Concavity of some entropies

Ioan RAŞA
Department of Mathematics, Technical University of Cluj-Napoca, Romania
Email address: Ioan.Rasa@math.utcluj.ro

Abstract
It is well-known that the Shannon entropies of some parameterized probability distributions are concave functions with respect to the parameter. In this paper we consider a family of such distributions (including the binomial, Poisson, and negative binomial distributions) and investigate the concavity of the Shannon, Rényi, and Tsallis entropies of them.

1 Introduction
Let $c \in \mathbb{R}$, $I_c := [0, -\frac{1}{c}]$ if $c < 0$, and $I_c := [0, +\infty)$ if $c \geq 0$.

For $\alpha \in \mathbb{R}$ and $k \in \mathbb{N}_0$ the binomial coefficients are defined as usual by
\[
\binom{\alpha}{k} := \frac{\alpha(\alpha-1)\ldots(\alpha-k+1)}{k!} \quad \text{if } k \in \mathbb{N}, \quad \text{and } \binom{\alpha}{0} := 1.
\]

Let $n > 0$ be a real number such that $n > c$ if $c \geq 0$, or $n = -cl$ with some $l \in \mathbb{N}$ if $c < 0$.

For $k \in \mathbb{N}_0$ and $x \in I_c$ define
\[
p_{n,k}^c(x) := (-1)^k \binom{-n}{k} (cx)^k (1 + cx)^{-\frac{n}{c} - k}, \quad \text{if } c \neq 0,
\]
\[
p_{n,k}^0(x) := \lim_{c \to 0} p_{n,k}^c(x) = \frac{(nx)^k}{k!} e^{-nx}.
\]

Details and historical notes concerning these functions can be found in [3], [7], [21] and the references therein. In particular,
\[
\frac{d}{dx} p_{n,k}^c(x) = n \left( p_{n+c,k-1}^c(x) - p_{n+c,k}^c(x) \right).
\]

Moreover, $\sum_{k=0}^{\infty} p_{n,k}^c(x) = 1$, so that $\left( p_{n,k}^c(x) \right)_{k \geq 0}$ is a parameterized probability distribution. Its associated Shannon entropy is
\[
H_{n,c}(x) := -\sum_{k=0}^{\infty} p_{n,k}^c(x) \log p_{n,k}^c(x),
\]
while the Rényi entropy of order 2 and the Tsallis entropy of order 2 are given, respectively, by (see [18], [20])

\[ R_{n,c}(x) := -\log S_{n,c}(x); \quad T_{n,c}(x) := 1 - S_{n,c}(x), \]

where

\[ S_{n,c}(x) := \sum_{k=0}^{\infty} \left( p_{n,k}^{|c|}(x) \right)^2, \quad x \in I_c. \]

The cases \( c = -1, c = 0, c = 1 \) correspond, respectively, to the binomial, Poisson, and negative binomial distributions.

2 Shannon entropy

\( H_{n,-1} \) is a concave function; this is a special case of the results of [13]; see also [6], [8], [9] and the references therein. \( H_{n,0} \) is also concave; moreover, \( H'_{n,0} \) is completely monotonic (see, e.g., [2, p. 2305]).

For the sake of completeness we present here the proof for the concavity of \( H_{n,c}, c \in \mathbb{R} \). Let us consider separately the cases \( c \geq 0 \) and \( c < 0 \).

**Theorem 2.1.** For \( c \geq 0 \), \( H_{n,c} \) is concave and increasing on \([0, +\infty)\).

**Proof.** Using (1.1), it is a matter of calculus to prove that

\[ H'_{n,c}(x) = n \sum_{k=0}^{\infty} p_{n+k,k}^{|c|}(x) \log \frac{p_{n+k}^{|c|}(x)}{p_{n+k+1}^{|c|}(x)}, \]

which leads to

\[ H'_{n,c}(x) = n \left( \log \frac{1 + cx}{x} + \sum_{k=0}^{\infty} p_{n+k,k}^{|c|}(x) \log \frac{k+1}{n+ck} \right), \quad (2.1) \]

and therefore to

\[ H''_{n,c}(x) = -\frac{n}{x(1 + cx)} + n(n+c) \sum_{k=0}^{\infty} p_{n+k+2,k}^{|c|}(x) \log \frac{(k+2)(n+ck)}{(k+1)(n+ck+c)}. \]

It follows that

\[ H''_{n,c}(x) > -\frac{n}{x(1 + cx)}, \quad x > 0. \quad (2.2) \]

Since \( \log t < t - 1, t > 1 \), we have also

\[ H''_{n,c}(x) < -\frac{n}{x(1 + cx)} + n(n+c) \sum_{k=0}^{\infty} p_{n+k+2,k}^{|c|}(x) \frac{n-c}{(k+1)(n+ck+c)} \]

\[ = -\frac{n}{x(1 + cx)} \left( \frac{c}{n} + \left( 1 - \frac{c}{n} \right) (1 + cx)^{-\frac{n}{c}} \right) < 0, \]

so that \( H_{n,c} \) is concave on \([0, +\infty)\); being positive, it is also increasing on \([0, +\infty)\). \( \square \)
Remark 2.2. From (2.2) it follows that the functions $H_{n,0}(x) + nx \log x$ and $H_{n,c}(x) + \frac{c}{e}(cx \log x - (1 + cx) \log(1 + cx))$, $c > 0$, are convex on $[0, +\infty)$.

Remark 2.3. The following inequalities are valid for $x > 0$ and $c \geq 0$:

$$\log \frac{x}{1+cx} \leq \sum_{k=0}^{\infty} \frac{[c]^{k+1}}{ck+n} \log \frac{k+1}{ck+n} \leq \log \frac{x+\frac{1}{c}}{1+cx}. \quad (2.3)$$

The first one follows from $H_{n,0}'(x) > 0$, taking into account (2.1), and the second is a consequence of Jensen’s inequality applied to the concave function $\log x$. In particular, for $c = 0$ and $n = 1$ we get:

$$\log x \leq \sum_{k=0}^{\infty} e^{-\frac{x}{k+1}} \log(k+1) \leq \log(x+1), \quad x > 0.$$

The case $c < 0$ can be studied with the same method as in Theorem 2.1, but we present here a different approach, based on an integral representation from [10].

Theorem 2.4. For $c < 0$, $H_{n,c}$ is concave on $[0, -\frac{1}{2c}]$, increasing on $[0, -\frac{1}{2c}]$, and decreasing on $[-\frac{1}{2c}, -\frac{1}{c}]$.

Proof. For $c < 0$ we have $n = -cl$ with $l \in \mathbb{N}$. Using [10 (2.5)] we get

$$H_{n,c}(x) = H_{l,-1}(cx) = -l \left[-(cx) \log(-cx) + (1 + cx) \log(1 + cx)\right] + \int_{0}^{1} \frac{s-1}{\log(s)} \left[(1+cx-ces)^l + ((1+cx)s-cx)^l - 1 - s^l\right] ds.$$

It is matter of calculus to prove that

$$H_{n,c}(x) = -\frac{n}{x(1+cx)} + e^2 l(l-1) \int_{0}^{1} \frac{s-1}{\log s} \left[(1+cx-ces)^l + ((1+cx)s-cx)^l - 1 - s^l\right] ds. \quad (2.4)$$

For $0 < s < 1$ we have $0 < \frac{s-1}{\log s} < 1$; moreover

$$\int_{0}^{1} \left[(1+cx-ces)^l + ((1+cx)s-cx)^l - 1 - s^l\right] ds = \frac{1}{l-1} \left[\frac{1 - (cx)^{l-1}}{1 + cx} - \frac{1 - (1 + cx)^{l-1}}{cx}\right].$$

Summing up, we get

$$H''_{n,c}(x) < -n \frac{(1+cx)^{-\frac{2}{l}}} {x(1+cx)} + \frac{(-cx)^{-\frac{2}{l}}} {1 + cx} < 0, \quad 0 < x < -\frac{1}{c}.$$  

Consequently, $H_{n,c}$ is concave on $[0, -\frac{1}{2c}]$. Since

$$H_{n,c} \left(-\frac{1}{2c} - t\right) = H_{n,c} \left(-\frac{1}{2c} + t\right), \quad t \in \left[0, -\frac{1}{2c}\right],$$

we conclude that $H_{n,c}$ is increasing on $[0, -\frac{1}{2c}]$ and decreasing on $[-\frac{1}{2c}, -\frac{1}{c}].$
Remark 2.5. Let $c < 0$ and $l \geq 2$. From (2.4) it follows that $H''_{n,c}(x) > -\frac{n}{x(1+cx)}$, and so the function

$$H_{n,c}(x) + \frac{n}{c}[cx \log x - (1 + cx) \log(1 + cx)]$$

is convex on $[0, -\frac{1}{c}]$.

Remark 2.6. For $c = -1$, the method used to prove (2.3) leads to

$$\log \frac{x}{1-x} < \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \frac{k+1}{n+1-k}$$

$$< \log \left( \frac{x}{1-x} + \frac{1 - (n+2)x^{n+1}}{(n+1)(1-x)} \right), \quad 0 < x < \frac{1}{2}.$$

3 $S_{n,c}$ and Heun functions

The following conjecture was formulated in [13]:

Conjecture 3.1. $S_{n,-1}$ is convex on $[0, 1]$.

Th. Neuschel [11] proved that $S_{n,-1}$ is decreasing on $[0, \frac{1}{2}]$ and increasing on $[\frac{1}{2}, 1]$. The conjecture and the result of Neuschel can be found also in [5].

A proof of the conjecture was given by G. Nikolov [12], who related it with some new inequalities involving Legendre polynomials. Another proof can be found in [4].

Using the important results of Elena Berdysheva [3], the following extension was obtained in [17]:

Theorem 3.2. ([17, Theorem 9]). For $c < 0$, $S_{n,c}$ is convex on $[0, -\frac{1}{c}]$.

A stronger conjecture was formulated in [14] and [17]:

Conjecture 3.3. For $c \in \mathbb{R}$, $S_{n,c}$ is logarithmically convex, i.e., $\log S_{n,c}$ is convex.

It was validated for $c \geq 0$ by U. Abel, W. Gawronski and Th. Neuschel [1], who proved a stronger result:

Theorem 3.4. ([1]). For $c \geq 0$, the function $S_{n,c}$ is completely monotonic, i.e.,

$$(-1)^m \left( \frac{d}{dx} \right)^m S_{n,c}(x) > 0, \quad x \geq 0, m \geq 0.$$

Consequently, for $c \geq 0$, $S_{n,c}$ is logarithmically convex, and hence convex.

On the other hand, according to [17, Th. 4], $S_{n,c}$ is a solution to the differential equation

$$x(1+cx)(1+2cx)y''(x) + (4(n+c)x(1+cx)+1)y'(x) + 2n(1+2cx)y(x) = 0.$$  \hspace{1cm} (3.1)
Consequently, for $c \neq 0$ the function $S_{n,c}(-\frac{r}{c})$ is a solution to the Heun equation

$$y''(x) + \left(\frac{1}{x} + \frac{1}{x-1} + \frac{2n}{x-\frac{1}{2}}\right)y'(x) + \frac{2nx - n}{x(x-1)(x-\frac{1}{2})}y(x) = 0,$$

and $S_{n,0}$ is a solution of the confluent Heun equation:

$$u''(x) + \left(4n + \frac{1}{x}\right)u'(x) + \frac{2nx - 2n}{x(x-1)}u(x) = 0.$$

For details, see [14]-[17].

4 Rényi entropy and Tsallis entropy

Theorem 4.1. (i) For $c \geq 0$, $R_{n,c}$ is concave and increasing on $[0, +\infty)$. 

(ii) $R_{n,c}'$, with $c \in \mathbb{R}$, is a solution to the Riccati equation

$$x(1 + cx)(1 + 2cx)u'(x) = x(1 + cx)(1 + 2cx)u^2(x) - (4(n + c)x(1 + cx) + 1)u(x) + 2n(1 + 2cx).$$

Proof. i) is a direct consequence of Theorem 3.4.

ii) We have $S_{n,c} = \exp(-R_{n,c})$ and (3.1) yields

$$x(1 + cx)(1 + 2cx)\left((R_{n,c}')^2 - R_{n,c}''\right) - (4(n + c)x(1 + cx) + 1)R_{n,c}' + 2n(1 + 2cx) = 0.$$

Setting $u = R_{n,c}'$, we conclude the proof.

Remark 4.2. As far as we know, Conjecture 3.3 is still open for $c < 0$, so that the concavity of $R_{n,c}$, $c < 0$, remains to be investigated.

Theorem 4.3. (i) $T_{n,c}$ is concave. For $c \geq 0$ it is increasing on $[0, +\infty)$. For $c < 0$ it is increasing on $[0, -\frac{1}{2c}]$ and decreasing on $[-\frac{1}{2c}, -\frac{1}{c}]$.

(ii) $T_{n,c}$ is a solution to the equation

$$x(1 + cx)(1 + 2cx)u''(x) + (4(n + c)x(1 + cx) + 1)u'(x) + 2n(1 + 2cx)u(x) = 2n(1 + 2cx).$$

Proof. (i) is a consequence of Theorems 3.2 and 3.4 while (ii) follows from (3.1).
5 Some inequalities

a) The explicit expression of $S_{n-1}, n \in \mathbb{N}$, is

$$S_{n-1}(x) = \sum_{k=0}^{n} \left( \binom{n}{k} x^k (1-x)^{n-k} \right)^2, \quad x \in [0,1].$$

Consider also the function

$$f_n(t) := \sum_{j=0}^{n-1} \binom{n-1}{j} \left( \binom{n}{j+1} t^{2j+1} - \binom{n}{j} t^{2j} \right), \quad t \geq 1.$$

Since $S_{n-1}(1-x) = S_{n-1}(x)$, it follows that

$$S_{n-1}^{(2j+1)} \left( \frac{1}{2} \right) = 0, \quad j = 0, 1, \ldots, n-1.$$

In relation with Conjecture 3.1, it was also conjectured in [13] that

$$S_{n-1}^{(2j)} \left( \frac{1}{2} \right) > 0, \quad j = 0, 1, \ldots, n, \quad (5.1)$$

$$f_n^{(i)}(t) \geq 0, \quad t \geq 1; \quad i = 0, 1, \ldots, 2n-1, \quad (5.2)$$

$$\sum_{j=[i/2]}^{n-1} \binom{n-1}{j} \left( \binom{n}{j+1} \binom{2j+1}{i} - \binom{n}{j} \binom{2j}{i} \right) \geq 0, \quad i = 0, 1, \ldots, 2n-1. \quad (5.3)$$

We shall prove here these inequalities\(^1\).

It can be proved directly that

$$S_{n-1} \left( \frac{1}{2} \right) = \frac{1}{4^n} \binom{2n}{n}; \quad S_{n-1}^{(2)} \left( \frac{1}{2} \right) = \frac{1}{4^{n-2}} \binom{2n-2}{n-1};$$

$$S_{n-1}^{(4)} \left( \frac{1}{2} \right) = \frac{9}{4^{n-4}} \binom{2n-4}{n-2}.$$

The following formula was obtained in [4]:

$$S_{n-1}(x) = \sum_{k=0}^{n} 4^{k-n} \binom{2n}{n} \left( \binom{n}{k} \right)^2 \left( \frac{2n}{2k} \right)^{-1} \left( x - \frac{1}{2} \right)^{2k}. \quad (5.4)$$

\(^1\)For (5.2) with $i = 0$, see also http://www.artofproblemsolving.com/community/c6h494060p2772738.
Using it we get
\[ S_{n-1}^{(2j)} \left( \frac{1}{2} \right) = (2j)!4^{j-n} \binom{2n}{n} \left( \frac{n}{j} \right)^2 \binom{2n}{2j}^{-1}, \quad j = 0, 1, \ldots, n, \]
and so (5.1) is proved.

On the other hand, it is not difficult to prove that
\[ f_n(t) = \frac{1}{2n} (t + 1)^{2n-1} S'_{n-1} \left( \frac{t}{t+1} \right), \quad t \geq 1. \quad (5.5) \]

From (5.4) we obtain
\[ S'_{n-1}(x) = \sum_{k=1}^{n} 2ka_{nk} \left( x - \frac{1}{2} \right)^{2k-1} \quad (5.6) \]
for certain \( a_{nk} > 0 \). Now (5.5) and (5.6) imply
\[ f_n(t) = \sum_{k=1}^{n} c_{nk}(t - 1)^{2k-1}(t + 1)^{2n-2k}, \quad t \geq 1, \]
with suitable \( c_{nk} > 0 \), and using Leibniz’ formula we get (5.2).

Finally, starting from the definition of \( f_n(t) \), it is not difficult to infer that
\[ f_n^{(i)}(1) = i! \sum_{j=\lceil i/2 \rceil}^{n-1} \binom{n-1}{j} \left( \binom{n}{j+1} \left( \binom{2j+1}{i} \right) - \binom{n}{j} \left( \binom{2j}{i} \right) \right). \]

Combined with (5.2), this proves (5.3).

b) Let \( B_m : C[0,1] \rightarrow C[0,1], B_m f(x) = \sum_{k=0}^{m} \binom{m}{k} x^k (1-x)^{m-k} \), be the classical Bernstein operators. It is well-known that if \( f \in C[0,1] \) is convex, then \( B_m f \) is also convex.

Now let \( w_n \in C[0,1] \) be piece-wise linear, such that \( w_n \left( \frac{2k-1}{2n} \right) = 0, \quad k = 1, \ldots, n; \ w_n \left( \frac{2k}{2n} \right) = \left( \binom{n}{k} \right)^2 / \left( \binom{2n}{2k} \right), \quad k = 0, 1, \ldots, n. \) Then \( B_{2n} w_n = S_{n-1} \), hence \( B_{2n} w_n \) is convex although the graph of \( w_n \) is “like a saw”.

References

[1] U. Abel, W. Gawronski, Th. Neuschel, Complete monotonicity and zeros of sums of squared Baskakov functions, Appl. Math. Comput. 258 (2015), 130-137.
[2] J.A. Adell, A. Lekuona and Y. Yu, Sharp bounds on the entropy of the Poisson Law and related quantities, IEEE Trans. Information Theory 56 (2010), 2299-2306.
[3] E. Berdysheva, Studying Baskakov-Durrmeyer operators and quasi-interpolants via special functions, J. Approx. Theory 149 (2007), 131-150.
[4] I. Gavrea, M. Ivan, On a conjecture concerning the sum of the squared Bernstein polynomials, Appl. Math. Comput. 241 (2014), 70-74.

[5] H. Gonska, I. Raşa, M.-D. Rusu, Chebyshev-Grüss-type inequalities via discrete oscillations, Bul. Acad. Stiinte Repub. Mold. Mat. no. 1 (74) (2014), 63-89; arxiv 1401.7908 [math.CA].

[6] P. Harremoës, Binomial and Poisson distributions as maximum entropy distributions, IEEE Trans. Information Theory 47 (2001), 2039 - 2041.

[7] M. Heilmann, Erhöhung der Konvergenzgeschwindigkeit bei der Approximation von Funktionen mit Hilfe von Linearkombinationen spezieller positiver linearer Operatoren, Habilitationsschrift, Universität Dortmund (1992).

[8] E. Hillion, Concavity of entropy along binomial convolutions, Electron. Commun. Probab. 17 (2012), 1-9.

[9] E. Hillon, O. Johnson, A proof of the Shepp-Olkin entropy concavity conjecture, arXiv: 1503.01570v1, 5 Mar 2015.

[10] C. Knessl, Integral representations and asymptotic expansions for Shannon and Rényi entropies, Appl. Math. Lett. 11 (1998), 69-74.

[11] Th. Neuschel, Unpublished manuscript (2012).

[12] G. Nikolov, Inequalities for ultraspherical polynomials. Proof of a conjecture of I. Raşa, J. Math. Anal. Appl. 418 (2014), 852-860.

[13] I. Raşa, Unpublished manuscripts (2012).

[14] I. Raşa, Special functions associated with positive linear operators, arxiv: 1409.1015v2.

[15] I. Raşa, Rényi entropy and Tsallis entropy associated with positive linear operators, arxiv: 1412.4971v1.

[16] I. Raşa, Entropies and the derivatives of some Heun functions, arxiv: 1502.05570v1.

[17] I. Raşa, Entropies and Heun functions associated with positive linear operators, Appl. Math. Comput. 268 (2015), 422-431.

[18] A. Rényi, On measures of entropy and information, Proc. Fourth Berkeley Symp. Math. Statist. Prob., Vol. 1, Univ. of California Press (1961), pp. 547-561.

[19] L.A. Shepp, I. Olkin, Entropy of the sum of independent Bernoulli random variables and of the multinomial distribution, in: Proc. Contributions to Probability, New York (1981), pp. 201-206.

[20] C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, J. Stat. Phys. 52 (1988), 479-487.

[21] M. Wagner, Quasi-Interpolaten zu geminen Baskakov-Durrmeyer-Typ Operatoren, Shaker Verlag, Aachen (2013).