

Entropies associated to q -distributions of probability

ADINA ELENA BĂRAR
(CLUJ-NAPOCA)

DANIELA IOANA INOAN
(CLUJ-NAPOCA)

IOAN RAȘA
(CLUJ-NAPOCA)

ABSTRACT. We consider the distributions of probability associated with the q -Bernstein, q -Szász-Mirakyan, q -Meyer-König and Zeller operators. The paper is devoted to investigating convexity-like properties of the Shannon, Rényi and Tsallis entropies of these distributions in the limiting case when $q \rightarrow 0$.

KEY WORDS: entropy, geometric distribution, shape properties.

MSC 2010: 94A17, 26A51

1 Introduction

Consider a probability distribution depending on a parameter $x \in I$, where I is an interval of the real axis:

$$p(x) = (p_0(x), p_1(x), \dots), \text{ with } p_j(x) \geq 0, \sum_j p_j(x) = 1, \text{ for } x \in I.$$

[◇]Adina Elena Bărar, Technical University of Cluj-Napoca, Romania,
email: ellena_sontica@yahoo.com

[◇]Daniela Ioana Inoan, Technical University of Cluj-Napoca, Romania,
email: daniela.inoan@math.utcluj.ro

[◇]Ioan Rașa, Technical University of Cluj-Napoca, Romania,
email: ioan.rasa@math.utcluj.ro

Related to it, several notions of entropies can be formulated (see for instance [6], [2], [5], [7]). We will be concerned with the following:

$$H = - \sum_j p_j(x) \log p_j(x), \text{ called the Shannon entropy,}$$

$$R = - \log \sum_j p_j^2(x), \text{ called the Rényi entropy,}$$

$$T = 1 - \sum_j p_j^2(x), \text{ called Tsallis entropy.}$$

The sum of squares $S = \sum_j p_j^2(x)$, appears in literature under the name index of coincidence.

In the theory of positive linear operators, various generalizations of Bernstein polynomials are used. For $q \in (0, 1]$, the q -Bernstein basis of order n is given by the polynomials

$$(1.1) \quad \begin{bmatrix} n \\ r \end{bmatrix}_q x^r \prod_{s=0}^{n-r-1} (1 - q^s x), \quad r = 0, 1, \dots, n, \quad x \in [0, 1].$$

where

$$\begin{aligned} \begin{bmatrix} n \\ 0 \end{bmatrix}_q &:= 1, \quad \begin{bmatrix} n \\ 1 \end{bmatrix}_q := 1 + q + \dots + q^{n-1}, \\ \begin{bmatrix} n \\ k \end{bmatrix}_q &:= \frac{(1 + q + \dots + q^{n-1})(1 + q + \dots + q^{n-2}) \dots (1 + \dots + q^{n-k})}{1(1 + q)(1 + q + q^2) \dots (1 + q + \dots + q^{k-1})}. \end{aligned}$$

The polynomials (1.1) form also a probability distribution. In the limit case $q = 0$, the distribution becomes

$$(1.2) \quad p_n(x) = (1 - x, x(1 - x), x^2(1 - x), \dots, x^{n-1}(1 - x), x^n), \quad x \in [0, 1]$$

Related to p_n is the geometric probability distribution with the parameter $1 - x$:

$$(1.3) \quad p(x) = (1 - x, x(1 - x), x^2(1 - x), \dots), \quad x \in [0, 1].$$

This is also the limiting case, when $q \rightarrow 0$, of the probability distributions associated with the q -Szász-Mirakjan operators (see [1], p. 23, (2.21)) and q -Meyer-König and Zeller operators (see [1], pp. 8-9, (1.10)).

In this paper we wish to study various entropies associated to (1.2) and (1.3). We are interested especially in the convexity, monotonicity of the functions and some limit properties. This kind of properties were investigated for a general family of distributions (including the binomial, Poisson, and negative binomial distributions) in [3], [4] and the references therein.

2 Shape properties of entropies

By direct computation one can obtain the index of coincidence and the entropies for $p_n(x)$:

$$S_n(x) = \frac{2x^{2n+1} - x + 1}{x + 1} = 2(1 - x + x^2 - \dots - x^{2n-1} + x^{2n}) - 1$$

$$R_n(x) = \log(1 + x) - \log(2x^{2n+1} - x + 1)$$

$$T_n(x) = \frac{2x}{1+x}(1 - x^{2n})$$

$$H_n(x) = \begin{cases} -\frac{1-x^n}{1-x} \left[(1-x) \log(1-x) + x \log x \right], & x \in (0, 1) \\ 0, & x \in \{0, 1\} \end{cases}$$

and for $p(x)$ respectively

$$S(x) = \frac{1-x}{1+x}$$

$$R(x) = -\log \frac{1-x}{1+x}$$

$$T(x) = \frac{2x}{1+x}$$

$$H(x) = -\frac{x}{1-x} \log x - \log(1-x)$$

To study the function $S_n(x)$ we calculate the first and second order derivatives:

$$S'_n(x) = 2 \frac{2nx^{2n+1} + (2n+1)x^{2n} - 1}{(x+1)^2}$$

$$S''_n(x) = 4 \frac{(2n^2 - n)x^{2n+1} + (4n^2 - 1)x^{2n} + (2n^2 + n)x^{2n-1} + 1}{(x+1)^3} > 0.$$

This implies

Theorem 2.1 *The function $S_n : [0, 1] \rightarrow \mathbf{R}$ defined above is strictly convex.*

To study the log-convexity of S_n , notice that

$$(\log S_n(x))'' = \frac{S''_n(x)S_n(x) - (S'_n(x))^2}{S_n^2(x)}.$$

Denote $\varphi_n(x) := \frac{(x+1)^4}{4} [S''_n S_n - S_n'^2](x)$, that is

$$\begin{aligned} \varphi_n(x) &= (2x^{2n+1} - x + 1)[n(2n-1)x^{2n+1} + (4n^2 - 1)x^{2n} \\ &\quad + n(2n+1)x^{2n-1} + 1] - (2nx^{2n+1} + (2n+1)x^{2n} - 1)^2. \end{aligned}$$

It is easy to see that $\varphi_n(0) = \varphi_n(1) = 0$.

For $n \geq 2$ we have $\varphi'_n(0) = -1$, so $\varphi_n(x) < 0$ in a neighborhood of 0. For $n = 1$ we have $(S''_1 S_1 - S_1'^2)(x) = 8x(1-x) \geq 0$. In conclusion:

Theorem 2.2 *The function $S_1 : [0, 1] \rightarrow \mathbf{R}$ is log-convex. For $n \geq 2$, S_n is not log-convex.*

This means also that the Rényi entropy R_1 is concave and for $n \geq 2$, R_n is not concave.

Denote by $g_n(x) = 2nx^{2n+1} + (2n+1)x^{2n} - 1$ the numerator of the first order derivative $S'_n(x)$. It is easy to see that $g_n(0) = -1$, $g_n(1) = 4n$ and g_n is strictly increasing. It follows that there exists a unique point $x_n \in (0, 1)$ such that $g_n(x_n) = 0$, that is $S'_n(x_n) = 0$. Let us remark that

$$2nx_n^{2n+1} + (2n+1)x_n^{2n} - 1 < 2nx_n^{2n+1} + (2n+1)x_n^{2n} - 1 = 0$$

implies

$$x_n < (4n + 1)^{-\frac{1}{2n+1}},$$

while

$$2nx_n^{2n} + (2n + 1)x_n^{2n} - 1 > 2nx_n^{2n+1} + (2n + 1)x_n^{2n} - 1 = 0$$

implies

$$x_n > (4n + 1)^{-\frac{1}{2n}}.$$

From the above we get

$$(2.1) \quad (4n + 1)^{-\frac{1}{2n}} < x_n < (4n + 1)^{-\frac{1}{2n+1}}.$$

As a consequence,

$$\lim_{n \rightarrow \infty} x_n = 1.$$

We have also

$$\min_{x \in [0,1]} S_n(x) = S_n(x_n) = \frac{2x_n^{2n+1} - x_n + 1}{x_n + 1}.$$

This fact, combined with (2.1), gives us the possibility to estimate the minimum value of S_n ; in particular,

$$(2.2) \quad \min_{x \in [0,1]} S_n(x) > \frac{1 - x_n}{2} > \frac{1}{2} \left(1 - (4n + 1)^{-\frac{1}{2n+1}} \right).$$

To conclude, we see that

- S_n is convex on $[0, 1]$.
- $S_n(0) = S_n(1) = 1$.
- x_n is a minimum point of S_n , satisfying (2.1).
- The minimum value of S_n satisfies (2.2).

In what follows we turn our attention to Shannon entropy H_n .

Theorem 2.3 *If $H_n : [0, 1] \rightarrow \mathbf{R}$ is the Shannon entropy associated to the 0-Bernstein distribution p_n , then*

$$\lim_{n \rightarrow \infty} \frac{H_n(1 - \frac{1}{n})}{\log n} = \frac{e-1}{e} \quad \text{and} \quad \lim_{n \rightarrow \infty} H_n(1 - \frac{1}{n^2}) = 0.$$

Proof: We have

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{H_n(1 - \frac{1}{n})}{\log n} &= \lim_{n \rightarrow \infty} \left[1 - \left(1 - \frac{1}{n}\right)^n \right] \frac{\log \frac{1}{n} + \frac{n-1}{n} \log \left(1 - \frac{1}{n}\right)^n}{\log \frac{1}{n}} \\ &= \left(1 - \frac{1}{e}\right) \left(1 + \frac{-1}{-\infty}\right) = \frac{e-1}{e}. \end{aligned}$$

$$\begin{aligned} \lim_{n \rightarrow \infty} H_n(1 - \frac{1}{n^2}) &= \lim_{n \rightarrow \infty} \left[\left(1 - \frac{1}{n^2}\right)^n - 1 \right] \left[\log \frac{1}{n^2} + \left(1 - \frac{1}{n^2}\right) \log \left(1 - \frac{1}{n^2}\right)^{n^2} \right] \\ &= \lim_{n \rightarrow \infty} \left[\left(1 - \frac{1}{n^2}\right)^n - 1 \right] \log \frac{1}{n^2} \\ &= \lim_{x \rightarrow 0} \left[(1 - x^2)^{\frac{1}{x}} - 1 \right] \log x^2 \\ &= 2 \lim_{x \rightarrow 0} \frac{(1 - x^2)^{\frac{1}{x}} - 1}{x} \cdot x \log x = 0. \end{aligned}$$

For $n \rightarrow \infty$, the entropies of the 0-Bernstein distribution p_n converge to those of p , the geometric distribution with parameter $1 - x$.

$$\lim_{n \rightarrow \infty} S_n(x) = S(x), \quad \lim_{n \rightarrow \infty} R_n(x) = R(x),$$

$$\lim_{n \rightarrow \infty} T_n(x) = T(x), \quad \lim_{n \rightarrow \infty} H_n(x) = H(x).$$

Remark 2.1 *The functions $S(x)$, $R(x)$, $T(x)$, $H(x)$ have rather simple graphs. This fact, combined with the above results, gives us an idea about the behaviour of the functions $S_n(x)$, $R_n(x)$, $T_n(x)$, $H_n(x)$ for large values of n .*

The above convergences are not uniform on the whole interval $[0, 1]$. Here are some estimates obtained using the norm in $L_1[0, 1]$.

Theorem 2.4 *For any $n \in \mathbf{N}$, the Shannon entropies H_n , H and the Tsallis entropies T_n , T satisfy the inequalities*

$$\begin{aligned}\|H_n - H\|_{L^1} &\leq \sqrt{\frac{2}{2n+1}} + \frac{1}{n+1}, \\ \|T_n - T\|_{L^1} &\leq \sqrt{\frac{2}{4n+3}}.\end{aligned}$$

Proof: We have

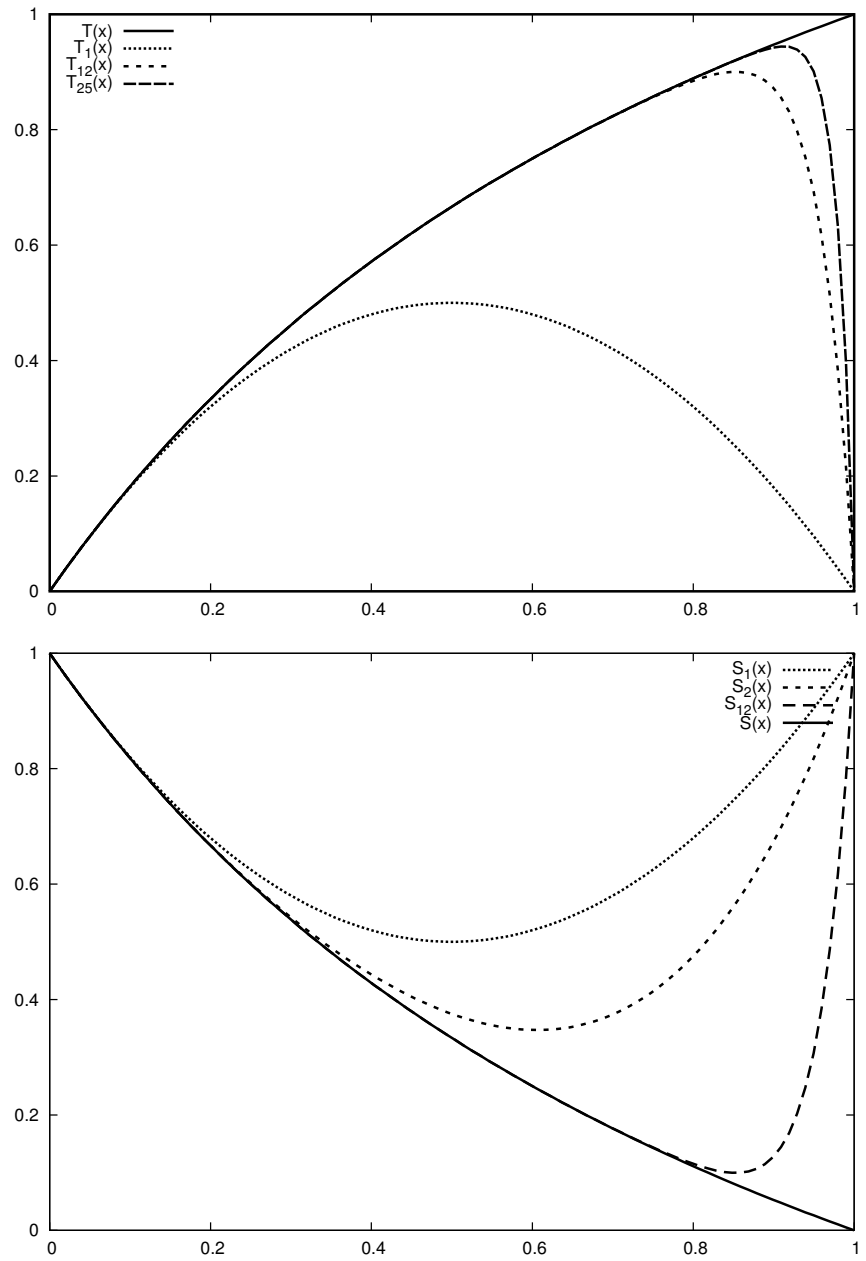
$$\begin{aligned}\|H_n - H\|_{L^1} &= \int_0^1 |H_n(x) - H(x)| dx \\ &\leq \int_0^1 |x^n \log(1-x)| dx + \int_0^1 x^n \left| \frac{x}{1-x} \log x \right| dx\end{aligned}$$

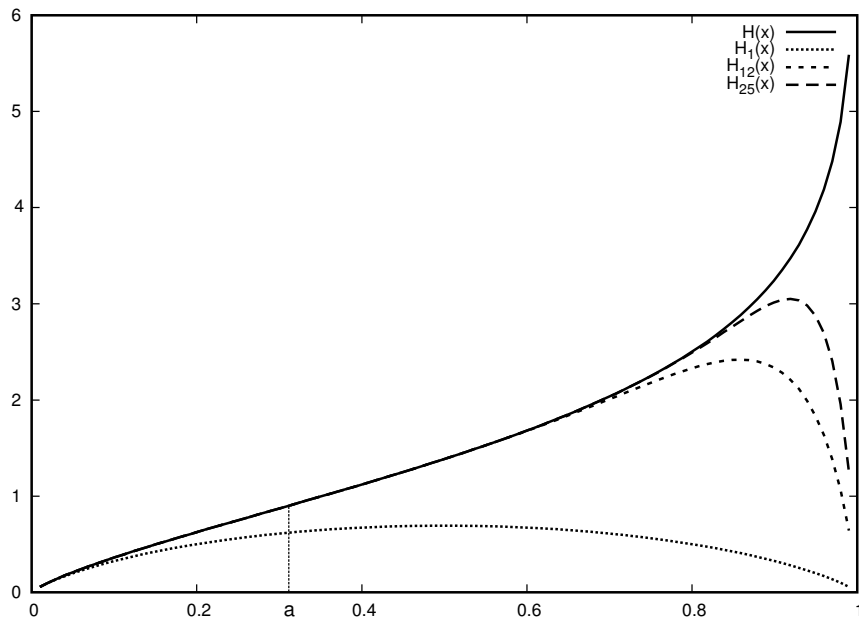
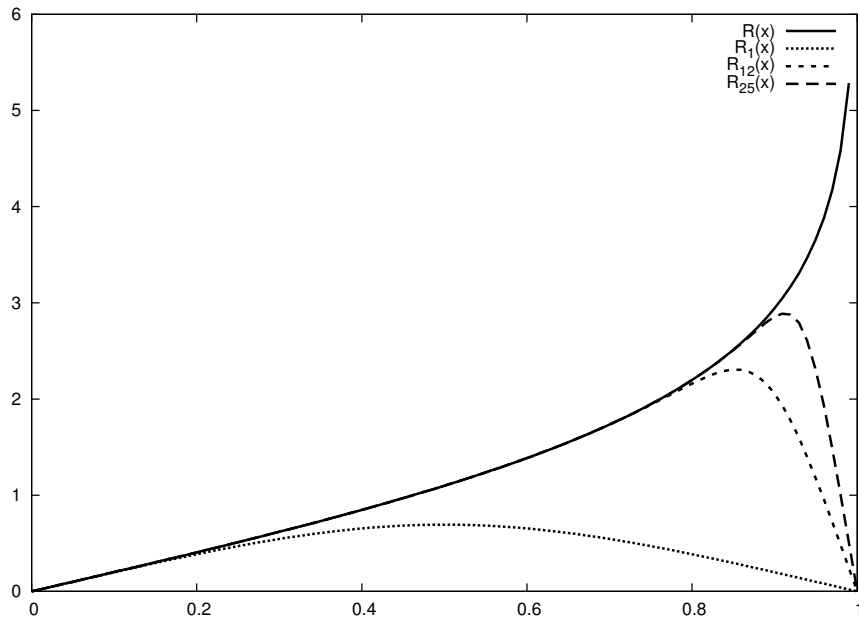
$$\int_0^1 |x^n \log(1-x)| dx \leq \sqrt{\int_0^1 x^{2n} dx} \cdot \sqrt{\int_0^1 \log^2(1-x) dx} = \frac{1}{\sqrt{2n+1}} \cdot \sqrt{2},$$

$$\int_0^1 x^n \left| \frac{x}{1-x} \log x \right| dx \leq \int_0^1 x^n dx = \frac{1}{n+1}.$$

Similarly,

$$\begin{aligned}\|T_n - T\|_{L^1} &= \int_0^1 \frac{2x^{2n+1}}{1+x} dx \leq 2 \sqrt{\int_0^1 x^{4n+2} dx} \cdot \sqrt{\int_0^1 \frac{1}{(1+x)^2} dx} \\ &= 2 \sqrt{\frac{1}{4n+3}} \cdot \sqrt{\frac{1}{2}}.\end{aligned}$$





References

- [1] A. Aral, V. Gupta, R.P. Agarwal. *Applications of q -Calculus in Operator Theory*. Springer 2013.
- [2] T.M. Cover, J.A. Thomas. *Elements of Information Theory*. John Wiley & Sons 1991.
- [3] I. Raşa. Entropies and Heun functions associated with positive linear operators. *Appl. Math. Comput.*, 268: 422-431, 2015.
- [4] I. Raşa. Complete monotonicity of some entropies. *Periodica Math. Hungar.*, to appear.
- [5] A. Rényi. On measures of entropy and information. Proc. Fourth Berkeley Symp. Math. Statist. Prob., vol. 1, Univ. California Press, 547-561, 1961.
- [6] C.E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal.*, 27 (3): 379-423, 1948.
- [7] C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.*, 52: 479-487, 1988.