Otto-von-Guericke-University Magdeburg
Faculty of Economics and Management

Non-uniqueness of non-extensive entropy under Rényi's recipe.

Sönke Hoffmann

FEMM Working Paper No. 11, March 2008

F E M M

Faculty of Economics and Management Magdeburg

Working Paper Series

Otto-von-Guericke-University Magdeburg
Faculty of Economics and Management
P.O. Box 4120
39016 Magdeburg, Germany
http://www.ww.uni-magdeburg.de/

Non–uniqueness of non–extensive entropy under Rényi's recipe.

S. Hoffmann [a,*],

[a] *Otto-von-Guericke University Magdeburg, Dept. of Economic Policy (VWL III), Universitätsplatz 2, D-39106 Magdeburg, Germany*

Abstract

In this note I show that Tsallis entropy (Tsallis, 1988) is not unique in the class of non–additive, selfweighted and quasilinear means. A characterization is given which disproves a result in Dukkipati et al. (2005a,b) and Dukkipati et al. (2006)

Key words: Non-extensive entropy; Non-additivity; Quasilinear means

1 Introduction

In one of his seminal contributions to information theory Rényi (1961) defined information measures as a self-weighted quasilinear mean $\langle h \rangle_\phi = \phi^{-1}\left(\sum_{i=1}^n p_i \phi\left(h_i\right)\right)$, with elementary information $h_i = H\left(p_i\right),\ h = \left(h_i\right)_{i=1\ldots n}$ and $\phi(x)$ being a continuous and strictly monotonic function on a real interval defined at $x = 0$. From the perspective of information theory H should be log–additive ("lad") such that $H\left(p_i q_j\right) = H\left(p_i\right) + H\left(q_j\right)$ for independent events of the discrete $n-$outcome random variable at positive probability $p_i \in (0, 1]$. Log–additivity uniquely characterizes $H^{\text{lad}}\left(p_i\right) := c \ln\left(1/p_i\right)$ as "elementary information" in this domain (Aczél and Daróczy, 1975) and the famous Shannon- or Boltzman-Gibbs statistic can be written as a self-weighted quasilinear mean of all $H^{\text{lad}}\left(p_i\right)$ for linear ϕ, i.e. $V^{\text{S}}\left(p\right) := \left\langle h^{\text{lad}} \right\rangle_{\phi^{\text{lin}}} = \sum_{i=1}^n p_i h_i^{\text{lad}}$. From this starting point Rényi derived the class of all linear and non–linear ϕ which maintain log–additivity of $\left\langle h^{\text{lad}} \right\rangle_\phi$, such that for any two independent

* Corresponding author.
Email address: soenke.hoffmann@ww.uni-magdeburg.de (S. Hoffmann).

Preprint submitted to FEMM Working Paper Series *March 27, 2008*

distributions p and q and their direct product $p \times q$

$$\left\langle h^{\mathrm{lad}}_{p\times q}\right\rangle_{\phi} = \left\langle h^{\mathrm{lad}}_{p}\right\rangle_{\phi} + \left\langle h^{\mathrm{lad}}_{q}\right\rangle_{\phi}. \tag{1}$$

Rényi found that, given (1), generating functions must be either exponential or linear, which is tantamount to $V^{\mathrm{R}}_{\alpha}(p) := \left\langle h^{\mathrm{nad}}\right\rangle_{\phi^{\exp}} = \ln\left(\left(\sum_{i=1}^{n} p_i^{1-\alpha}\right)^{\frac{1}{\alpha}}\right)$ or $\lim_{\alpha\to 0} V^{\mathrm{R}}(p) = V^{\mathrm{S}}(p)$.

Similar to this information theoretical foundation of entropy measures non–extensive statistical mechanics employs elementary quantities being non–additive of some degree δ ("nad"), i.e. $H(p_i q_j) = H(p_i) + H(q_j) + \delta H(p_i) H(q_j)$ (Tsallis, 1988). Again, this property is charactrerizing.

Proposition 1 (Non–additivity of degree δ) *Let $x, y \in (0,\infty)$ and f be a continuous and non–constant function, then for constant $c \neq 0$*

$$f(xy) = f(x) + f(y) + \delta f(x) f(y) \Leftrightarrow f(x) = \ln_{\gamma}(x) := \begin{cases} \frac{x^{\gamma}-1}{c\gamma} & ; \gamma \neq 0 \\ \frac{1}{c}\ln(x) & ; \gamma = 0 \end{cases} \tag{2}$$

where $\delta = c\gamma$.

Proof of Proposition 1. For non–additivity of degree $\delta \neq 0$, $g(x) := \delta f(x) + 1$ gives the Cauchy power equation $g(xy) = g(x)g(y)$ which has the most general continuous solution $g(x) = x^{\gamma}$, nonconstant for $\gamma \neq 0$ (Aczél, 1966). Then for a constant $c = \delta/\gamma$ (2) follows by resubstitution. Clearly, non–additivity of degree $\delta = 0$ is log–additivity and then the most general non–constant solution is the common Napier logarithm $\lim_{\gamma\to 0} \ln_{\gamma}(x) = \frac{1}{c}\ln(x)$, where c defines its basis (e.g. $c = \ln(2)$ in information theory). ■

In analogy to elementary information h^{lad}_i and (1) we can define $h^{\mathrm{nad}}_i = H^{\mathrm{nad}}(p_i) := \ln_{\delta,\gamma}(1/p_i)$ and derive the self-weighted linear mean $V^{\mathrm{T}}(p) := \left\langle h^{\mathrm{nad}}\right\rangle_{\phi^{\mathrm{lin}}}$, which is known to be non–additive of degree δ, such that

$$\left\langle h^{\mathrm{nad}}_{p\times q}\right\rangle_{\phi} = \left\langle h^{\mathrm{nad}}_{p}\right\rangle_{\phi} + \left\langle h^{\mathrm{nad}}_{q}\right\rangle_{\phi} + \delta\left\langle h^{\mathrm{nad}}_{p}\right\rangle_{\phi}\left\langle h^{\mathrm{nad}}_{q}\right\rangle_{\phi}. \tag{3}$$

Means generated by ϕ_{lin} maintain log–additivity as well as non–additivity of degree δ, once these properties are satisfied by some elementary quantity h. Thus, it seems natural to ask whether a generalization of ϕ_{lin} similar to the one by Rényi can be undertaken to find the most general $\left\langle h^{\mathrm{nad}}\right\rangle_{\phi}$ satisfying (3). Dukkipati et al. (2005b) call this approach "Rényi's recipe" (applied to non–additivity) and suggest that there is no non–linear ϕ satisfying (3). In what follows I give a non–linear counterexample and an alternative characterization.

2 Characterization of non–additivitve quasilinear means

Definition 1 *Let the solution* $\ln_\gamma(x)$ *of (2) be called degree-deformed logarithm then for* $c \neq 0$

$$\exp_\gamma(z) := \ln_\gamma^{-1}(x) = \begin{cases} (c\gamma z+1)^{\frac{1}{\gamma}} & ;\gamma \neq 0 \\ \exp(cz) & ;\gamma = 0 \end{cases}$$

is the corresponding degree-deformed exponential.

Lemma 2 *Let* $\langle h \rangle_\phi$ *satisfy (1), then* $g(\langle h \rangle_\phi)$ *satisfies (3) for all* $c \neq 0$ *where*

$$g(x) = \ln_\gamma(\exp(x)) = \begin{cases} \frac{\exp(\gamma x)-1}{c\gamma} & ;\gamma \neq 0 \\ \frac{x}{c} & ;\gamma = 0 \end{cases}.$$

Proof of Lemma 2. Let $c, \gamma \neq 0$ then

$$\begin{aligned} \langle h_{p\times q} \rangle_\phi &= \langle h_p \rangle_\phi + \langle h_q \rangle_\phi \qquad (4) \\ &\Leftrightarrow \exp\left(\gamma \langle h_{p\times q} \rangle_\phi\right) = \exp\left(\gamma \langle h_p \rangle_\phi\right) \exp\left(\gamma \langle h_q \rangle_\phi\right) \\ &\Leftrightarrow \exp\left(\gamma \langle h_{p\times q} \rangle_\phi\right) = \exp\left(\gamma \langle h_p \rangle_\phi\right) + \exp\left(\gamma \langle h_q \rangle_\phi\right) \\ &+ \left[\exp\left(\gamma \langle h_p \rangle_\phi\right) - 1\right] \left[\exp\left(\gamma \langle h_q \rangle_\phi\right) - 1\right] - 1 \\ &\Leftrightarrow \frac{\exp\left(\gamma \langle h_{p\times q} \rangle_\phi\right) - 1}{c\gamma} = \frac{\exp\left(\gamma \langle h_p \rangle_\phi\right) - 1}{c\gamma} + \frac{\exp\left(\gamma \langle h_q \rangle_\phi\right) - 1}{c\gamma} \\ &+ c\gamma \frac{\left[\exp\left(\gamma \langle h_p \rangle_\phi\right) - 1\right]}{c\gamma} \cdot \frac{\left[\exp\left(\gamma \langle h_q \rangle_\phi\right) - 1\right]}{c\gamma} \\ &\Leftrightarrow g\left(\langle h_{p\times q} \rangle_\phi\right) = g\left(\langle h_p \rangle_\phi\right) + g\left(\langle h_q \rangle_\phi\right) + \delta g\left(\langle h_p \rangle_\phi\right) g\left(\langle h_q \rangle_\phi\right), \qquad (5) \end{aligned}$$

which is the non–additivity (degree δ) condition for the function $g(\langle h \rangle_\phi)$. In the $\gamma \to 0$ limit g is linear and (5) reduces back to (4). ■

Proposition 3 *Let* $\hat{\phi}(x) := \ln\left(\exp_\gamma(x)\right)$ *then* $\left\langle h^{nad} \right\rangle_{\hat{\phi}}$ *satisfies (3) for all* γ.

Proof of Proposition 3. As $V^{\mathrm{S}}(p)$ is known to satisfy (1), $\left\langle h^{\mathrm{nad}} \right\rangle_{\hat{\phi}} = \frac{1}{c\gamma}\left(\exp\left(\gamma V^{\mathrm{S}}(p)\right) - 1\right) = \ln_\gamma\left(\exp\left(V^{\mathrm{S}}(p)\right)\right)$ must satisfy (3) due to Lemma 2. ■

Now we want to find the most general set of functions satisfying (3) under “Rényi’s recipe”. To this end the following classical Lemma is essential.

Lemma 4 (Hardy et al. (1934)) *Let $a \neq 0$ and b be constants then*

$$\phi'(x) = a\phi(x) + b \Leftrightarrow \langle h\rangle_{\phi'} = \langle h\rangle_{\phi}. \tag{6}$$

Proposition 5 (Non–additivity–preserving means) *Let $H(p)$ be non–additive of degree δ, then $\langle H(p)\rangle_{\phi}$ is non–additive of degree δ iff $\phi(x) = \tilde{\phi}(x) = a\phi^*(x) + b$, $a \neq 0$ where*

$$\phi^*(x) = \begin{cases} (c\gamma x+1)^{\frac{\alpha}{\gamma}} & ;\alpha \neq \gamma \neq 0 \\ \exp(c\alpha x) & ;\alpha \neq 0; \gamma = 0 \\ \ln(c\gamma x + 1) & ;\alpha = 0; \gamma \neq 0 \\ x & ;\alpha = \gamma \neq 0 \end{cases}. \tag{7}$$

Proof of Proposition 5. By Proposition 1 $H(p) = H^{\mathrm{T}}(p)$, which will be written H^{T} for convenience. First let $\alpha \neq \gamma \neq 0$ in (7) then $\left\langle H^{\mathrm{T}}\right\rangle_{\phi^*} = \frac{1}{\delta}\left(\left(\sum_{i=1}^{n} p_i^{1-a}\right)^{\frac{\gamma}{\alpha}} - 1\right) = \ln_\gamma\left(\exp\left(V_\alpha^{\mathrm{R}}(p)\right)\right)$. $V_\alpha^{\mathrm{R}}(p)$ is known to satisfy (1) for all real α (Rényi, 1961) thus $\left\langle H^{\mathrm{T}}\right\rangle_{\phi^*}$ must satisfy (3) for all α and γ due to Lemma 2.

Vice versa, let $\left\langle H^{\mathrm{T}}\right\rangle_{\phi}$ be non–additive of degree δ then (3) must hold. Now define $\psi(x) := \phi(\ln_\gamma(x))$ viz. $\phi^{-1}(z) = \ln_\gamma\left(\psi^{-1}(z)\right)$, $q_j = \frac{1}{m}$ for all $j = 1 \dots m$ and $p^{-1} := \left(p_i^{-1}\right)_{i=1\dots n}$ then (3) becomes

$$\ln_\gamma\left(\left\langle mp^{-1}\right\rangle_{\psi}\right) = \ln_\gamma\left(\left\langle p^{-1}\right\rangle_{\psi}\right) + \ln_\gamma(m) + c\gamma\ln_\gamma(m)\ln_\gamma\left(\left\langle p^{-1}\right\rangle_{\psi}\right)$$
$$\Leftrightarrow \left\langle mp^{-1}\right\rangle_{\psi} = m\left\langle p^{-1}\right\rangle_{\psi}. \tag{8}$$

In order to find out which ψ satisfy (8) we will use a slightly different notation of ψ. Let $\check{\psi}(x) = \psi\left(\frac{m}{x}\right)$ and $\tilde{\psi}(x) = \psi\left(\frac{1}{x}\right)$ then (8) is equivalent to $\langle p\rangle_{\check{\psi}} = \langle p\rangle_{\tilde{\psi}}$ which holds due to (6) iff $\check{\psi}$ and $\tilde{\psi}$ are affine maps of each other, such that for constants a and b (being, however, different for different m)

$$\check{\psi}(x) = \psi\left(\frac{m}{x}\right) = a(m)\psi\left(\frac{1}{x}\right) + b(m). \tag{9}$$

Note that the role of x as a variable and the one of m as a constant can be interchanged from the beginning of the proof without changing the solutions of (9). Neither would the assumption $\psi(0) = 0 \Rightarrow b(m) = \psi(m)$ do, as $\psi(0)$ is defined and we can, by (6), transform ψ linearly without changing the mean generated by ψ. Then, for $t = x^{-1}$ (9) can be rewritten as

$$\psi(tm) = a(m)\psi(t) + \psi(m) = a(t)\psi(m) + \psi(t)$$
$$\beta := \frac{a(m)-1}{\psi(m)} = \frac{a(t)-1}{\psi(t)}$$
$$\Leftrightarrow a(m) = \psi(m)\beta + 1 \tag{10}$$

with β being a constant. Substituting (10) into (9) one obtains the functional equation $\psi(tm) = \psi(t) + \psi(m) + \beta\psi(t)\psi(m)$, which has by Proposition 1 the most general non–constant and continuous solution

$$\psi(x) = \ln_\alpha(x) := \begin{cases} \frac{x^\alpha - 1}{c\alpha} & ; \alpha \neq 0 \\ \frac{1}{c}\ln(x) & ; a = 0 \end{cases}$$

Then, recalling that $\phi(y) = \psi(x) = \psi\left(\exp_\gamma(y)\right)$ we have

$$\phi(y) = \ln_\alpha(\exp_\gamma(y)) = \begin{cases} \frac{(c\gamma y+1)^{\frac{\alpha}{\gamma}}-1}{c\alpha} & ; \alpha \neq \gamma \neq 0 \\ \frac{\exp(c\alpha y)-1}{c\alpha} & ; \alpha \neq 0; \gamma = 0 \\ \frac{\ln(c\gamma y+1)}{c\gamma} & ; \alpha = 0; \gamma \neq 0 \\ y & ; \alpha = \gamma \end{cases}. \tag{11}$$

Finally, applying (6) gives the solution (7). Note that the $\alpha \neq 0; \gamma = 0$ case recovers non–additivity of degree zero, i.e. log–additivity. ■

References

Aczél, J., 1966. Lectures on Functional Equations and Their Applications. Academic Press, New York.

Aczél, J., Daróczy, Z., 1975. On measures of information and their characterizations. Academic Press, San Diego.

Dukkipati, A., et al., 2005a. Nongeneralizability of Tsallis entropy by means of Kolmogorov-Nagumo averages under pseudo-additivity. Working Paper, available at: arXIV:math-ph/0505078.

Dukkipati, A., et al., 2005b. Uniqueness of nonextensive entropy under Rényis recipe. Working Paper, available at: arXIV:cs.IT/0511078.

Dukkipati, A., et al., 2006. On generalized measures of information with maximum and minimum entropy prescriptions. Ph.D. thesis, Computer Science and Automation Indian Institute of Science, Bangalore, India.

Hardy, G., et al., 1934. Inequalities, 2nd Edition. Paperback Edition, reprinted 1999. Cambridge University Press, Cambridge.

Rényi, A., 1961. On measures of entropy and information. In: Neyman, J. (Ed.), Proceedings of the Fourth Berkeley Symposium on Mathematical

Statistics and Probability. Vol. 1. University of California Press, Berkeley, pp. 547–561.
Tsallis, C., 1988. Possible generalization of Boltzmann-Gibbs statistics. Journal of Statistical Physics 52, 479–487.