数理解析研究所講究録
第1999巻 2016年 13-16

On the information on skew-normal distributions

筑波大学　赤平昌文 (Masafumi Akahira)
(University of Tsukuba)

最近，歪正規分布，歪 t 分布等の研究が盛んに行われ，分布の性質，推定量の頑健性の研究等も行われている (Azzalini [A14]). 本稿では，RIMS 共同による研究会における田崎・小池 ([TK16]) の講演を機に情報量の観点から歪正規分布の歪みの度合を表す形状母数に関する情報について考える.

まず，確率変数 Z が p.d.f.

$$f_\alpha(x) := 2\phi(x)\Phi(\alpha x), \quad x \in \mathbb{R},\ \alpha \in \mathbb{R} \tag{1}$$

をもつとき Z は歪正規 (skew-normal) 分布 $SN(0,1,\alpha)$ に従うという．このとき，形状母数 α に関する Fisher 情報量は

$$I(\alpha) := E\left[\left\{\frac{\partial \log f_\alpha(X)}{\partial \alpha}\right\}^2\right] = 2\int_{-\infty}^{\infty} \frac{x^2\phi^2(\alpha x)}{\Phi(\alpha x)}\phi(x)dx \tag{2}$$

となる．ただし，$\phi(\cdot), \Phi(\cdot)$ はそれぞれ標準正規分布の p.d.f., c.d.f. とする．ここで，(2) より $I(\alpha) \to I(0) = 2/\pi$ $(\alpha \to 0)$ になり，また $I(\alpha) \to 0$ $(\alpha \to \infty)$ になる．このことは，歪みの度合を表す形状母数 α が大きくなると Z の歪度は単調増加する一方で，α に関する情報量 $I(\alpha)$ が小さくなることを意味し，やや意味不明になる.

そこで，たとえば正規分布 $N(\mu, \sigma^2)$ の μ に関する Fisher 情報量 $I(\mu) = 1/\sigma^2$ になり，これは位置母数 μ には依存しないが尺度母数 σ に依存し，これは分散 σ^2 を通して μ に関する情報を与えているとも考えられる．何故なら σ が小さくなれば $N(\mu, \sigma^2)$ は μ の周りにより集中することを意味するからである．いま，$Y := Z + \theta$ とすれば，Y は歪正規分布 $SN(\theta, 1, \alpha)$ $(\theta \in \mathbb{R},\ \alpha \in \mathbb{R})$ の p.d.f.

$$f_\alpha(x-\theta) = 2\phi(x-\theta)\Phi(\alpha(x-\theta)), \quad x \in \mathbb{R} \tag{3}$$

をもち，位置母数 θ に関する Fisher 情報量は

$$I(f_\alpha) := \int_{-\infty}^{\infty} \left\{\frac{d}{dx}\log f_\alpha(x)\right\}^2 f_\alpha(x)dx$$

になり，θ には無関係であるが，形状母数 α には依存する．この $I(f_\alpha)$ は α を通して θ に関する情報を与えていると考えられる.

このとき，次の定理が成り立つ.

定理　歪正規分布 $SN(\theta, 1, \alpha)$ $(\alpha > 0)$ の θ に関する Fisher 情報量 $I(f_\alpha)$ について

(i) $I(f_\alpha)$ は α に関して単調増加である.

(ii) $\lim_{\alpha\to 0} I(f_\alpha) = 1$, $\lim_{\alpha\to\infty} I(f_\alpha) = \infty$.

証明　まず

$$
\begin{aligned}
I(f_\alpha) &= 2\int_{-\infty}^{\infty}\left\{\frac{\alpha\phi(\alpha x)}{\Phi(\alpha x)} - x\right\}^2 \phi(x)\Phi(\alpha x)dx \\
&= 2\alpha^2\int_{-\infty}^{\infty}\frac{\phi^2(\alpha x)}{\Phi(\alpha x)}\phi(x)dx - 4\alpha\int_{-\infty}^{\infty}x\phi(\alpha x)\phi(x)dx + 2\int_{-\infty}^{\infty}x^2\phi(x)\Phi(\alpha x)dx \\
&= 2\alpha\int_{-\infty}^{\infty}\frac{\phi^2(t)}{\Phi(t)}\phi\left(\frac{t}{\alpha}\right)dt + 2\int_{-\infty}^{\infty}x^2\phi(x)\Phi(\alpha x)dx \qquad (4)
\end{aligned}
$$

となる．ここで

$$
\int_{-\infty}^{\infty}|u|^3\phi(x)dx < \infty
$$

であるから

$$
\frac{d}{d\alpha}\int_{-\infty}^{\infty}x^2\phi(x)\Phi(\alpha x)dx = \int_{-\infty}^{\infty}x^3\phi(\alpha x)\phi(x)dx = 0 \qquad (5)
$$

になる．また，(4) の右辺の第 1 項について，$\phi(t/\alpha)$ は指数型で $\phi^2(t)/\Phi(t)$ は有界であるから

$$
\frac{d}{d\alpha}\int_{-\infty}^{\infty}\frac{\phi^2(t)}{\Phi(t)}\phi\left(\frac{t}{\alpha}\right)dt = \frac{1}{\alpha^3}\int_{-\infty}^{\infty}\frac{t^2\phi^2(t)\phi(t/\alpha)}{\Phi(t)}dt \qquad (6)
$$

になり，さらに (4) を α について微分して (5) と (6) を代入すれば

$$
\begin{aligned}
\frac{d}{d\alpha}I(f_\alpha) &= 2\int_{-\infty}^{\infty}\frac{\phi^2(t)}{\Phi(t)}\phi\left(\frac{t}{\alpha}\right)dt + 2\alpha\frac{d}{d\alpha}\int_{-\infty}^{\infty}\frac{\phi^2(t)}{\Phi(t)}\phi\left(\frac{t}{\alpha}\right)dt \\
&= 2\int_{-\infty}^{\infty}\frac{\phi^2(t)}{\Phi(t)}\phi\left(\frac{t}{\alpha}\right)dt + \frac{2}{\alpha^2}\int_{-\infty}^{\infty}\frac{t^2\phi^2(t)}{\Phi(t)}\phi\left(\frac{t}{\alpha}\right)dt \\
&> 0
\end{aligned}
$$

となるから $I(f_\alpha)$ は α に関して単調増加であるので (i) が成り立つ．

次に，$\phi(t/\alpha)$ は α の単調増加関数なので

$$
\lim_{\alpha\to 0}\int_{-\infty}^{\infty}\frac{\phi^2(t)}{\Phi(t)}\phi\left(\frac{t}{\alpha}\right)dt = 0, \qquad (7)
$$

$$
\lim_{\alpha\to\infty}\int_{-\infty}^{\infty}\frac{\phi^2(t)}{\Phi(t)}\phi\left(\frac{t}{\alpha}\right)dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{\phi^2(t)}{\Phi(t)}dt > 0 \qquad (8)
$$

になり，(8) は有限確定にもなる．また，$0 < \Phi(\alpha x) < 1$ より

$$
\lim_{\alpha\to 0}\int_{-\infty}^{\infty}x^2\phi(x)\Phi(\alpha x)dx = \int_{-\infty}^{\infty}\frac{1}{2}x^2\phi(x)dx = \frac{1}{2}, \qquad (9)
$$

$$
\begin{aligned}
\lim_{\alpha\to\infty}\int_{-\infty}^{\infty}x^2\phi(x)\Phi(\alpha x)dx &= \lim_{\alpha\to\infty}\left\{\int_{-\infty}^{0}x^2\phi(x)\Phi(\alpha x)dx + \int_{0}^{\infty}x^2\phi(x)\Phi(\alpha x)dx\right\} \\
&= \int_{0}^{\infty}x^2\phi(x)dx = \frac{1}{2} \qquad (10)
\end{aligned}
$$

になる．よって (4), (7), (9) より

$$\lim_{\alpha\to 0} I(f_\alpha) = 1$$

になり，(4), (8), (10) より

$$\lim_{\alpha\to \infty} I(f_\alpha) = \infty$$

になる．ゆえに (ii) が示された． □

上記の定理では $\alpha > 0$ の場合を考えたが，$\alpha < 0$ の場合でも同様にできる．また，既に，$SN(\theta, 1, \alpha)$ は α が大きくなるにつれて，より歪んだ分布になることは知られている．実際，確率変数 Z, Y がそれぞれ $SN(0, 1, \alpha)$, $SN(\theta, 1, \alpha)$ に従うので，

$$\mu_Z := E(Z) = \sqrt{\frac{2}{\pi}}\delta, \quad \sigma_Z^2 := V(Z) = 1 - \mu_Z^2 = 1 - \frac{2}{\pi}\delta^2, \tag{11}$$

$$\mu_Y := E(Y) = \theta + \mu_Z, \quad \sigma_Y^2 := V(Y) = \sigma_Z^2, \tag{12}$$

になる．ただし，$\delta = \alpha/\sqrt{1+\alpha^2}$ とする．このとき Y の歪度，尖度はそれぞれ

$$\gamma_1(Y) = \frac{4-\pi}{2}\left(\frac{\mu_Z}{\sigma_Z}\right)^3, \quad \gamma_2(Y) = 2(\pi - 3)\left(\frac{\mu_Z}{\sigma_Z}\right)^4$$

になり，これらは Z の歪度，尖度に等しくいずれも μ_Z/σ_Z の関数になっていて，μ_Z/σ_Z は δ について単調増加し，$|\delta| < 1$ に対して $|\mu_Z/\sigma_Z| < \sqrt{2/(\pi-2)}$ になる ([A14])．また，$\gamma_1 = \gamma_1(Z)$ と $\gamma_2 = \gamma_2(Z)$ の関係は

$$\gamma_2 = 2(\pi - 3)\left(\frac{2}{4-\pi}\right)^{4/3} \gamma_1^{4/3} \quad (|\gamma_1| \le \sqrt{2}(4-\pi)(\pi-2)^{-3/2})$$

になり，歪度 γ_1 は α について単調増加，尖度 γ_2 は $|\alpha|$ について単調増加になる．さらに $|\gamma_1| = \sqrt{2}(4-\pi)(\pi-2)^{-3/2}$ のとき，γ_2 は最大値 $8(\pi-3)(\pi-2)^{-2}$ をとる．

ここで，(11), (12) より $\alpha > 0$ のとき，σ_Y^2 は α について単調減少で $1-(2/\pi) \le \sigma_Y^2 < 1$ になり，$\alpha \to \infty$ のとき σ_Y^2 はその最小値 $1-(2/\pi)$ に収束し，分布 $SN(\theta, 1, \alpha)$ は $\mu_Y = \theta + \mu_Z$ の最大値 $\theta + \sqrt{2/\pi}$ の周りに寄ってくるとともに歪度 γ_1 は最大値に近いことが分かる．実際には $\alpha = 20$ 位でも比較的良い近似を与えるように見える．

さて，$\alpha \to \infty$ のとき (3) で与えられる $SN(\theta, 1, \alpha)$ の p.d.f. は半正規分布 (half-normal distribution) の p.d.f. $f_\infty(x-\theta) := 2\phi(x-\theta)\chi_{(0,\infty)}(x-\theta)$ に収束する．ただし，$\chi_{(0,\infty)}(\cdot)$ は区間 $(0, \infty)$ の定義関数とする．このとき，θ の Pitman 推定量 (最良位置共変推定量) が求められる．実際，$Y_1, \ldots, Y_n$ をたがいに独立にいずれも p.d.f. $f_\infty(y-\theta)$ に従うとすれば，θ が下側切断母数にもなることに注意すると θ の Pitman 推定量は

$$\hat{\theta}_{PT} = \bar{Y} - \frac{\phi(\sqrt{n}(Y_{(1)} - \bar{Y}))}{\sqrt{n}\Phi(\sqrt{n}(Y_{(1)} - \bar{Y}))}$$

になる．ただし，$Y_{(1)} = \min_{1\le i\le n} Y_i$, $\bar{Y} = (1/n)\sum_{i=1}^{n} Y_i$ とする．また，$\hat{\theta}_{PT}$ は位置共変推定量なので θ の不偏推定量でもある．さらに $(Y_{(1)}, \bar{Y})$ は θ に対する十分統計量になるが,

$$\hat{\theta} := Y_{(1)} - 2^n \int_0^\infty (1 - \Phi(t))^n dt$$

もまた $(Y_{(1)}, \bar{Y})$ の関数で $\hat{\theta}_{PT}$ とは異なる θ の不偏推定量になることから完備性を持たない．ともかく，$\alpha \to \infty$ とすれば θ に関する情報が増えて，θ の Pitman 推定量が具体的に求められるという意味で θ の最良な推定を行うことができるので，上記の定理における $\lim_{\alpha\to\infty} I(f_\alpha) = \infty$ という結果は妥当のように思われる．実は歪正規分布 $SN(0, 1, \alpha)$ は $\alpha > 0$ のとき正則な分布であるが，$\alpha \to \infty$ のとき半正規分布に収束し，これは非正則分布になる．このような非正則分布族の位置母数の推定は [AOT07] において論じられている．

最後に，本稿について橋本真太郎氏 (広島大) と，小池健一氏 (筑波大) に貴重な御意見を頂いたことを感謝申し上げます．

参考文献

[AOT07] Akahira, M., Ohyauchi, N. and Takeuchi, K. (2007). On the Pitman estimator for a family of non-regular distributions. *Metron* **65**, 113–127.

[A14] Azzalini, A. (2014). *The Skew-Normal and Related Families.* Cambridge Univ. Press.

[TK16] 田崎雅裕，小池健一 (2016). Skew-q-gaussian distribution. 数理解析研究所講究録 (掲載予定).