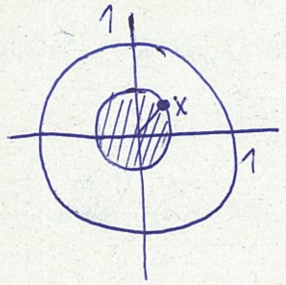
iný spôsob (lepší):

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\frac{\pi}{2}}^{x} \frac{1}{2}\cos t\,dt = \left[\frac{1}{2}\sin t\right]_{-\frac{\pi}{2}}^{x} = \frac{1}{2}\sin x - \frac{1}{2}\sin\left(-\frac{\pi}{2}\right) =$$

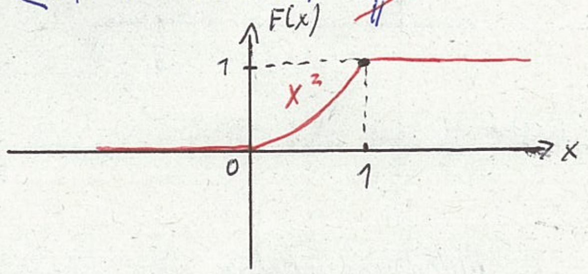
$$= \frac{1}{2}\sin x + \frac{1}{2}$$

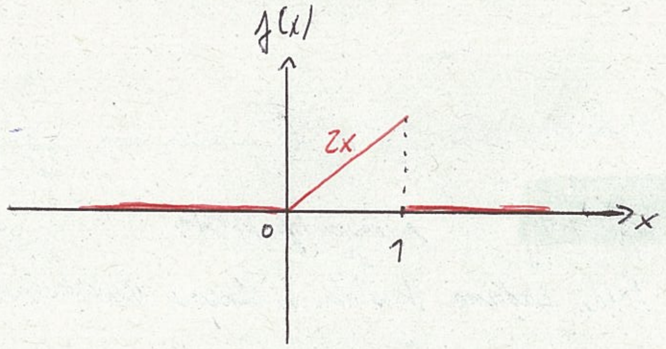
**Pr → z prednášky**



X = ~~náhodná~~ vzdialenosť náhodne vybraného bodu od počiatku

$x \in \langle 0, 1\rangle : P(X \le x) = \frac{\pi \cdot x^2}{\pi} = x^2$





**Stredná hodnota:**

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

**Rozptyl:**

$$DX = \int_{-\infty}^{\infty} \left(x - \frac{2}{3}\right)^2 f(x)\,dx = \int_0^1 \left(x - \frac{2}{3}\right)^2 \cdot 2x\,dx \ldots$$ → varianta 1

$$= \int_{-\infty}^{\infty} x^2 f(x)\,dx - \left(\frac{2}{3}\right)^2 = \int_0^1 x^2 \cdot 2x\,dx - \frac{4}{9} = \left[\frac{1}{2}x^4\right]_0^1 - \frac{4}{9} = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$$ → varianta 2 (lepšia!)

**Smerodajná odchýlka:**

$$\sigma = \sqrt{DX} = \frac{1}{\sqrt{18}} = \frac{1}{3\cdot\sqrt{2}}$$

## Exponenciálne rozdelenie

→ Nech platia rovnaké predpoklady ako u Poissonovho rozdelenia.

→ Náhodná veličina X, ktorá udáva dobu medzi dvoma výskytmi určitej udalosti (alebo tiež dobu čakania na ďalšiu udalosť), keď vieme že priemerne nastáva $\lambda$ udalostí za jednotku času, má exponenciálne rozdelenie pravdepodobnosti s parametrom $\lambda$, píšeme

$X \sim Exp(\lambda)$

za jednotku času ... $\lambda$ udalostí
za $x$ jednotiek času ... $\lambda x$ udalostí

$F(x) = P(X \le x) = 1 - P(X \ge x)$

↑

$X \sim Exp(\lambda)$

$P(0) = \frac{(\lambda x)^0}{0!} \cdot e^{-\lambda x}$ → v Poissonovom rozdelení

**Hustota exponenciálneho rozdelenia je:**

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{pre } x \ge 0 \\ 0 & \text{pre } x < 0 \end{cases}$$

**Distribučná funkcia je:**

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{pre } x \ge 0 \\ 0 & \text{pre } x < 0 \end{cases}$$

**Stredná hodnota exponenciálneho rozdelenia:**

$$EX = \frac{1}{\lambda}$$

**Rozptyl:**

$$DX = \frac{1}{x^2}$$

**Pr → z prednášky**

V továrni máme výrobní linku, ktorá sa z času na čas pokazí

za 8 hodín ... 3 poruchy

Aká je pravdepodobnosť že porucha nastane do $\frac{1}{2}$ hodiny

1 hodina ... $\frac{3}{8}$ poruchy, $\lambda = \frac{3}{8}$, $X \sim EXP\left(\frac{3}{8}\right)$

↑ čas čakania na poruchu

$$P\left(X \le \frac{1}{2}\right) = 1 - e^{-\frac{3}{8}\cdot\frac{1}{2}} = 1 - e^{-\frac{3}{16}}$$

26.11.2018

# Normální rozdelení

**Definícia:**

→ Náhodná veličina $X$, ktorá má hustotu $\boxed{f(x) = \frac{1}{\sigma\cdot\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}}$ pre $x \in \mathbb{R}$,

kde $\mu \in \mathbb{R}$ a $\sigma > 0$ sú konštanty, má normální rozdelenie pravdepodobnosti so strednou hodnotou

**Stredná hodnota:** $\boxed{EX = \mu}$ Rozptyl: $\boxed{DX = \sigma^2}$

píšeme: $\boxed{X \sim N(\mu, \sigma^2)}$

**Štandardizované normální rozdelenie:**

→ je rozdelenie so strednou hodnotou $\mu = 0$ a rozptylom $\sigma^2 = 1$.

→ Náhodnú veličinu so štandardizovaným normálnym rozdelením značíme $U$, platí teda: $U \sim N(0,1)$

→ Hustota štandardizovaného normálneho rozdelenia je: $\boxed{f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}}$ pre $x \in \mathbb{R}$

→ Distribučnú funkciu náhodnej veličiny $U$ značíme $\Phi$ a platí:

$$\boxed{\Phi(u) = P(U \leq u)}$$

**Vlastnosti distribučnej funkcie $\Phi$:**

→ $\Phi(u)$ sa nedá vyjadriť pomocou elementárnej funkcie

→ hodnoty $\Phi(u)$ pre $u \geq 0$ sú tabelované

→ platí $\boxed{\Phi(-u) = 1 - \Phi(u)}$, pre kvantily náhodnej veličiny $U$ platí $\boxed{u_\alpha = -u_{1-\alpha}}$

→ tento vzťah sa používa predovšetkým pre $\alpha < 0{,}5$

**Výpočty rôznych pravdepodobností pomocou $\Phi$**

→ $P(U \leq b) = \Phi(b)$

→ $P(U > a) = 1 - \Phi(a)$

→ $P(a < U \leq b) = \Phi(b) - \Phi(a)$

PZ: Pretože $U$ je spojitá náhodná veličina, všetky ostré nerovnosti lze nahradiť neostrými a naopak

**Transformácia na štandardizované normálne rozdelenie:**

- Ak $X \sim N(\mu, \sigma^2)$ potom náhodná veličina $\boxed{U = \frac{X-\mu}{\sigma}}$ má štandardizované normálne rozdelenie

**Výpočty rôznych pravdepodobností u normálneho rozdelenia všeobecne**

→ Ak $X \sim N(\mu, \sigma^2)$ potom

→ $P(X \leq b) = P\left(U \leq \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right)$

→ $P(X > a) = P\left(U > \frac{a-\mu}{\sigma}\right) = 1 - \Phi\left(\frac{a-\mu}{\sigma}\right)$

→ $P(a < X < b) = P\left(\frac{a-\mu}{\sigma} < U < \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$

PZ: všetky ostré nerovnosti sa dá nahradiť neostrými a naopak

**Súčet dvoch nezávislých náhodných veličín s normálnym rozdelením**

→ Ak $X \sim N(\mu_1, \sigma_1^2)$ a $Y \sim N(\mu_2, \sigma_2^2)$ sú dve nezávislé náhodné veličiny, potom

$$\boxed{X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)}$$

**Náhodný výber (všeobecne nie len pre normálne rozdelenie)**

- náhodný výber rozsahu $n$ je vektor $(X_1, X_2, \dots, X_n)$ kde $X_1, X_2, \dots, X_n$ sú navzájom nezávislé náhodné veličiny, ktoré majú všetky rovnaké rozdelenie pravdepodobnosti

**Výberový priemer (všeobecne, nie len pre normálne rozdelenie)**

→ výberový priemer z náhodného výberu $(X_1, X_2, \dots, X_n)$ je $\boxed{\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i}$, výberový priemer $\bar{X}$ je náhodná veličina

**Rozdelenie pravdepodobnosti výberového priemeru z $N(\mu, \sigma^2)$**

→ ak náhodný výber pochádza z normálneho rozdelenia so strednou hodnotou $\mu$ a rozptylom $\sigma^2$ potom výberový priemer $\bar{X}$ má tiež normálne rozdelenie $\boxed{\bar{X} \sim N(\mu_{\bar{X}}, \sigma^2_{\bar{X}})}$ kde $\boxed{\mu_{\bar{X}} = \mu}$, $\boxed{\sigma^2_{\bar{X}} = \frac{\sigma^2}{n}}$

**Pr → z prednášky**

z tabuľky → 283 str

$$P(0{,}03 \le U \le 0{,}1) = \Phi(0{,}1) - \Phi(0{,}03) = 0{,}5398 - 0{,}512 = \underline{\underline{0{,}0278}}$$

**Pr → z prednášky**

$X \sim N(2,9)$

$\mu = 2, \cancel{\sigma} = 3 \quad \sigma = 3$

$$P(3 \le X \le 3{,}5) = P\left(\frac{3-2}{3} \le \frac{X-2}{\sigma} \le \frac{3{,}5-2}{3}\right) = P(0{,}33 \le U \le 0{,}5) = \Phi(0{,}5) - \Phi(0{,}3) =$$

$$= 0{,}6915 - 0{,}6293 = \underline{\underline{0{,}0622}}$$

**Pr → z prednášky**

Maslo: 250 [g]

Norma: 245[g] − 255[g]

$\}$ X = hmotnosť masla, $X \sim (\mu, \sigma^2)$

$\mu = 248$ [g], $\sigma = 2{,}1$ [g]

↓ stredná hodnota ↓ odchýlka

Aká je pravdepod. že náhodne vybraná tabla masla bude v norme?

$$P(245 \le X \le 255) = P\left(\frac{245-248}{2{,}1} \le U \le \frac{255-248}{2{,}1}\right) = P(-1{,}429 \le U \le 3{,}333) = \Phi(3{,}33) - \Phi(-1{,}429) =$$

$$= 0{,}999 - (1 - \Phi(1{,}429)) = 0{,}999 - (1 - 0{,}924) = 0{,}999 - 0{,}076 = \underline{\underline{0{,}923}}$$

## Centrálna limitná veta

**Veta (Lindeberg-Lévyho)**

→ Ak $X_1, X_2, \dots, X_n$ sú navzájom nezávislé náhodné veličiny, ktoré majú rovnaké rozdelenie so strednou hodnotou $EX_i = \mu$ a konečným rozptylom $DX_i = \sigma^2$, potom pre súčet a priemer týchto náhodných veličín platí:

$$Y = \sum_{i=1}^{n} X_i \quad a \quad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{platí:}$$

$$\lim_{n\to\infty} P\left(\frac{Y - n\mu}{\sqrt{n}\cdot\sigma} \le u\right) = \Phi(u) \qquad \lim_{n\to\infty} P\left(\sqrt{n}\cdot\frac{\bar{X}-\mu}{\sigma} \le u\right) = \Phi(u)$$

pre každé $u \in \mathbb{R}$

**Iným slovami:**

Pre dostatočne veľké $n$ (za prijateľné sa považuje $n > 30$) majú $Y$ a $\bar{X}$ približne normálne rozdelenie

$$Y \sim N(n\mu, n\sigma^2), \quad \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

alebo:

$$\frac{Y - n\mu}{\sigma\cdot\sqrt{n}} \doteq U, \quad \sqrt{n}\cdot\frac{\bar{X}-\mu}{\sigma} \doteq U$$

## Aproximácia binomického rozdelenia normálnym

**Moivre-Laplaceova veta:**

→ Ak $X$ je náhodná veličina s binomickým rozdelením, $X \sim Bi(n,p)$, $n$ je veľké a $p$ nie je príliš blízko nule ani jednotke, potom $X$ sa dá aproximovať normálnym rozdelením so rovnakou strednou hodnotou a rozptylom, ako malo pôvodné bin. rozdelenie

$$X \sim Bi(n,p) \Rightarrow X \sim N(\mu, \sigma^2)$$

kde $\mu = n\cdot p$, $\sigma^2 = n\cdot p\cdot(1-p)$,

$$P(a < X \le b) \doteq P\left(\frac{a-\mu}{\sigma} < U \le \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

**Kedy sa dá aproximácia použiť? - empirická podmienka**

→ aproximácia je dobrá, ak $n\cdot p > 5$ a súčasne $n\cdot(1-p) > 5$

→ iná podmienka: $n\cdot p\,(1-p) > 9$

**Aproximácia s korekciou**

→ lepšie výsledky dáva aproximácia s korekciou:

↳ Ak $a, b$ sú celé čísla a $X \sim Bi(n,p)$, potom

$$P(a \le X \le b) \doteq P(a - 0{,}5 < X < b + 0{,}5) = \Phi\left(\frac{b+0{,}5-\mu}{\sigma}\right) - \Phi\left(\frac{a-0{,}5-\mu}{\sigma}\right)$$

**Pr → z prednášky**

→ 100x hodíme kockou, aká je prav., že $\boxed{6}$ padne 20x až 30x

→ X = počet úspechov, $X \sim Bi\left(100, \frac{1}{6}\right)$, $\mu = \frac{100}{6}$, $\sigma = \sqrt{100\cdot\frac{1}{6}\cdot\frac{5}{6}}$, → $\mu = \underline{\underline{16{,}67}}$, $\sigma = \underline{\underline{3{,}727}}$

$$P(20 \le X \le 30) = P(19{,}5 < X < 30{,}5) = P\left(\frac{19{,}5-16{,}67}{3{,}727} \le U \le \frac{30{,}5-16{,}67}{3{,}727}\right) = P(0{,}76 \le U \le 3{,}71) =$$

$$= \Phi(3{,}71) - \Phi(0{,}76) = 0{,}9999 - 0{,}7764 = \underline{\underline{0{,}2235}}$$

3.12.2018

# Testovanie hypotéz

## Všeobecný postup pri testovaní hypotéz:

→ Vyslovíme nulovú hypotézu $H_0$ a alternatívnu hypotézu $H_1$.

→ Zvolíme testové kritérium

→ Predpokladáme, že platí $H_0$. Za tohto predpokladu poznáme rozdelenie testového kritéria

→ nájdeme kritický obor pre testové kritérium – množinu (interval), do ktorej padne ho s pravdepodobnosťou $\alpha$.

→ Leží hodnota testového kritéria v kritickom obore? Pokiaľ áno nulovú hypotézu $H_0$ zamietneme. Hypotéza $H_1$ bola testom preukázaná. Pokiaľ nie, nulovú hypotézu $H_0$ nezamietneme. Hypotéza $H_1$ nebola preukázaná

## Možné výsledky testovania:

→ Pravdepodobnosť chyby 1. druhu je $\alpha$

→ Pravdepodobnosť chyby 2. druhu je $\beta$

Čím väčšie je $\alpha$ tým je menšie $\beta$ a naopak

## Sila testu:

→ Sila testu je pravdepodobnosť, že správne zamietneme $H_0$, keď v skutočnosti platí $H_1$

$$\boxed{\text{Sila testu} = 1-\beta}$$

## Jednostranné a obojstranné testy:

Nulová hypotéza býva spravidla typu „určitý parameter = určitá hodnota"

Ak je napríklad $H_0: \mu = \mu_0$

alternatívna hypotéza $H_1$ potom môže byť niekoľko rôznych typov:

→ $H_1: \mu \neq \mu_0$

↳ obojstranný test, kritický obor je tvaru $(-\infty, T_{K1}) \cup (T_{K2}, \infty)$

→ $H_1: \mu > \mu_0$

↳ jednostranný (pravostranný test), kritický obor je tvaru $(T_K, \infty)$

→ $H_1: \mu < \mu_0$

↳ jednostranný (ľavostranný test), kritický obor je tvaru $(-\infty, T_K)$

**Pr → z prednášky**

100x hodím kockou, 30x padne $\boxed{6}$ → nadpriemerne veľa, treba to overiť

$H_0$... "náhoda"

$H_1$... "nejedná sa o náhodný výsledok z binomického rozdelenia s parametrami $n = 100$, $p = \frac{1}{6}$"

musíme si stanoviť hladinu významnosti: $\alpha = 0{,}01$

$X \sim Bi(100, \frac{1}{6})$ ← nahradíme normálnym rozdelením

$\mu = \frac{100}{6} = 16{,}667$ ; $\sigma = \sqrt{100 \cdot \frac{1}{6} \cdot \frac{5}{6}} = 3{,}727$

$X \sim N(\mu, \sigma^2)$ ← nahradené binomické rozdelenie normálnym

$P(X \geq 30) = P(X > 29{,}5)$ → korekcia medzí

$P\left(U > \frac{29{,}5 - 16{,}667}{3{,}727}\right) \doteq P(U > 3{,}44) = 1 - \Phi(3{,}44) \doteq 1 - 0{,}9997 = 0{,}0003 < \alpha$ ← Test pomocou P-hodnoty

Výsledok: $H_0$ zamietame
$H_1$ prijímame

**Pr → z prednášky → pravostranný test**

Chceme nájsť medzu (prelomový bod)

$P(X \geq t) = 0{,}01$

$P(X \leq t) = 0{,}99$

$P(X < t + 0{,}5) = 0{,}99$

$P\left(U < \frac{t + 0{,}5 - 16{,}667}{3{,}727}\right) = 0{,}99$ → u = chceme nájsť

$\Phi(u) = P$ u... P-kvantil ⇒ $u_P$

$u_{0{,}99} = 2{,}33$

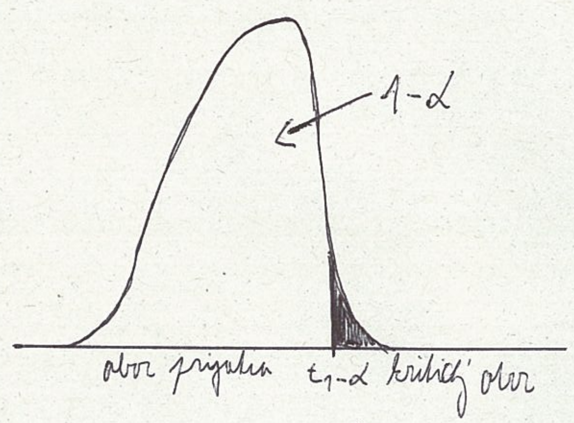
$\frac{t + 0{,}5 - 16{,}667}{3{,}727} = 2{,}33$

$t = 16{,}667 + 2{,}33 \cdot 3{,}727 - 0{,}5 =$

$t = 24{,}85$

$\underline{\underline{t = 25}}$ ← naša hľadaná hranica (medza)

počet $\boxed{6} > 25$ je už podozrivé




Pravostranný test