Centrální limitní věta

- x_i nezávislé náhodné proměnné s hustotami pravděpodobnosti $f_i(x_i)$
- očekávané hodnoty $E[x_i] = \mu_i$ a rozptyly $V[x_i] = \sigma_i^2$
- potom platí:

$$\boxed{y = \sum_{i=1}^{N} x_i \quad \text{pro} \quad N \to \infty \quad \text{je} \quad y \in N\left(\sum_{i=1}^{N} \mu_i, \sqrt{\sum_{i=1}^{N} \sigma_i^2}\right)}$$

$$\boxed{z = \frac{y - \sum\limits_{i=1}^{N} \mu_i}{\sqrt{\sum\limits_{i=1}^{N} \sigma_i^2}} \quad \text{pro} \quad N \to \infty \quad \text{je} \quad z \in N(0,1)}$$

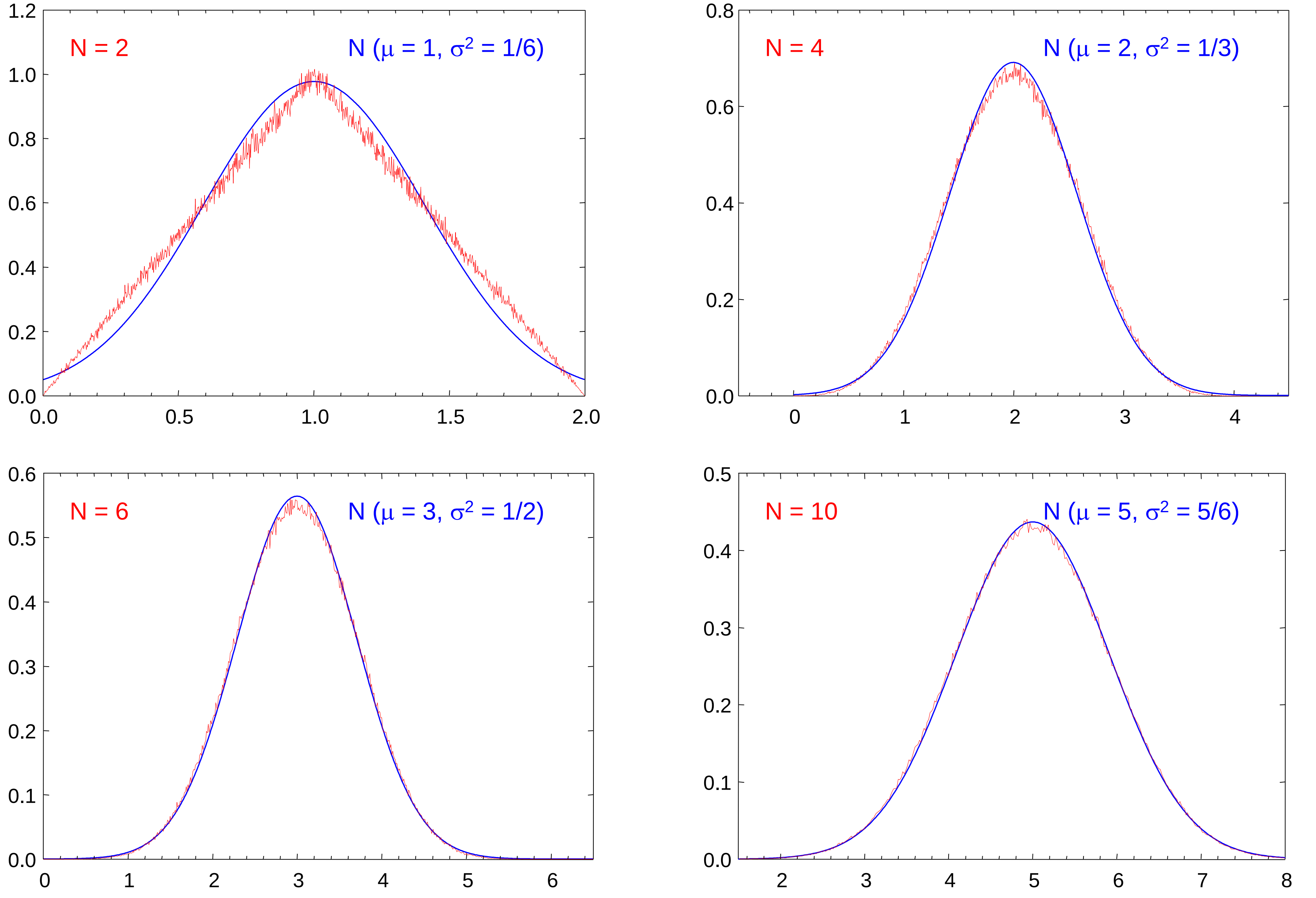
Centrální limitní věta

$$y = \sum_{i=1}^{N} x_i$$

$$x_i \in U(0,1)$$

$$\mu = \sum_{i=1}^{N} \mu_i$$

$$\sigma^2 = \sum_{i=1}^{N} \sigma_i^2$$



Hustota pravděpodobnosti – případ dvou proměnných

- **marginální hustoty pravděpodobnosti**

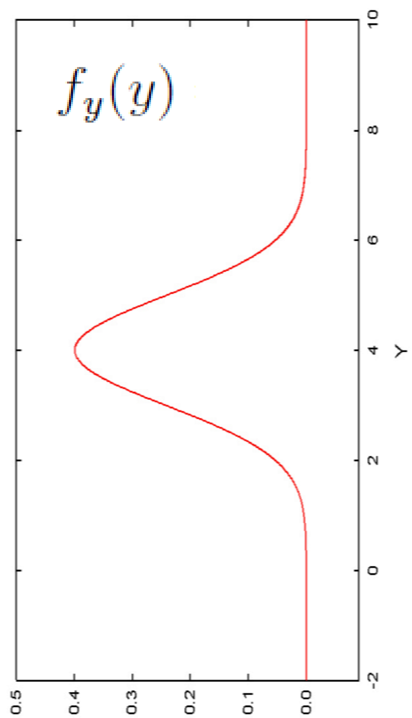
$$f_x(x) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}y$$

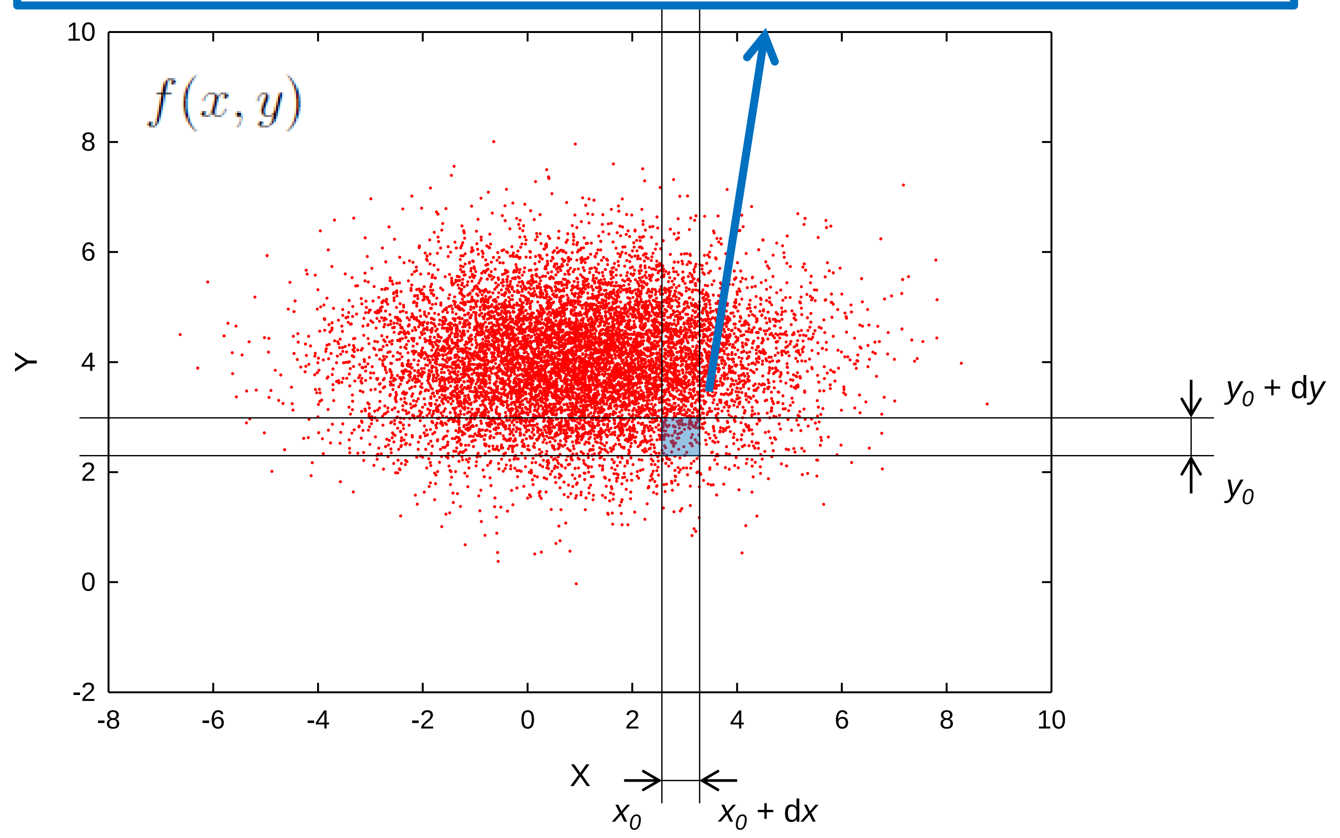
$$f_y(y) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}x$$

$A: \quad x \in [x_0, x_0 + \mathrm{d}x]$

$$P(A \cap B) = f(x_0, y_0)\mathrm{d}x\mathrm{d}y$$

$B: \quad y \in [y_0, y_0 + \mathrm{d}y]$





Hustota pravděpodobnosti – případ dvou proměnných

- operátor **očekávané hodnoty**

očekávaná hodnota náhodné proměnné *x*

$$\mu_x \equiv E[x] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f(x,y)\,\mathrm{d}x\mathrm{d}y$$

očekávaná hodnota náhodné proměnné *y*

$$\mu_y \equiv E[y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(x,y)\,\mathrm{d}x\mathrm{d}y$$

obecně:

$$E[g(x,y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f(x,y)\,\mathrm{d}x\mathrm{d}y$$

Hustota pravděpodobnosti – případ dvou proměnných

- operátor **rozptylu**

rozptyl náhodné proměnné *x*

$$\begin{aligned}\sigma_x^2 \equiv V[x] &= E\left[(x-\mu_x)^2\right] \\ &= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}(x-\mu_x)^2 f(x,y)\mathrm{d}x\mathrm{d}y\end{aligned}$$

rozptyl náhodné proměnné *y*

$$\begin{aligned}\sigma_y^2 \equiv V[y] &= E\left[(y-\mu_y)^2\right] \\ &= \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}(y-\mu_y)^2 f(x,y)\mathrm{d}x\mathrm{d}y\end{aligned}$$

Hustota pravděpodobnosti – případ dvou proměnných

- operátor **rozptylu**

kovariance náhodných proměnných *x* a *y*

$$\boxed{\mathrm{cov}(x,y) = E\left[(x-\mu_x)(y-\mu_y)\right] = E[xy] - E[x]E[y]}$$

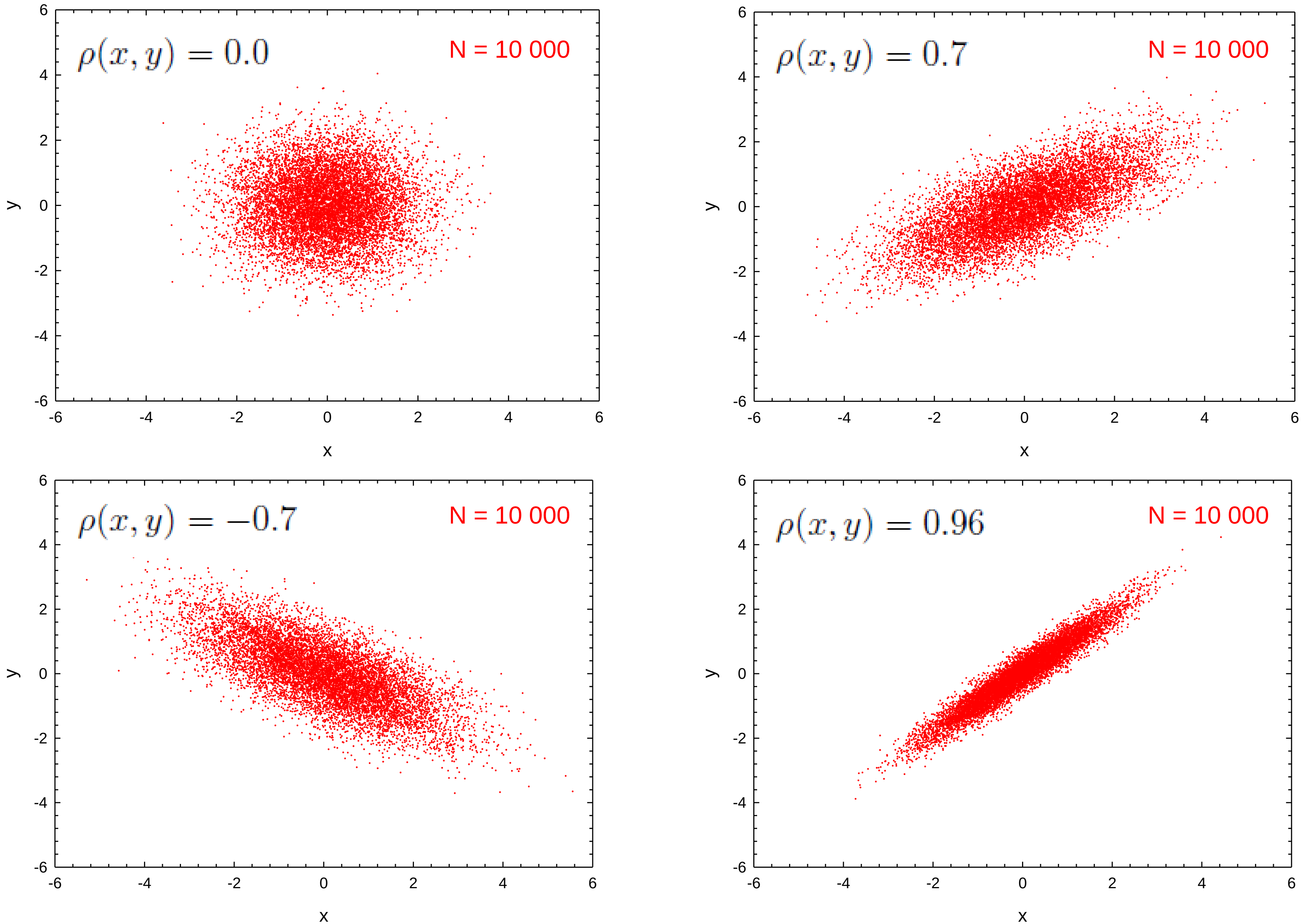
kovarianční matice

$$\begin{pmatrix} V[x] & \mathrm{cov}(x,y) \\ \mathrm{cov}(y,x) & V[y] \end{pmatrix} = \begin{pmatrix} \sigma_x^2 & \mathrm{cov}(x,y) \\ \mathrm{cov}(x,y) & \sigma_y^2 \end{pmatrix}$$

korelace náhodných proměnných *x* a *y*

$$\boxed{\rho(x,y) = \frac{\mathrm{cov}(x,y)}{\sigma_x \sigma_y}}$$

Korelace náhodných proměnných



Hustota pravděpodobnosti – případ dvou proměnných

- **nezávislé proměnné** *x* a *y*

$$\boxed{f(x,y) = f_x(x) f_y(y)}$$

rozptyl náhodné proměnné *x*

$$V[x] = E\left[(x-\mu_x)^2\right] = \int_{-\infty}^{\infty} (x-\mu_x)^2 f_x(x)\mathrm{d}x \int_{-\infty}^{\infty} f_y(y)\mathrm{d}y = \int_{-\infty}^{\infty} (x-\mu_x)^2 f_x(x)\mathrm{d}x$$

rozptyl náhodné proměnné *y*

$$V[y] = E\left[(y-\mu_y)^2\right] = \int_{-\infty}^{\infty} (y-\mu_y)^2 f_y(y)\mathrm{d}y \int_{-\infty}^{\infty} f_x(x)\mathrm{d}x = \int_{-\infty}^{\infty} (y-\mu_y)^2 f_y(y)\mathrm{d}y$$

kovariační matice

$$\begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}$$

$$\mathrm{cov}(x,y) = 0$$

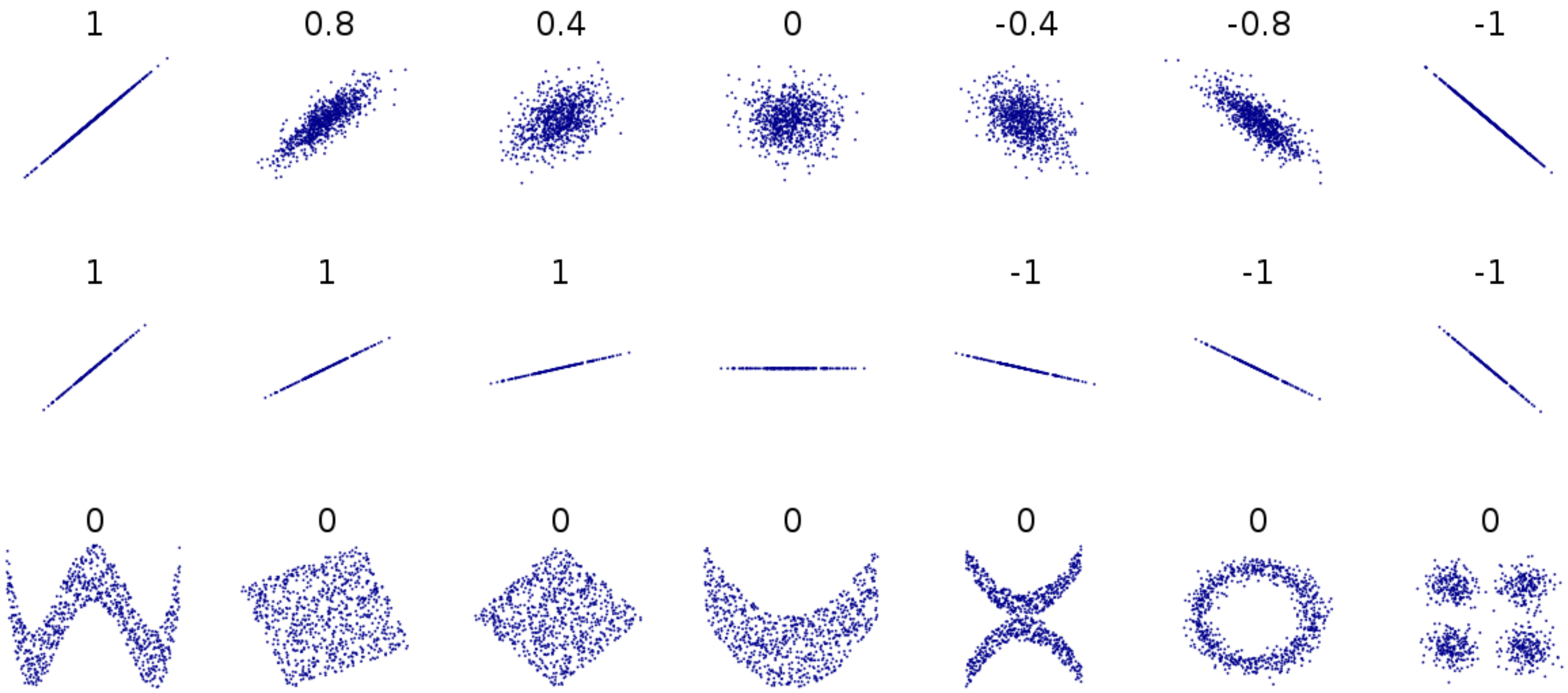
$$\rho(x,y) = 0$$

Hustota pravděpodobnosti – případ dvou proměnných

- **nezávislé proměnné** *x* a *y* ⇒ $\boxed{\mathrm{cov}(x,y) = 0 \qquad \rho(x,y) = 0}$

Obrácená implikace neplatí!

Nulová korelace je nutná, nikoli postačující podmínka nezávislosti proměnných.



Odhad kovariance a korelace

- máme náhodné proměnné *x* a *y*.
- naměříme hodnoty $x_1, x_2, \dots, x_N$ a $y_1, y_2, \dots, y_N$

$$\boxed{\hat{\mathrm{cov}}(x,y) = \langle xy \rangle - \langle x \rangle \langle y \rangle \qquad \hat{\rho}(x,y) = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\hat{\sigma_x}\hat{\sigma_y}} \qquad \sigma_\rho \approx \frac{1-\rho^2}{\sqrt{N-1}}}$$

$$\langle xy \rangle = \frac{1}{N}\sum_{i=1}^{N} x_i y_i$$

$$\langle x \rangle = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\langle y \rangle = \frac{1}{N}\sum_{i=1}^{N} y_i$$

$$\hat{\sigma}_x = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} (x_i - \langle x \rangle)^2}$$

$$\hat{\sigma}_y = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} (y_i - \langle y \rangle)^2}$$