7_Intervaly spolehlivosti

8_1Bodový a intervalový odhad populačního parametru

8_2Interval spolehlivosti pro odhad populační proporce

8_3Interval spolehlivosti pro odhad populačního průměru

Populace vs. vzorek

	Populace	Vzorek
Proporce	p	$\hat{p}$
Průměr	μ	$\bar{x}$
Rozptyl (variance)	σ^2	s^2
Směrodatná odchylka	σ	s

parametry

statistiky (odhady parametrů)

- Inference používá statistiky (průměr,proporce) z náhodně vybraného vzorku za účelem rozhodování o hodnotě parametrů v populaci

Souvislost s minulými tématy

- Statistická inference užívá počty pravděpodobnosti, které předpokládají že data byla získána náhodným výběrem
- Výpočty pravděpodobností se vztahují k výběrové distribuci statistiky, která se často blíží normálnímu rozložení

Dvě metody statistické inference

- Odhad populačních parametrů
  - Bodový
    - jaký je nejlepší odhad charakteristiky populace? (např. $\bar{x}$ je bodovým odhadem μ)
  - Intervalový
    - jaký je interval hodnot, který s vysokou pravděpodobností obsahuje charakteristiku populace?(např. μ leží s 99% pravděpodobností mezi 95 a 105) = interval spolehlivosti
- Testování hypotéz o hodnotách parametru
  - Za předpokladu že populační průměr 100, pravděpodobnost že vytáhnu vzorek s průměrem 110 a větším je dostatečně vysoká a proto nezamítám nulovou hypotézu že průměr populace je 100

Bodový odhad

- Jediné číslo (hodnota) která je nejlepším odhadem charakteristiky populace
- např. x¯ je bodovým odhadem μ, p^ je nejlepším odhadem p
- Vlastnosti dobrého estimátoru
  - Nevychýlenost (výběrová distribuce je vycentrovaná kolem parametru)
  - Nízká směrodatná odchylka
- Průměr vzorku je dobrým estimátorem populačního průměru

Intervalový odhad

- Interval hodnot, uvnitř kterého je pravděpodobné, že leží populační parametr
- Obsahuje hodnotu parametru s určitou zvolenou pravděpodobností (spolehlivostí) = interval spolehlivosti
- Interval spolehlivosti obsahuje nejpravděpodobnější hodnoty parametru
- Pravděpodobnost že tato metoda produkuje interval obsahující parametr = úroveň/míra spolehlivosti (nejčastěji 0.95 a 0.99)
  - 95% interval spolehlivosti = máme 95% pravděpodobnost, že interval obsahuje parametr
  - 99% interval spolehlivosti = máme 99% pravděpodobnost, že interval obsahuje parametr

Logika konstrukce intervalu spolehlivosti a výběrová chyba

- Prostřednictvím výběrové distribuce, která nám umožňuje určit pravděpodobnost s kterou bodový odhad leží uvnitř konkrétní vzdálenosti od parametru
- Např. z definice normálního rozložení víme, že s 95 % pravděpodobností proporce vzorku spadá do intervalu +-1.96 směrodatné odchylky (SE) od populační proporce
  - Vzdálenost 1.96 (SE) = výběrová chyba
- Výběrová chyba = měří jak přesný je bodový odhad parametru
  - Je násobkem směrodatné odchylky výběrové distribuce, např. 1.96 násobkem pokud je výběrová distribuce normálně rozložena
- Pokud hodnota proporce vzorku leží v rozmezí +- 1.96 (SE) od proporce populace, pak interval proporce vzorku +-1.96 (SE) obsahuje populační proporci. Tedy tento interval obsahuje populační proporci s 95 % pravděpodobností = je 95% intervalem spolehlivosti

Směrodatná chyba

- Směrodatná odchylka výběrové distribuce vychází z parametru , ten je však neznámý, proto za něj dosazujeme odhad = statistiku (např. p^ namísto p)
- Směrodatná chyba (SE) = odhadovaná směrodatná odchylka výběrové distribuce
- SE (proporce) = $\sqrt{\frac{p \wedge (1 - p \wedge)}{n}}$
- Směrodatná chyba průměru = s / √n
- Směrodatná chyba proporce = √ (p^ (1 - p^) / n)

Interval spolehlivosti pro populační proporci

- = p^ +- výběrová chyba
- = p^ +- z * SE
- = p^ +- z * √ (p^ (1 - p^) / n)
- Hodnotu (z) si volím podle požadované míry spolehlivosti
  - Pro 90, 95 a 99% míru spolehlivosti z = 1.645, 1.96 a 2.58

Vliv úrovně spolehlivosti a velikosti vzorku na výběrovou chybu a velikost intervalu spolehlivosti

- p^ +- z * √ (p^ (1 - p^) / n)
- Čím vyšší úroveň spolehlivosti (z), tím vyšší výběrová chyba a tedy i širší interval spolehlivosti a méně přesný odhad
- Čím větší velikost vzorku, tím menší výběrová chyba a tedy i užší interval spolehlivosti a přesnější odhad

Interpretace intervalu spolehlivosti

- Když budu opakovaně (nekonečně dlouho) tahat náhodně vzorek a vždy vytvořím 95% interval spolehlivosti, pak v 95% případů budou tyto intervaly obsahovat populační proporci – v 95% případů bude můj odhad správný, v 5% nikoli.
- Tedy existuje 95% pravděpodobnost, že můj konkrétní interval obsahuje skutečný populační parametr

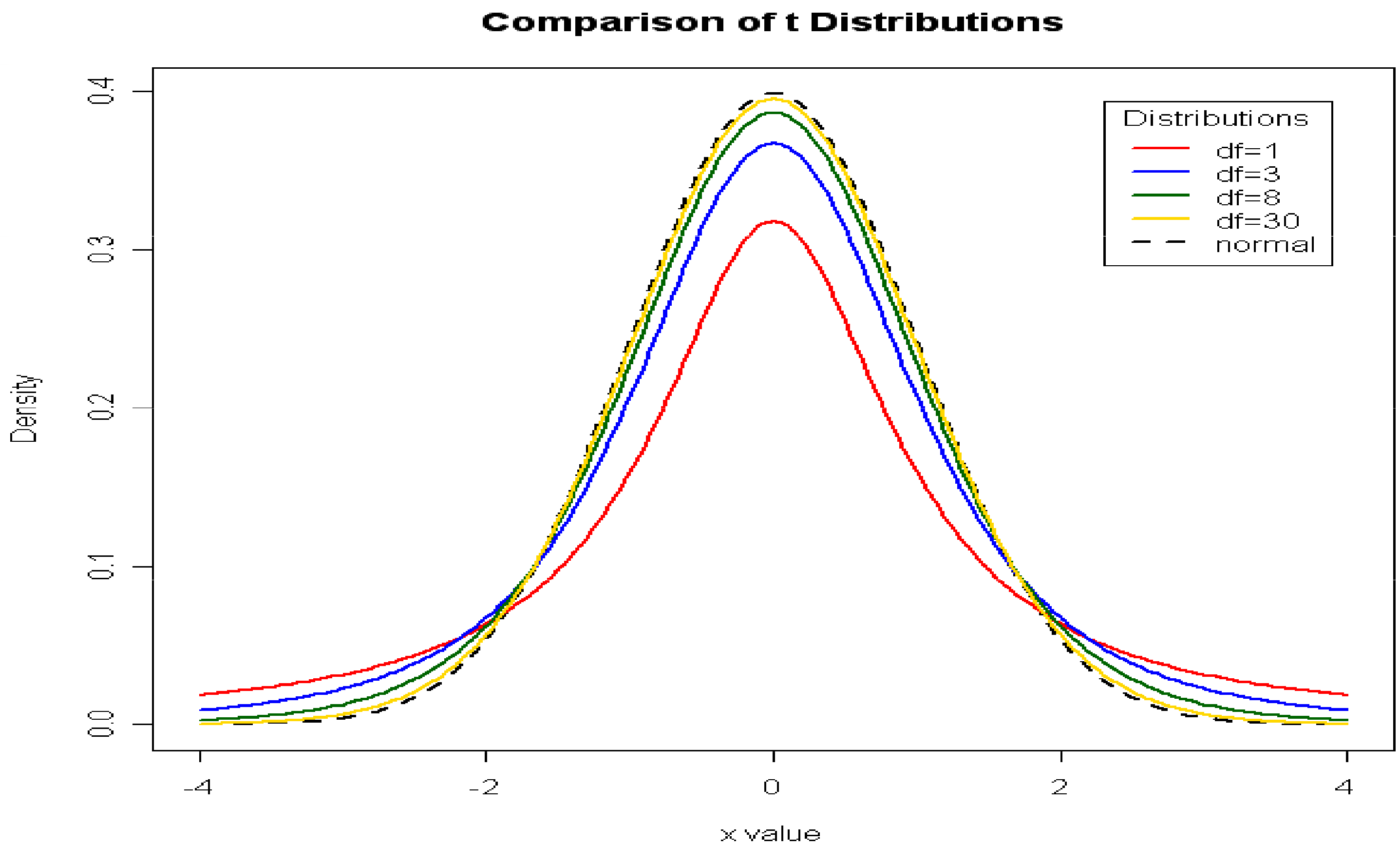
Interval spolehlivosti pro populační průměr

- Ekvivalentně k intervalu spolehlivosti pro populační proporci
- $= \bar{x} \pm t_{.025}\,(SE)$
  - kde $SE = s / \sqrt{n}$, kdy (s) představuje odchylku vzorku, která nahrazuje neznámou populační odchylku (σ), stejně jako ($\hat{p}$) nahrazuje (p) v případě SE proporce
  - t-statistika nahrazuje z-statistiku

T-distribuce a t-statistika

- Podobná normálnímu rozložení, širší konce
- Má zvonovitý tvar, symetrická kolem 0
- Rozložení pravděpodobností záleží na stupních volnosti (df) a tedy na velikosti vzorku (n) neboť
  - Df = n – 1
  - t-statistika tedy závisí na stupních volnosti
- Robustnější konce a větší variabilitu než normální rozložení

- Se zvětšující se velikostí vzorku se t-rozložení blíží normálnímu rozložení



T-skór a Z-skór

- Namísto populační odchylky (σ) používáme odchylku ze vzorku (s)

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

Zdroj: wiki

$$z = \frac{\bar{X} - \mu}{\sigma/\sqrt{N}}$$

Interval spolehlivosti pro populační průměr

- 95procentní interval spolehlivosti pro populační průměr je
  - $\bar{x}$ +- $t_{.025}$ (SE), kde SE = s / √n
- Hodnoty (t) volíme stejně jako hodnoty (z) podle požadované míry spolehlivosti (95 nebo 99%), přičemž s rostoucí velikostí vzorku se (t) hodnoty blíží (z) hodnotám (viz tabulka další snímek)

Hodnoty t-statistiky

Significance level = α

Degrees of Freedom	.005 (1-tail) .01 (2-tails)	.01 (1-tail) .02 (2-tails)	.025 (1-tail) .05 (2-tails)	.05 (1-tail) .10 (2-tails)	.10 (1-tail) .20 (2-tails)	.25 (1-tail) .50 (2-tails)
1	63.657	31.821	12.706	6.314	3.078	1.000
2	9.925	6.965	4.303	2.920	1.886	.816
3	5.841	4.541	3.182	2.353	1.638	.765
4	4.604	3.747	2.776	2.132	1.533	.741
5	4.032	3.365	2.571	2.015	1.476	.727
6	3.707	3.143	2.447	1.943	1.440	.718
7	3.500	2.998	2.365	1.895	1.415	.711
8	3.355	2.896	2.306	1.860	1.397	.706
9	3.250	2.821	2.262	1.833	1.383	.703
10	3.169	2.764	2.228	1.812	1.372	.700
11	3.106	2.718	2.201	1.796	1.363	.697
12	3.054	2.681	2.179	1.782	1.356	.696
13	3.012	2.650	2.160	1.771	1.350	.694
14	2.977	2.625	2.145	1.761	1.345	.692
15	2.947	2.602	2.132	1.753	1.341	.691
16	2.921	2.584	2.120	1.746	1.337	.690
17	2.898	2.567	2.110	1.740	1.333	.689
18	2.878	2.552	2.101	1.734	1.330	.688
19	2.861	2.540	2.093	1.729	1.328	.688
20	2.845	2.528	2.086	1.725	1.325	.687
21	2.831	2.518	2.080	1.721	1.323	.686
22	2.819	2.508	2.074	1.717	1.321	.686
23	2.807	2.500	2.069	1.714	1.320	.685
24	2.797	2.492	2.064	1.711	1.318	.685
25	2.878	2.485	2.060	1.708	1.316	.684
26	2.779	2.479	2.056	1.706	1.315	.684
27	2.771	2.473	2.052	1.703	1.314	.684
28	2.763	2.467	2.048	1.701	1.313	.683
29	2.756	2.462	2.045	1.699	1.311	.683
Large	2.575	2.327	1.960	1.645	1.282	.675

Př. Interval spolehlivosti pro populační průměr (použitím z-statistiky, neboť známe σ)

- Př. Výsledky IQ testu jsou aproximovány (blíží se) normálním rozložením o průměru $\mu = 100$ a $\sigma=16$. Třída 36 dětí dosáhne průměru 105 bodů.
  - Vypočítejte 95 % interval spolehlivosti.
- Postup:
  - vypočítám chybu průměru
    $\sigma\, m(\bar{x}) = \sigma / \sqrt{n} = 16 / \sqrt{36} = 2.67$
  - Vypočítám horní hranici intervalu 95:
    $1.96\, \sigma_{(\bar{x})} + \bar{x} = (1.96 * 2.67) + 105 = 105 + 5.23 = 110.23$
  - Vypočítám spodní hranici intervalu 95:
    $\bar{x} - 1.96\, \sigma_{(\bar{x})} = 105 - (1.96 * 2.67) = 105 – 5.23 = 99.\ 77$
- Interpretace: Když budu tahat vzorky nekonečně mnohokrát a pokaždé vytvořím 95 % interval, pak v 95% případů bude tento interval spolehlivosti obsahovat skutečný průměr
- Alternativní interpretace: Existuje 95 % pravděpodobnost že skutečný průměr leží v intervalu 99.77 až 110.23
- Analogicky vypočteme a interpretujeme 99 % interval spolehlivosti, namísto hodnoty 1.96 dosadíme hodnotu 2.58 (ověřte v tabulce z skoru)

Příklad proporce: zpět k parlamentním volbám ČR 2013

- Vztah mezi posledním průzkumem a skutečnými výsledky: cvičení

	předpověď	skutečnost
cssd	0,215	20,45
ano	0,175	18,65
kscm	0,12	14,91
top	0,105	11,99
ods	0,065	7,72
usvit	0,065	6,88
kdu	0,055	6,78
sz	0,045	3,19
spoz	0,025	1,51
ostatní	0,13	7,92
	1	100

Zdroj: ČT