JIRSS (2006)
Vol. 5, Nos. 1-2, pp 25-37

Relations between Renyi Distance and Fisher Information

M. Abbasnejad, N. R. Arghami

Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Iran. (arghami_nr@yahoo.com)

Abstract. In this paper, we first show that Renyi distance between any member of a parametric family and its perturbations, is proportional to its Fisher information. We, then, prove some relations between the Renyi distance of two distributions and the Fisher information of their exponentially twisted family of densities. Finally, we show that the partial ordering of families induced by Renyi distance is the same as that induced by Fisher information.

1 Introduction and Preliminaries

Consider a parametric density $p(x;\theta)$ defined over a probability space $\mathcal{X}$ parameterized by $\theta \in \Omega$. Renyi information (distance) between $p(x;\theta_0)$ and $p(x;\theta_1)$ is defined by

$$D^{\alpha}(p(x;\theta_1), p(x;\theta_0)) = \frac{1}{\alpha-1} \log \int_{\mathcal{X}} \left[\frac{p(x;\theta_1)}{p(x;\theta_0)}\right]^{\alpha-1} p(x;\theta_1)dx$$

for all $\alpha > 0$ $(\alpha \neq 1)$ (Renyi, 1961).

Kullback-Leibler (K-L) information is a limiting case of the above

Key words and phrases: Exponentialy twisted family, Fisher information, Renyi information.

as α tends to 1; that is,

$$\lim_{\alpha \to 1} D^{\alpha}(p_{\theta_1}, p_{\theta_0}) = D(p_{\theta_1}, p_{\theta_0}) = \int_{\mathcal{X}} p(x;\theta_1) \log \frac{p(x;\theta_1)}{p(x;\theta_0)} dx.$$

The paper is organized as follows: In section 2, we show that when $\theta_1 = \theta_0 + \delta$, where δ is a perturbation, Renyi distance between p_{θ_1} and p_{θ_0} is proportional to the family's Fisher information at θ_0.

In section 3, we generalize the relation between Renyi and Fisher information when the condition "δ small" may not hold and when we do not have parametric densities. Thus we prove that for two exponential families, their order in terms of Renyi information is the same as their order in terms of Fisher information. Finally, in section 4, we prove a general relation between Renyi and Fisher information, generalizing the result of Habibi *et al* (2006), who proved a general relation between K-L and Fisher information.

2 A link between Renyi and Fisher information in the case of parametric families

Suppose that θ_0 and $\theta_0 + \delta$ are neighboring points in the parameter space Ω. The following Theorem states in effect that for small δ, $D^{\alpha}(p_{\theta_0+\delta}, p_{\theta_0})$ and its derivatives are proportional to Fisher information at $\theta = \theta_0$.

Theorem 2.1. *Suppose that density $p(x;\theta)$ satisfies the following regularity conditions:*
1. For all $x \in \mathcal{X}$, $p(x;\theta)$ is twice differentiable with respect to θ and integral and derivatives can be interchanged.
2. Fisher information is finite.
3. For every $\theta \in \Omega$, $\left|\frac{\partial}{\partial\theta} p(x;\theta)\right| < K(x)$ and $\left|\frac{\partial^2}{\partial\theta^2} p(x;\theta)\right| < H(x)$, where $K(x)$ and $H(x)$ are integrable function over $\mathcal{X}$.
Then

$$(i) \qquad \lim_{\delta \to 0} \frac{1}{\delta^2} D^{\alpha}(p_{\theta_0+\delta}, p_{\theta_0}) = \frac{\alpha}{2} I_X(\theta_0)$$

$$(ii) \qquad \frac{\partial^2}{\partial\theta^2} D^{\alpha}(p_{\theta}, p_{\theta_0})|_{\theta=\theta_0} = \alpha I_X(\theta_0).$$

The above two relations for $\alpha = 1$ were proved by Kullback (1959, p 27).

Proof. (i) Assuming we can take the differentiation under the integral sign, we can write

$$\begin{aligned}
&\lim_{\delta\to 0}\frac{1}{\delta^2}D^\alpha(p_{\theta_0+\delta},p_{\theta_0})\\
&= \lim_{\delta\to 0}\frac{(1/(\alpha-1))\log\int p^\alpha_{\theta_0+\delta}(x)p^{1-\alpha}_{\theta_0}(x)dx}{\delta^2}\\
&= \lim_{\delta\to 0}\frac{\int \alpha p^{\alpha-1}_{\theta_0+\delta}(x)p^{1-\alpha}_{\theta_0}(x)\frac{\partial}{\partial\theta}p_{\theta_0+\delta}(x)dx}{2\delta(\alpha-1)\int p^\alpha_{\theta_0+\delta}(x)p^{1-\alpha}_{\theta_0}(x)dx}\\
&= \frac{\alpha}{2}\lim_{\delta\to 0}\frac{\int p^{\alpha-2}_{\theta_0+\delta}(x)[\frac{\partial p_{\theta_0+\delta}(x)}{\partial\delta}]^2 p^{1-\alpha}_{\theta_0}(x)dx}{\int p^{1-\alpha}_{\theta_0}(x)\left[p^\alpha_{\theta_0+\delta}(x)+\delta\alpha p^{\alpha-1}_{\theta_0+\delta}(x)\frac{\partial p_{\theta_0+\delta}(x)}{\partial\delta}\right]dx}\\
&+ \lim_{\delta\to 0}\frac{\alpha\int p^{\alpha-1}_{\theta_0+\delta}(x)\frac{\partial^2 p_{\theta_0+\delta}(x)}{\partial\delta^2}p^{1-\alpha}_{\theta_0}(x)dx}{2(\alpha-1)\int p^{1-\alpha}_{\theta_0}(x)\left[p^\alpha_{\theta_0+\delta}(x)+\delta\alpha p^{\alpha-1}_{\theta_0+\delta}(x)\frac{\partial p_{\theta_0+\delta}(x)}{\partial\delta}\right]dx},
\end{aligned}$$

by regularity conditions, it is possible to change the order of integration and taking limits. Hence

$$\begin{aligned}
\lim_{\delta\to 0}\frac{1}{\delta^2}D^\alpha(p_{\theta_0+\delta},p_{\theta_0}) &= \frac{\alpha}{2}\frac{\int\frac{[p'_{\theta_0}(x)]^2}{p_{\theta_0}(x)}dx}{\int p_{\theta_0}(x)dx}+\frac{\alpha\int p''_{\theta_0}(x)dx}{2(\alpha-1)\int p_{\theta_0}(x)dx}\\
&= \frac{\alpha}{2}\int\frac{(p'_{\theta_0}(x))^2}{p_{\theta_0}(x)}dx\\
&= \frac{\alpha}{2}I_X(\theta_0),
\end{aligned}$$

where p'_{θ_0} and p''_{θ_0} denote the first and the second derivatives of p_{θ_0}.
(ii) We have

$$D^\alpha(p_\theta,p_{\theta_0})=\frac{1}{\alpha-1}\log\int p^\alpha_\theta(x)p^{1-\alpha}_{\theta_0}(x)dx\ .$$

Then

$$\frac{\partial}{\partial\theta}D^\alpha(p_\theta,p_{\theta_0})=\frac{1}{\alpha-1}\frac{\int\alpha p^{\alpha-1}_\theta(x)p^{1-\alpha}_{\theta_0}(x)p'_\theta(x)dx}{\int p^\alpha_\theta(x)p^{1-\alpha}_{\theta_0}(x)dx}$$

and

$$\begin{aligned}
&\frac{\partial^2}{\partial\theta^2} D^\alpha(p_\theta, p_{\theta_0}) \\
&= \alpha \frac{[\int p_\theta^{\alpha-2}(x) p_{\theta_0}^{1-\alpha}(x)(p'_\theta(x))^2 dx][\int p_\theta^\alpha(x) p_{\theta_0}^{1-\alpha}(x)dx]}{\left[\int p_\theta^\alpha(x) p_{\theta_0}^{1-\alpha}(x)dx\right]^2} \\
&+ \frac{1}{\alpha-1} \frac{[\int \alpha p_\theta^{\alpha-1}(x) p_{\theta_0}^{1-\alpha}(x) p''_\theta(x)dx][\int p_\theta^\alpha(x) p_{\theta_0}^{1-\alpha}(x)dx]}{\left[\int p_\theta^\alpha(x) p_{\theta_0}^{1-\alpha}(x)dx\right]^2} \\
&- \frac{1}{\alpha-1} \frac{\left[\int \alpha p_\theta^{\alpha-1}(x) p_{\theta_0}^{1-\alpha}(x) p'_\theta(x)dx\right]^2}{\left[\int p_\theta^\alpha(x) p_{\theta_0}^{1-\alpha}(x)dx\right]^2}
\end{aligned}$$

and if $\theta = \theta_0$ we have

$$\begin{aligned}
\frac{\partial^2}{\partial\theta^2} D^\alpha(p_\theta, p_{\theta_0})|_{\theta=\theta_0} &= \alpha \frac{[\int \left(\frac{(p'_{\theta_0}(\mathbf{X}))^2}{p_{\theta_0}(\mathbf{X})} + \alpha p''_{\theta_0}(x)\right) dx][\int p_{\theta_0}(x)dx]}{[\int p_{\theta_0}(x)dx]^2} \\
&- \frac{[\int \alpha p'_{\theta_0}(x)dx]^2}{(\alpha-1)[\int p_{\theta_0}(x)dx]^2} \\
&= \alpha \int \frac{(p'_{\theta_0}(x))^2}{p_{\theta_0}(x)} dx \\
&= \alpha I_X(\theta_0). \qquad \square
\end{aligned}$$

3 Exact relations between Renyi distance and Fisher information

Consider two probability density functions $p_0(x)$ and $p_1(x)$ defined on a probability space $\mathcal{X}$. They could be arbitrary densities, not necessarily members of a parametric family of densities. Suppose p_1 and p_0 have common support. The following parametric family of densities is well defined.

Definition 3.1. Let $p_0(x)$ and $p_1(x)$ be two densities with common support $\mathcal{X}$. Then the family of densities defined by

$$p_t(x) = \frac{1}{N_t} p_0(x) \left[\frac{p_1(x)}{p_0(x)}\right]^t \qquad x \in \chi, \quad t > 0,$$

is called *the exponentially twisted family of densities of* p_0 *and* p_1 and is denoted by $ET(p_0, p_1)$.

Here

$$0 < N_t = \int [p_0(x)]^{1-t}[p_1(x)]^t dx < \infty,$$

where the first inequality above is obvious and the second inequality follows from the fact that $D^\alpha(p_1, p_0) = +\infty$ if and only if $Supp(p_1) \cap Supp(p_0) = \emptyset$, or $\alpha > 1$ and $Supp(p_1)$ is not a subset of $Supp(p_0)$, where $Supp(p) = \{x : p(x) > 0\}$ (Csiszar, 1995).

Since we can write

$$p_t(x) = \frac{1}{N_t} p_0(x) \exp\left\{t \log \frac{p_1(x)}{p_0(x)}\right\},$$

the above family is a one parameter exponential family (for any pair of densities p_0 and p_1), where t is considered as the parameter.

Fisher information of $p_t(x)$ is

$$I(t) = \int_{\mathcal{X}} \left(\frac{d \log p_t(x)}{dt}\right)^2 p_t(x) dx \qquad t > 0 \tag{3.1}$$

It is easy to show that

$$\frac{d \log p_t(x)}{dt} = \log \frac{p_1(x)}{p_0(x)} - \int_{\mathcal{X}} p_t(x) \log \frac{p_1(x)}{p_0(x)} dx.$$

Substituting the above in (3.1) and simplifying gives

$$I(t) = \int_{\mathcal{X}} p_t(x) \left(\log \frac{p_1(x)}{p_0(x)}\right)^2 dx - \left(\int_{\mathcal{X}} p_t(x) \log \frac{p_1(x)}{p_0(x)} dx\right)^2.$$

By noting that Renyi information is finite and using Jensen's inequality, it follows that $I(t) < \infty$ for $t > 0$. Theorem 3.1 below generalizes the following results by Dabak and Johnson (2002).

$$(i) \quad \frac{\partial D(p_t, p_u)}{\partial t} = (t - u) I(t) \quad \forall\ t, u \in [0, 1],$$

$$(ii) \quad \frac{\partial^2 D(p_t, p_u)}{\partial t^2}|_{t=u} = I(t),$$

$$(iii) \quad D(p_t, p_0) = \int_0^t u I(u) du.$$

Theorem 3.1. *Let $p_0(x)$ and $p_1(x)$ be two densities with common support $\mathcal{X}$. Then*

$$(i) \quad \frac{\partial D^{\alpha}(p_t, p_u)}{\partial t} = \frac{\alpha}{\alpha - 1} \int_t^{\alpha(t-u)+u} I(z) dz \qquad \forall\ t > u > 0 \tag{3.2}$$

$$(ii) \quad \frac{\partial^2 D^{\alpha}(p_t, p_u)}{\partial t^2}|_{t=u} = \alpha I(t) \tag{3.3}$$

$$(iii) \quad D^{\alpha}(p_t, p_0) = \frac{\alpha}{\alpha - 1} \left[\int_0^t \int_w^{\alpha w} I(z) dz\ dw \right]. \tag{3.4}$$

Proof. To prove the theorem, we need following relations:

$$\frac{dp_t(x)}{dt} = p_t(x) \left[\log \frac{p_1(x)}{p_0(x)} - \frac{d \log N_t}{dt} \right], \tag{3.5}$$

$$\frac{d \log N_t}{dt} = \int p_t(x) \log \frac{p_1(x)}{p_0(x)} dx, \tag{3.6}$$

Also,

$$\begin{aligned}
\frac{\partial \int p_t(x) \log \frac{p_1(x)}{p_0(x)} dx}{\partial t} &= \int \frac{dp_t(x)}{dt} \log \frac{p_1(x)}{p_0(x)} dx \\
&= \int p_t(x) \left[\log \frac{p_1(x)}{p_0(x)} - \frac{d \log N_t}{dt} \right] \log \frac{p_1(x)}{p_0(x)} dx \\
&= \int p_t(x) \left[\log \frac{p_1(x)}{p_0(x)} - \int p_t(x) \log \frac{p_1(x)}{p_0(x)} dx \right] \\
&\quad \times \log \frac{p_1(x)}{p_0(x)} dx \\
&= \int p_t(x) \left[\log \frac{p_1(x)}{p_0(x)} \right]^2 dx \\
&- \left[\int p_t(x) \log \frac{p_1(x)}{p_0(x)} dx \right]^2 \\
&= I(t),
\end{aligned} \tag{3.7}$$

and

$$\frac{\partial \int p_{\alpha(t-u)+u}(x) \log \frac{p_1(x)}{p_0(x)} dx}{\partial t} = \alpha I(\alpha(t-u)+u). \tag{3.8}$$

(i) Differentiating both sides of

$$D^{\alpha}(p_t, p_u) = \frac{1}{\alpha - 1} \log \int \frac{p_t^{\alpha}(x)}{p_u^{\alpha-1}(x)} dx$$

with respect to t, we have

$$\frac{\partial D^\alpha(p_t,p_u)}{\partial t} = \frac{1}{\alpha-1}\frac{\int \alpha \frac{dp_t(x)}{dt}\frac{p_t^{\alpha-1}(x)}{p_u^{\alpha-1}(x)}dx}{\int \frac{p_t^\alpha(x)}{p_u^{\alpha-1}(x)}dx}.$$

Using the equalities (3.5) and (3.6) we can write

$$\begin{aligned}
&\frac{\partial D^\alpha(p_t,p_u)}{\partial t}\\
&= \frac{\alpha}{\alpha-1}\frac{\int \left[\frac{p_t^\alpha(x)}{p_u^{\alpha-1}(x)}\log\frac{p_1(x)}{p_0(x)} - \frac{p_t^\alpha(x)}{p_u^{\alpha-1}(x)}\int p_t(x)\log\frac{p_1(x)}{p_0(x)}dx\right]dx}{\int \frac{p_t^\alpha(x)}{p_u^{\alpha-1}(x)}dx}\\
&= \frac{\alpha}{\alpha-1}\left[\frac{\int \frac{p_t^\alpha(x)}{p_u^{\alpha-1}(x)}\log\frac{p_1(x)}{p_0(x)}dx}{\int \frac{p_t^\alpha(x)}{p_u^{\alpha-1}(x)}dx} - \int p_t(x)\log\frac{p_1(x)}{p_0(x)dx}\right]\\
&= \frac{\alpha}{\alpha-1}\left[\frac{\int (\frac{p_1(x)}{p_0(x)})^{\alpha(t-u)+u}p_0(x)\log\frac{p_1(x)}{p_0(x)}dx}{\int (\frac{p_1(x)}{p_0(x)})^{\alpha(t-u)+u}p_0(x)dx}\right]\\
&- \frac{\alpha}{\alpha-1}\int p_t(x)\log\frac{p_1(x)}{p_0(x)}dx.
\end{aligned}$$

But

$$\int (\frac{p_1(x)}{p_0(x)})^{\alpha(t-u)+u}p_0(x)dx = N_{\alpha(t-u)+u}.$$

So, we have

$$\begin{aligned}
&\frac{\partial D^\alpha(p_t,p_u)}{\partial t} =\\
&\quad \frac{\alpha}{\alpha-1}\left[\int p_{\alpha(t-u)+u}(x)\log\frac{p_1(x)}{p_0(x)}dx - \int p_t(x)\log\frac{p_1(x)}{p_0(x)}dx\right]
\end{aligned}$$

Now using relations (3.7) and (3.8), it follows that

$$\begin{aligned}
\frac{\partial D^\alpha(p_t,p_u)}{\partial t} &= \frac{\alpha}{\alpha-1}\left[\int_u^t \alpha I(\alpha(z-u)+u)dz - \int_u^t I(z)dz\right]\\
&= \frac{\alpha}{\alpha-1}\int_t^{\alpha(t-u)+u} I(z)dz.
\end{aligned}$$

(ii) To prove the second relation, we differentiate both sides of (3.2) and obtain

$$\frac{\partial^2 D^\alpha(p_t,p_u)}{\partial t^2} = \frac{\alpha}{\alpha-1}\left[\alpha I(\alpha(t-u)+u) - I(t)\right].$$

Substituting $t = u$ gives (3.3).
(iii) Integrating (3.2) with respect to t and putting $u = 0$ proves the result. □

Example 3.1. Let $p_0(x) = \sqrt{\frac{2}{\pi}} e^{-\frac{x^2}{2}}$ $x > 0$ and $p_1(x) = x e^{-\frac{x^2}{2}}$ $x > 0$. Then $p_t(x) = \frac{1}{\Gamma(\frac{t+1}{2}) 2^{\frac{t-1}{2}}} x^t e^{-\frac{x^2}{2}}$ $x > 0$. It can be easily shown that $I(t) = \frac{1}{4}\psi'(\frac{t+1}{2})$ where $\psi'(t) = \frac{d^2}{dt^2}\log\Gamma(t)$ is the trigamma function. Using the relation (3.4), we get $D_\alpha(p_t, p_0) = \frac{1}{\alpha-1}[\log\Gamma(\frac{\alpha t+1}{2}) - \alpha\log\Gamma(\frac{t+1}{2})] + \frac{1}{2}\log\pi$.

Comparing two families of distributions with respect to their Fisher information is often easier than establishing their order in terms of Renyi distance.

Now we use the above established link between Renyi and Fisher information to prove a result for exponential families.

Lemma 3.1. *If $p_0(x) = p(x;\theta_0)$ and $p_1(x) = p(x;\theta_1)$ are two members of the exponential family $\mathcal{P}$=$\{p(x;\theta), \theta \in \Omega\}$, then $ET(p_0, p_1)$ is a sub family of $\mathcal{P}$.*

Proof. The proof is easy and is given in Habibi *et al* (2006). □

Theorem 3.3. *Let P =$\{p(x;\theta),\ x \in \boldsymbol{X},\ \theta \in \Omega\}$ and Q =$\{q(y;\theta),\ y \in \boldsymbol{Y},\ \theta \in \Omega\}$ be exponential families, then the following three statements are equivalent:*

$$\begin{array}{lccl} (i) \quad I_X(\theta) & > & (<,\ =) & I_Y(\theta),\ \forall\theta \in \Omega, \\ (ii)\ D^\alpha(p_{\theta_1}, p_{\theta_0}) & > & (<,\ =) & D^\alpha(q_{\theta_1}, q_{\theta_0}),\ \forall\theta_0,\ \theta_1 \in \Omega, \\ (iii)\ J^\alpha(p_{\theta_1}, p_{\theta_0}) & > & (<,\ =) & J^\alpha(q_{\theta_1}, q_{\theta_0}),\ \forall\theta_0,\ \theta_1 \in \Omega, \end{array}$$

where p_{θ_i} and q_{θ_i} $(i = 0, 1)$ are members of P and Q, respectively and $J^\alpha(p_{\theta_1}, p_{\theta_0}) = D^\alpha(p_{\theta_1}, p_{\theta_0}) + D^\alpha(p_{\theta_0}, p_{\theta_1})$.

Proof. Assuming $\theta_0 < \theta_1$, by (3.4) we can write

$$D^\alpha(p_t, p_0) = \frac{\alpha}{\alpha - 1}\left[\int_0^t \int_w^{\alpha w} I(z)dz\ dw\right].$$

But $I_X(z) = I_X^*(\theta)[\frac{d\theta}{dz}]$, where $\theta = \theta_0 + (\theta_1 - \theta_0)z$ (Lehmann,1983, p.118). Thus

$$\int_w^{\alpha w} I_X(z)dz = \int_{\theta_0+(\theta_1-\theta_0)w}^{\theta_0+(\theta_1-\theta_0)\alpha w} I_X^*(\theta)(\theta_1 - \theta_0)^2(\frac{1}{\theta_1 - \theta_0})d\theta$$

$$= \int_{\theta_0+(\theta_1-\theta_0)w}^{\theta_0+(\theta_1-\theta_0)\alpha w} I_X^*(\theta)(\theta_1-\theta_0)d\theta.$$

Thus the result follows. □

Example 3.2. Let X have a Weibull distribution with pdf

$$f(x;\theta,\beta) = \theta\beta x^{\beta-1}e^{-\theta x^\beta} \quad x>0, \quad \theta,\beta>0$$

and Y be a random variable with gamma density as

$$g(y;\theta,\beta) = \frac{\theta^\beta}{\Gamma(\beta)}x^{\beta-1}e^{-\theta x} \quad x>0, \quad \theta,\beta>0.$$

Here $I_X(\theta) = \frac{1}{\theta^2}$ and $I_Y(\theta) = \frac{\beta}{\theta^2}$. Using Theorem 3.2 we have

$$D_\alpha(g_{\theta_1}, g_{\theta_0}) > D_\alpha(f_{\theta_1}, f_{\theta_0}) \qquad \forall \;\; \theta_1 > \theta_0, \quad \beta > 1$$

$$D_\alpha(g_{\theta_1}, g_{\theta_0}) = D_\alpha(f_{\theta_1}, f_{\theta_0}) \qquad \forall \;\; \theta_1 > \theta_0, \quad \beta = 1$$

$$D_\alpha(g_{\theta_1}, g_{\theta_0}) < D_\alpha(f_{\theta_1}, f_{\theta_0}) \qquad \forall \;\; \theta_1 > \theta_0, \quad \beta < 1.$$

4 A general theorem

Classifying a distribution family with respect to Renyi distance is not usually (algebraically) easy, while it is relatively easier with respect to Fisher information.

In this section, we generalize relations between K-L distance and Fisher information established by Habibi *et al* (2006) to the case of Renyi distance which gives a partial ordering of parametric families of distributions in terms of D^α (and hence in terms of J^α).

Theorem 4.1. *Let $\{p_\theta, \; \theta \in \Omega\}$ and $\{q_\theta, \; \theta \in \Omega\}$ be two families of densities. Assume that both family have finite Fisher and Renyi information, continuous in θ. If the following conditions hold*

(i) $I_X(\theta) - I_Y(\theta) \geq d_0 > 0, \; \forall \theta \in I = [\theta_0, \theta_1]$
(ii) The third derivatives of $D_X^\alpha(p_{\theta+\delta}, p_\theta)$ and $D_Y^\alpha(q_{\theta+\delta}, q_\theta)$ with respect to (w.r.t) δ are bounded for every $\theta \in I$ and every δ in the neighborhood $I_0 = [0, c]$ of zero, then

$$D_X^\alpha(p_{\theta_1}, p_{\theta_0}) > D_Y^\alpha(q_{\theta_1}, q_{\theta_0})$$

To prove the above Theorem along the line of proof of a similar theorem in Habibi *et al* (2006), we need the following lemmas.

Lemma 4.1. *Let θ be fixed, and let*

$$h_\theta^\alpha(\delta) = D_X^\alpha(p_{\theta+\delta}, p_\theta) = \frac{1}{\alpha-1}\log\int_X \left[\frac{p_{\theta+\delta}(x)}{p_\theta(x)}\right]^{\alpha-1} p_{\theta+\delta}(x)dx.$$

for $\alpha > 0 \quad (\alpha \neq 1)$.
Then
(i) $h_\theta^\alpha(0) = 0$
(ii) $\frac{d\ h_\theta^\alpha(\delta)}{d\ \delta}|_{\delta=0} = 0$
(iii) $[H_\theta^\alpha(\delta)]' = \frac{1}{2}h_\theta^{\alpha(3)}(\xi_2) - \frac{1}{3}h_\theta^{\alpha(3)}(\xi_1)$, *for some* $0 < \xi_1, \xi_2 < \delta$,
where $H_\theta^\alpha(\delta) = \frac{1}{\delta^2}h_\theta^\alpha(\delta)$.

Proof. Relations (i) and (ii) are easy to see.
(iii) We have

$$[H_\theta^\alpha(\delta)]' = -\frac{2}{\delta^3}h_\theta^\alpha(\delta) + \frac{1}{\delta^2}\frac{d\ h_\theta^\alpha(\delta)}{d\ \delta}.$$

By Taylor expansion we have

$$\begin{aligned}
&[H_\theta^\alpha(\delta)]' \\
&= -\frac{2}{\delta^3}\left[h_\theta^\alpha(0) + \delta\frac{d\ h_\theta^\alpha(\delta)}{d\ \delta}|_{\delta=0} + \frac{\delta^2}{2}\frac{d^2\ h_\theta^\alpha(\delta)}{d\ \delta^2}|_{\delta=0} + \frac{\delta^3}{6}h_\theta^{(3)\alpha}(\xi_1)\right] \\
&+ \frac{1}{\delta^2}\left[\frac{d\ h_\theta^\alpha(\delta)}{d\ \delta}|_{\delta=0} + \delta\frac{d^2\ h_\theta^\alpha(\delta)}{d\ \delta^2}|_{\delta=0} + \frac{\delta^2}{3}h_\theta^{(3)\alpha}(\xi_2)\right] \\
&= -\frac{1}{3}h_\theta^{(3)\alpha}(\xi_1) + \frac{1}{2}h_\theta^{(3)\alpha}(\xi_2).
\end{aligned}$$

□

Corollary 4.1. *If the third derivative of $D_X^\alpha(p_{\theta+\delta}, p_\theta)$ w.r.t δ is bounded in $B = I_0 \otimes I$, then $H_\theta^\alpha(\delta)$ is bounded in B.*

Corollary 4.2. *If the third derivatives of $D_X^\alpha(p_{\theta+\delta}, p_\theta)$ and $D_Y^\alpha(p_{\theta+\delta}, p_\theta)$ w.r.t δ are bounded in B, then $G_\theta^\alpha(\delta)$ is bounded in B, where*

$$G_\theta^\alpha(\delta) = \frac{1}{\delta^2}\left[D_X^\alpha(p_{\theta+\delta}, p_\theta) - D_Y^\alpha(q_{\theta+\delta}, q_\theta)\right].$$

Lemma 4.2. *Under the assumptions of Theorem 4.1, there exists $\delta_0 > 0$ such that for every $\theta \in I$ and every $\delta \leq \delta_0$*

$$G_\theta^\alpha(\delta) > \frac{d_1}{8},$$

where $d_1 = \frac{\alpha d_0}{2}$.

Proof. Let

$$G_\theta^\alpha(\delta) = \frac{1}{\delta^2}\left[D_X^\alpha(p_{\theta+\delta}, p_\theta) - D_Y^\alpha(q_{\theta+\delta}, q_\theta)\right].$$

By part (i) of Theorem 2.1,

$$\lim_{\delta\to 0} G_\theta^\alpha(\delta) = \frac{\alpha}{2}[I_X(\theta) - I_Y(\theta)] \geq \frac{d_1}{2} > 0 \tag{4.1}$$

To prove the lemma by contradiction, suppose it doesn't hold. Then we have

$$\forall\ \delta > 0\ \ \exists\ \theta \in I\ \ s.t.\ \ G_\theta^\alpha(\delta) \leq \frac{d_1}{8}.$$

Thus for some $\delta_1 > 0$ in I_0, there exist $\theta_1 \in I$ such that $G_{\theta_1}^\alpha(\delta_1) \leq d_1/8$.
But by (4.1)

$$\exists\ \gamma_1 > 0\ \ s.t.\ \ G_{\theta_1}^\alpha(\gamma) > \frac{d_1}{4},\ \ \forall\ 0 < \gamma < \gamma_1,$$

where (obviously) $\gamma_1 < \delta_1$. Now, let $\delta_2 = \frac{1}{2}\gamma_1$, again

$$\exists\ \theta_2 \in I\ \ s.t.\ \ G_{\theta_2}^\alpha(\delta_2) \leq \frac{d_1}{8},$$

and again by (4.1)

$$\exists\ \gamma_2 > 0\ \ s.t.\ \ G_{\theta_2}^\alpha(\gamma) > \frac{d_1}{4},\ \ \forall\ 0 < \gamma < \gamma_2,$$

where $\gamma_2 < \delta_2$. If we continue in this manner, we shall have three sequences $\delta_1, \delta_2, \ldots, \theta_1, \theta_2, \ldots$ and $\gamma_1, \gamma_2, \ldots$ such that for every $n \geq 1$, we have $\gamma_n < \delta_n$ and

$$G_{\theta_n}^\alpha(\gamma_n) - G_{\theta_n}^\alpha(\delta_n) > \frac{d_1}{4} - \frac{d_1}{8} = \frac{d_1}{8}.$$

But by mean value theorem we have

$$G_{\theta_n}^\alpha(\gamma_n) - G_{\theta_n}^\alpha(\delta_n) = (\gamma_n - \delta_n)(G_{\theta_n}^\alpha(\xi_n))' > \frac{d_1}{8},$$

where $\gamma_n < \xi_n < \delta_n$. We know that $(\gamma_n - \delta_n) \to 0$, as $n \to \infty$, so $(G_{\theta_n}^\alpha(\xi_n))' \to \infty$ as $n \to \infty$. But this contradicts the assumption that $(G_{\theta_n}^\alpha)'$ is bounded. Thus the lemma is proved. □

Now by noting that we can always restrict c in condition (ii) of the Theorem 4.1 to be less than unity, we have the following

Corollary 4.3. *Under the conditions of Theorem 4.1 there exist $\delta_0 > 0$ such that for every $\theta \in I$ and every $\delta \leq \delta_0$*

$$D_X^\alpha(p_{\theta+\delta}, p_\theta) - D_Y^\alpha(q_{\theta+\delta}, q_\theta) > \frac{d_1}{8}.$$

Lemma 4.3. *Under the assumptions of Theorem 4.1, there exist $\delta' > 0$ such that for every $\theta \in I$ and every $\delta \leq \delta'$*

$$-\frac{d_1}{32} < D_X^\alpha(p_\theta, p_{\theta_0}) - D_X^\alpha(p_{\theta+\delta'}, p_{\theta_0}) + D_X^\alpha(p_{\theta+\delta'}, p_\theta) < \frac{d_1}{32}$$

$$-\frac{d_1}{32} < D_Y^\alpha(q_\theta, q_{\theta_0}) - D_Y^\alpha(q_{\theta+\delta'}, q_{\theta_0}) + D_Y^\alpha(q_{\theta+\delta'}, q_\theta) < \frac{d_1}{32}$$

Proof. By considering the fact that the expressions between inequality signs tend to zero when $\delta' \to 0$, the proof is obvious. □

Proof of Theorem 4.1. Let $\delta^* = Min[\delta_0, \delta']$ and $K = [\frac{\theta_1 - \theta_0}{\delta^*}]$, where δ_0 and δ' are as in Lemma 4.2 and Lemma 4.3, respectively, and $[u]$ means the smallest integer greater than or equal to u. By Lemma 3.2 the inequality

$$D_X^\alpha(p_{\theta_0+k\delta^*}, p_{\theta_0}) - D_Y^\alpha(q_{\theta_0+k\delta^*}, q_{\theta_0}) > 0, \tag{4.2}$$

holds for $k = 1$. Aiming to give a proof by induction, we assume (4.2) holds for k. Let

$$Q = D_X^\alpha(p_{\theta_0+(k+1)\delta^*}, p_{\theta_0}) - D_Y^\alpha(q_{\theta_0+(k+1)\delta^*}, q_{\theta_0})$$

It is easy to show that $Q = Q' + Q''$, where

$$Q' = D_X^\alpha(p_{\theta_0+(k+1)\delta^*}, p_{\theta_0+k\delta^*}) - D_Y^\alpha(q_{\theta_0+(k+1)\delta^*}, q_{\theta_0+k\delta^*}),$$

and

$$\begin{aligned} Q'' &= D_X^\alpha(p_{\theta_0+(k+1)\delta^*}, p_{\theta_0}) - D_X^\alpha(p_{\theta_0+(k+1)\delta^*}, p_{\theta_0+k\delta^*}) \\ &+ D_Y^\alpha(q_{\theta_0+(k+1)\delta^*}, q_{\theta_0}) - D_Y^\alpha(q_{\theta_0+(k+1)\delta^*}, q_{\theta_0+k\delta^*}) \end{aligned}$$

By Lemma 4.2 $Q' > \frac{d_1}{8}$ and by Lemma 4.3

$$Q'' > D_X^\alpha(p_{\theta_0+k\delta^*}, p_{\theta_0}) - D_Y^\alpha(q_{\theta_0+k\delta^*}, q_{\theta_0}) - \frac{d_1}{16}.$$

So, $Q > (d_1/8) - (d_1/16) > 0$. Thus (4.2) holds for $k+1$. The result follows by induction. □

Theorem 4.1 and its reverse (which is trivial, in view of Theorem 2.1 part (*ii*)) states in effect that the partial ordering of families of distributions (under the stated regularity conditions) with respect to Renyi information is the same as that with respect to Fisher information. This knowledge can be useful when comparing experiments' potential evidence in terms of one or the other of the two indices. There is more about this point in our other article, also in the present volume, where we present some applications and examples.

References

Csiszar, I. (1995), Generalized cutoff rates and Renyi's information measures. IEEE Transactions on information theory, **41**, 26–34.

Dabak, A. G. and Johnson, D. H. (2002), Relation between Kullback-Leibler distance and Fisher information. http://cmc.rice.edu/docs/docs/Dab2002Sep1Relationsb.pdf.

Habibi, A. Arghmi, N. R., and Ahmadi, J. (2006), Statistical evidence in experiments and in record values. Communications in Statistics - Theory and Methods, **35**, 1971-1983.

Kullback, S. (1959), Information Theory and Statistics. New York: John Wiley.

Renyi, A. (1961), On measures of entropy and information. Proc. Fourth. Berkeley Symp. Math. Stat. Prob., 1960, Vol I, University of California Press, Berkeley, 547.