Více náhodných veličin

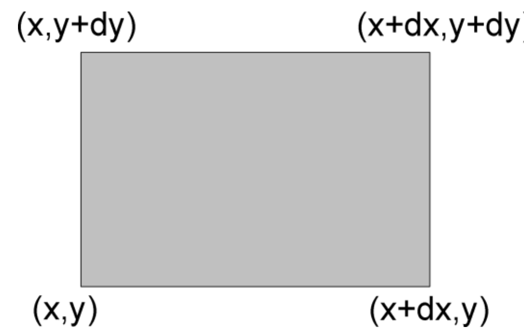
• Dvě náhodné proměnné x, y mají rozdělení pravděpodobnosti na intervalech V_x, V_y popsáno funkcemi $p(x)$, $q(y)$

• Jaká je pravděpodobnost, že

x se nachází v intervalu $(x, x+\mathrm{d}x)$

a **zároveň**

y se nachází v intervalu $(y, y+\mathrm{d}y)$?



$$P\begin{bmatrix} x \in (x, x+\mathrm{d}x) \\ y \in (y, y+\mathrm{d}y) \end{bmatrix} = \rho(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

$\rho(x, y)$ je rozdělení pravděpodobnosti dvou náhodných proměnných.

$$\left.\begin{array}{l} p(x) \text{ na } V_x \\ q(y) \text{ na } V_y \end{array}\right\} \rightarrow \rho(x,y) \text{ na } V = V_x \times V_y$$

Více náhodných veličin

- Rozdělení pravděpodobnosti dvou náhodných proměnných $\rho(x, y)$
  - funguje podobně jako v případě jedné proměnné:

- **střední hodnota:** $\mu_x = \langle x \rangle = \int\limits_V x\,\rho(x, y)\,\mathrm{d}x\,\mathrm{d}y \qquad \mu_y = \langle y \rangle = \int\limits_V y\,\rho(x, y)\,\mathrm{d}x\,\mathrm{d}y$

(obecně) $\langle f(x, y) \rangle = \int\limits_V f(x, y)\,\rho(x, y)\,\mathrm{d}x\,\mathrm{d}y$

- **momenty:**

$$\mu_{x,n} = \langle x^n \rangle = \int\limits_V x^n\,\rho(x, y)\,\mathrm{d}x\,\mathrm{d}y \qquad \mu_{y,n} = \langle y^n \rangle = \int\limits_V y^n\,\rho(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

- **centrální momenty:** $\mu'_{x,n} = \langle (x - \mu_x)^n \rangle = \int\limits_V (x - \mu_x)^n\,\rho(x, y)\,\mathrm{d}x\,\mathrm{d}y$

$$\mu'_{y,n} = \langle (y - \mu_x)^n \rangle = \int\limits_V (y - \mu_y)^n\,\rho(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

Kovariance, koeficient korelace

Jak vypadá rozdělení $\rho(x, y)$? x a y **nezávislé** $\rightarrow$ $\rho(x,y) = p(x)q(y)$

Obecně (např. nejsou-li nezávislé), vyjadřujeme míru jejich vztahu pomocí **kovariance**.

$$\mathrm{Cov}(x,y) = \left\langle (x-\mu_x)(y-\mu_y) \right\rangle = \int_V (x-\mu_x)(y-\mu_y)\,\rho(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

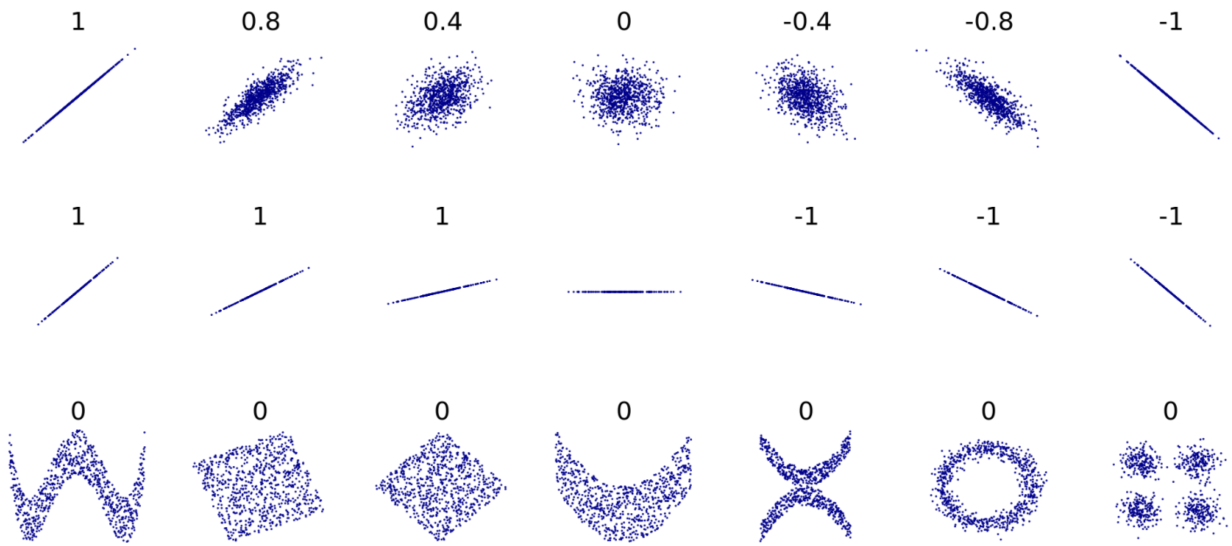
$$\mathrm{Cov}(x,y) = \langle xy \rangle - \langle x \rangle \langle y \rangle \quad \rightarrow \mathrm{Cov}(x,x) = \mathrm{V}(x) = \sigma_x^2$$ (ko-variance)

Koeficient korelace: $$\mathrm{r}(x,y) = \frac{\mathrm{Cov}(x,y)}{\sigma_x \sigma_y}$$

$$-1 \leq \mathrm{r}(x,y) \leq 1$$

antikorelované $= 0$ nezávislé korelované

příklady:



$$\sigma_r \approx \frac{1-r}{\sqrt{N-1}}$$

Kovariance, koeficient korelace

Příklad:

Veličiny x a y jsou lineárně závislé: $y = a.x + b$

$$\mathrm{r}(x, y) = ?$$

$$\mathrm{r}(x, y) = \frac{\mathrm{Cov}(x, y)}{\sigma_x \sigma_y} \qquad \mathrm{Cov}(x, y) = \langle xy \rangle - \langle x \rangle \langle y \rangle$$

n náhodných veličin

- Obecný případ pro n náhodných veličin: $x_1, x_2, ..., x_n$

  - rozdělení pravděpodobnosti: $\rho(x_1, x_2, ..., x_n)$

- Pro každou veličinu x_i lze opět psát:

  střední hodnotu, momenty, disperzi, ...

$$\langle f(x_i) \rangle = \int_V f(x_i)\, \rho(x_1, x_2, ..., x_n)\, \mathrm{d}x_1\, \mathrm{d}x_2 \ldots \mathrm{d}x_n$$

- Součet náhodných veličin: $s = \sum_{i=1}^{n} x_i$

... a jeho střední hodnota:

$$\langle s \rangle = \int_V s\, \rho(x_1, x_2, ..., x_n)\, \mathrm{d}x_1\, \mathrm{d}x_2 \ldots \mathrm{d}x_n \quad = \ldots = \sum_{i=1}^{n} \langle x_i \rangle$$

Aritmetický průměr - střední hodnota

- Střední hodnota součtu náhodných veličin:
  (je rovna součtu středních hodnot)

$$\left\langle \sum_{i=1}^{n} x_i \right\rangle = \sum_{i=1}^{n} \langle x_i \rangle$$

- Speciálně: pro n-násobné opakování veličiny x

$$\left\langle \sum_{i=1}^{n} x \right\rangle = \sum_{i=1}^{n} \langle x \rangle = n\langle x \rangle$$

- Aritmetický průměr: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

$$\langle \bar{x} \rangle = \left\langle \frac{1}{n}\sum_{i=1}^{n} x_i \right\rangle = \frac{1}{n}\sum_{i=1}^{n} \langle x_i \rangle = \frac{1}{n}\sum_{i=1}^{n} \langle x \rangle = \langle x \rangle$$

Zákon velkých čísel $\Pr\{\lim_{n\to\infty} \bar{x} = \langle x \rangle\} = 1$

(Čebyševova nerovnost)

Disperze aritmetického průměru

- A co disperze $\mathrm{V}(\bar{x})$?

- Disperze (variance) součtu náhodných veličin: $\mathrm{V}(s) = \langle s^2 \rangle - \langle s \rangle^2$

$$\mathrm{V}(s) = \mathrm{V}\left(\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} \mathrm{V}(x_i) + \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \mathrm{Cov}(x_i, x_j)$$

- Jsou-li x_i nezávislé, $\mathrm{Cov}(x_\mathrm{i}, x_\mathrm{j}) = 0$

- Pro aritmetický průměr: $\mathrm{V}(x) = \frac{1}{n}\mathrm{V}(x)$ $\qquad \sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$

(význam pro středování)

Centrální limitní věta

- Náhodná veličina *x* je popsána rozdělením pravděpodobnosti $p(x)$.
  - střední hodnota: $\mu = \langle \mathrm{x} \rangle$
  - disperze: $\mathrm{V} = \langle x^2 \rangle - \langle x \rangle^2$
- Aritmetický průměr při *n*-násobném nezávislém opakování veličiny *x*:
  - je popsán rozdělením $q(x)$
  - střední hodnota: $\mu_n = \langle \mathrm{x} \rangle$
  - disperze: $\mathrm{V}_n = \dfrac{\mathrm{V}}{n}$
- **CLV:** S rostoucím *n* se $q(x)$ blíží **normálnímu rozdělení** $\mathrm{N}\left(\mu, \dfrac{\mathrm{V}}{n}\right)$

$$\lim_{n \to \infty} q\left(\frac{x - \mu}{\sqrt{\mathrm{V}_n}}\right) = \mathrm{N}(0,1)$$

Na typu rozdělení $p(x)$ nezáleží!