## Význam hustoty pravděpodobnosti (hp)

Implicitní definice hp je

$$F(x) = \int_{-\infty}^{x} f(t)\, dt$$

tj. pokud existuje taková nezáporná funkce $f(x)$, pro kterou platí $\int_{-\infty}^{\infty} f(x)\, dx = 1$ a která splňuje předchozí vztah, pak je to hp náhodné veličiny s distribuční funkcí $F(x)$. Protože zároveň (podle definice) platí $F(x) = P(X \le x)$ dostaneme

$$P(X \le x) = \int_{-\infty}^{x} f(t)\, dt,$$

a dále

$$P(X \in (a,b)) = \int_{-\infty}^{b} f(t)\, dt - \int_{-\infty}^{a} f(t)\, dt = \int_{a}^{b} f(t)\, dt,$$

tedy, pravděpodobnost, že realizace náhodné veličiny padne do intervalu $(a,b)$ je rovna ploše pod hp na tomto intervalu.

Jestliže nyní budeme uvažovat nějaký elementární interval $(a,b) = \Delta$ - vždy stejně dlouhý, pak můžeme říci, že pravděpodobnost tohoto intervalu je úměrná hodnotě hustoty pravděpodobnosti v tomto místě (interval $\Delta$ uvažujeme tak malý, že hp můžeme na něm považovat za konstantní).

A to je přímá paralela s pravděpodobnostní funkcí - každá hodnota má svou pravděpodobnost a ta je dána hodnotou pravděpodobnostní funkce. Tady každý bod (reprezentovaný svým intervalem $\Delta$) má také svou pravděpodobnost a ta je úměrná hodnotě hp

### Příklad 1

Určete charakteristiky rozdělení s hp

$$f(x) = 0.5x, \text{ pro } x \in (0,2), \text{ jinak } 0$$

Zjevně $f(x) \ge 0$ na intervalu $(0,2)$ a $\int_0^2 f(x)\, dx = 1$ - zadání je tedy správné.

**Střední hodnota**

$$E[X] = \int_0^2 x f(x)\, dx = \int_0^2 x \cdot 0.5x dx = \int_0^2 0.5x^2 dx = \left[\frac{1}{6}x^3\right]_0^2 = \frac{4}{3}$$

**Rozptyl**

$$D[X] = \int_0^2 \left(x - \frac{4}{3}\right)^2 0.5x\, dx = \frac{2}{9}$$

**Medián** $\tilde{x}$

$$\int_0^{\tilde{x}} 0.5x\, dx = 0.5$$

po integraci

$$\frac{1}{4}\left[x^2\right]_0^{\tilde{x}} = 0.5$$

$$\tilde{x}^2 = 2$$
$$\tilde{x} = \sqrt{2}$$

Což odpovídá i geometrické představě: plocha pod hp je trojúhelník a my ho chceme svislou čárou rozdělit na dva kusy se stejnými plochami. Čára tedy musí být někde za jedničkou.

Pozn. V tomto jednoduchém příkladě jsme se dobrali výsledku. Jde ale o řešení implicitního vztahu, a to většinou nelze analyticky provést. Proto je třeba použít řešení numerické.

## Příklad 2

Je dána hp
$$f(x) = 1 - |x-2|$$
pro $x \in (1,3)$ a jinak 0. Určete distribuční funkci.

V grafu je tato hp "stříška", pověšená nad hodnotou 2. Je to tedy funkce po částech analytická, které můžeme zapsat takto

$$f(x) = \begin{cases} 0 & \text{pro } x < 1 \\ x-1 & \text{pro } x \in (1,2) \\ 3-x & \text{pro } x \in (2,3) \\ 0 & \text{pro } x > 3 \end{cases}$$

Distribuční funkce je dána integrálem

$$F(x) = \int_{-\infty}^{x} f(t)\, dt$$

a my musíme integrovat po částech na jednotlivých intervalech.

Na $x < 1$ je $f(x) = 0$ a tedy $F(x) = 0$

Na $x \in (1,2)$ je
$$F(x) = \int_{-\infty}^{x} f(t)\, dt = \int_{-\infty}^{1} f(t)\, dt + \int_{1}^{x} f(t)\, dt =$$
$$= 0 + \int_{1}^{x} (t-1)\, dt = \left[\frac{1}{2}f^2 - t\right]_1^x = \frac{1}{2}x^2 - x - \frac{1}{2} + 1 = \frac{1}{2}x^2 - x + \frac{1}{2}$$

Na $x \in (2,3)$ máme: do 1 je to 0, do 2 je to $F(x=2) = \frac{1}{2}$ (dosazeno do předchozího vztahu) a tedy
$$F(x) = 0 + \frac{1}{2} + \int_{2}^{x} (3-t)\, dt = \frac{1}{2} + \left[3t - \frac{1}{2}t^2\right]_2^x =$$
$$= \frac{1}{2} + 3x - \frac{1}{2}x^2 - 6 + 2 = -\frac{1}{2}x^2 + 3x - \frac{7}{2}$$

Kontrola - v $x = 3$ distribuční funkce končí, a tedy tam musí být $F(3) = 1$ - což je.

Na $x > 3$ je zase $f(x) = 0$ a tedy
$$F(x) = 1.$$

## Příklad 3

Je dán náhodný vektor $[X,Y]'$ se sdruženou hp

$$f(x) = 2x^2 - 4xy + 4y^2$$

pro $x, y \in (0,1)$. Určete marginály, podmíněné hp a kovarianci.

**Marginály**

$$f(x) = \int_{-\infty}^{\infty} f(x,y)\, dy = \int_{-\infty}^{\infty} \left(2x^2 - 4xy + 4y^2\right) dy = 2x^2 - 2x + \frac{4}{3}$$

$$f(y) = \int_{-\infty}^{\infty} \left(2x^2 - 4xy + 4y^2\right) dx = 4y^2 - 2y + \frac{2}{3}$$

**Podmíněné hp**

$$f(x|y) = \frac{f(x,y)}{f(y)}, \quad f(y|x) = \frac{f(x,y)}{f(x)}$$

$$f(x|y) = \frac{2x^2 - 4xy + 4y^2}{4y^2 - 2y + \frac{2}{3}}, \quad f(y|x) = \frac{2x^2 - 2x + \frac{4}{3}}{2x^2 - 4xy + 4y^2}$$

**Kovariance**

Nejprve musíme spočítat střední hodnoty. pomocí marginál

$$E[X] = \int_{-\infty}^{\infty} x \left(2x^2 - 2x + \frac{4}{3}\right) dx = \frac{1}{2}$$

$$E[Y] = \int_{-\infty}^{\infty} y \left(4y^2 - 2y + \frac{2}{3}\right) dx = \frac{2}{3}$$

$$C[X,Y] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - EX)(y - EY) f(x,y)\, dxdy =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(x - \frac{1}{2}\right) \left(y - \frac{2}{3}\right) \left(2x^2 - 4xy + 4y^2\right) dxdy = -\frac{1}{36}$$