Akademie věd České republiky
Ústav teorie informace a automatizace

Academy of Sciences of the Czech Republic
Institute of Information Theory and Automation

RESEARCH REPORT

Jiří Michálek:

Regulační diagramy založené na bayesovském přístupu

No. 2280 květen 2010

ÚTIA AV ČR, P. O. Box 18, 182 08 Prague, Czech Republic
E-mail: utia@utia.cas.cz

Regulační diagramy založené na bayesovském přístupu

Jiří Michálek

Obecný model

Nechť náhodné veličiny $X_1, X_2, \ldots, X_n$ mají zpočátku rozdělení pravděpodobnosti dané hustotou $f(\cdot)$, předpokládáme jejich vzájemnou nezávislost. Od jistého okamžiku $r \in \{1, 2, \ldots, n, n+1\}$ nastane změna v tom smyslu, že rozdělení se změní na hustotu $g(\cdot)$. Okamžik $r = n + 1$ znamená, že ke změně hustoty během časových okamžiků $1, 2, \ldots, n$ nedošlo.

Sdružená hustota veličin $X_1, X_2, \ldots, X_n$ má tedy obecně tvar

$$h_r(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{r-1} f(x_i) \prod_{j=r}^{n} g(x_j)$$

pro $r = 2, \ldots, n$. Pro krajní případy $r = 1$ a $r = n + 1$ dodefinujeme následovně:

$$\begin{aligned} h_1(x_1, \ldots, x_n) &= \prod_{i=1}^{n} g(x_i) \\ h_{n+1}(x_1, \ldots, x_n) &= \prod_{i=1}^{n} f(x_i). \end{aligned}$$

Parametr $r \in \{1, 2, \ldots, n+1\}$ náleží do diskrétního parametrického prostoru, na němž je zadáno apriorní rozdělení možného výskytu změny.

Nulová hypotéza znamená, že ke změně nedošlo, tedy $r = n + 1$. Alternativní hypotéza je složená a představuje tedy celkem n možností změny $r \in \{1, 2, \ldots, n\}$. Na základě pozorování $x_1, x_2, \ldots, x_n$ musíme učinit rozhodnutí, zda-li ke změně nedošlo či došlo. Rozhodnutí d_i znamená detekci změny v okamžiku i při volbě ztrátové funkce

$$e(i, d_j) = 1 - \Delta_{ij}$$

Hledáme tedy bayesovskou rozhodovací funkci $\{\varphi(i, \mathbf{x})\}$, kde

$$\{\varphi(i, \mathbf{x})\} = \text{Pst}\{\text{přijetí } d_i \,|\, x_1, x_2, \ldots, x_n\}$$

za podmínky

$$\sum_{i=1}^{n+1} \varphi(i, \mathbf{x}) = 1.$$

Označme $\{p_i\}_{i=1}^{n+1}$ apriorní rozdělení na $\{1, 2, \ldots, n+1\}$. Pak podmíněné riziko přijetí

$$\theta_i \longleftrightarrow h_i(x_1, x_2, \ldots, x_n), \quad i = 1, 2, \ldots, n+1$$

při rozhodovací funkci $\varphi(i, \mathbf{x})$ je rovno

$$\begin{aligned} R(\theta_i, \varphi(\cdot) \,|\, \mathbf{x}) &= \sum_{j=1}^{n+1} e(i, d_j)\, E_{\theta_i}\{\varphi(j \,|\, \mathbf{x})\} \\ &= 1 - E_{\theta_i}\{\varphi(i, \mathbf{x})\}. \end{aligned}$$

Pak střední riziko při apriorním rozdělení $\{p_i\}$ je dáno výrazem

$$r\left(\{p_i\},\, \varphi(\cdot)\right) = 1 - \sum_{i=1}^{n+1} p_i\, E_{\theta_i}\{\varphi(i, \mathbf{x})\}.$$

Bayesovská procedura, která minimalizuje střední riziko, má tvar

$$\begin{aligned} \Phi(j \,|\, \mathbf{x}) = 0 &\Leftrightarrow f_j\, h_j(\mathbf{x}) < p_{i_0} h_{i_0}(\mathbf{x}) \\ &= \max_{1 \le \ell \le n+1} \{p_\ell\, h_\ell(\mathbf{x})\} \\ \Phi(i_0 \,|\, \mathbf{x}) &= 1, \end{aligned}$$

neboli

$$\Phi(i_0 \,|\, \mathbf{x}) = 1 \Longleftrightarrow p_{i_0}\, h_{i_0}(\mathbf{x}) > p_j\, h_j(\mathbf{x}) \quad \text{pro všechna } j \neq i_0.$$

Tím je tedy detekována změna v okamžiku $r = i_0$. Je předpokládáno, že se objeví pouze jedna změna.

Nyní tento postup budeme aplikovat na detekci změn v chování střední hodnoty a směrodatné odchylky pro normální rozdělení, které se nejčastěji používá v praxi pro popis spojitých znaků jakosti na výrobcích.

Nejdříve ale zcela obecně.

Je-li tedy změna detekována v okamžiku $r = i_0$, pak musí současně platit:

a) pro $1 \le i \le i_0 - 1$ musí platit nerovnost

$$p_{i_0}\, h_{i_0}(\mathbf{x}) > p_i\, h_i(\mathbf{x})$$

tedy

$$p_{i_0} \prod_{\ell=1}^{i_0-1} f(x_\ell) \prod_{\ell=i_0}^{n} g(x_\ell) > p_i \prod_{\ell=1}^{i-1} f(x_\ell) \prod_{\ell=i}^{n} g(x_\ell),$$

což dává

$$p_{i_0} \prod_{\ell=i}^{i_0-1} f(x_\ell) > p_i \prod_{\ell=1}^{i_0-1} g(x_\ell).$$

Čili pro každé $i \in 1, 2, \ldots, i_0 - 1$ musí platit

$$\frac{p_{i_0}}{p_i} > \prod_{\ell=i}^{i_0-1} \frac{g(x_\ell)}{f(x_\ell)}.$$

b) pro $i_0 < i \leq n+1$

$$p_{i_0} \prod_{\ell=1}^{i_0-1} f(x_\ell) \prod_{\ell=i_0}^{n} g(x_\ell) > p_i \prod_{\ell=1}^{i-1} f(x_\ell) \prod_{\ell=i}^{n} g(x_\ell),$$

tedy při $i > i_0$

$$p_{i_0} \prod_{\ell=i_0}^{i-1} g(x_\ell) > p_i \prod_{i_0}^{i-1} f(x_\ell),$$

neboli

$$\prod_{\ell=i_0}^{i-1} \frac{g(x_\ell)}{f(x_\ell)} > \frac{p_i}{p_{i_0}}$$

pro každé $i \in \{i_0+1, i_0+2, \ldots, n, n+1\}$.

Změna nebude detekována v tom případě, když pro $i_0 = n+1$, což znamená splnění nerovností

$$p_{n+1} \prod_{\ell=1}^{n} f(x_\ell) > p_i \prod_{\ell=1}^{i=1} f(x_\ell) \prod_{\ell=i}^{n} g(x_\ell)$$

pro každé $i = 1, 2, \ldots, n$. To dává sadu nerovností:

$$p_{n+1} \prod_{\ell=i}^{n} f(x_\ell) > p_i \prod_{\ell=i}^{n} g(x_\ell),$$

neboli

$$\frac{p_{n+1}}{p_i} > \prod_{\ell=i}^{n} \frac{g(x_\ell)}{f(x_\ell)}$$

pro každé $i = 1, 2, \ldots, n$.

Tyto nerovnosti lze snadno převést na nerovnosti součtů, totiž

$$\ln \frac{p_{n+1}}{p_i} > \sum_{\ell=i}^{n} \ln \frac{g(x_\ell)}{f(x_\ell)},$$

pro každé $i = 1, 2, \ldots, n$, neboli musí platit souhrnně:

$$\begin{array}{rcl} \ln \frac{p_{n+1}}{p_1} & > & \sum_{\ell=1}^{n} \ln \frac{g(x_\ell)}{f(x_\ell)} \\ \ln \frac{p_{n+1}}{p_2} & > & \sum_{\ell=2}^{n} \ln \frac{g(x_\ell)}{f(x_\ell)} \\ \vdots & \vdots & \vdots \\ \ln \frac{p_{n+1}}{p_n} & > & \ln \frac{g(x_n)}{f(x_n)}. \end{array}$$

Například při rovnoměrném apriorním rozdělení $p_i = \frac{1}{n+1}$ to znamená, že všechny součty na pravých stranách musí být záporné, aby touto procedurou nebyla detekována změna v rozdělení pravděpodobnosti.

Označme

$$T_n(x_1, \ldots, x_n) = \max_{1 \le i \le n} \left\{ \sum_{\ell=1}^{n} \ln \frac{g(x_\ell)}{f(x_\ell)} + \ln \frac{p_i}{p_{n+1}} \right\}.$$

Když tedy bude $T_n(x_1, \ldots, x_n) < 0$, pak mezi pozorováními $x_1, x_2, \ldots, x_n$ není detekována změna. Této skutečnosti lze využít pro návrh algoritmu detekce změny ve smyslu on-line. Nechť $T_0(x_1, \ldots, x_n) = 0$. Pak lze snadno úplnou matematickou indukcí dokázat, že

$$\begin{aligned} T_{n+1}(x_1, \ldots, x_{n+1}) &= \max \Big\{ \ln \frac{g(x_{n+1})}{f(x_{n+1})} + \ln \frac{p_{n+1}}{p_{n+2}}, T_n(x_1, \ldots, x_n) \\ &\qquad + \ln \frac{g(x_{n+1})}{f(x_{n+1})} + \ln \frac{p_{n+1}}{p_{n+2}} \Big\}. \end{aligned}$$

Aby nebyla detekována změna ani v kroku $n+1$, za předpokladu, že nebyla doposud detekována, a tedy je

$$T_n(x_1, \ldots, x_n) < 0,$$

pak $T_{n+1}(x_1, x_2, \ldots, x_{n+1}) < 0$ buď při splnění nerovnosti

$$\ln \frac{g(x_{n+1})}{f(x_{n+1})} + \ln \frac{p_{n+1}}{p_{n+2}} < 0$$

či při splnění nerovnosti

$$T_n(x_1, x_2, \ldots, x_n) + \ln \frac{p_{n+1}}{p_{n+1}} + \ln \frac{p_{n+1}}{p_{n+2}} < 0.$$

Posloupnosti $\{p_i\}_{i=1}^{n+1}$ a $\{p_i\}_{i=1}^{n+2}$ hrají roli ladících parametrů v algoritmu. Pokud změna do okamžiku $n+1$ není detekována, pak je logické pravděpodobnost výskytu změny v čase $n+2$ maximalizovat. Když bychom výskytu změny přisoudili pro nejnovější okamžik stejnou apriorní pravděpodobnost, pak by bylo

$$\ln \frac{p_{n+1}}{p_{n+2}} = 0.$$

Obecně lze tedy hodnotu $\ln \frac{p_{n+1}}{p_{n+2}}$ považovat za nastavitelný parametr, který určuje statistické vlastnosti studovaného algoritmu. Tento parametr by měl být záporný, jelikož lze usuzovat na to, že by mělo být $p_{n+1} < p_{n+2}$, když do okamžiku $n+1$ změna nenastala.

Detekce změny by probíhala v režimu on-line následovně: Stanoví se vhodná mez, která hlídá nejnovější podíl

$$\ln \frac{g(x_{n+1})}{f(x_{n+1})} < \ln \frac{p_{n+2}}{p_{n+1}}$$

a současně se hlídá chování součtů

$$T_n(x_1, \ldots, x_n) + \ln \frac{g(x_{n+1})}{f(x_{n+1})} < \ln \frac{p_{n+2}}{p_{n+1}}.$$

Podíl $\ln \frac{p_{n+2}}{p_{n+1}}$ by měl být kladný, ale ne zase příliš velký.

Aby změna nebyla detekována, je tedy nutné, aby maximum z veličin

$$T_n(x_1, x_2, \ldots, x_n) + \ln \frac{g(x_{n+1})}{f(x_{n+1})} + \ln \frac{p_{n+1}}{p_{n+2}}$$

a

$$\ln \frac{g(x_{n+1})}{f(x_{n+1})} + \ln \frac{p_{n+1}}{p_{n+2}}$$

bylo záporné. Kdyby ale platila nerovnost

$$T_n(x_1, \ldots, x_n) + \ln \frac{g(x_{n+1})}{f(x_{n+1})} + \ln \frac{p_{n+1}}{p_{n+2}} > \ln \frac{g(x_{n+1})}{f(x_{n+1})} + \ln \frac{p_{n+1}}{p_{n+2}},$$

pak nemůže být splněna nerovnost $T_n(x_1, x_2, \ldots, x_n) < 0$, která musí platit při žádné změně do okamžiku n.

Tato skutečnost pak vede k požadavku na splnění nerovnosti

$$\ln \frac{g(x_{n+1})}{f(x_{n+1})} < \ln \frac{p_{n+2}}{p_{n+1}}.$$

Tedy on-line algoritmus je vymezen jedinou mezí, která má ladící charakter a její nastavení odpovídá riziku chyby 1. druhu, totiž

$$\text{Pst} \left\{ \ln \frac{g(x_{n+1})}{c(x_{n+1})} \geq \ln \frac{p_{n+2}}{p_{n+1}} \right\} = \alpha$$

při platnosti hypotézy $H_0 : f$. Z toho ihned plyne, že

$$\alpha = \int_{\rho}^{\infty} \frac{g(x)}{f(x)} f(x) \, \mathrm{d}x = \int_{\rho}^{\infty} g(x) \, \mathrm{d}x = 1 - \mathcal{G}(\rho)$$

za předpokladu, že $f(x) > 0$ skoro všude, $\rho = \frac{p_{n+2}}{p_{n+1}}$.

V čase $n + 1$ máme k dispozici pozorování x_{n+1} a změna nebude od okamžiku 1 do $n + 1$ detekována, pokud budou platit následující nerovnosti

$$\begin{array}{rlll}
1. & \ln \dfrac{g(x_{n+1})}{f(x_{n+1})} & + \ln \dfrac{p_{n+1}}{p_{n+2}} < 0 & \\
2. & \ln \dfrac{g(x_{n+1})}{f(x_{n+1})} & + \ln \dfrac{p_n}{p_n} + \ln \dfrac{p_n}{p_{n+2}} < 0 & \\
\vdots & \quad\vdots & \vdots & \\
n+1. & \displaystyle\sum_{\ell=1}^{n+1} \frac{g(x_\ell)}{f(x_\ell)} & + \dfrac{p_1}{p_{n+2}} < 0. &
\end{array}$$

Změna tedy bude detekována, pokud alespoň jedna nerovnost z výše uvedených bude opačná. Tedy pak bude existovat index i_0 takový, že

$$\sum_{\ell=i_0}^{n+1} \ln \frac{g(x_\ell)}{f(x_\ell)} + \frac{p_{i_0}}{p_{n+2}} \geq 0,$$

tedy

$$\sum_{\ell=i_0}^{n} \ln \frac{g(x_\ell)}{f(x_\ell)} + \ln \frac{g(x_{n+1})}{f(x_{n+1})} + \ln \frac{p_{i_0}}{p_{n+2}} \geq 0.$$

Tuto nerovnost rozepíšeme následovně:

$$\sum_{\ell=i_0}^{n} \ln \frac{g(x_\ell)}{f(x_\ell)} + \ln \frac{p_{i_0}}{p_{n+2}} + \ln \frac{g(x_{n+1})}{f(x_{n+1})} + \ln \frac{p_{i_0}}{p_{n+1}} - \ln \frac{p_{i_0}}{p_{n+1}} \geq 0.$$

Tedy

$$\sum_{\ell=i_0}^{n} \ln \frac{g(x_\ell)}{f(x_\ell)} + \ln \frac{p_{i_0}}{p_{n+1}} + \ln \frac{g(x_{n+1})}{f(x_{n+1})} + \ln \frac{p_{n+1}}{p_{n+2}} \geq 0.$$

Protože z předchozího víme, že v časech $1, 2, \ldots, n$ nebyla chyba detekována, pak musí platit, že

$$\sum_{\ell=i_0}^{n} \ln \frac{g(x_\ell)}{f(x_\ell)} + \ln \frac{p_{i_0}}{p_{n+1}} < 0,$$

což implikuje, že nutně musí být

$$\ln \frac{g(x_{n+1})}{f(x_{n+1})} + \ln \frac{p_{n+1}}{p_{n+2}} \geq 0.$$

Pokud by tedy byla změna detekována, musí se to nutně projevit na chování

$$\ln \frac{g(x_{n+1})}{f(x_{n+1})} + \ln \frac{p_{n+1}}{p_{n+2}} \geq 0.$$

Síla testu

Chyba 2. druhu znamená, že změna sice nastala, ale nebyla detekována, tedy jedná se pak o riziko takové chyby jako

$$\begin{aligned} & \mathrm{Pst}\left\{ \ln \frac{g(x_{n+1})}{f(x_{n+1})} < \ln \rho \,|\, A : g(\cdot) \right\} \\ = \; & \mathrm{Pst}\left\{ \frac{g(x_{n+1})}{f(x_{n+1})} < \rho \,|\, A : g(\cdot) \right\} \\ = \; & \int_{-\infty}^{\rho} \frac{g(x)}{f(x)}\, g(x)\, \mathrm{d}x = \int_{-\infty}^{\rho} \frac{g^2(x)}{f(x)}\, \mathrm{d}x = \beta. \end{aligned}$$

Odtud již snadno síla testu je

$$1 - \beta = \int_{\rho}^{\infty} \frac{g^2(x)}{f(x)}\, \mathrm{d}x,$$

samozřejmě při volbě rizika chyby 1. druhu

$$\alpha = \int_{\rho}^{\infty} g(x)\, \mathrm{d}x, \quad \text{tedy} \quad \rho = q_{1-\alpha}(g(\cdot)).$$

Zde je ihned vidět, že rizika α a β jsou svázána jako spojité nádoby volbou ladícího parametru ρ.

Detekce změny v chování střední hodnoty $N(\mu, \sigma^2)$ při konstantním rozptylu

V tomto případě máme, že
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2}$$
a
$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu_1)^2},$$
kde $\mu_1 = \mu + \Delta$. Pak
$$\ln \frac{g(x)}{f(x)} = \frac{1}{2\sigma^2}(2\Delta(x-\mu) - \Delta^2).$$
Aby nebyla tedy změna detekována poprvé, když doposud nenastala, musí platit
$$\frac{\Delta(x-\mu)}{\sigma^2} - \frac{\Delta^2}{2\sigma^2} < \ln \rho.$$
Při $\Delta > 0$ musí tedy platit nerovnost
$$\frac{x-\mu}{\sigma^2} < \frac{\ln \rho}{\Delta} + \frac{\Delta}{2\sigma^2}.$$
Když bychom uvažovali rovnoměrné rozdělení jako apriorní, pak dostáváme požadavek
$$x - \mu < \frac{\Delta}{2}.$$
Pro $\Delta < 0$ zcela analogicky musí platit
$$\frac{x-\mu}{\sigma^2} > \frac{\ln \rho}{\Delta} + g\frac{\Delta}{2\sigma^2},$$
tady při rovnoměrném apriorním rozdělení
$$x - \mu < +\frac{\Delta}{2}.$$
Pro oboustranný případ musí tedy platit nerovnost
$$\frac{\sigma^2 \ln \rho}{\Delta_1} + \mu + \frac{\Delta_1}{2} < x < \frac{\sigma^2 \ln \rho}{\Delta_2} + \mu + \frac{\Delta_2}{2},$$
kde $\Delta_1 < 0,\ \Delta_2 > 0$. Pokud změna nebyla v časech $1, 2, \ldots, n$ detekována, pak nebude ani v čase $n+1$, když platí nerovnost
$$\frac{\sigma^2 \ln \rho}{\Delta_1} + \mu + \frac{\Delta_1}{2} < x_{n+1} < \frac{\sigma^2 \ln \rho}{\Delta_2} + \mu + \frac{\Delta_2}{2}.$$
V případě, že bychom uvažovali rovnoměrné rozdělení jako apriorní, pak nastává zajímavá situace, že nemusíme znát směrodatnou odchylku σ.

Jak je ihned vidět, chování tohoto algoritmu pro detekci změny ve střední hodnotě je velice podobné s chováním Shewhartových regulačních diagramů pro aritmetické průměry z logických podskupin či individuálních dat v případě tzv. technických mezí, které jsou odvozeny od specifikace sledovaného znaku kvality.

Z tohoto lze i snadno vyvodit vztah mezi rizikem α chyby 1. druhu a volbou parametru Δ při dané či požadované směrodatné odchylce σ. V případě rovnoměrného apriorního rozdělení lze snadno zjistit, že parametr Δ musí pak splňovat vztah

$$\Delta = 2\sigma\, u_{1-\alpha/2},$$

v případě $\Delta_1 = -\Delta,\ \Delta_2 = \Delta > 0$. ($u_{1-\alpha/2}$ je kvantil $N(0,1)$.) Když nebudeme uvažovat rovnoměrné rozdělení, pak musí platit

$$\frac{\sigma \ln \rho}{\Delta} - \frac{\Delta}{2\sigma} < \frac{x_{n+1} - \mu}{\sigma} < \sigma \frac{\ln \rho}{\Delta} + \frac{\Delta}{2\sigma},$$

neboli

$$\sigma \frac{\ln \rho}{\Delta} + \frac{\Delta}{2\sigma} = u_{1-\alpha/2}.$$

Budeme-li znát σ a Δ bude stanoveno jako největší možné odchýlení od cílové hodnoty μ, pak platí pro parametr ρ:

$$\ln \rho = \frac{\Delta}{\sigma}\, u_{1-\alpha/2} - \frac{\Delta^2}{2\sigma^2},$$

tedy

$$\rho = \exp\left\{ \frac{\Delta}{\sigma}\, u_{1-\alpha/2} - \frac{\Delta^2}{2\sigma^2}. \right\}$$

U regulačních diagramů se obvykle posunutí Δ vyjadřuje v násobcích směrodatné odchylky σ, tedy uvažuje se poměr

$$\tau = \frac{\Delta}{\sigma}.$$

Pak

$$\rho = \exp\left\{ \tau\, u_{1-\alpha/2} - \frac{\tau^2}{2} \right\}.$$

Například pro $\alpha = 0,05$, kdy $u_{1-\alpha/2} = 1,96$ pak dostáváme vztah

$$\rho = \exp\left\{ 1,96\tau - \frac{\tau^2}{2} \right\},$$

pro $\alpha = 0,0027$ (což je riziko klasických Shewhartových diagramů vůči oběma regulačním mezím současně), pak máme vztah

$$\rho = \exp\left\{ 2,78\tau - \frac{\tau^2}{2} \right\}.$$

Tedy při $\tau = 1,5$, což opět je často aplikovaný poměr v praxi (viz metodologie Six Sigma), pak

$$\ln \rho \doteq 3,045.$$

Když parametry μ a σ neznáme, je přirozené je nahradit je jejich odhady $\overline{x}$ a s, které jsou spočítány ze všech předchozích dat, když nebyla doposud změna zaznamenána. Tím se pak dostáváme k tzv. přirozeným mezím:

$$\frac{s}{\Delta}\ln\rho - \frac{\Delta}{2s} < \frac{x_{n-1}-\overline{x}}{s} < \frac{s}{\Delta}\ln\rho + \frac{\Delta}{2s}.$$

Jaká je síla tohoto testu?

Nechť tedy nastala změna ve střední hodnotě a nejnovější pozorování x_{n+1} je z rozdělení $N(\mu+\Delta,\sigma^2)$, musí tedy platit, že při $\Delta > 0$

$$\frac{x_{n+1}-\mu}{\sigma} \le \frac{\sigma}{\Delta}\ln\rho - \frac{\Delta}{2\sigma}$$

či

$$\frac{x_{n+1}-\mu}{\sigma} \ge \frac{\sigma}{\Delta}\ln\rho + \frac{\Delta}{2\sigma}$$

při rozdělení $N(\mu+\Delta,\sigma^2)$. Snadným výpočtem zjistíme, že síla testu je rovna

$$1-\beta = \Phi\left(\frac{\ln\rho}{\tau} - \frac{3}{2}\tau\right) + 1 - \Phi\left(\frac{\ln\rho}{\tau} - \frac{\tau}{2}\right),$$

kde $\tau = \frac{\Delta}{\sigma}$ pro $\Delta > 0$. Tedy např. pro $\rho = 0$ a $\tau = 1$ je pak síla testu

$$1-\beta = \Phi\left(-\frac{3}{2}\right) + 1 - \Phi\left(-\frac{1}{2}\right) \doteq 0,7586.$$

Obecně v případě při $\rho = 0$ a volbě $\Delta = 2\sigma\, u_{1-\alpha/2}$ pak máme, že síla testu

$$1-\beta = \Phi\left(-3u_{1-\alpha/2}\right) + 1 - \Phi\left(-u_{1-\alpha/2}\right) \doteq \Phi\left(u_{1-\alpha/2}\right) = 1-\alpha/2.$$

Detekce změny v chování směrodatné odchylky při konstantní střední hodnotě u $N(\mu,\sigma^2)$

Bez újmy na obecnosti lze položit $\mu = 0$. Pak platí

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\frac{x^2}{\sigma^2}}, \qquad f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}.$$

Odtud snadno

$$\frac{g(x)}{f(x)} = \frac{1}{\sigma}e^{-\frac{1}{2}x^2\left(\frac{1}{\sigma^2-1}\right)},$$

a

$$\ln\frac{g(x)}{f(x)} = -\frac{1}{2}\ln\sigma^2 - \frac{1}{2}\left(\frac{1}{\sigma^2}-1\right).$$

Z obecného závěru vyplývá, že změna v parametru σ v okamžiku $(n+1)$ nebude detekována za předpokladu, že doposud detekována nebyla, když

$$-\frac{1}{2}\ln\sigma^2 - \frac{1}{2}\left(\frac{1}{\sigma^2}-1\right)x^2 < \ln\rho,$$

tedy

$$\ln \sigma^2 + \left(\frac{1}{\sigma^2 - 1}\right) x^2 < \ln \rho^2.$$

Nyní je třeba rozlišit, zdali se bude hlídat snížení rozptylu, což znamená $\sigma < 1$, či zvětšení rozptylu, neboli $\sigma > 1$.

Při $\sigma < 1$ dostáváme nerovnost

$$x_{n+1}^2 > \frac{-\sigma^2}{1 - \sigma^2} \ln \rho^2 \, \sigma^2,$$

pro $\sigma > 1$ je rozhodující nerovnost

$$x_{n+1}^2 < \frac{\sigma^2}{\sigma^2 - 1} \ln \rho^2 \, \sigma^2.$$

Pokud nebude změna, pak $x_{n+1} \sim N(0, 1)$ a tedy riziko α chyby 1. druhu je určeno pomocí kvantilu χ^2-rozdělení o 1 stupni volnosti, tedy při $\sigma < 1$ musí pak být

$$\alpha = \mathrm{Pst}\left\{x_{n+1}^2 \leq -\frac{\sigma^2}{1 - \sigma^2} \ln \rho^2 \, \sigma^2\right\},$$

tudíž

$$-\frac{\sigma^2}{1 - \sigma^2} \ln \rho^2 \, \sigma^2 = \chi_\alpha^2(1).$$

Analogicky při $\sigma > 1$ pak musí být

$$\frac{\sigma^2}{\sigma^2 - 1} \ln \rho^2 \, \sigma^2 = \chi_{1-\alpha}^2(1).$$

Jaká je síla tohoto testu? Při $\sigma < 1$ je změna detekována, když

$$x_{n+1}^2 \leq \frac{-\sigma^2}{1 - \sigma^2} \ln \rho^2 \, \sigma^2.$$

Za platnosti změny má podíl $\frac{x_{n+1}^2}{\sigma^2}$ rozdělení $\chi^2(1)$, tedy, aby test měl sílu $1 - \beta$, musí být

$$-\frac{1}{1 - \sigma^2} \ln \rho^2 \, \sigma^2 = \chi_{1-\beta}^2(1).$$

Obdobně při $\sigma > 1$ dospějeme k závěru, že test bude mít sílu $1 - \beta$ právě tehdy, když

$$\frac{1}{\sigma^2 - 1} \ln \rho^2 \, \sigma^2 = \chi_\beta^2(1).$$

Aby tedy test splňoval požadavek na hladinu významnosti α a současně měl sílu $1 - \beta$, musí být splněno zároveň, že (např. při $\sigma < 1$):

$$-\frac{\sigma^2}{1 - \sigma^2} \ln \rho^2 \, \sigma^2 = \chi_\alpha^2(1), \quad \frac{-1}{1 - \sigma^2} \ln \rho^2 \, \sigma^2 = \chi_{1-\beta}^2(1).$$

Je tedy důležitou otázkou, zdali se tento požadavek dá splnit v rozumných relacích pro praxi.

Pokud bychom například požadovali, aby $\alpha = 0,05$ a $\beta = 0,05$, pak dojdeme k požadavku při detekování snížení parametru σ, že musí být $\sigma \doteq 0,0323$. to je ovšem v praxi nerealistický požadavek. Pokud bychom chtěli hledat snížení parametru σ např. na $1/4$, pak síla testu je při $\alpha = 0,05$ přibližně pouze 0,20, a to ještě je nutno parametr ρ položit rovný hodnotě $3,8688$.

Z tohoto je patrno, že test na hledání úrovně variability u rozdělení $N(\mu, \sigma)$ nevykazuje dobré vlastnosti a lze jej klasifikovat jako test slabý.

Pokud sledovaný proces bude probíhat pod rozdělením pravděpodobnosti daném hustotou $f(\cdot)$, budeme říkat, že proces je statisticky zvládnut. Tvar hustoty $f(\cdot)$ obvykle v praxi vyplývá z technické specifikace sledovaného výrobku. Hustota $g(\cdot)$ pak představuje změnu ve vývoji procesu a výskyt nestability. Pro zvládnutelnost musí tedy platit:

$$\frac{p_{n+1}}{p_i} > \prod_{\ell=i}^{n} \frac{g(x_\ell)}{f(x_\ell)}$$

pro každé $i = 1, 2, \ldots, n$.

Hustotu $g(\cdot)$ je možno stanovit na základě volby dvou hustot $g^{-(\cdot)}$ a $g^{+}(\cdot)$, které představují možné mezní změny v chování sledovaného procesu.

Regulační diagram pro sledování střední hodnoty $N(\mu, \sigma^2)$

Pro jednoduchost předpokládejme, že parametr σ je znám a lze položit $\sigma = 1$ (jedná se pouze o změnu měřítka). Pak pro $f(\cdot) \sim N(\mu_0, 1)$ a $g(\cdot) \sim N(\mu_1, 1)$ máme, že změna z μ_0 na μ_1 není detekována v průběhu úseku $\{1, 2, 3, \ldots, n\}$, pokud

$$\ln \frac{p_{n+1}}{p_i} > \sum_{\ell=i}^{n} \frac{1}{2}(x_\ell - \mu_0)^2 - \frac{1}{2}(x_\ell - \mu_1)^2$$

pro $i = 1, 2, \ldots, n$. Označme pro jednoduchost $\frac{p_{n+1}}{p_i} = \rho_i$. Pak musí platit nerovnosti

$$\frac{\ln \rho_i}{\Delta} + \frac{\Delta + 2\mu_0}{2}(n - i + 1) > \sum_{\ell=i}^{n} x_\ell$$

v případě $\mu_1 - \mu_0 = \Delta > 0$. V opačném případě při $\Delta = 0$ pro $\mu_0 - \mu_1 > 0$ získáváme nerovnosti

$$-\frac{\ln \rho_i}{\Delta} - \frac{2\mu_0 + \Delta}{2}(n - i + 1) < \sum_{\ell=i}^{n} x_e ll$$

pro $i = 1, 2, \ldots, n$. V oboustranném případě pro $\mu_1^{+} = \mu_0 + \Delta$ a $\mu_1^{-} = \mu_0 - \Delta$, $\Delta > 0$ pak postupné součty $\sum_{\ell=i}^{n} x_\ell$ se musí objevit mezi horní a dolní hranicí, aby nebyla detekována změna v parametru μ větší v absolutní hodnotě nežli je zvolené $\Delta > 0$. Meze pro jednotlivé součty se liší pouze ve členu $\ln \rho_i/\Delta$. Když bude $\rho_i = \rho$ (např. při rovnoměrném apriorním rozdělení), pak navržený algoritmus má tvar tzv. V-masky známé z regulačního diagramu CUSUM, jedině s tím rozdílem, že u CUSUM diagramu se hlídají součty dopředného typu

$$S_i = \sum_{\ell-1}^{i} x_\ell, \quad i = 1, 2, \ldots, n,$$

kdežto u tohoto algoritmu se sledují součty zpětného typu

$$T_i = \sum_{\ell=i}^{n} x_\ell.$$

Jestliže položíme $\rho_i = 1$, pak změna není detekována pokud pro každé $i = 1, 2, \ldots, n$ platí

$$-\frac{2\mu_0 + \Delta}{2}(n - i + 1) < \sum_{\ell-i}^{n} x_\ell < \frac{2\mu_0 + \Delta}{2}(n - i + 1),$$

což vede na diagram pro zpětné aritmetické průměry

$$-\frac{2\mu_0 + \Delta}{2} < \frac{1}{(n - i + 1)} \sum_{\ell=i}^{n} x_\ell < \frac{2\mu_0 + \Delta}{2},$$

což silně připomíná klasické Shewkartovy regulační diagramy pro aritmetické průměry z podskupin. Rozdíl je v tom, že u klasických Shewhartových diagramů se předpokládá nezávislost mezi podskupinami, což u tohoto regulačního diagramu neplatí, pracuje s pamětí.

Regulační diagram pro rozptyl σ^2 u rozdělení pravděpodobnosti

Pro jednoduchost položme $\mu = 0$, pak při $f \sim N(0, \sigma_0^2)$ a $g \sim N(0, \sigma_1^2)$ máme:

$$\ln \frac{g(x)}{f(x)} = \frac{1}{2} \ln \frac{\sigma_0^2}{\sigma_1^2} + \frac{1}{2}\left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right) x^2.$$

Změna v parametru σ_0^2 nebude v úseku $\{1, 2, \ldots, n\}$ zaznamenána, když

$$\ln \rho_i^2 > (n - i + 1) \ln \frac{\sigma_0^2}{\sigma_1^2} + \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right) \sum_{\ell=i}^{n} x_\ell^2$$

pro každé $i = 1, 2, \ldots, n$.

Je vhodné oddělit případy $\sigma_1 > \sigma_0$ a $\sigma_1 < \sigma_0$. První případ znamená sledování možného zesílení úrovně variability. Zesílení není detekováno, pokud platí nerovnosti:

$$\ln \rho_i^2 \cdot \frac{\sigma_0^2\,\sigma_1^2}{\sigma_1^2 - \sigma_0^2} + (n - i + 1) \ln \frac{\sigma_1^2}{\sigma_0^2} > \sum_{\ell=i}^{n} x_\ell^2$$

pro každé $i = 1, 2, \ldots, n$. Opět je vidět, že zde horní hranice má klesající tendenci, při $\rho_i = 1$ pak dostáváme lineární diagram pro zpětné průměry

$$\frac{\sigma_0^2\,\sigma_1^2}{\sigma_1^2 - \sigma_0^2} \ln \frac{\sigma_1^2}{\sigma_0^2} > \frac{1}{(n-i+1)} \sum_{\ell=i}^{n} x_\ell^2.$$

Ve druhém případě zeslabení, tj. $\sigma_1 < \sigma_0$ je algoritmus detekce založen na spodní hranici tvaru

$$-\frac{\sigma_0^2\,\sigma_1^2}{\sigma_0^2 - \sigma_1^2} \ln \rho_i^2 + \frac{\sigma_0^2 - \sigma_1^2}{\sigma_0^2 - \sigma_1^2}(n-i+1) \ln \frac{\sigma_0^2}{\sigma_1^2} < \sum_{\ell=i}^{n} x_\ell^2,$$

$i = 1, 2, \ldots, n$. Bez újmy na obecnosti lze položit

$$\sigma_1 = t\imath, \sigma_0, \quad t < 1, \qquad \sigma_1 = s\,\sigma_0, \quad s < 1$$

a uvažovat oboustranný případ daný pak horní a dolní mezí, které vymezují prostor ve tvaru “trychtýře”, v němž se musí udržet zpětné součty kvadrátů, aby změna v σ^2 nebyla detekována.

Horní mez má tvar

$$\sigma_0^2 \ln \rho_i^2 \cdot \frac{t^2}{t^2 - 1} + \frac{t^2\,\sigma_0^2}{t^2 - 1}(n-i+1) \ln t^2 > \sum_{\ell=i}^{n} x_\ell^2$$

a dolní mez je ve tvaru

$$-\sigma_0^2 \ln \rho_i^2 \frac{s^2}{1 - s^2} - \frac{s^2\,\sigma_0^2}{1 - s^2}(n-i+1) \ln s^2 < \sum_{\ell=i}^{n} x_\ell^2.$$

Obě meze mají klesající tendenci a musí plati nerovnost

$$\frac{t^2 \ln t^2}{t^2 - 1} > -\frac{s^2 \ln s^2}{1 - s^2},$$

např. pro $s = 1/t$ nerovnost platí.

Reference

[1] J. Michálek, J. Skřivánek: A method of detecting changes in the behaviour of a random sequence based on the Bayes approach. Kybernetika, Vol. 29 (1993), No. 2, pp. 166–179.

[2] J. R. Thompson, J. Koronacki: Statistical Process Control. Chapman Hall/CRC, N. Y., 2002, 2nd edition.