有界な台をもつ尺度モデルでの Rényi divergence の漸近展開

日本大学・商学部　小池健一
Ken-ichi Koike
College of Commerce,
Nihon University

1　はじめに

$p = p(x)$ と $q = q(x)$ をルベーグ測度に関する pdf とする．このとき p から q への次数 s の Rényi divergence は

$$D_s(p||q) = \frac{1}{s-1}\log\int p(x)^s q(x)^{1-s}\mathrm{d}\mu(x)$$

で定義される ([4]).

次数を変えることで，Rényi divergence から様々な divergence や距離が導出される ([5])．例えば，Hellinger 距離の 2 乗

$$H^2(p||q) = \frac{1}{2}\int\left(\sqrt{p(x)} - \sqrt{q(x)}\right)^2\mathrm{d}\mu(x)$$

に対して，

$$D_{1/2}(p||q) = -2\log\left(1 - H^2(p||q)\right)$$

となる．また，正則条件の下で，$s \to 1$ とすれば，

$$\lim_{s\to 1} D_s(p||q) = D_{\mathrm{KL}}(p||q)$$

となる．ただし，$D_{\mathrm{KL}}(p||q)$ は Kullback-Leibler divergence

$$D_{\mathrm{KL}}(p||q) = \int p(x)\log\frac{p(x)}{q(x)}\mathrm{d}\mu(x)$$

とする．

また，Rényi divergence は統計的推測において重要な役割を果たすことが知られている (例えば，非正則な位置母数分布族に対する情報量損失については Akahira (1996), 非正則な位置母数分布族に対する大偏差型の漸近理論については Hayashi (2006, 2010) 等)．さらに，正則な確率分布族においては，Rényi divergence の摂動が Fisher 情報量を与えることが Akahira (1996) により示されている．

本稿では，台が有界な尺度母数分布族の下で，Rényi divergence の漸近展開を考える．

2 仮定

pdf $f(x \mid \theta) = \dfrac{1}{\theta} g\left(\dfrac{x}{\theta}\right)$ をもつ確率分布族を考える（$\theta > 0$ は尺度母数）．ただし，本稿においては pdf $g(x)$ に次の条件を課す．

条件

$$g(x) \begin{cases} > 0 & (a < x < b), \\ = 0 & (\text{それ以外}) \end{cases}$$

かつ，$a < x < b$ において $g(x)$ は十分滑らかな関数で

$$g(a) = \lim_{x \to a+0} g(x) > 0,\ g(b) = \lim_{x \to b-0} g(x) > 0$$

である．

また，簡単のため

$$g'(a) = \lim_{x \to a+0} g'(x),\ g'(b) = \lim_{x \to b-0} g'(x),$$
$$g''(a) = \lim_{x \to a+0} g''(x),\ g''(b) = \lim_{x \to b-0} g''(x)$$

のように表記をする．

3 Fisher 情報量の違い

pdf $f(x \mid \theta) = \dfrac{1}{\theta} g\left(\dfrac{x}{\theta}\right)$ をもつ確率分布に従うとする（$\theta > 0$ は尺度母数）．いま，

$$\begin{aligned}
\frac{\partial}{\partial \theta} \log f(x \mid \theta) &= \frac{\frac{\partial}{\partial \theta} f(x \mid \theta)}{f(x \mid \theta)} = \frac{\partial}{\partial \theta}\left\{\frac{1}{\theta} g\left(\frac{x}{\theta}\right)\right\} \Big/ \left\{\frac{1}{\theta} g\left(\frac{x}{\theta}\right)\right\} = \left\{-\frac{1}{\theta} - \frac{x g'\left(\frac{x}{\theta}\right)}{\theta^2 g\left(\frac{x}{\theta}\right)}\right\}, \\
\frac{\partial^2}{\partial \theta^2} \log f(x \mid \theta) &= \frac{\frac{\partial^2}{\partial \theta^2} f(x \mid \theta)}{f(x \mid \theta)} - \left\{\frac{\frac{\partial}{\partial \theta} f(x \mid \theta)}{f(x \mid \theta)}\right\}^2 \\
&= \frac{2}{\theta^2} + \frac{4 x g'\left(\frac{x}{\theta}\right)}{\theta^3 g\left(\frac{x}{\theta}\right)} + \frac{x^2 g''\left(\frac{x}{\theta}\right)}{\theta^4 g\left(\frac{x}{\theta}\right)} - \left\{-\frac{1}{\theta} - \frac{x g'\left(\frac{x}{\theta}\right)}{\theta^2 g\left(\frac{x}{\theta}\right)}\right\}^2 \\
&= \frac{1}{\theta^2} + \frac{-x^2 g'\left(\frac{x}{\theta}\right)^2}{\theta^4 g^2\left(\frac{x}{\theta}\right)} + \frac{2 x g'\left(\frac{x}{\theta}\right)}{\theta^3 g\left(\frac{x}{\theta}\right)} + \frac{x^2 g''\left(\frac{x}{\theta}\right)}{\theta^4 g\left(\frac{x}{\theta}\right)}
\end{aligned}$$

である．Fisher 情報量を

$$I_0(\theta) = \int_{S(\theta)} \left(\frac{\partial}{\partial \theta} \log f(x \mid \theta)\right)^2 f(x \mid \theta) \mathrm{d}x,$$
$$I(\theta) = -\int_{S(\theta)} \frac{\partial^2 \log f(x \mid \theta)}{\partial \theta^2} f(x \mid \theta) \mathrm{d}x$$

とおく[1)]．ただし，$S(\theta)$ は $f(x \mid \theta)$ の台で $S(\theta) = (a\theta, b\theta)$ である．

3.1　$I(\theta)$ と $I_0(\theta)$ の差

まず $I(\theta)$ と $I_0(\theta)$ の差を求める．$y = x/\theta$ とおくと

$$\begin{aligned} I(\theta) =& -\int_{a\theta}^{b\theta} \frac{\partial^2 \log f(x \mid \theta)}{\partial\theta^2} f(x \mid \theta) \mathrm{d}x \\ =& \int_{a\theta}^{b\theta} \left[-\frac{2}{\theta^2} - \frac{4xg'\left(\frac{x}{\theta}\right)}{\theta^3 g\left(\frac{x}{\theta}\right)} - \frac{x^2 g''\left(\frac{x}{\theta}\right)}{\theta^4 g\left(\frac{x}{\theta}\right)} + \left\{ -\frac{1}{\theta} - \frac{xg'\left(\frac{x}{\theta}\right)}{\theta^2 g\left(\frac{x}{\theta}\right)} \right\}^2 \right] \frac{1}{\theta} g\left(\frac{x}{\theta}\right) \mathrm{d}x \\ =& \frac{1}{\theta^2} \int_a^b \{-2g(y) - 4yg'(y) - y^2 g''(y)\} \mathrm{d}y + I_0(\theta) \end{aligned} \tag{3.1}$$

となる．ここで，部分積分より

$$\begin{aligned} & \int_a^b \{-2g(y) - 4yg'(y) - y^2 g''(y)\} \mathrm{d}y \\ =& -2\int_a^b g(y)\mathrm{d}y - 4\left\{ [yg(y)]_a^b - \int_a^b g(y)\mathrm{d}y \right\} - \left\{ [y^2 g'(y)]_a^b - 2[yg(y)]_a^b + 2\int_a^b g(y)\mathrm{d}y \right\} \\ =& -2\{bg(b) - ag(a)\} - \{b^2 g'(b) - a^2 g'(a)\} \end{aligned}$$

である．よって，(3.1) より

$$I(\theta) = \frac{-2\{bg(b) - ag(a)\} - \{b^2 g'(b) - a^2 g'(a)\}}{\theta^2} + I_0(\theta)$$

を得る．

3.2　$I_0(\theta)$ と $I_0'(\theta)$ の差

次に，

$$I_0'(\theta) = \int_S \left(\frac{\partial}{\partial\theta} \log f(x \mid \theta) \right)^2 f(x \mid \theta) \mathrm{d}x$$

とおいて，$I_0'(\theta)$ と $I_0(\theta)$ の関係式を求める．ただし，$S = S(\theta) \cap S(\theta + \varepsilon)$ とする．

$$\begin{aligned} I_0(\theta) =& \int_{a\theta}^{b\theta} \left(-\frac{1}{\theta} - \frac{x}{\theta^2} \frac{g'\left(\frac{x}{\theta}\right)}{g\left(\frac{x}{\theta}\right)} \right)^2 \frac{1}{\theta} g\left(\frac{x}{\theta}\right) \mathrm{d}x \\ =& \int_a^b \left(-\frac{1}{\theta} - \frac{y}{\theta} \frac{g'(y)}{g(y)} \right)^2 g(y)\,\mathrm{d}y \quad (y = x/\theta) \\ =& \frac{1}{\theta^2} \int_a^b \left(1 + \frac{yg'(y)}{g(y)} \right)^2 g(y)\,\mathrm{d}y = \frac{1}{\theta^2} I_0(1) \end{aligned}$$

となる．a, b や ε の正負により S は異なる区間になるので，場合分けをして考える．まず “$\varepsilon > 0$ かつ $0 < a < b$” または “$\varepsilon > 0$ かつ $a < b < 0$” のときを考える．このとき，$S = (a(\theta + \varepsilon), b\theta)$ となる

1）ただし，本稿のモデルでは pdf $f(\cdot \mid \theta)$ の台が母数 θ に依存するため，本来は Fisher 情報量は存在せず，正則な場合のときの Fisher 情報量の定義を当てはめたものである．

ので

$$
\begin{aligned}
I_0'(\theta) =& \int_S \left(\frac{\partial}{\partial\theta}\log f(x\mid\theta)\right)^2 f(x\mid\theta)\mathrm{d}x = \int_{a(\theta+\varepsilon)}^{b\theta}\left(-\frac{1}{\theta}-\frac{x}{\theta^2}\frac{g'\left(\frac{x}{\theta}\right)}{g\left(\frac{x}{\theta}\right)}\right)^2\frac{1}{\theta}g\left(\frac{x}{\theta}\right)\mathrm{d}x\\
=& \int_{a(1+(\varepsilon/\theta))}^{b}\left(-\frac{1}{\theta}-\frac{y}{\theta}\frac{g'(y)}{g(y)}\right)^2 g(y)\,\mathrm{d}y \quad (y=x/\theta)\\
=& \frac{1}{\theta^2}\int_a^b\left(1+\frac{yg'(y)}{g(y)}\right)^2 g(y)\,\mathrm{d}y-\frac{1}{\theta^2}\int_a^{a(1+(\varepsilon/\theta))}\left(1+\frac{yg'(y)}{g(y)}\right)^2 g(y)\,\mathrm{d}y\\
=& I_0(\theta)-\frac{1}{\theta^2}\left\{\frac{a\varepsilon\left(a^2g'(a)^2+2ag(a)g'(a)+g(a)^2\right)}{\theta g(a)}\right.\\
&+\frac{a^2\varepsilon^2\left(-a^2g'(a)^3+2a^2g(a)g'(a)g''(a)+2ag(a)^2g''(a)+2ag(a)g'(a)^2+3g(a)^2g'(a)\right)}{2\theta^2g(a)^2}\\
&\left.+O\left(\varepsilon^3\right)\right\}
\end{aligned}
$$

となる．次に “$\varepsilon>0$ かつ $a<0<b$” のときを考える．このとき，$S=(a\theta,b\theta)=S(\theta)$ となるので

$$
I_0'(\theta)=\int_S\left(\frac{\partial}{\partial\theta}\log f(x\mid\theta)\right)^2 f(x\mid\theta)\mathrm{d}x = I_0(\theta)
$$

となる．次に “$\varepsilon>0$ かつ $a<b<0$” または “$\varepsilon<0$ かつ $0<a<b$” のときを考える．このとき，$S=(a\theta,b(\theta+\varepsilon))$ となるので

$$
\begin{aligned}
I_0'(\theta) =& \int_S \left(\frac{\partial}{\partial\theta}\log f(x\mid\theta)\right)^2 f(x\mid\theta)\mathrm{d}x = \int_{a\theta}^{b(\theta+\varepsilon)}\left(-\frac{1}{\theta}-\frac{x}{\theta^2}\frac{g'\left(\frac{x}{\theta}\right)}{g\left(\frac{x}{\theta}\right)}\right)^2\frac{1}{\theta}g\left(\frac{x}{\theta}\right)\mathrm{d}x\\
=& \int_{a}^{b(1+(\varepsilon/\theta))}\left(-\frac{1}{\theta}-\frac{y}{\theta}\frac{g'(y)}{g(y)}\right)^2 g(y)\,\mathrm{d}y \quad (y=x/\theta)\\
=& \frac{1}{\theta^2}\int_a^b\left(1+\frac{yg'(y)}{g(y)}\right)^2 g(y)\,\mathrm{d}y-\frac{1}{\theta^2}\int_{b(1+(\varepsilon/\theta))}^{b}\left(1+\frac{yg'(y)}{g(y)}\right)^2 g(y)\,\mathrm{d}y\\
=& I_0(\theta)-\frac{1}{\theta^2}\left\{-\frac{b\varepsilon\left(b^2g'(b)^2+2bg(b)g'(b)+g(b)^2\right)}{\theta g(b)}\right.\\
&-\frac{\varepsilon^2\left(b^2\left(-b^2g'(b)^3+2b^2g(b)g'(b)g''(b)+2bg(b)^2g''(b)+2bg(b)g'(b)^2+3g(b)^2g'(b)\right)\right)}{2\theta^2g(b)^2}\\
&\left.+O\left(\varepsilon^3\right)\right\}\\
=& I_0(\theta)+\frac{1}{\theta^2}\left\{\frac{b\varepsilon\left(b^2g'(b)^2+2bg(b)g'(b)+g(b)^2\right)}{\theta g(b)}\right.\\
&+\frac{\varepsilon^2\left(b^2\left(-b^2g'(b)^3+2b^2g(b)g'(b)g''(b)+2bg(b)^2g''(b)+2bg(b)g'(b)^2+3g(b)^2g'(b)\right)\right)}{2\theta^2g(b)^2}\\
&\left.+O\left(\varepsilon^3\right)\right\}
\end{aligned}
$$

となる．

次に “$\varepsilon < 0$ かつ $a < 0 < b$” のときを考える．このとき，$S = (a(\theta+\varepsilon), b(\theta+\varepsilon))$ となるので

$$
\begin{aligned}
&I_0'(\theta)\\
=&\int_S \left(\frac{\partial}{\partial\theta}\log f(x \mid \theta)\right)^2 f(x \mid \theta)\mathrm{d}x = \int_{a(\theta+\varepsilon)}^{b(\theta+\varepsilon)} \left(-\frac{1}{\theta} - \frac{x}{\theta^2}\frac{g'\left(\frac{x}{\theta}\right)}{g\left(\frac{x}{\theta}\right)}\right)^2 \frac{1}{\theta}g\left(\frac{x}{\theta}\right)\mathrm{d}x\\
=&\int_{a(1+(\varepsilon/\theta))}^{b(1+(\varepsilon/\theta))} \left(-\frac{1}{\theta} - \frac{y}{\theta}\frac{g'(y)}{g(y)}\right)^2 g(y)\,\mathrm{d}y \quad (y = x/\theta)\\
=&\frac{1}{\theta^2}\int_a^b \left(1 + \frac{yg'(y)}{g(y)}\right)^2 g(y)\,\mathrm{d}y - \frac{1}{\theta^2}\int_{b(1+(\varepsilon/\theta))}^{b} \left(1 + \frac{yg'(y)}{g(y)}\right)^2 g(y)\,\mathrm{d}y\\
&- \frac{1}{\theta^2}\int_a^{a(1+(\varepsilon/\theta))} \left(1 + \frac{yg'(y)}{g(y)}\right)^2 g(y)\,\mathrm{d}y\\
=&I_0(\theta) - \frac{1}{\theta^2}\Bigg\{\varepsilon\left(\frac{a\left(a^2g'(a)^2 + 2ag(a)g'(a) + g(a)^2\right)}{\theta g(a)} - \frac{b\left(b^2g'(b)^2 + 2bg(b)g'(b) + g(b)^2\right)}{\theta g(b)}\right)\\
&+ \varepsilon^2\Bigg(\frac{a^2\left(-a^2g'(a)^3 + 2a^2g(a)g'(a)g''(a) + 2ag(a)^2g''(a) + 2ag(a)g'(a)^2 + 3g(a)^2g'(a)\right)}{2\theta^2g(a)^2}\\
&- \frac{b^2\left(-b^2g'(b)^3 + 2b^2g(b)g'(b)g''(b) + 2bg(b)^2g''(b) + 2bg(b)g'(b)^2 + 3g(b)^2g'(b)\right)}{2\theta^2g(b)^2}\Bigg)\\
&+ O\left(\varepsilon^3\right)\Bigg\}
\end{aligned}
$$

となる．

4 Rényi divergence の漸近展開

$f(x \mid \theta)$ の台を $S(\theta) = \{x \mid f(x \mid \theta) > 0\}$ とする．$0 < t < 1$ とする．pdf $f(x \mid \theta)$ と $f(x \mid \theta+\varepsilon)$ に対して

$$
H_t(\theta, \theta+\varepsilon) = \int_{S(\theta)\cap S(\theta+\varepsilon)} f^{1-t}(x \mid \theta) f^t(x \mid \theta+\varepsilon)\mathrm{d}x
$$

とおく．$\log f(x \mid \theta+\varepsilon) = \log\frac{1}{\theta+\varepsilon}g\left(\frac{x}{\theta+\varepsilon}\right) = -\log(\theta+\varepsilon) + \log g\left(\frac{x}{\theta+\varepsilon}\right)$ の θ の周りでの Taylor 展開より，$\varepsilon \to 0$ のとき

$$
\begin{aligned}
&\log f(x \mid \theta+\varepsilon)\\
=&\log f(x \mid \theta) + \frac{\partial}{\partial\theta}\log f(x \mid \theta)\varepsilon + \frac{1}{2}\frac{\partial^2}{\partial\theta^2}\log f(x \mid \theta)\varepsilon^2 + O\left(\varepsilon^3\right)\\
=&\log f(x \mid \theta) + \frac{\frac{\partial}{\partial\theta}f(x \mid \theta)}{f(x \mid \theta)}\varepsilon + \frac{1}{2}\left[\frac{\frac{\partial^2}{\partial\theta^2}f(x \mid \theta)}{f(x \mid \theta)} - \left\{\frac{\frac{\partial}{\partial\theta}f(x \mid \theta)}{f(x \mid \theta)}\right\}^2\right]\varepsilon^2 + O\left(\varepsilon^3\right)\\
=&\log f(x \mid \theta) + \left\{-\frac{1}{\theta} - \frac{xg'\left(\frac{x}{\theta}\right)}{\theta^2g\left(\frac{x}{\theta}\right)}\right\}\varepsilon + \frac{1}{2}\left[\frac{2}{\theta^2} + \frac{4xg'\left(\frac{x}{\theta}\right)}{\theta^3g\left(\frac{x}{\theta}\right)} + \frac{x^2g''\left(\frac{x}{\theta}\right)}{\theta^4g\left(\frac{x}{\theta}\right)} - \left\{-\frac{1}{\theta} - \frac{xg'\left(\frac{x}{\theta}\right)}{\theta^2g\left(\frac{x}{\theta}\right)}\right\}^2\right]\varepsilon^2 + O\left(\varepsilon^3\right)
\end{aligned}
$$

となるので，

$$
\begin{aligned}
&H_t(\theta,\theta+\varepsilon)\\
&=\int_S f^{1-t}(x\mid\theta)f^t(x\mid\theta+\varepsilon)\mathrm{d}x\\
&=\int_S \exp\{(1-t)\log f(x\mid\theta)+t\log f(x\mid\theta+\varepsilon)\}\,\mathrm{d}x\\
&=\int_S f(x\mid\theta)\Bigg[1+\frac{\partial}{\partial\theta}\log f(x\mid\theta)t\varepsilon+\frac{1}{2}\frac{\partial^2}{\partial\theta^2}\log f(x\mid\theta)t\varepsilon^2\\
&\quad+\frac{1}{2}\left\{\frac{\partial}{\partial\theta}\log f(x\mid\theta)\right\}^2t^2\varepsilon^2\Bigg]\mathrm{d}x+O\left(\varepsilon^3\right)\\
&=\int_S f(x\mid\theta)\Bigg[1+\frac{\frac{\partial}{\partial\theta}f(x\mid\theta)}{f(x\mid\theta)}t\varepsilon+\frac{1}{2}\left[\frac{\frac{\partial^2}{\partial\theta^2}f(x\mid\theta)}{f(x\mid\theta)}-\left\{\frac{\frac{\partial}{\partial\theta}f(x\mid\theta)}{f(x\mid\theta)}\right\}^2\right]t\varepsilon^2\\
&\quad+\frac{1}{2}\left\{\frac{\frac{\partial}{\partial\theta}f(x\mid\theta)}{f(x\mid\theta)}\right\}^2t^2\varepsilon^2\Bigg]+O\left(\varepsilon^3\right)\\
&=\int_S\frac{1}{\theta}g\left(\frac{x}{\theta}\right)\Bigg[1+\left\{-\frac{1}{\theta}-\frac{xg'\left(\frac{x}{\theta}\right)}{\theta^2g\left(\frac{x}{\theta}\right)}\right\}t\varepsilon+\frac{1}{2}\left[\frac{2}{\theta^2}+\frac{4xg'\left(\frac{x}{\theta}\right)}{\theta^3g\left(\frac{x}{\theta}\right)}+\frac{x^2g''\left(\frac{x}{\theta}\right)}{\theta^4g\left(\frac{x}{\theta}\right)}-\left\{-\frac{1}{\theta}-\frac{xg'\left(\frac{x}{\theta}\right)}{\theta^2g\left(\frac{x}{\theta}\right)}\right\}^2\right]t\varepsilon^2\\
&\quad+\frac{1}{2}\left\{-\frac{1}{\theta}-\frac{xg'\left(\frac{x}{\theta}\right)}{\theta^2g\left(\frac{x}{\theta}\right)}\right\}^2t^2\varepsilon^2\Bigg]\mathrm{d}x+O\left(\varepsilon^3\right)\\
&=\int_{S_0}g(y)\Bigg[1-\frac{1}{\theta}\left\{1+\frac{yg'(y)}{g(y)}\right\}t\varepsilon+\frac{1}{2\theta^2}\left[2+\frac{4yg'(y)}{g(y)}+\frac{y^2g''(y)}{g(y)}-\left\{1+\frac{yg'(y)}{g(y)}\right\}^2\right]t\varepsilon^2\\
&\quad+\frac{1}{2\theta^2}\left\{1+\frac{yg'(y)}{g(y)}\right\}^2t^2\varepsilon^2\Bigg]\mathrm{d}y+O\left(\varepsilon^3\right)\quad(\varepsilon\to 0)
\end{aligned}
\tag{4.1}
$$

を得る．ただし，$S_0=S(1)\cap S(1+(\varepsilon/\theta))$ とする．(4.1) は ε の冪により次のように分解できる．

$$
\begin{aligned}
&H_t(\theta,\theta+\varepsilon)\\
&=\int_{S_0}g(y)\mathrm{d}y-\frac{1}{\theta}\left\{\int_{S_0}g(y)\mathrm{d}y+\int_{S_0}yg'(y)\mathrm{d}y\right\}t\varepsilon\\
&\quad+\frac{1}{2\theta^2}\left[2\int_{S_0}g(y)\mathrm{d}y+4\int_{S_0}yg'(y)\mathrm{d}y+\int_{S_0}y^2g''(y)\mathrm{d}y-I_0'(1)\right]t\varepsilon^2+\frac{1}{2\theta^2}I_0'(1)t^2\varepsilon^2+O\left(\varepsilon^3\right)\\
&=\int_{S_0}g(y)\mathrm{d}y-\frac{1}{\theta}\left\{\int_{S_0}g(y)\mathrm{d}y+\int_{S_0}yg'(y)\mathrm{d}y\right\}t\varepsilon\\
&\quad+\frac{1}{2\theta^2}\left[2\int_{S_0}g(y)\mathrm{d}y+4\int_{S_0}yg'(y)\mathrm{d}y+\int_{S_0}y^2g''(y)\mathrm{d}y-\theta^2I_0'(\theta)\right]t\varepsilon^2+\frac{1}{2}I_0'(\theta)t^2\varepsilon^2+O\left(\varepsilon^3\right)\ (\varepsilon\to 0).
\end{aligned}
$$

この式は部分積分によって

$$
\begin{aligned}
H_t(\theta,\theta+\varepsilon)&=\int_{S_0}g(y)\mathrm{d}y-\frac{1}{\theta}[yg(y)]_{S_0}t\varepsilon\\
&\quad+\frac{1}{2\theta^2}\left(2[yg(y)]_{S_0}+[y^2g'(y)]_{S_0}-\theta^2I_0'(\theta)\right)t\varepsilon^2+\frac{1}{2}I_0'(\theta)t^2\varepsilon^2+O\left(\varepsilon^3\right)\ (\varepsilon\to 0)
\end{aligned}
\tag{4.2}
$$

と書き換えることができる．ただし，$S_0 = [c, d]$ であるとき，$[yg(y)]_{S_0} = [yg(y)]_c^d = dg(d) - cg(c)$, $[y^2 g'(y)]_{S_0} = [y^2 g'(y)]_c^d = d^2 g'(d) - c^2 g'(c)$ とする．

付記より，(4.2) は，“$\varepsilon > 0$ かつ $0 < a < b$” または “$\varepsilon < 0$ かつ $a < b < 0$” のとき

$$\begin{aligned}
&H_t(\theta, \theta + \varepsilon)\\
=&1 + \frac{\varepsilon\{atg(a) - g(a) - btg(b)\}}{\theta}\\
&+ \frac{\varepsilon^2 \left\{a^2 tg'(a) - g'(a) + b^2 tg'(b) + 2btg(b) + \theta^2 I_0'(\theta) t(t-1)\right\}}{2\theta^2} + O\left(\varepsilon^3\right)\\
=&1 + \frac{\varepsilon\{atg(a) - g(a) - btg(b)\}}{\theta}\\
&+ \frac{\varepsilon^2 \left\{a^2 tg'(a) - g'(a) + b^2 tg'(b) + 2btg(b) + \theta^2 I_0(\theta) t(t-1)\right\}}{2\theta^2} + O\left(\varepsilon^3\right) \ (\varepsilon \to 0)
\end{aligned}$$

となる．同様に，“$\varepsilon > 0$ かつ $a < 0 < b$” のとき

$$\begin{aligned}
H_t(\theta, \theta + \varepsilon) =&1 + \frac{\varepsilon}{\theta}\{atg(a) - btg(b)\}\\
&+ \frac{\varepsilon^2}{2\theta^2}\left\{2btg(b) - 2atg(a) + b^2 tg'(b) - a^2 tg'(a) + \theta^2 t(t-1) I_0(\theta)\right\} + O\left(\varepsilon^3\right) \ (\varepsilon \to 0)
\end{aligned}$$

となる．次に “$\varepsilon > 0$ かつ $a < b < 0$” または “$\varepsilon < 0$ かつ $0 < a < b$” のとき

$$\begin{aligned}
H_t(\theta, \theta + \varepsilon) =&1 + \frac{\varepsilon}{\theta}\{atg(a) - btg(b) + g(b)\}\\
&+ \frac{\varepsilon^2}{2\theta^2}\left\{a^2(-t)g'(a) - 2atg(a) - b^2 tg'(b) + g'(b) + \theta^2 t(t-1) I_0'(\theta)\right\} + O\left(\varepsilon^3\right)\\
=&1 + \frac{\varepsilon}{\theta}\{atg(a) - btg(b) + g(b)\}\\
&+ \frac{\varepsilon^2}{2\theta^2}\left\{-a^2 tg'(a) - 2atg(a) - b^2 tg'(b) + g'(b) + \theta^2 t(t-1) I_0(\theta)\right\} + O\left(\varepsilon^3\right) \ (\varepsilon \to 0)
\end{aligned}$$

となる．最後に “$\varepsilon < 0$ かつ $a < 0 < b$” のときを考える．

$$\begin{aligned}
H_t(\theta, \theta + \varepsilon) =&1 + \frac{\varepsilon}{\theta}\{atg(a) - g(a) - btg(b) + g(b)\}\\
&+ \frac{\varepsilon^2}{2\theta^2}\left\{2a^2 tg'(a) - g'(a) - 2b^2 tg'(b) + g'(b) + \theta^2 t(t-1) I_0'(\theta)\right\} + O\left(\varepsilon^3\right)\\
=&1 + \frac{\varepsilon}{\theta}\{atg(a) - g(a) - btg(b) + g(b)\}\\
&+ \frac{\varepsilon^2}{2\theta^2}\left\{2a^2 tg'(a) - g'(a) - 2b^2 tg'(b) + g'(b) + \theta^2 t(t-1) I_0(\theta)\right\} + O\left(\varepsilon^3\right) \ (\varepsilon \to 0)
\end{aligned}$$

となる．

特に，上記の $H_t(\theta, \theta+\varepsilon)$ の漸近展開より，$f(x \mid \theta)$ と $f(x \mid \theta+\varepsilon)$ の間の affinity $D_{1/2}(f(x \mid \theta) \,\|\, f(x \mid \theta + \varepsilon))$ の漸近展開を得ることもできる．

さらに，$H_t(\theta, \theta + \varepsilon)$ と Rényi divergence $D_s(f(\cdot \mid \theta) \,\|\, f(\cdot \mid \theta + \varepsilon))$ の間には

$$D_s(f(\cdot \mid \theta) \,\|\, f(\cdot \mid \theta + \varepsilon)) = \frac{1}{s-1} \log H_{1-s}(\theta, \theta + \varepsilon)$$

の関係があるから，Rényi divergence の漸近展開は以下のようになる．“$\varepsilon > 0$ かつ $0 < a < b$” または “$\varepsilon < 0$ かつ $a < b < 0$” のとき

$$
\begin{aligned}
&D_s(f(\cdot \mid \theta) \parallel f(\cdot \mid \theta + \varepsilon)) \\
=&\frac{\varepsilon}{\theta(1-s)}\{-a(1-s)g(a) + g(a) + b(1-s)g(b)\} \\
&- \frac{\varepsilon^2}{2\theta^2(1-s)}\Big\{(g(a) - a(1-s)g(a) + b(1-s)g(b))^2 \\
&+ g'(a) - a^2(1-s)g'(a) - b^2(1-s)g'(b) - 2btg(b) + \theta^2 s(s-1)I_0(\theta)\Big\} + O\left(\varepsilon^3\right) \ (\varepsilon \to 0),
\end{aligned}
$$

“$\varepsilon > 0$ かつ $a < 0 < b$” のとき

$$
\begin{aligned}
&D_s(f(\cdot \mid \theta) \parallel f(\cdot \mid \theta + \varepsilon)) \\
=&\frac{\varepsilon}{\theta}\{bg(b) - ag(a)\} + \frac{\varepsilon^2(1-s)}{2\theta^2}\Big\{2bg(b) - 2ag(a) + b^2 g'(b) - a^2 g'(a) - s\theta^2 I_0(\theta)\Big\} + O\left(\epsilon^3\right) \ (\varepsilon \to 0),
\end{aligned}
$$

“$\varepsilon > 0$ かつ $a < b < 0$” または “$\varepsilon < 0$ かつ $0 < a < b$” のとき

$$
\begin{aligned}
&D_s(f(\cdot \mid \theta) \parallel f(\cdot \mid \theta + \varepsilon)) \\
=&\frac{\varepsilon}{\theta(1-s)}\{-a(1-s)g(a) + b(1-s)g(b) - g(b)\} \\
&+ \frac{\varepsilon^2}{2(1-s)\theta^2}\Big\{(a(1-s)g(a) - b(1-s)g(b) + g(b))^2 \\
&+ a^2(1-s)g'(a) + 2a(1-s)g(a) + b^2(1-s)g'(b) - g'(b) - \theta^2 s(1-s)\theta^2 I_0(\theta)\Big\} + O\left(\varepsilon^3\right) \ (\varepsilon \to 0),
\end{aligned}
$$

“$\varepsilon < 0$ かつ $a < 0 < b$” のとき

$$
\begin{aligned}
&D_s(f(\cdot \mid \theta) \parallel f(\cdot \mid \theta + \varepsilon)) \\
=&\frac{\varepsilon}{\theta(1-s)}\{-a(1-s)g(a) + g(a) + b(1-s)g(b) - g(b)\} \\
&+ \frac{\varepsilon^2}{2t\theta^2}\Big\{-2a^2(1-s)g'(a) + g'(a) + 2b^2(1-s)g'(b) - g'(b) - s(1-s)\theta^2 I_0(\theta) \\
&+ (a(1-s)g(a) - g(a) - b(1-s)g(b) + g(b))^2\Big\} + O\left(\varepsilon^3\right) \ (\varepsilon \to 0)
\end{aligned}
$$

となる．

謝辞

本研究は，京都大学数理解析研究所の国際共同利用・共同研究拠点事業の支援を受けた．

参考文献

[1] Akahira, M. (1996). Loss of information of a statistic for a family of non-regular distributions. *Ann. Inst. Stat. Math.*, **48** (2), 349–364.

[2] Hayashi, M. (2010). Two non-regular extensions of large deviation bound. *Commun. Statist. Theory Meth.*, **39**, 1404–1423.

[3] Hayashi, M. (2010). Limiting behavior of relative Rényi entropy in a non-regular location shift family. *Ann. Inst. Stat. Math.*, **62**, 547–569.

[4] Rényi, A. (1961). On measures of entropy and information. *Proc. Fourth Berkeley Symp. on Math. Stat. and Prob.*, **1**, 547–561.

[5] Van Erven, T. and Harremoës, P. (2014). Rényi divergence and Kullback-Leibler divergence. *IEEE Trans. Inf. Theory*, **60**(7), 3797–3820.

付記：(4.1) の各項の漸近展開

$g(x) = f(x \mid 1)$ の台が (a,b) のとき，$f(x \mid 1+(\varepsilon/\theta))$ の台は $S(1+(\varepsilon/\theta)) = (a(1+(\varepsilon/\theta)), b(1+(\varepsilon/\theta)))$ となる．$|\varepsilon|$ が十分小さくなるように ε をとると，$\varepsilon > 0$ のとき，a, b の正負によって，$S_0 = S(1) \cap S(1+(\varepsilon/\theta))$ は，$\varepsilon > 0$ のとき

$$S_0 = \begin{cases} (a(1+(\varepsilon/\theta)), b) & (0 < a < b), \\ (a, b) & (a < 0 < b), \\ (a, b(1+(\varepsilon/\theta))) & (a < b < 0) \end{cases}$$

であり，$\varepsilon < 0$ のとき

$$S_0 = \begin{cases} (a, b(1+(\varepsilon/\theta))) & (0 < a < b), \\ (a(1+(\varepsilon/\theta)), b(1+(\varepsilon/\theta))) & (a < 0 < b), \\ (a(1+(\varepsilon/\theta)), b) & (a < b < 0) \end{cases}$$

である．ここで，部分積分より

$$\int_{S_0} x g'(x)\mathrm{d}x = [x g(x)]_{S_0} - \int_{S_0} g(x)\mathrm{d}x,$$
$$\int_{S_0} x^2 g''(x)\mathrm{d}x = \left[x^2 g'(x)\right]_{S_0} - 2\,[x g(x)]_{S_0} + 2\int_{S_0} g(x)\mathrm{d}x$$

である．(4.1) の各項は ε や S_0 の端点の正負に応じて場合分けされる．例えば，$\varepsilon > 0$ の場合は，$\varepsilon \to +0$ のとき

$$\int_{S_0} g(x)\mathrm{d}x$$
$$= \begin{cases} \int_{a(1+(\varepsilon/\theta))}^{b} g(x)\mathrm{d}x = \int_a^b g(x)\mathrm{d}x - \int_a^{a(1+(\varepsilon/\theta))} g(x)\mathrm{d}x & \\ = 1 - g(a)a(\varepsilon/\theta) - \frac{g'(a)}{2}a^2(\varepsilon/\theta)^2 + o(\varepsilon^2) & (0 < a < b), \\ \int_a^b g(x)\mathrm{d}x = 1 & (a < 0 < b), \\ \int_a^{b(1+(\varepsilon/\theta))} g(x)\mathrm{d}x = \int_a^b g(x)\mathrm{d}x - \int_{b(1+(\varepsilon/\theta))}^{b} g(x)\mathrm{d}x & \\ = 1 + g(b)b(\varepsilon/\theta) + \frac{g'(b)}{2}b^2(\varepsilon/\theta)^2 + o\left(\varepsilon^2\right) & (a < b < 0) \end{cases}$$

となる．また，$\varepsilon < 0$ の場合は，$\varepsilon \to -0$ のとき

$$
\begin{aligned}
&\int_{S_0} g(x)\mathrm{d}x \\
&= \begin{cases}
\int_a^{b(1+(\varepsilon/\theta))} g(x)\mathrm{d}x = \int_a^b g(x)\mathrm{d}x - \int_{b(1+(\varepsilon/\theta))}^b g(x)\mathrm{d}x & \\
= 1 + g(b)b(\varepsilon/\theta) + \frac{g'(b)}{2} b^2(\varepsilon/\theta)^2 + o(\varepsilon^2) & (0 < a < b), \\
\int_{a(1+(\varepsilon/\theta))}^{b(1+(\varepsilon/\theta))} g(x)\mathrm{d}x = \int_a^b g(x)\mathrm{d}x - \int_a^{a(1+(\varepsilon/\theta))} g(x)\mathrm{d}x - \int_{b(1+(\varepsilon/\theta))}^b g(x)\mathrm{d}x & \\
= 1 - g(a)a(\varepsilon/\theta) + g(b)b(\varepsilon/\theta) - \frac{g'(a)}{2} a^2(\varepsilon/\theta)^2 + \frac{g'(b)}{2} b^2(\varepsilon/\theta)^2 + o(\varepsilon^2) & (a < 0 < b), \\
\int_{a(1+(\varepsilon/\theta))}^b g(x)\mathrm{d}x = \int_a^b g(x)\mathrm{d}x - \int_a^{a(1+(\varepsilon/\theta))} g(x)\mathrm{d}x & \\
= 1 - g(a)a(\varepsilon/\theta) - \frac{g'(a)}{2} a^2(\varepsilon/\theta)^2 + o(\varepsilon^2) & (a < b < 0)
\end{cases}
\end{aligned}
$$

となる．また，$\varepsilon \to +0$ のとき

$$
\begin{aligned}
&[xg(x)]_{S_0} \\
&= \begin{cases}
bg(b) - a(1 + (\varepsilon/\theta))g(a(1 + (\varepsilon/\theta))) & \\
= (bg(b) - ag(a)) - a(\varepsilon/\theta)\left(ag'(a) + g(a)\right) + a(\varepsilon/\theta)^2 \left(-\frac{1}{2}a^2 g''(a) - ag'(a)\right) + O\left(\varepsilon^3\right) & (0 < a < b), \\
bg(b) - ag(a) & (a < 0 < b), \\
b(1 + (\varepsilon/\theta))g(b(1 + \varepsilon)) - ag(a) & \\
= (bg(b) - ag(a)) + b(\varepsilon/\theta)\left(bg'(b) + g(b)\right) + b(\varepsilon/\theta)^2 \left(\frac{1}{2}b^2 g''(b) + bg'(b)\right) + O\left(\varepsilon^3\right) & (a < b < 0),
\end{cases}
\end{aligned}
$$

$$
\begin{aligned}
&[x^2 g'(x)]_{S_0} \\
&= \begin{cases}
b^2 g'(b) - a^2(1 + (\varepsilon/\theta))^2 g'(a(1 + (\varepsilon/\theta))) & \\
= \left(b^2 g'(b) - a^2 g'(a)\right) + a^2(\varepsilon/\theta)\left(-ag''(a) - 2g'(a)\right) & \\
+a^2(\varepsilon/\theta)^2 \left(-\frac{1}{2}a^2 g^{(3)}(a) - 2ag''(a) - g'(a)\right) + O\left(\varepsilon^3\right) & (0 < a < b), \\
b^2 g'(b) - a^2 g'(a) & (a < 0 < b), \\
b^2(1 + (\varepsilon/\theta))^2 g'(b(1 + (\varepsilon/\theta))) - a^2 g'(a) & \\
= \left(b^2 g'(b) - a^2 g'(a)\right) + b^2 \varepsilon \left(bg''(b) + 2g'(b)\right) & \\
+b^2(\varepsilon/\theta)^2 \left(\frac{1}{2}b^2 g^{(3)}(b) + 2bg''(b) + g'(b)\right) + O\left(\varepsilon^3\right) & (a < b < 0)
\end{cases}
\end{aligned}
$$

となる．$\varepsilon \to -0$ のとき

$$
\begin{aligned}
&[xg(x)]_{S_0}\\
&=\begin{cases}
b(1+\varepsilon)g(b(1+(\varepsilon/\theta)))-ag(a)\\
=(bg(b)-ag(a))+b(\varepsilon/\theta)\left(bg'(b)+g(b)\right)+b(\varepsilon/\theta)^2\left(\frac{1}{2}b^2g''(b)+bg'(b)\right)+O\left(\varepsilon^3\right) & (0<a<b),\\
b(1+(\varepsilon/\theta))g(b(1+(\varepsilon/\theta)))-a(1+(\varepsilon/\theta))g(a(1+(\varepsilon/\theta)))\\
=(bg(b)-ag(a))+(\varepsilon/\theta)\left(b\left(bg'(b)+g(b)\right)-a\left(ag'(a)+g(a)\right)\right)\\
+\frac{1}{2}(\varepsilon/\theta)^2\left(a^3\left(-g''(a)\right)-2a^2g'(a)+b^3g''(b)+2b^2g'(b)\right)+O\left(\varepsilon^3\right) & (a<0<b),\\
bg(b)-a(1+(\varepsilon/\theta))g(a(1+(\varepsilon/\theta)))\\
=(bg(b)-ag(a))-a(\varepsilon/\theta)\left(ag'(a)+g(a)\right)+a(\varepsilon/\theta)^2\left(-\frac{1}{2}a^2g''(a)-ag'(a)\right)+O\left(\varepsilon^3\right) & (a<b<0),
\end{cases}\\
&[x^2g'(x)]_{S_0}\\
&=\begin{cases}
b^2(1+(\varepsilon/\theta))^2g'(b(1+(\varepsilon/\theta)))-a^2g'(a)\\
=\left(b^2g'(b)-a^2g'(a)\right)+b^2(\varepsilon/\theta)\left(bg''(b)+2g'(b)\right)\\
+b^2(\varepsilon/\theta)^2\left(\frac{1}{2}b^2g^{(3)}(b)+2bg''(b)+g'(b)\right)+O\left(\varepsilon^3\right) & (0<a<b),\\
b^2(1+(\varepsilon/\theta))^2g'(b(1+(\varepsilon/\theta)))-a^2(1+(\varepsilon/\theta))^2g'(a(1+(\varepsilon/\theta)))\\
=\left(b^2g'(b)-a^2g'(a)\right)+(\varepsilon/\theta)\left(a^2\left(-ag''(a)-2g'(a)\right)+b^2\left(bg''(b)+2g'(b)\right)\right)\\
+(\varepsilon/\theta)^2\Bigg(a^2\left(-\frac{1}{2}a^2g^{(3)}(a)-2ag''(a)-g'(a)\right)\\
+b^2\left(\frac{1}{2}b^2g^{(3)}(b)+2bg''(b)+g'(b)\right)\Bigg)+O\left(\varepsilon^3\right) & (a<0<b),\\
b^2g'(b)-a^2(1+(\varepsilon/\theta))^2g'(a(1+(\varepsilon/\theta)))\\
=\left(b^2g'(b)-a^2g'(a)\right)+a^2(\varepsilon/\theta)\left(-ag''(a)-2g'(a)\right)\\
+a^2(\varepsilon/\theta)^2\left(-\frac{1}{2}a^2g^{(3)}(a)-2ag''(a)-g'(a)\right)+O\left(\varepsilon^3\right) & (a<b<0)
\end{cases}
\end{aligned}
$$

となる．