Statistika

(MD360P03Z, MD360P03U)
ak. rok 2007/2008

Karel Zvára

karel.zvara@mff.cuni.cz
http://www.karlin.mff.cuni.cz/∼zvara

(naposledy upraveno 20. listopadu 2007)

změnila se za deset roků výška desetiletých hochů?

- v roce 1951 byla průměrná výška desetiletých hochů 136,1 cm (zjištěno z velkého výběru o tisících měření)
- v roce 1961 bylo změřeno 15 náhodně vybraných desetiletých hochů: 127 130 133 136 136 138 139 139 139 140 141 142 147 149 151
- $\bar{X} = 139{,}13$ cm, $n = 15$
- znamená to, že za těch deset roků jsou desetiletí opravdu vyšší?
- stačí k důkazu, že 10 hochů je větších než 136,1 cm a jen 5 menších než 36,1 cm?
- stačí k důkazu, že nový průměr je o 3 cm vyšší?

změnila se za deset roků výška desetiletých hochů?

- v roce 1951 byla průměrná výška desetiletých hochů 136,1 cm (zjištěno z velkého výběru o tisících měření)
- v roce 1961 bylo změřeno 15 náhodně vybraných desetiletých hochů: 127 130 133 136 136 138 139 139 139 140 141 142 147 149 151
- $\bar{X} = 139{,}13$ cm, $n = 15$
- znamená to, že za těch deset roků jsou desetiletí opravdu vyšší?
- stačí k důkazu, že 10 hochů je větších než 136,1 cm a jen 5 menších než 36,1 cm?
- stačí k důkazu, že nový průměr je o 3 cm vyšší?

změnila se za deset roků výška desetiletých hochů?

- v roce 1951 byla průměrná výška desetiletých hochů 136,1 cm (zjištěno z velkého výběru o tisících měření)
- v roce 1961 bylo změřeno 15 náhodně vybraných desetiletých hochů: 127 130 133 136 136 138 139 139 139 140 141 142 147 149 151
- $\bar{X} = 139{,}13$ cm, $n = 15$
- znamená to, že za těch deset roků jsou desetiletí opravdu vyšší?
- stačí k důkazu, že 10 hochů je větších než 136,1 cm a jen 5 menších než 36,1 cm?
- stačí k důkazu, že nový průměr je o 3 cm vyšší?

změnila se za deset roků výška desetiletých hochů?

- v roce 1951 byla průměrná výška desetiletých hochů 136,1 cm (zjištěno z velkého výběru o tisících měření)
- v roce 1961 bylo změřeno 15 náhodně vybraných desetiletých hochů: 127 130 133 136 136 138 139 139 139 140 141 142 147 149 151
- $\bar{X} = 139{,}13$ cm, $n = 15$
- znamená to, že za těch deset roků jsou desetiletí opravdu vyšší?
- stačí k důkazu, že 10 hochů je větších než 136,1 cm a jen 5 menších než 36,1 cm?
- stačí k důkazu, že nový průměr je o 3 cm vyšší?

změnila se za deset roků výška desetiletých hochů?

- v roce 1951 byla průměrná výška desetiletých hochů 136,1 cm (zjištěno z velkého výběru o tisících měření)
- v roce 1961 bylo změřeno 15 náhodně vybraných desetiletých hochů: 127 130 133 136 136 138 139 139 139 140 141 142 147 149 151
- $\bar{X} = 139{,}13$ cm, $n = 15$
- znamená to, že za těch deset roků jsou desetiletí opravdu vyšší?
- stačí k důkazu, že 10 hochů je větších než 136,1 cm a jen 5 menších než 36,1 cm?
- stačí k důkazu, že nový průměr je o 3 cm vyšší?

změnila se za deset roků výška desetiletých hochů?

- v roce 1951 byla průměrná výška desetiletých hochů 136,1 cm (zjištěno z velkého výběru o tisících měření)
- v roce 1961 bylo změřeno 15 náhodně vybraných desetiletých hochů: 127 130 133 136 136 138 139 139 139 140 141 142 147 149 151
- $\bar{X} = 139{,}13$ cm, $n = 15$
- znamená to, že za těch deset roků jsou desetiletí opravdu vyšší?
- stačí k důkazu, že 10 hochů je větších než 136,1 cm a jen 5 menších než 36,1 cm?
- stačí k důkazu, že nový průměr je o 3 cm vyšší?

test o střední hodnotě μ normálního rozdělení

- předpokládáme $X_1, X_2, \ldots, X_n \sim \mathrm{N}\left(\mu, \sigma^2\right)$, nezávislé
- $\sigma > 0$ odhadneme pomocí $s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2}$
- rozptyl $\bar{X}$ odhadneme pomocí s_x^2/n, střední chyba $\bar{X}$ (odmocnina z rozptylu) je tedy $\mathrm{S.E.}(\bar{X}) = s_x/\sqrt{n}$
- $\mathrm{H}_0 : \mu = \mu_0$ (μ_0 známá konstanta)

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

statistka T má za H_0 Studentovo t-rozdělení s $n-1$ st. vol.

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

test o střední hodnotě μ normálního rozdělení

- předpokládáme $X_1, X_2, \ldots, X_n \sim \mathrm{N}(\mu, \sigma^2)$, nezávislé
- $\sigma > 0$ odhadneme pomocí $s_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$
- rozptyl $\bar{X}$ odhadneme pomocí s_x^2/n, střední chyba $\bar{X}$ (odmocnina z rozptylu) je tedy $\mathrm{S.E.}(\bar{X}) = s_x/\sqrt{n}$
- $\mathrm{H}_0 : \mu = \mu_0$ (μ_0 známá konstanta)

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

statistka T má za H_0 Studentovo t-rozdělení s $n-1$ st. vol.

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

test o střední hodnotě μ normálního rozdělení

- předpokládáme $X_1, X_2, \ldots, X_n \sim \mathrm{N}\big(\mu, \sigma^2\big)$, nezávislé
- $\sigma > 0$ odhadneme pomocí $s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2}$
- rozptyl $\bar{X}$ odhadneme pomocí s_x^2/n, střední chyba $\bar{X}$ (odmocnina z rozptylu) je tedy $\mathrm{S.E.}(\bar{X}) = s_x/\sqrt{n}$
- $\mathrm{H}_0 : \mu = \mu_0$ (μ_0 známá konstanta)

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

  statistka T má za H_0 Studentovo t-rozdělení s $n-1$ st. vol.
- kdy hypotézu H_0 zamítáme (kritický obor):
  - $\mathrm{H}_1 : \mu \neq \mu_0$ (oboustranná alternativa) $|T| \geq t_{n-1}(\alpha)$
  - $\mathrm{H}_1 : \mu > \mu_0$ (jednostranná alternativa) $T \geq t_{n-1}(2\alpha)$
  - $\mathrm{H}_1 : \mu < \mu_0$ (jednostranná alternativa) $T \leq -t_{n-1}(2\alpha)$

test o střední hodnotě μ normálního rozdělení

- předpokládáme $X_1, X_2, \ldots, X_n \sim \mathrm{N}(\mu, \sigma^2)$, nezávislé
- $\sigma > 0$ odhadneme pomocí $s_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$
- rozptyl $\bar{X}$ odhadneme pomocí s_x^2/n, střední chyba $\bar{X}$ (odmocnina z rozptylu) je tedy $\mathrm{S.E.}(\bar{X}) = s_x/\sqrt{n}$
- $\mathrm{H}_0 : \mu = \mu_0$ (μ_0 známá konstanta)

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

statistka T má za H_0 Studentovo t-rozdělení s $n-1$ st. vol.

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

test o střední hodnotě μ normálního rozdělení

- předpokládáme $X_1, X_2, \dots, X_n \sim \mathrm{N}\left(\mu, \sigma^2\right)$, nezávislé
- $\sigma > 0$ odhadneme pomocí $s_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$
- rozptyl $\bar{X}$ odhadneme pomocí s_x^2/n, střední chyba $\bar{X}$ (odmocnina z rozptylu) je tedy $\mathrm{S.E.}(\bar{X}) = s_x/\sqrt{n}$
- $\mathrm{H}_0 : \mu = \mu_0$ (μ_0 známá konstanta)

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

statistka T má za H_0 Studentovo t-rozdělení s $n-1$ st. vol.

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

test o střední hodnotě μ normálního rozdělení

- předpokládáme $X_1, X_2, \ldots, X_n \sim \mathrm{N}\left(\mu, \sigma^2\right)$, nezávislé
- $\sigma > 0$ odhadneme pomocí $s_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$
- rozptyl $\bar{X}$ odhadneme pomocí s_x^2/n, střední chyba $\bar{X}$ (odmocnina z rozptylu) je tedy $\mathrm{S.E.}(\bar{X}) = s_x/\sqrt{n}$
- $\mathrm{H}_0 : \mu = \mu_0$ (μ_0 známá konstanta)

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

statistka T má za H_0 Studentovo t-rozdělení s $n-1$ st. vol.

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

test o střední hodnotě μ normálního rozdělení

- předpokládáme $X_1, X_2, \ldots, X_n \sim \mathrm{N}\big(\mu, \sigma^2\big)$, nezávislé
- $\sigma > 0$ odhadneme pomocí $s_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$
- rozptyl $\bar{X}$ odhadneme pomocí s_x^2/n, střední chyba $\bar{X}$ (odmocnina z rozptylu) je tedy $\mathrm{S.E.}(\bar{X}) = s_x/\sqrt{n}$
- $\mathrm{H}_0 : \mu = \mu_0$ (μ_0 známá konstanta)

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

  statistka T má za H_0 Studentovo t-rozdělení s $n-1$ st. vol.

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

test o střední hodnotě μ normálního rozdělení

- předpokládáme $X_1, X_2, \ldots, X_n \sim \mathrm{N}(\mu, \sigma^2)$, nezávislé
- $\sigma > 0$ odhadneme pomocí $s_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$
- rozptyl $\bar{X}$ odhadneme pomocí s_x^2/n, střední chyba $\bar{X}$ (odmocnina z rozptylu) je tedy $\mathrm{S.E.}(\bar{X}) = s_x/\sqrt{n}$
- $\mathrm{H}_0 : \mu = \mu_0$ (μ_0 známá konstanta)

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

  statistka T má za H_0 Studentovo t-rozdělení s $n-1$ st. vol.
- kdy hypotézu H_0 zamítáme (kritický obor):
  - $\mathrm{H}_1 : \mu \neq \mu_0$ (oboustranná alternativa) $|T| \geq t_{n-1}(\alpha)$
  - $\mathrm{H}_1 : \mu > \mu_0$ (jednostranná alternativa) $T \geq t_{n-1}(2\alpha)$
  - $\mathrm{H}_1 : \mu < \mu_0$ (jednostranná alternativa) $T \leq -t_{n-1}(2\alpha)$

souvislost s intervalem spolehlivosti

- připomeňme interval spolehlivosti pro μ

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

$$\bar{X} - \frac{s_x}{\sqrt{n}} t_{n-1}(\alpha) < \mu < \bar{X} + \frac{s_x}{\sqrt{n}} t_{n-1}(\alpha)$$

- 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 obsahuje takové hodnoty μ_0, které bychom jako hypotézu nezamítli**

souvislost s intervalem spolehlivosti

- připomeňme interval spolehlivosti pro μ

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

- 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 obsahuje takové hodnoty μ_0, které bychom jako hypotézu nezamítli**

souvislost s intervalem spolehlivosti

- připomeňme interval spolehlivosti pro μ

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

$$\bar{X} - \frac{s_x}{\sqrt{n}} t_{n-1}(\alpha) < \mu < \bar{X} + \frac{s_x}{\sqrt{n}} t_{n-1}(\alpha)$$

- 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 obsahuje takové hodnoty μ_0, které bychom jako hypotézu nezamítli**

souvislost s intervalem spolehlivosti

- připomeňme interval spolehlivosti pro μ

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

- lze přepsat jako

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

- $\mathrm{H}_0 : \mu = \mu_0$ tedy **nezamítneme** na hladině α při oboustranné alternativě, právě když μ_0 leží v $100(1-\alpha)\%$ intervalu spolehlivosti
- **interval spolehlivosti obsahuje takové hodnoty μ_0, které bychom jako hypotézu nezamítli**

příklad: výšky desetiletých hochů (σ^2 neznámé)

- kritický obor: $\bar{X}$ se příliš liší od μ_0 ve směru zvolené alternativy
- spočítáme [t.test(hosi,mu=136.1,alternative="greater")]

$$T = \frac{139{,}13 - 136{,}1}{6{,}56}\sqrt{15} = 1{,}79$$

- na 5% hladině při jednostranné alternativě $\mu > \mu_0$ hypotézu zamítáme, neboť $t_{14}(0{,}10) = 1{,}76$ ($p = 4{,}7$ %)
- na 5% hladině jsme **prokázali**, že výška desetiletých vzrostla
- 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: $(135{,}5; 142{,}8)$

příklad: výšky desetiletých hochů (σ^2 neznámé)

- kritický obor: $\bar{X}$ se příliš liší od μ_0 ve směru zvolené alternativy
- spočítáme [t.test(hosi,mu=136.1,alternative="greater")]

$$T = \frac{139{,}13 - 136{,}1}{6{,}56}\sqrt{15} = 1{,}79$$

- na 5% hladině při jednostranné alternativě $\mu > \mu_0$ hypotézu zamítáme, neboť $t_{14}(0{,}10) = 1{,}76$ ($p = 4{,}7$ %)
- na 5% hladině jsme **prokázali**, že výška desetiletých vzrostla
- 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: $(135{,}5; 142{,}8)$

příklad: výšky desetiletých hochů (σ^2 neznámé)

- kritický obor: $\bar{X}$ se příliš liší od μ_0 ve směru zvolené alternativy
- spočítáme [t.test(hosi,mu=136.1,alternative="greater")]

$$T = \frac{139{,}13 - 136{,}1}{6{,}56}\sqrt{15} = 1{,}79$$

- na 5% hladině při jednostranné alternativě $\mu > \mu_0$ hypotézu zamítáme, neboť $t_{14}(0{,}10) = 1{,}76$ ($p = 4{,}7$ %)
- na 5% hladině jsme **prokázali**, že výška desetiletých vzrostla
- 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: $(135{,}5; 142{,}8)$

příklad: výšky desetiletých hochů (σ^2 neznámé)

- kritický obor: $\bar{X}$ se příliš liší od μ_0 ve směru zvolené alternativy
- spočítáme [t.test(hosi,mu=136.1,alternative="greater")]

$$T = \frac{139{,}13 - 136{,}1}{6{,}56}\sqrt{15} = 1{,}79$$

- na 5% hladině při jednostranné alternativě $\mu > \mu_0$ hypotézu zamítáme, neboť $t_{14}(0{,}10) = 1{,}76$ ($p = 4{,}7$ %)
- na 5% hladině jsme **prokázali**, že výška desetiletých vzrostla
- 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: $(135{,}5; 142{,}8)$

příklad: výšky desetiletých hochů (σ^2 neznámé)

- kritický obor: $\bar{X}$ se příliš liší od μ_0 ve směru zvolené alternativy
- spočítáme [t.test(hosi,mu=136.1,alternative="greater")]

$$T = \frac{139{,}13 - 136{,}1}{6{,}56}\sqrt{15} = 1{,}79$$

- na 5% hladině při jednostranné alternativě $\mu > \mu_0$ hypotézu zamítáme, neboť $t_{14}(0{,}10) = 1{,}76$ ($p = 4{,}7$ %)
- na 5% hladině jsme **prokázali**, že výška desetiletých vzrostla
- 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: $(135{,}5; 142{,}8)$

příklad: výšky desetiletých hochů (σ^2 neznámé)

- kritický obor: $\bar{X}$ se příliš liší od μ_0 ve směru zvolené alternativy
- spočítáme [t.test(hosi,mu=136.1,alternative="greater")]

$$T = \frac{139{,}13 - 136{,}1}{6{,}56}\sqrt{15} = 1{,}79$$

- na 5% hladině při jednostranné alternativě $\mu > \mu_0$ hypotézu zamítáme, neboť $t_{14}(0{,}10) = 1{,}76$ ($p = 4{,}7$ %)
- na 5% hladině jsme **prokázali**, že výška desetiletých vzrostla
- 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: $(135{,}5; 142{,}8)$

použití Excelu

přednáška	Excel	hoši
průměr	Stř. hodnota	139,13
střední chyba	Chyba stř. hodnoty	1,693
medián	Medián	139
modus	Modus	139
s	Směr. odchylka	6,56
s^2	Rozptyl výběru	42,98
špičatost	Špičatost	0,006
šikmost	Šikmost	0,090
rozpětí	Rozdíl max-min	24
minimum	Minimum	127
maximum	Maximum	151
součet	Součet	2087
rozsah výběru n	Počet	15
pol. šířka int. spol.	Hladina spol.	3,63

- 139,13 – 3,63 = 135,50
- 139,13 + 3,63 = 142,76
- 95% interval spolehlivosti: (135,5; 142,8)
- $\mu_0 = 136{,}1$ je v int. spolehlivosti
- při oboustranné alternativě jsme nezamítli H_0

použití Excelu

přednáška	Excel	hoši
průměr	Stř. hodnota	139,13
střední chyba	Chyba stř. hodnoty	1,693
medián	Medián	139
modus	Modus	139
s	Směr. odchylka	6,56
s^2	Rozptyl výběru	42,98
špičatost	Špičatost	0,006
šikmost	Šikmost	0,090
rozpětí	Rozdíl max-min	24
minimum	Minimum	127
maximum	Maximum	151
součet	Součet	2087
rozsah výběru n	Počet	15
pol. šířka int. spol.	Hladina spol.	3,63

- 139,13 – 3,63 $= 135{,}50$
- 139,13 $+$ 3,63 $= 142{,}76$
- 95% interval spolehlivosti: (135,5; 142,8)
- $\mu_0 = 136{,}1$ je v int. spolehlivosti
- při oboustranné alternativě jsme nezamítli H_0

použití Excelu

přednáška	Excel	hoši
průměr	Stř. hodnota	139,13
střední chyba	Chyba stř. hodnoty	1,693
medián	Medián	139
modus	Modus	139
s	Směr. odchylka	6,56
s^2	Rozptyl výběru	42,98
špičatost	Špičatost	0,006
šikmost	Šikmost	0,090
rozpětí	Rozdíl max-min	24
minimum	Minimum	127
maximum	Maximum	151
součet	Součet	2087
rozsah výběru n	Počet	15
pol. šířka int. spol.	Hladina spol.	3,63

- 139,13 – 3,63 = 135,50
- 139,13 + 3,63 = 142,76
- 95% interval spolehlivosti: (135,5; 142,8)
- $\mu_0 = 136{,}1$ je v int. spolehlivosti
- při oboustranné alternativě jsme nezamítli H_0

použití Excelu

přednáška	Excel	hoši
průměr	Stř. hodnota	139,13
střední chyba	Chyba stř. hodnoty	1,693
medián	Medián	139
modus	Modus	139
s	Směr. odchylka	6,56
s^2	Rozptyl výběru	42,98
špičatost	Špičatost	0,006
šikmost	Šikmost	0,090
rozpětí	Rozdíl max-min	24
minimum	Minimum	127
maximum	Maximum	151
součet	Součet	2087
rozsah výběru n	Počet	15
pol. šířka int. spol.	Hladina spol.	3,63

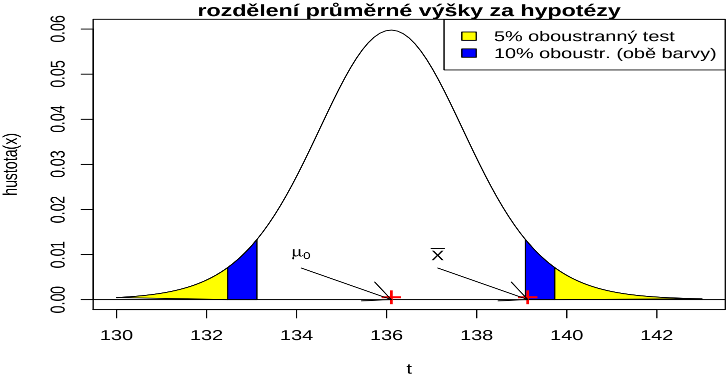
- 139,13 – 3,63 $=$ 135,50
- 139,13 $+$ 3,63 $=$ 142,76
- 95% interval spolehlivosti: (135,5; 142,8)
- $\mu_0 = 136{,}1$ je v int. spolehlivosti
- při oboustranné alternativě jsme nezamítli H_0

použití Excelu

přednáška	Excel	hoši
průměr	Stř. hodnota	139,13
střední chyba	Chyba stř. hodnoty	1,693
medián	Medián	139
modus	Modus	139
s	Směr. odchylka	6,56
s^2	Rozptyl výběru	42,98
špičatost	Špičatost	0,006
šikmost	Šikmost	0,090
rozpětí	Rozdíl max-min	24
minimum	Minimum	127
maximum	Maximum	151
součet	Součet	2087
rozsah výběru n	Počet	15
pol. šířka int. spol.	Hladina spol.	3,63

- 139,13 – 3,63 $= 135{,}50$
- $139{,}13 + 3{,}63 = 142{,}76$
- 95% interval spolehlivosti: (135,5; 142,8)
- $\mu_0 = 136{,}1$ je v int. spolehlivosti
- při oboustranné alternativě jsme nezamítli H_0

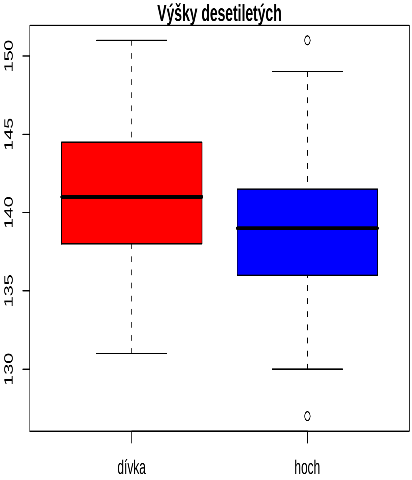
kritický obor pro $\bar{X}$



- při jednostr. alternativě $\mu > \mu_0$ je 5% kritický obor označen oběma barvami na pravé straně

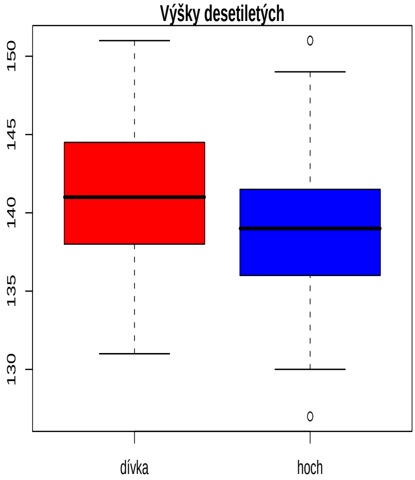
porovnání dvou populací (dvouvýběrový t-test)

- příklad: liší se desetileté dívky výškou postavy od desetiletých hochů?
- výšky hochů známe, $\bar{X} = 139{,}13$ cm, $s_x = 6{,}56$, $n_x = 15$
- výšky dívek: 131, 132, 135, 141, 141, 141, 141, 142, 143, 146, 146, 151
- $\bar{Y} = 140{,}83$, $s_y = 5{,}84$, $n_y = 12$



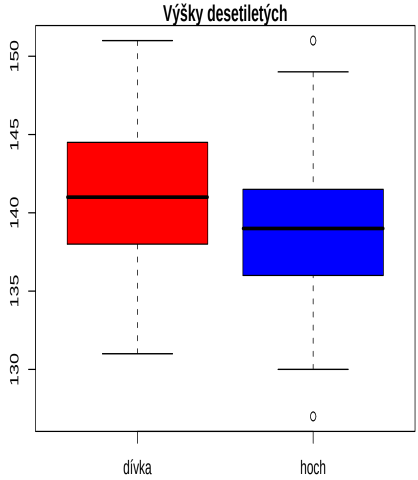
porovnání dvou populací (dvouvýběrový t-test)

- příklad: liší se desetileté dívky výškou postavy od desetiletých hochů?
- výšky hochů známe, $\bar{X} = 139{,}13$ cm, $s_x = 6{,}56$, $n_x = 15$
- výšky dívek: 131, 132, 135, 141, 141, 141, 141, 142, 143, 146, 146, 151
- $\bar{Y} = 140{,}83$, $s_y = 5{,}84$, $n_y = 12$



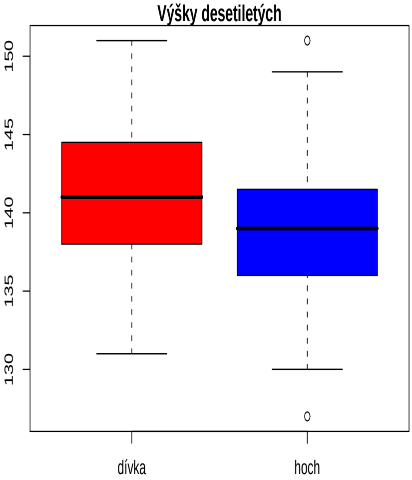
porovnání dvou populací (dvouvýběrový *t*-test)

- příklad: liší se desetileté dívky výškou postavy od desetiletých hochů?
- výšky hochů známe, $\bar{X} = 139{,}13$ cm, $s_x = 6{,}56$, $n_x = 15$
- výšky dívek: 131, 132, 135, 141, 141, 141, 141, 142, 143, 146, 146, 151
- $\bar{Y} = 140{,}83$, $s_y = 5{,}84$, $n_y = 12$



porovnání dvou populací (dvouvýběrový t-test)

- příklad: liší se desetileté dívky výškou postavy od desetiletých hochů?
- výšky hochů známe, $\bar{X} = 139{,}13$ cm, $s_x = 6{,}56$, $n_x = 15$
- výšky dívek: 131, 132, 135, 141, 141, 141, 141, 142, 143, 146, 146, 151
- $\bar{Y} = 140{,}83$, $s_y = 5{,}84$, $n_y = 12$



dvouvýběrový t.test

- lze předpokládat, že výšky náhodně vybraných hochů mají normální rozdělení

$$X_i \sim \mathrm{N}\left(\mu_x, \sigma^2\right), \qquad \text{nezávislé}, \quad i = 1, \ldots, n_x$$

- lze předpokládat, že výšky náhodně vybraných dívek mají normální rozdělení

$$Y_i \sim \mathrm{N}\left(\mu_y, \sigma^2\right), \qquad \text{nezávislé}, \quad i = 1, \ldots, n_y$$

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

dvouvýběrový t.test

- lze předpokládat, že výšky náhodně vybraných hochů mají normální rozdělení

$$X_i \sim \mathrm{N}\big(\mu_x, \sigma^2\big), \qquad \text{nezávislé}, \quad i = 1, \ldots, n_x$$

- lze předpokládat, že výšky náhodně vybraných dívek mají normální rozdělení

$$Y_i \sim \mathrm{N}\big(\mu_y, \sigma^2\big), \qquad \text{nezávislé}, \quad i = 1, \ldots, n_y$$

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

dvouvýběrový *t*.test

- lze předpokládat, že výšky náhodně vybraných hochů mají normální rozdělení

$$X_i \sim \mathrm{N}\left(\mu_x, \sigma^2\right), \qquad \text{nezávislé}, \quad i = 1, \ldots, n_x$$

- lze předpokládat, že výšky náhodně vybraných dívek mají normální rozdělení

$$Y_i \sim \mathrm{N}\left(\mu_y, \sigma^2\right), \qquad \text{nezávislé}, \quad i = 1, \ldots, n_y$$

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

dvouvýběrový t.test

- lze předpokládat, že výšky náhodně vybraných hochů mají normální rozdělení

$$X_i \sim \mathrm{N}\big(\mu_x, \sigma^2\big), \qquad \text{nezávislé}, \quad i = 1, \ldots, n_x$$

- lze předpokládat, že výšky náhodně vybraných dívek mají normální rozdělení

$$Y_i \sim \mathrm{N}\big(\mu_y, \sigma^2\big), \qquad \text{nezávislé}, \quad i = 1, \ldots, n_y$$

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

porovnání středních hodnot nezávislých výběrů

- $H_0 : \mu_x = \mu_y$ (není rozdíl, **nulová** hypotéza)
  zřejmě totéž jako $\mu_x - \mu_y = 0$ (nulový rozdíl stř. hodnot)
  (hoši a dívky se v deseti letech co do výšky neliší)
- možné alternativy
  - $H_1 : \mu_x \neq \mu_y$ (není-li důvod k jednostranné alternativě)
  - $H_1 : \mu_x > \mu_y$ (bylo cílem dokázat, že hoši jsou větší než dívky)
  - $H_1 : \mu_x < \mu_y$ (bylo cílem dokázat, že hoši jsou menší než dívky)
- rozhodování založeno na porovnání průměrů $\bar{X}$ a $\bar{Y}$; čím více se liší „správným směrem“, tím spíše zamítnout hypotézu
- je třeba porovnat s mírou přesnosti, s jakou rozdíl průměrů $\bar{X} - \bar{Y}$ odhadne skutečný rozdíl populačních průměrů $\mu_x - \mu_y$

porovnání středních hodnot nezávislých výběrů

- $\mathrm{H}_0 : \mu_x = \mu_y$ (není rozdíl, **nulová** hypotéza) zřejmě totéž jako $\mu_x - \mu_y = 0$ (nulový rozdíl stř. hodnot) (hoši a dívky se v deseti letech co do výšky neliší)
- možné alternativy
  - $\mathrm{H}_1 : \mu_x \neq \mu_y$ (není-li důvod k jednostranné alternativě)
  - $\mathrm{H}_1 : \mu_x > \mu_y$ (bylo cílem dokázat, že hoši jsou větší než dívky)
  - $\mathrm{H}_1 : \mu_x < \mu_y$ (bylo cílem dokázat, že hoši jsou menší než dívky)
- rozhodování založeno na porovnání průměrů $\bar{X}$ a $\bar{Y}$; čím více se liší „správným směrem“, tím spíše zamítnout hypotézu
- je třeba porovnat s mírou přesnosti, s jakou rozdíl průměrů $\bar{X} - \bar{Y}$ odhadne skutečný rozdíl populačních průměrů $\mu_x - \mu_y$

porovnání středních hodnot nezávislých výběrů

- $H_0 : \mu_x = \mu_y$ (není rozdíl, **nulová** hypotéza)
  zřejmě totéž jako $\mu_x - \mu_y = 0$ (nulový rozdíl stř. hodnot)
  (hoši a dívky se v deseti letech co do výšky neliší)
- možné alternativy
  - $H_1 : \mu_x \neq \mu_y$ (není-li důvod k jednostranné alternativě)
  - $H_1 : \mu_x > \mu_y$ (bylo cílem dokázat, že hoši jsou větší než dívky)
  - $H_1 : \mu_x < \mu_y$ (bylo cílem dokázat, že hoši jsou menší než dívky)
- rozhodování založeno na porovnání průměrů $\bar{X}$ a $\bar{Y}$; čím více se liší „správným směrem“, tím spíše zamítnout hypotézu
- je třeba porovnat s mírou přesnosti, s jakou rozdíl průměrů $\bar{X} - \bar{Y}$ odhadne skutečný rozdíl populačních průměrů $\mu_x - \mu_y$

porovnání středních hodnot nezávislých výběrů

- $H_0 : \mu_x = \mu_y$ (není rozdíl, **nulová** hypotéza)
  zřejmě totéž jako $\mu_x - \mu_y = 0$ (nulový rozdíl stř. hodnot)
  (hoši a dívky se v deseti letech co do výšky neliší)
- možné alternativy
  - $H_1 : \mu_x \neq \mu_y$ (není-li důvod k jednostranné alternativě)
  - $H_1 : \mu_x > \mu_y$ (bylo cílem dokázat, že hoši jsou větší než dívky)
  - $H_1 : \mu_x < \mu_y$ (bylo cílem dokázat, že hoši jsou menší než dívky)
- rozhodování založeno na porovnání průměrů $\bar{X}$ a $\bar{Y}$; čím více se liší „správným směrem“, tím spíše zamítnout hypotézu
- je třeba porovnat s mírou přesnosti, s jakou rozdíl průměrů $\bar{X} - \bar{Y}$ odhadne skutečný rozdíl populačních průměrů $\mu_x - \mu_y$

porovnání středních hodnot nezávislých výběrů

- $H_0 : \mu_x = \mu_y$ (není rozdíl, **nulová** hypotéza)
  zřejmě totéž jako $\mu_x - \mu_y = 0$ (nulový rozdíl stř. hodnot)
  (hoši a dívky se v deseti letech co do výšky neliší)
- možné alternativy
  - $H_1 : \mu_x \neq \mu_y$ (není-li důvod k jednostranné alternativě)
  - $H_1 : \mu_x > \mu_y$ (bylo cílem dokázat, že hoši jsou větší než dívky)
  - $H_1 : \mu_x < \mu_y$ (bylo cílem dokázat, že hoši jsou menší než dívky)
- rozhodování založeno na porovnání průměrů $\bar{X}$ a $\bar{Y}$; čím více se liší „správným směrem", tím spíše zamítnout hypotézu
- je třeba porovnat s mírou přesnosti, s jakou rozdíl průměrů $\bar{X} - \bar{Y}$ odhadne skutečný rozdíl populačních průměrů $\mu_x - \mu_y$

porovnání středních hodnot nezávislých výběrů

- $H_0 : \mu_x = \mu_y$ (není rozdíl, **nulová** hypotéza)
  zřejmě totéž jako $\mu_x - \mu_y = 0$ (nulový rozdíl stř. hodnot)
  (hoši a dívky se v deseti letech co do výšky neliší)
- možné alternativy
  - $H_1 : \mu_x \neq \mu_y$ (není-li důvod k jednostranné alternativě)
  - $H_1 : \mu_x > \mu_y$ (bylo cílem dokázat, že hoši jsou větší než dívky)
  - $H_1 : \mu_x < \mu_y$ (bylo cílem dokázat, že hoši jsou menší než dívky)
- rozhodování založeno na porovnání průměrů $\bar{X}$ a $\bar{Y}$; čím více se liší „správným směrem“, tím spíše zamítnout hypotézu
- je třeba porovnat s mírou přesnosti, s jakou rozdíl průměrů $\bar{X} - \bar{Y}$ odhadne skutečný rozdíl populačních průměrů $\mu_x - \mu_y$

porovnání středních hodnot nezávislých výběrů

- $H_0 : \mu_x = \mu_y$ (není rozdíl, **nulová** hypotéza)
  zřejmě totéž jako $\mu_x - \mu_y = 0$ (nulový rozdíl stř. hodnot)
  (hoši a dívky se v deseti letech co do výšky neliší)
- možné alternativy
  - $H_1 : \mu_x \neq \mu_y$ (není-li důvod k jednostranné alternativě)
  - $H_1 : \mu_x > \mu_y$ (bylo cílem dokázat, že hoši jsou větší než dívky)
  - $H_1 : \mu_x < \mu_y$ (bylo cílem dokázat, že hoši jsou menší než dívky)
- rozhodování založeno na porovnání průměrů $\bar{X}$ a $\bar{Y}$; čím více se liší „správným směrem“, tím spíše zamítnout hypotézu
- je třeba porovnat s mírou přesnosti, s jakou rozdíl průměrů $\bar{X} - \bar{Y}$ odhadne skutečný rozdíl populačních průměrů $\mu_x - \mu_y$

odhad σ^2

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

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

(vážený průměr odhadů rozptylu v obou výběrech)

- výška desetiletých dětí: $n_x = 15$, $n_y = 12$, $\bar{X} = 139{,}13$, $\bar{Y} = 140{,}83$, $s_x^2 = 42{,}98$, $s_y^2 = 33{,}79$, tudíž

$$s^2 = \frac{14}{25} \cdot 42{,}98 + \frac{11}{25} \cdot 33{,}79 = 38{,}94 = 6{,}24^2$$

odhad σ^2

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

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

  (vážený průměr odhadů rozptylu v obou výběrech)
- výška desetiletých dětí: $n_x = 15$, $n_y = 12$, $\bar{X} = 139{,}13$, $\bar{Y} = 140{,}83$, $s_x^2 = 42{,}98$, $s_y^2 = 33{,}79$, tudíž

$$s^2 = \frac{14}{25} \cdot 42{,}98 + \frac{11}{25} \cdot 33{,}79 = 38{,}94 = 6{,}24^2$$

kritický obor

- o hypotéze $H_0 : \mu_1 = \mu_2$ se rozhoduje pomocí

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

- $H_1 : \mu_x \neq \mu_y$ zamítáme pokud $|T| \geq t_{n_1+n_2-2}(\alpha)$
- $H_1 : \mu_x > \mu_y$ zamítáme pokud $T \geq t_{n_1+n_2-2}(2\alpha)$
- $H_1 : \mu_x < \mu_y$ zamítáme pokud $T \leq -t_{n_1+n_2-2}(2\alpha)$
- výšky desetiletých: $T = -0{,}70 \Rightarrow$
  $$|-0{,}70| < 2{,}06 = t_{15+12-2}(0{,}05)$$
- na 5% hladině jsme **neprokázali** rozdíl mezi výškami desetiletých hochů a dívek ($p = 48{,}8$ %)

[t.test(vyska~Divka,var.equal=TRUE)]

[TTEST(A14:A28;A2:A13;2;2)]

kritický obor

- o hypotéze $H_0 : \mu_1 = \mu_2$ se rozhoduje pomocí

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

- $H_1 : \mu_x \neq \mu_y$ zamítáme pokud $|T| \geq t_{n_1+n_2-2}(\alpha)$
- $H_1 : \mu_x > \mu_y$ zamítáme pokud $T \geq t_{n_1+n_2-2}(2\alpha)$
- $H_1 : \mu_x < \mu_y$ zamítáme pokud $T \leq -t_{n_1+n_2-2}(2\alpha)$
- výšky desetiletých: $T = -0{,}70 \Rightarrow$
  $$|-0{,}70| < 2{,}06 = t_{15+12-2}(0{,}05)$$
- na 5% hladině jsme **neprokázali** rozdíl mezi výškami desetiletých hochů a dívek ($p = 48{,}8$ %)

[t.test(vyska∼Divka,var.equal=TRUE)]

[TTEST(A14:A28;A2:A13;2;2)]

kritický obor

- o hypotéze $H_0 : \mu_1 = \mu_2$ se rozhoduje pomocí

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

- $H_1 : \mu_x \neq \mu_y$ zamítáme pokud $|T| \geq t_{n_1+n_2-2}(\alpha)$
- $H_1 : \mu_x > \mu_y$ zamítáme pokud $T \geq t_{n_1+n_2-2}(2\alpha)$
- $H_1 : \mu_x < \mu_y$ zamítáme pokud $T \leq -t_{n_1+n_2-2}(2\alpha)$
- výšky desetiletých: $T = -0{,}70 \Rightarrow$
$$|-0{,}70| < 2{,}06 = t_{15+12-2}(0{,}05)$$
- na 5% hladině jsme **neprokázali** rozdíl mezi výškami desetiletých hochů a dívek ($p = 48{,}8$ %)

[t.test(vyska∼Divka,var.equal=TRUE)]
[TTEST(A14:A28;A2:A13;2;2)]

kritický obor

- o hypotéze $H_0 : \mu_1 = \mu_2$ se rozhoduje pomocí

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

- $H_1 : \mu_x \neq \mu_y$ zamítáme pokud $|T| \geq t_{n_1+n_2-2}(\alpha)$
- $H_1 : \mu_x > \mu_y$ zamítáme pokud $T \geq t_{n_1+n_2-2}(2\alpha)$
- $H_1 : \mu_x < \mu_y$ zamítáme pokud $T \leq -t_{n_1+n_2-2}(2\alpha)$
- výšky desetiletých: $T = -0{,}70 \Rightarrow$
  $$|-0{,}70| < 2{,}06 = t_{15+12-2}(0{,}05)$$
- na 5% hladině jsme **neprokázali** rozdíl mezi výškami desetiletých hochů a dívek ($p = 48{,}8$ %)

[t.test(vyska~Divka,var.equal=TRUE)]

[TTEST(A14:A28;A2:A13;2;2)]

kritický obor

- o hypotéze $H_0 : \mu_1 = \mu_2$ se rozhoduje pomocí

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

- $H_1 : \mu_x \neq \mu_y$ zamítáme pokud $|T| \geq t_{n_1+n_2-2}(\alpha)$
- $H_1 : \mu_x > \mu_y$ zamítáme pokud $T \geq t_{n_1+n_2-2}(2\alpha)$
- $H_1 : \mu_x < \mu_y$ zamítáme pokud $T \leq -t_{n_1+n_2-2}(2\alpha)$
- výšky desetiletých: $T = -0{,}70 \Rightarrow$
  $$|-0{,}70| < 2{,}06 = t_{15+12-2}(0{,}05)$$
- na 5% hladině jsme **neprokázali** rozdíl mezi výškami desetiletých hochů a dívek ($p = 48{,}8$ %)

[t.test(vyska~Divka,var.equal=TRUE)]
[TTEST(A14:A28;A2:A13;2;2)]

kritický obor

- o hypotéze $H_0 : \mu_1 = \mu_2$ se rozhoduje pomocí

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

- $H_1 : \mu_x \neq \mu_y$ zamítáme pokud $|T| \geq t_{n_1+n_2-2}(\alpha)$
- $H_1 : \mu_x > \mu_y$ zamítáme pokud $T \geq t_{n_1+n_2-2}(2\alpha)$
- $H_1 : \mu_x < \mu_y$ zamítáme pokud $T \leq -t_{n_1+n_2-2}(2\alpha)$
- výšky desetiletých: $T = -0{,}70 \Rightarrow$
  $$|-0{,}70| < 2{,}06 = t_{15+12-2}(0{,}05)$$
- na 5% hladině jsme **neprokázali** rozdíl mezi výškami desetiletých hochů a dívek ($p = 48{,}8$ %)

[t.test(vyska~Divka,var.equal=TRUE)]
[TTEST(A14:A28;A2:A13;2;2)]

souvislost s intervalem spolehlivosti

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

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

H_0 zamítáme právě tehdy, když nula **není** v int. spol. pro δ

- při porovnání výšek hochů a dívek je 95% interval pro δ

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

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

souvislost s intervalem spolehlivosti

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

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

  H_0 zamítáme právě tehdy, když nula **není** v int. spol. pro δ
- při porovnání výšek hochů a dívek je 95% interval pro δ

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

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

souvislost s intervalem spolehlivosti

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

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

H_0 zamítáme právě tehdy, když nula **není** v int. spol. pro δ

- při porovnání výšek hochů a dívek je 95% interval pro δ

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

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

souvislost s intervalem spolehlivosti

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

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

H$_0$ zamítáme právě tehdy, když nula **není** v int. spol. pro δ

- při porovnání výšek hochů a dívek je 95% interval pro δ

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

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

shrnutí

- důležité předpoklady
  - nezávislé výběry
  - stejné (populační) rozptyly (lze testovat)
  - normální rozdělení (lze testovat)
- existuje varianta bez předpokladu stejných rozptylů
- pro velká n_x, n_y na normalitě tolik nezáleží (CLV)
- je-li problém s normalitou, lze použít jiný test (Mann-Whittney)

shrnutí

- důležité předpoklady
  - nezávislé výběry
  - stejné (populační) rozptyly (lze testovat)
  - normální rozdělení (lze testovat)
- existuje varianta bez předpokladu stejných rozptylů
- pro velká n_x, n_y na normalitě tolik nezáleží (CLV)
- je-li problém s normalitou, lze použít jiný test (Mann-Whittney)

shrnutí

- důležité předpoklady
  - nezávislé výběry
  - stejné (populační) rozptyly (lze testovat)
  - normální rozdělení (lze testovat)
- existuje varianta bez předpokladu stejných rozptylů
- pro velká n_x, n_y na normalitě tolik nezáleží (CLV)
- je-li problém s normalitou, lze použít jiný test (Mann-Whittney)

shrnutí

- důležité předpoklady
  - nezávislé výběry
  - stejné (populační) rozptyly (lze testovat)
  - normální rozdělení (lze testovat)
- existuje varianta bez předpokladu stejných rozptylů
- pro velká n_x, n_y na normalitě tolik nezáleží (CLV)
- je-li problém s normalitou, lze použít jiný test (Mann-Whittney)

shrnutí

- důležité předpoklady
  - nezávislé výběry
  - stejné (populační) rozptyly (lze testovat)
  - normální rozdělení (lze testovat)
- existuje varianta bez předpokladu stejných rozptylů
- pro velká n_x, n_y na normalitě tolik nezáleží (CLV)
- je-li problém s normalitou, lze použít jiný test (Mann-Whittney)

shrnutí

- důležité předpoklady
  - nezávislé výběry
  - stejné (populační) rozptyly (lze testovat)
  - normální rozdělení (lze testovat)
- existuje varianta bez předpokladu stejných rozptylů
- pro velká n_x, n_y na normalitě tolik nezáleží (CLV)
- je-li problém s normalitou, lze použít jiný test (Mann-Whittney)

shrnutí

- důležité předpoklady
  - nezávislé výběry
  - stejné (populační) rozptyly (lze testovat)
  - normální rozdělení (lze testovat)
- existuje varianta bez předpokladu stejných rozptylů
- pro velká n_x, n_y na normalitě tolik nezáleží (CLV)
- je-li problém s normalitou, lze použít jiný test (Mann-Whittney)

provedení v MS Excelu (stejné rozptyly)

přednáška	Excel	Soubor 1	Soubor 2
průměr	Stř. hodnota	139.133	140.833
rozptyl	Rozptyl	42.981	33.788
rozsah výběru	Pozorování	15	12
spol. odhad rozpt.	Společný rozptyl	38.936	
$\mathrm{H}_0 : \mu_1 - \mu_2 =$	Hyp. rozdíl stř. hodnot	0	
stupně vol.	Rozdíl	25	
T	t stat	-0.733	
p jednostr. testu	P(T<=t) (1)	0.244	jen někdy!
$t_{n_1+n_2-2}(2\alpha)$	t krit (1)	1.708	
p oboustr. testu	P(T<=t) (2)	0.488	
$t_{n_1+n_2-2}(\alpha)$	t krit (2)	2.060	

při oboustranné alternativě nelze nulovou hypotézu zamítnout

problém nestejných rozptylů

- předpoklad o stejném rozptylu v obou souborech nemusí být ve skutečnosti splněn, lze jej ověřit porovnáním odhadů rozptylu F-testem $F = \frac{s_x^2}{s_y^2}$
- hypotéza $H_0 : \sigma_x^2 = \sigma_y^2$ se proti $H_1 : \sigma_x^2 \neq \sigma_y^2$ zamítá, když je

  buď $F = \frac{s_x^2}{s_y^2} \geq F_{n_1-1,n_2-1}(\alpha/2)$ nebo $\frac{1}{F} = \frac{s_y^2}{s_x^2} \geq F_{n_2-1,n_1-1}(\alpha/2)$
- vlastně se větší odhad rozptylu dělí menším odhadem, k tomu se musí zvolit správné pořadí stupňů volnosti a hladina
- příklad výšky desetiletých dětí: $F = \frac{42{,}98}{38{,}94} = 1{,}27 < F_{14,11}(0{,}025) = 3{,}36$
- [var.test(vyska~Divka)]

problém nestejných rozptylů

- předpoklad o stejném rozptylu v obou souborech nemusí být ve skutečnosti splněn, lze jej ověřit porovnáním odhadů rozptylu F-testem $F = \frac{s_x^2}{s_y^2}$
- hypotéza $\mathrm{H}_0 : \sigma_x^2 = \sigma_y^2$ se proti $\mathrm{H}_1 : \sigma_x^2 \neq \sigma_y^2$ zamítá, když je

$$\text{buď } F = \frac{s_x^2}{s_y^2} \geq F_{n_1-1,n_2-1}(\alpha/2) \text{ nebo } \frac{1}{F} = \frac{s_y^2}{s_x^2} \geq F_{n_2-1,n_1-1}(\alpha/2)$$

- vlastně se větší odhad rozptylu dělí menším odhadem, k tomu se musí zvolit správné pořadí stupňů volnosti a hladina
- příklad výšky desetiletých dětí: $F = \frac{42{,}98}{38{,}94} = 1{,}27 < F_{14,11}(0{,}025) = 3{,}36$
- [var.test(vyska~Divka)]

problém nestejných rozptylů

- předpoklad o stejném rozptylu v obou souborech nemusí být ve skutečnosti splněn, lze jej ověřit porovnáním odhadů rozptylu F-testem $F = \frac{s_x^2}{s_y^2}$
- hypotéza $H_0 : \sigma_x^2 = \sigma_y^2$ se proti $H_1 : \sigma_x^2 \neq \sigma_y^2$ zamítá, když je

$$\text{buď } F = \frac{s_x^2}{s_y^2} \geq F_{n_1-1,n_2-1}(\alpha/2) \text{ nebo } \frac{1}{F} = \frac{s_y^2}{s_x^2} \geq F_{n_2-1,n_1-1}(\alpha/2)$$

- vlastně se větší odhad rozptylu dělí menším odhadem, k tomu se musí zvolit správné pořadí stupňů volnosti a hladina
- příklad výšky desetiletých dětí: $F = \frac{42{,}98}{38{,}94} = 1{,}27 < F_{14,11}(0{,}025) = 3{,}36$
- [var.test(vyska~Divka)]

problém nestejných rozptylů

- předpoklad o stejném rozptylu v obou souborech nemusí být ve skutečnosti splněn, lze jej ověřit porovnáním odhadů rozptylu F-testem $F = \frac{s_x^2}{s_y^2}$
- hypotéza $\mathrm{H}_0 : \sigma_x^2 = \sigma_y^2$ se proti $\mathrm{H}_1 : \sigma_x^2 \neq \sigma_y^2$ zamítá, když je

$$\text{buď } F = \frac{s_x^2}{s_y^2} \geq F_{n_1-1,n_2-1}(\alpha/2) \text{ nebo } \frac{1}{F} = \frac{s_y^2}{s_x^2} \geq F_{n_2-1,n_1-1}(\alpha/2)$$

- vlastně se větší odhad rozptylu dělí menším odhadem, k tomu se musí zvolit správné pořadí stupňů volnosti a hladina
- příklad výšky desetiletých dětí: $F = \frac{42{,}98}{38{,}94} = 1{,}27 < F_{14,11}(0{,}025) = 3{,}36$
- [var.test(vyska~Divka)]

problém nestejných rozptylů

- předpoklad o stejném rozptylu v obou souborech nemusí být ve skutečnosti splněn, lze jej ověřit porovnáním odhadů rozptylu F-testem $F = \frac{s_x^2}{s_y^2}$
- hypotéza $H_0 : \sigma_x^2 = \sigma_y^2$ se proti $H_1 : \sigma_x^2 \neq \sigma_y^2$ zamítá, když je

$$\text{buď } F = \frac{s_x^2}{s_y^2} \geq F_{n_1-1,n_2-1}(\alpha/2) \text{ nebo } \frac{1}{F} = \frac{s_y^2}{s_x^2} \geq F_{n_2-1,n_1-1}(\alpha/2)$$

- vlastně se větší odhad rozptylu dělí menším odhadem, k tomu se musí zvolit správné pořadí stupňů volnosti a hladina
- příklad výšky desetiletých dětí:
  $F = \frac{42{,}98}{38{,}94} = 1{,}27 < F_{14,11}(0{,}025) = 3{,}36$
- [var.test(vyska∼Divka)]

MS Excel: Dvouvýběrový F-test pro rozptyl

přednáška	Excel	Soubor 1	Soubor 2
průměr	Stř. hodnota	139.13	140.83
rozptyl	Rozptyl	42.98	33.79
rozsah	Pozorování	15	12
stupně vol.	Rozdíl	14	11
F	F	1.27	
p	$\mathrm{P}(F <= f)$ (1)	0.349	
	F krit (1)	2.739	

pozor Excel pracuje **špatně**: uvádí kritickou hodnotu a p-hodnotu pro jednostrannou alternativu odvozenou z hodnoty statistiky F; při oboustranné alternativě je třeba p-hodnotu vynásobit dvěma ve skutečnosti je $\mathrm{P}(F > 1{,}27) = 0{,}349$, takže $p = 2 \cdot 0{,}349 = 0{,}698$ pro oboustrannou alternativu mělo být použito $F_{14,11}(0{,}025) = 3{,}359$

provedení v MS Excelu (nestejné rozptyly)

		Soubor 1	Soubor 2
průměr	Stř. hodnota	139.133	140.833
rozptyl	Rozptyl	42.981	33.788
rozsah	Pozorování	15	12
$H_0 : \mu_1 - \mu_2 =$	Hyp. rozdíl stř. hodnot	0	
stupně vol. f	Rozdíl	25	
T	t stat	-0.713	
p jednostr. testu	P(T<=t) (1)	0.241	
$t_f(2\alpha)$	t krit (1)	1.708	
p oboustr. testu	P(T<=t) (2)	0.482	
$t_f(\alpha)$	t krit (2)	2.060	

při oboustranné alternativě nelze nulovou hypotézu zamítnout

párové testy

- není-li předpoklad **nezávislosti** porovnávaných výběrů splněn, dá dvouvýběrový t-test nesprávný výsledek
- typické porušení předpokladu nezávislosti je u párových dat
  - měření na stejných objektech ve dvou různých časech
  - měření na stejných objektech před zásahem a po něm (ošetření)
  - měření na rodičích
- postup
  - spočítají se a hodnotí rozdíly (změny)
  - přejde se k úloze s jediným výběrem
  - mají-li rozdíly normální rozdělení, pak párový t-test

párové testy

- není-li předpoklad **nezávislosti** porovnávaných výběrů splněn, dá dvouvýběrový t-test nesprávný výsledek
- typické porušení předpokladu nezávislosti je u párových dat
  - měření na stejných objektech ve dvou různých časech
  - měření na stejných objektech před zásahem a po něm (ošetření)
  - měření na rodičích
- postup
  - spočítají se a hodnotí rozdíly (změny)
  - přejde se k úloze s jediným výběrem
  - mají-li rozdíly normální rozdělení, pak párový t-test

párové testy

- není-li předpoklad **nezávislosti** porovnávaných výběrů splněn, dá dvouvýběrový t-test nesprávný výsledek
- typické porušení předpokladu nezávislosti je u párových dat
  - měření na stejných objektech ve dvou různých časech
  - měření na stejných objektech před zásahem a po něm (ošetření)
  - měření na rodičích
- postup
  - spočítají se a hodnotí rozdíly (změny)
  - přejde se k úloze s jediným výběrem
  - mají-li rozdíly normální rozdělení, pak párový t-test

párové testy

- není-li předpoklad **nezávislosti** porovnávaných výběrů splněn, dá dvouvýběrový t-test nesprávný výsledek
- typické porušení předpokladu nezávislosti je u párových dat
  - měření na stejných objektech ve dvou různých časech
  - měření na stejných objektech před zásahem a po něm (ošetření)
  - měření na rodičích
- postup
  - spočítají se a hodnotí rozdíly (změny)
  - přejde se k úloze s jediným výběrem
  - mají-li rozdíly normální rozdělení, pak párový t-test

párové testy

- není-li předpoklad **nezávislosti** porovnávaných výběrů splněn, dá dvouvýběrový t-test nesprávný výsledek
- typické porušení předpokladu nezávislosti je u párových dat
  - měření na stejných objektech ve dvou různých časech
  - měření na stejných objektech před zásahem a po něm (ošetření)
  - měření na rodičích
- postup
  - spočítají se a hodnotí rozdíly (změny)
  - přejde se k úloze s jediným výběrem
  - mají-li rozdíly normální rozdělení, pak párový t-test

párové testy

- není-li předpoklad **nezávislosti** porovnávaných výběrů splněn, dá dvouvýběrový t-test nesprávný výsledek
- typické porušení předpokladu nezávislosti je u párových dat
  - měření na stejných objektech ve dvou různých časech
  - měření na stejných objektech před zásahem a po něm (ošetření)
  - měření na rodičích
- postup
  - spočítají se a hodnotí rozdíly (změny)
  - přejde se k úloze s jediným výběrem
  - mají-li rozdíly normální rozdělení, pak párový t-test

párové testy

- není-li předpoklad **nezávislosti** porovnávaných výběrů splněn, dá dvouvýběrový t-test nesprávný výsledek
- typické porušení předpokladu nezávislosti je u párových dat
  - měření na stejných objektech ve dvou různých časech
  - měření na stejných objektech před zásahem a po něm (ošetření)
  - měření na rodičích
- postup
  - spočítají se a hodnotí rozdíly (změny)
  - přejde se k úloze s jediným výběrem
  - mají-li rozdíly normální rozdělení, pak párový t-test

párové testy

- není-li předpoklad **nezávislosti** porovnávaných výběrů splněn, dá dvouvýběrový t-test nesprávný výsledek
- typické porušení předpokladu nezávislosti je u párových dat
  - měření na stejných objektech ve dvou různých časech
  - měření na stejných objektech před zásahem a po něm (ošetření)
  - měření na rodičích
- postup
  - spočítají se a hodnotí rozdíly (změny)
  - přejde se k úloze s jediným výběrem
  - mají-li rozdíly normální rozdělení, pak párový t-test

párové testy

- není-li předpoklad **nezávislosti** porovnávaných výběrů splněn, dá dvouvýběrový t-test nesprávný výsledek
- typické porušení předpokladu nezávislosti je u párových dat
  - měření na stejných objektech ve dvou různých časech
  - měření na stejných objektech před zásahem a po něm (ošetření)
  - měření na rodičích
- postup
  - spočítají se a hodnotí rozdíly (změny)
  - přejde se k úloze s jediným výběrem
  - mají-li rozdíly normální rozdělení, pak párový t-test

příklad: výška rodičů

- rozhodnout o tvrzení, že populační průměr výšek otců je právě o 10 cm větší než populační průměr výšek matek
- otcové: $\bar{Y} = 179{,}26, s_Y = 6{,}78, n_1 = 99$
  matky: $\bar{Z} = 166{,}97, s_Z = 6{,}11, n_2 = 99$
- otcové jsou (ve výběru) v průměru o $\bar{Y} - \bar{Z} = 12{,}29$ cm vyšší
- směrodatná odchylka **rozdílů** je 8,14 (méně, než kdyby byly výšky rodičů nezávislé … $6{,}78^2 + 6{,}11^2 = 9{,}13^2$)
- **střední chyba** rozdílu průměrů je $8{,}14/\sqrt{99} = 0{,}819$
- rozhodneme podle statistiky [t.test(vyska.o-vyska.m,mu=10)]

$$T = \left| \frac{12{,}29 - 10}{0{,}819} \right| = 2{,}801 > 1{,}984 = t_{98}(0{,}05) \qquad p = 0{,}6\ \%$$

příklad: výška rodičů

- rozhodnout o tvrzení, že populační průměr výšek otců je právě o 10 cm větší než populační průměr výšek matek
- otcové: $\bar{Y} = 179{,}26$, $s_Y = 6{,}78$, $n_1 = 99$
  matky: $\bar{Z} = 166{,}97$, $s_Z = 6{,}11$, $n_2 = 99$
- otcové jsou (ve výběru) v průměru o $\bar{Y} - \bar{Z} = 12{,}29$ cm vyšší
- směrodatná odchylka **rozdílů** je 8,14 (méně, než kdyby byly výšky rodičů nezávislé … $6{,}78^2 + 6{,}11^2 = 9{,}13^2$)
- **střední chyba** rozdílu průměrů je $8{,}14/\sqrt{99} = 0{,}819$
- rozhodneme podle statistiky [t.test(vyska.o-vyska.m,mu=10)]

$$T = \left| \frac{12{,}29 - 10}{0{,}819} \right| = 2{,}801 > 1{,}984 = t_{98}(0{,}05) \qquad p = 0{,}6\ \%$$

příklad: výška rodičů

- rozhodnout o tvrzení, že populační průměr výšek otců je právě o 10 cm větší než populační průměr výšek matek
- otcové: $\bar{Y} = 179{,}26, s_Y = 6{,}78, n_1 = 99$
  matky: $\bar{Z} = 166{,}97, s_Z = 6{,}11, n_2 = 99$
- otcové jsou (ve výběru) v průměru o $\bar{Y} - \bar{Z} = 12{,}29$ cm vyšší
- směrodatná odchylka **rozdílů** je 8,14 (méně, než kdyby byly výšky rodičů nezávislé … $6{,}78^2 + 6{,}11^2 = 9{,}13^2$)
- **střední chyba** rozdílu průměrů je $8{,}14/\sqrt{99} = 0{,}819$
- rozhodneme podle statistiky [t.test(vyska.o-vyska.m,mu=10)]

$$T = \left|\frac{12{,}29 - 10}{0{,}819}\right| = 2{,}801 > 1{,}984 = t_{98}(0{,}05) \qquad p = 0{,}6\ \%$$

příklad: výška rodičů

- rozhodnout o tvrzení, že populační průměr výšek otců je právě o 10 cm větší než populační průměr výšek matek
- otcové: $\bar{Y} = 179{,}26, s_Y = 6{,}78, n_1 = 99$
  matky: $\bar{Z} = 166{,}97, s_Z = 6{,}11, n_2 = 99$
- otcové jsou (ve výběru) v průměru o $\bar{Y} - \bar{Z} = 12{,}29$ cm vyšší
- směrodatná odchylka **rozdílů** je 8,14 (méně, než kdyby byly výšky rodičů nezávislé … $6{,}78^2 + 6{,}11^2 = 9{,}13^2$)
- **střední chyba** rozdílu průměrů je $8{,}14/\sqrt{99} = 0{,}819$
- rozhodneme podle statistiky [t.test(vyska.o-vyska.m,mu=10)]

$$T = \left|\frac{12{,}29 - 10}{0{,}819}\right| = 2{,}801 > 1{,}984 = t_{98}(0{,}05) \qquad p = 0{,}6\ \%$$

příklad: výška rodičů

- rozhodnout o tvrzení, že populační průměr výšek otců je právě o 10 cm větší než populační průměr výšek matek
- otcové: $\bar{Y} = 179{,}26,\ s_Y = 6{,}78,\ n_1 = 99$
  matky: $\bar{Z} = 166{,}97,\ s_Z = 6{,}11,\ n_2 = 99$
- otcové jsou (ve výběru) v průměru o $\bar{Y} - \bar{Z} = 12{,}29$ cm vyšší
- směrodatná odchylka **rozdílů** je 8,14 (méně, než kdyby byly výšky rodičů nezávislé … $6{,}78^2 + 6{,}11^2 = 9{,}13^2$)
- **střední chyba** rozdílu průměrů je $8{,}14/\sqrt{99} = 0{,}819$
- rozhodneme podle statistiky [t.test(vyska.o-vyska.m,mu=10)]

$$T = \left| \frac{12{,}29 - 10}{0{,}819} \right| = 2{,}801 > 1{,}984 = t_{98}(0{,}05) \qquad p = 0{,}6\ \%$$

příklad: výška rodičů

- rozhodnout o tvrzení, že populační průměr výšek otců je právě o 10 cm větší než populační průměr výšek matek
- otcové: $\bar{Y} = 179{,}26$, $s_Y = 6{,}78$, $n_1 = 99$
  matky: $\bar{Z} = 166{,}97$, $s_Z = 6{,}11$, $n_2 = 99$
- otcové jsou (ve výběru) v průměru o $\bar{Y} - \bar{Z} = 12{,}29$ cm vyšší
- směrodatná odchylka **rozdílů** je 8,14 (méně, než kdyby byly výšky rodičů nezávislé . . . $6{,}78^2+6{,}11^2=9{,}13^2$)
- **střední chyba** rozdílu průměrů je $8{,}14/\sqrt{99} = 0{,}819$
- rozhodneme podle statistiky [t.test(vyska.o-vyska.m,mu=10)]

$$T = \left|\frac{12{,}29 - 10}{0{,}819}\right| = 2{,}801 > 1{,}984 = t_{98}(0{,}05) \qquad p = 0{,}6\ \%$$