Sekvenční testy

Je jich celá řada, existují sekvenční obdoby snad všech známějších klasických testů (t-testu, F-testu), i neparametrických (o korelačním koeficientu). Jejich odvození vyžaduje dost teorie, takže uvedu na ukázku jeden (asi nejjednodušší).

Waldův sekvenční test poměrem pravděpodobností

Mám 2 jednoduché hypotézy H_0 (hypotéza)

H_1 (alternativa)

týkající se posloupnosti náh. veličin $X_1, X_2, \ldots$

Nechť $f(x_1, \ldots, x_m / H_0)$ je hustota prvních m pozorování za platnosti H_0

$f(x_1, \ldots, x_m / H_1)$ —"— H_1

Rozhodování se provádí na základě poměru

$$R_m = \frac{f(x_1, \ldots, x_m / H_1)}{f(x_1, \ldots, x_m / H_0)}$$

tak, že malé hodnoty R_m svědčí ve prospěch H_0, velké ve prospěch H_1.

Konkrétně:

ukončíme experimenty a přijmeme H_0, když $R_m \leq B$

—"— H_1, když $R_m \geq A$

provedeme další experiment x_{m+1}, když $B < R_m < A$

A, B jsou konstanty takové, že $0 < B < 1 < A < \infty$, které se stanoví v závislosti na pravděpodobnostech chyb α, β.

$\alpha = P$(zamítneme H_0 když H_0 platí) .. pst chyby 1. druhu

$\beta = P$(přijmeme H_0 (tj. zamítneme H_1) když H_1 platí). pst chyby 2. druhu

Problémy:

a) jak stanovit konstanty A, B

b) vede tento postup alespoň s pstí 1 po konečném počtu pokusů k definitivnímu rozhodnutí?

c) jaká je očekávaná délka výběru, potřebná ke konečnému rozhodnutí?

ad a) doporučená volba konstant je $A = \frac{1-\beta}{\alpha}$, $B = \frac{\beta}{1-\alpha}$

Při nich sice test nemá přesně psti chyb rovny α a β, ale jejich hodnoty jsou jim hodně blízké.

Přesné hodnoty chyb jsou α', β' a platí pro ně:

$$\alpha' \leq \frac{\alpha}{1-\beta}, \quad \beta' \leq \frac{\beta}{1-\alpha}, \quad (\alpha'+\beta') \leq (\alpha+\beta).$$

ad b) Za předpokladu, že pozorování $X_1, X_2, \ldots$ jsou navzájem nezávislá a mají stejné rozdělení, platí $P(n<\infty) = 1$.

ad c) Za týchž předpokladů existuje konečná střední hodnota potřebného počtu pozorování a platí pro ni

$$E(n|H_0) \doteq \frac{b(1-\alpha)+a\alpha}{E(z|H_0)}, \quad E(n|H_1) \doteq \frac{b\beta+a(1-\beta)}{E(z|H_1)}$$

kde $a = \log A$, $b = \log B$, $z(x) = \log \frac{f(x|H_1)}{f(x|H_0)}$, kde bude $f(x|H_0)$, resp. $f(x|H_1)$, je hustota jednotlivého pozorování za platnosti H_0, resp. H_1.

Za předpokladu nezávislosti a stejného rozdělení platí

$$R_m = \frac{\prod_{i=1}^{m} f(x_i|H_1)}{\prod_{i=1}^{m} f(x_i|H_0)} = \prod_{i=1}^{m} \frac{f(x_i|H_1)}{f(x_i|H_0)} = \cancel{\prod_{i=1}^{m} z(x_i)}.$$

Z toho $\log R_m = \sum_{i=1}^{m} z(x_i)$.

Další experimenty se dělají, dokud je $B < R_m < A$, tj. $b < \log R_m < a$.

Postup testování:

Stanovím si α, β, podle nich vypočtu A, B.

Provádím postupně experimenty, tj. získávám pozorování $x_1, x_2, \dots$, dokud je $b < z(x_1) + z(x_2) + \dots + z(x_i) < a$.

Jakmile součet vyjde mimo interval (b, a), experimenty ukončím a přijmu buď H_0 nebo H_1.

Vzorec pro $E(n)$ nebudeme dokazovat, místo toho radši uvedeme příklad.

Příklad: Nechť $X_1, X_2, \dots$ jsou navzájem nezávislá pozorování náh. veličiny X, která má za platnosti H_0 rozdělení $N(\mu_0, \sigma^2)$ a za platnosti H_1 rozdělení $N(\mu_1, \sigma^2)$, kde $\mu_1 > \mu_0$ a σ^2 je známé.

Testujeme tedy $H_0: \mu = \mu_0$ proti $H_1: \mu = \mu_1$ při známém σ^2.

$$R_m = \frac{f(x_1, \dots, x_m \mid \mu_1)}{f(x_1, \dots, x_m \mid \mu_0)} = \prod_{i=1}^{m} \frac{f(x_i \mid \mu_1)}{f(x_i \mid \mu_0)}$$

$$\log R_m = \sum_{i=1}^{m} \log \frac{f(x_i \mid \mu_1)}{f(x_i \mid \mu_0)} = \sum_{i=1}^{m} \log \frac{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\mu_1)^2}{2\sigma^2}}}{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\mu_0)^2}{2\sigma^2}}} =$$

$$= \sum_{i=1}^{m} \left(-\frac{(x_i-\mu_1)^2}{2\sigma^2} + \frac{(x_i-\mu_0)^2}{2\sigma^2} \right) = \frac{1}{2\sigma^2} \sum_{i=1}^{m} (-x_i^2 + 2x_i\mu_1 - \mu_1^2 +$$

$$+ x_i^2 - 2x_i\mu_0 + \mu_0^2) =$$

$$= \frac{\mu_1-\mu_0}{2\sigma^2} \sum_{i=1}^{m} \left(2x_i - (\mu_1+\mu_0)\right) = \frac{\mu_1-\mu_0}{2\sigma^2} \left(2\sum_{i=1}^{m} x_i - m(\mu_1+\mu_0)\right).$$

a) Jestliže $R_m \le B = \frac{\beta}{1-\alpha}$, tj. $\log R_m \le \log B = b$, experiment ukončíme a přijmeme H_0.

$$\frac{\mu_1-\mu_0}{2\sigma^2} \left(2\sum_{i=1}^{m} x_i - m(\mu_1+\mu_0)\right) \le b$$

$$\sum_{i=1}^{m} x_i \le \frac{b\sigma^2}{\mu_1-\mu_0} + \frac{m(\mu_1+\mu_0)}{2}$$

b) Jestliže $R_m \geq A = \frac{1-\beta}{\alpha}$, tj. $\log R_m \geq \log A = a$, experimenty ukončíme a zamítneme H_0.

$$\sum_{i=1}^{m} X_i \geq \frac{a\sigma^2}{\mu_1-\mu_0} + \frac{m(\mu_1+\mu_0)}{2}$$

c) Jestliže $B < R_m < A$, tj. $b < \log R_m < a$, pokračujeme v experimentech.

$$\frac{b\sigma^2}{\mu_1-\mu_0} + \frac{m(\mu_1+\mu_0)}{2} < \sum_{i=1}^{m} X_i < \frac{a\sigma^2}{\mu_1-\mu_0} + \frac{m(\mu_1+\mu_0)}{2}$$

Očekávaná velikost výběru:

$$E(n/H_0) \doteq \frac{b(1-\alpha)+a\alpha}{E(\varkappa/H_0)} \qquad E(n/H_1) \doteq \frac{b\beta+a(1-\beta)}{E(\varkappa/H_1)}$$

$$\varkappa(x) = \log\frac{f(x/H_1)}{f(x/H_0)} = \log\frac{\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu_1)^2}{2\sigma^2}}}{\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu_0)^2}{2\sigma^2}}} = -\frac{(x-\mu_1)^2}{2\sigma^2} + \frac{(x-\mu_0)^2}{2\sigma^2} =$$

$$= \frac{(x-\mu_0)^2-(x-\mu_1)^2}{2\sigma^2}$$

$$E(\varkappa/H_0) = -\frac{(\mu_0-\mu_1)^2}{2\sigma^2} \qquad E(\varkappa/H_1) = \frac{(\mu_1-\mu_0)^2}{2\sigma^2}$$

$$E(n/H_0) = -\frac{b(1-\alpha)+a\alpha}{(\mu_0-\mu_1)^2/2\sigma^2} \qquad E(n/H_1) = \frac{b\beta+a(1-\beta)}{(\mu_1-\mu_0)^2/2\sigma^2}$$

Zajímavé je porovnání s klasickým nesekvenčním testem:

zamítáme H_0, když $\frac{\bar{X}-\mu_0}{\sigma}\sqrt{n} > u_{1-\alpha}$

síla testu při H_1 je

$$\gamma(\mu_1) = \Phi\left(u_\alpha + \frac{\mu_1-\mu_0}{\sigma}\sqrt{n}\right) = 1-\beta$$

rozsah výběru nutný k dosažení této síly je $n^* = \dfrac{(u_{1-\alpha}+u_{1-\beta})^2\sigma^2}{(\mu_1-\mu_0)^2}$

Porovnáme to se středním rozsahem výběru sekvenčního testu:

za platnosti H_0:

$$\frac{E(n|H_0)}{n^*} = -\frac{2\sigma^2(b(1-\alpha)+a\alpha)}{(\mu_0-\mu_1)^2}\cdot\frac{(\mu_1-\mu_0)^2}{\sigma^2(u_{1-\alpha}+u_{1-\beta})^2} = -\frac{2(b(1-\alpha)+a\alpha)}{(u_{1-\alpha}+u_{1-\beta})^2}$$

konkrétně pro $\alpha = 0{,}05$, $\beta = 0{,}05$

$$u_{0,95} = 1{,}645$$

$$a = \log\frac{1-\beta}{\alpha} = \log\frac{0{,}95}{0{,}05} = \log 19 = 2{,}9444$$

$$b = \log\frac{\beta}{1-\alpha} = \log\frac{0{,}05}{0{,}95} = \log\frac{1}{19} = -\log 19 = -2{,}9444$$

$$\frac{E(n|H_0)}{n^*} = -\frac{2(-2{,}9444\cdot 0{,}95 + 2{,}9444\cdot 0{,}05)}{(1{,}645+1{,}645)^2} = \frac{5{,}29996}{10{,}8241} = 0{,}4896$$

za platnosti H_1:

$$\frac{E(n|H_1)}{n^*} = \frac{2\sigma^2(b\beta+a(1-\beta))}{(\mu_1-\mu_0)^2}\cdot\frac{(\mu_1-\mu_0)^2}{\sigma^2(u_{1-\alpha}+u_{1-\beta})^2} = \frac{2(b\beta+a(1-\beta))}{(u_{1-\alpha}+u_{1-\beta})^2} = 0{,}48$$

Sekvenčním postupem se ušetří polovina experimentů. %

Pokud by skutečná hodnota μ byla mezi μ_0 a μ_1, byl by očekávaný rozsah výběru sekvenčního testu asi vyšší.

12649 Svépomoc II/91-100 – TZ 13-322-91-G

Příklad: Quicksort, varianta M1, měříme čas

máme $n = 25$, $\sum_{i=1}^{25} X_i = 201,1$, $\bar{X} = 8,044$, $S = 0,115$

testujeme $H_0: \mu = 8$ proti $A: \mu = 8,115$

(tj. jako bychom si mysleli, že S je skutečná hodnota parametru σ (známe), rozdíl mezi hypotetickou a alternativní hodnotou μ je roven σ)

zvolíme $\alpha = 0,05$, $\beta = 0,1$

$$A = \frac{1-\beta}{\alpha} = \frac{0,9}{0,05} = 18 \qquad \ln A = 2,89 = a$$

$$B = \frac{\beta}{1-\alpha} = \frac{0,1}{0,95} = 0,105 \qquad \ln B = -2,25 = b$$

$$\frac{\sigma^2}{\mu_1 - \mu_0} = 0,115$$

$$\frac{n(\mu_1 + \mu_0)}{2} = \frac{25 \cdot 16,115}{2} = 201,4375$$

dolní mez: $-2,25 \cdot 0,115 + 201,4375 = 201,178$

horní mez: $2,89 \cdot 0,115 + 201,4375 = 201,769$

platí $\sum X_i <$ dolní mez, experimenty ukončíme a přijmeme H

Pro srovnání: podle tabulek odvozených ze silofunkce nám má 25 měření stačit na rozpoznání rozdílu $0{,}6\sigma$ (pro $\alpha = 0{,}05$ a síle testu $0{,}9$)

zkusme tedy $H: \mu = 8$ proti $A: \mu = 8 + 0{,}6\sigma = 8{,}069$

$$\frac{\sigma^2}{\mu_1 - \mu_0} = \frac{\sigma^2}{0{,}6\sigma} = \frac{\sigma}{0{,}6} = \frac{0{,}115}{0{,}6} = 0{,}192$$

$$\frac{n(\mu_1 + \mu_0)}{2} = \frac{25 \cdot 16{,}069}{2} = 200{,}8625$$

dolní mez: $-2{,}25 \cdot 0{,}192 + 200{,}8625 = 200{,}8145$

horní mez: $2{,}89 \cdot 0{,}192 + 200{,}8625 = 201{,}417$

platí: dolní mez $< \sum X_i <$ horní mez

takže musíme provést další experimenty

V tomto případě jsme tedy neušetřili polovinu experimentů, možná proto, že $\bar{X} = 8{,}044$ naznačuje, že skutečná hodnota μ by opravdu mohla být mezi μ_0 a μ_1.

Jaká je vlastně očekávaná hodnota velikosti výběru pro tuto situaci?

$$E(n \mid \mu = 8) = -\frac{b(1-\alpha) + a\alpha}{(\mu_0 - \mu_1)^2 / 2\sigma^2} = \frac{2{,}25 \cdot 0{,}95 + 2{,}89 \cdot 0{,}05}{\frac{(0{,}6\sigma)^2}{2\sigma^2}} = \frac{2{,}1375 + 0{,}1445}{0{,}18} =$$

$= 12{,}67$ (což je skutečně polovina z 25)

$$E(n \mid \mu = 8{,}069) = \frac{b\beta + a(1-\beta)}{0{,}18} = \frac{-2{,}25 \cdot 0{,}1 + 2{,}89 \cdot 0{,}9}{0{,}18} = 13{,}2 \text{ (to taky)}$$

Zkusíme ještě ten samý rozdíl, tj. $0{,}6\sigma$, ale tak, aby skutečná hodnota μ neležela mezi μ_0 a μ_1, např.:

$H: \mu = 8{,}069$ proti $A: \mu = 8{,}138$

zůstává: $a = 2{,}89$

$b = -2{,}25$

$$\frac{\sigma^2}{\mu_1 - \mu_0} = \frac{\sigma^2}{0{,}6\sigma} = \frac{\sigma}{0{,}6} = \frac{0{,}115}{0{,}6} = 0{,}192$$

mění se: $$\frac{n(\mu_1 + \mu_0)}{2} = \frac{25 \cdot 16{,}207}{2} = 202{,}5875$$

dolní mez: $-2{,}25 \cdot 0{,}192 + 202{,}5875 = 202{,}1555$

horní mez: $2{,}89 \cdot 0{,}192 + 202{,}5875 = 203{,}14238$

$\sum X_i = 201{,}1 <$ dolní mez

experimenty ukončíme a přijmeme H