Přednáška 7 – Testy hypotéz pro dva výběry

- parametrické testy s předpokladem normality
- neparametrické testy bez předpokladu normality

Párové výběry

- dvojice $\{X_1, Y_1\}, \{X_2, Y_2\}, \ldots$
- vždy stejný počet dat

Příklad:

- počet nakažených minulý týden a tento týden
- měření teploty ráno a večer
- známky skupiny studentů z matematiky a fyziky

Nepárové výběry

- hodnoty se měří nezávisle na sobě
- výběry mohou mít různý rozsah

Příklad:

- rychlosti vozidel na dálnici z Prahy a do Prahy
- počet nakažených z krajů v ČR a spolkových zemí v Německu
- známky dvou skupin studentů z matematiky

Tabulka testů hypotéz:

http://staff.utia.cas.cz/uglickich/pdfka/volbaTH.pdf

Parametrické testy pro dva výběry s předpokladem normality

Test o shodě středních hodnot při známých rozptylech (z-test 2)

Příklad:

- vliv nového léku na krevní tlak
- 978 pacientů – lék
- 936 pacientů – placebo

Měření tlaku – dva výběry:

$X_{\text{lék}} = [120/60 \;\; 143/92 \;\; 117/92 \;\; \ldots]$

$X_{\text{placebo}} = [137/86 \;\; 112/82 \;\; \ldots]$

Tvrzení: krevní tlak po použití léku je v průměru stejný jako při použití placebo (lék nemá vliv)

Jak správně zvolit test:

- dva výběry?
- mají oba výběry normální rozdělení? – parametrické testy
- tlak je v průměru stejný – test o shodě dvou středních hodnot
- Pro velké výběry, výběrové rozptyly $\Rightarrow$ rozptyly souborů

$H_0 : \mu_{\text{lék}} = \mu_{\text{placebo}}$ – podle tvrzení

$H_A : \mu_{\text{lék}} \neq \mu_{\text{placebo}}$ – opačné tvrzení

oboustranný test

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

Pro jeden výběr: $H_0 : \mu_{\text{lék}} = \mu_0$, T - jenom $\bar{X}$, σ^2

Test o shodě středních hodnot při neznámých rozptylech pro dva nepárové výběry – nepárový t-test

Příklad:

Jazyková škola má dvouletý kurz angličtiny. Škola tvrdí, že 1.ročník má lepší výsledky.

Nepárové výběry: $n_1 \neq n_2$

- bodové hodnocení 1.ročníku

$X_1 = [78 \;\; 83 \;\; 91 \;\; 56 \;\; \ldots]$ – rozsah n_1

- bodové hodnocení 2.ročníku

$X_2 = [91 \;\; 63 \;\; 72 \;\; \ldots]$ – rozsah n_2

Nulová hypotéza:

$H_0 : \mu_1 = \mu_2$ (nebo $\mu_1 > \mu_2$) – podle tvrzení
bodové hodnocení studentů 1.ročníku je vyšší

Alternativní hypotéza – opačné tvrzení:

$H_A : \mu_1 < \mu_2$ – bodové hodnocení studentů 1.ročníku je nižší
levostranný test

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim \text{Studentovo rozdělení}$$

Pro jeden výběr:
$H_0 : \mu_1 = \mu_0$, T - jenom $\bar{X}$, s^2

Test o shodě středních hodnot při neznámých rozptylech pro dva párové výběry – párový t-test

Příklad:

Jazyková škola má dvouletý kurz angličtiny. Škola tvrdí, že se studenti ve 2.ročníku zhorší.

Párové výběry:

- hodnocení 1.ročníku

$X_1 = [88 \;\; 66 \;\; 99 \;\; 56 \;\; \ldots]$ – rozsah n

- hodnocení stejného ročníku za rok

$X_2 = [76 \;\; 78 \;\; 91 \;\; 62 \;\; \ldots]$ – rozsah n

Nulová hypotéza:

$H_0 : \mu_1 = \mu_2$ (nebo $\mu_1 > \mu_2$) – podle tvrzení
bodové hodnocení studentů v 1.ročníku je vyšší

Alternativní hypotéza – opačné tvrzení:

$H_A : \mu_1 < \mu_2$ – bodové hodnocení studentů v 1.ročníku je nižší
levostranný test

$T = \frac{\bar{X}_d}{s_d}\sqrt{n} \sim$ Studentovo rozdělení,

$\bar{X}_d = \frac{1}{n}\sum_{i=1}^{n}(X_{i(1)} - X_{i(2)})$ – průměr odchylek dvojic z výběrů

$s_d = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_{i(d)} - \bar{X}_d)^2}$ – směrodatná odchylka odchylek

Test o shodě dvou rozptylů (var-test-2)

Příklad:

Jazyková škola má dvouletý kurz angličtiny. Řeší se, jaký ročník poslat na jazykovou olympiádu. Škola tvrdí, že stabilita výsledků 1.ročníku je vyšší.

výkonnost, spolehlivost, stabilita, variabilita, proměnlivost

Výběry: $n_1 \neq n_2$

- bodové hodnocení 1.ročníku

$X_1 = [78 \;\; 83 \;\; 91 \;\; 56 \;\; \ldots]$ – rozsah n_1

- bodové hodnocení 2.ročníku

$X_2 = [91 \;\; 63 \;\; 72 \;\; \ldots]$ – rozsah n_2

Nulová hypotéza:

$H_0 : \sigma_1^2 = \sigma_2^2$ (nebo $\sigma_1^2 < \sigma_2^2$) – podle tvrzení (rozptyl je menší)
stabilita výsledků 1.ročníku je vyšší

Alternativní hypotéza – opačné tvrzení:

$H_A : \sigma_1^2 > \sigma_2^2$ – stabilita 1.ročníku je nižší (rozptyl je větší)
pravostranný test

Pro jeden výběr: $H_0 : \sigma^2 = \sigma_0^2$

$T = \frac{s_1^2}{s_2^2} \sim$ Fisherovo rozdělení

Mann-Whitneyův test (ranksum):
- dva nepárové výběry
- alespoň jeden je bez normality

Příklad:

Skupina pacientů byla léčena na koronavirus odlišnými metodami: nadějným lékem A nebo nadějným lékem B. Počty dní, za kolik se pacienti uzdravili jsou v tabulce. Je doba léčení stejná?

Doba léčení lékem A	14	17	23	7	11	...	...		
Doba léčení lékem B	28	10	16	4	19	9	14	6	...

Nulová hypotéza: pacienti se vyléčili za stejnou dobu

$H_0 : \tilde{X}_{0,5(A)} = \tilde{X}_{0,5(B)}$ – mediány výběrů jsou stejné

Alternativní hypotéza: pacienti se nevyléčili za stejnou dobu

$H_A : \tilde{X}_{0,5(A)} \neq \tilde{X}_{0,5(B)}$ – oboustranný test

$T = \min\{U_A, U_B\}$

$T_A,\ T_B$ – součty pořadí výběrů

$$U_A = n_A n_B + \frac{n_A(n_A+1)}{2} - T_A$$

$$U_B = n_A n_B + \frac{n_B(n_B+1)}{2} - T_B$$

Test mediánu (znaménkový test) (signtest)

- dva párové výběry
- alespoň jeden výběr je bez normality

Příklad:

Testujeme, zda má skupina 11 studentů stejný počet vyřešených příkladů ze statistiky na začátku a na konci semestru. Nepředpokládáme normalitu dat.

Počet příkladů na začátku semestru:
$X_1 = [1\ \ 3\ \ 4\ \ 0\ \ 1\ \ 2\ \ 3\ \ 5\ \ 4\ \ 2\ \ 3]$
Počet příkladů na konci semestru
$X_2 = [5\ \ 14\ \ 1\ \ 15\ \ 0\ \ 12\ \ 4\ \ 9\ \ 4\ \ 3\ \ 25]$

Nulová hypotéza: počty příkladů jsou stejné

$H_0 : \tilde{X}_{0,5(1)} = \tilde{X}_{0,5(2)}$ – mediány výběrů jsou stejné

Alternativní hypotéza: počty příkladů nejsou stejné

$H_A : \tilde{X}_{0,5(1)} \neq \tilde{X}_{0,5(2)}$ – oboustranný test

Výpočet statistiky:

$X_{i(1)}$	$X_{i(2)}$	$X_{i(1)} - X_{i(2)}$	znaménka
1	5	1-5 = - **4**	-
3	14	3-14 = - **11**	-
4	1	4-1 = **3**	+
0	15	0-15 = - **15**	-
...	...	...	...

- Počet znamének = nový počet dat
- Menší počet znamének = b

Pro jeden výběr:

$H_0 : \tilde{X}_{0,5} = \tilde{X}_{0,5(0)}$

v tabulce:

$X_i - \tilde{X}_{0,5}$

$$T = \frac{2b-n}{\sqrt{n}} \sim N(0,1)$$

Pro tento příklad:

- p-hodnota $= 0.1094 > 0.05$
- nezamítáme nulovou hypotézu, že studenti vyřeší stejný počet příkladů na začátku a na konci semestru
- pouze pro předběžnou analýzu
- je slabší – porovnání s Wilcoxonovým testem

Wilcoxonův test (signrank) dva párové výběry, alespoň jeden výběr je bez normality

Stejný příklad: stejne H_0 a H_A. Stejné použití pro jeden výběr

Výpočet statistiky s použitím pořadí:

$X_{i(1)}$	$X_{i(2)}$	$X_{i(1)} - X_{i(2)}$	$\lvert X_{i(1)} - X_{i(2)} \rvert$	s	r
1	5	1-5 = - **4**	**4**	- (0)	2
3	14	3-14 = - **11**	**11**	-(0)	3
4	1	4-1 = **3**	**3**	+(1)	1
0	15	0-15 = - **15**	**15**	- (0)	4
...	...	...	...	...	...

- s – znaménka, $+ = 1$, $- = 0$
- r – pořadí absolutních hodnot

$$W = \sum_{i=1}^{n} s_i r_i$$

Pro tento příklad:

- p-hodnota $= 0.0273 < 0.05$
- Zamítáme nulovou hypotézu, že studenti vyřeší stejný počet příkladů na začátku a na konci semestru
- na cvičení v Matlabu – Wilcoxonův test