# Detekce strukturálních změn

Marie Hušková

## 1 Úvod

Cílem článku je vyložit základní principy konstrukce testů a podat přehled základních typů úloh, které jsou v anglicky psané literatuře nazývány "testing constancy of regression relationship over time" a "change point problem".

Soustředíme se na jednodušší případy, kdy máme nezávislé náhodné veličiny $X_1, X_2, \ldots, X_n$ pozorované v po sobě jdoucích časových okamžicích $t_1 < t_2 < \cdots < t_n$ a $X_i$ má distribuční funkci $F_i$, $i = 1, \ldots, n$.

*Detekcí změny regrese* budeme rozumět úlohu testovat hypotézu

$$(1.1) \qquad H_0 : \underset{\sim}{\theta}_1 = \cdots = \underset{\sim}{\theta}_n \text{ proti } H_1 : \exists_{m<n} \underset{\sim}{\theta}_1 = \cdots = \underset{\sim}{\theta}_m \neq \underset{\sim}{\theta}_{m+1} = \cdots \underset{\sim}{\theta}_n$$

kde vycházíme ze situace, že $X_i$ má distribuční funkci $F(x - \underset{\sim}{c}_i' \underset{\sim}{\theta}_i)$, $i = 1, \ldots, n$, $\underset{\sim}{c} = (c_{i1}, \ldots, c_{ip})'$ je vektor neznámých parametrů, $i = 1, \ldots, n$.

*Detekcí změny rozdělení* budeme rozumět úlohu testovat hypotézu

$$(1.2) \qquad H_0 : F_1 = \cdots = F_n \text{ proti } H_1 : \exists_{m<n} F_1 = \cdots = F_m \neq F_{m+1} = \cdots = F_n.$$

Tedy hypotéza $H_1$ v (1.1) znamená, že až do určitého časového okamžiku $\tau \in (t_m, t_{m+1})$, který obecně neznáme, se pozorování řídila jedním regresním modelem popsaným parametrem $\underset{\sim}{\theta}_1$ a po něm dochází ke změně a pozorování se řídí jiným regresním modelem popsaným parametrem $\underset{\sim}{\theta}_{m+1}$. Časový okamžik $\tau$ se nazývá bodem změny (change point). Podobně je tomu s hypotézou $H_1$. Kromě verifikace platnosti hypotéz nás může zajímat odhad bodu změny $\tau$ (popř. $m$). Při speciálních úlohách mohou být některé z hodnot parametrů známy.

Nyní si uvedeme některé modifikace zmíněných úloh. Při detekci změny regrese můžeme do hypotézy $H_1$ zahrnout i změnu měřítka (viz (3.1) dole). Další možnost je uvažovat v alternativě více než jednu změnu (tj. existenci bodů změn $\tau_1 < \tau_2 < \cdots < \tau_\nu$, $\nu > 1$). Dále změna hodnot parametrů může proběhnout náhle, což odpovídá (1.1), nebo může probíhat spojitě, s tím, že začne v okamžiku $\tau_1^*$, pak v časovém intervalu $(\tau_1^*, \tau_2^*)$ změna pokračuje (jde o jakýsi přechodný stav) a teprve v $\tau_2^*$ se situace ustálí a sledované charakteristiky se už nemění.

Se všemi těmito úlohami se můžeme při aplikacích setkat. Uvedeme si aspoň některé z nich.

Při kontrole jakosti sledujeme ukazatele kvality výrobků a zajímá nás, zda došlo ke změně ukazatele či nikoli, a pokud ke změně došlo, co nejdříve ji odhalit (většinou sledujeme podíl vadných, např. Page (1955, 1957)).

$X_i$ mají normální rozdělení a k sestavení testové statistiky se používají odhady získané metodou nejmenších čtverců.

Při robustním přístupu se předpokládá, že distribuční funkce $F_i$ splňují nějaké podmínky regularity, ale jinak jsou neznámé. Při sestavování testové statistiky se používají tzv. robustní odhady ($R$,-, $M$-odhady, $U$-statistiky), popř. statistiky Kolmogorova-Smirnovova typu.

Pochopitelně existují postupy, které jsou kombinací přístupů výše uvedených, nebo které bychom těžko klasifikovali podle tohoto třídění.

Byla napsána celá řada přehledných článků, většinou zaměřených jen na určité typy přístupů a postupů. Patří mezi ně následující články: Zacks (1983), Shaban (1980), Schulze (1986), Hackl a Westlund (1985), Deshayes a Picard (1986), Hušková a Sen (1989), Wolfe a Schechtman (1984), Csörgö a Horváth (1988), Krishnaiah a Miao (1988). Pravděpodobně největší popularitu získal článek Browna, Durbina a Evanse (1975).

Zbytek článku je členěn následovně. Paragraf 2 je věnován problému posunutí (tj. změna se projeví v posunutí). Je rozdělen do dvou částí. V první jsou pro případ nezávislých normálně rozdělených $X_1, \ldots, X_n$ vyloženy základní typy testů (bayesovský, podílem věrohodností, CUSUM a MOSUM). V druhé části jsou pak jejich robustní analogie založené na pořadích a $M$-odhadech. Třetí paragraf je věnován detekci změny regrese a je členěn obdobně jako druhý. Čtvrtý paragraf obsahuje poznámky k postupům v předchozích paragrafech, o odhadech bodu změny a dalších testech.

## 2 Detekce změny posunutí

### 2.1 Testy při normálním rozdělení

Předpokládejme, že $X_1, \ldots, X_n$ jsou nezávislé veličiny, $X_i$ má rozdělení $N(\theta_i, 1)$ $i = 1, \ldots, n$. V tomto případě lze formulovat jako úlohu (1.1) s $p = 1$, $c_{i1} = 1$ nebo úlohu (1.2).

Bayesovské řešení problému (bod změny $\tau$ je náhodná veličina s rozdělením $P(\tau \in (t_m, t_{m+1})) = p_m$, $m = 1, \ldots, n-1$, $p_n = P(\tau > t_n)$) a pro jednostrannou alternativu, tj. pro úlohu

(2.1 a) $\quad H_0 : \theta_1 = \cdots = \theta_n = \theta_0$ proti $H_1 : \exists_{m<n} \theta_0 = \theta_1 = \cdots = \theta_m \neq \theta_{m+1} = \cdots = \theta_n,$

navrhli Chernoff a Zacks (1964) a Gardner (1969) pro dvoustrannou alternativu:

(2.1 b) $\quad H_0 : \theta_1 = \cdots = \theta_n = \theta_0$ proti $H_1 : \exists_{m<n} \theta_0 = \theta_1 = \cdots = \theta_m \neq \theta_{m+1} = \cdots = \theta_n,$

Pro jednostrannou alternativu (při rovnoměrném apriorním rozdělení $\tau$ tj. $p_m = 1/n$, $m = 1, \ldots, n$) má testová statistika tvar

$$T_{n1} = \sum_{i=1}^{n-1} (i+1)(X_i - \theta_0) \quad \text{při } \theta_0 \text{ známém} \tag{2.2}$$

$$T_{n2} = \sum_{i=1}^{n-1} (i+1)(X_i - \overline{X}_n) \quad \text{při } \theta_0 \text{ neznámém,} \tag{2.3}$$

Při sledování vývoje ekonomiky nás zajímá, zda došlo ke změně strukturálních parametrů či nikoli. Známé jsou aplikace v ekonomických časových řadách (Hsu (1979)), vývoj nezaměstnanosti v Německu, vývoj cen na trhu (Deshayes a Picard (1986)). Dále jsou známy aplikace na hydrologické časové řady (Cobb (1978)), aplikace v biologii (změny chování zvířat – Dienske a Metz (1977), vývoj stavu zvířat Deshayes a Picard (1986)), v medicíně (detekce změny výsledků opakovaně prováděných vyšetření) a použití při rozpoznávání obrazců.

Pro řešení výše zformulovaných úloh (podobně jako u většiny statistických úloh) byly použity různé přístupy. V zásadě se dají roztřídit do několika skupin:

## A. Sekvenční a nesekvenční přístup.

Při sekvenčním přístupu vycházíme z posloupnosti nezávislých náhodných veličin $\{X_j\}_{j=1}^{\infty}$, pozorovaných v časových okamžicích $t_1 < t_2 \cdots, X_i$ má distribuční funkci $F_i$. Zajímá nás, zda existuje $m_0$ přirozené takové, že $F_1 = \cdots = F_{m_0} \neq F_{m_0+1} + \cdots$. V tomto případě je rozsah výběru $N$ náhodný a volba testového kritéria úzce souvisí s volbou pravidla o zastavení (ukončení) výběru (stopping rule). Je–li $m_0$ konečné, požadujeme, aby $N - m_0$ bylo co nejmenší (tj. chceme odhalit změnu co nejdříve), je–li $m_0 = +\infty$ (tedy ke změně v konečném čase nedojde), pak by pravděpodobnost, že $N < +\infty$ měla být malá. Popis testů spolu s dalšími podrobnostmi lze najít v pracech Širjajeva (1963), (1978). V dalším se tímto obecným případem zabývat nebudeme.

Budeme se však zabývat tzv. kvazisekvenčními postupy, při kterých je předem pevně stanoven maximální možný počet pozorování $n_0$. Po provedení $i$–tého ($i < n_0$) pozorování se na základě $i$ pozorování (dosud provedených) rozhodneme buď pro platnost alternativní hypotézy $H_1$ (popř. $H_1^*$) a ukončíme výběr, nebo pro další pozorování. Provedeme–li však $n_0$ pozorování, rozhodneme se na základě $n_0$ pozorování pro $H_0$ nebo $H_1$ (popř. $H_0^*$ nebo $H_1^*$). Tyto postupy úzce souvisí s tzv. rekurzivními postupy založené na rekurzivních reziduích. Testová statistika je zadána rekurentně a v $i$–tém kroku její hodnota spočteme z její hodnoty v $(i-1)$ním kroku a $X_i$. Tyto postupy nalezly široké uplatnění v praxi.

Při nesekvenčním postupu není rozsah výběru náhodný. V tomto případě máme k dispozici výsledky všech $n$ pozorování a chceme vědět, zda změna nastala a popř. kdy k ní došlo.

## B. Bayesovský a nebayesovský přístup

Při nebayesovském přístupu považujeme bod změny $\tau$ za neznámou konstantu, zatímco při bayesovském považujeme $\tau$ za náhodnou veličinu popsanou rozděleními $P(\tau \in \langle t_m, t_{m+1}\rangle) = p_m$, $m = 1, \ldots, n-1$ a $p_n = P(\tau > t_m)$. Tohoto přístupu se využívá, pokud máme apriorní informaci, kdy může nastat změna.

## C. Klasický a robustní přístup

Klasickým přístupem budeme rozumět případ, kdy se předpokládá znalost distribuční funkce $F_i$ až na neznámé parametry a většinou se používá testová statistika odvozená metodou podílem věrohodností popř. nějaké její modifikace. Často se předpokládá, že kde $\overline{X}_n = n^{-1}\sum_{i=1}^{n} X_i$. Obě statistiky mají při nulové i alternativní hypotéze normální rozdělení. Kritické oblasti s hladinou $\alpha$ mají tvar

$$T_{n1} > \Phi^{-1}(1-\alpha)\left(\sum_{i=1}^{n-1}(i+1)^2\right)^{\frac{1}{2}} \tag{2.4}$$

a

$$T_{n2} > \Phi^{-1}(1-\alpha)\left(\sum_{i=1}^{n-1}(i+1)^2 - \frac{(n+1)}{n^2}\sum_{j=1}^{n-1}(j+1)^2\right)^{\frac{1}{2}} \tag{2.5}$$

kde $\Phi^{-1}(\cdot)$ je kvantilová funkce normovaného normálního rozdělení. Pro oboustrannou alternativu můžeme testovou statistiku vyjádřit následovně:

$$T_{n3} = \sum_{t=1}^{n-1} p_t \left(\sum_{j=t}^{n-1}(X_{j+1} - \overline{X}_n)\right)^2 = \sum_{t=1}^{n-1} p_t (n-t)^2 \left(\overline{X}^*_{n-t} - \overline{X}_n\right)^2, \tag{2.6}$$

kde $p_m = P\left(\tau \in (t_m, t_{t+1})\right)$, $m = 1, \ldots, n-1$ a

$$\overline{X}^*_{n-t} = (n-t)^{-1}\sum_{j=n-t+1}^{n} X_j. \tag{2.7}$$

Gardner (1969) odvodil přesné i asymptotické rozdělení $T_{n3}$ při $H_0$ a ukázal, že $6nT_{n3}(n^2-1)^{-1}$ má asymptotické rozdělení jako $\sum_{k=1}^{\infty} 6U_k^2(\pi k)^{-2}$, kde $\{U_k\}_{k=1}^{\infty}$ jsou nezávislé náhodné veličiny s normovaným normálním rozdělením.

Metoda podílem věrohodností vede pro jednostrannou alternativu k testové statistice (2.8)

$$T_{n4} = \max_{1\le t\le n-1}\left\{\left(\overline{X}^*_{n-t} - \overline{X}_t\right)(t(n-t)/n)^{\frac{1}{2}}\right\} = \max_{1\le t\le n-1}\left\{-\sum_{i=1}^{t}\left(X_i - \overline{X}_n\right)/(t(n-t)/n)^{\frac{1}{2}}\right\}.$$

Pro oboustrannou alternativu Hawkins (1977) navrhl testovou statistiku

$$T_{n5} = \max_{1\le t\le n-1}\left\{\left|\sum_{i=1}^{t}\left(X_i - \overline{X}_n\right)\right|/(t(n-t)/n)^{\frac{1}{2}}\right\}. \tag{2.9}$$

a odvodil rekurentní vztahy pro rozdělení $T_{n5}$.

Známé jsou postupy založené na tzv. rekurzivních reziduích

$$W_k = \frac{X_k - \overline{X}_{k-1}}{\operatorname{var}(X_k - \overline{X}_{k-1})} = (X_k - \overline{X}_{k-1})(1-k^{-1})^{-\frac{1}{2}} \qquad k = 2, 3, \ldots, n, \tag{2.10}$$

která jsou nezávislá a všechna mají rozdělení $N(0,1)$. Při metodě známé pod názvem CUSUM (cumulative sums of recursive reziduals) pak postupně počítáme

$$M_t = \sum_{r=2}^{t} W_r = M_{t-1} + W_t, \qquad t = 2, \ldots, n, \tag{2.11}$$

$$M_1 = W_1 = 0.$$

Test pro jednostrannou alternativu (2.1 a) provádíme následovně: děláme postupně pozorování, po každém spočteme veličinu $M_t$ a

(2.12) je-li $M_t < d^+_{\alpha,c}(t)$ a $t < n$ děláme další pozorování,
je-li $M_t \geq d^+_{\alpha,c}(t)$, zamítneme $H_0$ a neděláme další pozorování, při $M_n < d^+_{\alpha,c}(t)$ přijmeme $H_0$.

Zde $d^+_{\alpha,c}(t)$ jsou kritické hodnoty určené tak, aby

$$P\left(\exists_{2\leq t<n} M_t \geq d^+_{\alpha,c}(t) \,|\, H_0\right) = \alpha. \tag{2.13}$$

Pro oboustrannou alternativu postupně děláme pozorování a

(2.14) je-li $|M_t| < d^+_{\alpha,c}(t)$ a $t < n$ děláme další pozorování,
je-li $|M_t| \geq d^+_{\alpha,c}(t)$, zamítneme $H_0$ a neděláme další pozorování, při $|M_n| < d^+_{\alpha,c}(t)$ přijeme $H_0$.

Hodnoty $d_{\alpha,c}(t)$ jsou určeny tak, aby

$$P\left(\exists_{2\leq t<n} |M_t| \geq d_{\alpha,c}(t) \,|\, H_0\right) = \alpha. \tag{2.15}$$

Touto rovnicí nejsou hodnoty $d_{\alpha,c}(t)$ určeny jednoznačně. Můžeme je např. (dle Brown a kol. (1975)) aproximovat následovně:

$$d_{\alpha,c}(t) \cong h_\alpha \sqrt{n-1}(2t+n-6)/(n-2), \qquad t = 2, \ldots, n, \tag{2.16}$$

kde $h_\alpha$ je určeno vztahem:

$$1 - \Phi(3h_\alpha) = \Phi(h_\alpha)\exp\{-4h^2_\alpha\} = \alpha/2. \tag{2.17}$$

Speciálně dostáváme $h_{0,01} = 1,143$; $h_{0,05} = 0,948$; $h_{0,1} = 0,850$. O dalších možnostech je poznámka v závěru paragrafu.

Dalším rekurzivním postupem je tzv. metoda MOSUM (moving sums of recursive reziduals). Postup se shoduje s CUSUM, jen místo veličin $M_t$ se používají veličiny

$$R_t = \sum_{j=t-G+1}^{t} W_j G^{-\frac{1}{2}}, \qquad t = 1+G, \ldots, n, \tag{2.18}$$

kde $G$ pevně volíme (Hackl doporučuje $5 \leq G \leq 20$) a místo kritických hodnot $d^+_{\alpha,c}(t)$ (popř. $d_{\alpha,c}(t)$ u dvoustranné alternativy) se použije $d^+_{\alpha,M}$ (popř. $d_{\alpha,M}$) odpovídající tomuto postupu.

Toto jsou nejpoužívanější testy pro problémy (2.1 a) a (2.1 b).

Pokud se týče kritických hodnot pro uvedené testy, nejjednodušší je situace pro bayesovské testy založené na $T_{n1}$ a $T_{n2}$, neboť obě statistiky mají normální rozdělení (kritické obory jsou popsány vzorci (2.4) a (2.5)). Pro $T_{n3}$ je lepší využít asymptotické rozdělení. Pokud se týče testů založených na $T_{n4}$, $T_{n5}$ lze použít Bonferroniho nerovnost a pak dostaneme kritickou oblast ($\overline{X}_t - \overline{X}^*_{n-t}$ má při $H_0$ rozdělení $N(0, t^{-1} + (n-t)^{-1})$.

$$T_{n4} > \Phi^{-1}\left(1 - \alpha(n-1)^{-1}\right) \tag{2.19}$$

$$T_{n5} > \Phi^{-1}\left(1 - \alpha(2(n-1))^{-1}\right). \tag{2.20}$$

Další možnost je použít Šidákovu nerovnost, která dává kritickou oblast tvaru:

$$T_{n5} > \Phi^{-1}\left(\left(1+(1-\alpha)\right)^{(n-1)^{-1}}/2\right). \tag{2.21}$$

Kritické nerovnosti se dají určit též na základě tzv. vylepšené Bonferroniho nerovnosti:

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i) - \sum_{i=1}^{n-1} P\left(A_{i+1} \cap A_i\right), \tag{2.22}$$

kde $A_1, \ldots, A_n$ jsou libovolné jevy (Worsley (1982)).

Pro výpočet kritických hodnot lze též použít limitní rozdělení při $H_0$, pro které platí:

$$\mathcal{L}\left(\max_{1\leq k\leq n}\left\{n^{-\frac{1}{2}}\left|\sum_{i=1}^{k}(X_i = \overline{X}_n)\right| g(k/n)\right\}\right) \longrightarrow \mathcal{L}\left(\sup_{t\in(0,1)}\{|B(t)|g(t)\}\right) \text{ pro } n\to\infty, \tag{2.23}$$

kde $\{B(t),)t \in (0,1)\}$ je Brownův most a $g$ je nezáporná funkce na (0,1) integrovatelná se čtvercem. Je vidět, že statistiky $T_{n4}$ a $T_{n5}$ nejsou příslušného tvaru $\left(\int_0^1 (u(1-u))^{-1}du = +\infty\right)$. Je však možné použít buď test s kritickou oblastí

$$\max_{1\leq k\leq n}\left\{-n^{-\frac{1}{2}}\sum_{i=1}^{k}(X_i - \overline{X}_n)\right\} > b_\alpha^+ \tag{2.24}$$

pro jednostrannou alternativu, nebo

$$\max_{1\leq k\leq n}\left\{n^{-\frac{1}{2}}\left|\sum_{i=1}^{k}(X_i - \overline{X}_n)\right| g(k/N)\right\} > b_\alpha(g) \tag{2.25}$$

pro oboustrannou, kde $b_\alpha^+$ a $b_\alpha(g)$ jsou definovány vztahy

$$P\left(\sup_{t\in(0,1)}\{B(t)\} > b_\alpha^+\right) = \alpha, \tag{2.26}$$

$$P\left(\sup_{t\in(0,1)}\{|B(t)|g(t)\} > b_\alpha(g)\right) = \alpha. \tag{2.27}$$

Tabulka hodnot $b_\alpha^+$ a $b_\alpha(g)$; $g \equiv 1$ je připojena na konci článku.

James a kol. (1987) odvodili aproximaci

$$\begin{aligned} P\left(\max_{t_0\leq t\leq t_1}\left\{|B(t)|/(t(1-t))^{-\frac{1}{2}}\right\} \geq b\right) &= \\ = \sqrt{2}\frac{b\exp\{-b^2/2\}}{\Gamma(\frac{1}{2})}&\left\{\frac{1}{2}(1-b^{-2})\log\frac{(1-t_0)t_1}{t_0(1-t_1)} + 2b^{-2} + o(b^{-2})\right\} \end{aligned} \tag{2.28}$$

pro $b \to \infty$, kde $0 < t_0 < t_1 < 1$, což spolu s (2.23) vede k testu s kritickou oblastí

$$\max_{k_0\leq k\leq k_1}\left\{\left|\sum_{i=1}^{k}(X_i - \overline{X}_N)\right| \left(N^{-1}\left(\sum_{i=1}^{k}(X_i - \overline{X}_k)^2 + \sum_{i=k+1}^{n}(X_i - \overline{X}^*_{n-k})^2\right)^{\frac{1}{2}}\right)\right\} \geq b_\alpha$$

kde $b_\alpha$ je řešení rovnice, kterou získáme, jestliže pravou stranu v (2.28) položíme rovnou $\alpha$ a $[nt_0] = k_0$, $[nt_1] = k_1$.

Pro metodu CUSUM a oboustrannou alternativu je možné použít buď kritické hodnoty dané (2.16), nebo opět použít Bonferroniho nebo Šidákovu nerovnost, které dávají

(2.29) $$d_{\alpha,c}(t) \cong \sqrt{t-1}\Phi^{-1}\left(1 - \frac{\alpha}{2(n-1)}\right) \qquad t = 2, \ldots, n$$

respektive

(2.30) $$d_{\alpha,c}(t) \cong \sqrt{t-1}\Phi^{-1}\left(1 + (1-\alpha)^{(n-1)^{-1}}/2\right) \qquad t = 2, \ldots, n.$$

Pro jednostrannou alternativu Bonferroniho nerovnost dává:

$$d_{\alpha,c}(t) \cong \sqrt{t-1}\Phi^{-1}\left(1 - \frac{\alpha}{n-1}\right).$$

Dále lze využít limitní rozdělení. Při $H_0$ totiž platí

(2.31) $$\mathcal{L}\left(\max_{2\le t\le n}\left\{\frac{1}{\sqrt{n}}|M_t|\right\}\right) \longrightarrow \mathcal{L}\left(\max_{0<t<1}\{|W(t)|\}\right),$$

(2.32) $$P\left(\max_{2\le t\le n}\left\{|M_t|t^{-\frac{1}{2}}\right\}\right)(2\log\log n)^{-\frac{1}{2}} - 2\log\log n -$$
$$-\frac{1}{2}\log\log\log n + \frac{1}{2}\log(4\pi) \le y = \exp\{-2\exp\{-y\}\}, \qquad y \in R_1,$$

kde $\{W(t),\ t \in (0,1)\}$ je normovaný Wienerův proces. Kritické oblasti pak mají tvar

(2.33) $$\max_{2\le t\le N}\{|M_t|\} \ge n^{\frac{1}{2}} w_\alpha,$$

respektive

(2.34) $$\max_{2\le t\le N}\left\{|M_t|t^{-\frac{1}{2}}\right\}(2\log\log n)^{\frac{1}{2}} - 2\log\log n -$$
$$-\frac{1}{2}\log\log\log n + \frac{1}{2}\log(4\pi) > -\log\log(1-\alpha)^{-\frac{1}{2}},$$

kde $w_\alpha$ je definováno následovně

(2.35) $$P\left(\sup_{t\in(0,1)}\{|W(t)|\} > w_\alpha\right) = \alpha.$$

Kritické hodnoty (2.29), (2.30) vedou ke konzervativním testům, test s kritickou oblastí (2.33) není citlivý vůči změně při malém $m$ (tj., $m/n$ je malé). Konvergence (2.32) je velmi pomalá, takže jako nejvhodnější při nepříliš velkém $n$ se jeví test s kritickou oblastí (2.16).

Pro metodu MOSUM plyne z Bonferroniho a Šidákovy nerovnosti následující aproximace:

(2.36) $$d_{\alpha,M}(t) \cong \Phi^{-1}\left(1 - \frac{\alpha}{2(n-G)}\right) \qquad t = 1 + G, \ldots, m$$

respektive

(2.37) $$d_{\alpha,M}(t) \cong \Phi^{-1}\left(1+(1-\alpha)^{(n-G)^{-1}}/2\right) \qquad t = 1+G,\ldots,n.$$

Limitní rozdělení $\max_{G<k\leq n}|R_t|$ vede k testu s kritickou oblastí

(2.38)
$$\max_{G<t\leq n}|R_t| > \left(2\log\frac{n}{G}\right)^{\frac{1}{2}} + \left(\log\log\frac{n}{G} + \log(4/\pi) - 2\log\log(1-\alpha)^{-1}\right)\left(8\log\log\frac{n}{G}\right)^{-\frac{1}{2}}.$$

Je nutné si uvědomit, že kritické hodnoty získané přes Bonferroniho a Šidákovu nerovnost vedou ke konzervativním testům (tj. skutečná hladina testu je nižší než $\alpha$), ale simulace ukazují, že skutečné kritické hodnoty jsou sice menší, ale rozdíl je přijatelný (podrobněji viz Hackl (1980)).

Ve všech případech lze získat kritické hodnoty simulacemi.

Testy lze celkem snadno modifikovat na případ, kdy $X_i$ mají rozdělení $N(\theta_i,\sigma^2)$, kde $\sigma^2 > 0$ neznáme, stačí nahradit $\sigma^2$ odhadem, což vede ke změně rozdělení exaktního, ale ne limitního. Analogicky lze sestavit testy pro změnu rozptylu, popř. současné změny posunutí a rozptylu. Např. jsou-li $X_1,\ldots,X_n$ nezávislé, $X_i$ má rozdělení $N(\theta_0,\sigma_i^2)$, $i = 1,\ldots,n$ a uvažujeme-li úlohu testovat hypotézu $H_0^\sigma : \sigma_1^2 = \cdots = \sigma_n^2$ proti $H_1^\sigma : \exists_{m<n}\sigma_1^2 = \cdots = \sigma_M^2 \neq \sigma_{m+1}^2 = \cdots = \sigma_n^2$, pak CUSUM test pro změnu rozptylu je založen na statistikách $\sum_{r=2}^{t} W_r^2((n-1)^{-1}\sum_{k=2}^{n} W_k^2)^{-1}$, $t = 2,\ldots n$. Podrobnosti o těchto postupech lze najít v Hacklovi (1980).

## 2.2 Robustní postupy

V tomto paragrafu budeme vycházet z předpokladu, že náhodné veličiny $X_i$ jsou nezávislé, $X_i$ má distribuční funkci $F(x-\theta_i)$, $i = 1,\ldots,n$. Budeme uvažovat jen dvoustrannou alternativu

(2.35) $$H_0 : \theta_1 = \cdots = \theta_n = \theta_0 \text{ proti } H_1 : \exists_{m<n}\theta_0 = \theta_1 = \cdots = \theta_m \neq \theta_{m+1} = \cdots = \theta_n.$$

Modifikace pro jednostrannou alternativu plyne z předchozího a postupů pro dvoustrannou alternativu. Při sestavování robustních testů pro tuto úlohu vycházeli autoři většinou z testů uvedených v předchozích paragrafech. Největší pozornost byla věnována testům založeným na pořadích.

Bayesovské testy založené na pořadí navrhli Bhattacharya a Johnson (1968). Pro případ $\theta_0$ známého a $F$ spojité symetrické kolem 0 má testová charakteristika tvar

(2.36)
$$T_{n6} = \sum_{i=1}^{n} D_i \operatorname{sign}(X_i-\theta_0)a_n^+(R_i^+) = \\ = \sum_{i=1}^{n} p_i \sum_{k=1}^{i} \operatorname{sign}(X_k-\theta_0)a_n^+(R_k^+)$$

kde $(p_1,\ldots,p_n)$ je apriorní rozdělení náhodné veličiny $\tau$ (bod změny), $D_k = \sum_{i=k}^{n} p_i$; $i = 1,\ldots,n$, $R_i^+$ je pořadí $|X_i-\theta_0|$ mezi $|X_1-\theta_0|,\ldots,|X_n-\theta_0|$ a $(a_n^+(1),\ldots,a_n^+(n))$ je vektor skórů užívaných pro test hypotézy symetrie. Za povšimnutí stojí fakt, že statistika $T_{n6}$ je lineární kombinací pořadových statistik pro test hypotézy symetrie při rozsahu výběru

$i$, $i = 1, \dots, n$. Často se používá test s $a_n^+(i) = 1$ (tedy obdoba znaménkového testu). Pro případ $\theta_0$ neznámého a $F$ spojité (ne nutně symetrické) navrhli autoři testovou statistiku

$$T_{n7} = \sum_{i=1}^{n} D_i a_n(R_i) = \sum_{i=1}^{n} p_i \sum_{k=1}^{i} a_n(R_k), \tag{2.37}$$

kde význam $D_i$ a $p_i$ je stejný jako u předchozího testu, $R_i$ je pořadí $X_i$ mezi $X_1, \dots, X_n$, $i = 1, \dots, n$ $(a_n(1), \dots, a_n(n))$ je vektor skórů pro test hypotézy náhodnosti (tj. $X_j$, $j = 1, \dots, n$ mají distribuční funkci $F(x - \theta_0)$ proti alternativě $X_1, \dots, X_i$ mají distribuční funkcí $F(x - \theta_0)$ a $X_{i+1}, \dots, X_n$ mají distribuční funkci $F(x - \theta_n)$, $\theta_0 \neq \theta_n$). Optimální volby skórů se v obou případech shodují s optimální volbou pro test hypotézy symetrie popř. test hypotézy náhodnosti. Pokud se týče kritických hodnot (pro jedno i oboustrannou alternativu), stačí si uvědomit, že $T_{n6}$ a $T_{n7}$ jsou opět lineární pořadové statistiky a tedy lze použít kritické hodnoty pro příslušné pořadové testy založené na exaktním rozdělení nebo limitním rozdělení.

Vyjdeme-li z $M$-odhadů (při distribuční funkci $F$ symetrické), můžeme sestavit následující bayesovské testové statistiky

$$T_{n8} = \sum_{i=1}^{n} D_i \psi(X_i - \theta_0) \tag{2.38}$$

a

$$T_{n9} = \sum_{i=1}^{n} D_i \psi(X_i - \theta_n(\psi)) \tag{2.39}$$

při případ $\theta_0$ známého respektive neznámého, kde $\theta_n(\psi)$ je $M$-odhad $\theta_0$ generovaný funkcí $\psi, \psi$ je monotónní lichá funkce splňující obvyklé požadavky. Pro nalezení kritických hodnot se obvykle využívá limitní rozdělení (obě testové statistiky mají při $H_0$ asymptoticky normální rozdělení s parametry $(0, \sum_{i=1}^{n} D_i^2 \int \psi^2(x) dF(x))$.

Maximálně věrohodnému přístupu zde odpovídá tzv. pseudodvouvýběrový pořadový test, jehož idea je zřejmá z testové statistiky:

$$T_{n10}(g) = \max_{2 \le k \le n} \left\{ \left| \sum_{i=1}^{k} (a_n(R_i) - \bar{a}_n) \right| g(k/n) \right\}, \tag{2.40}$$

kde $\bar{a}_n = \sum_{i=1}^{n} a_n(i) n^{-1}$ a za $g$ volili autoři:

$$g_1(k/n) = \left( \mathrm{var}_{H_0} \left\{ \sum_{i=1}^{k} a_n(R_i) \right\} \right)^{-\frac{1}{2}}, \quad k = 2, \dots, n \tag{2.41}$$

(Sen A. a Srivastava (1975)) nebo

$$g_2(k/n) = \left( \mathrm{var}_{H_0} \left\{ \sum_{i=1}^{n/2} a_n(R_i) \right\} \right)^{-\frac{1}{2}}, \quad k = 2, \dots, n \tag{2.42}$$

(Pettit (1979)) nebo funkce tzv. $U$-tvaru $g(t) = (t(1-t))^{-\beta}$, $0 \le \beta < \frac{1}{2}$ (Sen P. K. (1978)). Kritické hodnoty lze nalézt buď simulacemi, nebo využitím limitního rozdělení, a to statistik $\sum_{i=1}^{k}(a_n(R_i)-\bar{a}_n)g(k/n)$, $k=2,\dots,n-1$ (pro $k\to\infty$ a $n-k\to\infty$ mají asymptoticky normální rozdělení) spolu s Bonferroniho nebo Šidákovou nerovností. Další možnost je použít tvrzení:

(2.43)
$$\mathcal{L}\left(T_{n10}(g)\left(\sum_{i=1}^{n}(a_n(i)-\bar{a}_n)^2\right)^{-\frac{1}{2}}\right) \longrightarrow \mathcal{L}\left(\sup_{t\in(0,1)}\{|B(t)|\,g(t)\}\right) \quad \text{pro } k\to\infty,$$

kde $\{B(t),\, t\in(0,1)\}$ je Brownův most a $g$ nezáporná funkce na $(0,1)$ integrovatelná se čtvercem.

Zcela analogicky lze sestavit testy založené na $M$–odhadech. Testová statistika má tvar

(2.44)
$$T_{n11}(g)=\max_{2\le k\le n}\left\{\left|\sum_{i=1}^{k}(\psi(X_i-\theta_n(\psi))\,g(k/n)\right| \left\{\sum_{i=1}^{n}\psi^2(X_i-\theta_n(\psi))\right\}^{-\frac{1}{2}}\right\},$$

kde $\theta_n(\psi)$ je $M$–odhad založený na $X_1,\dots,X_n$ generovaný funkcí $\psi$. Pokud se týče funkce $g$ a kritických hodnot, platí totéž co pro pořadové testy.

Robustní postup odpovídající CUSUM metodě založený na $M$–odhadech vychází ze statistik

(2.45)
$$W_i^*=\psi\left(X_i-\theta_{i-1}(\psi)\right), \qquad i=2,\dots,n$$

(2.46)
$$\sigma_k^{*2}=k^{-1}\sum_{i=1}^{k}\psi^2\left(X_i-\theta_{i-1}(\psi)\right), \qquad k=2,\dots,n$$

(2.47)
$$M_k^*=W_k^*+M_{k-1}^*, \quad M_1^*=0, \qquad k=2,\dots,n$$

kde $\theta_{i-1}(\psi)$ je $M$–odhad založený na pozorováních $X_1,\dots X_{i-1}$ a generovaný funkcí $\psi$. Veličina $\sigma_k^{*2}$ je odhadem $\int\psi^2(x)dF(x)$. Test lze popsat následovně

(2.48) je–li na $k$–tém kroku $n^{-\frac{1}{2}}|M_k^*|\sigma_k^{*-1}\ge f_c(\alpha,k)$ $k\le n$, pak zamítneme $H_0$ a neděláme další pozorování,
je–li $n^{-\frac{1}{2}}|M_k^*|\sigma_k^{*-1}<f_c(\alpha,k)$ a $k<n$, děláme další pozorování,
je–li $k=n$ a $n^{-\frac{1}{2}}|M_k^*|\sigma_k^{*-1}\le f_c(\alpha,k)$ přijmeme $H_0$, jinak $H_0$ zamítneme,

kde $f_c(\alpha,k)$ jsou kritické hodnoty odpovídající hladině $\alpha$. Pro získání $f_c(\alpha,k)$ se dá využít limitní rozdělení, neboť při $H_0$ (a splnění určitých podmínek regularity) platí:

(2.49)
$$\mathcal{L}\left(\max_{2\le k\le n}\left\{\left(n\sigma_k^{*2}\right)^{-\frac{1}{2}}|M_k^*|\right\}\right)\longrightarrow\mathcal{L}\left(\sup_{t\in(0,1)}\{|W(t)|\}\right) \quad \text{při } n\to\infty,$$

kde $\{W(t);\ t\in(0,1)\}$ je normovaný Wienerův proces. Platí též (2.32), kde $M_t$ je nahrazen $M_k^*\sigma_k^{*-\frac{1}{2}}$. Další podrobnosti lze nalézt v Hušková (1990a). Souvislost s CUSUM vyloženou v předchozím paragrafu je vidět, položíme-li $\psi(x) = x$ a $\theta_k(\psi) = \overline{X}_k$ (pozor v 2.1 jsme znali var $\left(X_k - \overline{X}_{k-1}\right) = \sigma^2(k+1)/k$).

Zcela analogicky lze definovat robustní versi MOSUM. Pokud se týče kritických hodnot, lze použít Bonferroniho nebo Šidákovu nerovnost event. test s kritickou oblastí (2.37), kde $R_t$ je nahrazeno $\sum_{i=k-G}^{k} W_i^* \sigma_k^{*-1}$, ale v tomto případě potřebujeme $G$ i $n$ velké (kvůli asymptotické normalitě).

Z výpočetního hlediska je vhodnější v robustních verzích CUSUM a MOSUM použít místo $M$–odhadů (ty se totiž počítají pomocí iteračních metod) tzv. rekurzivní $M$–odhady nebo jednokrokové rekurzivní $M$–odhady. Rekurzivní $M$–odhady jsou definovány následovně:

$$\theta_i^0(\psi) = \theta_{i-1}^0(\psi) + (i\gamma_i^0)^{-1}\psi\left(X_i - \theta_{i-1}^0(\psi)\right) \qquad i = 2,\ldots,n, \tag{2.50}$$

$$\gamma_i^0 = (2ti)^{-1}\sum_{j=2}^{i}\left(\psi\left(X_j - \theta_{j-2}^0(\psi) + tj^{-\beta}\right) - \psi\left(X_j - \theta_{j-1}^0(\psi) - tj^{-\beta}\right)\right)j^{\beta} \tag{2.51}$$

kde $\theta_1^0(\psi)$ je rozumný počáteční odhad, $0 < \beta \le \frac{1}{2}$, $t \neq 0$ pevné. Jednokrokový rekurzivní $M$–odhad je definován následovně

$$\theta_i^*(\psi) = \theta_{i-1}^*(\psi) + (i\gamma_i^*)^{-1}\sum_{j=2}^{i}\psi\left(X_j - \theta_{j-1}^*(\psi)\right), \qquad i = 2,\ldots,m \tag{2.52}$$

$$\gamma_i^* = \frac{1}{2t\sqrt{i}}\sum_{j=2}^{i}\left(\psi\left(X_j - \theta_i^*(\psi) + ti^{-\frac{1}{2}}\right) - \psi\left(X_i - \theta_i^*(\psi) - ti^{-\frac{1}{2}}\right)\right), \tag{2.53}$$

kde $t \neq 0$, $\theta_1^*(\psi)$ je rozumný počáteční odhad. Veličiny $\gamma_i^0$ a $\gamma_i$ jsou odhadem $\partial/\partial a\left(\int \psi(x-a)\,dF(x)\right)$.

V obou případech je nutné $\gamma_i^0$ a $\gamma_i^*$ při výpočtu modifikovat (ořezávat). Podrobněji se lze s těmito odhady seznámit v literatuře o rekurzivních odhadech nebo stochastických aproximacích (např. Novovičová (1984), Hušková (1990 c)). Pokud se týče volby skórové funkce $\psi$, platí pro ni stejné doporučení jako u $M$–odhadů parametru posunutí. Podrobnější poznámku nalezne čtenář na konci 3.2.

# 3 Detekce změn regrese

## 3.1 Postup pro normální rozdělení

Předpokládejme, že $X_1,\ldots,X_n$ jsou nezávislé veličiny, $X_i$ má normální rozdělení $N\left(\underset{\sim}{c}_i' \cdot \underset{\sim}{\theta}_i, \sigma_i^2\right)$ $i = 1,\ldots,n$.

Nejprve se zmíníme o různých modifikacích. Za právě uvedených předpokladů lze dále uvažovat úlohu (1.1) při $\sigma_i^2 = \sigma^2$, $i = 1,\ldots,n$ ($\sigma^2$ známé nebo neznámé) nebo studovat obecnější úlohu:

$$\begin{aligned} H_0^\sigma &: \left(\underset{\sim}{\theta}_1, \sigma_1^2\right) = \cdots = \left(\underset{\sim}{\theta}_n, \sigma_n^2\right) \qquad \text{proti} \\ H_1^\sigma &: \exists_{m<n} \left(\underset{\sim}{\theta}_1, \sigma_1^2\right) = \cdots = \left(\underset{\sim}{\theta}_m, \sigma_m^2\right) \neq \left(\underset{\sim}{\theta}_{m+1}, \sigma_{m+1}^2\right) = \cdots = \left(\underset{\sim}{\theta}_n, \sigma_n^2\right). \end{aligned} \tag{3.1}$$

Další možnost je uvažovat jen změnu rozptylu (tj. $\underset{\sim}{\theta}_1 = \cdots = \underset{\sim}{\theta}_n$ platí při alternativě). Zde se soustředíme na úlohu (1.1). Informace o ostatních lze nalézt v Hacklovi (1980).

Dále se někdy místo úlohy (1.1) uvažuje úloha nazývaná "time-trending regression" formulovaná jako diskriminace mezi hypotézami

$$
\begin{aligned}
(3.2)\qquad & H_{(0)} : E\,X_i = \underset{\sim}{c}_i'\underset{\sim}{\theta}_{(0)} && \forall i\\
& H_{(1)} : E\,X_i = \underset{\sim}{c}_i'\left(\underset{\sim}{\theta}_{(0)} + \underset{\sim}{\theta}_{(1)}t_i\right) && \forall i\\
& H_{(g)} : E\,X_i = \underset{\sim}{c}_i'\left(\underset{\sim}{\theta}_{(0)} + \underset{\sim}{\theta}_{(1)}t_i + \cdots \underset{\sim}{\theta}_{(g)}t_i^g\right) && \forall i
\end{aligned}
$$

kde všechna $\underset{\sim}{\theta}_{(v)}$, $(v = 1,\ldots,g)$ jsou $p$-rozměrné vektory, $t_1 < t_2 < \ldots$ jsou časové okamžiky, kdy se provádějí pozorování. Informace o úlohách tohoto typu lze nalézt např. v Farley a Hinich (1970). Při úloze (1.1) někdy předpokládáme, že v alternativě $E\,X_i = h(t_i)$, $i = 1,\ldots,n$ kde $h$ je nějaká funkce, a rozlišujeme, zda jde o změnu skokem ("abrupt change") nebo o spojitou změnu (tj. zda-li má $h(t)$ v bodě změny skok nebo je spojitá). Informace o tomto přístupu lze nalézt např. v Hawkins (1980).

Bayesovská řešení problému lze nalézt např. v Ferreira (1975), Esterby a El Shaarawi (1981), Choy a Broemling (1980) a Broemling a Tsurumi (1987). Testy jsou analogií bayesovských testů popsaných v 2.1. Autoři se však více soustřeďují na odhady než na testování.

Test podílem věrohodnosti pro úlohu (3.1) dává testovou statistiku (nazývanou Quandtovou):

$$(3.3)\qquad \max_{p<m<n-p}\left\{\frac{1}{2}\left(-m\ln\tilde{\sigma}_m^2 - (n-m)\ln\tilde{\sigma}_{n-m}^{*2} + n\ln\tilde{\sigma}_n^2\right)\right\},$$

kde $\tilde{\sigma}_m^2$, $\tilde{\sigma}_{n-m}^{*2}$, $\tilde{\sigma}_n^2$ jsou maximálně věrohodné odhady $\sigma^2$ odpovídající pozorováním $(X_1,\ldots,X_m)$, $(X_{m+1},\ldots,X_n)$, respektive $(X_1,\ldots,X_n)$. Problém je nalézt kritické hodnoty. Lepší je situace pro případ $\sigma^2 = \sigma_i^2$, $i = 1,\ldots,n$, známé. Hawkins (1977) a Worsley (1983) ukázali, že metoda podílem věrohodností vede k testové statistice

$$(3.4)\qquad \max_{p<k<n-p}\{V_{n,k}\},$$

$$
\begin{aligned}
(3.5)\qquad V_{n,k} &= \left(\hat{\underset{\sim}{\theta}}_k - \hat{\underset{\sim}{\theta}}^*_{n-k}\right)'\left(\underset{\sim}{C}_k^{-1} + \underset{\sim}{C}_{n-k}^{*-1}\right)^{-1}\left(\hat{\underset{\sim}{\theta}}_k - \hat{\underset{\sim}{\theta}}^*_{n-k}\right)\frac{1}{\sigma^2} =\\
&= \left(\hat{\underset{\sim}{\theta}}_k - \hat{\underset{\sim}{\theta}}_n\right)'\left(\underset{\sim}{C}_k^{-1} - \underset{\sim}{C}_n^{-1}\right)^{-1}\left(\hat{\underset{\sim}{\theta}}_k - \hat{\underset{\sim}{\theta}}_n\right)\frac{1}{\sigma^2}
\end{aligned}
$$

kde $\hat{\underset{\sim}{\theta}}_k$ a $\hat{\underset{\sim}{\theta}}_{n-k}$ jsou odhady $\underset{\sim}{\theta}$ metodou nejmenších čtverců založené na $(X_1,\ldots,X_k)$ respektive $(X_{k+1},\ldots,X_n)$ a $\underset{\sim}{C}_k = \sum_{i=1}^k \underset{\sim}{c}_i\underset{\sim}{c}_i'$, $\underset{\sim}{C}_{n-k} = \sum_{i=k+1}^n \underset{\sim}{c}_i\underset{\sim}{c}_i'$, $k = 1,\ldots n$.

Odhady $\hat{\underset{\sim}{\theta}}_k$ a $\hat{\underset{\sim}{\theta}}_{n-k}$ jsou nezávislé, mají normální rozdělení var $\left(\hat{\underset{\sim}{\theta}}_k - \hat{\underset{\sim}{\theta}}^*_{n-k}\right) = \sigma^2\left(\underset{\sim}{C}_k^{-1} + \underset{\sim}{C}_{n-k}^{-1}\right)$ a var $\left(\hat{\underset{\sim}{\theta}}_k - \hat{\underset{\sim}{\theta}}_n\right) = \left(\underset{\sim}{C}_k^{-1} - \underset{\sim}{C}_n^{-1}\right)\sigma^2$.

Bonferroniho nerovnost pak vede ke kritické hodnotě $\chi^2_{\alpha^*}(p)$, kde $\alpha^* = \alpha(n-2p)^{-1}$ a $\psi^2_{\alpha^*}(p)$ je 100 $\alpha^*$ % kritická hodnota $\chi^2$ rozdělení o $p$ st. v.

Za předpokladu, že

(3.6) $$n^{-1}\underset{\sim}{C}_{[ns]} \longrightarrow s\underset{\sim}{C} \quad \text{pro} \quad n \to \infty,\ s \in \langle 0,1\rangle$$

(3.7) $$n^{-1}\sum_{i=1}^{n} c_{ij}^4 = 0(1) \quad \text{pro} \quad n \to \infty,\ j = 1,\dots,p$$

(3.8) $$\max_{1\le i\le n}\left\{c_{ij}^4\right\} n^{-1} \longrightarrow 0 \quad \text{pro} \quad n \to \infty,\ j = 1;\dots p,$$

kde $\underset{\sim}{C}$ je pozitivně definitní matice, platí

(3.9) $$\mathcal{L}\left(\max_{p<k<n-p}\left\{V_{n,k}g(k/n)k(n-k)/n^2\right\}\right.$$
$$\longrightarrow \mathcal{L}\left(\sup_{t\in(0,1)}\left\{B_p^*(t)g(t)\right\}\right) \qquad \text{pro } n \to \infty$$

kde $B_p^*(t) = \sum_{i=1}^{p} B_i^2(p)$, $t \in (0,1)$, $\{B_i(t), t \in (0,1)\}$, $i = 1,\dots,p$ jsou nezávislé Brownovy mosty, $g$ je nezáporná funkce na $(0,1)$ integrovatelná se čtvercem. Chceme-li tedy využít limitní rozdělení, pak bychom měli použít test s kritickým oborem

(3.10) $$\max_{p<k<n-p}\left\{V_{n,k}k(n-k)/n^2\right\} > b_{\alpha,p}^*$$

popř. asymptoticky ekvivalentní test s kritickým oborem

(3.11) $$\max_{p<k<n-p}\left\{\left(\hat{\underset{\sim}{\theta}}_k - \hat{\underset{\sim}{\theta}}_n\right)' \underset{\sim}{C}_k \underset{\sim}{C}_n^{-1} \underset{\sim}{C}_k \left(\hat{\underset{\sim}{\theta}}_k - \hat{\underset{\sim}{\theta}}_n\right)\sigma^{-2}\right\} > b_\alpha^*,$$

kde $b_\alpha^*(p)$ je definováno následovně:

(3.12) $$P\left(\sup_{t\in(0,1)}\{B^*(t)\} > b_{\alpha,p}^*\right) = \alpha.$$

Tabulky $b_\alpha^*(p)$ jsou v tabulce na konci článku.

James a kol. (1987) dokázali, že

(3.13) $$P\left(\sup_{t_0<t<1-t_1}\left\{\sum_{i=1}^{p} B_i^2(t)\left(t(1-t)^{-1}\right)\right\}^{\frac{1}{2}} \ge b\right) =$$
$$= \frac{b^p \exp\{-b^2/2\}}{2^{(p-2)/2}\Gamma(p/2)}\left\{\frac{1}{2}(1-p/b^2)\log\frac{(1-t_1)(1-t_0)}{t_0 t_1} + \right.$$
$$\left. + \ 2b^{-2} + o(b^{-2})\right\} \qquad \text{pro } b \to \infty,$$

kde $\{B_i(t), t \in (0,1)\}$, $i = 1,\dots,p$, jsou nezávislé Brownovy mosty, což umožňuje snadno spočítat aproximaci pro kritickou hodnotu pro test založený na $\max_{nt_0\le k\le n(1-t_1)}\{V_{n,k}\sigma^{-2}\}$.

Metoda CUSUM (vyložená pro případ posunutí v 2.1) je asi nejznámější. Vychází z rekurzivních reziduí:

(3.14) $$W_k = \frac{X_k - \underset{\sim}{c}_k'\underset{\sim}{\theta}_{k-1}}{\left(\operatorname{var}(X_k - \underset{\sim}{c}_k'\underset{\sim}{\theta}_{k-1})\right)^{\frac{1}{2}}}$$
$$= \frac{X_k - \underset{\sim}{c}_k'\underset{\sim}{C}_{k-1}^{-1}(\underset{\sim}{c}_1,\dots,\underset{\sim}{c}_{k-1})\underset{\sim}{X}_{k-1}}{(\sigma^2(1+\underset{\sim}{c}_k'\underset{\sim}{C}_{k-1}^{-1}\underset{\sim}{c}_k))^{\frac{1}{2}}} \qquad k = p+1,\dots,n.$$

Pro výpočet $\underset{\sim}{\theta}_k$ a $\underset{\sim}{C}_k^{-1}$ lze použít následující rekurentní vzorce:

$$\begin{aligned} \underset{\sim}{C}_k^{-1} &= \underset{\sim}{C}_{k-1}^{-1} - \frac{\underset{\sim}{C}_{k-1}^{-1}\underset{\sim}{c}_k\underset{\sim}{c}_k'\underset{\sim}{C}_{k-1}^{-1}}{1+\underset{\sim}{c}_k'\underset{\sim}{C}_{k-1}^{-1}\underset{\sim}{c}_k} \\ \theta_k &= \underset{\sim}{\theta}_{k-1} + \underset{\sim}{C}_k^{-1}\underset{\sim}{c}_k\left(X_k - \underset{\sim}{c}_k'\underset{\sim}{\theta}_{k-1}\right). \end{aligned}$$

Z rekurzivních reziduí počítáme postupně

$$M_t = M_{t-1} + W_t \qquad t = p+1,\dots,n \tag{3.15}$$

$W_p = 0$ a test provádíme dle (2.14) přičemž za $M_t$ a $W_t$ dosazujeme z (3.14) a (3.15). Kritické hodnoty lze aproximovat stejně jako pro posunutí (tj. dle (2.28) a (2.20) nebo (2.16)).

Metoda MOSUM pro regresní model se shoduje s MOSUM pro posunutí, přičemž do (2.18) dosazujeme za $W_j$ z (3.14). Další podrobnosti lze nalézt v Hacklovi (1980).

## 3.2 Robustní přístup

V tomto paragrafu budeme uvažovat problém (1.1), přičemž budeme předpokládat, že distribuční funkce $F$ splňuje nějaké podmínky regularity (obvykle konečnost Fisherovy informace) a jinak je neznámá a pro limitní rozdělení platnost (3.6 - 3.8).

Uvedeme si zde testy založené na $M$–odhadech, které jsou "odvozeny" z testů uvedených v předchozím paragrafu. (Testy založené na pořadích jsou příliš komplikované a pro praktické využití nejsou atraktivní).

Vyjdeme–li z testu (3.4 - 3.5) (tj. z testu podílem věrohodností), pak můžeme zkonstruovat robustní testy založené na statistice (všechny jsou asymptoticky ekvivalentní)

$$\max_{p<k<n-p}\{Z_{n,k}\} \tag{3.16}$$

kde za $Z_{n,k}$ lze dosadit např. jednu z následujících statistik

$$Z_{n,k}^{(1)} = \left(\underset{\sim}{\theta}_k(\psi) - \underset{\sim}{\theta}_n(\psi)\right)'\underset{\sim}{C}_k\underset{\sim}{C}_n^{-1}\underset{\sim}{C}_k\left(\underset{\sim}{\theta}_k(\psi) - \underset{\sim}{\theta}_n(\psi)\right)\hat{\sigma}^{-2}, \tag{3.17 a}$$

$$Z_{n,k}^{(2)} = \left(\sum_{i=1}^{k}\underset{\sim}{c}_i\psi\left(X_i - \underset{\sim}{c}_i'\underset{\sim}{\theta}_n(\psi)\right)\right)'\underset{\sim}{C}_n^{-1}\left(\sum_{i=1}^{k}\underset{\sim}{c}_i\psi\left(X_i - \underset{\sim}{c}_i'\underset{\sim}{\theta}_n(\psi)\right)\right)\hat{\sigma}^{-2}, \tag{3.17 b}$$

pro $p < k < n-p$, kde $\underset{\sim}{\theta}_k(\psi)$ je $M$–odhad generovaný funkcí $\psi$ založený na $X_1,\dots,X_k$ a

$$\hat{\sigma}^2 = n^{-1}\sum_{i=1}^{n}\psi^2\left(X_i - \underset{\sim}{c}_i'\underset{\sim}{\theta}_n(\psi)\right), \tag{3.18}$$

což je odhad $\int\psi^2(x)\,dF(x)$. Jestliže platí (3.6 - 3.8) a $\psi$ je lichá monotónní s ohraničenou první derivací, pak při $H_0$

$$\mathcal{L}\left(\max_{p<k<n-p}\left\{Z_{n,k}^{(i)}\right\}\right) \longrightarrow \mathcal{L}\left(\sup_{t\in(0,1)}\left\{B_p^*(t))\right\}\right) \qquad \text{pro } n\to\infty. \tag{3.19}$$

$i = 1, 2$, kde $\{B_p^*(t),\ t \in (0,1)\}$ je proces popsaný pod (3.9). Povšiměte si, že pro $\psi(x) = x$, je-li $\underset{\sim}{\theta}_k(\psi)$ odhad metodou nejmenších čtverců a $g(k/n) = 1$, pak $\max_{p<k<n-p}\{V_{n,k}k(n-k)/n^2\}$ je asymptoticky ekvivalentní statistice (3.16) ($V_{n,k}$ je def. (3.4)).

Využijeme-li (3.19), pak hypotézu $H_0$ v (1.1) zamítneme, jestliže

$$\max_{p<k<n-p}\left\{Z_{n,k}^{(i)}\right\} > b_{\alpha,p}^*, \tag{3.20}$$

$i = 1, 2$, kde $b_\alpha^*(p)$ je definováno (3.12) (příslušné tabulky jsou na konci článku). Robustní analogií testu založeného na $\max_{nt_0\le k\le n(1-t_1)}\{V_{n,k}\sigma^{-2}\}$ je test založený na $\max_{nt_0\le k\le (1-t_1)n}\left\{Z_{n,k}^{(3)}\right\}$, kde

$$Z_{n,k}^{(3)} = \left(\underset{\sim}{\theta}_k(\psi) - \underset{\sim}{\theta}_n(\psi)\right)'\left(\underset{\sim}{C}_k^{-1} - \underset{\sim}{C}_n^{-1}\right)^{-1}\left(\underset{\sim}{\theta}_k(\psi) - \underset{\sim}{\theta}_n(\psi)\right)\hat{\sigma}^{-2}.$$

Uvědomíme-li si, že $U_k = \left(\underset{\sim}{\theta}_k(\psi) - \underset{\sim}{\theta}_{n-k}^*(\psi)\right)'\left(\underset{\sim}{C}_k^{-1} + \underset{\sim}{C}_{n-k}^{-1}\right)^{-1}\left(\underset{\sim}{\theta}_k(\psi) - \underset{\sim}{\theta}_{n-k}^*(\psi)\right)\hat{\sigma}^{-2}$ má pro $k \to \infty$, $n-k \to \infty$ asymptoticky $\chi^2$-rozdělení o $p$ st. v., pak můžeme (využijeme-li Bonferroniho nerovnost) test provést následovně:

(3.21) je-li $\max_{p<k<n-k}\{U_k\} > \chi^2_{\alpha^*}(p)$, zamítneme $H_0$, jinak $H_0$ přijmeme

kde $\alpha^* = \alpha(n-2p)^{-1}$. Obdobný test dostaneme, využijeme-li Šidákovu nerovnost.

Při robustní verzi metody CUSUM založené na $M$-odhadech počítáme $M$-rekurzivní rezidua (odpovídající $W_k$ def. (3.13)) a odhady $\int \psi^2(x)\,dF(x)$, tj. počítáme

$$W_i^* = \psi\left(X_i - \underset{\sim}{c}_i'\underset{\sim}{\theta}_{i-1}(\psi)\right) \qquad i = p+1, \ldots, n, \tag{3.22}$$

$$\sigma_i^{*2} = i^{-1}\sum_{j=1}^{i}\psi^2\left(X_j - \underset{\sim}{c}_j'\underset{\sim}{\theta}_{j-1}(\psi)\right) \qquad i = p+1, \ldots, n. \tag{3.23}$$

$$M_k^* = M_{k-1}^* + W_k^*, \qquad k = p+1, \ldots, n,\ W_p^* = 0. \tag{3.24}$$

Pro případ posunutí dostaneme $W_i^*$, $\sigma_i^{*2}$ a $M_k^*$ definované vztahy (2.45 - 2.47). Test provádíme dle (2.48) jen za $W_i^*$, $\sigma_i^{*2}$ a $M_k^*$ dosazujeme z (3.20 - 3.22). Pro limitní rozdělení platí (2.49).

Zcela analogicky lze sestavit robustní verzi MOSUM, tj. v (2.18) dosazujeme za $W_j$ veličiny $W_j^*\sigma_j^{*-1}$ definované (3.20) a (3.21). Kritickou hodnotu můžeme získat využitím limitního rozdělení (pro $G$ dost velké při $H_0$ má $G^{-\frac{1}{2}}\sum_{j=t-G}^{t}W_j^*\sigma_j^{*-1}$ přibližně normované normální rozdělení) a můžeme použít buď Bonferroniho nebo Šidákovu nerovnost eventuálně, test s kritickou oblastí (2.37), kde $R_t$ je nahrazeno $\sum_{i=k+G}^{k}W_i^*\sigma_k^{*-1}$ ($W_i^*$ a $\sigma_k^*$ jsou dány (3.20) resp. (3.21)).

$M$-odhady jsou definovány implicitně, k jejich výpočtu se používají iterační postupy, což vede v našem případě ke zdlouhavým výpočtům. Výpočetně méně náročně jsou tzv.

rekurzivní $M$–odhady $\underset{\sim}{\theta}^0_k(\psi)$ a rekurzivní jednokrokové $M$–odhady $\underset{\sim}{\theta}^*_k(\psi)$. Rekurzivní $M$–odhad je definován následovně

$$(3.25)\qquad \underset{\sim}{\theta}^0_{k+1}(\psi) = \underset{\sim}{\theta}^0_k(\psi) - \frac{1}{\gamma^0_k(k+1)}\underset{\sim}{C}^{-1}_{k+1}\underset{\sim}{c}_{k+1}\psi\left(X_{k+1} - \underset{\sim}{c}'_{k+1}\underset{\sim}{\theta}^0_k(\psi)\right) \qquad k > p$$

$$\gamma^0_k = \frac{1}{2kt}\sum_{i=1}^{k}\left(\psi\left(X_i - \underset{\sim}{c}'_i\underset{\sim}{\theta}^0_{i-1}(\psi) + i^{-\beta}t\right) - \right.$$
$$\left. - \ \psi\left(X_i - \underset{\sim}{c}'_i\underset{\sim}{\theta}^0_{i-1}(\psi) + i^{-\beta}t\right)i^{\beta}\right), \qquad k > p$$

$t \neq 0$, kde $\underset{\sim}{\theta}^0_p(\psi)$ je vhodný počáteční odhad, matice $\underset{\sim}{C}$ je pozitivně definitní matice definovaná (3.6). S vlastnostmi odhadu $\underset{\sim}{\theta}^0_{k+1}(\psi)$ se čtenář může seznámit např. v Novovičová, Robust (1982). Rekurzivní jednokroková verze $M$–odhadů parametru $\underset{\sim}{\theta}$ je definována jak následuje:

$$\begin{aligned}
\underset{\sim}{\theta}^*_{k+1}(\psi) &= \underset{\sim}{\theta}^*_k(\psi) - \underset{\sim}{C}^{-1}_{k+1}\gamma^{*-1}_k\sum_{i=1}^{k+1}\underset{\sim}{c}_i\psi\left(X_i - \underset{\sim}{c}'_i\underset{\sim}{\theta}^*_k(\psi)\right), \qquad k > p, \\
\gamma^*_k &= \frac{1}{2\sqrt{kt}}\sum_{i=p+1}^{k}\left(\psi\left(X_i - \underset{\sim}{c}'_i\underset{\sim}{\theta}^*_k(\psi) + k^{-\frac{1}{2}}t\right) - \psi\left(X_i - \underset{\sim}{c}'_i\underset{\sim}{\theta}^*_k(\psi) + k^{-\frac{1}{2}}t\right)\right), \qquad k > p
\end{aligned}$$

$t \neq 0$ kde $\underset{\sim}{\theta}^*_p(\psi)$ je vhodný počáteční odhad, $\underset{\sim}{C}_k = \sum_{i=1}^k \underset{\sim}{c}_i\underset{\sim}{c}'_i$. Veličiny $\gamma^0_k$ $\gamma^*_k$ je nutno při výpočtu poněkud modifikovat (ořezávat).

Pokud se týče volby $\psi$, platí pro ni stejné doporučení jako při problému $M$–odhadu. Vzhledem k tomu, že $M$–odhady ani $M$–rekurzivní rezidua nejsou invariantní vůči změně měřítka, doporučuje se nahradit funkci $\psi(\cdot)$ funkcí $\psi(\cdot/s)$, kde $s$ je odhad měřítka podle konkrétního postupu založený na všech $n$ pozorováních nebo na pozorováních provedených do okamžiku mezivýpočtu. Další podrobnosti viz např. Antoch (1984).

# 4 Závěrečné poznámky

## 4.1 Poznámky k postupů v paragrafech 2 a 3

Podíváme–li se zpět na postupy vyložené v paragrafech 2 a 3, je vidět, že se setkáváme se třemi typy testových statistik. První je součtem náhodných veličin (přesněji lineární kombinace testových statistik užívaných pro uvažovaný problém při známém bodu změny $\tau$ (popř. $m$)). Další typ je maximum náhodných veličin (maximum testových statistik užívaných pro uvažovaný problém při známém $m$). Třetím případem je opět typově maximum náhodných veličin, přesněji máme testové statistiky $\max_{p<i\leq k}\{M_i\}$, $k = p + 1, \ldots$, a hypotézu zamítáme, jakmile poprvé $M_k$ překročí kritickou hodnotu, a neděláme další pozorování.

Sledujeme–li jejich limitní chování, pak v prvním případě dostaneme jako limitní rozdělení normální, ve druhém se shoduje s rozdělením $\sup_{t\in(0,1)}\left\{B^*_p(t)\right\}$, kde $B^*_p(t)$ je součet kvadrátů $p$ nezávislých Brownových mostů. Ve třetím případě se shoduje s rozdělením $\sup_{t\in(0,1)}\{|W(t)|\}$, kde $\{W(t),\ t \in (0,1\}$ je normovaný Wienerův proces.

Nyní se podíváme na porovnání různých testů. V literatuře lze najít porovnání hlavně pro případ změny posunutí. Nejprve uvedeme porovnání testů v 2.1. Lze očekávat, že výsledky pro změnu regrese jsou obdobné. Bayesovské testy (2.4), (2.5) jsou optimální z bayesovského hlediska (je však potřebná znalost apriorního rozdělení $\tau$). Z hlediska Bahadurovy vydatnosti je test podílem věrohodností optimální (potřebujeme však přesně znát distribuční funkci $F$). Test optimální ve smyslu Pitmanově neexistuje. Simulace toto potvrzují a dále ukazují, že bayesovské testy jsou poněkud lepší, jen, je-li změna $\tau$ přibližně kolem $t_{[n/2]}$ a že pro malé $\tau$ CUSUM detekuje alternativu stejně dobře jako test podílem věrohodností (CUSUM však většinou vyžaduje menší rozsah výběru).

Pokud se týče srovnání CUSUM a MOSUM, ukazuje se, že MOSUM detekuje alternativu dříve než CUSUM. Další informace lze nalézt např. v Hacklovi (1980), Deshayes a Picard (1986), Praagman (1987) a Haccou a kol. (1985).

Pro porovnání postupů uvedených v (2.1, 3.1) a (2.2, 3.2) platí to co např. pro dvouvýběrový problém. Je-li rozdělení náhodných veličin skutečně normální, jsou postupy uvedené v 2.1 a 3.1 lepší. Je-li však normální rozdělení kontaminované nebo máme značně neúplnou informaci o rozdělení, je lepší použít některý z robustních postupů. Je nutné si však uvědomit, že vliv prvních pozorování po změně bude u robustních postupů potlačen (budou klasifikována jako odlehlá pozorování), což způsobí pozdější zamítnutí $H_0$ (lze tedy očekávat, že odhad bodu změny bude vyšší než skutečná hodnota).

## 4.2 Odhad bodu změny

V literatuře se můžeme setkat s několika typy odhadů bodu změny $\tau$. Při některé formulaci budeme vlastně odhadovat $m$ (viz formulace alternativy).

Při byesovském přístupu odhadneme $\tau$ na základě aposteriorního rozdělení parametru $\tau$ (apriorní jsme značili $p_m = P(\tau \in (t_m, t_{m+1}))$ $m = 1, \ldots, n-1$, $p_n = P(\tau > t_n)$) buď jako medián, nebo střední hodnotu aposteriorního rozdělení (v závislosti na volbě ztrátové funkce). Vzorce pro tyto odhady nejsou jednoduché. Pro některá rozdělení existují tabulky (viz např. Zacks (1983), Broemling a Tsurumi (1987) pro další informace).

Maximálně věrohodný odhad $\tau$ získáme jako bod $t_m$, kde $m$ je bod, pro který dosáhne testová statistika maxima (statistiky def. (2.8), (2.9), (3.3), (3.4)). Vlastnosti odhadů pro některé případy jsou uvedeny např. v Deshayes a Picard (1986). Další informace lze též získat v článku Krishnaiah a Miao (1988).

Při metodě CUSUM a MOSUM odhadneme $\tau$ (přesněji řečeno $m$) následovně:

$$\min\{k;\ |M_k| \geq d_{\alpha,c}(k)\} \tag{4.1}$$

respektive

$$\min\{k;\ |R_k| \geq d_{\alpha,M}(k)\} \tag{4.2}$$

kde $M_k$ je dáno v závislosti na uvažovaném modelu (2.11) nebo (3.14) a $R_j$ dáno (2.18) $W_j$ z (2.10) nebo (3.13). Další informace jsou např. v Hacklovi (1980).

Při robustním přístupu můžeme $\tau$ odhadnout zcela analogicky.

## 4.3 Některé další postupy

Testy uvedené v 2.1 a 3.1 se týkaly změn parametrů normálního rozdělení, což byl případ nejčastěji uvažovaný v literatuře. Bayesovský přístup i test podílem věrohodností lze použít obecně (informace lze nalézt např. v Schulze (1986)). Největší problém je nalezení kritických hodnot.

Při kontrole jakosti výrobků je obvykle problém formulován jako (1.2), kde $P(X_i = a) = p_i$ a $P(X_i = b) = 1 - p_i$, $i = 1, \dots, n$ a $E_{H_0} X_i = 0$. Kromě bayesovského testu a testu podílem věrohodnosti byly navrženy Pagem (1955) testy založené na statistice

$$T = \max_{0 \le k \le n} \left\{ \sum_{i=1}^{s} X_i - \min_{0 < s < k} \sum_{j=1}^{k} X_j \right\}. \tag{4.3}$$

Příslušné kritické hodnoty jsou tabelovány v citovaném článku.

Dalším rozdělením, se kterým se setkáme hlavně v biologických aplikacích, je exponenciální rozdělení (další informace jsou v Haccou a kol. (1985)). Používají se testy podílem věrohodnosti, popř. jejich modifikace.

Z dalších testů stojí za zmínku testy Kolmogorova-Smirnova typu pro úlohu (1.2). Jsou založeny na statistice

$$\begin{aligned} K_n &= \max_{1 \le k \le n} \sup_y \left\{ \left| \hat{F}_k(y) - \hat{F}_{n-k}(y) \right| g(k/n) k/n \left(1 - \frac{k}{n}\right) \right\} \sqrt{n} = \\ & \max_{1 \le k \le n} \sup_y \left\{ \left| \hat{F}_k(y) - \hat{F}_n(y) \right| g(k/n) \right\} \sqrt{n}, \end{aligned} \tag{4.4}$$

kde $\hat{F}_k$ a $\hat{F}_{n-k}$ jsou empirické distribuční funkce odpovídající pozorováním $X_1, \dots, X_k$ respektive $X_{k+1}, \dots, X_n$ a $g$ je nezáporná funkce na $(0,1)$ integrovatelná se čtvercem. Je-li $X_1, \dots, X_n$ náhodný výběr z rozdělení se spojitou distribuční funkcí $F$, pak

$$\mathcal{L}(K_n) \longrightarrow \mathcal{L}\left( \sup_{t \in (0,1)} \sup_{u \in (0,1)} \{ |B(u,t)| g(u) \} \right) \qquad \text{pro } n \to \infty, \tag{4.5}$$

kde $\{B(u,t), (u,t) \in (0,1)^2\}$ je dvourozměrný Brownův most (tj. gaussovský proces na $(0,1)^2$ a kovarianční funkcí $E\,B(u_1,t_1) B(u_2,t_2) = (\min(t_1,t_2) - t_1 t_2) \cdot (\min(u_1,u_2) - u_1 u_2)$ pro $u_i$, $t_i \in (0,1)$, $i = 1, 2$).

Vlastnostmi statistik typu (4.4) se zabývali např. Deshayes a Picard (1986), Csörgö, M. a Horváth (1988), Hawkins (1988).

Sen (1983 b) navrhl test založený na $U$-statistikách pro případ, že se změna rozdělení projeví ve funkcionálu $\theta(F) = E\,h(X_1, \dots, X_s)$ kde $h$ je funkce symetrická v proměnných $x_1, \dots, x_s$ integerovatelná se čtvercem, $s < n$.

Existuje poměrně rozsáhlá literatura týkající se problému (1.1) popř. (1.2) pro případ, že pozorování jsou závislá (např. časové řady, AR, ARMA, ARIMA). Uvedeme zde aspoň které citace: Bagshaw a Johnson (1977), Johnson a Bagshaw (1974), Sastri, Flores a Valdéz (1989), Segen a Sanderson (1980), Picard (1985), Deshayes a Picard (1986), Hawkins (1984, 1985).

Tabulka kritických hodnot příslušných Wienerovu procesu a Brownovu mostu:

| $\alpha$ | $w_\alpha$ | $b_\alpha$ | $b^*_{\alpha,1}$ | $b^*_{\alpha,2}$ | $b^*_{\alpha,3}$ | $b^*_{\alpha,4}$ | $b^*_{\alpha,5}$ |
|---|---|---|---|---|---|---|---|
| 0,01 | 2,807 | 1,517 | 2,650 | 3,392 | 4,004 | 4,550 | 5,054 |
| 0,05 | 2,241 | 1,224 | 1,844 | 2,510 | 3,056 | 3,542 | 4,000 |
| 0,1 | 1,96 | 1,073 | 1,498 | 2,114 | 2,615 | 2,084 | 3,516 |

Připomeňme definice jednotlivých kritických hodnot:

$$P\left(\sup_{t\in(0,1)}\{|W(t)|\}\geq w_\alpha\right)=\alpha$$

$$P\left(\sup_{t\in(0,1)}\{B_1(t)\}>b_\alpha^+\right)=\alpha$$

$$P\left(\sup_{t\in(0,1)}\left\{\sum_{j=1}^{p}B_j^2(t)\right\}\geq b_{\alpha,p}^*\right)=\alpha,$$

kde $\{W(t),\, t\in(0,1)\}$ je normovaný Wienerův proces, $\{B_j(t),\, t\in(0,1)\}$, $j=1,\dots p$ jsou nezávislé Brownovy mosty.

## Literatura


Antoch, J. (1984). Některé postupy pro výpočet robustních odhadů v lineárním modelu. Robust 1984, 1–9.

Antoch, J. (1990). Recursive $M$–tests for the change–point problem Monte Carlo study. Zasláno k otištění.

Antoch, J. a Hušková, M. (1989). Some $M$–tests for detection of a change in linear models. Proceedings of the Fourth Prague Symposium on Asymptotic Statistics, ed. Mandl, P., Hušková, M., 123–136.

Bagshaw, M. a Johnson, R. A. (1977). Sequential procedures for detecting parameter change in a time series models. JASA, 593–597.

Bhattacharya, G. K. a Johnson R. A. (1968). Nonparametric shift at an unknown time point. Ann. Math. Statist. 39, 1731–1743.

Broemling, L. D. a Tsurumi, H. (1987). Econometrics and structural changes. New York, M. Dekker. Ser. Statistics, vol. 74.

Brown, R. L., Durbin, J. a Evans, J. M. (1975). Techniques for testing the constancy of regression relationship over time (w. discussion). J. Roy. Statist. Soc. B 37, 149–192.

Chernoff, H. a Zacks, S. (1964). Estimating the current mean of a normal distribution which is subject to changes in time. Ann. Math. Statist. 35, 999–1018.

Chin Choy, C. J. H. a Broemling, L. D. (1980). Some Bayesian inferences for a changing linear model. Technometrics 22, 71–78.

Cobb, G. W. (1978). The problem of the Nile conditional solution to a change point problem. Biometrika 65, 243–251.

Csörgö, M. a Horváth, L. (1988). Nonparametric methods for change–point problems. Handbook of Statistics, vol. 7, editoři Krishnaiah, P. R. a Rao, C. R., Elsevier Science Publishers, B. V., 403–429.

Dienske, H. a Metz, J. A. J. (1977). Mother–infant body contact in Macaques; a time interval analysis. Biology of Behaviour 2, 3–37.

Deshayes, J. a Picard, D. (1986). Off–line statistical analysis of change–point models using non–parametric and likelihood methods, 103–168, v Lecture Notes in Control and Information Sciences 77 – Detection of abrupt changes in signals and dynamical systems.

Esterby, S. S. a El. Shaarawi (1981). Inference about the point of change in a regression model. Appl. Statistics 30, 277–285.

Farley, J. N. a Hinich, M. J. (1970). A test for a shifting slope coefficient in a linear model. JASA 65, 1320–1329.

Ferreira, P. E. (1975). A Bayesian analysis of switching regression model: known number of regimes. JASA 70, 370–374.

Gardner, L. A., Jr. (1969). On detecting changes in the mean of normal variates. Ann. Math. Statist. 40, 116–126.

Haccou, P., Dienske, E. a Meelis, E. (1983). Analysis of time–inhomogenity in Markov chains applied to mother–infant interactions of rheus monkeys. Animal Behaviour 31, 927–945.

Haccou, P., Meelis, E. a van der Geer Sara (1985). On the likelihood ratio test for a change point in a sequence of independent exponentially distributed random variables. Report MS–R 8507 (Amsterdam).

Hackl, P. (1980). Testing the constancy of regression relationships over time. Coettingen. Vanden Hoeck und Ruprecht.

Hackl, P. a Westlund, A. (1985). Statistical analysis of structural change, an annotated Bibliography. Zpráva IIASA CP–85–31.

Hawkins, D. M. (1980). A note on continuous and discontinuous segmented regression. Technometrics 22, 443–444.

Hawkins, D. L. (1984). Sequential procedures of detecting deviations in the parameters of autoregressive model from specified target. Sequential Analysis 3, 121–154.

Hawkins, D. L. (1985). Sequential procedures for monitoring the parameters of the autoregressive model relative to unspecified target. Sequential Analysis 4, 275–310.

Hawkins, D. L. (1988). Retrospective and sequential tests for a change in distribution based on Kolmogorov–Smirnov type statistics. Sequential Analysis 7, 23–51.

Hinkley, D. V. (1972). Time–ordered classification. Biometrika 59, 509–523.

Hsu, D. A. (1979). Detecting shifts of parameter in gamma sequences with applications to stock price and air traffic flow analysis. JASA 74, 31–40.

Hušková, M. (1988). Adaptive procedures for detection of change. Statistics and Decisions 6, 137–148.

Hušková, M. (1988). Recursive $M$–test for detection of change. Sequential Analysis 7, 75–90.

Hušková, M. (1990 a). Some asymptotic results for robust procedures for testing the constancy of regression models over time. Kybernetika 26, 392–403.

Hušková, M. (1990 b). Asymptotic for robust MOSUM. CMUC 31, 345–356.

Hušková, M. (1990 c). Stochastic approximation type estimators in linear models. V tisku.

Hušková, M. a Sen, P. K. (1989). Nonparametric tests for shift and change in regression at an unknown time point. Statistical Analysis and Forecasting of Economic Structural Change, editor P. Hackl, Springer-Verlag, 71-86.

James, B., James, L. J. a Siegmund, D. (1987). Tests for a change point, Biometrika 74, 71-83.

Johnson, R. A. a Bagshaw, M. (1974). The effect of serial correlation on the performance of CUSUM tests. Technometrics 16, 103-112.

Kiefer, J. (1959). $K$-sample analogue of the Kolmogorov-Smirnov and Cramér-von Mises tests. Ann. Math. Statist. 30, 420-447.

Krishnaiah, P. R. a Miao, B. Q. (1988). Review about estimation of change points. Handbook of Statistics vol. 7, editoři Krishnaiah, P. R. a Rao, R. C., Elsevier Science Publishers B. V., 375-402.

De Long, D. (1981). Crossing probabilities for a square root boundary by a Bessel process. Comm. Statist. Theor. Math. A 10, 2197-2213.

Novovičová, J. (1982). Rekurentní algoritmy odhadování v lineárním regresním modelu. Robust 1982, 60-67.

Page, E. S. (1955). A test for a change in a parameter occuring at an unknown time point. Biometrika 42, 248-252.

Page, E. S. (1957). On problems in which a change of parameters occurs at an unknown time point. Biometrika 44, 248-252.

Pettit, A. N. (1979). A nonparametric approach to the change-point problem. Appl. Statist. 28, 126-135.

Picard, O. (1985). Testing and estimating change-points in time series. Journal of Appl. Probab. 17, 841-867.

Praagman, J. (1988). Bahadur efficiency of rank tests for the change-point problem. Annals of Statist. 16, 198-217.

Sastri, T., Flores, B. a Valdéz, J. (1988). Detecting points of change in time series. Computers Opus Res. 16, 271-293.

Schechtman, E., a Wolfe, D. A. (1981). Distribution-free tests for the change-point problem. Techn. Report, Ohio State Univ.

Schulze, U. (1986). Identifikacija momentu zmiany rozkladu w ciagu zmiennych losowych. Roczniki polskiego towarzstwa matematycznego ser. III. Math. Stosowana XXVII.

Segen, J. a Sanderson, A. C. (1980). Detecting change in a time series. IEEE transactions on information theory, vol. 26, 246-255.

Sen, A. a Srivastava, M. S. (1975). On tests for detecting change in mean. Ann. Statist. 3, 98-108.

Sen, P. K. (1980). Asymptotic theory of some tests for a possible change in the regression slope occuring at an unknown time point. Z. Wahrscheinl. verw. Gebiete 52, 203-218.

Sen, P. K. (1982). Asymptotic theory of some tests for constancy of regression relationshops over time. Statistics 13, 21-31.

Sen, P. K. (1983 a). On some recursive residual rank tests for change-points. In Rizvi, M. H. (Ed.) Recent Advances in Statistics: Papers in Honor of Herman Chernoff's Sixtieth Birthday. New York, Acad. Press, 371–391.

Sen, P. K. (1983 b). Test for change-points based on recursive $U$-statistics. Sequential Analysis 1, 263–284.

Sen, P. K. (1984). Recursive $M$-tests for the constancy of multivariate regression relationships over time. Sequential analysis 3, 191–211.

Shaban, S. A. (1980). Change-point problem and two phase regression: an annotated bibliography. Inst. Statist. Review 48, 83–93.

Shirjajev, A. N. (1963). On optimum methods in quicktest detection problems. Theor. Prob. Appl. 10, 348–354.

Shirjajev, A. N. (1978). Optimal Stopping Rules. Springer-Verlag, New York.

Wolfe, D. A. a Schechtman, E. (1984). Nonparametric statistical procedures for the changepoint problem. J. Stat. Planning and Inference 9, 389–396.

Worsley, K. J. (1982). An improved Bonferroni inequality and applications. Biometrika 62, 297–302.

Worsley, K. J. (1983). Testing for a two-phase multiple regression. Technometrics 25, 35–42.

Zacks, S. (1983). Survey of classical and Bayesian approaches to the change-point problem: fixed sample and sequential procedures of testing and estimation. Recent Advances in Statistics: Paper in Honor of Herman Chernoff's Sixtieth Birthday, New York, Acad. Press, 245–269.