Statistika

(D360P03Z, D360P03U)

Karel Zvára

22. listopadu 2004

schéma testování hypotéz

rozhodnutí	H_0 platí	H_0 neplatí
H_0 zamítnout	chyba 1. druhu (pst $\leq \alpha$) hladina testu	správné rozhodnutí (pst $1-\beta$) síla testu
H_0 nezamítnout (přijmout)	správné rozhodnutí (pst $\geq 1-\alpha$)	chyba 2. druhu (pst β)

příklad: výšky desetiletých hochů

- velký výběr v roce 1951 dal průměr $136{,}1$ cm, rozptyl $6{,}4^2$ cm^2
- v roce 1961 naměřeno v náhodném výběru $n = 15$ hodnot s průměrem $\bar{X} = 139{,}13$ cm (lze předpokládat nezměněný rozptyl)
- prokázali jsme na 5% hladině představu, že desetiletí hoši jsou (co do populačního průměru) v roce 1961 *větší* než desetiletí hoši v roce 1951?
- hypotéza $\mathsf{H}_0 : \mu = \mu_0 = 136{,}1$ (nebo $\mu \leq 136{,}1$, postup by byl stejný)
- alternativa $\mathsf{H}_1 : \mu > 136{,}1$

výšky desetiletých hochů

- alternativě nasvědčují průměry o hodně větší než $\mu = \mu_0 = 136{,}1$
- kritický obor: $\bar{X} \geq x_0$, kde x_0 je zvoleno tak, aby za platnosti hypotézy bylo překročeno s pstí nejvýš 5 %
- platí (za platnosti hypotézy)

$$\bar{X} \sim \mathsf{N}\Big(136{,}1, 6{,}4^2/15\Big) \qquad \Rightarrow Z = \frac{\bar{X} - 136{,}1}{6{,}4}\sqrt{15} = \frac{\bar{X} - 136{,}1}{\mathsf{S.E.}(\bar{X})} \sim \mathsf{N}(0,1)$$

- proto hypotézu zamítáme, je-li $Z > z(0{,}05) = 1{,}645$
- v našem příkladu je $Z_0 = \frac{139{,}13-136{,}1}{6{,}4}\sqrt{15} = 1{,}82 > 1{,}645$, takže na 5% hladině hypotézu **zamítáme ve prospěch jednostranné alternativy, že populační průměr za deset roků vzrostl**

obecně (jednostranná alternativa)

- $\mathsf{H}_0 : \mu = \mu_0$
- $\mathsf{H}_1 : \mu > \mu_0$
- kritický obor: $\bar{X} \geq x_0$, kde x_0 je zvoleno tak, aby za platnosti hypotézy bylo překročeno s pstí nejvýš 5 %
- platí (za platnosti hypotézy)

$$\bar{X} \sim \mathsf{N}\Big(\mu_0, \sigma^2/n\Big) \qquad \Rightarrow Z = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n} \sim \mathsf{N}(0,1)$$

- proto hypotézu zamítáme na hladině α, je-li $Z > z(\alpha)$
- $\mathsf{H}_1 : \mu < \mu_0$, podobně jako výše hypotézu zamítáme na hladině α, je-li $Z < -z(\alpha)$

obecně

- $\mathsf{H}_0 : \mu = \mu_0$
- $\mathsf{H}_1 : \mu \neq \mu_0$
- kritický obor: $\bar{X}$ je příliš daleko od μ_0,
- platí (za platnosti hypotézy)

$$\bar{X} \sim \mathsf{N}\Big(\mu_0, \sigma^2/n\Big) \qquad \Rightarrow Z = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n} \sim \mathsf{N}(0,1)$$

- protože hladinu musíme rozdělit na dvě části ($\bar{X} << \mu_0$ a $\bar{X} >> \mu_0$ hypotézu zamítáme na hladině α, je-li $|Z| > z(\alpha/2)$

výšky desetiletých hochů

- výpočet p-hodnoty: (v Z odečítáme vždy skutečně platnou střední hodnotu, uvedeme ji jako dolní index u P)
  za H_0 je $Z = \frac{\bar{X}-136{,}1}{6{,}4}\sqrt{15} \sim \mathsf{N}(0,1)$

$$\begin{aligned} p &= \mathsf{P}_{136{,}1}\,(\bar{X} \geq 139{,}1) \\ &= \mathsf{P}_{136{,}1}\left(\frac{\bar{X}-136{,}1}{6{,}4}\sqrt{15} \geq \frac{139{,}13-136{,}1}{6{,}4}\sqrt{15}\right) \\ &= \mathsf{P}(Z \geq 1{,}82) = 1-\Phi(1{,}82) = 1-0{,}965 = 0{,}035 < 0{,}05 \end{aligned}$$

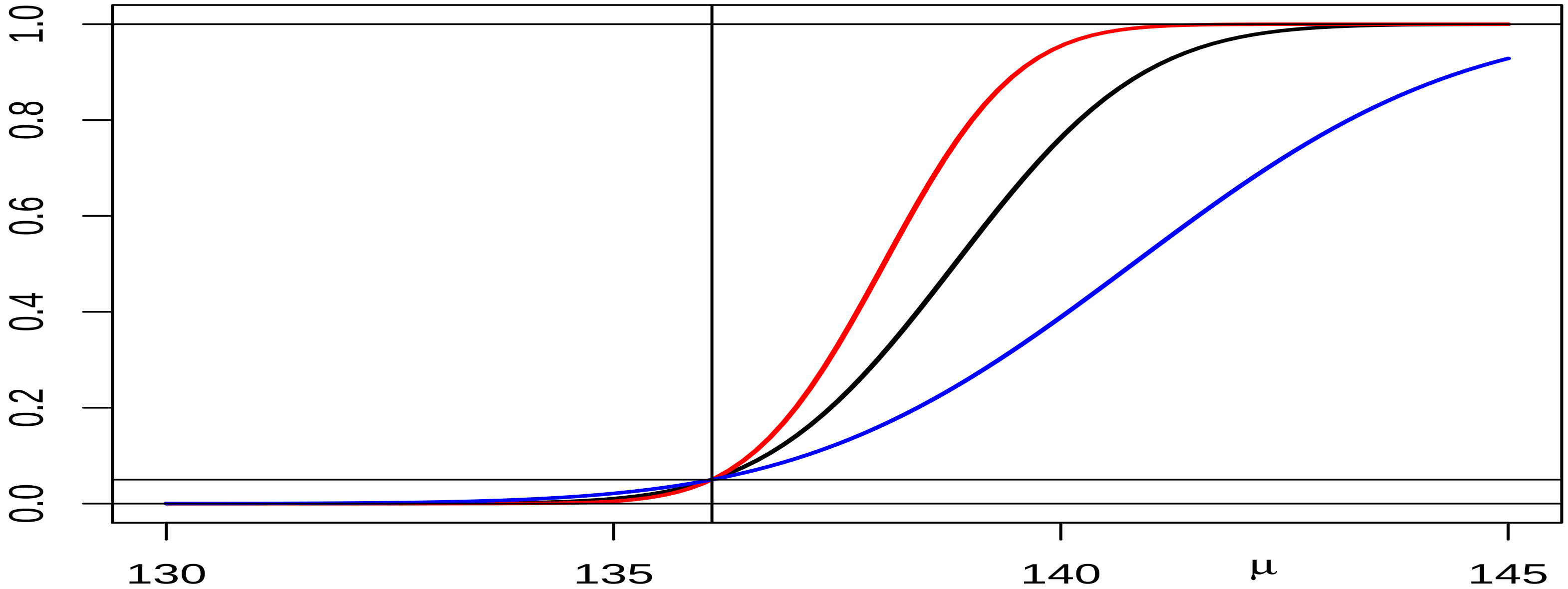
- na 5% hladině jsme zamítli hypotézu ve prospěch jednostranné alternativy (kterou jsme zvolili předem, bez znalosti dat!)
- prokázali jsme na 5% hladině vzrůst populačního průměru

síla testu pro $\mu = 140$

- síla testu = pst(zamítnout hypotézu, když tato neplatí)
- musíme vzít v úvahu, že $\mu = 140$

$$\begin{aligned}
1-\beta(140) &= \mathsf{P}_{140}\left(\frac{\bar{X}-136{,}1}{6{,}4}\sqrt{15} > 1,645\right)\\
&= \mathsf{P}_{140}\left(\frac{\bar{X}-140}{6{,}4}\sqrt{15} + \frac{140-136{,}1}{6{,}4}\sqrt{15} > 1,645\right)\\
&= \mathsf{P}\left(Z > 1{,}645 - \frac{140-136{,}1}{6{,}4}\sqrt{15}\right) = \mathsf{P}(Z \geq -0{,}715)\\
&= 1-\Phi(-0{,}715) = 1-0{,}237 = 0{,}763
\end{aligned}$$

síla testu v závislosti na μ, n =15 (30, 5)



shrnutí: $X_1, \ldots, X_n \sim \mathsf{N}\left(\mu, \sigma^2\right)$, nezávislé

- předpokládáme, že $\sigma > 0$ známe
- $\mathsf{H}_0 : \mu = \mu_0$ (μ_0 známá konstanta)

$$Z = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n} = \frac{\bar{X} - \mu_0}{\mathsf{S.E.}(\bar{X})}$$

- kdy hypotézu H_0 zamítáme (kritický obor):
  - $\mathsf{H}_1 : \mu \neq \mu_0$ (oboustranná alternativa) $\quad |Z| \geq z(\alpha/2)$
  - $\mathsf{H}_1 : \mu > \mu_0$ (jednostranná alternativa) $\quad Z \geq z(\alpha)$
  - $\mathsf{H}_1 : \mu < \mu_0$ (jednostranná alternativa) $\quad Z \leq -z(\alpha)$

častěji: $X_1, \ldots, X_n \sim \mathsf{N}\left(\mu, \sigma^2\right)$, nezávislé

- neznámé $\sigma > 0$ odhadneme pomocí $s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2}$
- $\mathsf{H}_0 : \mu = \mu_0$ (μ_0 známá konstanta)

$$T = \frac{\bar{X} - \mu_0}{s} \sqrt{n} = \frac{\bar{X} - \mu_0}{\widehat{\mathsf{S.E.}}(\bar{X})}$$

- kdy hypotézu H_0 zamítáme (kritický obor):
  - $\mathsf{H}_1 : \mu \neq \mu_0$ (oboustranná alternativa) $\quad |T| \geq t_{n-1}(\alpha)$
  - $\mathsf{H}_1 : \mu > \mu_0$ (jednostranná alternativa) $\quad T \geq t_{n-1}(2\alpha)$
  - $\mathsf{H}_1 : \mu < \mu_0$ (jednostranná alternativa) $\quad T \leq -t_{n-1}(2\alpha)$

souvislost s intervalem spolehlivosti

- připomeňme interval spolehlivosti pro μ

$$\bar{X} - \frac{s_x}{\sqrt{n}} t_{n-1}(\alpha) < \mu < \bar{X} + \frac{s_x}{\sqrt{n}} t_{n-1}(\alpha)$$
$$\bar{X} - \widehat{\mathsf{S.E.}}(\bar{X}) \cdot t_{n-1}(\alpha) < \mu < \bar{X} + \widehat{\mathsf{S.E.}}(\bar{X}) \cdot t_{n-1}(\alpha)$$

  což lze přepsat jako

$$|T| = \left| \frac{\bar{X} - \mu}{s_x} \sqrt{n} \right| < t_{n-1}(\alpha)$$

- H$_0$: $\mu = \mu_0$ tedy **nezamítneme** na hladině α při oboustranné alternativě, právě když μ_0 leží v $100(1-\alpha)$% intervalu spolehlivosti

- interval spolehlivosti tedy obsahuje takové hodnoty μ_0, které bychom jako hypotézu nezamítli

výšky desetiletých hochů (σ^2 neznámé)

- kritický obor: $\bar{X}$ se příliš liší od μ_0 ve směru zvolené alternativy
- spočítáme

$$s = \sqrt{\frac{1}{15-1}((130-139{,}13)^2 + \ldots + (141-139{,}13)^2)} = \sqrt{42{,}98} = 6{,}56$$

$$T = \frac{\bar{X} - 136{,}1}{6{,}56}\sqrt{15} = 1{,}79$$

- na 5% hladině při jednostranné alternativě $\mu > \mu_0$ hypotézu zamítáme, neboť $t_{14}(0{,}10) = 1{,}76$ ($p = 4{,}7$ %)
- na 5% hladině při oboustranné alternativě hypotézu nezamítáme, neboť $t_{14}(0{,}05) = 2{,}14$ ($p = 9{,}5$ %)
- 95% int. spolehlivosti pro populační průměr výšek hochů: $(135{,}5; 142{,}8)$

nová úloha: porovnání dvou populací

- liší se desetileté dívky výškou postavy od desetiletých hoch?
- lze předpokládat, že výšky hochů

$$X_i \sim \mathsf{N}\Big(\mu_1, \sigma^2\Big), \quad i = 1, \ldots, n_1$$

- lze předpokládat, že výšky dívek

$$Y_i \sim \mathsf{N}\Big(\mu_2, \sigma^2\Big), \quad i = 1, \ldots, n_2$$

- předpoklad stejných rozptylů bývá splněn, lze jej ověřit
- musí jít o **nezávislé** náhodné výběry, nelze např. vybírat sourozenecké dvojice

porovnání středních hodnot nezávislých výběrů

- zřejmě $\mathsf{H}_0 : \mu_1 = \mu_2$ (není rozdíl, **nulová** hypotéza)
- alternativy
  - $\mathsf{H}_1 : \mu_1 \neq \mu_2$ (není-li důvod k jednostranné alternativě)
  - $\mathsf{H}_1 : \mu_1 > \mu_2$ (bylo cílem dokázat, že hoši větší dívek)
  - $\mathsf{H}_1 : \mu_1 < \mu_2$ (bylo cílem dokázat, že hoši menší dívek)
- rozhodování založeno na porovnání průměrů $\bar{X}$ a $\bar{Y}$; čím více se liší, tím spíše zamítnout hypotézu
- je třeba porovnat s mírou přesnosti, s jakou průměry $\bar{X}, \bar{Y}$ odhadnou skutečné populační průměry μ_1, μ_2

porovnání středních hodnot nezáv. výběrů (2)

- k tomu je třeba odhadnout také neznámé σ^2 pomocí

$$s^2 = \frac{1}{n_1+n_2-2}\left(\sum_{i=1}^{n_1}(X_i-\bar{X})^2+\sum_{i=1}^{n_2}(Y_i-\bar{Y})^2\right)$$
$$= \frac{n_1-1}{n_1+n_2-2}s_X^2+\frac{n_2-1}{n_1+n_2-2}s_Y^2$$

  (vážený průměr odhadů rozptylu v obou výběrech)

- výška desetiletých dětí: $n_1 = 15$, $n_2 = 12$, $\bar{X} = 139{,}13$, $\bar{Y} = 140{,}83$, $s_X^2 = 42{,}98$, $s_Y^2 = 33{,}79$, tudíž $s^2 = 38{,}94 = 6{,}24^2$

kritické obory

- o $\mathsf{H}_0 : \mu_1 = \mu_2$ se rozhoduje pomocí

$$T = \frac{\bar{X} - \bar{Y}}{s}\sqrt{\frac{n_1 n_2}{n_1 + n_2}} = \frac{\bar{X} - \bar{Y}}{\widehat{\mathsf{S.E.}}(\bar{X} - \bar{Y})}$$

- $\mathsf{H}_1 : \mu_1 \neq \mu_2$ zamítáme pokud $|T| \geq t_{n_1+n_2-2}(\alpha)$
- $\mathsf{H}_1 : \mu_1 > \mu_2$ zamítáme pokud $T \geq t_{n_1+n_2-2}(2\alpha)$
- $\mathsf{H}_1 : \mu_1 < \mu_2$ zamítáme pokud $T \leq -t_{n_1+n_2-2}(2\alpha)$
- výšky desetiletých: $T = -0{,}70 < 2{,}06 = t_{15+12-2}(0{,}05)$
- na 5% hladině jsme **neprokázali** rozdíl mezi výškami desetiletých hochů a dívek ($p = 48{,}8$ %)

souvislost s intervalem spolehlivosti

- $\mu_1 - \mu_2 = \delta$ o kolik se liší populační průměrné výšky
- odhadem pro δ je $d = \bar{X} - \bar{Y} = -1{,}7$
- interval spolehlivosti pro delta je

  $(\bar{X}-\bar{Y}) - \widehat{\mathrm{S.E.}}(\bar{X}-\bar{Y}) \cdot t_{n_1+n_2-2}(\alpha) < \delta < (\bar{X}-\bar{Y}) + \widehat{\mathrm{S.E.}}(\bar{X}-\bar{Y}) \cdot t_{n_1+n_2-2}(\alpha)$

  H$_0$ zamítáme právě tehdy, když nula **není** v int. spol. pro δ
- při porovnání výšek hochů a dívek je 95% interval pro δ

$$\left(-1{,}7 - 6{,}24\sqrt{\frac{1}{15}+\frac{1}{12}} \cdot 2{,}06;\, -1{,}7 + 6{,}24\sqrt{\frac{1}{15}+\frac{1}{12}} \cdot 2{,}06\right)$$

$$(-3{,}3; 6{,}7)$$

příklad: přijímačky na MFF

- liší se úrovní znalosti matematiky uchazeči o studium matematiky a fyziky?
- $n_1 = 104$, $\bar{X} = 34{,}6$, $s_x = 11{,}4 = \sqrt{129{,}4}$,
  $n_2 = 114$, $\bar{Y} = 31{,}2$, $s_x = 10{,}8 = \sqrt{117{,}2}$
  $s = 11{,}1 = \sqrt{123{,}0}$
- dostaneme tedy

$$T = \frac{34{,}6 - 31{,}2}{11{,}1}\sqrt{\frac{104 \cdot 114}{104 + 114}} = 2{,}25$$

- na 5% hladině jsme prokázali rozdíl mezi dvěma skupinami uchazečů ($p = 2{,}6$ %)
- 95% interval spolehlivosti pro rozdíl populačních průměrů je $(0{,}4; 6{,}3)$