Kapitola 7. : Pořadové testy o mediánech

7.1. Motivace

Při používání t-testů či analýzy rozptylu by měl být splněn předpoklad normality dat. Pro výběry větších rozsahů ($n \geq 30$) nemá mírné porušení normality závažný dopad na výsledky. Někdy se však setkáváme s výběry malých rozsahů, které pocházejí z výrazně nenormálních rozložení. Pro práci s nimi byly vytvořeny tzv. neparametrické testy, které nevyžadují předpoklad o konkrétním typu rozložení (např. normálním), stačí např. předpokládat, že distribuční funkce rozložení, z něhož náhodný výběr pochází, je spojitá.

Tyto neparametrické testy se rovněž používají v situacích, kdy zkoumaná data nemají intervalový či poměrový charakter, ale pouze ordinální charakter.

Ve srovnání s klasickými parametrickými testy jsou však neparametrické testy slabší, tzn., že nepravdivou hypotézu zamítají s menší pravděpodobností než testy parametrické.

V této kapitole se omezíme na ty neparametrické testy, které jsou založeny na pořadí a týkají se mediánů. Nazývají se pořadové testy.

7.2. Pořadí čísla v posloupnosti čísel

7.2.1. Pojem pořadí

Nechť $x_1, \dots, x_n$ je posloupnost reálných čísel.
a) Jsou-li čísla navzájem různá, pak pořadím R_i čísla x_i rozumíme počet těch čísel $x_1, \dots, x_n$, která jsou menší nebo rovna číslu x_i.
b) Vyskytují-li se mezi danými čísly skupinky stejných čísel, pak každé takové skupince přiřadíme průměrné pořadí.

7.2.2. Příklad

a) Jsou dána čísla 9, 4, 5, 7, 3, 1. Stanovte pořadí těchto čísel.
b) Jsou dána čísla 6, 7, 7, 9, 6, 10, 8, 6, 6, 9.

Řešení

ad a)

usp. čísla	1	3	4	5	7	9
pořadí	1	2	3	4	5	6

ad b)

usp. čísla	6	6	6	6	7	7	8	9	9	10
pořadí	1	2	3	4	5	6	7	8	9	10
prům. pořadí	2,25	2,25	2,25	2,25	5,5	5,5	7	8,5	8,5	10

7.3. Jednovýběrové pořadové testy

Jde o neparametrické obdoby jednovýběrového t-testu a párového t-testu.

7.3.1. Jednovýběrový Wilcoxonův test

Nechť $X_1, ..., X_n$ je náhodný výběr ze spojitého rozložení s hustotou $\varphi(x)$, která je symetrická kolem mediánu $x_{0,50}$, tj. $\varphi(x_{0,50} + x) = \varphi(x_{0,50} - x)$. Nechť c je reálná konstanta. Testujeme hypotézu H_0: $x_{0,50} = c$ proti oboustranné alternativě H_1: $x_{0,50} \neq c$ (resp. proti levostranné alternativě H_1: $x_{0,50} < c$ resp. proti pravostranné alternativě H_1: $x_{0,50} > c$).

Utvoříme rozdíly $Y_i = X_i - c$, $i = 1, ..., n$. (Jsou-li některé rozdíly nulové, pak za n bereme jen počet nenulových hodnot.)

Absolutní hodnoty $|Y_i|$ uspořádáme vzestupně podle velikosti a spočteme pořadí R_i.

Zavedeme statistiku $S_W^+ = \sum_{Y_i>0} R_i^+$, což je součet pořadí přes kladné hodnoty Y_i. Analogicky zavedeme statistiku $S_W^- = \sum_{Y_i<0} R_i^-$, což je součet pořadí přes záporné hodnoty Y_i.

Přitom platí, že součet $S_W^+ + S_W^- = n(n+1)/2$. Za platnosti H_0 statistika S_W^+ má střední hodnotu $E(S_W^+) = n(n+1)/4$ a rozptyl $D(S_W^+) = n(n+1)(2n+1)/24$.

H_0 zamítáme na hladině významnosti α, když testová statistika je menší nebo rovna tabelované kritické hodnotě. Testová statistika = $\min(S_W^+, S_W^-)$ pro oboustrannou alternativu, = S_W^+ pro levostrannou alternativu, = S_W^- pro pravostrannou alternativu.

Pro $n \geq 30$ lze využít asymptotické normality statistiky S_W^+. Platí-li H_0, pak $U_0 = \frac{S_W^+ - E(S_W^+)}{\sqrt{D(S_W^+)}} = \frac{S_W^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \approx N(0,1)$. Kritický obor pro oboustrannou alternativu má tvar: $W = \left(-\infty, -u_{1-\alpha/2}\right\rangle \cup \left\langle u_{1-\alpha/2}, \infty\right)$. (Analogicky pro jednostranné alternativy.) H_0 zamítáme na asymptotické hladině významnosti α, když $U_0 \in W$.

Wilcoxonův test se hodí jen pro výběr ze symetrického rozložení. Není-li tento předpoklad splněn, lze použít např. znaménkový test (viz doporučená literatura, str. 193).

7.3.2. Příklad

U 12 náhodně zemí bylo zjištěno procento populace starší 60 let: 4,9 6,0 6,9 17,6 4,5 12,3 5,7 5,3 9,6 13,5 15,7 7,7. Na hladině významnosti 0,05 testujte hypotézu, že medián procenta populace starší 60 let je 12 proti oboustranné alternativě.

Řešení

Vypočteme rozdíly pozorovaných hodnot od čísla 12: -7,1 -6,0 -5,1 5,6 -7,5 0,3 -6,3 -6,7 -2,4 1,5 3,7 -4,3. Absolutní hodnoty těchto rozdílů uspořádáme vzestupně podle velikosti. Kladné rozdíly přitom označíme tučně:

usp. $\lvert x_i - 12 \rvert$	**0,3**	**1,5**	2,4	**3,7**	-4,3	-5,1	**5,6**	-6	-6,3	-6,7	-7,1	-7,5
pořadí	**1**	**2**	3	**4**	5	6	**7**	8	9	10	11	12

$S_W^+ = 14$, $S_W^- = 64$, $n = 12$, $\alpha = 0,05$, tabelovaná kritická hodnota = 13, testová statistika = $\min(S_W^+, S_W^-) = \min(14,64) = 14$. Protože $14 > 13$, H_0 nezamítáme na hladině významnosti 0,05.

7.3.3. Párový Wilcoxonův test

Nechť $(X_1, Y_1), ..., (X_n Y_n)$ je náhodný výběr ze spojitého dvourozměrného rozložení. Testujeme H_0: $x_{0,50} - y_{,50} = c$ proti H_1: $x_{0,50} - y_{0,50} \neq c$ (resp. proti jednostranným alternativám). Utvoříme rozdíly $Z_i = X_i - Y_i$, $i = 1, ..., n$ a testujeme hypotézu o mediánu $z_{0,50}$, tj. H_0: $z_{0,50} = c$ proti H_1: $z_{0,50} \neq c$.

7.3.4. Příklad

K zjištění cenových rozdílů mezi určitými dvěma druhy zboží bylo náhodně vybráno 15 prodejen a byly zjištěny ceny zboží A a ceny zboží B: (11,10), (14,11), (11,9), (13,9), (11,9), (10,9), (12,10), (10,8), (12,11), (11,9), (13,10), (14,10), (14,12), (19,15), (14,12). Na hladině významnosti 0,05 je třeba testovat hypotézu, že medián cenových rozdílů činí 3 Kč.

Řešení:

Jedná se o párový test. Vypočteme rozdíly mezi cenou zboží A a cenou zboží B, čímž úlohu převedeme na jednovýběrový test. Výpočty uspořádáme do tabulky:

č. prodejny	cena zboží A	cena zboží B	rozdíl	|rozdíl-medián|	pořadí
1	11	10	1	2	12
2	14	11	3	0	-
3	11	9	2	1	5,5
4	13	9	4	1	**5,5**
5	11	9	2	1	5,5
6	10	9	1	2	12
7	12	10	2	1	5,5
8	10	8	2	1	5,5
9	12	11	1	2	12
10	11	9	2	1	5,5
11	13	10	3	0	-
12	14	10	4	1	**5,5**
13	14	12	2	1	5,5
14	19	15	4	1	**5,5**
15	14	12	2	1	5,5

Tučně jsou vytištěna pořadí pro kladné hodnoty rozdíl-medián.
$S_W^+ = 16{,}5$, $S_W^- = 74{,}5$, $n = 13$, $\alpha = 0{,}05$, tabelovaná kritická hodnota = 17, testová statistika = $\min(S_W^+, S_W^-) = \min(16{,}5; 74{,}5) = 16{,}5$. Protože $16{,}5 \le 17$, H_0 zamítáme na hladině významnosti 0,05.

7.4. Dvouvýběrové pořadové testy

Jedná se o neparametrickou obdobu dvouvýběrového t-testu.

7.4.1. Dvouvýběrový Wilcoxonův test

Nechť $X_1, ..., X_n$ a $Y_1, ..., Y_m$ jsou dva nezávislé náhodné výběry ze dvou spojitých rozložení, jejichž distribuční funkce se mohou lišit pouze posunutím. Označme $x_{0,50}$ medián prvního rozložení a $y_{0,50}$ medián druhého rozložení. Testujeme hypotézu, že distribuční funkce těchto rozložení jsou shodné neboli mediány jsou shodné proti alternativě, že jsou rozdílné.

Všech n + m hodnot $X_1, ..., X_n$ a $Y_1, ..., Y_m$ uspořádáme vzestupně podle velikosti. Zjistíme součet pořadí hodnot $X_1, ..., X_n$ a označíme ho T_1. Součet pořadí hodnot $Y_1, ..., Y_m$ označíme T_2. Vypočteme statistiky $U_1 = mn + n(n+1)/2 - T_1$, $U_2 = mn + m(m+1)/2 - T_2$.
Přitom platí $U_1 + U_2 = mn$. Pokud $\min(U_1,U_2) \le$ tabelovaná kritická hodnota (pro dané rozsahy výběrů m, n a dané α), pak nulovou hypotézu o totožnosti obou distribučních funkcí zamítáme na hladině významnosti α.

Pro velká n, m (prakticky n, m > 30) lze využít asymptotické normality statistiky U_1.
V případě platnosti H_0 má statistika $U_0 = \dfrac{U_1 - \frac{mn}{2}}{\sqrt{\frac{mn(m+n+1)}{12}}}$ asymptoticky rozložení N(0,1). Kritický obor pro oboustrannou alternativu má tvar: $W = \left(-\infty, -u_{1-\alpha/2}\right\rangle \cup \left\langle u_{1-\alpha/2}, \infty\right)$. (Analogicky pro jednostranné alternativy.) H_0 zamítáme na asymptotické hladině významnosti α, když $U_0 \in W$.

Dvouvýběrový Wilcoxonův test se používá v situacích, kdy distribuční funkce rozložení, z nichž dané dva nezávislé náhodné výběry pocházejí, se mohou lišit pouze posunutím.

7.4.2. Příklad

Bylo vybráno 10 polí stejné kvality. Na čtyřech z nich se zkoušel nový způsob hnojení, zbylých šest bylo ošetřeno starým způsobem. Pole byla oseta pšenicí a sledoval se její hektarový výnos. Je třeba zjistit, zda nový způsob hnojení má týž vliv na průměrné hektarové výnosy pšenice jako starý způsob hnojení.

hektarové výnosy při novém způsobu: 51 52 49 55

hektarové výnosy při starém způsobu: 45 54 48 44 53 50

Řešení

usp. hodnoty	44	45	48	**49**	50	**51**	**52**	53	54	**55**
pořadí x-ových hodnot				4		6	7			10
pořadí y-ových hodnot	1	2	3		5			8	9	

$T_1 = 4 + 6 + 7 + 10 = 27$, $T_2 = 1 + 2 + 3 + 5 + 8 + 9 = 28$

$U_1 = 4.6 + 4.5/2 - 27 = 7$, $U_2 = 4.6 + 6.7/2 - 28 = 17$

Kritická hodnota pro $\alpha = 0{,}05$, min(4,6) = 4, max(4,6) = 6 je 2. Protože min(7,17) > 2, nemůžeme na hladině významnosti 0,05 zamítnout hypotézu, že nový způsob hnojení má na hektarové výnosy pšenice stejný vliv jako starý způsob.

Upozornění: Ve STATISTICE je dvouvýběrový Wilcoxonův test uveden pod názvem Mannův – Whitneyův test.

7.5. Kruskalův – Wallisův test a mediánový test (neparametrické obdoby analýzy rozptylu jednoduchého třídění)

7.5.1. Formulace problému

Nechť je dáno $r \geq 3$ nezávislých náhodných výběrů o rozsazích $n_1, \dots, n_r$. Předpokládáme, že tyto výběry pocházejí ze spojitých rozložení. Označme $n = n_1 + \dots + n_r$. Chceme testovat hypotézu, že všechny tyto výběry pocházejí z téhož rozložení.

7.5.2. Kruskalův – Wallisův test

Všech n hodnot seřadíme do rostoucí posloupnosti a určíme pořadí každé hodnoty. Označme T_j součet pořadí těch hodnot, které patří do j-tého výběru, $j = 1, \dots, r$ (kontrola: musí platit $T_1 + \dots + T_r = n(n+1)/2$). Testová statistika má tvar: $Q = \frac{12}{n(n+1)}\sum_{j=1}^{r}\frac{T_j^2}{n_j} - 3(n+1)$. Platí-li H_0, má statistika Q asymptoticky rozložení $\chi^2(r-1)$. H_0 tedy zamítneme na asymptotické hladině významnosti α, když $Q \geq \chi_{1-\alpha}^2(r-1)$.

7.5.3. Mediánový test

Testová statistika má tvar $Q_M = 4\sum_{j=1}^{r}\frac{P_j^2}{n_j} - n$, kde P_j je počet hodnot v j-tém výběru, které jsou větší nebo rovny mediánu vypočtenému ze všech n hodnot. Platí-li H_0, má statistika Q_M asymptoticky rozložení $\chi^2(r-1)$. H_0 tedy zamítneme na asymptotické hladině významnosti α, když $Q_M \geq \chi_{1-\alpha}^2(r-1)$.

7.5.4. Metody mnohonásobného porovnávání

Zamítneme-li H_0, zajímá nás, které dvojice náhodných výběrů se liší na zvolené hladině významnosti.

a) Neményiho metoda

Používá se v případě, že všechny výběry mají týž rozsah p. Je-li $|T_l - T_k| \geq$ tabelovaná kritická hodnota (pro dané p, r, α), pak na hladině významnosti α zamítáme hypotézu, že l-tý a k-tý výběr pocházejí z téhož rozložení.

b) Obecná metoda mnohonásobného porovnávání

Jestliže $\left|T_l - T_k\right| \geq \sqrt{\frac{1}{12}\left(\frac{1}{n_l}+\frac{1}{n_k}\right)n(n+1)h_{KW}(\alpha)}$, pak na hladině významnosti α zamítáme hypotézu, že l-tý a k-tý výběr pocházejí z téhož rozložení. Kritickou hodnotu $h_{KW}(\alpha)$ najdeme ve speciálních statistických tabulkách. Při větších rozsazích výběrů je možno ji nahradit kvantilem $\chi_{1-\alpha}{}^2(r-1)$.

7.5.5. Příklad

V roce 1980 byly získány tři nezávislé výběry obsahující údaje o průměrných ročních příjmech (v tisících dolarů) čtyř sociálních skupin ve třech různých oblastech USA.

jižní oblast: 6 10 15 29

pacifická oblast: 11 13 17 131

severovýchodní oblast: 7 14 28 25

Na hladině významnosti 0,05 testujte hypotézu, že příjmy v těchto oblastech se neliší. Zamítnete-li nulovou hypotézu, vyšetřete, které dvojice výběrů se od sebe liší na hladině významnosti 0,05.

Řešení

Kruskalův – Wallisův test

Usp.hodnoty	6	7	10	11	13	14	15	17	25	28	29	131
Pořadí 1.výběru	1		3				7				11	
Pořadí 2.výběru				4	5			8				12
Pořadí 3.výběru		2				6			9	10		

$T_1 = 22$, $T_2 = 29$, $T_3 = 27$, $Q = \frac{12}{12\cdot 13}\left(\frac{22^2}{4}+\frac{29^2}{4}+\frac{27^2}{4}\right)-3\cdot 13 = 0{,}5$, $\chi_{0,95}{}^2(2) = 5{,}991$.

Protože $Q < 5{,}991$, H_0 nezamítáme na asymptotické hladině významnosti 0,05. Rozdíly mezi průměrnými ročními příjmy v uvedených třech oblastech se neprokázaly.

Mediánový test

Medián všech 12 hodnot je 14,5. V 1. výběru leží nad mediánem 2 hodnoty, ve 2. výběru 2 hodnoty, ve 3. výběru 2 hodnoty. Testová statistika $Q_M = 4\left[\frac{1}{4}\left(2^2+2^2+2^2\right)\right]-12 = 0$, odpovídající kvantil $\chi_{0,95}{}^2(2) = 5{,}991$. Protože $Q_M < 5{,}991$, H_0 nezamítáme na asymptotické hladině významnosti 0,05.

Kontrolní otázky

1. V jakých situacích používáme neparametrické testy?
2. Jaká je nevýhoda neparametrických testů oproti testům parametrickým?
3. Jak vypočítáme pořadí čísla v dané posloupnostiu čísel?
4. Popište rozdíl mezi jednovýběrovým a párovým Wilcoxonovým testem.
5. Jaké podmínky musí být splněny pro dvouvýběrový Wilcoxonův test?
6. K čemu slouží Kruskalův-Wallisův test?
7. Jak provedeme mediánový test?
8. Které metody mnohonásobného porovnávání znáte?

Příklady

1. U 10 náhodně vybraných vzorků benzínu byly zjištěny následující hodnoty oktanového čísla: 98,2 96,8 96,3 99,8 96,9 98,6 95,6 97,1 97,7 98,0. Na hladině významnosti 0,05 testujte hypotézu, že medián oktanového čísla je 98 proti oboustranné alternativě.
Výsledek:
Použijeme jednovýběrový Wilcoxonův test. Testová statistika se realizuje hodnotou 12, tabelovaná kritická hodnota pro $\alpha = 0{,}05$ a $a = 9$ je 5. Protože $12 > 5$, H_0 nezamítáme na hladině významnosti 0,05.

2. Výrobce určitého výrobku se má rozhodnout mezi dvěma dodavateli polotovarů vyrábějících je různými technologiemi. Rozhodující je procentní obsah určité látky.
1. technologie: 1,52 1,57 1,71 1,34 1,68
2. technologie: 1,75 1,67 1,56 1,66 1,72 1,79 1,64 1,55
Na hladině významnosti 0,05 posuďte pomocí dvouvýběrového Wilcoxonova testu, zda je oprávněný předpoklad, že obě technologie poskytují stejné procento účinné látky.
Výsledek:
Testová statistika se realizuje hodnotou 12, tabelovaná kritická hodnota pro $\alpha = 0{,}05$, min(5,8) = 5, max(5,8) = 8 je 6. Protože min(28,12) > 2, nemůžeme na hladině významnosti 0,05 zamítnout hypotézu, že obě technologie poskytují stejné procento účinné látky.

3. Výrobce koláčů v prášku má 4 nové recepty a chce zjistit, zda se jejich kvalita liší. Upekl proto 5 koláčů z každého druhu a dal je porotě k ohodnocení.
recept A: 72 88 70 87 71,
recept B: 85 89 86 82 88,
recept C: 94 94 88 87 89,
recept D: 91 93 92 95 94.
Na hladině významnosti 0,05 testujte hypotézu, že recepty se neliší.
Výsledek:
Použijeme Kruskalův – Wallisův test. Všech 20 hodnot uspořádáme vzestupně podle velikosti a stanovíme součet pořadí pro recepty A, B, C, D: $T_1 = 23{,}5$, $T_2 = 37{,}5$, $T_3 = 66$, $T_4 = 83$. Testová statistika:
$Q = \frac{12}{20 \cdot 21}\left(\frac{23{,}5^2}{5} + \frac{37{,}5^2}{5} + \frac{66^2}{5} + \frac{83^2}{5}\right) - 3 \cdot 21 = 12{,}45$, $\chi_{0,95}^{\ 2}(3) = 7{,}81$. Protože $Q \geq 7{,}81$, H_0 zamítáme na asymptotické hladině významnosti 0,05. Neményiho metoda prokázala, že na hladině významnosti 0,05 se liší recepty A a D.

4. (S) U osmi osob byl změřen systolický krevní tlak před pokusem a po něm.

č. osoby	1	2	3	4	5	6	7	8
tlak před	130	185	162	136	147	181	128	139
tlak po	139	190	175	135	155	175	158	149

Na hladině významnosti 0,05 testujte hypotézu, že pokus neovlivní systolický krevní tlak
Výsledek:
Párový Wilcoxonův test poskytl p-hodnotu 0,04995, tedy H_0 zamítáme na hladině významnosti 0,05.

5. (S) Majitel obchodu chtěl zjistit, zda velikost nákupů (v dolarech) placených kreditními kartami Master/EuroCard a Visa jsou přibližně stejné. Náhodně vybral 7 nákupů placených Master/EuroCard: 42 77 46 73 78 33 37 a 9 placených Visou: 39 10 119 68 76 126 53 79 102. Lze na hladině významnosti 0,05 tvrdit, že mediány nákupů placených těmito dvěma typy karet se shodují?
Výsledek:
Dvouvýběrový Wilcoxonův test poskytl p-hodnotu 0,2523, H_0 tedy nezamítáme na hladině významnosti 0,05.

6. (S) Z produkce tří podniků vyrábějících televizory bylo vylosováno 10, 8 a 12 kusů. Byly získány následující výsledky zjišťování citlivosti těchto televizorů v mikrovoltech:
1.podnik: 420 560 600 490 550 570 340 480 510 460
2.podnik: 400 420 580 470 470 500 520 530
3.podnik: 450 700 630 590 420 590 610 540 740 690 540 670
Ověřte na hladině významnosti 0,05 hypotézu o shodě úrovně citlivosti televizorů v jednotlivých podnicích.
Výsledek:
K-W test poskytl testovou statistiku 3,2043, počet stupňů volnosti = 2, odpovídající p-hodnota = 0,0165, H_0 zamítáme na asymptotické hladině významnosti 0,05. Liší se výrobky podniků 2 a 3.