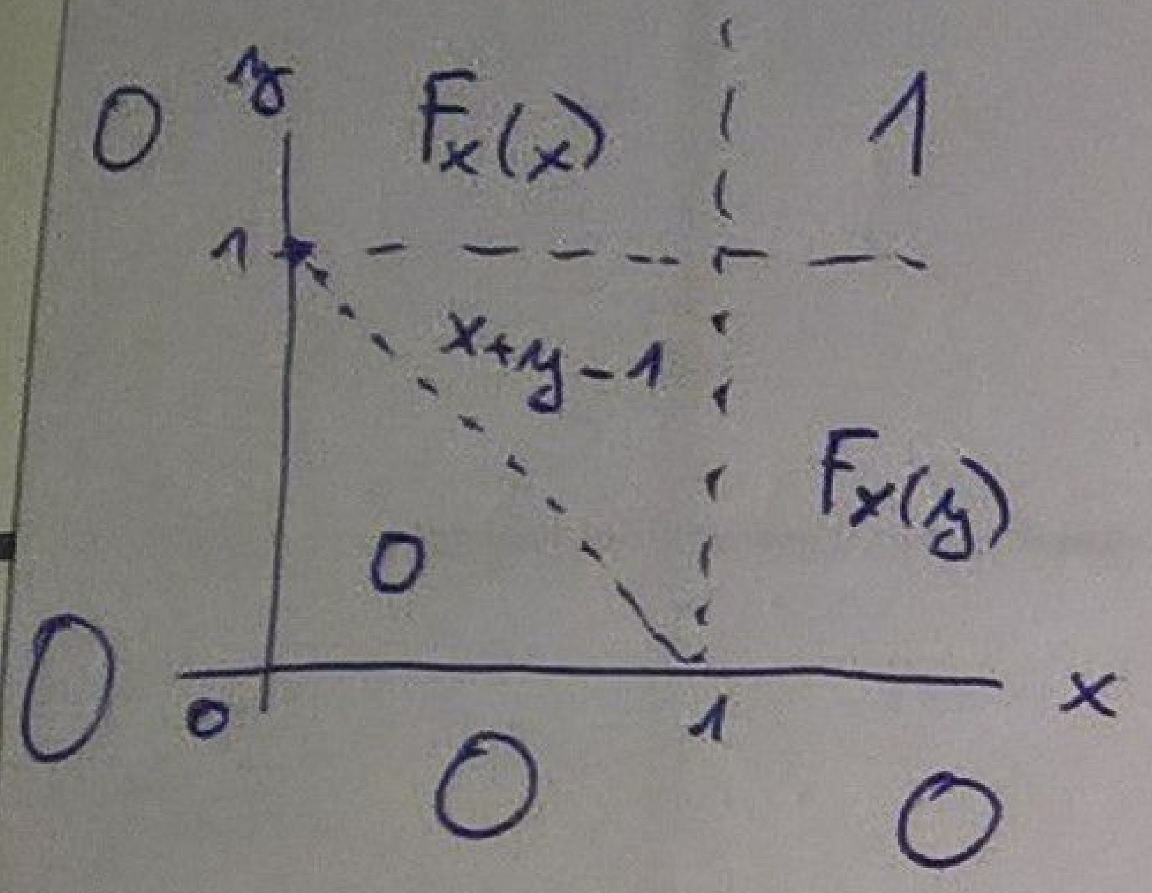
b) $F(x,y) = \min\{x,y\} \quad , 0\le x\le 1,\ 0\le y\le 1$



$$F_X(x) = \lim_{y\to\infty} F(x,y) = \begin{cases} 0 & , x<0 \\ x & , 0\le x\le 1 \\ 1 & , x>1 \end{cases} \qquad X\sim Rs(0,1)$$

podobně $F_Y(y) = \lim\limits_{x \to \infty} F(x,y) = \begin{cases} 0 & , y<0 \\ y & , 0 \le y \le 1 \\ 1 & , y > 1 \end{cases}$ $\quad Y \sim Rs(0,1)$ $\quad$ X a Y nejsou nezávislé, navíc $Y = X$.

c) $F(x,y) = \max\{x+y-1, 0\}$ $, 0 \le x \le 1, \; 0 \le y \le 1$



$F_X(x) = \lim\limits_{y \to \infty} F(x,y) = \begin{cases} 0 & , x<0 \\ x & , 0 \le x \le 1 \\ 1 & , x > 1 \end{cases}$ $\quad X \sim Rs(0,1)$

podobně $F_Y(y) = \begin{cases} 0 & , y<0 \\ y & , 0 \le y \le 1 \\ 1 & , y > 1 \end{cases}$ $\quad Y \sim Rs(0,1)$ $\quad$ X a Y nejsou nezávislé, navíc $Y = 1 - X$.

<u>Poznámka</u>

Sdružené rozdělení náh. vektoru jednoznačně určuje marginální rozdělení jeho složek.

Opačně to však neplatí!

Cíl: Popsat, jak vypadají všechna sdružená rozdělení s předem danými marginálními rozděleními?

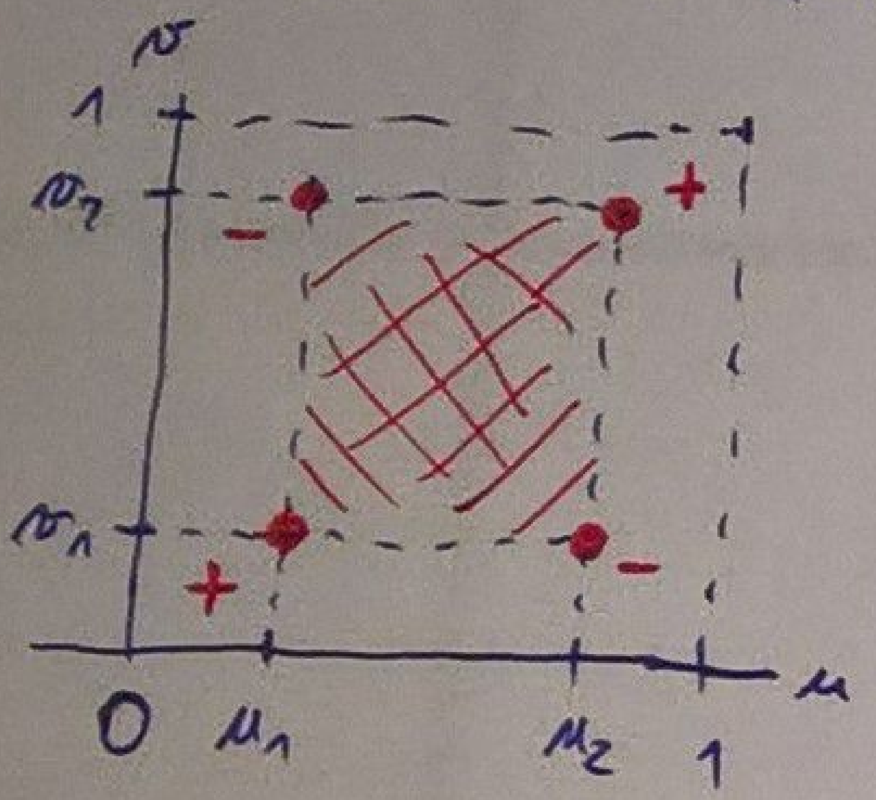
Definice (kopula)

Funkce $C:[0,1]\times[0,1]\to[0,1]$ se nazývá kopula, jestliže:

(i) $C(u,0)=C(0,u)=0, \quad \forall u\in[0,1]$

(ii) $C(u,1)=C(1,u)=u, \quad \forall u\in[0,1]$

(iii) $C(u_1,v_1)-C(u_1,v_2)-C(u_2,v_1)+C(u_2,v_2)\geq 0, \quad \forall\, 0\leq u_1\leq u_2\leq 1, \ 0\leq v_1\leq v_2\leq 1$ (C je 2-rostoucí)



Jiná definice

Funkce $C:[0,1]\times[0,1]\to[0,1]$ je kopula, jestliže existuje pravd. prostor $(\Omega,\mathcal{A},P)$ a na něm náhodný vektor $(U,V)^T$ takový, že:

(i) $U\sim Rs(0,1)$

(ii) $V\sim Rs(0,1)$

(iii) $C(u,v)=P(U\leq u, V\leq v)$ je sdružená distribuční funkce náh. vektoru $(U,V)^T$.

Příklady

(i) $C(u,v)=u\cdot v=\Pi(u,v)$ ... součinová (nezávislá) kopula (U,V nezávislé)

(ii) $C(u,v) = \min\{u,v\} = M(u,v)$ ... horní kopula $(V=U)$

(iii) $C(u,v) = \max\{u+v-1, 0\} = W(u,v)$ ... dolní kopula $(V=1-U)$

<u>Tvrzení</u> (Fréchetovy - Höffdingovy meze)

Pro každou kopulu C platí: $W(u,v) \leq C(u,v) \leq M(u,v)$, $\forall u,v \in [0,1]$

A co když chceme marginální rozdělení jiná než rovnoměrná?

<u>Věta (Sklar)</u>

Nechť náh. vektor (X,Y) má sdruženou distribuční funkci F a marginální distribuční funkce F_1 a F_2. Pak existuje taková kopula C, že

$F(x,y) = C(F_1(x), F_2(y))$, $\forall x,y \in \mathbb{R}$. Je-li F spojitá, pak C je určena jednoznačně.

<u>Poznámka</u>

Předchozí věta nám dává návod, jak modelovat dvourozměrná data (rozdělení) - předepíšeme si marginální rozdělení a zvolíme vhodnou kopulu, která popisuje závislost složek (nezávisle na jejich rozdělení).

<u>Vztah kopul a korelačních koeficientů</u>

1) Pearsonův korelační koeficient $\overset{\rho(X,Y)}{\rho} = \dfrac{C(X,Y)}{\sqrt{D(X)\cdot D(Y)}}$ ... závisí na marginálním rozdělení n.v. X a Y

nelze vyjádřit jen pomocí kopuly

náh. vektor $(X,Y)^T$ ... $F(x,y)$

$\rightarrow$ kopula C

2) Spearmanův korelační koeficient

označme F_1 marg. distr. fce n. v. X, F_2 -||- Y $\Big\}$ $F_1(X) \sim Rs(0,1)$ a $F_2(Y) \sim Rs(0,1)$

$$\underset{\rho_S(X,Y)}{\rho_S} = \rho\left(F_1(X), F_2(Y)\right) = \frac{E[F_1(X)\cdot F_2(Y)] - \underbrace{[EF_1(X)]}_{=\frac{1}{2}}\cdot\underbrace{[EF_2(Y)]}_{=\frac{1}{2}}}{\sqrt{\underbrace{DF_1(X)}_{=\frac{1}{12}}\cdot\underbrace{DF_2(Y)}_{=\frac{1}{12}}}} = 12\cdot E[F_1(X)\cdot F_2(Y)] - 3.$$

$$\iint_{\mathbb{R}^2} F_1(x)\cdot F_2(y)\cdot \underbrace{dF(x,y)}_{dC(F_1(x),F_2(y))} = \int_0^1\int_0^1 u\cdot v\; dC(u,v) = \int_0^1\int_0^1 C(u,v)\,du\,dv = E[F_1(X)\cdot F_2(Y)]$$

$F_1(x) = u$

$F_2(y) = v$

$$\rho_S = 12\cdot\int_0^1\int_0^1 C(u,v)\,du\,dv - 3.$$

3) Kendallovo τ

Nechť $(X_1,Y_1)^T$ a $(X_2,Y_2)^T$ jsou dvě nezávislé kopie náh. vektoru (X,Y)

$$\underset{\tau(X,Y)}{\tau} = P((X_1-X_2)\cdot(Y_1-Y_2)>0) - P((X_1-X_2)\cdot(Y_1-Y_2)<0) = E\,\mathrm{sign}\,(X_1-X_2)\cdot(Y_1-Y_2) = 4\cdot\int_0^1\int_0^1 C(u,v)\,dC(u,v) - 1 =$$

$$4\int_0^1\int_0^1 C(u,v)\,c(u,v)\,du\,dv - 1,$$ kde $c(u,v) = \dfrac{\partial^2 C(u,v)}{\partial u\,\partial v}$ je hustota kopuly.

Třídy používaných kopul

1.) Archimédovské kopuly

- dají se většinou vyjádřit v uzavřeném tvaru
- většinou závisí na jednom parametru

$$C(u,v) = \phi^{-1}(\phi(u)+\phi(v)), \quad 0\le u\le 1,\ 0\le v\le 1, \text{ kde } \phi:[0,1]\to\mathbb{R}^+ \text{ je generátor.}$$

- ϕ je rostoucí, konvexní a spojitá
- $\phi(1)=0$

lze ukázat, že $\tau = 1+4\int_0^1 \frac{\phi(u)}{\phi'(u)}\,du$.

a) součinová (nezávislá) kopula

$$\left.\begin{aligned}\phi(u) &= -\log u\\ \phi^{-1}(x) &= e^{-x}\end{aligned}\right\} C(u,v) = \phi^{-1}(\phi(u)+\phi(v)) = e^{-(-\log u-\log v)} = e^{\log u+\log v} = e^{\log u\cdot v} = u\cdot v$$

$$C(u,v) = u\cdot v$$

$$\tau = 1+4\cdot\int_0^1 \frac{\phi(u)}{\phi'(u)}\,du = 1+4\cdot\int_0^1 \frac{-\log u}{-\frac{1}{u}}\,du = 1+4\underbrace{\int_0^1 u\cdot\log u\,du}_{=-\frac{1}{4}} = 0.$$

b) Gumbelova (Gumbelova–Hougaardova) kopula

$$\left.\begin{array}{l}\phi(u) = (-\log u)^{\theta}, \quad \theta \geq 1 \text{ je parametr} \\ \phi^{-1}(x) = \exp\left\{-x^{\frac{1}{\theta}}\right\}\end{array}\right\} C(u,v) = \exp\left\{-\left[(-\log u)^{\theta} + (-\log v)^{\theta}\right]^{\frac{1}{\theta}}\right\}.$$

$$\tau = 1 - \frac{1}{\theta}.$$

<u>Poznámka</u>

Pro $\theta = 1$ je Gumbelova kopula rovna nezávislé.

Pro $\theta \to \infty$ se Gumbelova kopula blíží horní kopule.

c) Joeova kopula

$$\left.\begin{array}{l}\phi(u) = -\log\left[1-(1-u)^{\theta}\right], \quad \theta \geq 1 \\ \phi^{-1}(x) = 1-(1-e^{-x})^{\frac{1}{\theta}}\end{array}\right\} C(u,v) = 1-\left[(1-u)^{\theta} + (1-v)^{\theta} - (1-u)^{\theta}\cdot(1-v)^{\theta}\right]^{\frac{1}{\theta}}.$$

<u>Poznámka</u>

Pro $\theta = 1$ je Joeova kopula rovna nezávislé a pro $\theta \to \infty$ se blíží horní kopule.

d) Claytonova kopula

$$\left.\begin{array}{l}\phi(u) = \frac{1}{\theta}\cdot\left(u^{-\theta}-1\right) \\[2ex] \phi^{-1}(x) = \left(\theta x+1\right)^{-\frac{1}{\theta}}\end{array}\right\} C(u,v) = \max\left\{\left[u^{-\theta}+v^{-\theta}-1\right]^{-\frac{1}{\theta}}, 0\right\}$$

, kde $\theta \in [-1, \infty) \setminus \{0\}$ je parametr.

$$\tau = \frac{\theta}{2+\theta}.$$

Poznámka

$\theta=-1$ ... dolní kopula, $\theta\to\infty$ ... horní kopula a pro $\theta\to 0$ ... součinová

e) BB6 kopula

$$\phi(u) = \left\{-\log\left[1-(1-u)^{\theta}\right]\right\}^{\delta}$$

, $\theta \geq 1$, $\delta \geq 1$ jsou parametry

$$C(u,v) = 1-\left\{1-\exp\left\{-\left(\left[-\log\left(1-(1-u)^{\theta}\right)\right]^{\delta}+\left[-\log\left[1-(1-v)^{\theta}\right]\right]^{\delta}\right)^{\frac{1}{\delta}}\right\}\right\}^{\frac{1}{\theta}}.$$

Poznámka

$\delta=1$ ... Joeova kopula

f) Frankova kopula

$$\phi(u) = -\log\left(\frac{e^{-\theta u}-1}{e^{-\theta}-1}\right)$$

$$C(u,v) = -\frac{1}{\theta}\cdot\log\left(1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right)$$, $\theta \in \mathbb{R}\setminus\{0\}$ parametr

<u>Poznámka</u>

$\theta \to \infty$ ... horní kopula , $\theta \to -\infty$ ... dolní kopula , $\theta \to 0$ ... součinová kopula

2) Eliptické kopuly

- vhodné pro modelování elipticky symetrických rozdělení

a) normální (gaussovská) kopula

$$C(u,v) = \Phi_\rho\left(\Phi^{-1}(u), \Phi^{-1}(v)\right)$$

, kde Φ^{-1} je kvantilová funkce stand. normálního rozdělení $N(0,1)$

$\Phi_\rho(x,y)$ je sdružená distribuční funkce 2-rozměrného normálního rozdělení $N_2(\mathbb{0}, \Sigma)$, kde

$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ je jeho varianční matice,

$\rho \in [-1, 1]$ je parametr.

$$\tau = \frac{2}{\pi}\arcsin\rho$$

<u>Poznámka</u>

$\rho = 1$ ... horní kopula , $\rho = -1$ ... dolní kopula , $\rho = 0$ ... součinová kopula

b) Studentova t kopula

$C(u,v) = t_{\nu,\rho}\left(t_\nu^{-1}(u),\, t_\nu^{-1}(v)\right)$, kde t_ν^{-1} je kvantilová funkce t-rozdělení s ν stupni volnosti,

$t_{\nu,\rho}$ je sdružená distr. fce 2-rozměrného t-rozdělení s ν stupni volnosti a variančí maticí $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$,

$\rho \in [-1,1]$ a $\nu \in [1,\infty)$ jsou parametry

$\tau = \frac{2}{\pi} \cdot \arcsin \rho$.