KAPITOLA 4.

Testy hypotézy o symetrii jednorozměrného a dvourozměrného rozdělení

Představme si, že chceme zjistit, zda nový druh ošetření má kladný efekt nebo zda je lepší než standardní ošetření. Abychom co nejvíce vyloučili vlivy, které nesouvisí s ošetřením, uspořádáme náhodný pokus tak, že pokusné jednotky rozdělíme do dvojic; přitom dvojice volíme tak, aby byly co nejvíce homogenní. Na jednoho člena dvojice aplikujeme zkoumané ošetření, zatímco druhý slouží ke kontrole (přiřazení uvnitř dvojice často provádíme náhodně). Jindy pokus uspořádáme tak, že téhož jedince pozorujeme např. před a po podání léku.

Dvojice pozorování $(X_1,Y_1),\dots,(X_N,Y_N)$, kde Y_i odpovídá ošetření a X_i slouží ke kontrole, můžeme považovat za náhodný výběr ze dvourozměrného rozdělení s distribuční funkcí $F(x,y)$. Na tuto funkci neklademe žádné zvláštní předpoklady kromě toho, že je spojitá.

Hypotéza, že ošetření nemá významný vliv, je ekvivalentní předpokladu, že distribuční funkce $F(x,y)$ je symetrická podle přímky $y = x$, tj. $H_1 : F(x,y) = F(y,x)$ pro všecka $x \in R^1$, $y \in R^1$. Alternativa kladného vlivu ošetření obecně znamená, že rozdělení vektoru (X,Y) je posunuto směrem k polorovině $y > x$.

4.1. Párový t-test

Jestliže lze předpokládat, že $F(x,y)$ je distribuční funkce dvourozměrného normálního rozdělení se středem (μ_1, μ_2) a s kovarianční maticí $\begin{pmatrix} \sigma^2 & \sigma^2\rho \\ \sigma^2\rho & \sigma^2 \end{pmatrix}$, $\sigma^2 > 0, -1 \leq \rho \leq 1$, pak problém testu významného vlivu ošetření se redukuje na test hypotézy $\mu_1 = \mu_2$ proti alternativě $\mu_2 > \mu_1$. Nejvhodnějším testem pro tuto situaci je párový t-test. Označíme-li $Z_i = Y_i - X_i$, $i=1,\dots,N$, pak test má kritický obor

$$\frac{\sqrt{N}\ \bar{Z}}{\left[\frac{1}{N-1}\sum_{i=1}^{N}(Z_i-\bar{Z})^2\right]^{1/2}} > t_{N-1}(\alpha) \tag{4.1}$$

kde $t_{N-1}(\alpha)$ je kritická hodnota t-rozdělení o $(N-1)$ stupních volnosti.

4.2. Testy H_1 založené na pořadích

Proveďme transformaci

$$Z_i = Y_i - X_i, \qquad W_i = X_i + Y_i, \qquad i=1,\dots,N.$$

Pak $(Z_1,W_1),\dots,(Z_N,W_N)$ tvoří opět výběr z dvourozměného rozdělení se spojitou distribuční funkcí. Za hypotézy H_1 je toto rozdělení symetrické vzhledem k ose w, zatímco za alternativy je posunuto ve směru kladné poloosy z. Problém je invariantní vzhledem k transformacím proměnných $w_1,\dots,w_N$ typu $w_i' = g(w_i)$, kde g je vzájemně jednoznačná funkce, která má nejvýše konečně mnoho bodů nespojitosti. Maximální invariantou vzhledem k těmto transformacím je vektor $Z_1,\dots,Z_N$. Invariantní testy tedy budou záviset jen na $Z_1,\dots Z_N$.

$Z_1,\dots,Z_N$ tvoří náhodný výběr z nějakého jednorozměrného rozdělení se spojitou distribuční funkcí D. Problém testu H_1 je pak ekvivalentní testu hypotézy

(4.2) $H_1' : D(z) + D(-z) = 1$ pro vš. $z \in R^1$,

že rozdělení D je symetrické kolem 0, proti alternativě

$$K_1' : D(z+\Delta) + D(-z+\Delta) = 1 \quad \text{pro vš.} \quad z \in R^1, \Delta > 0.$$

Alternativa odpovídá tomu, že rozdělení D je posunuto směrem ke kladným hodnotám z .

Rozdělení D je jednoznačně určeno trojicí (p, F_1, F_2), kde $p = P(Z<0)$, $F_1(z) = P(|Z|<z \mid Z<0)$ a $F_2(z) = P(Z<z \mid Z>0)$. Hypotéza H_1' se pak dá ekvivalentně vyjádřit ve tvaru

$$H_1'' : p = \tfrac{1}{2}, \quad F_2 = F_1$$

a chceme ji testovat proti alternativě

$$K_1'' : p < \tfrac{1}{2}, \quad F_2 \leq F_1.$$

Tento problém testu je invariantní vzhledem ke grupě $G : z_i' = g(z_i)$, $i=1,\dots,N$, kde g je spojitá, lichá a rostoucí funkce. Najdeme maximální invariantu vzhledem ke G : nechť $z_{i_1}, \dots, z_{i_m} < 0 < z_{j_1}, \dots, z_{j_n}$, kde $i_1 < \dots < i_m$ a $j_1 < \dots < j_n$; nechť $S_1,\dots,S_m$ jsou pořadí $|z_{i_1}|,\dots,|z_{i_m}|$ mezi $|z_1|,\dots,|z_N|$ a $R_1,\dots,R_n$ pořadí $|z_{j_1}|,\dots,|z_{j_n}|$ mezi $|z_1|,\dots,|z_N|$. Pak $(S_1,\dots,S_m, R_1,\dots,R_n)$ je maximální invarianta. Skutečně, nechť $z_1,\dots,z_N$ a $z_1',\dots,z_N'$ jsou 2 body takové, že $m'=m$, $n'=n$ a se stejnými S_i a R_j. Pak existuje spojitá rostoucí funkce g na kladné poloose taková, že $z_i' = g(z_i)$, $i=1,\dots,N$ a $g(0) = 0$. Jestliže dodefinujeme $g(-z) = -g(z)$, pak $g \in G$ a $z_i' = g(z_i)$, $i=1,\dots,N$.

Postačující statistikou pro vektor $(S_1,\dots,S_m; R_1,\dots,R_n)$ jsou vektory uspořádaných pořadí $S_1' < \dots < S_m'$ absolutních hodnot záporných pozorování a $R_1' < \dots < R_n'$ kladných pozorování mezi $|z_1|,\dots,|z_N|$; tyto vektory jsou dále jednoznačně určeny jedním z nich, např. $R_1' < \dots < R_n'$.

Invariantní testy hypotézy H_1 (nebo H_1') se pak redukují na testy, závislé jen na uspořádaných pořadích $R_1' < \dots < R_n'$.

Nechť ν je počet kladných pozorování mezi $Z_1,\dots,Z_N$. Pak ν je náhodná veličina s binomickým rozdělením; za platnosti H_1 je parametr binomického rozdělení roven $\frac{1}{2}$. Pro libovolné pevné n, $1 \leq n \leq N$ dostaneme s použitím věty 4 kapitoly 2

$$P(R_1' = r_1,\dots,R_\nu' = r_\nu, \nu = n) =$$

$$(4.3) \quad = P_{H_1}(R_1'=r_1,\dots,R_\nu'=r_\nu \mid \nu = n) \cdot P_{H_1}(\nu = n) =$$

$$= \frac{1}{\binom{N}{n}} \cdot \binom{N}{n} \left(\tfrac{1}{2}\right)^N = \left(\tfrac{1}{2}\right)^N$$

pro libovolnou z $\sum_{n=0}^{N} \binom{N}{n} = 2^N$ n-tic $(r_1,\dots,r_n)$ takových, že $1 \leq r_1 < \dots < r_n \leq N$. Kritický obor libovolného pořadového testu velikosti $\alpha = \frac{k}{2^N}$ obsahuje právě k takových bodů $(r_1,\dots,r_n)$. Mezi těmito testy neexistuje takový, který by byl stejnoměrně nejsilnější pro hypotézu H_1'' proti alternativě K_1''.

Obvykle se uvažují testy s kritickým oborem

(4.4) $$h(R_1 + \dots + h(R_\nu) > C$$

kde ν už není konstanta jako u H_0, ale náhodná veličina, která závisí na pozorováních; h je vhodná neklesající funkce. Z testů typu (4.4) probereme 2 nejběžnější. Budou se lišit volbou funkce h a budou vhodné proti různým alternativám.

4.2.1 Wilcoxonův test symetrie

Ve (4.4) položme $h(i) = i$, $i=1,\dots,N$. Výsledná testová statistika

(4.5) $$W_N^+ = \sum_{i=1}^{\nu} R_i$$

je rovna součtu pořadí kladných Z_i mezi $|Z_1|,\dots,|Z_N|$. Statistiku W_N^+ lze také vyjádřit v následujícím tvaru

(4.6) $$W_N^+ = \frac{1}{2}\sum_{i=1}^{N} \operatorname{sign} Z_i . R_i^+ + \frac{1}{4}N(N+1)$$

kde R_i^+ je pořadí $|Z_i|$ mezi $|Z_1|,\dots,|Z_N|$.
Wilcoxonův test zamítá H_1, jestliže W_N^+ překročí příslušnou kritickou hodnotu. Tabulky kritických hodnot lze nalézt např. v pracech:

F.Wilcoxon(1947):"Probability tables for individual comparisons by ranking methods". Biometrics 3,119-122.(N=6(1)20);
J.Hájek(1955):"Některá pořadová rozdělení a jejich použití". Čas.pro pěst.matematiky 80,17-31.
R.L.McCornack(1965):"Extended tables of the Wilcoxon matched pairs signed rank statistics".J.Amer.Statist.Assoc.60,864-871.

Pro velká N použijeme normální aproximace rozdělení W_N^+. Střední hodnotu a rozptyl W_N^+ vypočteme např. ze (4.6) s použitím následujícího lemmatu:

Lemma 4.1. Nechť náhodná veličina Z má spojitou distribuční funkci F takovou, že $F(z)+F(-z)=1$, $z \in R^1$. Pak Z a sign Z jsou nezávislé.

Důkaz. Zřejmě platí $P(\text{sign } Z=1) = P(\text{sign } Z=-1) = \frac{1}{2}$. Dále platí

$P(\text{sign } Z=1, |Z|<z) = P(0<Z<z) = P(-z<Z<0) = P(\text{sign } Z=-1, |Z|<z) = \frac{1}{2} P(|Z|<z)$.

Ze (4.6) a z lemmatu 4.1 dostaneme ($(Z_1, \ldots, Z_N)$ je náhodný výběr ze spojitého rozdělení, tedy rozdělení náhodného vektoru $R_1^+, \ldots, R_N^+$ je dáno větou 4 kapitoly 2):

$$E\, W_N^+ = \frac{1}{4} N(N+1),$$
$$\text{var } W_N^+ = \frac{1}{4} \sum_{i=1}^{N} E(R_i^+)^2 = \frac{1}{24} N(N+1)(2N+1). \tag{4.7}$$

Stejně jako dvouvýběrový test, i Wilcoxonův test symetrie je zvláště vhodný, mají-li Z_i rozdělení logistického typu.

4.2.2 Znaménkový test

Uvažujme obecnější situaci, kdy $Z_1, \ldots, Z_N$ jsou nezávislé veličiny, kde Z_i má spojitou distribuční funkci D_i, přičemž $D_1, \ldots, D_N$ nemusí být shodné. Tato situace nastane, jestliže různá srovnávání provádíme za různých experimentálních podmínek nebo různými metodami.

Za těchto předpokladů chceme testovat hypotézu

$$H_1''' : D_i(z) + D_i(-z) = 1, \quad i=1,\dots,N$$

proti alternativě, že rozdělení jsou posunutá směrem ke kladným hodnotám.

Problém je invariantní vzhledem ke všem transformacím typu $z_i' = f_i(z_i)$, $i=1,\dots,N$, kde f_i je spojitá, rostoucí a lichá funkce. Maximální invarianta je počet n kladných pozorování. Mezi všemi invariantními testy (tj. mezi testy závislými jen na počtu kladných pozorování) existuje stejnoměrně nejsilnější, který má tvar

$$(4.8) \qquad \Phi(n) = \begin{cases} 1 & \dots \; n > C \\ \gamma & \dots \; n = C \\ 0 & \dots \; n < C \end{cases}$$

kde C je určeno vztahem

$$\sum_{n>C} \binom{N}{n} \left(\tfrac{1}{2}\right)^N + \gamma \binom{N}{C} \left(\tfrac{1}{2}\right)^N = \alpha .$$

Test (4.8) zamítá, je-li počet kladných pozorování příliš velký; je to tzv. <u>znaménkový test</u>.

Nejprve musíme ověřit, že (4.8) je skutečně stejnoměrně nejsilnější mezi invariantními testy. Nechť $p_i = P(Z_i > 0) = 1-q_i$, $i=1,\dots,N$. Pak

$$P(\nu=n) = q_1 \dots q_N \sum{}' \frac{p_{i_1}}{q_{i_1}} \dots \frac{p_{i_n}}{q_{i_n}}$$

kde sčítáme přes všech $\binom{N}{n}$ kombinací $i_1 < \dots < i_n$ z $1,\dots,N$.

Nejsilnější test pro hypotézu $H_1''' : p_1 = \dots = p_N = \frac{1}{2}$ proti pevné alternativě $(p_1',\dots,p_N')$, $p_i' > \frac{1}{2}$, $i=1,\dots,N$ zamítá H_1''' pro taková n, pro která $\binom{N}{n}^{-1} P(\nu=n \mid p_1',\dots,p_n') > C_1$, neboli když platí

$$(4.9) \qquad \overset{*}{f}(n) = \frac{1}{\binom{N}{n}} \Sigma' \frac{p'_{i_1}}{q'_{i_1}} \cdots \frac{p'_{i_n}}{q'_{i_n}} > C_2 .$$

Dokážeme-li, že $\overset{*}{f}(n)$ je rostoucí v n, bude (4.9) ekvivalentní nerovnosti $n > C_3$ pro nějaké C_3, tedy kritickému oboru znaménkového testu. Označme $a_i = \frac{p'_i}{q'_i} > 1$, $i=1,\dots,N$. Pak

$$\overset{*}{f}(n+1) = \frac{1}{\binom{N}{n+1}} \Sigma' a_{i_1} \cdots a_{i_{n+1}} =$$

$$= \frac{1}{(n+1)\binom{N}{n+1}} \cdot \sum_{j \neq i_1,\dots,i_n} a_j \Sigma' a_{i_1} \cdots a_{i_n} >$$

$$> \frac{1}{(N-n)\binom{N}{n}} \sum_{j \neq i_1,\dots,i_n} \Sigma' a_{i_1} \cdots a_{i_n} =$$

$$= \frac{1}{\binom{N}{n}} \Sigma' a_{i_1} \cdots a_{i_n} = \overset{*}{f}(n),$$

což bylo třeba dokázat.

Znaménkový test je také testem typu (4.4); odpovídající funkce je $h(t) \equiv 1$. Testová statistika ν má za platnosti H_1 binomické rozdělení s parametrem $\frac{1}{2}$, odkud vyplývají kritické hodnoty, střední hodnota i rozptyl ν za platnosti H_1 :

$$E\,\nu = \frac{N}{2} , \quad \operatorname{var} \nu = \frac{N}{4} .$$

Pro velká N lze užít normální aproximace.

Znaménkového testu se pro jeho jednoduchost často používá i pro testování hypotézy H'_1 proti alternativě

$$K'_1 : D_1 = \dots = D_N \equiv D ;$$

$$D(z+\Delta) + D(-z+\Delta) = 1 \quad \text{pro vš. } z \in R^1, \ \Delta > 0.$$

Proti těmto alternativám už znaménkový test není stejnoměrně nejsilnějším invariantním testem. Naproti tomu dovedeme vypočítat sílu testu proti pevné alternativě z K_1' : za alternativy má $\mathcal{V}$ opět binomické rozdělení, tentokrát s parametrem

$$P(Z > 0) = 1 - D(0)$$

a tedy

$$\beta_{\Phi}(D) = \sum_{n > C} \binom{N}{n}(1-D(0))^n (D(0))^{N-n} + \gamma \binom{N}{C}(1-D(0))^C (D(0))^C ,$$

kde C je kritická hodnota testu.

Později uvidíme, že znaménkový test je lokálně nejsilnějším pořadovým testem pro H_1' proti K_1', je-li D distribuční funkce dvojitě exponenciálního typu s hustotou

$$d(x) = \frac{1}{2} e^{-|x-\Delta|} , \quad x \in R^1 .$$

<u>Poznámka.</u> K provedení znaménkového testu není třeba znát přesné hodnoty X_i, Y_i, $i=1,\dots,N$, ale stačí vědět, zdali je rozdíl $Y_i - X_i$ kladný nebo záporný. Proto je znaménkový test použitelný i v případě, kdy jsou k dispozici pouze kvalitativní srovnání jednotlivých ošetření, např. výroky typu "droga B utišuje bolesti lépe než droga A". Při podobných kvalitativních srovnáních ve skutečnosti nemáme k dispozici jiný než znaménkový test.

<u>4.2.3. Shody a nuly v pořadových testech symetrie</u>

Shoda dvou nebo více diferencí Z_i nemá vliv na hodnotu testové statistiky znaménkového testu; u Wilcoxonova testu řešíme tuto situaci podobně jako u dvouvýběrového Wilcoxonova testu.

U testů symetrie se však může vyskytnout další jev, který vede k nejednoznačnosti, a to jsou-li některé rozdíly Z_i rovny nule. Např. u kvalitativních srovnání je $Z_i=0$ tehdy, není-li subjekt schopen se rozhodnout, které z ošetření mělo příznivější vliv.

Uvažujme nejprve znaménkový test. Nechť $Z_1,\dots,Z_N$ jsou nezávislé a stejně rozdělené a nechť $p_+ = P(Z_1 > 0)$, $p_- = P(Z_1 < 0)$ a $p_0 = P(Z_1=0)$. Pak počet nul ν_0 mezi $Z_1,\dots,Z_N$ je náhodná veličina s binomickým rozdělením $b(p_0,N)$. V zásadě jsou 3 možnosti jak upravit znaménkový test: (1) uvažovat modifikovanou testovou statistiku ve tvaru $\nu + \frac{1}{2}\nu_0$; (2) každou nulovou hodnotu Z_i považovat s pravděpodobností $\frac{1}{2}$ za kladnou a s pravděpodobností $\frac{1}{2}$ za zápornou; (3) nulové hodnoty Z_i vynechat. Hemelrijk (1952) a Putter (1955) ukázali, že varianta (3) je z hlediska silofunkce nejlepší.

Podobně postupujeme u Wilcoxonova testu. Zde ovšem musíme kombinovat zpracování nulových hodnot se zpracováním shodných hodnot (viz 3.7).

4.3. Problémy a cvičení

(1) (R.F.Harell (1943):"Effect of Added Thiamine on Learning", Contrib.Educ.877,table 10). Na skupině 24 dětí byl vyšetřován vliv vitaminu B_1 na pokroky v učení. Děti byly rozděleny do 12 homogenních dvojic; náhodně zvolené dítě v každé dvojici dostávalo pravidelné dávky vitaminu B_1, druhé dítě dostávalo neutrální látku a sloužilo ke kontrole. Všechny děti prošly testem IQ před a po provedení pokusu, který trval

6 týdnů. Následující tabulka udává přírůstky IQ u všech dětí.

Dvojice	1	2	3	4	5	6	7	8	9	10	11	12
Ošetřené	14	18	2	4	-5	14	-3	-1	1	6	3	3
Kontrolní	8	26	-7	-1	2	9	0	-4	13	3	3	4

Pomocí Wilcoxonova a znaménkového testu ověřte, zda vitamin B_1 má prokazatelný vliv na pokrok v učení dětí.

(2) Lehmann(1975). Vyšetřovala se účinnost nového prostředku proti bolestem hlavy. 15 pacientů trpících bolestmi hlavy dostalo stejné množství tablet nového léku a standardního léku ve dvou lahvičkách označených náhodně A a B. Pacienti dostali pokyn, aby brali po jedné tabletě při každé bolesti hlavy, střídavě z lahviček A a B, až do využívání všech tablet a pak sdělili lékaři, který z prostředků považují za účinnější (lékař má zaznamenáno přidělení léků do lahviček A a B pro každého pacienta). 10 pacientů se vyjádřilo ve prospěch nového léku. Pomocí znaménkového testu ověřte, zda tento počet potvrzuje vyšší účinnost nového léku.

(3) Ukažte, že pro velká N lze sílu znaménkového testu aproximovat hodnotou

$$\Phi\left(\frac{\frac{1}{2}c_\alpha + \sqrt{N}\,(p-\frac{1}{2})}{\sqrt{p(1-p)}}\right)$$

kde $p=P(Z_i > 0)$, $i=1,\dots,N$; Φ je distribuční funkce standardního normálního rozdělení.

(4) Dokažte, že pro Wilcoxonovu statistiku platí

$$W_N^+ = \sum_{j=1}^{N} \sum_{i=1}^{j} u(Z_i + Z_j)$$

kde $u(t) = 1$ pro $t \geq 0$ a $u(t) = 0$ pro $t < 0$.
Pomocí tohoto vztahu dokažte, že za platnosti H_1 je rozdělení statistiky W_N^+ symetrické kolem hodnoty $\frac{1}{4} N(N+1)$.

KAPITOLA 5.

Testy hypotézy o shodnosti několika populací (ošetření)

5.1. Model jednoduchého třídění

Chceme porovnat p různých typů ošetření nebo populací na základě p nezávislých výběrů $X_{i1},\dots,X_{in_i}$, $i=1,\dots,p$, po jednom z každé populace. Přesněji řečeno, $X_{i1},\dots,X_{in_i}$ je náhodný výběr z rozdělení s distribuční funkcí F_i, o které předpokládáme, že je spojitá, $i=1,\dots,p$; výběry jsou vzájemně nezávislé. Označme $N = \sum_{i=1}^{p} n_i$ celkový počet pozorování.

Chceme testovat hypotézu

$$(5.1) \qquad H_2 : F_1(x) = F_2(x) = \dots = F_p(x), \quad x \in R^1,$$

která tvrdí, že není významný rozdíl mezi jednotlivými typy ošetření, buď proti obecné alternativě

$$(5.2) \qquad K_2 : F_i \neq F_j \text{ alespoň pro jednu dvojici } i,j$$

nebo proti speciálnější alternativě

$$(5.3) \qquad K_2' : F_i(x) = F(x-\Delta_i), \quad i=1,\dots,p$$
$$\Delta_i \neq \Delta_j \text{ alespoň pro jednu dvojici } i,j\,,$$

která tvrdí, že ošetření mají lineární vliv na hodnotu pozorované veličiny a alespoň dvě ošetření se významně liší.

5.1.1. F-test

Jestliže můžeme předpokládat, že F_i je distribuční funkce normálního rozdělení $N(\mu+\alpha_i, \sigma^2)$, $i=1,\ldots,p$, dostáváme obvyklý model analýzy rozptylu při jednoduchém třídění:

$$(5.4)\begin{cases} X_{ij} = \mu + \alpha_i + E_{ij}, & j=1,\ldots,n_i; \quad i=1,\ldots,p, \\ E_{ij} & \text{jsou nezávislé náhodné veličiny s rozdělením } N(0,\sigma^2). \end{cases}$$

Hypotéza H_2 v tomto případě nabývá tvaru

$$H_2' : \alpha_1 = \alpha_2 = \ldots = \alpha_p = 0.$$

Jak je známo, vhodným testem parametrické hypotézy H_2', nejsilnějším v určité třídě invariantních testů, je F-test s kritickým oborem

$$(5.5)\quad F = \frac{N-p}{p-1}\;\frac{\sum_{i=1}^{p} n_i(\bar{X}_{i.} - \bar{X}_{..})^2}{\sum_{i=1}^{p}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_{i.})^2} > C_\alpha,$$

kde kritickou hodnotu C_α najdeme v tabulkách F-rozdělení o $(p-1, N-p)$ stupních volnosti; přitom

$$\bar{X}_{i.} = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij} \quad \text{a} \quad \bar{X}_{..} = \frac{1}{N}\sum_{i=1}^{p}\sum_{j=1}^{n_i} X_{ij}.$$

5.1.2. Kruskal-Wallisův pořadový test

Jestliže nemůžeme předpokládat, že rozdělení pozorování jsou normální, použijeme testu založeného na pořadích. Situace je podobná jako u testování shodnosti dvou populací založeného na dvou nezávislých výběrech.

Nechť $R_{i1},\dots,R_{in_i}$ jsou pořadí $X_{i1},\dots,X_{in_i}$, stanovená vzhledem k vektoru všech pozorování

$$(X_{11},\dots,X_{n_1};\ X_{21},\dots,X_{2n_2};\dots;\ X_{p1},\dots,X_{pn_p}).$$

Nechť $R'_{i1}<\dots<R'_{in_i}$ jsou tatáž pořadí uspořádaná podle velikosti, $i=1,\dots,p$. Pak za platnosti H_2 pro libovolnou permutaci $(r_{11},\dots,r_{1n_1};\ \dots;\ r_{p1},\dots,r_{pn_p})$ čísel $1,\dots,N$ takovou, že $r_{i1}<\dots<r_{in_i}$; $i=1,\dots,p$, platí

$$P_{H_2}(R'_{11}=r_{11},\dots,R'_{1n_1}=r_{1n_1};\ \dots;\ R'_{p1}=r_{p1},\dots,$$
$$\dots,R'_{pn_p}=r_{pn_p}) = \frac{n_1!\dots n_p!}{N!}\ . \tag{5.6}$$

Označme $R_{i.} = \frac{1}{n_i}\sum_{j=1}^{n_i} R_{ij}$, $i=1,\dots,p$ a

$$R_{..} = \frac{1}{N}\sum_{i=1}^{p}\sum_{j=1}^{n_i} R_{ij} = \frac{N+1}{2}\ .$$

Dosadíme-li do testové statistiky (5.5) F-testu R_{ij} místo X_{ij}, $R_{i.}$ místo $\bar{X}_{i.}$ a $R_{..}$ místo $\bar{X}_{..}$, dostaneme

$$\frac{N-p}{p-1}\ \frac{\sum_{i=1}^{p} n_i(R_{i.}-R_{..})^2}{\sum_{i=1}^{p}\sum_{j=1}^{n_i}(R_{ij}-R_{i.})^2}\ ,$$

z čehož po úpravě a vynásobení vhodnou konstantou dostaneme testovou statistiku <u>Kruskal-Wallisova testu</u>

$$K = \frac{12}{N(N+1)}\sum_{i=1}^{p} n_i\left(R_{i.}-\frac{N+1}{2}\right)^2 =$$
$$= \frac{12}{N(N+1)}\sum_{i=1}^{p} n_i\,R_{i.}^2 - 3(N+1). \tag{5.7}$$

Jestliže je $p=2$, redukuje se statistika (5.7) na statistiku oboustranného Wilcoxonova testu. Test zamítá hypotézu H_2, jestliže platí

$$K \geq C_\alpha \,, \tag{5.8}$$

kde C_α je kritická hodnota. Kritické hodnoty lze stanovit z rozdělení pravděpodobností (5.6); jejich výpočet je však velmi pracný. Kritické hodnoty Kruskal-Wallisova testu tabelovali W.H.Kruskal a W.A.Wallis(1952): "Use of ranks in one-criterion variance analysis". J.Amer.Statist.Assoc.47,582-612 ($p=3$, $n_i=5$);
C.Kraft a C.vanEeden(1968):"A Nonparametric Introduction to Statistics".Macmillan,N.York.

Pro $p>3$ a $n_i \geq 5$ používáme přibližných kritických hodnot stanovených na základě asymptotického rozdělení. Dá se ukázat, že za platnosti H_2 a pro velká $n_1,\dots,n_p$ mají náhodné veličiny $\sqrt{12\, n_i}\,\frac{R_{i.}}{N+1}$, $i=1,\dots,p$ přibližně stejné sdružené rozdělení jako veličiny $\sqrt{n_i}(\bar{Z}_{i.}-\bar{Z}_{..})$, $i=1,\dots,p$, kde Z_{ij}, $i=1,\dots,p$; $j=1,\dots,n_i$ jsou nezávislé s rozdělením $N(0,1)$. Z toho plyne, že K má přibližně stejné rozdělení jako $\sum_{i=1}^{p} n_i(\bar{Z}_{i.}-\bar{Z}_{..})^2$, což je rozdělení χ^2 o $p-1$ stupních volnosti.

<u>Poznámka 1</u>. Jestliže platí H_2 a společné rozdělení veličin X_{ij} má konečný rozptyl, pak pro velká $n_1,\dots,n_p$ má i statistika

$$\frac{(N-p)\sum_{i=1}^{p} n_i(\bar{X}_{i.}-\bar{X}_{..})^2}{\sum_{i=1}^{p}\sum_{j=1}^{n_i}(X_{ij}-X_{i.})^2}$$

přibližné rozdělení χ^2_{p-1}. To znamená, že velikost F-testu (5.5) zůstane přibližně zachována, i když skutečné rozdělení dat není normální.

Poznámka 2. Jestliže se vyskytnou shodná pozorování, užijeme metody průměrných pořadí, podobně jako u Wilcoxonova testu.

5.1.3. Mediánový test

Nechť M je medián spojeného výběru $(X_{11},\dots,X_{1n_1},\dots,X_{p1},\dots,X_{pn_p})$. Předpokládejme pro jednoduchost, že N je sudé.

Označme A_j počet pozorování j-tého výběru, která jsou větší než M, $j=1,\dots,p$. Pak náhodný vektor $(A_1,\dots,A_p)$ má za platnosti H_2 rozdělení pravděpodobností

$$(5.9)\qquad P_{H_2}(A_1=a_1,\dots,A_p=a_p) = \frac{\binom{n_1}{a_1}\binom{n_2}{a_2}\cdots\binom{n_p}{a_p}}{\binom{N}{a}}$$

kde $a = \sum_{i=1}^{p} a_i = \frac{N}{2}$.

Testová statistika mediánového testu má tvar

$$(5.10)\qquad Q = 4\sum_{i=1}^{p}\frac{1}{n_i}\left(A_i-\frac{n_i}{2}\right)^2 = 4\sum_{i=1}^{p}\frac{A_i^2}{n_i} - N$$

a test zamítá H_2 při velkých hodnotách Q.

5.2. Model dvojného třídění (náhodné bloky)

Chceme-li porovnávat účinnost p různých ošetření a pozorovaná data vykazují velkou variabilitu způsobenou různými dalšími vlivy, je vhodné uspořádat experiment tak, že pozorované subjekty rozdělíme do n co nejvíce homogenních skupin, tzv. bloků, a srovnáváme účinnost ošetření pouze uvnitř bloků; jednotlivá ošetření přiřadíme jednotlivým členům bloku náhodně.

Budeme uvažovat nejjednodušší z těchto modelů, ve kterém pozorované subjekty rozdělíme do n bloků o p členech a každé ošetření aplikujeme v každém bloku právě jednou. Předpokládáme, že bloky jsou vzájemně nezávislé.

Formální popis modelu: máme $n.p$ pozorování, která uspořádáme do tabulky:

(5.11)

Blok \ Ošetření	1	2	...	p
1	X_{11}	X_{12}	...	X_{1p}
2	X_{21}	X_{22}	...	X_{2p}
⋮	⋮	⋮		
n	X_{n1}	X_{n2}	...	X_{np} .

Pozorování X_{ij} odpovídá i-tému bloku a j-tému ošetření. Předpokládáme, že náhodné veličiny X_{ij} jsou vzájemně nezávislé a X_{ij} má spojitou distribuční funkci F_{ij}, $i=1,\dots,n$; $j=1,\dots,p$. Naším úkolem je testovat hypotézu, že není významný rozdíl mezi ošetřeními, tedy

(5.12) $$H_3 : F_{i1}(x) = F_{i2}(x) = \dots = F_{ip}(x), \quad x \in R^1;$$
$i=1,\dots,n$, proti alternativě

(5.13) $$K_3 : F_{ij} \neq F_{ik}$$ alespoň pro jedno i a alespoň pro jednu dvojici j,k ;

nebo proti méně obecné alternativě

(5.14) $\quad K_3' : F_{ij}(x) = F_i(x - \Delta_j); \quad j=1,\dots,p; \; i=1,\dots,n$

$\Delta_j \neq \Delta_k$ alespoň pro jednu dvojici j,k .

5.2.1. F-test

Jestliže můžeme předpokládat, že veličiny X_{ij} vyhovují vztahům

(5.15) $\quad X_{ij} = \mu + \alpha_i + \beta_j + E_{ij}; \quad i=1,\dots,n; \quad j=1,\dots,p,$

kde E_{ij} jsou nezávislé náhodné veličiny s rozdělením $N(0,\sigma^2)$ a μ, α_i, β_j jsou neznámé parametry (μ-hlavní aditivní efekt, α_i -efekt i-tého bloku a β_j -efekt j-tého ošetření), pak hypotéza H_3 nabývá tvaru

(5.16) $\quad \beta_1 = \beta_2 = \dots = \beta_p .$

Kritický obor testu, vhodného pro hypotézu (5.16), má tvar

$$(5.17) \quad F = \frac{(n-1)n \sum_{j=1}^{p} (\bar{X}_{.j} - \bar{X}_{..})^2}{\sum_{j=1}^{p} \sum_{i=1}^{n} (X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..})^2} > C_\alpha \quad ,$$

kde C_α je kritická hodnota F-rozdělení o p-1 a (p-1)(n-1) stupních volnosti.

5.2.2. Friedmanův pořadový test

Uspořádejme pozorování v každém bloku podle velikosti a označme příslušná pořadí $R_{i1},\dots,R_{ip}$; $i=1,\dots,n$. Pořadí můžeme uspořádat do tabulky

Ošetření / Blok	1	2	...	p	Řádkové průměry
1	R_{11}	R_{12}	...	R_{1p}	$\frac{p+1}{2}$
2	R_{21}	R_{22}	...	R_{2p}	$\frac{p+1}{2}$
⋮	⋮	⋮		⋮	⋮
n	R_{n1}	R_{n2}	...	R_{np}	$\frac{p+1}{2}$
Sloupcové průměry	$R_{\cdot 1}$	$R_{\cdot 2}$	...	$R_{\cdot p}$	Celkový průměr $R_{\cdot\cdot} = \frac{p+1}{2}$

kde $R_{\cdot j} = \frac{1}{n} \sum_{i=1}^{n} R_{ij}$, $\quad R_{\cdot\cdot} = \frac{1}{np} \sum_{i=1}^{n} \sum_{j=1}^{p} R_{ij}$.

Friedmanův test je založen na statistice

$$Q = \frac{12\,n}{p(p+1)} \sum_{j=1}^{p} \left(R_{\cdot j} - \frac{p+1}{2}\right)^2 = \tag{5.18}$$

$$= \frac{12\,n}{p(p+1)} \sum_{j=1}^{p} R^2_{\cdot j} - 3n(p+1)$$

a zamítá H_3, jestliže $Q \geq C_\alpha$. Kritické hodnoty Friedmanova testu tabelovali

M.Friedman (1937): "The use of ranks to avoid the assumption of normality implicit in the analysis of variance". J.Amer. Statist.Assoc.32, 675-701;

D.B.Owen (1962):" Handbook of Statistical Tables", Adison-Wesley,Mass.;

C.Kraft a C.van Eeden (1968):"A Nonparametric Introduction to Statistics", Macmillan,N.York;

M.G.Kendall (1970):"Rank Correlation Methods", 4.vydání; Griffin,London.

Pro větší hodnoty p a n můžeme opět použít přibližných kritických hodnot. Jestliže $n \to \infty$, konverguje rozdělení statistiky (5.18) za platnosti H_3 k rozdělení χ^2 o $(p-1)$ stupních volnosti.

Jestliže $p=2$, redukuje se model náhodných bloků na model párových srovnávání, který jsme uvažovali v kapitole 4. Friedmanův test se pak redukuje na oboustranný znaménkový test

Podobně jako znaménkového testu pro porovnání 2 ošetření lze i Friedmanova testu pro porovnání p ošetření použít i tehdy, nejsou-li k dispozici přesná měření, ale jen uspořádání podle účinnosti. Podobně jako u znaménkového testu je výhodou i tohoto testu snadné provedení. Nevýhodou je, jak uvidíme, jeho nízká asymptotická vydatnost.

5.3. Problémy a cvičení

(1) Kruskal-Wallisův test při výskytu shodných pozorování. Nechť mezi pozorováními $(X_{11},\dots,X_{1n_1},\dots,X_{p1},\dots,X_{pn_p})$ je právě e hodnot různých, přičemž t_1 pozorování je rovno nejmenší z nich, atd., až t_e pozorování je rovno největší z nich. Nechť $R^*_{i1},\dots,R^*_{in_i}$ jsou průměrná pořadí i-tého výběru, $R^*_{i.} = \frac{1}{n_i} \sum_{j=1}^{n_i} R^*_{ij}$. Pak modifikovaná Kruskal-Wallisova statistika má tvar

$$K^* = \frac{12}{N(N+1)} \sum_{i=1}^{p} n_i (R^*_{i.} - \frac{N+1}{2})^2 \cdot \left[1 - \frac{\sum_{k=1}^{e}(t_k^3 - t_k)}{N^3 - N}\right].$$

Podmíněné rozdělení K^* při daných $t_1,\dots,t_e$ je přibližně χ^2_{p-1}.

(2) Srovnání 4 laboratoří (Mandel (1964):" The Statistical Analysis of Experimental Data", J.Wiley).

4 různé laboratoře měřily hladkost určitého typu papíru. Následující tabulka udává po 8 měřeních z každé laboratoře.

Laboratoř								
A	38.7	41.5	43.8	44.5	45.5	46.0	58.0	47.7
B	39.2	39.3	39.7	41.4	41.8	42.9	45.8	43.3
C	34.0	35.0	39.0	40.0	43.0	43.0	45.0	44.0
D	34.0	34.8	34.8	35.4	37.2	37.8	42.8	41.2

Pomocí Kruskal-Wallisova testu (modifikace uvedená ve cv.(1)) testujte hypotézu, že nejsou systematické rozdíly v práci jednotlivých laboratoří.

(3) (Beecher (1959):"Measurement of Subjective Responses", Oxford University Press), 7 pacientů trpících kašlem postupně obdrželo neutrální látku a 3 uklidňující prostředky. Následující tabulka udává počet zakašlání za den u jednotlivých pacientů při jednotlivých ošetřeních.

Pacient / Ošetření	1	2	3	4	5	6	7
Heroin,5mg	251	126	49	45	233	291	1385
Dextromethorphan 10 mg	207	180	123	85	232	208	1204
Codein,10mg	167	104	63	147	233	158	1611
Neutrální látka	301	120	186	100	250	183	1913

Pomocí Friedmanova testu rozhodněte, zda je významný rozdíl mezi jednotlivými typy ošetření.

KAPITOLA 6.

Testy hypotézy nezávislosti ve dvourozměrné populaci

Nechť $(X_1,Y_1),\dots,(X_n,Y_n)$ je náhodný výběr z dvourozměrného rozdělení pravděpodobností se spojitou distribuční funkcí $F(x,y)$. Chceme testovat hypotézu

(6.1) $H_5 : F(x,y) = F_1(x)F_2(y)$,
kde F_1 a F_2 jsou libovolné distribuční funkce, tj. že náhodné veličiny X a Y jsou nezávislé.
Hypotézu H_5 uvažujeme proti nejrůznějším alternativám závislosti X a Y. Nejčastější je alternativa kladné (záporné) závislosti X a Y. Nechť Y_x značí náhodnou veličinu, jejíž rozdělení je shodné s podmíněným rozdělením Y za podmínky $X = x$. Pak alternativa kladné závislosti znamená

(6.2) $K_5 : x < x' \Rightarrow Y_{x'}$ je stochasticky větší než Y_x.
Speciální případ alternativy kladné závislosti nastane, jestliže $F(x,y)$ má hustotu $f_\Delta(x,y)$ tvaru

(6.3) $K_5' : f_\Delta(x,y) = \int f_1(x-\Delta z)f_2(y-\Delta z)dM(z),\ \Delta>0$
kde $M(z)$ je libovolná nedegenerovaná distribuční funkce s konečným rozptylem a f_1,f_2 jsou libovolné hustoty. (6.3) znamená, že platí

$$X_i = X_i^0 + \Delta Z_i$$
$$Y_i = Y_i^0 + \Delta Z_i, \qquad i=1,\dots,n,$$

přičemž veličiny X_i^0, Y_i^0, Z_i, $i=1,\dots,n$ jsou nezávislé a jejich rozdělení nezávisí na i. Je-li $\Delta = 0$, jsou X_i a Y_i nezávislé, $i=1,\dots,n$.

Pořadové testy, se kterými se setkáme, jsou vhodné právě proti alternativám (6.3).

6.1. t-test

Předpokládejme, že $F(x,y)$ je distribuční funkce dvourozměrného normálního rozdělení, tj. má hustotu

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{1}{\sigma_1^2}(x-\mu_1)^2 + \frac{1}{\sigma_2^2}(y-\mu_2)^2 - \frac{2\rho}{\sigma_1\sigma_2}(x-\mu_1)(y-\mu_2)\right]\right\} \tag{6.4}$$

$x,y \in R^1$
$\sigma_1,\sigma_2 > 0$
$\mu_1,\mu_2 \in R^1$
$|\rho| < 1$,

kde μ_1, μ_2 jsou střední hodnoty a σ_1^2, σ_2^2 rozptyly X a Y a ρ je korelační koeficient X a Y. Za tohoto předpokladu lze hypotézu H_5 přepsat ve tvaru

$$H_5' : \rho = 0$$

a alternativa kladné závislosti má tvar $\rho > 0$.

Stejnoměrně nejsilnější nestranný kritický obor pro H_5' proti jednostranným alternativám má tvar

$$\sqrt{n-2}\,\frac{R_c}{\sqrt{1-R_c^2}} > t_\alpha(n-2), \tag{6.5}$$

kde

$$R_c = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\left[\sum_{i=1}^{n}(X_i-\bar{X})^2 \sum_{i=1}^{n}(Y_i-\bar{Y})^2\right]^{1/2}} \tag{6.6}$$

je výběrový korelační koeficient a $t_{\alpha}(n-2)$ je kritická hodnota rozdělení t o (n-2) stupních volnosti.

6.2. Permutační t-test

Chceme testovat hypotézu H_5 proti alternativě kladné závislosti; máme podezření, že $F(x,y)$ je distribuční funkce normálního rozdělení, ale nejsme si tím stoprocentně jisti. Podobnou situaci jsme uvažovali v § 3.3. Stejně jako tam bude vhodné hledat test, nejsilnější proti normálním alternativám, ale mezi testy, jejichž velikost nepřekročí předepsanou hladinu významnosti α nejen pro normální rozdělení, ale i pro všecka rozdělení absolutně spojitého typu vyhovující hypotéze nezávislosti.

Nechť $X^{(1)} < \ldots < X^{(n)}$ a $Y^{(1)} < \ldots < Y^{(n)}$ jsou pořádkové statistiky odpovídající výběrům $X_1,\ldots,X_n$ a $Y_1,\ldots,\ldots,Y_n$. Nechť $\mathcal{F}$ je systém všech dvourozměrných hustot tvaru

$$f(x,y) = f_1(x).f_2(y) \qquad x,y \in R^1,$$

kde f_1, f_2 jsou libovolné skoro všude spojité hustoty. Podle věty 1 § 3.3 test $\Phi = \Phi(x_1,\ldots,x_n;\, y_1,\ldots,y_n)$ splňuje rovnost

$$(6.7)\quad \int\!\!\ldots\!\!\int \Phi(x_1,\ldots,x_n;\, y_1,\ldots,y_n) f_1(x_1) \ldots f_1(x_n) f_2(y_1) \ldots \ldots f_2(y_n)\, d\underset{\sim}{x}\, d\underset{\sim}{y} = \alpha$$

pro vš. $f \in \mathcal{F}$ právě tehdy, jestliže

$$(6.8)\quad \frac{1}{(n!)^2} \sum \Phi(X^{(i_1)},\ldots,X^{(i_n)};\, Y^{(j_1)},\ldots,Y^{(j_n)}) = \alpha$$

platí skoro jistě vzhledem k $\mathcal{F}$, kde sčítáme přes všech $(n!)^2$ permutací $(i_1,\dots,i_n;\ j_1,\dots,\dots,j_n)$ bodu $(1,\dots,n;\ 1,\dots,n)$. Mezi všemi testy, vyhovujícími (6.8), hledáme stejnoměrně nejsilnější proti normálním alternativám, za kterých má vektor $(X_1,\dots,X_n;\ Y_1,\dots,\dots,Y_n)$ hustotu

$$(6.9)\quad \frac{1}{(2\pi\sigma_1\sigma_2\sqrt{1-\varrho^2})^n}\cdot\exp\left\{-\frac{1}{2(1-\varrho^2)}\left[\frac{1}{\sigma_1^2}\sum_{i=1}^{n}(x_i-\mu_1)^2+\right.\right.$$

$$\left.\left.+\frac{1}{\sigma_2^2}\sum_{i=1}^{n}(y_i-\mu_2)^2-\frac{2\varrho}{\sigma_1\sigma_2}\sum_{i=1}^{n}(x_i-\mu_1)(y_i-\mu_2)\right]\right\},\quad \varrho>0.$$

Protože $\sum_{i=1}^{n} x_i^2$, $\sum_{i=1}^{n} y_i^2$, $\sum_{i=1}^{n} x_i$, $\sum_{i=1}^{n} y_i$ jsou konstantní na množině všech $(n!)^2$ permutací bodu $(X^{(1)},\dots,X^{(n)};\ Y^{(1)},\dots,Y^{(n)})$, vyplývá odtud, že nejsilnější test zamítá H_5 pro M permutací $(i_1,\dots,i_n;\ j_1,\dots,j_n)$, kterým přísluší nejvyšší hodnoty výrazu

$$(6.10)\qquad \sum_{k=1}^{n} X^{(i_k)}\ Y^{(j_k)}$$

kde M je určeno tak, aby platilo $\frac{M}{(n!)^2}=\alpha$ (jestliže tato rovnost neplatí pro žádné M, je nutné bod na rozhraní randomizovat). Mezi hodnotami (6.10) je pouze $n!$ různých: můžeme tedy říci, že nejsilnější test zamítá pro M' permutací $j_1,\dots,j_n$ čísel $1,\dots,n$, vedoucí k největším hodnotám výrazu $\sum_{i=1}^{n} X^{(i)}\ Y^{(j_i)}$, kde $\frac{M'}{n!}=\alpha$.

Prakticky provedeme test takto: získáme data $(x_1,y_1),\dots,\dots,(x_n,y_n)$; stanovíme $x^{(1)}<\dots<x^{(n)}$, uspořádáme $n!$ hodnot $\sum_{i=1}^{n} x^{(i)}y^{(j_i)}$ podle velikosti a stanovíme kritický

obor (podmíněný jevem $X^{(1)} = x^{(1)},\dots,X^{(n)} = x^{(n)}$; $Y^{(1)} = y^{(1)},\dots, Y^{(n)} = y^{(n)}$), který obsahuje M' permutací odpovídající nejvyšším hodnotám $\sum_{i=1}^{n} x^{(i)} y^{(j_i)}$. Jestliže pak napozorovaná hodnota $\sum_{i=1}^{n} x_i y_i$ překročí hranici kritického oboru, zamítneme H_5 ve prospěch alternativy kladné závislosti.

Test je tedy podmíněný vektorem pořádkových statistik $T(\underset{\sim}{X},\underset{\sim}{Y}) = (X^{(1)},\dots,X^{(n)}; Y^{(1)},\dots,Y^{(n)})$ a kritický obor permutačního testu lze psát ve tvaru

$$\sum_{i=1}^{n} X_i Y_i > C(T(\underset{\sim}{X},\underset{\sim}{Y}))$$

nebo v ekvivalentním tvaru

$$R_c > C'(T(\underset{\sim}{X},\underset{\sim}{Y}));$$

tedy permutační test je verze standardního t-testu významnosti korelačního koeficientu, podmíněná vektorem pořádkových statistik. Dá se ukázat, že pokud $\sigma_1^2 > 0$, $\sigma_2^2 > 0$, $E|X|^3 < \infty$ a $E|Y|^3 < \infty$, jsou oba testy asymptoticky ekvivalentní při $n \to \infty$. Pro velká n, kdy provedení permutačního testu znamená velkou řadu výpočtů, tedy můžeme permutační test aproximovat standardním t-testem.

6.3. Pořadové testy nezávislosti

Nechť $(X_1,Y_1),\dots,(X_n,Y_n)$ je náhodný výběr z dvourozměrného rozdělení se spojitou distribuční funkcí $F(x,y)$. Označme $R_1,\dots,R_n$ pořadí $X_1,\dots,X_n$ a $S_1,\dots,S_n$ pořadí $Y_1,\dots,Y_n$. Jestliže platí hypotéza H_5, jsou vektory $(X_1,\dots,X_n)$ a $(Y_1,\dots,Y_n)$ nezávislé a tedy i vektory pořadí $(R_1,\ldots,R_n)$ a $(S_1,\ldots,S_n)$ jsou nezávislé; podle věty 4 kapitoly 2, má každý z těchto vektorů rovnoměrné rozdělení na množině $\mathcal{R}$ permutací $(1,\ldots,n)$.

6.3.1. Spearmanův korelační koeficient

Za hypotézy H_5 jsou vektory pořadí $(R_1,\ldots,R_n)$ a $(S_1,\ldots,S_n)$ nezávislé. Dosaďme do výrazu (6.6) pro výběrový korelační koeficient místo X_i a Y_i pořadí R_i, S_i, $i = 1,\ldots,n$. Dostaneme Spearmanův korelační koeficient

$$(6.11)\quad r_s = \frac{\frac{1}{n}\sum_{i=1}^{n} R_i S_i - \bar{R}\,\bar{S}}{\left[\frac{1}{n}\sum_{i=1}^{n}(R_i-\bar{R})^2 \cdot \frac{1}{n}\sum_{i=1}^{n}(S_i-\bar{S})^2\right]^{1/2}} .$$

Protože $\bar{R} = \bar{S} = \frac{n+1}{2}$ a $\frac{1}{n}\sum_{i=1}^{n}(R_i-\bar{R})^2 = \frac{1}{n}\sum_{i=1}^{n}(S_i-\bar{S})^2 = \frac{1}{n}\sum_{i=1}^{n} i^2 - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}$,

můžeme (6.11) vyjádřit také ve tvaru

$$(6.12)\qquad r_s = \frac{12}{n(n^2-1)}\sum_{i=1}^{n} R_i S_i - \frac{3(n+1)}{n-1} .$$

Spearmanův test zamítá H_5 ve prospěch alternativy kladné závislosti, jestliže platí

$$r_s > k_\alpha$$

nebo, což je ekvivalentní, jestliže

$$(6.13)\qquad \mathcal{S} = \sum_{i=1}^{n} R_i S_i > k'_\alpha .$$

r_s lze také vyjádřit ve tvaru

(6.14) $$r_s = 1 - \frac{6}{n^3-n} \sum_{i=1}^{n} (R_i - S_i)^2 ,$$

ze kterého vyplývá, že při testování můžeme též použít ekvivalentní statistiky

(6.15) $$\wp' = \sum_{i=1}^{n} (R_i - S_i)^2 ,$$

přičemž test zamítá pro malé hodnoty $\wp'$.

Kritické hodnoty Spearmanova testu lze nalézt např. v pracech G.J.Glasser a R.F.Winter (1961):"Critical values of the coefficient of rank correlation for testing the hypothesis of independence", Biometrika 48, 444-448 (n=4(1)30); D.B. Owen: Handbook of Statistical Tables (ruský překlad Moskva 1966).

Pro velká n stanovíme kritické hodnoty pomocí normální aproximace; střední hodnota a rozptyl $\wp$ za H_5 jsou

(6.16) $$E\wp = \frac{1}{4}\, n(n+1)^2$$

$$\operatorname{var} \wp = \frac{n^2(n+1)^2(n-1)}{144} .$$

Test (6.13) je lokálně nejsilnější pro H_5 proti alternativám (6.3), kde f_1 a f_2 jsou hustoty logistického typu (viz kapitola 7).

6.3.2. Kvadrantový test

Tento test je založen na statistice

$$(6.17) \qquad S^{*} = \frac{1}{4} \sum_{i=1}^{n} \left[\operatorname{sign}\left(R_i - \frac{n+1}{2}\right)+1\right] . \left[\operatorname{sign}\left(S_i - \frac{n+1}{2}\right)+ 1\right]$$

a zamítá při velkých hodnotách S^{*}. Jestliže n je sudé, je S^{*} rovno počtu dvojic (X_i, Y_i), pro které X_i je větší než medián $(X_1,\dots,X_n)$ a zároveň Y_i je větší než medián $(Y_1,\dots,Y_n)$. Statistika S^{*} má pak za platnosti H_5 hypergeometrické rozdělení

$$P(S^{*} = s) = \frac{\binom{m}{s}\binom{m}{m-s}}{\binom{n}{m}} \; ; \qquad s=0,\dots,m ,$$

kde $m = \frac{n}{2}$.

Odtud můžeme stanovit kritické hodnoty; také můžeme použít tabulek

G.J. Lieberman, D.B.Owen (1961): "Tables of the hypergeometrical probability distribution". Stanford Univ.Press;

pro velká n použijeme normální aproximace s parametry

$$ES^{*} = \frac{n}{4}$$

$$\operatorname{var} S^{*} = \begin{cases} \frac{1}{16} \frac{n^2}{n-1} & \dots \; n \text{ sudé} \\ \frac{1}{16} (n-1) & \dots \; n \text{ liché.} \end{cases}$$

6.3.3. Kendallův pořadový korelační koeficient

Další jednoduchý pořadový test nezávislosti je založen na

Kendallově korelačním koeficientu, který je dán vztahem

$$(6.18) \qquad \tau = \frac{1}{\binom{n}{2}} \sum_{i<j}\sum \operatorname{sign}(R_i - R_j)\operatorname{sign}(S_i - S_j).$$

Zřejmě $-1 \leq \tau \leq 1$. Hypotézu H_5 zamítáme ve prospěch alternativy kladné závislosti, jestliže $\tau > k_\alpha$, kde k_α je kritická hodnota; kritické hodnoty tabeloval M.G.Kendall (1948):"Rank correlation methods".Griffin & Co., London (3.vydání 1962).

Pro velká n použijeme přibližných kritických hodnot založených na asymptoticky normálním rozdělení τ s parametry

$$E\tau = 0, \qquad \operatorname{var}\tau = \frac{2(2n+5)}{9n(n-1)} .$$

6.4. Problémy a cvičení

(1) Následující tabulka udává výšku a obvod hlavy 16 chlapců ve věku 48 týdnů (Thompson (1951):"Data on the Growth of Children during the First Year of Life", Human Biol.23, 75--92).

Výška (mm)	773	730	717	796	754	776	720
Obvod hlavy	475	469	463	475	474	471	473

Výška	764	756	705	716	733	709	750
Obvod hlavy	482	464	482	482	450	461	474

Testujte hypotézu nezávislosti proti alternativě kladné závislosti.

(2) Nechť (X,Y) má dvourozměrné normální rozdělení s korelačním koeficientem $\rho > 0$. Pak X,Y jsou kladně závislé ve smyslu definice (6.2).

Návod : Podmíněné rozdělení Y při daném $X = x$ je normální se střední hodnotou

$$\mu_2 + \rho \frac{\sigma_2}{\sigma_1} (x - \mu_1)$$

a rozptylem $\sigma_2^2(1-\rho^2)$. Přičteme-li k veličině s tímto rozdělením kladnou hodnotu $\rho \frac{\sigma_2}{\sigma_1}(x'-x)$, dostaneme veličinu, jejíž rozdělení je rovno podmíněnému rozdělení Y při daném $X = x' > x$.

(3) Spearmanův test a Kendallův test jsou nestranné proti alternativám kladné závislosti.