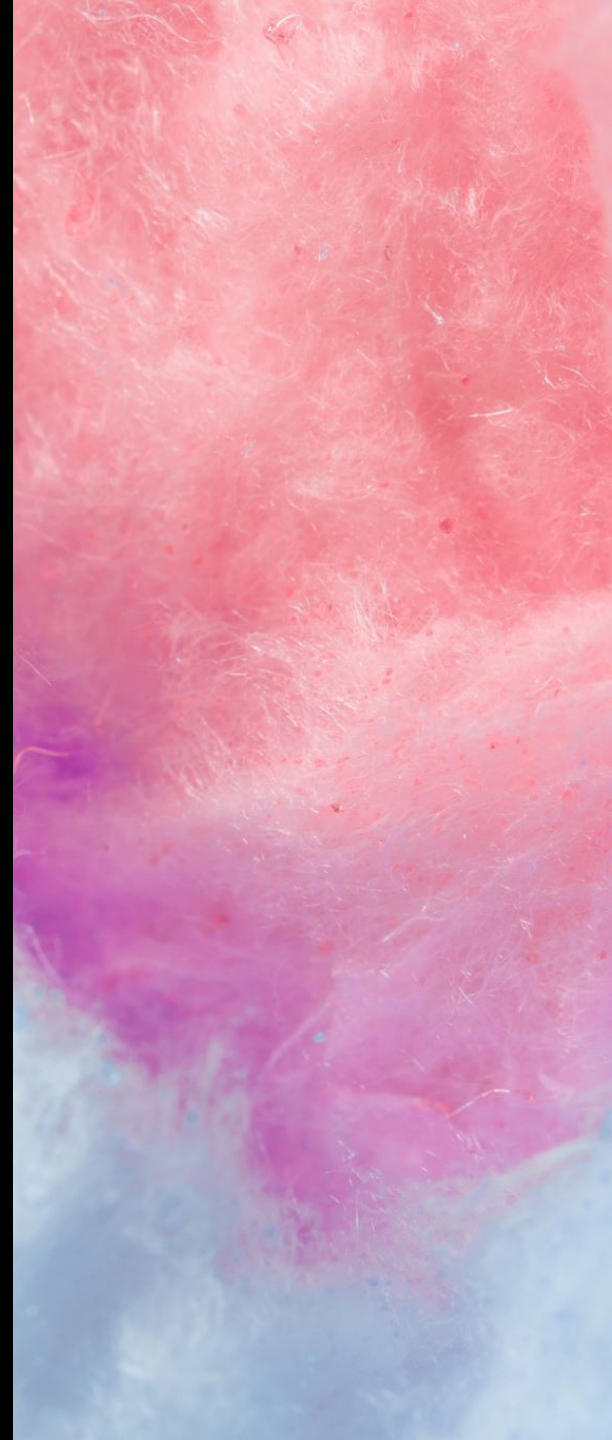




# Hypotézy, testování hypotéz

---

7. 11. 2020



# Statistická inference

- Odhad populačního parametru ze vzorku
- Př: průměr a interval spolehlivosti
  
- Hypotézy a jejich testování



# Hypotéza

- Vědecké tvrzení
- Ověřitelné

# Statistická hypotéza

- Převod naší otázky do jazyka statistického testování
- Testujeme hypotézu že předpokládaný efekt existuje oproti hypotéze, že předpokládaný efekt neexistuje
- ALTERNATIVNÍ HYPOTÉZA (H1)
- NULOVÁ HYPOTÉZA (H0)
- Průměrná výška mužů se liší od průměrné výšky žen (H1)
- Průměrná výška mužů se neliší od průměrné výšky žen (H0)

# Jednostranná vs. dvoustranná hypotéza

- Dvoustranná, neudává směr efektu
  - Průměrná výška mužů se liší od průměrné výšky žen ( $x_m - x_z \neq 0$ )
- Jednostranná předpokládá i směr efektu
  - Průměrná výška mužů je větší než průměrná výška žen ( $x_m - x_z > 0$ )

# Výsledek testu

- Na základě dat neumíme přijmout hypotézu
- Umíme odmítnout nulovou hypotézu
- Pokud se průměrná výška mužů a žen liší:
  - Zamítáme nulovou hypotézu, že se výška mužů a žen neliší.
- Pokud nám data dají dostatečnou jistotu, že můžeme zamítnout nulovou hypotézu, znamená to, že naše data podporují alternativní hypotézu (hypotéza se ale nepřijímá!)
- Naše data jsou v této podobě velmi nepravděpodobná, pokud by nulová hypotéza platila.

# Testování hypotéz

- $H_0$  a  $H_1$  se navzájem vylučují
- Vždy předpokládáme, že  $H_0$  je pravdivá (platí)
- Pokud nemáme dostatečně silné důkazy, tak  $H_0$  zamítáme
- Fakticky  $H_0$  nepřijímáme.

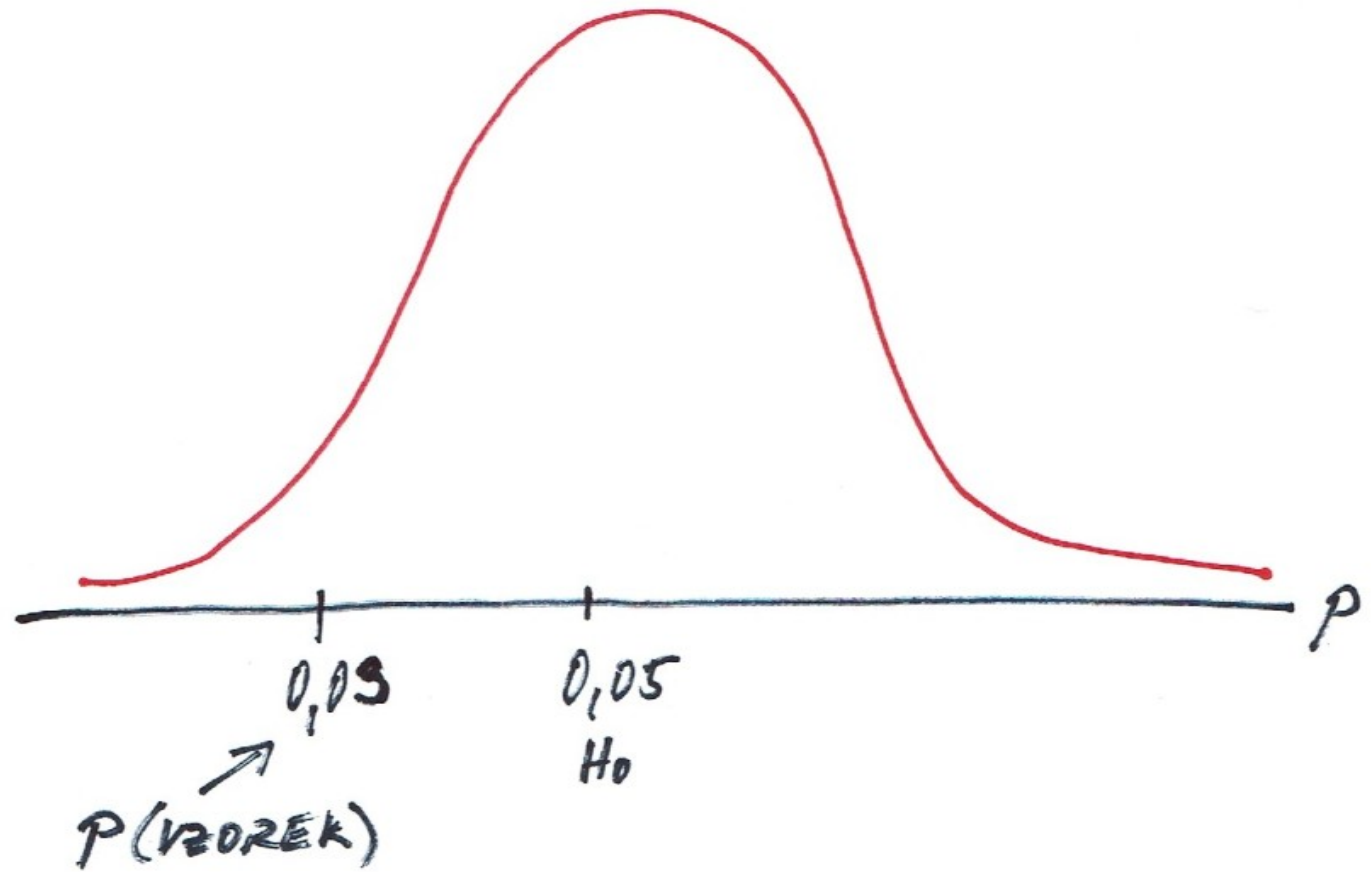
# Předpoklad: v populaci voličů ČR je méně než 5 % voličů Trikolóry

- $H_1: \pi < 0.05$
- $H_0: \pi = 0.05$
- Změříme preference na vzorku 1000 lidí. Ve vzorku by Trikolóru volily 3 % respondentů.
- $\pi_0 = 0.05$
- $p = 0.03$
- $n = 1000$

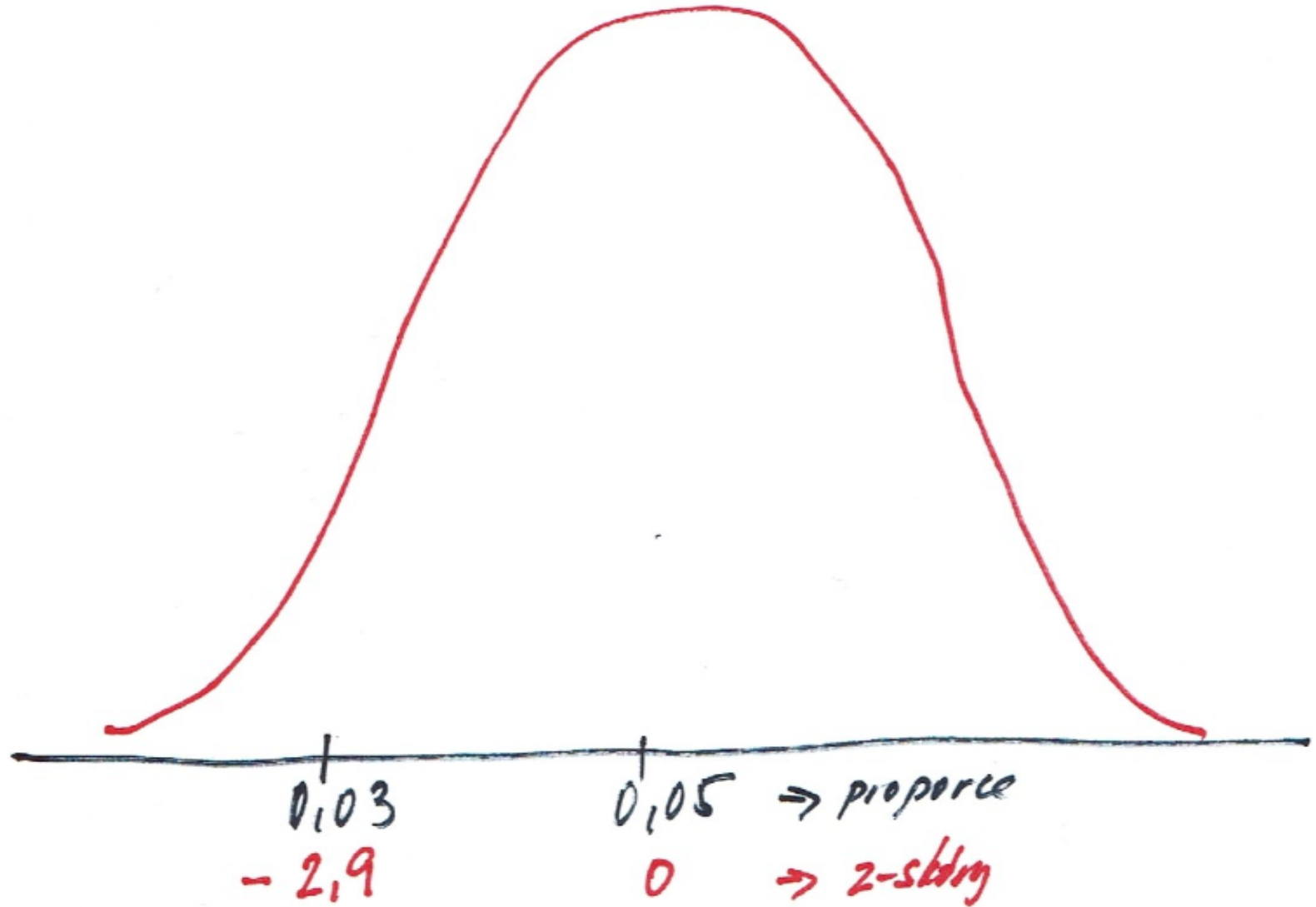
$$Z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1 - \pi_0)}{n}}} \rightarrow \text{Směrodatná chyba}$$

$$Z = \frac{-0,02}{0.00689} = -2.90$$





$$Z = \frac{p - \pi_0}{\frac{\sqrt{\pi_0 (1 - \pi_0)}}{n}} = \frac{-0,02}{0,00689} = -2,9$$



Za předpokladu, že platí nulová hypotéza, je pravděpodobnost, že budu mít z-skór v této hodnotě 0,00187.

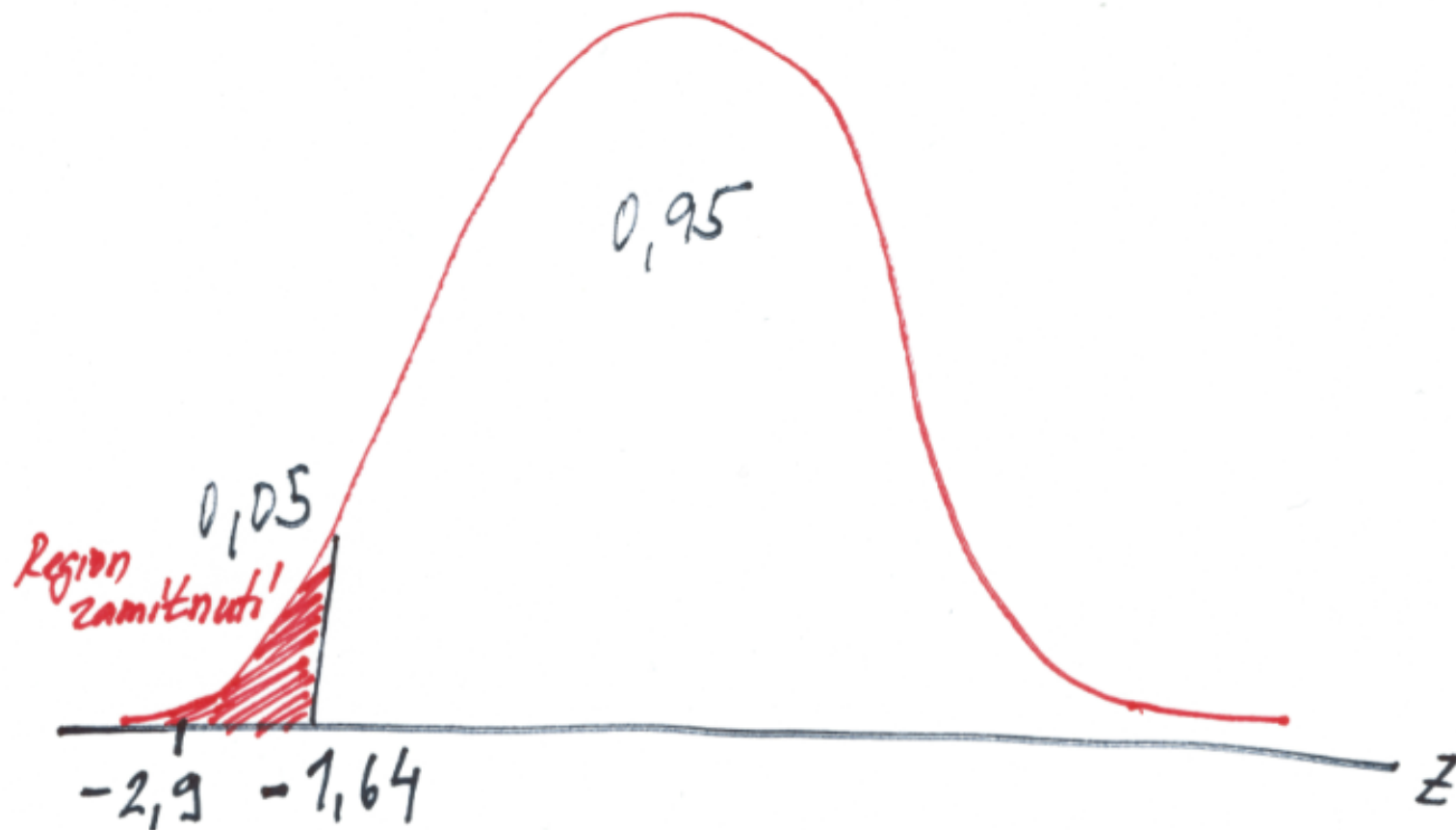
Této pravděpodobnosti se říká P-HODNOTA (p-value)

z		0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-0	.50000	.49601	.49202	.48803	.48405	.48006	.47608	.47210	.46812	.46414
-0.1	.46017	.45620	.45224	.44828	.44433	.44034	.43640	.43251	.42858	.42465
-0.2	.42074	.41683	.41294	.40905	.40517	.40129	.39743	.39358	.38974	.38591
-0.3	.38209	.37828	.37448	.37070	.36693	.36317	.35942	.35569	.35197	.34827
-0.4	.34458	.34090	.33724	.33360	.32997	.32636	.32276	.31918	.31561	.31207
-0.5	.30854	.30503	.30153	.29806	.29460	.29116	.28774	.28434	.28096	.27760
-0.6	.27425	.27093	.26763	.26435	.26109	.25785	.25463	.25143	.24825	.24510
-0.7	.24196	.23885	.23576	.23270	.22965	.22663	.22363	.22065	.21770	.21476
-0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673
-0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
-1	.15866	.15625	.15386	.15151	.14917	.14686	.14457	.14231	.14007	.13786
-1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
-1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
-1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08692	.08534	.08379	.08226
-1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
-1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
-1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
-1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
-1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938
-1.9	.02872	.02807	.02743	.02680	.02619	.02559	.02500	.02442	.02385	.02330
-2	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
-2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
-2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
-2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
-2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
-2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
-2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
-2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
-2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
-2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139
-3	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
-3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
-3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050

# Je tato pravděpodobnost dostatečně malá k odmítnutí $H_0$ ?

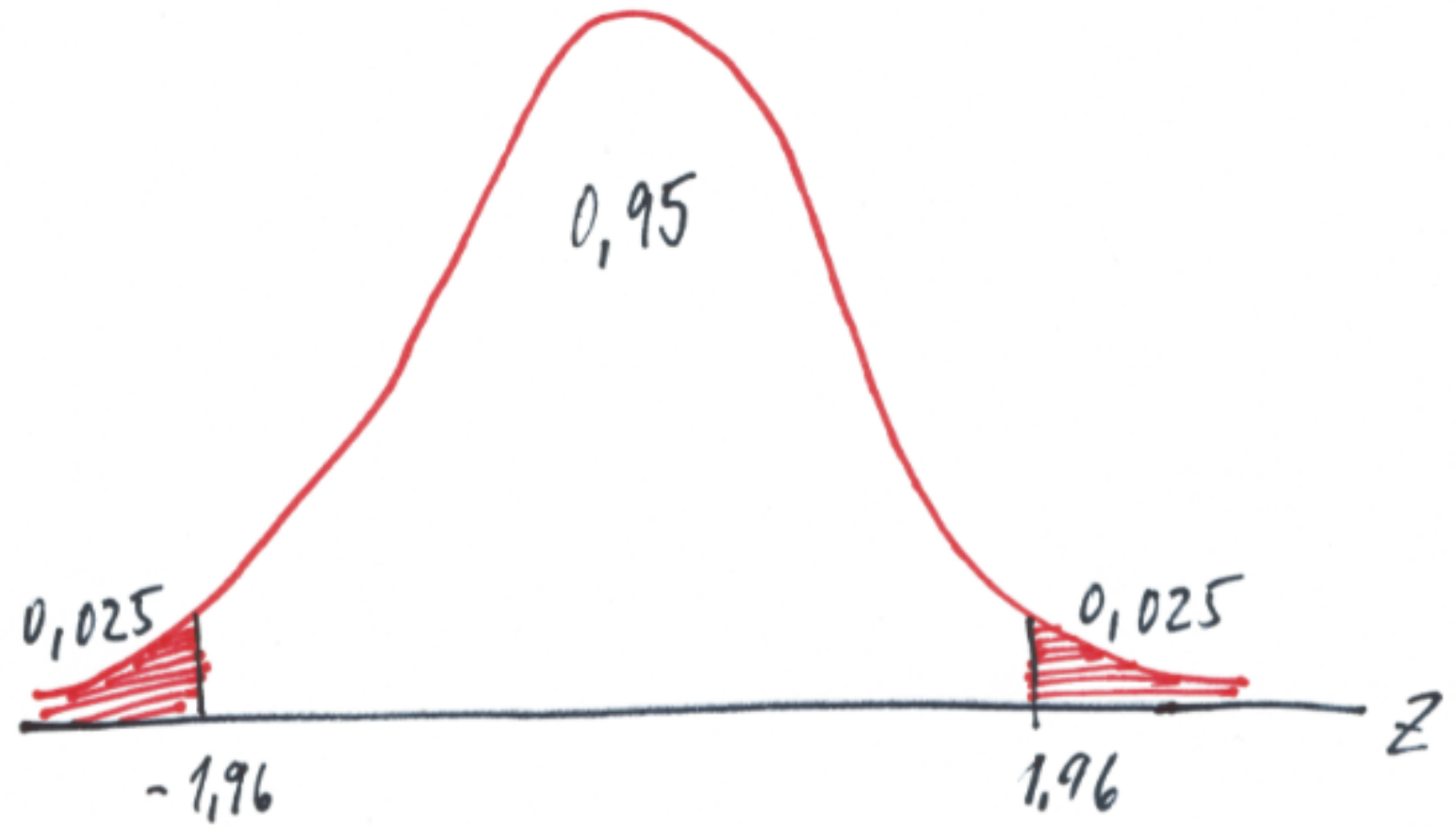
- Záleží na tom, s čím ji porovnávám.
- Arbitrárně si určím hranici.
- Typicky: 0,05
- Je-li p-hodnota stejně velká nebo menší, mám tedy dostatečně silný důkaz, že moje data jsou velmi málo pravděpodobná na to, aby platila  $H_0$ . Proto  $H_0$  zamítám.

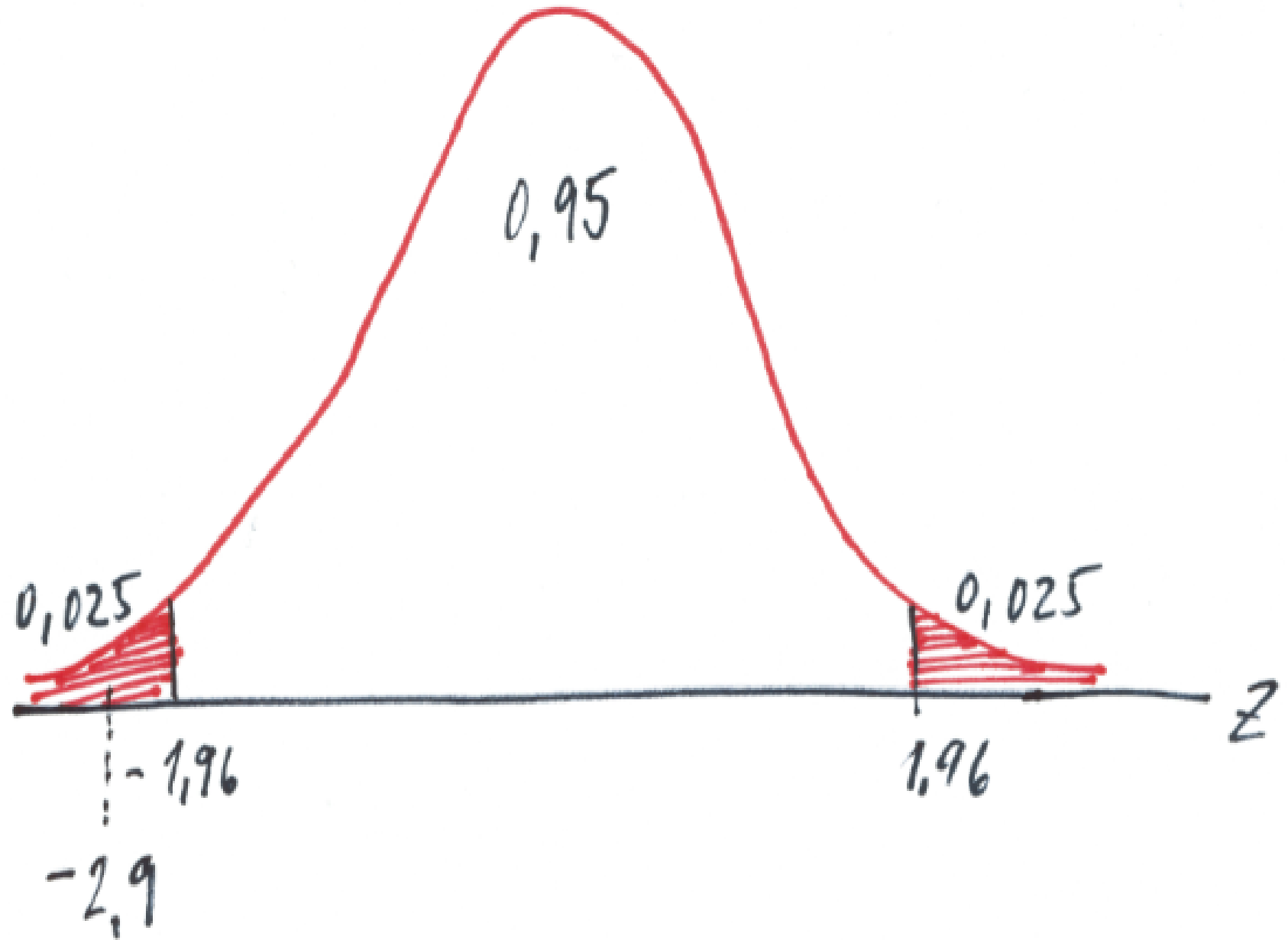
**Kde leží  
moje  
hodnota -  
2,9?**



# Co když mám dvoustrannou hypotézu?

- Teď jsme testovali alternativní hypotézu, která měla směr (méně než 0,05)
- Dvousměrná hypotéza: Proporce voličů Trikolóry se nerovná 0,05.
- $H_1: \pi \neq 0,05$  a  $H_0: \pi = 0,05$







# Jak statistická významnost souvisí s intervaly spolehlivosti?

- Jak se počítá CI pro populační proporcii?

- **Horní hranice CI:**  $p + 1,96 \times \sqrt{p(1-p)/n}$

- **Dolní hranice CI:**  $p - 1,96 \times \sqrt{p(1-p)/n}$

- $p$  = pozorovaná **relativní četnost v našem vzorku** (%)

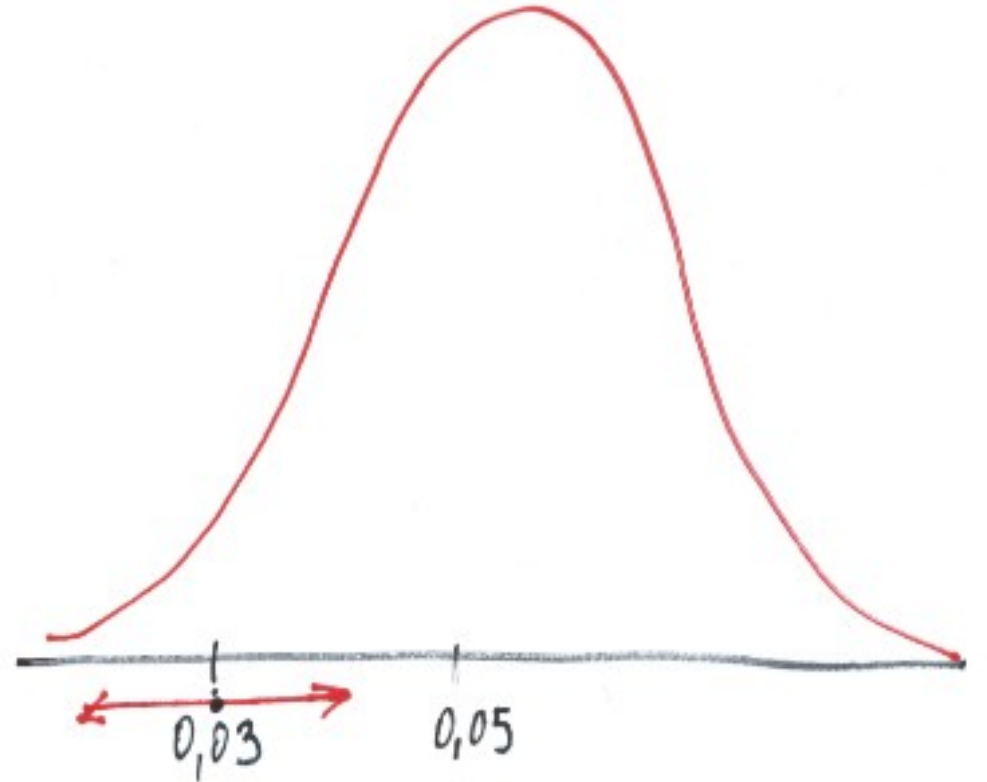
- $n$  = velikost vzorku

V našem případě:

$n = 1000$

$p = 0,03$

- Interval spolehlivosti: 0,019 – 0,041
- Neleží v něm hodnota  $H_0$  (0,05)
- Pokud je p-hodnota oboustranného testu menší nebo rovna 0,05, pak CI nezahrnuje hodnotu  $H_0$ .
- Pokud je p-hodnota oboustranného testu větší než 0,05, pak CI zahrnuje hodnotu  $H_0$ .



$$p \pm 1,96 \cdot (\sqrt{p(1-p)/n})$$
$$0,03 \pm 1,96 \cdot 0,00539442 = CI(0,019; 0,041)$$

# Důležité poznámky na závěr

- Hladina statistické významnosti 0,05 je čistě arbitrární (Field, s. 54)
- Statistická významnost výsledku nemusí implikovat, že je výsledný efekt (rozdíl, síla vztahu..) podstatný.
- Interpretace faktické velikosti efektu (Field, s. 57-58).
- Testové statistiky jsou různé (ne jen z-skóry).
- Odhady, CI, p-hodnoty = pokud děláme inference ze vzorku na populaci.
- Pro populační data není třeba.