O B S A H

UNIVERZITA KARLOVA V PRAZE

Fakulta matematicko-fyzikální

TESTY PARAMETRICKÝCH HYPOTÉZ

RNDr. Jana Jurečková, CSc.

Státní pedagogické nakladatelství
Praha

Kapitola 2

Stejnoměrně nejsilnější testy

2.1. Formulace problému

Nechť X je pozorovatelná náhodná veličina nebo náhodný vektor; víme, že rozdělení pravděpodobnosti X tvoří systém $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ pravděpodobnostních měr na $(\mathcal{X}, \mathcal{B})$. Nechť H a K jsou dva disjunktní podsystémy $\mathcal{P}$ takové, že $H \cup K = \mathcal{P}$. Řekneme, že vektor X splňuje hypotézu, jestliže rozdělení pravděpodobnosti X patří do H a že splňuje alternativu, jestliže jeho rozdělení patří do K. Pro hypotézu a alternativu použijeme rovněž symbolu H a K; tedy H i K označují jak výrok, tak systém rozdělení výrok splňujících.

Předpokládejme, že známe-li hodnotu θ, jednoznačně víme, zda $P_\theta \in H$ nebo $P_\theta \in K$. Označme $\Theta_H = \{\theta \in \Theta;\ P_\theta \in H\}$ a $\Theta_K = \{\theta \in \Theta : P_\theta \in K\}$ množiny bodů parametrického prostoru, odpovídající hypotéze nebo alternativě.

Stojíme před následujícím problémem: na základě pozorování x náhodného vektoru X máme rozhodnout, zda platí hypotéza H nebo alternativa K. Rozhodnutí o přijetí H označme d_0 a rozhodnutí o zamítnutí H označme d_1. Každé pravidlo, které každému bodu $x \in \mathcal{X}$ přiřadí právě jedno z rozhodnutí d_o a d_1, nazveme (nerandomizovaným) testem hypotézy H proti alternativě K. Takový test rozdělí výběrový prostor $\mathcal{X}$ na dvě komplementární části: kritický obor (obor zamítnutí) A_K a obor přijetí A_H. Jestliže $x \in A_K$, test hypotézu zamítá a v případě $x \in A_H$ ji nezamítá.

Jestliže provádíme nějaký test, může být naše rozhodnutí správné, nebo se můžeme dopustit jedné ze dvou druhů chyb:

(1) zamítneme H, i když H platí (chyba I.druhu)

(2) přijmeme H, i když H neplatí (chyba II.druhu).

Je žádoucí použít takový test, který má co nejnižší pravděpodobnosti obou druhů chyb. Pravděpodobnost chyby I.druhu, je-li $\theta \in \Theta_H$, je rovna

(2.1) $$P_\theta(X \in A_K) \; ; \; \theta \in \Theta_H$$

Číslo $\sup\limits_{\theta \in \Theta_H} P_\theta(X \in A_K)$ nazýváme velikostí testu (nebo velikostí kritického oboru A_K). Pravděpodobnost chyby II.druhu při $\theta \in \Theta_K$ je rovna

(2.2) $$P_\theta(X \in A_H) = 1 - P_\theta(X \in A_K), \quad \theta \in \Theta_K.$$

Hodnotu pravděpodobnosti

(2.3) $$\beta(\theta) = P_\theta(X \in A_K), \quad \theta \in \Theta_K$$

nazýváme sílou testu proti alternativě P_θ (nebo stručně θ), $\theta \in \Theta_K$. Funkci $\beta(\theta) : \Theta \longrightarrow (0,1)$ nazýváme silofunkcí příslušného testu.

Jestliže je pevně stanoven počet pozorování, nemůžeme současně minimalizovat pravděpodobnosti obou druhů chyb. Test, který by měl minimální pravděpodobnost chyby I.druhu, by odpovídal prázdnému kritickému oboru, a jeho chyba II.druhu by měla pravděpodobnost 1. Proto obvykle omezujeme pravděpodobnost chyby I.druhu pevně zvoleným malým číslem α , $0 < \alpha < 1$, tzv. <u>hladinou významnosti</u>, tj. omezíme se na testy splňující

$$(2.4) \qquad P_\theta(X \in A_K) \leq \alpha \quad \text{pro vš. } \theta \in \Theta_H,$$

a mezi všemi testy, splňujícími (1.4), hledáme takový, který má stejnoměrně nejmenší pravděpodobnost chyby II.druhu, neboli pro který platí

$$(2.5) \qquad P_\theta(X \in A_K) = \max \quad \text{pro vš. } \theta \in \Theta_K.$$

Volba hladiny významnosti je více méně libovolná; její hodnota znamená, jakou hranici pravděpodobnosti chyby I.druhu jsme ochotni tolerovat. Obvykle volíme jednu z hodnot 0,005; 0,01; 0,05; výhodou takové konvence je, že pro tyto hodnoty jsou sestrojeny různé tabulky, potřebné k provedení testu.

Optimální kritický obor splňující (2.4) a (2.5) nemusí vždy existovat. Teorie testování a hledání optimálního testu se značně zjednoduší, rozšíříme-li množinu testů o tzv. <u>randomizované testy</u>. Randomizovaný test zamítá H při daném pozorování x s pravděpodobností $\Phi(x)$ a přijímá s pravděpodobností $1-\Phi(x)$, $0 \leq \Phi(x) \leq 1$, $x \in \mathcal{X}$. Každý test je tak charakterizován funkcí $\Phi : \mathcal{X} \rightarrow \langle 0,1 \rangle$, zvanou <u>testová</u> (nebo <u>kritická) funkce</u>, takovou, že $0 \leq \Phi \leq 1$. U nerandomizovaných testů $\Phi(x)$ nabývá jen hodnot 0 (pro $x \in A_H$)

a 1 (pro $x \in A_K$); $\bar{\Phi}(x)$ je pak indikátorem kritického oboru.

Množina všech testů se pak ztotožňuje s množinou všech funkcí $\{\bar{\Phi}(x) : 0 \leq \bar{\Phi} \leq 1\}$. Tato množina je zřejmě konvexní a kromě toho kompaktní vzhledem ke slabé konvergenci (viz např. Lehman [6]).

Hodnota testové funkce $\bar{\Phi}(x)$ je vlastně podmíněná pravděpodobnost zamítnutí H při daném X = x. Jestliže P_θ je skutečné rozdělení X, pak celková pravděpodobnost zamítnutí hypotézy testem $\bar{\Phi}$ (tj. silofunkce testu v bodě θ) je rovna

$$(2.6) \qquad \beta_{\bar{\Phi}}(\theta) = \int \bar{\Phi}(x)\,dP_\theta(x) = E_\theta\,\bar{\Phi}(X).$$

Optimální volba testu je pak ekvivalentní následující maximalizační úloze: mezi všemi testy $\bar{\Phi}$ splňujícími

$$(2.7) \qquad \beta_{\bar{\Phi}}(\theta) = E_\theta\,\bar{\Phi}(X) \leq \alpha \qquad \text{pro vš. } \theta \in \Theta_H$$

nalézt takový, pro který

$$(2.8) \qquad \beta_{\bar{\Phi}}(\theta) = E_\theta\,\bar{\Phi}(X) = \max \qquad \text{stejnoměrně pro } \theta \in \Theta_K.$$

Test, který splňuje (2.7) a (2.8), se nazývá <u>stejnoměrně nejsilnější test velikosti $\leq \alpha$</u> (stejnoměrně nejsilnější α-test, nebo SNT velikosti $\leq \alpha$).

Řešení úlohy (2.7)-(2.8) nemusí obecně existovat. Obecně test, který maximalizuje sílu proti pevné alternativě $\theta \in \Theta_K$, závisí na této alternativě, a není tedy nejsilnější stejnoměrně pro celou alternativu. Teorie testování statistických hypo-

Nestranný test: $\beta_{\bar{\Phi}}(\theta) \leq \alpha \quad \forall \theta \in \Theta_H$
$\geq \alpha \quad \forall \theta \in \Theta_K$

též se zabývá otázkou existence optimálních testů, jejich hledáním, a vhodným řešením problému v případě, že stejnoměrně nejsilnější test neexistuje.

2.2. Neyman-Pearsonovo fundamentální lemma

Systém rozdělení pravděpodobnosti nazveme jednoduchý, jestliže obsahuje právě jeden prvek. Systém, který není jednoduchý, nazveme složený.

Jestliže je hypotéza i alternativa jednoduchá, odpadá problém stejnoměrnosti v úloze (2.7)-(2.8). Úloha pak má řešení za velmi obecných podmínek a řešení má jednoduchý explicitní tvar.

Následující věta je základem teorie testování statistických hypotéz; má četné aplikace nejen při hledání optimálního testu, ale i při jiných optimalizačních úlohách.

VĚTA 2.1. (Neyman-Pearsonovo lemma). Nechť P_0 a P_1 jsou dvě rozdělení pravděpodobnosti na $(\mathfrak{X},\mathfrak{A})$, která mají hustoty p_0 a p_1 vzhledem k míře μ (může být i $\mu = P_0+P_1$).

(1) Pak pro lib. $\alpha\in(0,1)$ existuje test Φ hypotézy $H : \{P_0\}$ proti alternativě $K : \{P_1\}$, který splňuje

$$(2.9) \qquad \Phi(x) = \begin{cases} 1 & \text{když} \quad p_1(x) > k\, p_0(x) \\ 0 & \text{když} \quad p_1(x) < k\, p_0(x) \end{cases}$$

kde konstanta k a hodnoty $\Phi(x)$ na $\{x : p_1(x) = k.p_0(x)\}$ jsou určeny tak, že platí

$$(2.10) \qquad E_0\, \Phi(X) = \alpha \ .$$

Test Φ je nejsilnějším testem H proti K velikosti $\leq \alpha$.
(2) Je-li Φ nejsilnější test H proti K velikosti $\leq \alpha$, pak Φ vyhovuje vztahu (2.9) pro nějakou konstantu k s.v.$[\mu]$. Pokud neexistuje test velikosti $< \alpha$ a síly 1, pak Φ vyhovuje také (2.10).

Důkaz. (1) Uvažujme funkci $\alpha(c) = P_0\left\{p_1(X) > c\, p_0(X)\right\}$; tato funkce je nerostoucí, zprava spojitá a taková, že $\lim\limits_{c \to -0} \alpha(c) = 1$ a $\lim\limits_{c \to \infty} \alpha(c) = 0$. Pak k libovolnému $\alpha \in (0,1)$ existuje $k \geq 0$ tak, že

$$(2.11) \qquad \alpha(k-0) \geq \alpha \geq \alpha(k)$$

kde jsme označili $\alpha(k-0) = \lim\limits_{c \to k-} \alpha(c)$.

Uvažujme test

$$(2.12) \qquad \Phi(x) = \begin{cases} 1 & \text{když} \quad p_1(x) > k\, p_0(x) \\ \dfrac{\alpha - \alpha(k)}{\alpha(k-0) - \alpha(k)} & \text{když} \quad p_1(x) = k\, p_0(x) \\ 0 & \text{když} \quad p_1(x) < k\, p_0(x)\,. \end{cases}$$

Střední výraz na (2.12) není definován, pokud je k bodem spojitosti $\alpha(c)$; pak je však $P_0\left\{p_1(X) = k\, p_0(x)\right\} = 0$.
Velikost testu (2.12) je

$$(2.13) \qquad E_0 \Phi(X) = P_0\left\{\frac{p_1(X)}{p_0(X)} > k\right\} + \frac{\alpha - \alpha(k)}{\alpha(k-0) - \alpha(k)}\, P_0\left\{\frac{p_1(X)}{p_0(X)} = k\right\} = \alpha$$

a tedy test Φ splňuje (2.9) a (2.10).
Zbývá dokázat, že test Φ je nejsilnější. Nechť Φ^* je libovolný jiný test takový, že $E_0 \Phi^*(X) \leq \alpha$. Nechť $S^+ = \left\{x \in \mathfrak{X}: \Phi(x) - \Phi^*(x) > 0\right\}$ a $S^- = \left\{x \in \mathfrak{X}: \Phi(x) - \Phi^*(x) < 0\right\}$.

Jestliže $x \in S^+$, je $\Phi(x) > 0$, a tedy $p_1(x) \geq k\, p_0(x)$; podobně je $p_1(x) \leq k\, p_0(x)$ pro $x \in S^-$; odtud plyne

$$(2.14) \qquad \int (\Phi - \Phi^*)(p_1 - k\, p_0)\, d\mu = \int_{S^+ \cup S^-} (\Phi - \Phi^*)(p_1 - k p_0)\, d\mu \geq 0$$

a tedy

$$\int (\Phi - \Phi^*) p_1\, d\mu \geq k \int (\Phi - \Phi^*)\, p_0\, d\mu \geq 0$$

a test Φ^* je nejvýše stejně silný jako Φ.

(2) Nechť Φ^* je nejsilnější test H proti K velikosti $\leq \alpha$ a nechť Φ vyhovuje (2.9) a (2.10). Nechť S^+ a S^- jsou definovány stejně jako v důkazu části (1) a nechť $S = (S^+ \cup S^-) \cap \{x : p_1(x) \neq k\, p_0(x)\}$.

Předpokládejme $\mu(S) > 0$. Protože je $(\Phi(x) - \Phi^*(x)) \cdot (p_1(x) - k\, p_0(x)) > 0$ pro $x \in S$, platí

$$\int_{S^+ \cup S^-} (\Phi - \Phi^*)(p_1 - k\, p_0)\, d\mu = \int_S (\Phi - \Phi^*)(p_1 - k\, p_0)\, d\mu > 0,$$

a tedy

$$E_1\left[\Phi(x) - \Phi^*(x)\right] > k\, E_0\left[\Phi(x) - \Phi^*(x)\right] \geq 0$$

a tedy Φ je silnější než Φ^* proti alternativě K. To je však spor, a tedy $\mu(S) = 0$, což bylo třeba dokázat.

Kdyby velikost Φ^* byla $< \alpha$ a síla < 1, přidáním dalších bodů (nebo části bodů) do kritického oboru bychom zvyšovali sílu Φ^* tak dlouho, dokud by buď velikost nedosáhla α nebo síla nedosáhla 1. Odtud plyne, že platí buď $E_0 \Phi^*(x) = \alpha$ nebo $E_1 \Phi^*(x) = 1$. Q.E.D.

Důsledek. Nechť Φ je nejsilnější test $H : \{P_0\}$ proti $K : \{P_1\}$ velikosti $\leq \alpha$, $0 < \alpha < 1$. Pak platí buď $\beta =$

$= E_1 \overline{\Phi}(X) > \alpha$ nebo $P_0 = P_1$.

Důkaz: Test $\overline{\Phi}_1(x) \equiv \alpha$ má velikost α i sílu α; test $\overline{\Phi}$ je alespoň stejně silný jako $\overline{\Phi}_1$, a tedy $\beta = E_1 \overline{\Phi}(X) \geq \alpha$. Jestliže je $\alpha = \beta < 1$, je test $\overline{\Phi}_1(x) \equiv \alpha$ nejsilnější a podle věty 2.1, část (2), musí vyhovovat (2.9). Odtud plyne, že $p_0(x) = p_1(x)$ s.v. $[\mu]$, a tedy $P_0 = P_1$.

2.3. Testy jednostranné hypotézy proti jednostranné alternativě

Jestliže $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ je systém rozdělení závislý na reálném nebo vektorovém parametru a Θ je interval, má test jednoduché hypotézy proti jednoduché alternativě spíše teoretický význam. První krok k obecnějšímu modelu je případ systému závislého na jediném reálném parametru θ, ve kterém chceme testovat tzv. jednostrannou hypotézu $H_1 : \theta \leq \theta_0$ proti jednostranné alternativě $K_1 : \theta > \theta_0$. Ani v tomto jednoduchém případě nemusí obecně existovat stejnoměrně nejsilnější test. Ukážeme, že SN test existuje, pokud je $\mathcal{P}$ dominovaný systém, jehož věrohodnostní poměr je monotonní funkcí vhodné statistiky $T(x)$. Tuto vlastnost má např. expoenciální systém s jediným reálným parametrem.

Definice 2.1. Nechť $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ je systém dominovaný σ-konečnou mírou μ, kde Θ je podmnožina R^1. Řekneme, že $\mathcal{P}$ má monotonní poměr věrohodnosti, jestliže existuje reálná statistika $T(x)$ taková, že pro lib. $\theta < \theta'$ jsou rozdělení P_θ a $P_{\theta'}$ různá a poměr $p_{\theta'}(x)/p_\theta(x)$ jejich hustot je neklesajici funkci $T(x)$.

VĚTA 2.2. <u>Nechť systém rozdělení</u> $\mathcal{P} = \{P_\theta,\ \theta \in \Theta\}$ <u>náhodné veličiny nebo náhodného vektoru</u> X <u>má monotonní poměr věrohodnosti vzhledem ke statistice</u> $T(x)$.

(1) <u>Pak existuje stejnoměrně nejsilnější α-test</u> $H_1 : \theta \leq \theta_0$ <u>proti</u> $K_1 : \theta > \theta_0$, <u>který má tvar</u>

$$(2.15)\qquad \Phi(x) = \begin{cases} 1 & \text{když } T(x) > c \\ \gamma & \text{když } T(x) = c \\ 0 & \text{když } T(x) < c \end{cases}$$

<u>kde konstanty</u> c a γ <u>jsou určeny tak, aby platilo</u>

$$(2.16)\qquad E_{\theta_0} \Phi(X) = \alpha .$$

(2) <u>Silofunkce</u> $\beta(\theta) = E_\theta \Phi(X)$ <u>testu</u> Φ <u>je rostoucí na množině</u> $\{\theta : \beta(\theta) < 1\}$.

<u>Důkaz :</u> Uvažujme nejprve hypotézu $H_0 : \theta = \theta_0$ a nějakou jednoduchou alternativu $\theta_1 > \theta_0$. Podle věty 2.1 existuje nejsilnější test takový, že

$$\Phi(x) = \begin{cases} 1 & \text{když } \dfrac{p_{\theta_1}(x)}{p_{\theta_0}(x)} > c_1 \text{, neboli} & T(x) > c \\ \gamma & \text{když} & T(x) = c \\ 0 & \text{když} & T(x) < c \end{cases}$$

kde c a γ jsou určena tak, aby platilo $E_{\theta_0} \Phi(X) = \alpha$. Podle věty 2.1, část (1) dále platí, že test Φ je také nejsilnější pro hypotézu $P_{\theta'}$ proti alternativě $P_{\theta''}$ na hladině

$\alpha' = \beta'(\theta')$ pro všecka θ', θ'' ; $\theta' < \theta''$. Z důsledku věty 2.1 dále plyne, že silofunkce $\beta(\theta)$ testu Φ je rostoucí v θ . Z této monotonie dále plyne

$$(2.17) \qquad E_{\theta}\Phi(X) \leq \alpha \qquad \text{pro } \theta \leq \theta_0 .$$

Třída A testů splňujících (2.17) je částí třídy B testů splňujících $E_{\theta_0}\Phi(X) \leq \alpha$; proto test Φ , který maximalizuje $\beta(\theta_1)$ na třídě B, tím spíše maximalizuje $\beta(\theta_1)$ na třídě A. Tím je zatím dokázáno, že Φ je nejsilnějším α -testem hypotézy $H : \theta \leq \theta_0$ proti jednoduché alternativě θ_1. Protože však test Φ ve skutečnosti nezávisí na speciální alternativě θ_1, je stejnoměrně nejsilnější pro $\theta < \theta_0$. Q.E.D.

Příklad 1. Nechť partie N výrobků obsahuje M zmetků. Z partie vybereme náhodně n výrobků, které zkontrolujeme. Nechť X je počet zmetků zjištěných ve výběru. Chceme testovat hypotézu $H : M \leq M_0$ proti $K : M > M_0$. X má hypergeometrické rozdělení

$$(2.18) \qquad P_M(X=x) = P_M(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} , \quad x \text{ celé, } a \leq x \leq b$$

$$= 0 \qquad \text{jinak}$$

kde $a = \max(M+n-N, 0)$ a $b = \min(M, n)$.

Interpretujeme $P_M(x)$ jako hustotu vzhledem k čítací míře, která přiřazuje libovolné podmnožině R^1 počet nezáporných celých čísel v ní obsažených. Pak platí

$$\frac{P_{M+1}(x)}{P_M(x)} = \frac{M+1}{N-M} \cdot \frac{N-M-n+x}{M+1-x}$$

a systém má monotonní poměr věrohodnosti vzhledem k $T(x)=x$. Stejnoměrně nejsilnější test H proti K zamítá H, když $x > c$.

Poznámka. Uvažujme duální problém $H: \theta \geq \theta_0$ proti $K: \theta < \theta_0$. Pak SN test dostaneme převrácením nerovnosti ve (2.15).

VĚTA 2.3. Nechť systém rozdělení náhodného vektoru X je jednoparametrický exponenciální systém s reálným parametrem θ a hustotami (vzhledem k nějaké míře μ) tvaru

$$(2.19) \qquad p_\theta(x) = \exp\left\{c(\theta)T(x) + A(\theta) + B(x)\right\}$$

kde $c(\theta)$ je ryze monotonní funkce. Pak existuje stejnoměrně nejsilnější test Φ hypotézy $H: \theta \leq \theta_0$ proti $K: \theta > \theta_0$. Jestliže je $c(\theta)$ rostoucí, má test tvar

$$(2.20) \qquad \Phi(x) = \begin{cases} 1 & \text{když} \quad T(x) < C \\ \gamma & \text{když} \quad T(x) = C \\ 0 & \text{když} \quad T(x) > C \end{cases}$$

kde C a γ jsou určena tak, že $E_{\theta_0}\Phi(X) = \alpha$.
Jestliže je $c(\theta)$ klesající, je nejsilnější test analogický (2.20); příslušné nerovnosti jsou převrácené.

Důkaz. Snadno vidíme, že systém (2.19) má monotonní poměr věrohodnosti vzhledem k $T(x)$. Věta pak plyne z věty 2.2.

Příklad 2. Systém binomických rozdělení $\left\{b(p,n),\ 0<p<1\right\}$ s pravděpodobnostmi

$$P_p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x=0,1,\dots,n$$

vyhovuje (2.19) s $\theta = p$, $T(x)=x$, $c(p)=\log\frac{p}{1-p}$. SN test hypotézy $H: p \leq p_0$ proti $K: p > p_0$ zamítá H, jestliže počet úspěchů X překročí vhodnou konstantu C.

Podobnou hypotézu můžeme také testovat na základě jinak uspořádaného pokusu. Opakovaně provádíme alternativní pokus s pravděpodobností úspěchu p, dokud nedosáhneme celkového počtu m úspěchů. Nechť Y_i je počet pokusů, které provedeme mezi $(i-1)$-ním a i-tým úspěchem. Pak $P(Y_i=y)=p(1-p)^y$, $y=0,1,\dots,$ a tedy sdružené rozdělení $Y_1,\dots,Y_m$ je

$$P_p(y_1,\dots,y_m) = p^m(1-p)^{\sum_{i=1}^{m} y_i}, \quad y_k=0,1,\dots\ ; \quad k=1,\dots,m.$$

To je exponenciální systém s $T(y) = \sum_{i=1}^{m} y_i$ a $c(p) = \log(1-p)$. Protože $c(p)$ je klesající v p, SN test pro $H: p \leq p_0$ proti $K: p > p_0$ zamítá, když T je příliš malé. $T(y)$ má negativně binomické rozdělení

$$P(t) = \binom{m+t-1}{m-1} p^m (1-p)^t, \qquad t=0,1,\dots \quad .$$

<u>Příklad 3.</u> Nechť $X_1,\dots,X_n$ je nezávislý náhodný výběr z Poissonova rozdělení s $EX_1 = \lambda\tau$. Pak sdružené rozdělení $X_1,\dots,X_n$ je

$$P(x_1,\dots,x_n) = \frac{\lambda^{\sum x_i}\,\tau^{\sum x_i}}{x_1!\dots x_n!}\, e^{-n\lambda\tau},$$

a tvoří tedy exponenciální systém, ve kterém je $T(x) = \sum_{i=1}^{n} x_i$, $c(\lambda) = \log\lambda$. Např. λ může označovat hustotu provozu telefonní centrály a X_i jsou počty signálů, které dojdou do centrály během n stejně dlouhých nepřekrývajících se časových intervalů délky τ. SN test hypotézy $H: \lambda \leq \lambda_0$ zamítá H při velkých hodnotách $T(x) = \sum_{i=1}^{n} x_i$. Statistika

$T(x)$ má rovněž Poissonovo rozdělení s parametrem $n\lambda\tau$.

Hypotézu můžeme také testovat na základě inverzně uspořádaného pokusu: sledujeme centrálu tak dlouho, dokud nedojde m signálů. Nechť $Y_1, Y_2, \dots, Y_m$ jsou délky intervalů do prvního signálu, od prvního do druhého signálu, atd. Pak $Y_1, \dots, Y_m$ jsou nezávislé a Y_i má exponenciální hustotu $\lambda e^{-\lambda y}$, $y \geq 0$, $i=1,\dots,m$.

Sdružená hustota $Y_1,\dots,Y_m$ je

$$p\,(y_1,\dots,y_m) = \lambda^m \exp\left\{-\lambda \sum_{i=1}^{m} y_i\right\}, \quad y_1,\dots,y_m \geq 0,$$

a tvoří exponenciální systém, kde $T(y_1,\dots,y_m) = \sum_{i=1}^{m} y_i$ a $c(\lambda) = -\lambda$. SN test pro $H: \lambda \leq \lambda_0$ zamítá H při malých hodnotách $T = \sum_{i=1}^{m} Y_i$. Protože $2\lambda Y_i$ má hustotu $\frac{1}{2} e^{-u/2}$ pro $u \geq 0$, což je χ^2-rozdělení o 2 stupních volnosti, má $2\lambda T$ rozdělení χ^2 o $2m$ stupních volnosti. Kritickou hodnotu testu tedy můžeme najít v tabulkách χ^2 rozdělení.

2.4. Zobecnění Neyman-Pearsonova lemmatu

Následující věta je zobecněním Neyman-Pearsonova lemmatu; podle této věty existuje test hypotézy, složené z konečně mnoha rozdělení, proti jednoduché alternativě. Věta rovněž udává explicitní tvar testu.

VĚTA 2.4. Nechť $P_1,\dots,P_{m+1}$ jsou rozdělení pravděpodobnosti na euklidovském výběrovém prostoru $(\mathcal{X}, \mathcal{B})$ s hustotami $f_1,\dots f_{m+1}$ vzhledem k míře μ. Nechť $\alpha_1,\dots,\alpha_m$ jsou dané konstanty, $0<\alpha_i<1$, $i=1,\dots,m$; nechť existuje alespoň jedna testová funkce vyhovující

$$(2.21)\qquad \int \Phi(x) f_i(x)\, d\mu = \alpha_i, \quad i=1,\dots,m.$$

Označme $\mathcal{E}$ množinu všech testových funkcí splňujících (2.21). Pak

(i) existuje $\Phi \in \mathcal{E}$, která maximalizuje integrál

$$(2.22)\qquad \int \Phi(x)\, f_{m+1}(x)\, d\mu .$$

(ii) Jestliže test $\Phi \in \mathcal{E}$ má tvar

$$(2.23)\qquad \Phi(x) = \begin{cases} 1 & \text{když} \quad f_{m+1}(x) > \sum_{i=1}^{m} k_i f_i(x) \\ 0 & \text{když} \quad f_{m+1}(x) < \sum_{i=1}^{m} k_i f_i(x) \end{cases}$$

kde $k_1,\dots,k_m$ jsou konstanty, pak Φ maximalizuje (2.22) na $\mathcal{E}$. Jestliže kromě toho platí $k_1,\dots,k_m \geq 0$, pak Φ maximalizuje (2.22) na množině všech testových funkcí splňujících

$$(2.24)\qquad \int \Phi(x)\, f_i(x)\, d\mu \leq \alpha_i, \quad i=1,\dots,m.$$

(iii) Množina $M \subset R^m$, daná výrazem

$$(2.25)\qquad M = \left\{ \left(\int \Phi(x) f_1(x)\, d\mu, \dots, \int \Phi(x) f_m(x)\, d\mu \right) : \Phi \in \mathcal{D} \right\}$$

kde $\mathcal{D}$ je množina všech testových funkcí, je konvexní a uzavřená. Jestliže $(\alpha_1,\dots,\alpha_m)$ je vnitřní bod množiny M, pak existuje optimální test Φ vyhovující vztahům (2.21), (2.23), který maximalizuje (2.22).

Důkaz: (i) Nechť $\{\Phi_n\}$ je posloupnost testových funkcí taková, že

$$\lim_{n\to\infty}\int \Phi_n(x) f_{m+1}(x)\,d\mu = \sup_{\Phi\in\mathcal{E}} \int \Phi(x) f_{m+1}(x)\,d\mu .$$

Protože množina testových funkcí je kompaktní vzhledem ke slabé konvergenci, existuje podposloupnost $\{\Phi_{n_k}\}$ a testová funkce Φ tak, že

$$\lim_{k\to\infty}\int \Phi_{n_k} f_j\,d\mu = \int \Phi f_j\,d\mu , \qquad j=1,\dots,m+1.$$

Z (2.21) plyne, že $\Phi\in\mathcal{E}$ a že Φ maximalizuje integrál (2.22) na $\mathcal{E}$.

(ii) Nechť test Φ je tvaru (2.23) a nechť Φ^* je libovolný jiný test. Nechť $S^+ = \{x\in\mathfrak{X} : \Phi - \Phi^* > 0\}$ a $S^- = \{x\in\mathfrak{X} : \Phi - \Phi^* < 0\}$. Pro $x\in S^+$ je $f_{m+1}(x) \geqq \sum_{i=1}^{m} k_i f_i(x)$ a pro $x\in S^-$ je $f_{m+1}(x) \leqq \sum_{i=1}^{m} k_i f_i(x)$; tedy

$$\int(\Phi-\Phi^*)(f_{m+1}-\sum_{i=1}^{m}k_i f_i)\,d\mu =$$

$$= \int_{S^+\cup S^-}(\Phi-\Phi^*)(f_{m+1}-\sum_{i=1}^{m}k_i f_i)\,d\mu \geqq 0 ,$$

odkud vyplývá

$$\int(\Phi-\Phi^*) f_{m+1}\,d\mu \geqq \sum_{i=1}^{m} k_i \int(\Phi-\Phi^*) f_i\,d\mu . \tag{2.26}$$

Jestliže $\Phi^*\in\mathcal{E}$, je pravá strana (2.26) rovna 0 pro libovolná $k_1,\dots,k_m$; při obecných $k_1,\dots,k_m$ tedy test Φ maximalizuje (2.22) na $\mathcal{E}$. Jestliže kromě toho $k_1,\dots,k_m \geqq 0$, je pravá strana (2.26) nezáporná pro libovolný test Φ^* splňující (2.24). Tím je dokázána část (ii).

Jestliže $k_i \geqq 0$, stačí $\int \Phi^* f_i\,d\mu \leqq \alpha_i$

(iii) Konvexita množiny M je zřejmá a uzavřenost M plyne z toho, že množina $\mathcal{D}$ všech testových funkcí je kompaktní vzhledem ke slabé konvergenci. Nechť $N \subset R^{m+1}$, kde

$$N = \left\{ \left(\int \Phi f_1 d\mu , \dots\dots , \int \Phi f_{m+1} d\mu \right) : \Phi \in \mathcal{D} \right\}.$$

Pak N je také konvexní a uzavřená a množina

$$\left\{ \alpha \in R^1 : (\alpha_1, \dots, \alpha_m, \alpha) \in N \right\}$$

je uzavřený interval $\langle \alpha^*, \alpha^{**} \rangle$. Uvažujme zvlášť případ $\alpha^* < \alpha^{**}$ a $\alpha^* = \alpha^{**}$.

a) Nechť $\alpha^* < \alpha^{**}$. Pak bod $(\alpha_1, \dots, \alpha_m, \alpha^{**})$ leží na hranici N, a tedy existuje nadrovina s rovnicí

$$\sum_{i=1}^{m+1} k_i u_i = \sum_{i=1}^{m} k_i \alpha_i + k_{m+1} \alpha^{**} ,$$

která prochází bodem $(\alpha_1, \dots, \alpha_m, \alpha^{**})$ a taková, že N leží právě v jednom z poloprostorů nadrovinou určených. Protože $(\alpha_1, \dots, \alpha_m)$ je vnitřní bod M, musí být $k_{m+1} \neq 0$. Bez újmy obecnosti můžeme položit $k_{m+1} = -1$. Pak pro každý bod množiny M platí

$$u_{m+1} - \sum_{i=1}^{m} k_i u_i \leq \alpha^{**} - \sum_{i=1}^{m} k_i \alpha_i . \tag{2.27}$$

Nechť Φ^{**} je test takový, že $\int \Phi^{**} f_i \, d\mu = \alpha_i$, $i = 1, \dots, m$, a $\int \Phi^{**} f \, d\mu = \alpha^{**}$. Pak pro libovolný jiný test vzhledem ke (2.27) platí

$$\int \Phi \left(f_{m+1} - \sum_{i=1}^{m} k_i f_i \right) d\mu \leq \int \Phi^{**} \left(f_{m+1} - \sum_{i=1}^{m} k_i f_i \right) d\mu , \tag{2.28}$$

a tedy Φ^{**} maximalizuje integrál na levé straně (2.28) přes všecky testové funkce. Tento integrál je však maximální, je-li Φ^{**} s.v. $[\mu]$ rovno

$$\Phi^{**}(x) = \begin{cases} 1 & \text{když} \quad f_{m+1}(x) > \sum_{i=1}^{m} k_i f_i(x) \\ 0 & \text{když} \quad f_{m+1}(x) < \sum_{i=1}^{m} k_i f_i(x) \end{cases}$$

což je shodné s (2.23).

b) Nechť $\alpha^* = \alpha^{**}$. Dokážeme, že v tomto případě množina N leží v nadrovině procházející počátkem, která není rovnoběžná s osou u_{m+1}.

Nechť $(u_1,\dots,u_m) \in M$, $(u_1,\dots,u_m) \neq (\alpha_1,\dots,\alpha_m)$. Předpokládejme, že existují $\underline{u}$, $\overline{u}$, $\underline{u} < \overline{u}$ tak, že $(u_1,\dots,u_m,\underline{u}) \in N$ a $(u_1,\dots,u_m,\overline{u}) \in N$. Protože $(\alpha_1,\dots,\alpha_m)$ je vnitřní bod M, existuje bod $(u_1',\dots,u_m')$ tak, že $(\alpha_1,\dots,\alpha_m)$ je vnitřním bodem úsečky $\langle(u_1,\dots,u_m),(u_1',\dots,u_m')\rangle$. Dále existuje u_{m+1}' tak, že $(u_1',\dots,u_m',u_{m+1}') \in N$; a tedy konvexní obal tří bodů $(u_1,\dots,u_m,\underline{u})$, $(u_1,\dots,u_m,\overline{u})$ a $(u_1',\dots,u_{m+1}')$ leží v N a obsahuje body $(\alpha_1,\dots,\alpha_m,\underline{\alpha})$ a $(\alpha_1,\dots,\alpha_m,\overline{\alpha})$, kde $\underline{\alpha} < \overline{\alpha}$, což je spor.

To znamená, že existují konstanty $k_1,\dots,k_m$ tak, že pro libovolný test Φ platí

$$\int \Phi f_{m+1}\, d\mu = \sum_{i=1}^{m} k_i \int \Phi f_i \, d\mu \; .$$

To ovšem může být splněno jen v triviálním případě

$$f_{m+1}(x) = \sum_{i=1}^{m} k_i f_i(x) \qquad \text{s.v.} [\mu]$$

a (2.23) je splněno triviálně. Q.E.D.

Důsledek. Nechť $f_1,\dots,f_m,f_{m+1}$ jsou hustoty vzhledem k míře μ takové, že neplatí $f_{m+1} = \sum_{i=1}^{m} k_i f_i$ s.v.$[\mu]$; nechť $0 < \alpha < 1$. Pak existuje test Φ tak, že $E_i\,\Phi(X) = \alpha$ $(i=1,\dots,m)$ a $E_{m+1}\,\Phi(X) > \alpha$.

Důkaz provedeme indukcí podle m. Pro $m=1$ plyne tvrzení v důsledku věty 2.1. Předpokládejme, že tvrzení platí pro lib. systém m rozdělení a nechť jsou dány lineárně nezávislé hustoty $f_1,\dots,f_{m+1}$. Pak jsou i $f_1,\dots,f_m$ lineárně nezávislé a pro každé $j=1,\dots,m$ podle indukčního předpokladu existují testy Φ_j a Φ_j' tak, že $E_i\,\Phi_j(X) = E_i\,\Phi_j'(X) = \alpha$ pro $i=1,2,\dots,j-1,j+1,\dots,m$ a $E_j\,\Phi_j(X) < \alpha < E_j\,\Phi_j'(X)$. Odtud vyplývá, že bod $(\alpha_1,\dots,\alpha) \in R^m$ je vnitřním bodem množiny M definované v části (iii) věty (2.4), a platí tvrzení (iii) věty 2.4. Protože test $\Phi(x) \equiv \alpha$ splňuje podmínku $E_i\,\Phi(X) = \alpha$, $i=1,\dots,m+1$, existuje alespoň jeden test, pro který je $E_{m+1}\Phi(X) \geq \alpha$. Kdyby tato nerovnost platila pouze jako rovnost pro všecky testy, byl by test $\Phi(x) \equiv \alpha$ optimální a vyhovoval by (2.23). Protože $0 < \alpha < 1$, dostáváme, že musí platit $f_{m+1} = \sum_{i=1}^{m} k_i f_i$ s.v.$[\mu]$, což je spor. Q.E.D.

2.5. Testy oboustranných hypotéz

Jestliže θ je reálný parametr, pak kromě jednostranné hypotézy uvažované ve 2.3 přicházejí v úvahu další přirozené hypotézy a alternativy:

$H_2 : \theta \leq \theta_1$ nebo $\theta \geq \theta_2$ $(\theta_1 < \theta_2)$

proti $K_2 : \theta_1 < \theta < \theta_2$;

$H_3 : \theta_1 \leq \theta \leq \theta_2$ proti $K_3 : \theta < \theta_1$ nebo $\theta > \theta_2$

$H_4 : \theta = \theta_0$ proti $K_4 : \theta \neq \theta_0$.

Jestliže uvažovaný systém rozdělení je jednoparametrický exponenciální systém, dovedeme nalézt stejnoměrně nejsilnější test hypotézy H_2 proti alternativě K_2; k tomu použijeme větu 2.4. Bohužel neexistují stejnoměrně nejsilnější testy H_3 proti K_3 ani H_4 proti K_4. Z věty 2.2 vyplývá, že testy , nesilnější proti alternativám $\theta > \theta_2$ mají rostoucí silofukci, zatímco testy nejsilnější proti alternativám $\theta < \theta_1$ mají klesající silofunkci; neexistují tedy testy stejnoměrně nejsilnější proti všem alternativám $\theta < \theta_1$ nebo $\theta > \theta_2$. Chceme-li nalézt vhodné testy hypotéz H_3 a H_4, musíme omezit třídu testů, mezi kterými hledáme optimum, např. na třídu nestranných testů. O tomto problému bude pojednávat následující kapitola.

V této kapitole zkonstruujeme stejnoměrně nejsilnější test H_2 proti K_2 v jednoparametrickém exponenciálním systému.

VĚTA 2.5. Nechť rozdělení pravděpodobnosti náhodného vektoru X tvoří jednoparametrický exponenciální systém s reálným parametrem θ a s hustotami (vzhledem k nějaké míře μ)

$$p_\theta(x) = \exp\left\{c(\theta)T(x) + A(\theta) + B(x)\right\} \tag{2.29}$$

kde $c(\theta)$ je ryze rostoucí funkce.

(i) Pak existuje stejnoměrně nejsilnější test hypotézy $H_2 : \theta \leq \theta_1$ nebo $\theta \geq \theta_2$ $(\theta_1 < \theta_2)$ proti alternativě $K_2 : \theta_1 < \theta < \theta_2$ daný vztahem

$$\Phi(x) = \begin{cases} 1 & \text{když} \quad c_1 < T(x) < c_2 \qquad (c_1 < c_2) \\ \gamma_i & \text{když} \quad T(x) = c_i, \quad i = 1,2 \\ 0 & \text{když} \quad T(x) < c_1 \quad \text{nebo} \quad T(x) > c_2 \end{cases} \tag{2.30}$$

kde c_1, c_2, γ_1 a γ_2 jsou určeny tak, aby platilo

$$E_{\theta_1}\Phi(X) = E_{\theta_2}\Phi(X) = \alpha, \qquad 0 < \alpha < 1. \tag{2.31}$$

(ii) Tento test minimalizuje $E_\theta \Phi(X)$ za podmínky (2.31) pro vš. $\theta < \theta_1$ a $\theta > \theta_2$.

Důkaz. Protože T je postačující statistika, stačí hledat nejsilnější test mezi testy, které jsou funkcemi $\psi(T)$ postačující statistiky, $0 \leq \psi(T) \leq 1$. Hustotu (2.29) můžeme psát jako hustotu vzhledem k míře $e^{B(x)} d\mu(x)$ ve tvaru $\exp\{c(\theta)T(x) + A(\theta)\}$. Přenosem integrace odtud dostáváme, že statistika $T : (\mathfrak{X}, \mathcal{B}) \rightarrow (\mathcal{T}, \mathcal{A})$ má hustotu

$$p_\theta(t) = \exp\{c(\theta)t + A(\theta)\} \tag{2.32}$$

vzhledem k nějaké míře ν na $\mathcal{A}$.

Zvolme pevně nějakou alternativu $\theta' \in (\theta_1, \theta_2)$ a hledejme test $\psi(T)$, který maximalizuje $E_{\theta'}\psi(T)$ mezi všemi testy vyhovujícími $E_{\theta_1}\psi(T) = E_{\theta_2}\psi(T) = \alpha$. Nechť M je množina všech bodů $(E_{\theta_1}\psi(T), E_{\theta_2}\psi(T))$, kde ψ probíhá všecky testové funkce. Podle tvrzení (iii) věty 2.4 musíme ověřit, zda (α,α) je vnitřní bod M. Podle důsledku věty 2.4 však existují u_1, u_2; $0 < u_1 < \alpha < u_2 < 1$, tak, že $(\alpha, u_1) \in M$ a $(\alpha, u_2) \in M$; kromě toho $(u,u) \in M$ pro vš. $u \in (0,1) \Longrightarrow (\alpha,\alpha)$ je skutečně vnitřním bodem M. Podle části (iii) věty 2.4 pak existují konstanty k_1, k_2 a test $\psi_0(t)$ tak, že $\Phi_0(x) = \psi_0(T(x))$ vyhovuje (2.31) a $\psi_0(t) = 1$ když

$$k_1 \exp\left\{c(\theta_1)t + A(\theta_1)\right\} + k_2 \exp\left\{c(\theta_2)t + A(\theta_2)\right\} < \\ < \exp\left\{c(\theta')t + A(\theta')\right\}$$

neboli

$$(2.33) \qquad a_1 e^{b_1 t} + a_2 e^{b_2 t} < 1 \qquad (b_1 < 0 < b_2)$$

a $\psi_0(t) = 0$ když $a_1 e^{b_1 t} + a_2 e^{b_2 t} > 1.$

Dokážeme sporem, že $a_1 > 0$ i $a_2 > 0$: kdyby $a_1, a_2 \leq 0$, test by pořád zamítal; kdyby $a_1 \leq 0 < a_2$ nebo $a_2 \leq 0 < a_1$, byla by levá strana (2.33) ryze monotonní v t a výsledný test by byl jednostranného typu uvažovaný ve větě (2.3), jehož silofunkce by však byla ryze monotonní a který by tedy nemohl vyhovovat podmínce (2.31). Tedy kritický obor má tvar (2.33), kde $a_1, a_2 > 0$ a $b_1 < 0 < b_2$, což vede k testu (2.30).

Test nezávisí na speciálně zvolené alternativě θ'; tím jsme zatím dokázali, že (2.30) je nejsilnější test pro hypotézu $H_2' : \theta = \theta_1$ nebo $\theta = \theta_2$ proti $K_2 : \theta_1 < \theta < \theta_2$.

Podle (ii) věty 2.4 je test ψ také nejsilnější pro H_2' mezi všemi testy vyhovujícími $E_{\theta_i}\psi(T) \leq \alpha$ $(i=1,2)$.

Abychom ověřili, že ψ je také nejsilnější pro celou hypotézu H, zbývá dokázat, že $E_\theta\psi(T) \leq \alpha$ pro $\theta \leq \theta_1$ a $\theta \geq \theta_2$. Z věty 2.4 však plyne, že test minimalizuje sílu $E_\theta\psi(T)$ v každém bodě θ, $\theta < \theta_1$ nebo $\theta > \theta_2$, což je tvrzení(ii). Srovnáním s testem $\psi'(t) \equiv \alpha$ pak dostáváme $E_\theta\psi(T) \leq \alpha$ pro $\theta \leq \theta_1$ a $\theta \geq \theta_2$. Tím je dokázáno (i). Q.E.D.

K tomu, abychom určili $c_1, \gamma, c_2, \gamma_2$, musíme ovšem stanovit rozdělení testové statistiky T. Ale ani pak není volba $c_1, c_2, \gamma_1, \gamma_2$ zcela průhledná.
Podmínky (2.31) můžeme pak přepsat ve tvaru

$$\beta(\theta_i) = \int_{c_1}^{c_2} dP_{\theta_i}(t) + \gamma_1 P_{\theta_i}(T=c_1) + \gamma_2 P_{\theta_i}(T=c_2) = \alpha, \quad i=1,2.$$

V praxi můžeme postupovat tak, že vyjdeme z nějakých počátačních hodnot c_1^*, γ_1^*, a k nim určíme c_2^*, γ_2^* tak, že $\beta^*(\theta_1) = \alpha$. Pak vypočteme $\beta^*(\theta_2)$ a porovnáme s α. Jestliže je $\beta^*(\theta_2) < \alpha$, musí buď platit $c_1 > c_1^*$ nebo $c_1 = c_1^*$ a $\gamma_1 < \gamma_1^*$, což znamená, že skutečný kritický obor leží zpravo od předpokládaného. Jestliže je $\beta^*(\theta_2) > \alpha$, musí platit opačné nerovnosti.

2.6. Nejméně příznivá rozdělení

Podle věty 2.1 můžeme zkonstruovat test jednoduché hypotézy proti jednoduché alternativě. Věta 2.4 rozšiřuje tuto konstrukci na testy hypotézy složené z konečně mnoha hodnot proti jednoduché alternativě.

Nechť $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ je dominovaný systém rozdělení a uvažujme problém testu složené hypotézy $H : \theta \in \Theta_0$ proti jednoduché alternativě $K : \theta = \theta_1$. Ptáme se, za jakých okolností můžeme složenou hypotézu H nahradit jednoduchou tak, že nejsilnější test nové hypotézy proti K poskytuje zároveň nejsilnější test H proti K.

Očekáváme, že jednoduchá hypotéza, která nahrazuje H, bude váženým průměrem všech rozdělení z H. Označíme-li $f_\theta(x)$ hustotou P_θ, $\theta \in \Theta$, pak nová hypotéza bude mít tvar

$$H_\lambda : h_\lambda(x) = \int_{\Theta_0} f_\theta(x) d\lambda(\theta) \tag{2.34}$$

kde λ je vhodné rozdělení pravděpodobnosti na Θ_0. Jestliže jsme stanovili λ a tím i H_λ, dovedeme pomocí věty 2.1 nalézt nejsilnější test H_λ proti K. Otevřenou zbývá otázka, jak vhodně stanovit λ, což je vlastně Lagrangeův multiplikátor.

Protože H mezi sebou nerozlišuje hodnoty $\theta \in \Theta_0$, má i H_λ být stejně vhodné pro všecka $\theta \in \Theta_0$, i pro hodnoty nejbližší alternativě θ_1. Intuitivně je zřejmé, že rozdělení λ musí být <u>nejméně příznivé</u>, tj. pro libovolné jiné rozdělení musí platit $\beta_\lambda \leq \beta_{\lambda'}$, kde β_λ $[\beta_{\lambda'}]$ je síla nejsilnějšího testu H_λ $[H_{\lambda'}]$ proti K. Tuto domněnku potvrzuje i následující věta.

VĚTA 2.6. <u>Nechť $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ je dominovaný systém rozdělení s hustotami</u> $f_\theta(x)$ <u>vzhledem k</u> σ -konečné míře μ. <u>Uvažujme problém testování hypotézy</u> $H : \theta \in \Theta_0$ <u>proti</u> $K : \theta = \theta_1 \notin \Theta_0$.

Nechť existuje σ-algebra $\mathcal{D}$ podmnožin Θ_0 taková, že hustoty $f_\theta(x)$ jsou zároveň měřitelné v θ i v x. Předpokládejme, že existuje rozdělení pravděpodobnosti λ na $\mathcal{D}$ takové, že nejsilnější α-test Φ_λ pro testování H_λ proti K má velikost $\leq \alpha$ také jako test H proti K. Pak

(a) Test Φ_λ je nejsilnější α-test H proti K.

(b) Jestliže Φ_λ je jediným nejsilnějším α-testem pro H_λ proti K, je i jediným nejsilnějším α-testem H proti K.

(c) Rozdělení λ je nejméně příznivé.

Důkaz. Z Fubiniho věty plyne, že h_λ je hustotou vzhledem k μ, tj. $\int h_\lambda(x)\,d\mu(x) = 1$.

Nechť Φ_λ má velikost $\leq \alpha$ i pro H proti K a nechť Φ^* je libovolný jiný α-test H proti K. Pak

$$\int \Phi^*(x) h_\lambda(x)\,d\mu = \int_{\Theta_0} E_\theta \Phi^*(X)\,d\lambda(\theta) \leq \alpha ,$$

protože $E_\theta \Phi^*(X) \leq \alpha \ \forall \theta \in \Theta_0$. To znamená, že Φ^* je α-testem i pro H_λ proti K, a tedy jeho síla nemůže překročit sílu Φ_λ. Tím je dokázáno (a) a (b).

Nechť λ' je libovolné rozdělení na Θ_0. Pak

$$\int \Phi_\lambda(x) h_{\lambda'}(x)\,d\mu = \int_{\Theta_0} E_\theta \Phi_\lambda(X)\,d\lambda'(\theta) \leq \alpha ,$$

tedy Φ_λ je α-testem i pro hypotézu $H_{\lambda'}$ proti K. Jeho síla β_λ pak nemůže překročit sílu $\beta_{\lambda'}$ nejsilnějšího testu $H_{\lambda'}$ proti K, tedy λ je nejméně příznivé. Q.E.D.

Jestliže je velikost testu Φ_λ rovna α, můžeme podmínky věty 2.6 poněkud zjednodušit. Stačí si uvědomit, že $\int_{\Theta_0} E_\theta \Phi_\lambda(X) d\lambda(\theta) = \alpha$ a $E_\theta \Phi_\lambda(X) \leq \alpha \ \forall \theta \in \Theta_0$ mohou zároveň nastat pouze v případě, že $\lambda\{\theta : E_\theta \Phi_\lambda(X) = \alpha\} = 1$.

Důsledek. *Nechť λ je rozdělení pravděpodobnosti na Θ_0 a nechť Θ_0' je podmnožina Θ_0 taková, že $\lambda(\Theta_0') = 1$. Nechť Φ_λ je test daný vztahy*

$$(2.35)\quad \Phi_\lambda(x) = \begin{cases} 1 & \text{jestliže} \quad f_{\theta_1}(x) > k \int f_\theta(x) d\lambda(\theta) \\ 0 & \text{jestliže} \quad f_\theta(x) < k \int f_\theta(x) d\lambda(\theta). \end{cases}$$

Jestliže test Φ_λ splňuje

$$(2.36)\qquad E_{\theta'} \Phi_\lambda(X) = \sup_{\theta \in \Theta_0} E_\theta \Phi_\lambda(X) = \alpha \quad \forall \theta' \in \Theta_0',$$

pak je nejsilnějším α-testem H proti K.

Poznámka. Věty 2.3 a 2.5 jsou jednoduchými aplikacemi věty 2.6. Množina Θ_0', na které je soustředěno nejméně příznivé rozdělení, je v prvním případě složena z jediného bodu θ_0 a ve druhém případě ze dvou bodů θ_1 a θ_2.

2.7. Aplikace na testování hypotéz o rozptylu normálního rozdělení

Nechť $X_1, \ldots, X_n$ je náhodný výběr z normálního rozdělení $N(\xi, \sigma^2)$, $n > 3$. Uvažujme hypotézy

$$H_1 : \{(\xi,\sigma): \sigma \geq \sigma_0\} \quad \text{proti} \quad K_1 : \{(\xi,\sigma): \sigma < \sigma_0\}$$

a

$$H_2 : \{(\xi,\sigma): \sigma \leq \sigma_0\} \quad \text{proti} \quad K_2 : \{(\xi,\sigma): \sigma > \sigma_0\}.$$

Hypotézy se týkají pouze parametru σ, zatímco ξ je rušivým parametrem. Obvykle píšeme hypotézy a alternativy ve zkráceném tvaru $H_1 : \sigma \geq \sigma_0$ apod.

S použitím teorie dosud odvozené v této kapitole můžeme odvodit stejnoměrně nejsilnější test hypotézy H_2 proti K_2 a ukázat, že neexistuje stejnoměrně nejsilnější test H_1 proti K_1.

Uvažujme nejprve hypotézu H_1 proti jednoduché alternativě $K_1' : \xi = \xi_1, \sigma = \sigma_1 \ (\sigma_1 < \sigma_0)$. Budeme hledat nejméně příznivé rozdělení pro hypotézu H_1 vzhledem ke K_1'. Očekáváme, že rozdělení v rovině (ξ,σ), nejméně příznivé vzhleke K_1', bude soustředěno na přímce $\{(\xi,\sigma): \sigma = \sigma_0\}$. Nejméně příznivé rozdělení λ má být takové, že rozdělení $(X_1,\dots,X_n)$ za hypotézy H_λ je co nejblíže rozdělení $(X_1,\dots,X_n)$ za alternativy K_1', které má hustotu

$$(2.37) \qquad \sigma_1^{-n}(2\pi)^{-n/2} \exp\left\{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{n}(x_i-\xi_1)^2\right\}.$$

Pomocí důsledku věty 2.6 dokážeme, že nejméně příznivé rozdělení λ přiřazuje pravděpodobnost 1 bodu (ξ_1,σ_0). Skutečně, dosadíme-li takové λ do (2.35), dostaneme

$$(2.38) \quad \Phi_\lambda(x) = \begin{cases} 1 & \text{jestliže } (2\pi\sigma_1^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_1^2}\sum_{i=1}^n (x_i-\xi_1)^2\right\} > \\ & > k(2\pi\sigma_0^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_0^2}\sum_{i=1}^n (x_i-\xi_1)^2\right\} \\ 0 & \text{jestliže platí opačná nerovnost.} \end{cases}$$

Test Φ_λ můžeme přepsat ve tvaru

$$(2.39) \quad \Phi_\lambda(x) = \begin{cases} 1 & \text{jestliže } \sum_{i=1}^n (x_i-\xi_1)^2 \le c \\ 0 & \text{jestliže } \sum_{i=1}^n (x_i-\xi_1)^2 > c. \end{cases}$$

Podle důsledku věty 2.6 je rozdělení λ nejméně příznivé, jestliže platí

$$(2.40) \quad P_{\xi_1,\sigma_0}\left\{\sum_{i=1}^n (X_i-\xi_1)^2 \le c\right\} = \sup_{\substack{\xi\in R^1 \\ \sigma\ge\sigma_0}} P_{\xi,\sigma}\left\{\sum_{i=1}^n (X_i-\xi_1)^2 \le c\right\}.$$

Při libovolném pevném σ je $P_{\xi,\sigma}\left\{\sum_{i=1}^n (X_i-\xi_1)^2 \le c\right\}$ rovna pravděpodobnosti, že náhodný výběr z $N(\mu, \sigma^2)$ padne do koule se středem ξ_1 a poloměrem $\sqrt{c}$. Maximum této pravděpodobnosti při pevném σ nastane pro $\xi = \xi_1$ a je rovno

$$(2.41) \quad P_{\xi_1,\sigma}\left\{\sum_{i=1}^n \left(\frac{X_i-\xi_1}{\sigma}\right)^2 \le \frac{c}{\sigma^2}\right\} = P_{0,1}\left\{\sum_{i=1}^n (X_i^0)^2 \le \frac{c}{\sigma^2}\right\},$$

kde $X_1^0,\dots,X_n^0$ jsou nezávislé náhodné veličiny s rozdělením $N(0,1)$. Pravá strana (2.42) je klesající v σ a tedy nabývá maxima pro $\sigma = \sigma_0$; konstanta c je dána vztahem

$c = \sigma_o^2 \cdot \chi_n^2(1-\alpha)$, kde $\chi_n^2(1-\alpha)$ je $100(1-\alpha)$-procentní kritická hodnota rozdělení χ^2 o n stupních volnosti. Ze (2.40) a z důsledku věty (2.6) vyplývá, že test Φ_λ (2.39) je nejsilnějším testem hypotézy $H_1 : \sigma \geq \sigma_o$ proti alternativě $K_1' : \xi = \xi_1$, $\sigma = \sigma_1 (< \sigma_o)$. Protože Φ_λ nezávisí na zvoleném σ_1, je i stejnoměrně nejsilnějším testem hypotézy $H_1 : \sigma \geq \sigma_o$ proti alternativě $K_1'' : \xi = \xi_1$, $\sigma < \sigma_o$. Na druhé straně, protože Φ_λ závisí na zvolené hodnotě ξ_1, neexistuje stejnoměrně nejsilnější test H_1 proti K_1.

Nyní uvažujme hypotézu $H_2 : \sigma \leq \sigma_o$ proti alternativě $K_2' : \xi = \xi_1$, $\sigma = \sigma_1$ $(\sigma_1 > \sigma_o)$. Stejně jako v prvním případě očekáváme, že nejméně příznivé rozdělení bude soustředěno na přímce $\sigma = \sigma_o$.

Z faktorizační věty plyne, že statistiky $Y = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$ a $Z = \sum_{i=1}^{n} (X_i - \overline{X})^2$ jsou postačující pro parametr (ξ, σ). Stačí tedy omezit své úvahy na tyto postačující statistiky. Jak je známo (viz např. Anděl [1], kap.V, věta 18), jsou Y a Z nezávislé, Y má rozdělení $N(\xi, \frac{\sigma^2}{n})$ a $\frac{Z}{\sigma^2}$ má rozdělení $\chi^2(n-1)$. Jestliže $\lambda(\xi)$ je rozdělení v rovině (ξ, σ), soustředěné na přímce $\sigma = \sigma_o$, pak sdružené rozdělení Y, Z za platnosti hypotézy H_λ má hustotu

$$(2.42)\quad \left(\frac{n}{2\pi}\right)^{1/2} \left(\frac{2}{\sigma_o}\right)^{(n-1)/2} \Gamma\left(\frac{n-1}{2}\right) z^{(n-3)/2} \exp\left\{-\frac{z}{2\sigma_o^2}\right\} \cdot \int \exp\left\{-\frac{n}{2\sigma_o^2}(y-\xi)^2\right\} d\lambda(\xi)$$

zatímco za platnosti K_2' má hustotu

$$(2.43)\quad \left(\frac{n}{2\pi}\right)^{1/2} 2^{(n-1)/2}\,\Gamma\left(\frac{n-1}{2}\right)\left(\frac{z}{\sigma_1}\right)^{(n-3)/2}\exp\left\{-\frac{z}{2\sigma_1^2}\right\}\cdot$$

$$\cdot\ \exp\left\{-\frac{n}{2\sigma_1^2}(y-\xi_1)^2\right\}.$$

Ukážeme, že nejméně příznivé rozdělení $\lambda(\xi)$ hypotézy H_2 vzhledem ke K_2' je normální $N(\xi_1,(\sigma_1^2-\sigma_0^2)/n)$. Dosadíme-li toto λ do (2.42), dostaneme za integrálem konvoluci dvou normálních rozdělení (až na normalizační konstantu) $N(0,\sigma_0^2/n)$ a $N(\xi_1,(\sigma_1^2-\sigma_0^2)/n)$, což je $N(\xi_1,\sigma_1^2/n)$. Sdružené rozdělení (Y,Z) za H_λ má pak hustotu

$$\left(\frac{n}{2\pi}\right)^{1/2} 2^{(n-1)/2}\,\Gamma\left(\frac{n-1}{2}\right)\left(\frac{z}{\sigma_0}\right)^{(n-3)/2} e^{-z/2\sigma_0^2}\,\frac{1}{\sigma_1}$$

$$\exp\left\{-\frac{n}{2\sigma_1^2}(y-\xi_1)^2\right\}$$

Test Φ_λ, vyjádřený v závislosti na postačujících statistikách Y,Z, má pak tvar

$$\Phi_\lambda(y,z) = 1 \qquad \text{jestliže} \qquad z \geq C,$$

neboli

$$(2.44)\qquad \Phi_\lambda(\underset{\sim}{x}) = 1 \qquad \text{jestliže} \qquad \sum_{i=1}^{n}(x_i-\bar{x})^2 \geq C.$$

Pravděpodobnost $P_{\xi,\sigma}\left\{\sum_{i=1}^{n}(x_i-\bar{x})^2 \geq C\right\} = P_{\xi,\sigma}\left\{\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2 \geq \frac{C}{\sigma^2}\right\}$ nezávisí na ξ a je rostoucí v σ. Jsou tedy splněny předpoklady důsledku věty (2.6) a test Φ_λ (2.44) je stejnoměrně nejsilnějším testem $H_2 : \sigma \leq \sigma_0$ proti $K_2 : \sigma > \sigma_0$; přitom $C = \sigma_0^2 \cdot \chi^2_{n-1}(\alpha)$, kde $\chi^2_{n-1}(\alpha)$ je 100α-procentní kritická hodnota rozdělení χ^2 o $(n-1)$ stupních volnosti.

2.8. Doplňky a cvičení

(1) Nechť $X_1,\dots,X_m$ a $Y_1,\dots,Y_n$ jsou nezávislé výběry z $N(\xi,1)$ a $N(\eta,1)$. Uvažujme hypotézu $H: \eta \leq \xi$ proti alternativě $K: \eta > \xi$. Pak existuje stejnoměrně nejsilnější test, který zamítá H *při velkých hodnotách* $\bar{Y} - \bar{X}$.

[Návod : Zvolíme-li pevnou alternativu $K': \xi_1, \eta_1; \xi_1 < \eta_1$, pak existuje nejméně příznivé rozdělení H vzhledem ke K', které přiřazuje pravděpodobnost 1 bodu $\xi = \eta = \dfrac{m\xi_1 + n\eta_1}{m+1}$.]

(2) Nechť $X_1,\dots,X_m$ a $Y_1,\dots,Y_n$ jsou dva nezávislé výběry z $N(\xi_1, \sigma_1^2)$ a $N(\xi_2, \sigma_2^2)$. Uvažujme hypotézu $H: \sigma_2^2 \leq \sigma_1^2$ proti $K: \sigma_2^2 > \sigma_1^2$.

(a) Jestliže jsou známy hodnoty ξ_1 a ξ_2, existuje stejnoměrně nejsilnější test s kritickým oborem

$$\frac{\sum_{j=1}^{n}(Y_j - \xi_2)^2}{\sum_{i=1}^{n}(X_i - \xi_1)^2} \geq c.$$

(b) Jestliže hodnoty ξ_1 a ξ_2 nejsou známy, neexistuje stejnoměrně nejsilnější test.

(3) Stanovte silofunkci testu (2.45) hypotézy $H_2: \sigma \leq \sigma_0$ proti $K_2: \sigma > \sigma_0$ o rozptylu normálního rozdělení.

(4) Nechť $X_1,\dots,X_n$ je náhodný výběr z rozdělení gama s hustotou

$$g_{p,\lambda}(x) = \frac{\lambda^p x^{p-1} e^{-\lambda x}}{\Gamma(p)}, \quad x>0,\ \lambda>0,\ p>0$$

$$= 0 \qquad \text{jinde.}$$

Předpokládejme, že p je známo a $\frac{1}{\lambda}$ je neznámý parametr měřítka. Ukažte, že stejnoměrně nejsilnější test $H : \frac{1}{\lambda} \leq \frac{1}{\lambda_0}$ proti $K : \frac{1}{\lambda} > \frac{1}{\lambda_0}$ má kritický obor $\sum_{i=1}^{n} x_i \geq c$, kde $c = \frac{1}{\lambda_0} g_{np}(1-\alpha)$ a $g_{np}(1-\alpha)$ je $100(1-\alpha)$-procentní kritická hodnota rozdělení gama s parametry $np,1$.