Testowanie hipotezy o wzajemnym położeniu względem siebie rozkładów uporządkowanych dyspersyjnie

Magdalena Frąszczak

Wisła, 6 grudnia 2011

Oznaczenia

- $X \sim F$, $Y \sim G$ zmienne losowe o gęstościach f i g odpowiednio;
- $\bar{F} = 1 - F$ - funkcja przeżycia (ogon rozkładu);
- $F^{-1}(u) = \inf\{x : F(x) \geqslant u\}$, $u \in (0,1)$ - funkcja kwantylowa;
- $\frac{d}{du}F^{-1}(u) = 1/fF^{-1}(u)$ – gęstość kwantylowa rozkładu F;
- $\varphi\psi$ oznacza złożenie dwóch funkcji φ i ψ.
- $\mathcal{F}$ – klasa absolutnie ciągłych rozkładów F określonych na $[0,\infty)$, takich że $F(0) = 0$;
- $\mathbf{\Phi}$ – klasa ciągłych, dodatnich funkcji φ określonych na odcinku $(0,1)$.

Uogólniona transformata TTT

- Transformata TTT (*total time on test transform*)

$$H_F^{-1}(t) = \int_0^{F^{-1}(t)} \bar{F}(x)dx$$

- Uogólniona transformata TTT (transformata GTTT) (*generalized total time on test transform*) rozkładu F indukowaną przez funkcje $\varphi \in \mathbf{\Phi}$ (klasa funkcji ciągłych, dodatnich, określonych na $(0,1)$)

$$H_F^{-1}(t,\varphi) = \int_0^{F^{-1}(t)} \varphi F(x)dx$$

Porządek dyspersyjny

Definicja 1.

$F \leqslant_{\mathrm{disp}} G \ \equiv\ G^{-1}F(x) - x$ jest rosnąca;

Porządek dyspersyjny

Definicja 1.

$F \leqslant_{\mathrm{disp}} G \;\equiv\; G^{-1}F(x) - x$ jest rosnąca;

Lemat 1.

Niech F i G będą absolutnie ciągłymi rozkładami z gęstościami f i g odpowiednio oraz $F(0) = G(0) = 0$. Wówczas
$F \leqslant_{\mathrm{disp}} G \Leftrightarrow fF^{-1}(u)/gG^{-1}(u) \geqslant 1, \quad u \in (0,1)$.

Porządek dyspersyjny

Definicja 1.

$F \leqslant_{\mathrm{disp}} G \;\equiv\; G^{-1}F(x) - x$ jest rosnąca;

Lemat 1.

Niech F i G będą absolutnie ciągłymi rozkładami z gęstościami f i g odpowiednio oraz $F(0) = G(0) = 0$. Wówczas
$F \leqslant_{\mathrm{disp}} G \Leftrightarrow fF^{-1}(u)/gG^{-1}(u) \geqslant 1, \quad u \in (0,1)$.

Twierdzenie 1.

Niech $F, G \in \mathcal{F}$ oraz $\varphi_1, \varphi_2 \in \mathbf{\Phi}$. Wtedy
jeśli $F \leqslant_{\mathrm{disp}} G$ i $\varphi_2(u)/\varphi_1(u) \geqslant 1$ dla $u \in (0,1)$, to
$H_F(\cdot;\varphi_1) \leqslant_{\mathrm{disp}} H_F(\cdot;\varphi_2) \leqslant_{\mathrm{disp}} H_G(\cdot;\varphi_2)$.

Porządki stochastyczne niezmiennicze względem transformat GTTT

Niech $\mathcal{F}$ będzie klasą absolutnie ciągłych rozkładów F określonych na $[0, \infty)$, takich że $F(0) = 0$
Niech $\mathbf{\Phi}$ będzie klasą ciągłych, dodatnich funkcji φ określonych na odcinku $(0, 1)$.
Niech S oznacza porządki stochastyczne dające się scharakteryzować przez złożenie funkcji $G^{-1}F$

Definicja 2.

Półporządek $\mathcal{S}$ jest niezmienniczy względem transformat GTTT, jeśli
$F \leqslant_S G \quad \Rightarrow \quad H_F(\cdot;\varphi) \leqslant_S H_G(\cdot;\varphi)$ dla wszystkich $\varphi \in \mathbf{\Phi}$.

$$\mathbf{O}(F,G) = \{(H_F(\cdot;\varphi), H_G(\cdot;\varphi)), \varphi \in \mathbf{\Phi}\}$$

Twierdzenie 2.

Funkcja

$$\kappa(u) = \frac{fF^{-1}(u)}{gG^{-1}(u)}, \qquad u \in (0,1),$$

jest maksymalnym niezmiennikiem względem transformat GTTT indukowanych przez klasę $\mathbf{\Phi}$.

Wzajemne położenie względem siebie uporządkowanych rozkładów

Niech $\kappa_i = f_i F_i^{-1} / g_i G_i^{-1}$, $i = 1, 2$

Definicja 3.

Niech $F_1 \leqslant_S G_1$, $F_2 \leqslant_S G_2$. Wówczas G_2 jest bardziej na prawo od F_2 w porządku S niż G_1 od F_1, gdy

$$\frac{\kappa_2}{\kappa_1} \in \mathbf{K}_S.$$

W szczególności, $\mathbf{K}_{disp}$ jest klasą funkcji przyjmujących wartości nie mniejsze niż jeden.

Wzajemne położenie względem siebie rozkładów uporządkowanych dyspersyjnie

Niech będzie dana czwórka rozkładów $(F_1, G_1; F_2, G_2)$, uporządkowanych parami w porządku dyspersyjnym, tzn. niech $F_1 \leqslant_{disp} G_1$, $F_2 \leqslant_{disp} G_2$. Wówczas:
G_2 jest bardziej na prawo od F_2 w porządku S niż G_1 od F_1, gdy

$$\frac{f_2 F_2^{-1}}{f_1 F_1^{-1}} \cdot \frac{g_1 G_1^{-1}}{g_2 G_2^{-1}} \geqslant 1$$

Testowanie hipotezy o wzajemnym położeniu względem siebie rozkładów

Niech będą dane rozkłady absolutnie ciągłe F_0 i G_0, takie, że $F_0 \leqslant_{\mathrm{disp}} G_0$.
Niech $\mathbf{X} = (X_1, X_2, \ldots X_n)'$ i $\mathbf{Y} = (Y_1, Y_2, \ldots Y_n)'$ będą próbami z nieznanych rozkładów absolutnie ciągłych F i G odpowiednio.

Testowanie hipotezy o wzajemnym położeniu względem siebie rozkładów

Niech będą dane rozkłady absolutnie ciągłe F_0 i G_0, takie, że $F_0 \leqslant_{\text{disp}} G_0$.
Niech $\mathbf{X} = (X_1, X_2, \ldots X_n)'$ i $\mathbf{Y} = (Y_1, Y_2, \ldots Y_n)'$ będą próbami z nieznanych rozkładów absolutnie ciągłych F i G odpowiednio.

Wiemy, że istnieje funkcja $\varphi_0 \in \mathbf{\Phi}$, taka że $F = H_{F_0}(\cdot, \varphi_0)$ i jest ona postaci $\varphi_0 = f_0 F_0^{-1}/fF^{-1}$.

Problem testowania hipotez

$$\mathcal{H}_0: \quad G = H_{G_0}(\cdot, \varphi_0),$$
$$\mathcal{H}_1: \quad G \geqslant_{\text{disp}} H_{G_0}(\cdot, \varphi_0), \qquad G \neq H_{G_0}(\cdot, \varphi_0).$$

Testowanie hipotezy o wzajemnym położeniu względem siebie rozkładów

Równoważny problem testowania hipotez

$$\mathcal{H}_0' \colon \frac{fF^{-1}(t)}{gG^{-1}(t)} - c(t) = 0, \quad \text{dla każdego } t \in (0,1),$$

przy alternatywie

$$\mathcal{H}_1' \colon \frac{fF^{-1}(t)}{gG^{-1}(t)} - c(t) \geqslant 0, \quad \text{dla każdego } t \in (0,1),$$

przy czym dla pewnego $t' \in (0,1)$, $\dfrac{fF^{-1}(t')}{gG^{-1}(t')} - c(t') > 0$, gdzie $c(t) = \dfrac{f_0 F_0^{-1}(t)}{g_0 G_0^{-1}(t)}$.

Rozważamy przypadek, gdy F_0 i G_0 są rozkładami różniącymi się parametrami skali, tj. $F_0(x) = M(x/\lambda_F)$ i $G_0(x) = M(x/\lambda_G)$, $\lambda_F \neq \lambda_G$, gdzie M jest pewnym rozkładem absolutnie ciągłym. Wówczas

$$c(t) = \frac{f_0 F_0^{-1}(t)}{g_0 G_0^{-1}(t)} = \frac{\lambda_G}{\lambda_F} = \exp\{d_0(F_0, G_0)\} = c = \mathrm{const}.$$

Statystyka testowa

$$T(\mathbf{X}, \mathbf{Y}) = \max_{1 \leqslant i \leqslant n-1} \frac{Y_{i+1:n} - Y_{i:n}}{X_{i+1:n} - X_{i:n}} - c.$$

Test na poziomie istotności α dla testowania hipotezy $\mathcal{H}_0'$

$$\varphi(\mathbf{X}, \mathbf{Y}) = \begin{cases} 1, & \text{gdy} \quad T(\mathbf{X}, \mathbf{Y}) \geqslant C_0, \\ 0, & \text{gdy} \quad T(\mathbf{X}, \mathbf{Y}) < C_0. \end{cases}$$

Przy prawdziwości hipotezy $\mathcal{H}_0'$

$$\lim_{n\to\infty} P\left(\frac{1}{n} \max_{1\leqslant i\leqslant n-1} \frac{Y_{i+1:n} - Y_{i:n}}{X_{i+1:n} - X_{i:n}} \leqslant x\right) = e^{-c/x}, \quad x \geqslant 0.$$

Test asymptotyczny odrzuca hipotezę $\mathcal{H}_0'$ na rzecz alternatywy $\mathcal{H}_1'$ na poziomie istotności w przybliżeniu α, gdy

$$\max_{1\leqslant i\leqslant n-1} \frac{Y_{i+1:n} - Y_{i:n}}{X_{i+1:n} - X_{i:n}} \geqslant -\frac{c \cdot n}{\log \alpha}.$$

Szkic dowodu

- Niech $\Lambda = F^{-1}K$, gdzie $K(t) = 1 - e^{-t}$, $t > 0$, wówczas

$$X_{i+1:n} - X_{i:n} = \Lambda K^{-1}F(X_{i+1:n}) - \Lambda K^{-1}F(X_{i:n}).$$

Szkic dowodu

- Niech $\Lambda = F^{-1}K$, gdzie $K(t) = 1 - e^{-t}$, $t > 0$, wówczas

$$X_{i+1:n} - X_{i:n} = \Lambda K^{-1}F(X_{i+1:n}) - \Lambda K^{-1}F(X_{i:n}).$$

- Stosując twierdzenie o wartości średniej Lagrange'a do funkcji Λ, otrzymujemy

$$X_{i+1:n} - X_{i:n} = \frac{S_i}{(n-i+1)\tilde{r}_F(\theta^F_{i:n})},$$

gdzie

$$S_i = (n-i+1)[K^{-1}F(X_{i+1:n}) - K^{-1}F(X_{i:n})],$$

$$\tilde{r}_F(x) = r(F^{-1}(x)), \quad F(X_{i:n}) \leqslant \theta^F_{i:n} \leqslant F(X_{i+1:n}).$$

Szkic dowodu:

- Przy hipotezie $\mathcal{H}_0'$ mamy $\tilde{r}_F = c\tilde{r}_G$. Stąd

$$\frac{Y_{i+1:n} - Y_{i:n}}{X_{i+1:n} - X_{i:n}} = \frac{T_i}{S_i}\frac{\tilde{r}_F(\theta^F_{i:n})}{\tilde{r}_G(\theta^G_{i:n})} = c\frac{T_i}{S_i}\frac{\tilde{r}_F(\theta^F_{i:n})}{\tilde{r}_F(\theta^G_{i:n})},$$

Szkic dowodu:

- Przy hipotezie $\mathcal{H}_0'$ mamy $\tilde{r}_F = c\tilde{r}_G$. Stąd

$$\frac{Y_{i+1:n} - Y_{i:n}}{X_{i+1:n} - X_{i:n}} = \frac{T_i}{S_i}\frac{\tilde{r}_F(\theta^F_{i:n})}{\tilde{r}_G(\theta^G_{i:n})} = c\frac{T_i}{S_i}\frac{\tilde{r}_F(\theta^F_{i:n})}{\tilde{r}_F(\theta^G_{i:n})},$$

- Korzystamy z tego, że $F(X_{i:n}) =_{st} U_{i:n}$, $K^{-1}U_{i:n} =_{st} Z_{i:n}$ oraz unormowane odstępy $nZ_{1:n}, \ldots, (n-i+1)(Z_{i+1:n} - Z_{i:n}), \ldots,$ $Z_{n:n} - Z_{n-1:n}$ są niezależnymi zmiennymi losowymi o rozkładzie wykładniczym. Zatem wektor $\mathbf{S} = (S_1, S_2, \ldots S_n)'$ jest próbą z rozkładu wykładniczego ze średnią 1.

Szkic dowodu:

- Przy hipotezie $\mathcal{H}_0'$ mamy $\tilde{r}_F = c\tilde{r}_G$. Stąd

$$\frac{Y_{i+1:n} - Y_{i:n}}{X_{i+1:n} - X_{i:n}} = \frac{T_i}{S_i}\frac{\tilde{r}_F(\theta^F_{i:n})}{\tilde{r}_G(\theta^G_{i:n})} = c\frac{T_i}{S_i}\frac{\tilde{r}_F(\theta^F_{i:n})}{\tilde{r}_F(\theta^G_{i:n})},$$

- Korzystamy z tego, że $F(X_{i:n}) =_{st} U_{i:n}$, $K^{-1}U_{i:n} =_{st} Z_{i:n}$ oraz unormowane odstępy $nZ_{1:n}, \dots, (n-i+1)(Z_{i+1:n} - Z_{i:n}), \dots, Z_{n:n} - Z_{n-1:n}$ są niezależnymi zmiennymi losowymi o rozkładzie wykładniczym. Zatem wektor $\mathbf{S} = (S_1, S_2, \dots S_n)'$ jest próbą z rozkładu wykładniczego ze średnią 1.
- Wyznaczamy ograniczenia: $\bar{r}_{in} \geqslant \tilde{r}_F(\theta^F_{i:n})$, $\underline{r}_{in} \leqslant \tilde{r}_F(\theta^F_{i:n})$

Szkic dowodu:

- Pokazujemy, że

$$\lim_{n\to\infty} P\left(\max_{1\leqslant i\leqslant n-1} cW_i \leqslant nx\frac{\underline{r}_{in}}{\overline{r}_{in}}\right) = \lim_{n\to\infty}\prod_{i=1}^{n-1}\left(1-\frac{c/x}{n(c/nx+\underline{r}_{in}/\overline{r}_{in})}\right)$$

$$= \lim_{n\to\infty} P\left(\max_{1\leqslant i\leqslant n-1} cW_i \leqslant nx\frac{\overline{r}_{in}}{\underline{r}_{in}}\right) =$$

$$\lim_{n\to\infty}\prod_{i=1}^{n-1}\left(1-\frac{c/x}{n(c/nx+\overline{r}_{in}/\underline{r}_{in})}\right) = e^{-c/x},$$

a stąd

$$\lim_{n\to\infty} P\left(\frac{1}{n}\max_{1\leqslant i\leqslant n-1}\frac{Y_{i+1:n}-Y_{i:n}}{X_{i+1:n}-X_{i:n}} \leqslant x\right) =$$

$$\lim_{n\to\infty} P\left(\frac{1}{n}\max_{1\leqslant i\leqslant n-1} W_i\frac{r_F(\theta^F_{i:n})}{r_G(\theta^G_{i:n})} \leqslant x\right) = e^{-c/x}.$$

References

- J. Bartoszewicz, M. Benduch, (2009), *Some properties of the generalized TTT transform*, J. Statist. Plann. Inference, **139**, 2008–2017.
- E.L. Lehmann i J. Rojo, (1992), *Invariant directional orderings*, Ann. Statist. **20**, 2100-2110.
- X. Li,M. Shaked, (2007), *A general family of univariate stochastic orders*, J. Statist. Plann. Inference **137**, 3601-3610.