Skalární skórová funkce

Zdeněk Fabián
Ústav informatiky AVČR Praha

ROBUST
2016

Prototypy a transformovaná rozdělení na $\mathbb{R}$

- Y má na $\mathbb{R}$ rozdělení G (G_θ) s hustotou $g(y)$
  $\eta : \mathbb{R} \to \mathbb{R}$ hladká rostoucí, $\Rightarrow \eta^{-1}$ taky

  Transformovaná r.v. $X = \eta^{-1}(Y)$ má distribuční funkci

$$F(x) = G(\eta(x))$$

a hustotu

$$f(x) = g(\eta(x))\eta'(x)$$

Prototypy a transformovaná rozdělení na $\mathbb{R}$

- Y má na $\mathbb{R}$ rozdělení G (G_θ) s hustotou $g(y)$
  $\eta : \mathbb{R} \to \mathbb{R}$ hladká rostoucí, $\Rightarrow \eta^{-1}$ taky

  Transformovaná r.v. $X = \eta^{-1}(Y)$ má distribuční funkci

  $$F(x) = G(\eta(x))$$

  a hustotu $$f(x) = g(\eta(x))\eta'(x)$$

-

$$f(x) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{1+x^2}} e^{-\frac{1}{2}(\sinh^{-1} x)^2} \qquad g(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$$

Prototypy a transformovaná rozdělení na $\mathbb{R}$

- Y má na $\mathbb{R}$ rozdělení G (G_θ) s hustotou $g(y)$
  $\eta : \mathbb{R} \to \mathbb{R}$ hladká rostoucí, $\Rightarrow \eta^{-1}$ taky

  Transformovaná r.v. $X = \eta^{-1}(Y)$ má distribuční funkci

  $$F(x) = G(\eta(x))$$

  a hustotu $$f(x) = g(\eta(x))\eta'(x)$$

- $$f(x) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{1+x^2}}e^{-\frac{1}{2}(\sinh^{-1}x)^2} \qquad g(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}$$

- Nesymetrie: $g(y)$ je hustota *prototypu* transformovaného F

Skórová funkce prototypu

- $$S_G(y) = -\frac{g'(y)}{g(y)}$$

Skórová funkce prototypu

- $$S_G(y) = -\frac{g'(y)}{g(y)}$$

- popisuje vliv $y \in \mathbb{R}$ vzhledem ke 'středu' G
  pro $g(y-\mu)$ (*třída Ia*) platí
  $$\frac{\partial}{\partial \mu} \log g(y-\mu) = S_G(y-\mu)$$

Skórová funkce prototypu

- $$S_G(y) = -\frac{g'(y)}{g(y)}$$
- popisuje vliv $y \in \mathbb{R}$ vzhledem ke 'středu' G
  pro $g(y-\mu)$ (*třída Ia*) platí
  $$\frac{\partial}{\partial\mu}\log g(y-\mu) = S_G(y-\mu)$$
- popisuje vliv $y \in \mathbb{R}$ vzhledem ke 'středu' G i prototypu s θ bez parametru polohy, 'středem' je mód $y^* : S_G(y;\theta) = 0$
  $$g(y) = \frac{1}{B(p,q)}\frac{e^{py}}{(e^y+1)^{p+q}} \qquad S_G(y) = \frac{qe^y - p}{e^y + 1}$$

Tabulka: typy skórových funkcí

Typ	$S_G(y)$	$g(y)$	rozdělení	ES_G^2
NE	$\frac{e^y - e^{-y}}{2}$	$\frac{1}{2K_0(1)} e^{-\cosh y}$	-	1.43
NG	y	$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$	normal	1
ON	$e^y - 1$	$e^y e^{-e^y}$	Gumbel	1
NO	$1 - e^{-y}$	$e^{-y} e^{-e^{-y}}$	extreme value	1
OO	$\frac{e^y - 1}{e^y + 1}$	$\frac{e^y}{(1+e^y)^2}$	logistic	$1/3$
OR	$\frac{2y}{1+y^2}$	$\frac{1}{\pi} \frac{1}{1+y^2}$	Cauchy	$1/2$

Skórové momenty: ES_G^2 je (zobec.) Fisherova informace pro mód, ES_G^3 charakt. nesymetrie, ES_G^4 'placatosti'

Tabulka: typy skórových funkcí

Typ	$S_G(y)$	$g(y)$	rozdělení	ES_G^2
NE	$\frac{e^y - e^{-y}}{2}$	$\frac{1}{2K_0(1)} e^{-\cosh y}$	-	1.43
NG	y	$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$	normal	1
ON	$e^y - 1$	$e^y e^{-e^y}$	Gumbel	1
NO	$1 - e^{-y}$	$e^{-y} e^{-e^{-y}}$	extreme value	1
OO	$\frac{e^y - 1}{e^y + 1}$	$\frac{e^y}{(1+e^y)^2}$	logistic	$1/3$
OR	$\frac{2y}{1+y^2}$	$\frac{1}{\pi} \frac{1}{1+y^2}$	Cauchy	$1/2$

Skórové momenty: ES_G^2 je (zobec.) Fisherova informace pro mód, ES_G^3 charakt. nesymetrie, ES_G^4 'placatosti'

- Momentové rovnice

$$\frac{1}{n} \sum_{i=1}^{n} S_G^k(y_i; \theta) = ES_G^k(\theta), \qquad k = 1, ...$$

Transformovaná rozdělení

- Buď $\mathcal{X} \subseteq \mathbb{R}$ otevřený interval a $F(x) = G(\eta(x))$ na $\mathcal{X}$

Transformovaná rozdělení

- Buď $\mathcal{X} \subseteq \mathbb{R}$ otevřený interval a $F(x) = G(\eta(x))$ na $\mathcal{X}$
- I: Proč by pro rozdělení na polopřímce nebo intervalu neměla existovat obdoba skórové funkce prototypu, t.j. reálná (skalární) skórová funkce (vlivová funkce vzhledem k nějakému 'středu'). Protože $S(x) = -\frac{f'(x)}{f(x)}$ pro tato rozdělení nefunguje, musí být definovaná jiným vzorcem, který se jako v případě prototypů shoduje pro nějakou Ia třídu s jejich Fisherovým skórem pro nějaký parametr

Transformovaná rozdělení

- Buď $\mathcal{X} \subseteq \mathbb{R}$ otevřený interval a $F(x) = G(\eta(x))$ na $\mathcal{X}$
- I: Proč by pro rozdělení na polopřímce nebo intervalu neměla existovat obdoba skórové funkce prototypu, t.j. reálná (skalární) skórová funkce (vlivová funkce vzhledem k nějakému 'středu'). Protože $S(x) = -\frac{f'(x)}{f(x)}$ pro tato rozdělení nefunguje, musí být definovaná jiným vzorcem, který se jako v případě prototypů shoduje pro nějakou Ia třídu s jejich Fisherovým skórem pro nějaký parametr
- II. Transformované F na $\mathbb{R}$ poznáme a na $\mathcal{X} \neq \mathbb{R}$ jsou všechna transformovaná

Obdoba skórové funkce prototypu: t-skór

- $\eta(x) : \mathcal{X} \to \mathbb{R}$. Definujeme t-skór (transformation-based score) transformovaného rozdělení F jako

$$T_F(x) = S_G(\eta(x))$$

Obdoba skórové funkce prototypu: t-skór

- $\eta(x) : \mathcal{X} \to \mathbb{R}$. Definujeme t-skór (transformation-based score) transformovaného rozdělení F jako

$$T_F(x) = S_G(\eta(x))$$

- Ukáže se, že když $f(x) = g(\eta(x))\eta'(x)$,

$$T_F(x) = -\frac{1}{f(x)}\frac{d}{dx}[g(\eta(x))] = -\frac{1}{f(x)}\frac{d}{dx}[\frac{1}{\eta'(x)}f(x)]$$

Obdoba skórové funkce prototypu: t-skór

- $\eta(x) : \mathcal{X} \to \mathbb{R}$. Definujeme t-skór (transformation-based score) transformovaného rozdělení F jako

$$T_F(x) = S_G(\eta(x))$$

- Ukáže se, že když $f(x) = g(\eta(x))\eta'(x)$,

$$T_F(x) = -\frac{1}{f(x)}\frac{d}{dx}[g(\eta(x))] = -\frac{1}{f(x)}\frac{d}{dx}[\frac{1}{\eta'(x)}f(x)]$$

- a je dokázáno, že pro rozdělení s transformovaným parametrem polohy $\tau = \eta^{-1}(\mu)$ (rozdělení *třídy Ia*)

$$\frac{\partial}{\partial\tau}\log f(x;\tau) = konst. \times T_F(x;\tau)$$

t-skór rozdělení log-gamma

- $\mathcal{X} = (1, \infty)$ $\qquad f(x) = \frac{c^\alpha}{\Gamma(\alpha)} (\log x)^{\alpha-1} \frac{1}{x^{c+1}}$

t-skór rozdělení log-gamma

- $\mathcal{X} = (1, \infty)$ $\quad f(x) = \frac{c^\alpha}{\Gamma(\alpha)} (\log x)^{\alpha-1} \frac{1}{x^{c+1}}$
- Položme $\eta'(x) = \frac{1}{x \log x}$, $\eta(x) = \log \log x$,

$$T_{1F}(x) = -\frac{1}{f(x)} \frac{d}{dx} [x \log x \; f(x)] = c \log x - \alpha$$

a momentové rovnice

$$\sum_{i=1}^{n} (c \log x_i - \alpha) = 0 \qquad \frac{1}{n} \sum_{i=1}^{n} (c \log x_i - \alpha)^2 = \alpha$$

t-skór rozdělení log-gamma

- $\mathcal{X} = (1, \infty) \qquad f(x) = \frac{c^\alpha}{\Gamma(\alpha)}(\log x)^{\alpha-1}\frac{1}{x^{c+1}}$
- Položme $\eta'(x) = \frac{1}{x \log x}$, $\eta(x) = \log\log x$,

$$T_{1F}(x) = -\frac{1}{f(x)}\frac{d}{dx}[x \log x \, f(x)] = c \log x - \alpha$$

a momentové rovnice

$$\sum_{i=1}^{n}(c \log x_i - \alpha) = 0 \qquad \frac{1}{n}\sum_{i=1}^{n}(c \log x_i - \alpha)^2 = \alpha$$

- Zvolme $\eta(x) = \log(x-1)$,

$$T_{2F}(x) = -\frac{1}{f(x)}\frac{d}{dx}[(x-1)f(x)] = c - \frac{(\alpha-1)(x-1)}{\log x} - \frac{c+1}{x}$$

Pro $\alpha = 1$ se $f(x)$ i T_{2F} redukuje na hustotu a t-skór Pareta

Smysl má uvažovat dva druhy t-skórů

- t-skór přirozený: když je ve vzorci pro hustotu 'viditelné' $\eta(x)$ nebo $\eta'(x)$

Smysl má uvažovat dva druhy t-skórů

- t-skór přirozený: když je ve vzorci pro hustotu 'viditelné' $\eta(x)$ nebo $\eta'(x)$
- t-skór univerzální: založený na log-transformacích (Johnsonovy transformace)

$$\eta(x) = \begin{cases} \log(x-a) & \mathcal{X} = (a, \infty) \\ \log \frac{(x-a)}{(b-x)} & \mathcal{X} = (a, b) \end{cases}$$

Takto lze interpretovat většinu prakticky používaných rozdělení:

$$f(x) = e^{-x} = xe^{-x}\frac{1}{x}$$

$$T_F(x) = -\frac{1}{f(x)}\frac{d}{dx}[xf(x)] = e^x \frac{d}{dx}[xe^{-x}] = x - 1$$

Přirozený t-skór: pro momentové odhady

- Buď $\theta = (\theta_1, \theta_2)$. Protože $T_F(x) = S_G(\eta(x))$ a $ET_F^k = ES_G^k$, momentové rovnice

$$\sum_{i=1}^{n} S_G(\eta(x_i);\theta) = 0 \qquad \frac{1}{n}\sum_{i=1}^{n} S_G^2(\eta(x_i);\theta) = ES_G^2$$

jsou invariantní vůči transformacím proměnné (6 tříd rozdělení z Tabulky)

Přirozený t-skór: pro momentové odhady

- Buď $\theta = (\theta_1, \theta_2)$. Protože $T_F(x) = S_G(\eta(x))$ a $ET_F^k = ES_G^k$, momentové rovnice

$$\sum_{i=1}^{n} S_G(\eta(x_i);\theta) = 0 \qquad \frac{1}{n}\sum_{i=1}^{n} S_G^2(\eta(x_i);\theta) = ES_G^2$$

  jsou invariantní vůči transformacím proměnné (6 tříd rozdělení z Tabulky)
- Když S_F je omezená nebo 'huberovsky oříznutá'', estimátor je robustní pro všechny parametry

Přirozený t-skór: pro momentové odhady

- Buď $\theta = (\theta_1, \theta_2)$. Protože $T_F(x) = S_G(\eta(x))$ a $ET_F^k = ES_G^k$, momentové rovnice

$$\sum_{i=1}^{n} S_G(\eta(x_i); \theta) = 0 \qquad \frac{1}{n}\sum_{i=1}^{n} S_G^2(\eta(x_i); \theta) = ES_G^2$$

  jsou invariantní vůči transformacím proměnné (6 tříd rozdělení z Tabulky)
- Když S_F je omezená nebo 'huberovsky oříznutá'', estimátor je robustní pro všechny parametry
- Protože $T_F(x; \hat{\theta}) = S_G(\eta(x); \hat{\theta})$ je reálná funkce, lze ji např. použít v korelační a regresní analýze

Univerzální t-skór: Skórová funkce rozdělení SFD

- $$S_F(x) = \eta'(x^*)T_F(x)$$

  kde $\eta(x)$ je Johnsonova transformace a x^* je těžiště: řešení rovnice
  $$T_F(x) = 0$$
  (x^* je obraz módu prototypu).

Univerzální t-skór: Skórová funkce rozdělení SFD

- $$S_F(x) = \eta'(x^*)T_F(x)$$

  kde $\eta(x)$ je Johnsonova transformace a x^* je těžiště: řešení rovnice

  $$T_F(x) = 0$$

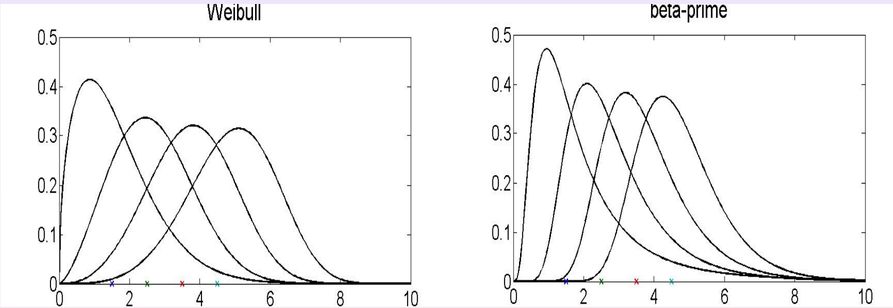
  (x^* je obraz módu prototypu).
- ES_F^2 zobecněná Fisherova informace (vzhledem k x^*)

  Variabilita rozdělení: skórová variance

  $$\omega^2 = \frac{1}{ES_F^2}$$

  Numerické charakteristiky rozdělení: $x^*(\theta), \omega(\theta) = \sqrt{\omega^2(\theta)}$

Rozdělení na $(0, \infty)$ s týmž ω

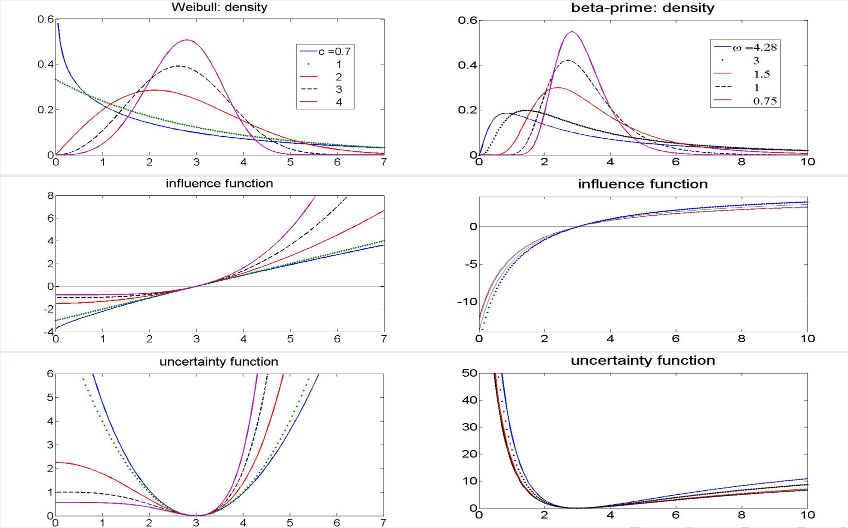


Kam jsme došli

K danému F na intervalu $\mathcal{X} \subseteq \mathbb{R}$ umíme jednoznečně přiřadit sfd $S_F(x)$, jejíž hodnota popisuje relativní vliv $x \in \mathcal{X}$ na střed x^* of F.

Popis náhodné veličiny X s rozdělením F na $\mathcal{X} \subseteq \mathbb{R}$

$F(x)$	cdf
$f(x)$	pdf
$S(x)$	sfd
$x^* : S(x) = 0$	těžiště
$S^2(x)$	informační funkce
$w(x) = \frac{dS(x)}{dx}$	vahová funkce
ES^2	střední (Fisherova) informace
$\omega^2 = 1/ES^2$	skórová variance
$\frac{S(x)}{ES^2} = \omega^2 S(x)$	influence function
$\frac{S^2(x)}{[ES^2]^2}$	uncertainty function



Skalární skórová funkce

Metrika ve výběrovém prostoru $\mathcal{X}$

$$
\begin{aligned}
&S_F(x) && \text{sfd} \\
&w_F(x) = \frac{dS_F(x)}{dx} && \text{vahová funkce} \\
&\omega_F = 1/\sqrt{ES_F^2}
\end{aligned}
$$

Vzdálenost $x_1, x_2 \in \mathcal{X}$

$$\rho_F(x_1, x_2) = \omega_F \int_{x_1}^{x_2} w_F(x)\, dx = \omega_F |S_F(x_2) - S_F(x_1)|$$

$$S_F(x) \Leftrightarrow f(x)$$

$$(\mathcal{X}, F) \Leftrightarrow (\mathcal{X}, \rho_F)$$

Centrální limitní věta pro sfd

$S_F(x)$	sfd
$x^* : S_F(x) = 0$	těžiště
$w_F(x) = \frac{dS_F(x)}{dx}$	vahová funkce
ES_F^2	střední (Fisherova) informace
$\omega^2 = 1/ES_F^2$	skórová variance

$X_1, ..., X_n$ iid podle F

Central limit theorem: $\bar{S} = \frac{1}{n}\sum_{i=1}^n S_F(X_i) \to \mathcal{N}(0; \omega^2/n)$

Výběrové těžiště:

$$\hat{x}^* = S_F^{-1}(\bar{S}) \to \mathcal{N}(x^*; [(S_F^{-1}(\bar{S}))']^2\omega^2/n)$$

Odhady pro rozdílné modely jdou porovnávat pomocí

$$\hat{x}^* = x^*(\hat{\theta}), \hat{\omega}^* = \omega(\hat{\theta})$$

Skórový průměr

Někdy je výběrové těžiště známou statistikou:

rozdělení	$\hat{x}^*$
normální, gamma, beta	$\bar{x}$
lognormální	$\bar{x}_{Geom}$
Weibull (c const.)	$\frac{1}{n}(\sum x_i^c)^{1/c}$
Pareto and jiná heavy-tailed	$\bar{x}_{Harm}$
beta-prime	$\sum \frac{x_i}{x_i+1} / \sum \frac{1}{x_i+1}$

Reference:

Fabián, Z. (2001). Induced cores and their use in robust parametric estimation. Communication in Statistics, Theory-Methods 30, 537-556.

Fabián, Z. (2010). Score moment estimators. In Lechavilier, Y.,Saporta, G.(Eds.), Proc. of conf. COMPSTAT, Physica-Verlag, pp.975-982.

Fabián Z. (2016). Score function of distribution and revival of the moment method. Communication in Statistics, Theory-Methods 45,4, 1118-1136.

Děkuji za pozornost

SFD rozdělení na intervalu

Př. Useknutá exponenciela $\mathcal{X} = (0, 1)$, $\mathbf{f}(\mathbf{x}) = \mathbf{b}\mathbf{e}^{-\lambda \mathbf{x}}$

Johnsonova transformace $\eta(x) = \log \frac{x}{1-x}, \eta'(x) = \frac{1}{x(1-x)}$ Pišme

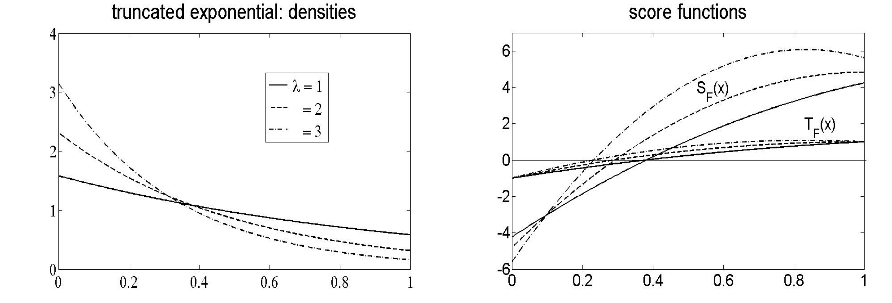
$$f(x) = bx(1-x)e^{-\lambda x}\frac{1}{x(1-x)},$$

takže

$$T_F(x) = -e^{\lambda x}\frac{d}{dx}[x(1-x)e^{-\lambda x}] = (2+\lambda)x - \lambda x^2 - 1$$

Pak $x^* : \quad T_F(x^*) = 0,$

$$S_F(x) = \frac{1}{x^*(1-x^*)}T_F(x)$$


truncated exponential: densities
λ = 1
= 2
= 3
score functions
S_F(x)
T_F(x)

Vztah ω_F a diferenciální entropie

$$h_F = -\int f(x) \log f(x)\, dx$$

F	e^{h_F}	ω_F
normal	$\sqrt{2\pi e}\sigma$	σ
Cauchy	$4\pi\sigma$	$\sqrt{2}\sigma$
exponential	$e\tau$	τ
lognormal	$\sqrt{2\pi e}\,\tau/c$	τ/c
Weibull	$e^{(1+\epsilon(1-1/c))}\,\tau/c$	τ/c
uniform$_{(a,b)}$	$b-a$	$\frac{\sqrt{3}}{4}(b-a)^2$