TRIEDA TRANSFORMOVANÝCH LAMBERT *W* $\times$ *GAMA* NÁHODNÝCH VELIČÍN A ICH APLIKÁCIE

Gejza Wimmer, Viktor Witkovský

ROBUST 2014, 19-24.I.2014, Jetřichovice

- **Triedy náhodných veličín s Lambertovým $W \times F$ rozdelením**

- **Testy založené na funkcii podielu vierohodností**

- **Trieda $LW^{\theta}_{\mathcal{F}_\theta}$ transformovaných náhodných veličín s Lambertovým $W \times F_{Gama}$ rozdelením**

  - **Distribúcia náhodnej veličiny $Y \sim LW^{\theta}_{\mathcal{F}_\theta}$ ak $X \sim \chi^2_\nu$**

  - **Distribúcia náhodnej veličiny $Y \sim LW^{\theta}_{\mathcal{F}_{\chi^2_\nu}}$ ak $\boldsymbol{\theta} = (\nu(\ln \nu - 1), \nu, 1)$**

Triedy náhodných veličín s Lambertovým $W \times F$ rozdelením

Georg M. Goerg zaviedol vo veľmi zaujímavom článku [2] triedy zovšeobecnených rozdelení. Pomocou parametrickej nelineárnej transformácie sa zmení náhodná veličina X na tzv. Lambertovu náhodnú veličinu Y, ktorá umožňuje veľmi flexibilný prístup k modelovaniu zošikmených (skeewed) rozdelení.

Iste ste sa všetci stretli so situáciami

- Dáta pochádzajú z rozdelenia "skoro" normálneho, ale zošikmeného doprava, alebo

- Návratnosť aktív má ťažké chvosty, ale údaje sú príliš zošikmené, aby malo zmysel uvažovať $t-$rozdelenie.

Definícia. Nech U je spojitá náhodná veličina s distribučnou funkciou $F_U(u) = P\{U < u\},\ u \in \mathcal{R}$ a hustotou $f_U(u)$. Potom

$$Z = Ue^{\gamma U}, \quad \gamma \in \mathcal{R} \tag{1}$$

je náhodná veličina s Lambertovým $W \times F_U$ rozdelením s parametrom šikmosti γ. Keď má náhodná veličina U dané rozdelenie pravdepodobnosti (napríklad $Gama(a, b)$ rozdelenie), tak túto náhodnú veličinu budeme volať $W \times Gama(a, b)$ náhodná veličina.

Poznámka. Mnohohodnotová funkcia $W(z)$, pre ktorú platí

$$W(z)e^{W(z)} = z, \quad z \in \mathcal{C}$$

sa nazýva Lambertova W funkcia.
Detailnejšie pozri napr. v [1].

Testy založené na funkcii podielu vierohodností

Majme parametrický štatistický model reprezentovaný triedou hustôt

$$\mathcal{M} = \{f(\mathbf{y}|\boldsymbol{\theta}) : \ \mathbf{y} \in \mathcal{Y}, \ \boldsymbol{\theta} \in \Theta\}, \qquad (2)$$

kde $\mathcal{Y}$ je výberový priestor (v našom prípade $\mathcal{R}^n$), $\Theta \subseteq \mathcal{R}^d$ je parametrický priestor, $\mathbf{y}$ je realizácia observačného (náhodného) vektora $\mathbf{Y}$ s hustotou $f(\cdot|\boldsymbol{\theta})$. Logaritmická vierohodnostná funkcia je

$$l(\boldsymbol{\theta}|\mathbf{y}) = \ln f(\mathbf{y}|\boldsymbol{\theta}).$$

Majme nulovú hypotézu $H_0: \ \boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')' \in \Theta_0, \ (\boldsymbol{\theta}_1 \in \mathcal{R}^t, \ \boldsymbol{\theta}_2 \in \mathcal{R}^{d-t})$
a alternatívnu hypotézu $H_a: \ \boldsymbol{\theta} = (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')' \in \Theta_a$, pričom $\Theta_0 \cup \Theta_a = \Theta$.
Testovať $H_0 \leftthreetimes H_a$ môžeme testovacou štatistikou

$$-2\left(\sup_{(\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')' \in \Theta_0} l(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|\mathbf{Y}) - \sup_{(\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')' \in \Theta} l(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|\mathbf{Y}))\right).$$

Nech $l(\hat{\boldsymbol{\theta}}_{1,0}, \hat{\boldsymbol{\theta}}_{2,0}|\mathbf{y}) = \sup_{(\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')' \in \Theta_0} l(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|\mathbf{Y})$. $l(\hat{\boldsymbol{\theta}}_{1,0}, \hat{\boldsymbol{\theta}}_{2,0}|\mathbf{y})$ reprezentuje najlepšie vysvetlenie pre namerané údaje $\mathbf{y}$ pre všetky $(\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')' \in \Theta_0$. Podobne $l(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2|\mathbf{y}) = \sup_{(\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')' \in \Theta} l(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|\mathbf{Y})$ reprezentuje najlepšie vysvetlenie pre namerané údaje $\mathbf{y}$ pre všetky $(\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')' \in \Theta$. Ak $l(\hat{\boldsymbol{\theta}}_{1,0}, \hat{\boldsymbol{\theta}}_{2,0}|\mathbf{y}) = l(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2|\mathbf{y})$, tak najlepšie vysvetlenie pre namerané údaje sa nachádza v Θ_0 a nezamietneme H_0. Ale ak $l(\hat{\boldsymbol{\theta}}_{1,0}, \hat{\boldsymbol{\theta}}_{2,0}|\mathbf{y}) < l(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2|\mathbf{y})$, tak najlepšie vysvetlenie pre namerané údaje sa nachádza v Θ_a a musíme uvažovať zamietnutie H_0 v prospech H_a. Pre štatistiku

$$\lambda(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|\mathbf{Y}) = -2\left(\sup_{(\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')' \in \Theta_0} l(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|\mathbf{Y}) - \sup_{(\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')' \in \Theta} l(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|\mathbf{Y}))\right) \tag{3}$$

platí $0 \le \lambda(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|\mathbf{Y}) < \infty$. Nulová hodnota $\lambda(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|\mathbf{y})$ indukuje, že H_0 nezamietame (pri danej realizácii $\mathbf{y}$). Veľká hodnota $\lambda(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|\mathbf{y})$ indukuje, že najlepšie vysvetlenie pre namerané údaje $\mathbf{y}$ sa nachádza v Θ_a a H_0 zamietame v prospech H_a.

Ak tento test má mať hladinu významnosti α, tak kritická oblasť je

$$U_\alpha = \{\mathbf{y} : \ \lambda(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|\mathbf{y}) > \kappa\},$$

pričom κ spočítame tak, aby za platnosti H_0

$$P\{\lambda(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2|\mathbf{Y}) > \kappa\} = \alpha.$$

Príklad 1. Majme náhodný výber $Y_1, Y_2, ..., Y_n, \quad Y_i \sim N(\mu, \sigma^2)$, pričom μ poznáme, $\sigma > 0$ nepoznáme. Testujeme hypotézu

$$H_0 : \ \sigma = \sigma_0 \qquad \times \qquad H_a : \ \sigma \neq \sigma_0$$

V tomto prípade

$$\Theta = \{(\mu, \sigma) : \ \sigma \in (0, \infty)\}, \qquad \Theta_0 = \{(\mu, \sigma_0)\},$$

$$\Theta_a = \{(\mu, \sigma) : \ \sigma \in (0, \infty), \ \sigma \neq \sigma_0\}$$

a

$$l(\mu, \sigma|\mathbf{Y}) = -\frac{n}{2}\ln 2\pi - n\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \mu)^2.$$

Z rovnice $\left.\frac{\partial l}{\partial \sigma}\right|_{\sigma=\hat{\sigma}} = 0$ dostávame $\frac{\partial l}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(Y_i-\mu)^2 = 0$, čiže

$$\sup_{(\mu,\sigma)\in\Theta} l(\mu,\sigma|\mathbf{Y}) = -\frac{n}{2}\ln 2\pi - n\ln\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i-\mu)^2} - \frac{\sum_{i=1}^{n}(Y_i-\mu)^2}{2\frac{1}{n}\sum_{i=1}^{n}(Y_i-\mu)^2} =$$

$$= -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\frac{1}{n}\sum_{i=1}^{n}(Y_i-\mu)^2 - \frac{n}{2},$$

$$\sup_{(\mu,\sigma)\in\Theta_0} l(\mu,\sigma|\mathbf{Y}) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma_0^2 - \frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(Y_i-\mu)^2.$$

Za platnosti nulovej hypotézy dostávame exaktné rozdelenie $\lambda(\mu,\sigma_0|\mathbf{Y})$:

$$\lambda(\mu,\sigma_0|\mathbf{Y}) = -n\ln\frac{\sum_{i=1}^{n}(Y_i-\mu)^2}{n\sigma_0^2} - n + \frac{\sum_{i=1}^{n}(Y_i-\mu)^2}{\sigma_0^2} \sim Q_n - n - n\ln\frac{Q_n}{n},$$

kde $Q_n \sim \chi_n^2$.

Príklad 2. Majme náhodný výber $Y_1, Y_2, ..., Y_n, \quad Y_i \sim N(\mu, \sigma^2)$, pričom $\mu \in (-\infty, \infty)$ nepoznáme ani $\sigma > 0$ nepoznáme. Testujeme hypotézu

$$H_0: \; \mu = \mu_0, \sigma = \sigma_0 \qquad \asymp \qquad H_a: \; (\mu, \sigma) \neq (\mu_0, \sigma_0).$$

Teraz

$$\Theta = \{(\mu, \sigma): \; -\infty < \mu < \infty, \sigma \in (0, \infty)\},$$

$$\Theta_0 = \{(\mu_0, \sigma_0)\}, \qquad \Theta_a = \{(\mu, \sigma): \; -\infty < \mu < \infty, \sigma \in (0, \infty), \mu \neq \mu_0, \sigma \neq \sigma_0\}$$

a

$$l(\mu, \sigma|\mathbf{Y}) = -\frac{n}{2}\ln 2\pi - n\ln\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \mu)^2.$$

Z rovníc $\left.\frac{\partial l}{\partial \mu}\right|_{\substack{\mu=\hat{\mu}\\ \sigma=\hat{\sigma}}} = 0,\ \left.\frac{\partial l}{\partial \sigma}\right|_{\substack{\mu=\hat{\mu}\\ \sigma=\hat{\sigma}}} = 0$ dostávame $\hat{\mu} = \overline{Y},\ \ \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (Y_i - \overline{Y})^2$ a teda

$$\sup_{(\mu,\sigma)\in\Theta} l(\mu,\sigma|\mathbf{Y}) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\frac{1}{n}\sum_{i=1}^n (Y_i - \overline{Y})^2 - \frac{n}{2},$$

$$\sup_{(\mu,\sigma)\in\Theta_0} l(\mu,\sigma|\mathbf{Y}) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma_0^2 - \frac{1}{2\sigma_0^2}\sum_{i=1}^n (Y_i - \mu_0)^2.$$

Za platnosti nulovej hypotézy dostávame exaktné rozdelenie $\lambda(\mu_0,\sigma_0|\mathbf{Y})$:

$$\lambda(\mu_0,\sigma_0|\mathbf{Y}) = \frac{\sum_{i=1}^n (Y_i - \overline{Y})^2}{\sigma_0^2} - n - n\ln\frac{\sum_{i=1}^n (Y_i - \overline{Y})^2}{n\sigma_0^2} + \frac{n(\overline{Y} - \mu_0)^2}{\sigma_0^2} \sim$$

$$\sim Q_{n-1} - n - n\ln\frac{Q_{n-1}}{n} + Q_1,$$

kde $Q_{n-1} \sim \chi^2_{n-1},\ Q_1 \sim \chi^2_1$ a $Q_{n-1},\ Q_1$ sú nezávislé.

Príklad 3. Majme lineárny regresný model $\mathbf{Y} = \mathbf{X}\beta + \sigma\varepsilon$, $\mathbf{Y} \sim N_n(\mathbf{X}_{n,k}\boldsymbol{\beta}, \sigma^2\mathbf{I}), \boldsymbol{\beta} \in \mathcal{R}^k, \sigma > 0,\ h(\mathbf{X}) = k < n$. Testujeme hypotézu

$$H_0 :\ (\boldsymbol{\beta}, \sigma) = (\boldsymbol{\beta}_0, \sigma_0) \qquad \asymp \qquad H_a :\ (\boldsymbol{\beta}, \sigma) \neq (\boldsymbol{\beta}_0, \sigma_0).$$

Za platnosti nulovej hypotézy možno odvodiť (pozri [3]) exaktné rozdelenie $\lambda(\boldsymbol{\beta}_0, \sigma_0|\mathbf{Y})$:

$$\lambda(\boldsymbol{\beta}_0, \sigma_0|\mathbf{Y}) \sim Q_{n-k} - n - n \ln \frac{Q_{n-k}}{n} + Q_k,$$

kde $Q_{n-k} \sim \chi^2_{n-k},\ Q_k \sim \chi^2_k$ a $Q_{n-k},\ Q_k$ sú nezávislé.

Príklad 4. Uvažujme lineárny zmiešaný model s dvomi variančnými komponentami

$$\mathbf{Y} = \mathbf{X}_{n,k}\boldsymbol{\beta} + \sigma_A \mathbf{U}_{n,q}\boldsymbol{\eta} + \sigma\boldsymbol{\varepsilon},$$

$\sigma_A^2 \geq 0,\ \sigma^2 > 0,\ \boldsymbol{\eta} \sim N_q(\mathbf{0}, \mathbf{A}_{q,q}),\ \boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \mathbf{I}),\ \mathbf{X}, \mathbf{U}$ a $\mathbf{A}$ sú známe matice, $\boldsymbol{\eta}, \boldsymbol{\varepsilon}$ sú nezávislé, $h(\mathbf{X}) = \nu$. Inferencia o (σ_A^2, σ^2) môže byť založená na vhodne transformovanom vektore $\widetilde{\mathbf{Y}} = \mathbf{B}_{\nu,n}\mathbf{Y}$, pričom $\widetilde{\mathbf{Y}} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} = \sigma_A^2 \mathbf{BVB}' + \sigma^2 \mathbf{I}_\nu = \sum_{i=1}^r (\sigma_A^2 \varrho_i + \sigma^2)\mathbf{E}_i$, $\mathbf{V} = \mathbf{UAU}'$, $\mathbf{BVB}' = \sum_{i=1}^r \varrho_i \mathbf{E}_i$ je spektrálny rozklad matice $\mathbf{BVB}'$ s vlastnými číslami ϱ_i, $(\varrho_1 > \varrho_2 > ... > \varrho_r \geq 0)$ a ich násobky sú $\nu_i,\ \sum_{i=1}^r \nu_i = \nu$. Označme kanonické parametre $\theta_i = (\sigma_A^2 \varrho_i + \sigma^2) > 0,\ i = 1, 2, ..., r,\ \boldsymbol{\theta} = (\theta_1, ..., \theta_r)',\ T_i = \widetilde{\mathbf{Y}}'\mathbf{E}_i\widetilde{\mathbf{Y}},\ i = 1, 2, ..., r,\ \mathbf{T} = (T_1, ..., T_r)'$. Testujeme hypotézu

$$H_0: \ \boldsymbol{\theta} = \boldsymbol{\theta}_0 \qquad \asymp \qquad H_a: \ \boldsymbol{\theta} \neq \boldsymbol{\theta}_0.$$

Za platnosti nulovej hypotézy možno odvodiť (pozri [3]) exaktné rozdelenie $\lambda(\boldsymbol{\theta}_0|\mathbf{T})$:

$$\lambda(\boldsymbol{\theta}_0, \sigma_0|\mathbf{T}) \sim \sum_{i=1}^r \left(Q_{\nu_i} - \nu_i - \nu_i \ln \frac{Q_{\nu_i}}{\nu_i} \right),$$

kde $Q_{\nu_i} \sim \chi^2_{\nu_i}$ sú navzájom nezávislé.

Trieda transformovaných náhodných veličín s Lambertovým $W \times F_{Gama}$ rozdelením

V každom z predchádzajúcich príkladov sme sa stretli s náhodnou veličinou

$$Y = (Q_\nu - c) - c \ln \frac{Q_\nu}{c}, \tag{4}$$

kde $Q_\nu \sim \chi^2_\nu$, $c > 0$. Po úprave

$$e^{-\frac{Y+c}{c}} = \frac{Q_\nu}{c} e^{-\frac{Q_\nu}{c}}.$$

Pri označení $\frac{Q_\nu}{c} = U \sim Gama(\frac{\nu}{2}, \frac{1}{2c})$, $\gamma = -1$ dostávame, že $Z = e^{-\frac{Y+c}{c}} = Ue^{-U}$ patrí do triedy náhodných veličín s Lambertovým $W \times F_{Gama}$ rozdelením a náhodná veličina

$$Y = -c - c \ln U + cU$$

je (vhodne) transformovaná náhodná veličina s Lambertovým $W \times F_{Gama}$ rozdelením.

V ďalšom budeme skúmať triedu náhodných veličín

$$Y = \theta_1 - \theta_2 \ln X + \theta_3 X\,, \tag{5}$$

kde X je spojitá náhodná veličina, ktorá nadobúda nezáporné hodnoty a závisí od parametrov ϑ, $\theta_1 \in \mathcal{R}$, $\theta_2 > 0$, $\theta_3 > 0$. Jej distribučnú funkciu označme $\mathcal{F}_\vartheta$. Náhodná veličina Y nadobúda reálne hodnoty $y \in \langle y_{min}, \infty)$, kde $y_{min} = \theta_1 + \theta_2 - \theta_2 \ln\left(\frac{\theta_2}{\theta_3}\right)$ v bode $x_{y_{min}} = \frac{\theta_2}{\theta_3}$, $0 < x_{y_{min}} < \infty$. Aplikáciou Lambertovej W funkcie dostaneme explicitnú inverznú transformáciu funkcie (5) a takto stanovíme exaktnú distribúciu náhodnej premennej Y ak máme distribúciu X. Distribúciu náhodnej veličiny Y označíme $LW^{\theta}_{\mathcal{F}_\vartheta}$, $Y \sim LW^{\theta}_{\mathcal{F}_\vartheta}$. Ak $x^{\theta}_L(y)$ a $x^{\theta}_U(y)$ sú dve rôzne riešenia rovnice $y = \theta_1 - \theta_2 \ln x + \theta_3 x$, pričom $x^{\theta}_L(y)$ je riešenie na intervale $(0, x_{y_{min}})$ a $x^{\theta}_U(y)$ je riešenie na intervale $(x_{y_{min}}, \infty)$ tak distribučná funkcia Y označená ako $F_{LW^{\theta}_{\mathcal{F}_\vartheta}}(\cdot)$ v bode y sa rovná

$$F_{LW^{\theta}_{\mathcal{F}_\vartheta}}(y) = F_{LW^{\theta}_{\mathcal{F}_\vartheta}}(x^{\theta}_U(y)) - F_{LW^{\theta}_{\mathcal{F}_\vartheta}}(x^{\theta}_L(y)). \tag{6}$$

Konkrétne

$$x_L^\theta(y) = -\frac{\theta_2}{\theta_3} W_0\left(-\frac{\theta_3}{\theta_2} e^{-\frac{y-\theta_1}{\theta_2}}\right), \quad x_U^\theta(y) = -\frac{\theta_2}{\theta_3} W_{-1}\left(-\frac{\theta_3}{\theta_2} e^{-\frac{y-\theta_1}{\theta_2}}\right), \tag{7}$$

kde $W_0(\cdot)$ a $W_{-1}(\cdot)$ sú dve reálne vetvy multihodnotovej Lamberovej W funkcie ($z = W(z)e^{W(z)}$) (viac pozri napr. [1]). Poznamenávame, že

$$\frac{d}{dy} x_L^\theta(y) = \frac{x_L^\theta(y)}{\theta_3 x_L^\theta(y) - \theta_2}, \quad \frac{d}{dy} x_U^\theta(y) = \frac{x_U^\theta(y)}{\theta_3 x_U^\theta(y) - \theta_2}. \tag{8}$$

Z (6) a (8) priamo dostávame hustotu náhodnej premennej $Y \sim LW^{\theta}_{\mathcal{F}_{\vartheta}}$:

$$f_{LW^{\theta}_{\mathcal{F}_{\vartheta}}}(y) = \frac{x^{\theta}_{L}(y)}{\theta_2 - \theta_3 x^{\theta}_{L}(y) x} f_{\mathcal{F}_{\vartheta}}\left(x^{\theta}_{L}(y)\right) + \frac{x^{\theta}_{U}(y)}{\theta_3 x^{\theta}_{U}(y) x - \theta_2} f_{\mathcal{F}_{\vartheta}}\left(x^{\theta}_{U}(y)\right), \qquad (9)$$

kde $f_{\mathcal{F}_{\vartheta}}(\cdot)$ je hustota náhodnej veličiny X. Ak $y_{1-\alpha}$ je $(1-\alpha)-$kvantil náhodnej veličiny Y, tak

$$y_{1-\alpha} = F^{-1}_{LW^{\theta}_{\mathcal{F}_{\vartheta}}}(1-\alpha),$$

kde $F^{-1}_{LW^{\theta}_{\mathcal{F}_{\vartheta}}}(\cdot)$ je kvantilová funkcia náhodnej premennej Y.

Distribúcia náhodnej veličiny $Y \sim LW^{\theta}_{\mathcal{F}_\vartheta}$ ak $X \sim \chi^2_\nu$

Pretože distribučnú funkciu $F_{\chi^2_\nu}(\cdot)$ náhodnej veličiny s $X \sim \chi^2_\nu$ vieme vyjadriť pomocou neúplnej Gama funkcie $\Gamma(n, a, b) = \int_a^b t^{n-1} e^{-t} dt$, priamo zo vzťahu (6) dostávame

$$F_{LW^{\theta}_{\mathcal{F}_\vartheta}}(y) = F_{LW^{\theta}_{\mathcal{F}_{\chi^2_\nu}}}(y) = \frac{1}{\Gamma\left(\frac{\nu}{2}\right)} \Gamma\left(\frac{\nu}{2}, x^{\theta}_L(y), x^{\theta}_U(y)\right).$$

Pretože hustota $f_{\chi^2_\nu}(\cdot)$ náhodnej veličiny $X \sim \chi^2_\nu$ je

$$f_{\chi^2_\nu}(x) = \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}, \; x > 0, \qquad f\chi^2_\nu(x) = 0, \; x \leq 0,$$

priamo z (9) dostávame

$$f_{LW^{\theta}_{\mathcal{F}_{\chi^2_\nu}}}(y) = \frac{1}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)} \left(\frac{\left(x^{\theta}_L(y)\right)^{\frac{\nu}{2}} e^{-\frac{1}{2} x^{\theta}_L(y)}}{\theta_2 - \theta_3 x^{\theta}_L(y)} + \frac{\left(x^{\theta}_U(y)\right)^{\frac{\nu}{2}} e^{-\frac{1}{2} x^{\theta}_U(y)}}{\theta_3 x^{\theta}_U(y) - \theta_2} \right).$$

Charakteristická funkcia $chf_{LW^{\theta}_{\mathcal{F}_{\chi^2_\nu}}}(\cdot)$ náhodnej veličiny $Y \sim LW^{\theta}_{\mathcal{F}_{\chi^2_\nu}}$ je

$$chf_{LW^{\theta}_{\mathcal{F}_{\chi^2_\nu}}}(t) = \mathcal{E}\left(e^{itY}\right) = \frac{e^{it\theta_1}\Gamma\left(\frac{\nu}{2} - it\theta_2\right)}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)\left(\frac{1}{2} - it\theta_3\right)^{\frac{\nu}{2} - it\theta_2}}.$$

Momentová generujúca funkcia

$$mgf_{LW^{\theta}_{\mathcal{F}_{\chi^2_\nu}}}(t) = chf_{LW^{\theta}_{\mathcal{F}_{\chi^2_\nu}}}(it) \quad \text{pre} \quad t < \frac{1}{2}\min\left(\frac{\nu}{\theta_2}, \frac{1}{\theta_3}\right).$$

Kumulanty κ_j, $j = 1, 2, ...$ náhodnej veličiny $Y \sim LW^{\theta}_{\mathcal{F}_{\chi^2_\nu}}$ dostaneme rozvojom logaritmu momentovej vytvárajúcej funkcie:

$$\kappa_1 = \theta_1 - \theta_2 \ln 2 + \nu\theta_3 - \theta_2\psi^{(0)}\left(\frac{\nu}{2}\right),$$

$$\kappa_j = 2^{j-1}\Gamma(j-1)\theta_3^{j-1}(-j\theta_2 + (j-1)\nu\theta_3) + (-1)^j\theta_2^j\psi^{(j-1)}\left(\frac{\nu}{2}\right), \quad j = 2, 3, ...,$$

pričom $\psi^{(m)}(\cdot)$ je polygamma funkcia $m-$tého rádu (t.j. $(m+1)-$vá derivácia logaritmu Gama funkcie).

Takto dostávame strednú hodnotu náhodnej premennej Y

$$\mathcal{E}(Y) = \kappa_1 = \theta_1 - \theta_2 \ln 2 + \nu\theta_3 - \theta_2\psi^{(0)}\left(\frac{\nu}{2}\right),$$

rozptyl

$$\mathcal{D}(Y) = \kappa_2 = 2\theta_3(-2\theta_2 + \nu\theta_3) + \theta_2^2\psi^{(1)}\left(\frac{\nu}{2}\right),$$

koeficient šikmosti

$$\frac{\kappa_3}{\kappa_2^{\frac{3}{2}}} = \frac{4\theta_3^2(-3\theta_2 + 2\nu\theta_3) - \theta_2^3\psi^{(2)}\left(\frac{\nu}{2}\right)}{\left(2\theta_3(-2\theta_2 + \nu\theta_3) + \theta_2^2\psi^{(1)}\left(\frac{\nu}{2}\right)\right)^{\frac{3}{2}}},$$

a koeficient špicatosti

$$\frac{\kappa_4}{\kappa_2^2} = \frac{16\theta_3^3(-4\theta_2 + 3\nu\theta_3) + \theta_2^4\psi^{(3)}\left(\frac{\nu}{2}\right)}{\left(2\theta_3(-2\theta_2 + \nu\theta_3) + \theta_2^2\psi^{(1)}\left(\frac{\nu}{2}\right)\right)^2}.$$

Poznamenávame len, že pre náhodnú veličina Y danú v (4), ak $X \sim \chi^2_\nu$, platí $Y \sim LW^{\boldsymbol{\theta}}_{\mathcal{F}_{\chi^2_\nu}}$, kde $\boldsymbol{\theta} = (c(\ln c - 1), c, 1)'$.

Distribúcia náhodnej veličiny $Y \sim LW^{\theta}_{\mathcal{F}_{\chi^2_\nu}}$ ak $\boldsymbol{\theta} = (\nu(\ln \nu - 1), \nu, 1)$

Ústrednú roľu v štatistických inferenciách založených na pomere vierohodností pre normálne rozdelené dáta hrá rozdelenie náhodnej veličiny

$$Y = (Q_\nu - \nu) - \nu \ln\left(\frac{Q_\nu}{\nu}\right),$$

teda $Y \sim LW^{\theta}_{\mathcal{F}_{\chi^2_\nu}}$, kde $\boldsymbol{\theta} = (\nu(\ln \nu - 1), \nu, 1)$.

Základné pravdepodobnostné charakteristiky náhodnej veličiny $Y \sim LW^{\theta}_{\mathcal{F}_{\chi^2_\nu}}, \ \boldsymbol{\theta} = (\nu(\ln \nu - 1), \nu, 1)$:

distribučná funkcia

$$F_{LW^{\theta}_{\mathcal{F}_{\chi^2_\nu}}}(y) = {}_{\nu}F(x) = \begin{cases} \dfrac{1}{\Gamma\left(\frac{\nu}{2}\right)}\left[\gamma\left(\dfrac{\nu}{2}, \dfrac{x_{2,y}}{2}\right) - \gamma\left(\dfrac{\nu}{2}, \dfrac{x_{1,y}}{2}\right)\right], & \text{ak } x > 0, \\ 0, & \text{ak } x \leq 0, \end{cases}$$

kde $\Gamma(a, 0, x) = \int_0^x t^{a-1}e^{-t}dt,\ a > 0,\ x > 0$ (neúplná Gama funkcia), $x_{1,y}$ je riešenie rovnice

$$y = -\nu \ln \frac{x}{\nu} + x - \nu \tag{10}$$

pre $x \in (0, \nu)$ a $x_{2,y}$ je riešenie rovnice (10) pre $x \in (\nu, \infty)$, *hustota*

$$ {}_{\nu}f(y) = \begin{cases} \left|\dfrac{x_{1,y}}{x_{1,y}-\nu}\right| f_{\chi^2_\nu}(x_{1,y}) + \left|\dfrac{x_{2,y}}{x_{2,y}-\nu}\right| f_{\chi^2_\nu}(x_{2,y}), & \text{ak } y > 0, \\ 0, & \text{ak } y \le 0, \end{cases} $$

charakteristická funkcia (chf)

$$ {}_{\nu}\,chf(t) = \frac{\Gamma\left(\frac{\nu}{2} - \nu t\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(\frac{\nu}{2e}\right)^{\nu t} (1-2it)^{\nu t - \frac{\nu}{2}}, $$

momentová generujúca funkcia (mgf)

$$
{}_{\nu}\, mgf(t) = \frac{\Gamma\left(\frac{\nu}{2} - \nu t\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(\frac{\nu}{2e}\right)^{\nu t} (1-2t)^{\nu t - \frac{\nu}{2}}, \quad 0 \le t < \frac{1}{2}, \tag{11}
$$

počiatočné momenty

$$
\mu_1' = \mu = -\nu \frac{\Gamma'\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} + \nu \ln \tfrac{\nu}{2},
$$

$$
\mu_2' = \nu^2 \frac{\Gamma''\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} + \nu^2 (\ln \tfrac{\nu}{2})^2 - 2\nu - 2\nu^2 \frac{\Gamma'\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \ln \tfrac{\nu}{2},
$$

$$
\mu_3' = \nu^3 \frac{\Gamma'''\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} + \nu^3 (\ln \tfrac{\nu}{2})^3 - 4\nu + 3\nu^3 \frac{\Gamma''\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \ln \tfrac{\nu}{2} - 3\nu^3 \frac{\Gamma'\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} (\ln \tfrac{\nu}{2})^2 +
$$

$$
+ 6\nu^2 \frac{\Gamma'\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} - 6\nu^2 \ln \tfrac{\nu}{2},
$$

centrálne momenty

$$\mu_2 = -\nu \left(2 - \nu \left\{ \frac{\Gamma'(z)}{\Gamma(z)} \right\}'_{z=\frac{\nu}{2}} \right),$$

$$\mu_3 = -\nu \left(4 - \nu^2 \left\{ \frac{\Gamma'(z)}{\Gamma(z)} \right\}''_{z=\frac{\nu}{2}} \right),$$

kumulanty

$$\kappa_2 = -2,$$

$$\kappa_k = -\nu \left(a_k + (-\nu)^{k-1} \ {}_{\nu}\psi^{(k-1)} \left(\frac{\nu}{2} \right) \right), \ a_2 = 2, \ a_k = 2a_{k-1}(k-1) \ k = 3, 4, \ldots \ .$$

Poznámka. Po zdĺhavých výpočtoch sa dá ukázať, že momentová generujúca funkcia (11) konverguje pre $\nu \to \infty$ k momentovej generujúcej funkcii χ_1^2 rozdelenia, t.j. k funkcii $(1-2t)^{-1/2}$. Náhodná veličina $Y \sim LW^{\theta}_{\mathcal{F}_{\chi^2_\nu}}$, $\boldsymbol{\theta} = (\nu(\ln \nu - 1), \nu, 1)$ konverguje v distribúcii k náhodnej veličine s χ_1^2 rozdelením.

Poznámka. Invertovaním testu pomerom vierohodnosti priamo dostaneme realizáciu exaktnej simultánnej $(1-\alpha)-$konfidenčnej oblasti pre parametre modelu (2), ktorá (oblasť) je založená na funkcii vierohodnosti.

Poďakovanie. Výskum bol podporený grantami APVV/0096/10 a VEGA 2/0038/12.

Literatúra

Corless, R. M., Gonnet, G. H., Hare, D. E. G., Jeffery, G. J., Knuth, D. E.:
On the Lambert W function.
Advances in Computational Mathematics **5** (1996), 329–359.

Goerg, G., M.:
Lambert W Random Variables - a New Family of Generalized skewed Distributions with Applications to Risk Estimation.
The Annals of Applied Statistics **5**(3) (2011), 1–34.

Witkovský, V.:
O niektorých exaktných simultánnych konfidenčných oblastiach založených na funkcii vierohodnosti pre parametre normálneho LRM.
Sborník XIX. letní školy biometriky, Lednica 6.-10.9.2010.

Ďakujeme za pozornosť