DISKRÉTNÍ ROZDĚLENÍ

Náhodná veličina X má diskrétní rozdělení, pokud existují čísla $p_i > 0$

taková, že $$\sum_{i \in M} p_i = 1,$$

kde M je spočetná či konečná množina,

a čísla $x_i \in \mathbb{R}$ tak,

že $$p_i = P[X = x_i]$$

tedy: $$P[a \le X \le b] = \sum_{i;\, x_i \in \langle a, b \rangle} p_i$$

Střední hodnota:

$$EX = \sum_{i \in M} x_i p_i$$

Je-li $g: \mathbb{R} \to \mathbb{R}$ měřitelná funkce,

pak: $$E(g(X)) = \sum_{i \in M} g(x_i)\, p_i$$

SPOJITÉ ROZDĚLENÍ

Náhodná veličina X má spojité rozdělení, pokud existuje funkce $f \ge 0$

taková, že $$\int_{\mathbb{R}} f(x)\, dx = 1,$$

a platí:

$$P[a < X < b] = \int_a^b f(x)\, dx$$

pro všechna $a, b \in \mathbb{R}$, $a < b$

Střední hodnota

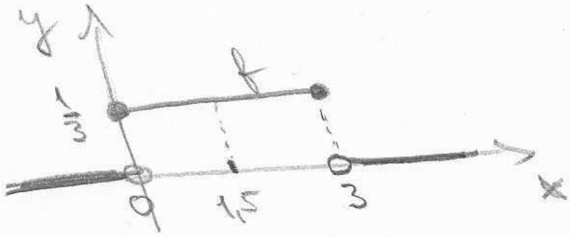
$$EX = \int_{\mathbb{R}} x f(x)\, dx$$

$$E(g(X)) = \int_{\mathbb{R}} g(x) f(x)\, dx$$

Rovnoměrné rozdělení

příklad: Doba odpoledního spánku dítěte je náhodná veličina s rovnoměrným rozdělením na intervalu $(0,3)$ hodiny.

Hustota této veličiny X je $f(x) = \begin{cases} \frac{1}{3}, & x \in \langle 0,3 \rangle \\ 0, & x \notin \langle 0,3 \rangle \end{cases}$



střední hodnota:

$$EX = \int_{\mathbb{R}} x\, f(x)\, dx = \int_0^3 x \cdot \frac{1}{3}\, dx = \frac{1}{3} \cdot \left[\frac{x^2}{2} \right]_0^3 = \frac{3}{2}$$

2. centrální moment:

$$EX^2 = \int_{\mathbb{R}} x^2 f(x)\, dx = \int_0^3 x^2 \cdot \frac{1}{3}\, dx = \frac{1}{3} \cdot \left[\frac{x^3}{3} \right]_0^3 = 3$$

Rozptyl: $\operatorname{var} X = EX^2 - (EX)^2 = 3 - \left(\frac{3}{2}\right)^2 = \frac{3}{4}$

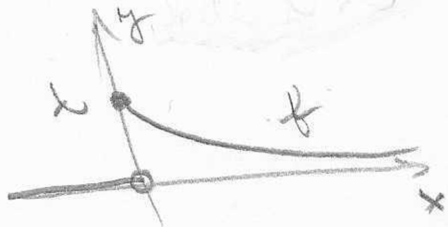
směrodatná odchylka: $\sigma = \sqrt{\operatorname{var} X} = \frac{\sqrt{3}}{2}$

interpretace: „očekávaná" doba spánku bude z intervalu $(EX - \sigma, EX + \sigma) = \left(\frac{3}{2} - \frac{\sqrt{3}}{2}, \frac{3}{2} + \frac{\sqrt{3}}{2}\right)$

Exponenciální rozdělení

– model pro náhodnou veličinu udávající dobu čekání na nějakou událost, dobu životnosti součástky, apod.

hustota: $$f(x) = \begin{cases} \lambda e^{-\lambda x} & , x \geq 0 \\ 0 & , x < 0 \end{cases}$$



kde $\lambda > 0$ je parametr.

střední hodnota: $$EX = \int_{\mathbb{R}} x f(x)\, dx = \int_0^\infty x \cdot \lambda e^{-\lambda x}\, dx =$$

PER-PARTES

$$\overset{\text{PER-PARTES}}{=} \left| \begin{matrix} u = x & v' = \lambda e^{-\lambda x} \\ u' = 1 & v = -e^{-\lambda x} \end{matrix} \right| \underbrace{\left[x \cdot (-e^{-\lambda x}) \right]_0^\infty}_{\lim\limits_{x \to \infty} \left(-\frac{x}{e^{\lambda x}}\right) - 0} + \int_0^\infty e^{-\lambda x}\, dx =$$

$$\boxed{\int u v' = uv - \int u' v}$$

$$\underbrace{\lim_{x \to \infty} \left(-\frac{x}{e^{\lambda x}}\right)}_{} \overset{\text{L.P.}}{=} \lim_{x \to \infty} \left(-\frac{1}{\lambda e^{\lambda x}}\right) = 0$$

$$= 0 + \left[-\frac{1}{\lambda} e^{-\lambda x} \right]_0^\infty = \left(-\frac{1}{\lambda}\right) \Big(\underbrace{\lim_{x \to \infty} \frac{1}{e^{\lambda x}}}_{0} - 1 \Big) = \underline{\underline{\frac{1}{\lambda}}}$$

2. centrální moment: $$EX^2 = \int_{\mathbb{R}} x^2 f(x)\, dx = \int_0^\infty x^2 \lambda e^{-\lambda x}\, dx =$$

$$\overset{\text{P.P.}}{=} \left| \begin{matrix} u = x^2 & v' = \lambda e^{-\lambda x} \\ u' = 2x & v = -e^{-\lambda x} \end{matrix} \right| \underbrace{\left[-x^2 e^{-\lambda x} \right]_0^\infty}_{\lim\limits_{x \to \infty} \left(-\frac{x^2}{e^{\lambda x}}\right) = 0} + \int_0^\infty 2x e^{-\lambda x}\, dx = 2 \cdot \int_0^\infty x e^{-\lambda x}\, dx$$

$$= 2 \cdot \frac{1}{\lambda^2}$$

Rozptyl: $$\operatorname{var} X = EX^2 - (EX)^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda^2}\right) = \frac{1}{\lambda^2}$$

tedy pro $X \sim Exp(\lambda)$ je $f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$,

kde $\lambda > 0$ je parametr a platí: $EX = \frac{1}{\lambda}$

a $\operatorname{var} X = \frac{1}{\lambda^2}$

směrodatná odchylka $\sigma = \sqrt{\operatorname{var} X} = \frac{1}{\lambda}$

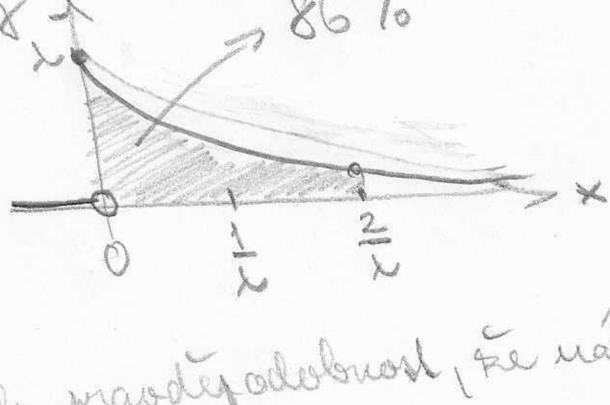
„očekávané" hodnoty budou v intervalu

$$(EX - \sigma, EX + \sigma) = \left(\frac{1}{\lambda} - \frac{1}{\lambda}, \frac{1}{\lambda} + \frac{1}{\lambda}\right) = \left(0, \frac{2}{\lambda}\right)$$

$$P\left[EX - \sigma < X < EX + \sigma\right] = \int_0^{\frac{2}{\lambda}} \lambda \cdot e^{-\lambda x} \, dx =$$

$$= \left[-e^{-\lambda x}\right]_0^{\frac{2}{\lambda}} = -e^{-\lambda \cdot \frac{2}{\lambda}} + e^0$$

$$= \underline{1 - e^{-2}} \doteq 0{,}86$$

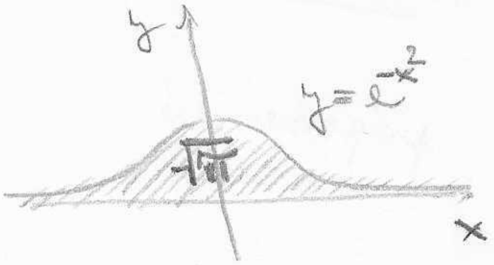


tedy pravděpodobnost, že náhodná veličina s exponenciálním rozdělením bude nabývat hodnot z intervalu „střední hodnota ± směrodatná odchylka"

tj. $(EX - \sigma, EX + \sigma)$ je $1 - e^{-2} \doteq 0{,}86$ nezávisle na parametru λ

Normální rozdělení

připomenutí: $\boxed{\int_{\mathbb{R}} e^{-x^2}\,dx = \sqrt{\pi}}$



Odvození: Označme $I := \int_{\mathbb{R}} e^{-x^2}\,dx$

Víme, že tento integrál je konvergentní, např. podle limitního srovnávacího kritéria, tedy číslo I je konečné.
K funkci e^{-x^2} nenajdeme ale primitivní funkci.

$$I^2 = I\cdot I = \left(\int_{\mathbb{R}} e^{-x^2}\,dx\right)\cdot\left(\int_{\mathbb{R}} e^{-y^2}\,dy\right) = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)}\,dx\,dy$$

FUBINIOVA VĚTA

$$= \int_0^{\infty}\int_0^{2\pi} r\cdot e^{-r^2}\,d\varphi\,dr = \int_0^{\infty} r e^{-r^2}\,dr\cdot\underbrace{\int_0^{2\pi} 1\,d\varphi}_{=2\pi}$$

substituce do polárních souřadnic

$$\left|\begin{array}{l} x = r\cos\varphi \\ y = r\sin\varphi \end{array}\right| \qquad J = r$$

$$= 2\pi\cdot\int_0^{\infty}\frac{1}{2}e^{-t}\,dt = 2\pi\cdot\frac{1}{2}\cdot\underbrace{\left[-e^{-t}\right]_0^{\infty}}_{\lim_{t\to\infty}(-e^{-t})-(-e^{0})}$$

$$\left|\begin{array}{l} r^2 = t \\ 2r\,dr = dt \\ r\,dr = \frac{1}{2}\,dt \end{array}\right|$$

$$= \pi\cdot\left(0-(-1)\right) = \pi \quad\Rightarrow\quad I = \sqrt{\pi}$$

Ukažme, že funkce $f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$ je hustota:

$$\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} \, dx = \frac{1}{\sqrt{2\pi}} \cdot \int_{\mathbb{R}} e^{-\left(\frac{x}{\sqrt{2}}\right)^2} dx =$$

$$= \frac{1}{\sqrt{2\pi}} \cdot \int_{\mathbb{R}} e^{-t^2} \cdot \sqrt{2} \, dt = \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2} \cdot \underbrace{\int_{\mathbb{R}} e^{-t^2} dt}_{\sqrt{\pi}} = 1$$

substituce: $\left| \begin{array}{l} \frac{x}{\sqrt{2}} = t \\ \frac{1}{\sqrt{2}} dx = dt \\ dx = \sqrt{2}\, dt \end{array} \right|$

Tato funkce je hustotou normovaného normálního rozdělení $N(0,1)$ se střední hodnotou 0 a rozptylem 1

Zápis $X \sim N(0,1)$ znamená, že hustota n.v. X je funkce $f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$

snadno ukážeme, že $EX = 0$.

$$EX = \int_{\mathbb{R}} \underbrace{x \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}_{} dx = 0$$

lichá funkce, spojitá na $\mathbb{R}$, jejíž integrál je konvergentní

Distribuční funkci pro $X \sim N(0,1)$ značíme $\Phi(x)$.

Nechť $Z \sim N(0,1)$ a $Y = \mu + \sigma Z$,

kde $\mu \in \mathbb{R}$, $\sigma > 0$ jsou konstanty.

Určeme distribuční funkci této veličiny Y.

$$F_Y(y) = P[Y < y] = P[\mu + \sigma Z < y] =$$

$$= P\left[Z < \frac{y-\mu}{\sigma}\right] = F_Z\left(\frac{y-\mu}{\sigma}\right) =$$

$$= \int_{-\infty}^{\frac{y-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}}\, dt = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx$$

substituce

$$\left| \begin{aligned} x &= \mu + \sigma t \\ dx &= \sigma\, dt \end{aligned} \right| \Rightarrow t = \frac{x-\mu}{\sigma}$$

(dolní mez pro x: $\mu + \sigma\cdot(-\infty) = -\infty$
horní mez pro x: $\mu + \sigma\cdot\frac{y-\mu}{\sigma} = y$)

Náhodná veličina Y má normální rozdělení se střední hodnotou μ a rozptylem σ^2,

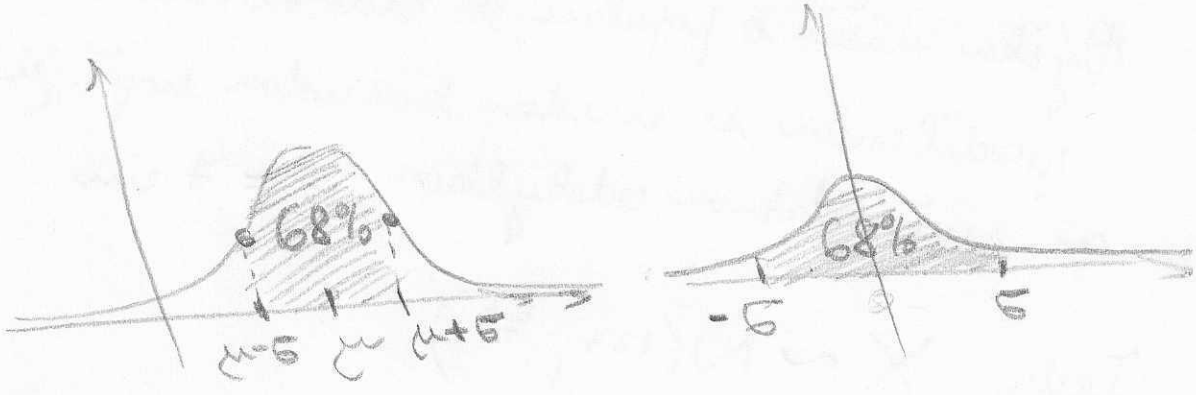
píšeme $Y \sim N(\mu, \sigma^2)$

distribuční funkce: $F_Y(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx$

hustota: $f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

pro každé normální rozdělení $X \sim N(\mu, \sigma^2)$

platí, že $P[\mu-\sigma < X < \mu+\sigma] \doteq 68\,\%$



a „pravidlo dvou sigma":

$$P[\mu-2\sigma < X < \mu+2\sigma] \doteq 95{,}45\,\%$$

a „pravidlo tří sigma":

$$P[\mu-3\sigma < X < \mu+3\sigma] \doteq 99{,}73\,\%$$

příklad: Kolik procent mužů v ČR je menších než 186 cm?

Výška mužů v populaci se řídí normálním rozdělením se střední hodnotou např. $\mu = 177$ cm a směrodatnou odchylkou $\sigma = 7$ cm.

Tedy $X \sim N(177, 7^2)$

$$P[X \leq 186] = P\Big[\underbrace{\frac{X-177}{7}}_{\sim N(0,1)} \leq \frac{186-177}{7}\Big] =$$

$$= \Phi\left(\frac{9}{7}\right) \doteq \underbrace{\Phi(1{,}286)}_{} \doteq \underline{0{,}9}$$

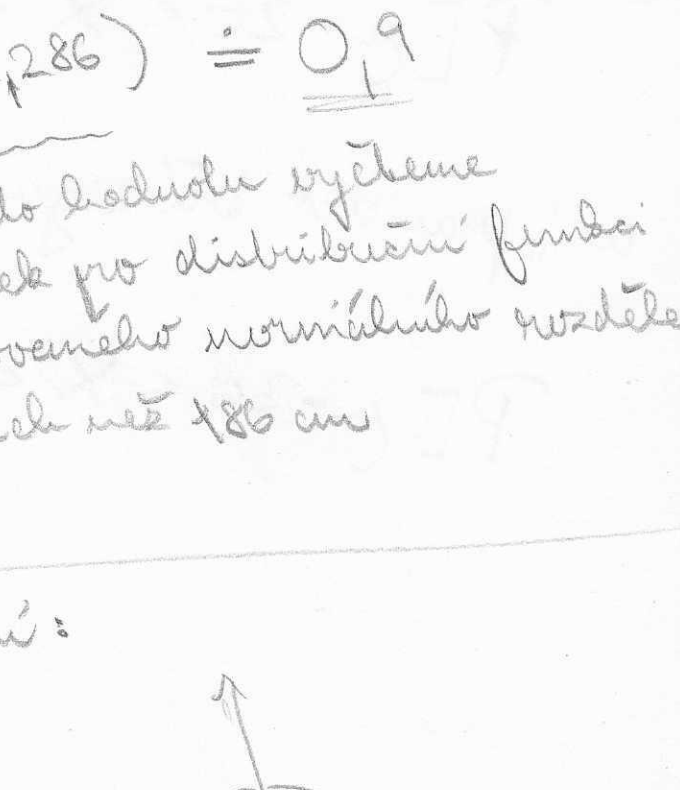
tuto hodnotu vyčteme z tabulek pro distribuční funkci normovaného normálního rozdělení

$\Rightarrow$ 90% mužů je menších než 186 cm

Vlastnosti normálního rozdělení:

$\forall x \in \mathbb{R}: \Phi(x) + \Phi(-x) = 1$

(ze symetrie hustoty)



tato plocha $\Phi(-x)$ je stejně velká jako tato

Nezávislost jevů

Definice: jevy A, B jsou nezávislé, pokud $P(A \cap B) = P(A) \cdot P(B)$

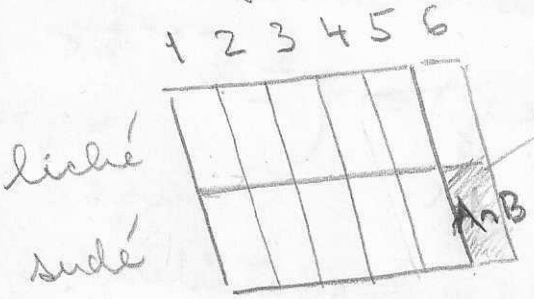
příklad: hodíme 2 kostkami a pozorujeme jevy

A ... na 1. kostce padne 6, $P(A) = \frac{1}{6}$

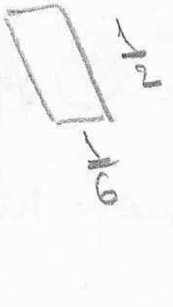
B ... na 2. kostce padne sudé číslo, $P(B) = \frac{1}{2}$

$$\frac{1}{12} = P(A \cap B) = P(A) \cdot P(B) = \frac{1}{6} \cdot \frac{1}{2}$$

jevy A, B jsou nezávislé.



obsah obdélníku



je $\frac{1}{12}$

Co je to nezávislost náhodných veličin?

Tušíme, že náhodné veličiny X ... číslo padlé na 1. kostce

a Y ... na 2. kostce padlo sudé číslo

budou nezávislé, tedy

$\forall k \in \{1,2,3,4,5,6\}$ $\forall m \in \{0,1\}$

platí: $P[(X=k) \cap (Y=m)] = P[X=k] \cdot P[Y=m]$

Definice: Náhodné veličiny X, Y jsou nezávislé, pokud $\forall$ intervaly $I, J \subset \mathbb{R}$: $P[(X\in I)\cap(Y\in J)] = P[X\in I]\cdot P[Y\in J]$

Ekvivalentně: $F(x,y) = F_X(x)\cdot F_Y(y)$ skoro všude

$F(x,y)$ – sdružená distribuční funkce náhodného vektoru (X,Y)

$F_X(x)$, $F_Y(y)$ – marginální distribuční funkce veličin X, Y

nebo pro spojité veličiny: $f(x,y) = f_X(x)\cdot f_Y(y)$ s.v.

$f(x,y)$ – sdružená hustota

$f_X(x)$, $f_Y(y)$ – marginální hustoty

příklad: Dvojice součástek má dobu životnosti popsanou sdruženou hustotou $f(x,y) = \begin{cases} \frac{1}{2}e^{-x-\frac{1}{2}y} & x,y\geq 0 \\ 0, & x,y<0 \end{cases}$

Jaké je rozdělení dob životnosti jednotlivých součástek? Jsou tyto doby nezávislé?

$$f_X(x) = \int_{\mathbb{R}} f(x,y)\,dy = \int_0^\infty \frac{1}{2}e^{-x-\frac{1}{2}y}\,dy = \frac{1}{2}e^{-x}\cdot\left[-2e^{-\frac{1}{2}y}\right]_0^\infty =$$

$$= \frac{1}{2}e^{-x}\cdot\Big((-2)\big(\underbrace{\lim_{y\to\infty} e^{-\frac{1}{2}y}}_{0} - 1\big)\Big) = \frac{1}{2}e^{-x}\cdot 2 = \underline{\underline{e^{-x}}}$$

$$f_Y(y) = \int_{\mathbb{R}} f(x,y)\,dx = \int_0^\infty \frac{1}{2}e^{-x-\frac{1}{2}y}\,dx = \frac{1}{2}e^{-\frac{1}{2}y}\cdot\underbrace{\left[-e^{-x}\right]_0^\infty}_{\lim_{x\to\infty}(-e^{-x})-(-1)} = \underline{\underline{\frac{1}{2}e^{-\frac{y}{2}}}}$$

Ano, veličiny X, Y jsou nezávislé, neboť $f(x,y) = f_X(x)\cdot f_Y(y)$

$$\frac{1}{2}e^{-x-\frac{y}{2}} = e^{-x}\cdot\frac{1}{2}e^{-\frac{y}{2}}$$