Testování hypotéz

1. vymezení základních pojmů
2. testování hypotéz o rozdílu průměrů
3. jednovýběrový t-test

Testování hypotéz

- proces, kterým rozhodujeme, zda přijmeme nebo zamítneme *nulovou hypotézu*

Obecný postup testování hypotéz

- 1. Určení statistické hypotézy
- 2. Určení hladiny chyby α
- 3. Výpočet testovací statistiky
- 4. Rozhodnutí

Nulová hypotéza

- hypotéza, kterou se snažíme vyvrátit (falzifikovat)
- Karl Popper (1968) tvrdil, že platnost hypotézy nemůže být nikdy prokázána pouhou generalizací příkladů, které ji potvrzují
  - jak říká filozof Bertrand Russel, krocan-vědec by mohl zobecnit tvrzení "každý den mě krmí", protože tato hypotéza je potvrzována den po dni celý jeho život. Tato generalizace ovšem neposkytuje žádnou jistotu, že krocan bude nakrmen i další den - některý den se pravděpodobně on sám stane pokrmem

Nulová hypotéza

- Popper došel k závěru, že jedinou možnou metodou je *falsifikace* hypotézy - nalezení jednoho příkladu, který stačí k jejímu vyvrácení
- vědci se proto snaží své hypotézy vyvrátit a tak potvrdit hypotézy opačné - alternativní

Nulová hypotéza

- nulová hypotéza je opakem naší výzkumné hypotézy
- obvykle zní: mezi dvěma průměry není rozdíl, korelace je nulová apod.
- např. *průměrná výška mužů a žen se neliší*
- označuje se H_0

Alternativní hypotéza

- H_1
- alternativní vzhledem k nulové, tj. naše výzkumná hypotéza
- např.
  - *průměrná výška mužů a žen se liší* (tzv. **oboustranná** hypotéza)

  nebo
  - *průměrná výška mužů je větší než průměrná výška žen* (tzv. **jednostranná** hypotéza)

Příklad testování hypotéz o rozdílu průměrů

- chceme zjistit, jaký vliv má v raném věku dítěte(<6 měsíců) hospitalizace bez matky na IQ dítěte v 7 letech
- vyšetříme vzorek 36 dětí náhodně vybraných z této populace
- zjistíme průměrné IQ 96 se směrodatnou odchylkou 15 bodů

Příklad testování hypotéz o rozdílu průměrů

- Můžeme na základě těchto výsledků tvrdit, že průměrné IQ dětí hospitalizovaných v raném věku bez matky se liší od průměrného IQ populace všech dětí (=100)?

Příklad testování hypotéz o rozdílu průměrů

- **Nulová hypotéza (H_0):**
průměrné IQ dětí hospitalizovaných v raném věku bez matky je stejné jako průměrné IQ populace všech dětí

- *jinými slovy:* ***není*** *nepravděpodobné, že vzorek 36 dětí má čistě náhodou průměr 96, pokud je průměr populace 100 a směrodatná odchylka 15*

Příklad testování hypotéz o rozdílu průměrů

- **Alternativní hypotéza (H_1):** průměrné IQ dětí hospitalizovaných v raném věku bez matky je nižší než průměrné IQ populace všech dětí
- *půjde o jednostranné testování hypotéz*

Hladina významnosti

- hladina významnosti je úroveň pravděpodobnosti, kterou používáme při rozhodování, zda zamítnout nebo přijmout nulovou hypotézu
- označuje se alfa (α)
- obvyklá hladina významnosti je 5% nebo 1% - volíme podle vlastního uvážení

Chyba I. druhu

- zvolíme-li hladinu významnosti 5%, pak se rozhodneme zamítnout nulovou hypotézu tehdy, když existuje pouze 5% pravděpodobnost našich dat v případě, že H_0 platí
- jde vlastně o 5% riziko, že nulová hypotéza platí a my ji přitom zamítneme –uděláme tzv. chybu I. druhu

Chyba II. druhu

- opak chyby I. druhu – riziko, že nezamítneme nulovou hypotézu, která ve skutečnosti neplatí
- označuje se beta (β)

Chyby typu I a II

	nulová hypotéza **platí**	nulová hypotéza **neplatí**
zamítneme nulovou hypotézu	**chyba I. druhu**	správné rozhodnutí
nezamítneme nulovou hypotézu	správné rozhodnutí	**chyba II. druhu**

Příklad testování hypotéz o rozdílu průměrů

- **Hladina významnosti:** v našem příkladu použijeme α **=5% = 0,05**
- pokud je pravděpodobnost získání vzorku o průměru 96 z populace o průměru 100 menší než 5%, pak zamítneme H_0
- pokud je pravděpodobnost získání vzorku o průměru 96 větší než 5%, pak H_0 nezamítneme

Výpočet testovací statistiky

- závisí na povaze dat a hypotéze
- pro testy hypotéz o rozdílu průměrů se používá standardizovaná vzdálenost odhadu od nulové hypotézy
- **testovací statistika =**

  (bodový odhad – hypotetická hodnota) / směrodatná chyba odhadu

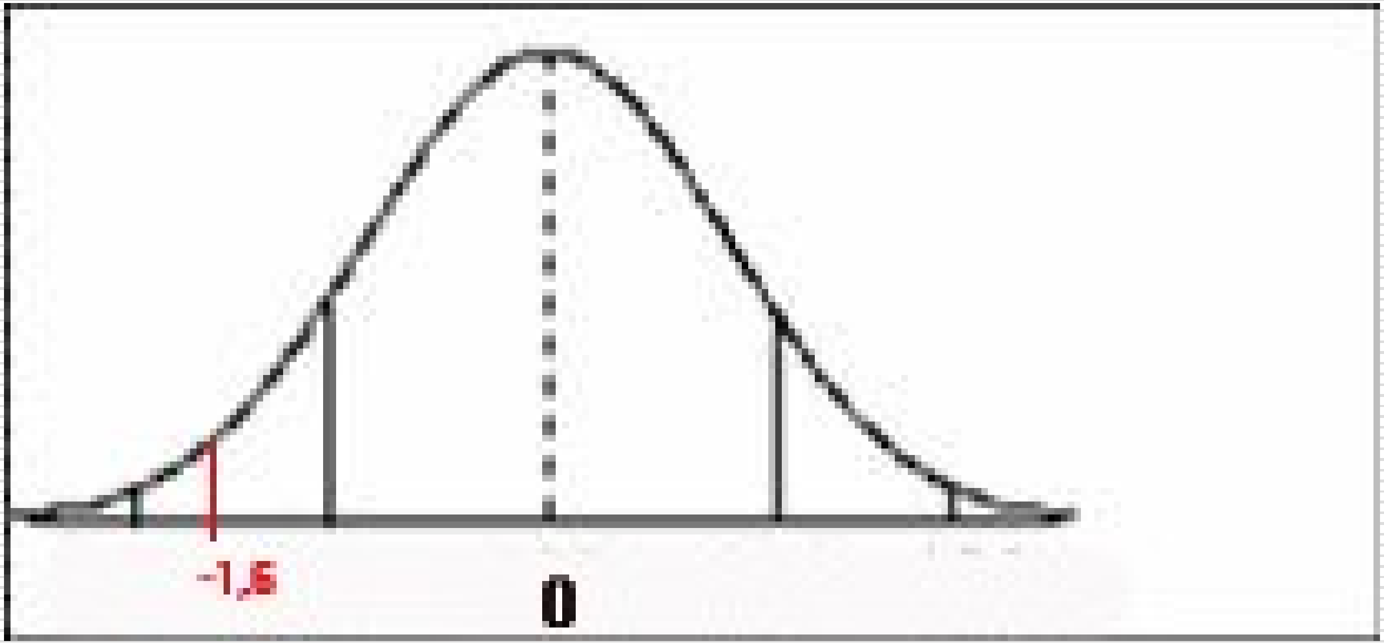
Příklad testování hypotéz o rozdílu průměrů

- $z = (\bar{x} - \mu_0) / \sigma_{\bar{x}}$
- $z = (\bar{x} - \mu_0) / (\sigma / \sqrt{n})$

- $z = (96-100) / (15/\sqrt{36})$
- $z = -4 / 2{,}5$
- $\mathbf{z = -1{,}6}$

Příklad testování hypotéz o rozdílu průměrů

- Rozdělení výběrových průměrů



Rozhodnutí o závěru testování hypotéz

2 možnosti

- 1) převedeme testovací statistiku na tzv. hodnotu významnosti *p*

nebo

- 2) srovnáme testovací statistiku s tzv. kritickou mezí

Hodnota významnosti *p*

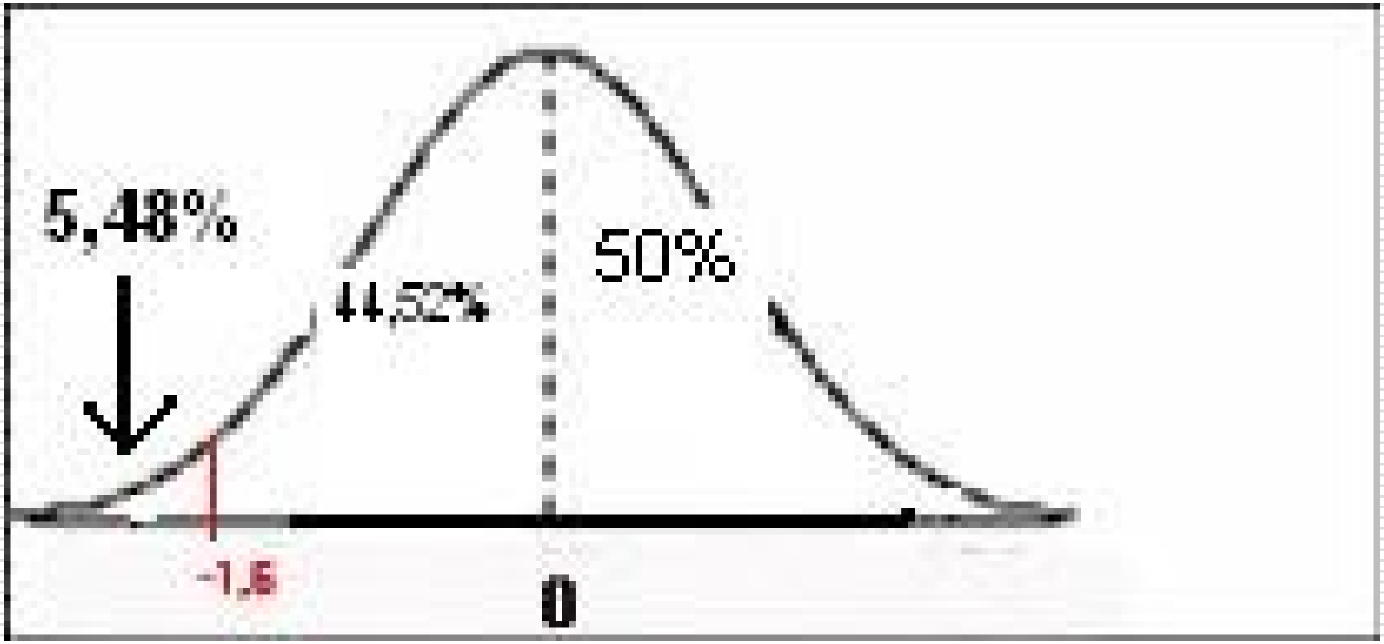
- pravděpodobnost realizace testovací statistiky za předpokladu, že platí nulová hypotéza („jestliže platí H_0, jaká je pravděpodobnost, že získáme tuto vypočítanou hodnotu?")
- pokud je p menší než hladina významnosti α nebo stejná, pak můžeme nulovou hypotézu zamítnout

Příklad testování hypotéz o rozdílu průměrů

- jaká je hodnota významnosti *p* pro $z = -1{,}6$?
- v tabulce z-rozdělení najdeme pravděpodobnost pro $z \leq -1{,}6$

Příklad testování hypotéz o rozdílu průměrů

- Rozdělení výběrových průměrů



Příklad testování hypotéz o rozdílu průměrů

- $p = 0{,}0548$
- $p > \alpha$
- nemáme dostatečné důkazy pro to, abychom zamítli nulovou hypotézu

Srovnání s kritickou mezí

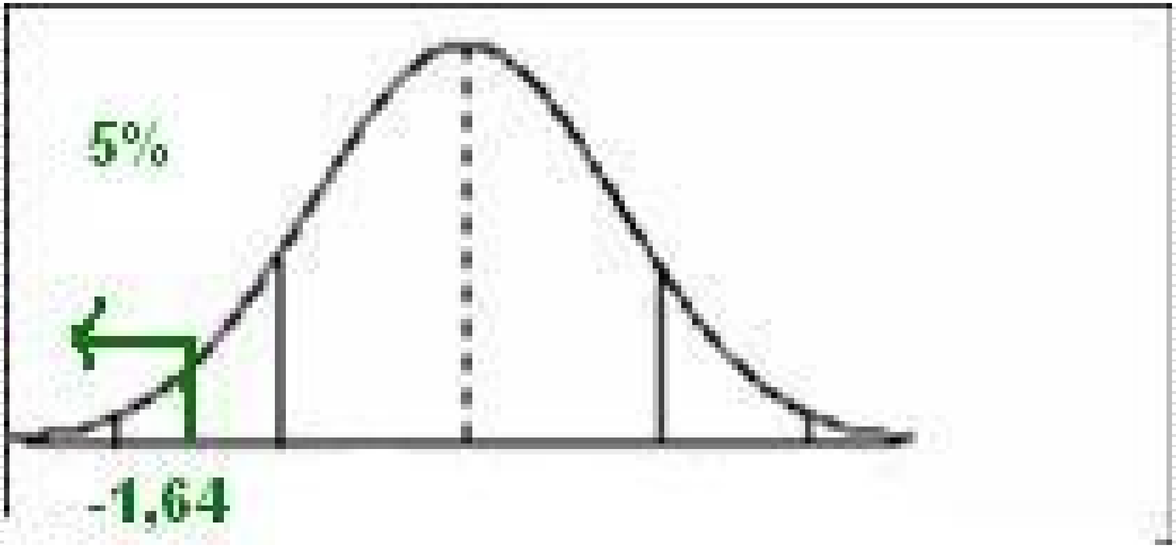
- kritická mez se stanoví na základě hladiny významnosti (α)
  - tzv. kritická oblast nebo oblast zamítnutí
- jestliže je testovací statistika v této kritické oblasti, pak můžeme zamítnout nulovou hypotézu

Příklad testování hypotéz o rozdílu průměrů

- najdeme v tabulce z-rozdělení hodnotu z, která odděluje nejnižších 5% případů
- **z = -1,64**
- = kritická mez pro jednostranný test hypotézy při α = 0,05

Příklad testování hypotéz o rozdílu průměrů

- Rozdělení výběrových průměrů

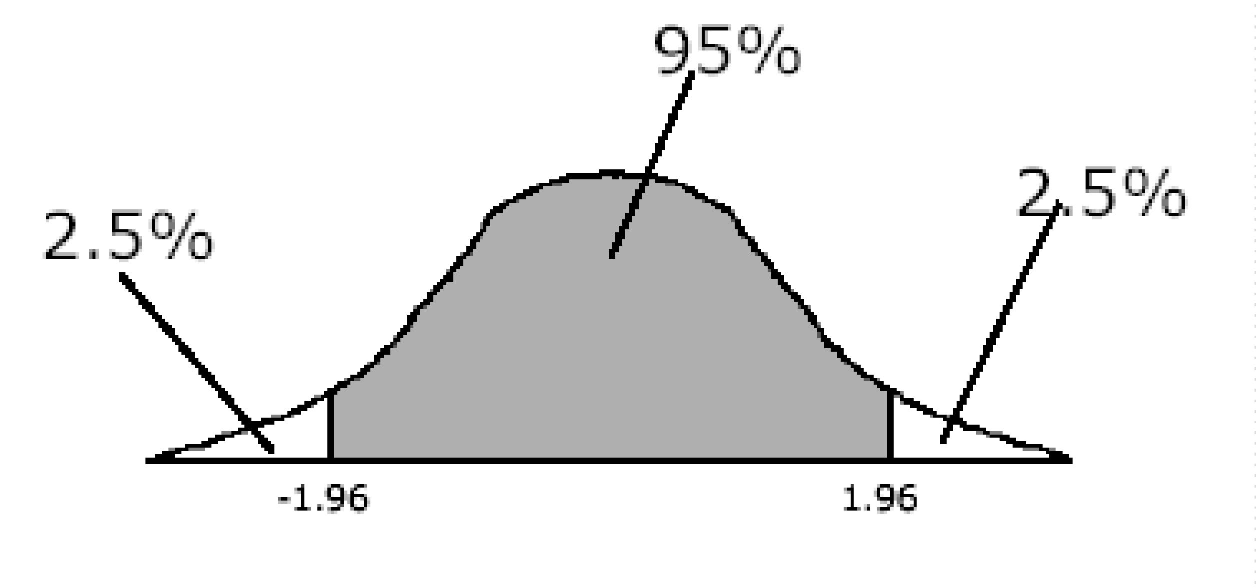


Příklad testování hypotéz o rozdílu průměrů

- vypočítaná hodnota z = -1,6 nespadá do kritické oblasti
- nemůžeme tedy zamítnout nulovou hypotézu

Kritická mez pro oboustranný test

- z = -1,96 a z = +1,96



Rozhodnutí o závěru testování hypotéz

- nemůžeme-li nulovou hypotézu zamítnout, neznamená to nutně, že platí – pouze nemáme dostatek důkazů pro její zamítnutí
- hodnota významnosti *p* není pravděpodobnost, že nulová hypotéza platí

Testování hypotéz o rozdílu průměrů

- 4 možné typy problémů:
  - porovnáváme **průměr vzorku s průměrem populace** → jednovýběrový t-test
  - porovnáváme **průměry dvou vzorků** → t-test pro nezávislé výběry
  - porovnáváme **dva průměry jednoho vzorku** → t-test pro závislé výběry (tzv. párový t-test)
  - porovnáváme více průměrů → analýza rozptylu

Jednovýběrový t-test - příklad

- Rozhodujeme se mezi jazykovými školami v Brně. Podaří se nám zjistit, že při zkouškách na Britské radě získávají absolventi různých jazykových škol průměrně 85 bodů, ale neznáme směrodatnou odchylku průměru.
- Jedna ze škol – ABC - se chlubí, že její absolventi dosahují nadprůměrných výsledků.

Jednovýběrový t-test - příklad

- Zjistíme, že posledních zkoušek se účastnilo 10 absolventů školy ABC s těmito výsledky:

  80 91 92 87 89 88 86 80 90 89

- Můžeme na základě výsledků tohoto vzorku 10 absolventů dojít k závěru, že škola ABC má lepší průměrné výsledky než ostatní školy v Brně?

Jednovýběrový t-test

- průměr vzorku je 87.2
- směrodatná odchylka 4.18
- známe průměr populace (μ=85), ale nikoli směrodatnou odchylku populace (místo ní použijeme jako odhad směrodatnou odchylku vzorku)

Jednovýběrový t-test - příklad

- **Nulová hypotéza:** průměrné výsledky absolventů školy ABC se neliší od výsledků absolventů ostatních škol
- jinými slovy: **není** nepravděpodobné, že vzorek má čistě náhodou průměr 87.2, pokud je průměr populace 85 a směrodatná odchylka 4.18

Jednovýběrový t-test

- **Alternativní hypotéza:** průměrné výsledky absolventů školy ABC jsou lepší než výsledky absolventů ostatních škol

Jednovýběrový t-test

- **Hladina významnosti:** použijeme α =5%
- pokud je pravděpodobnost získání vzorku o průměru 87.2 menší než 5%, pak zamítneme H_0
- pokud je pravděpodobnost získání vzorku o průměru 87.2 větší než 5%, pak H_0 nezamítneme

Jednovýběrový t-test

- potřebujeme spočítat, jaká je pravděpodobnost získání vzorku (n=10) o průměru 87.2 z populace o průměru 85 a směrodatné odchylce 4.18
- vzhledem k tomu, že velikost směrodatné odchylky jsme odhadli ze vzorku, nemůžeme pro rozdělení výběrových průměrů použít z-rozdělení, ale *Studentovo rozdělení t*

Studentovo rozdělení

- pokud **za σ nahradíme *s*** (směr. odchylku výběrového průměru), pak musíme při konstrukci rozdělení výběrových průměrů místo *z* rozdělení použít tzv. **Studentovo t rozdělení**

Rozdělení výběrových průměrů

pro **neznámé** hodnoty směrodatné odchylky v populaci:

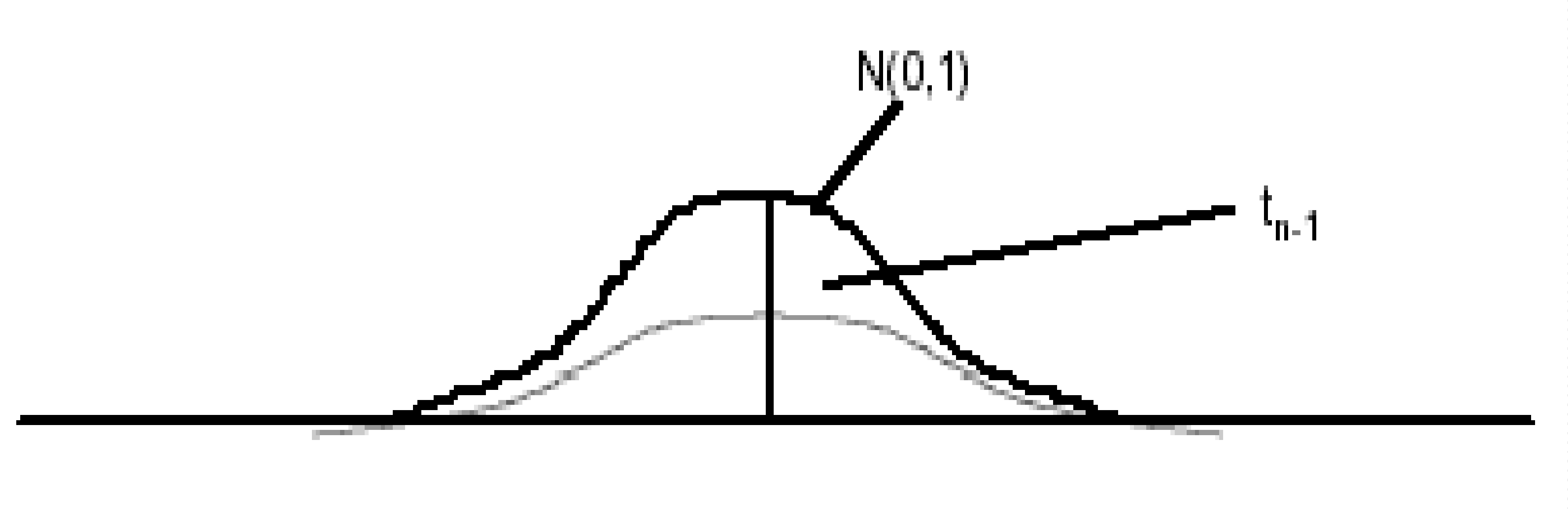
$$t = \frac{\overline{x} - \mu}{\frac{s}{\sqrt{n}}}$$

Studentovo rozdělení

- má také zvonovitý tvar, ale je více ploché než normální rozdělení
- je symetrické kolem průměru (0)
- pro každou velikost výběru (počet stupňů volnosti, df) existuje odlišné t rozdělení
  df = n-1

Studentovo rozdělení

srovnání s normálním rozdělením



Studentovo rozdělení

- srovnání s normálním rozdělením:
  - t rozdělení má vyšší variabilitu
  - více plochy na okrajích, méně ve středu
  - vzhledem k vyšší variabilitě budou intervaly spolehlivosti širší než u normálního rozdělení
  - jsou uváděny df obvykle jen do 100, protože pro n=100 se t rozdělení blíží normálnímu rozdělení

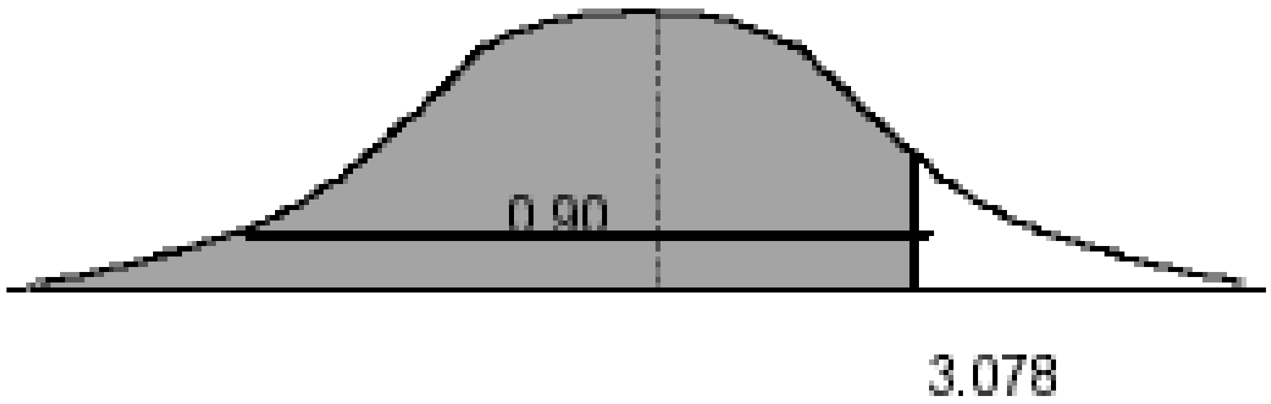
Studentovo rozdělení

- tabulka t-rozdělení:
  - každý řádek udává hodnoty t pro celé rozdělení pro daný počet stupňů volnosti (tj. n-1)
  - sloupce pro nejdůležitější percentily

Studentovo rozdělení

d.f.	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
1	3.078	6.3138	12.706	31.821	63.657
2	1.886	2.92	4.3027	6.965	9.9248
3	1.638	2.3534	3.1825	4.541	5.8409

Studentovo rozdělení



Jednovýběrový t-test

- potřebujeme spočítat, jaká je pravděpodobnost získání vzorku (n=10) o průměru 87.2 z populace o průměru 85 a směrodatné odchylce 4.18
- vzhledem k tomu, že velikost směrodatné odchylky jsme odhadli ze vzorku, nemůžeme použít z-rozdělení, ale *Studentovo rozdělení t*

Jednovýběrový t-test

$$t = \frac{\overline{X} - \mu}{\frac{s}{\sqrt{N}}}$$

Jednovýběrový t-test

- t = (87.2-85) / (4.18/$\sqrt{10}$)
  t = 2.2/1.32
  t = 1.66
- df = n-1 = 10 – 1 = **9**
  (počet stupňů volnosti pro vyhledání pravděpodobnosti v tabulce t-rozdělení)

Jednovýběrový t-test

- kritická hodnota t pro α=5% je 1,833
- získaná hodnota t je 1,66

Tabulka t-rozdělení

df	$\alpha = 0.1$	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	**6.314**	12.706	31.821	63.656	318.289	636.578
2	1.886	**2.920**	4.303	6.965	9.925	22.328	31.600
3	1.638	**2.353**	3.182	4.541	5.841	10.214	12.924
4	1.533	**2.132**	2.776	3.747	4.604	7.173	8.610
5	1.476	**2.015**	2.571	3.365	4.032	5.894	6.869
6	1.440	**1.943**	2.447	3.143	3.707	5.208	5.959
7	1.415	**1.895**	2.365	2.998	3.499	4.785	5.408
8	1.397	**1.860**	2.306	2.896	3.355	4.501	5.041
9	1.383	**1.833**	2.262	2.821	3.250	4.297	4.781
10	1.372	**1.812**	2.228	2.764	3.169	4.144	4.587
11	1.363	**1.796**	2.201	2.718	3.106	4.025	4.437

Jednovýběrový t-test

- v našem příkladě je 1,66<1,883
- tj. **nemůžeme zamítnout nulovou hypotézu** (rozdíl průměrů není tzv. statisticky významný)
- a náš závěr: nemůžeme tvrdit, že výsledky absolventů školy ABC se liší od průměru brněnských škol (je vyšší než 5% pravděpodobnost, že průměrný výsledek 87,2 deseti jejích absolventů je lepší jen náhodou)

Jednovýběrový t-test v SPSS

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
body	10	87,20	4,185	1,323

One-Sample Test

	Test Value = 85					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
body	1,663	9	,131	2,200	-,79	5,19

Kontrolní otázky

- vysvětlete pojmy
  - *nulová a alternativní hypotéza*
  - *testování hypotéz*
  - *chyba I. druhu a chyba II. druhu*
- jaké testy se používají pro testování hypotéz o rozdílu průměrů?