VARMA-TSALLIS ENTROPY: PROPERTIES AND APPLICATIONS

Romică TRANDAFIR[*], Vasile PREDA[**], Sorin DEMETRIU[*], Ion MIERLUŞ-MAZILU[*]

[*]Technical University of Civil Engineering, Bucharest, Romania (romica.trandafir@utcb.ro, demetriu@utcb.ro, mmi@utcb.ro),
[**]University of Bucharest, Romania (vasilepreda0@gmail.com)

DOI: 10.19062/1842-9238.2018.16.2.8

Abstract: *Entropy represents a universal concept in science suitable for quantifying the uncertainty of a series of random events. In this paper we obtain a new type of entropy named Varma-Tsallis entropy starting from Tsallis entropy and Varma entropy. Some properties and applications of the proposed entropy in water engineering are presented.*

Keywords: *entropy, principle of maximum entropy, Lagrange multipliers, maximum annual discharge*

1. INTRODUCTION

In this paper we present some types of entropy, the connections between them and we propose a new type of entropy starting from Tsallis entropy and Varma entropy, namely, the Varma-Tsallis entropy. For this proposed entropy we present some properties and a procedure that shows how it can be applied in practical applications. An application of the proposed entropy for the determination of the cumulative distribution function (cdf) for the recorded annual discharges of the Prut river and the Somes river.

2. TYPES OF ENTROPY

2.1. Boltzmann-gibbs-shannon entropy (referred to as the shannon entropy, 1948) [7]

- **Discrete case**

Let X be a random variable that takes on values x_i , $i=\overline{1,N}$, that occur with probabilities p_i , $0 \le p_i \le 1$, $i=\overline{1,N}$ and $\sum_{i=1}^{N} p_i = 1$. The information gain from the occurrence of any event x_i , is given by

$$\Delta H(x_i) = -\log_2(p_i) \tag{1}$$

i.e. the information gained is the logarithm of inverse of the probability of occurrence. For the all N events the average of information gain H_S can be expressed as

$$S = H_S = \sum_{i=1}^{N} p_i H(x_i) = -\sum_{i=1}^{N} p_i \log_2(p_i) \tag{2}$$

Equation (2) is the Shannon entropy, also called *informational entropy*.

- **Continuous case**

If the random variable is non-negative continuous with a probability density function (pdf) $f(x)$, the the Shannon entropy can be writen as:

$$S = H_S(X) = H_S(f) = -\int_0^\infty f(x)\log_2\left(f(x)\right)dx \qquad (2')$$

2.2. Renyi entropy (1961) [6]

Renyi proposed a generalized entropy of order α as

- **Discrete case**

$$R = H_\alpha(X) = \frac{1}{1-\alpha}\log\left(\sum_{i=1}^{N} p_i^\alpha\right) \text{ , } \alpha > 0 \text{ , } \alpha \neq 1 \qquad (3)$$

Remark. Renyi's entropy contains the Shannon entropy as a special case ($\lim_{\alpha\to 1} H_\alpha(X) = S$).

- **Continuous case**

$$R = H_\alpha(X) = \frac{1}{1-\alpha}\log\left(\int_0^\infty f_X(x)^\alpha dx\right) \text{ , } \alpha > 0 \text{ , } \alpha \neq 1 \qquad (3')$$

2.3. Varma entropy (1966) [12]

- **Discrete case**

$$V = H_\alpha^\beta(X) = \frac{1}{\beta-\alpha}\sum_{i=1}^{N} p_i^{\alpha+\beta-1} \text{, } \beta-1<\alpha<\beta \text{ , } \beta \geq 1 \qquad (4)$$

- **Continuous case**

$$V = H_\alpha^\beta(X) = \frac{1}{\beta-\alpha}\ln\left(\int_0^\infty f_X(x)^{\alpha+\beta-1}dx\right) \text{ , } \beta-1<\alpha<\beta \text{ , } \beta \geq 1 \qquad (4')$$

Remark:

The Varma entropy includes, as particular cases, the Renyi entropy.

$$\lim_{\beta\to 1} H_\alpha^\beta(X) = R_\alpha(X) = -\frac{1}{\alpha-1}\log\int_0^\infty \left(f(x)\right)^\alpha dx \qquad (5)$$

and the Shannon entropy

$$\lim_{\substack{\beta\to 1\\ \alpha\to 1}} H_\alpha^\beta(X) = S = -\int_0^\infty f(x)\cdot\log\left(f(x)\right)dx \qquad (5')$$

2.4. Kapur entropy (1967) [3]

- **Discrete case**

$$K = H_{\alpha,\beta}(X) = \frac{1}{\beta-\alpha}\frac{\sum_{i=1}^{N} p_i^{\alpha+\beta-1}}{\sum_{i=1}^{N} p_i^\beta} \text{ , } \alpha \neq \beta \text{ , } \alpha > 0,\ \beta > 0 \qquad (6)$$

- **Continuous case**

$$K = H_{\alpha,\beta}(X) = \frac{1}{\beta-\alpha}\frac{\int_0^\infty \left(f_X(x)\right)^\alpha dx}{\int_0^\infty \left(f_X(x)\right)^\beta dx} \text{ , } \alpha \neq \beta \text{ , } \alpha > 0,\ \beta > 0 \qquad (6')$$

2.5. Tsallis entropy (1988) [10]

- **Discrete case**

$$T_m = H_m(X) = \frac{1}{m-1}\sum_{i=1}^{N} p_i\left(1-p_i^{m-1}\right),\ m \neq 1,\ m \in R \tag{7}$$

- **Continuous case**

$$T_m = H_m(X) = H_m(f) = \frac{1}{m-1}\int_0^{\infty}\left\{f_X(x) - \left[f(x)\right]^m\right\}dx,\ \ m \neq 1,\ m \in R \tag{7'}$$

Remarks:

1. For $m \to 1$ Tsallis entropy converges to Shannon entropy.
2. For $m < 0$ Tsallis entropy is concave and for $m > 0$ Tsallis entropy is convex.
3. For all *m* the Tsallis entropy decreases as *m* increases.

2.6. Varma-Tsallis entropy

If we denote

$$m = m + r - 1 \tag{8}$$

- **Discrete case**

$$VT_{m,r} = H_{m,r} = \frac{1}{m-r}\sum_{i=1}^{N} p_{i\left(1-p_i^{m+r-2}\right)} = \frac{m-1}{m-r}\left[\frac{1}{m-1}\sum_{i=1}^{N} p_i\left(1-p_i^{m-1}\right)\right], m \neq 1, r \neq 1$$

$$VT_{m,r} = \frac{m-1}{m-r}T_m, m \neq 1, r \neq 1 \tag{9}$$

- **Continuous case**

$$VT_{m,r} = H_{m,r}(X) = \frac{1}{m-r}\left(1 - \int_0^{\infty} f_x(x)^{m+r-1}dx\right), m, r \neq 1 \tag{9'}$$

The equation (9') becomes

$$VT_{m,r}(X) = \frac{m-1}{m-r}\left\{\frac{1}{m-1}\left(1 - \int_0^{\infty} f_x(x)^m\, dx\right)\right\}, m \neq 1, r \neq 1 \tag{9''}$$

Thus

$$VT_{m,r} = \frac{m-1}{m-r}T_m, m \neq 1, r \neq 1 \tag{10}$$

Let $\alpha = \frac{m-r}{m+r-2}$ be a parameter. Then

$$T_m = \alpha VT_{m,r}, m \neq 1, r \neq 1 \tag{10'}$$

3. PROPERTIES OF THE VARMA-TSALLIS ENTROPY

3.1. Concavity, convexity

It can be shown that for

$$0 < a < 1,\ P = \{p_i\}_{1 \le i \le N},\ Q = \{q_i\}_{1 \le i \le N},\ G = \{g_i = ap_i + (1-a)q_i\}_{1 \le i \le N} \tag{11}$$

then

i)

$$\alpha \cdot VT_{m,r}(G) \geq a \cdot \alpha \cdot VT_{m,r}(P) + (1-a)\alpha \cdot VT_{m,r}(Q) \tag{12}$$

for $m.\alpha > 0$,

ii)

$$\alpha \cdot VT_{m,r}(G) \le \alpha \cdot \alpha \cdot VT_{m,r}(P) + (1-\alpha)\alpha \cdot VT_{m,r}(Q) \quad (13)$$

for $m > 0, \alpha < 0,$

iii)

$$\alpha \cdot VT_{m,r}(G) \le \alpha \cdot \alpha \cdot VT_{m,r}(P) + (1-\alpha)\alpha \cdot VT_{m,r}(Q) \quad (14)$$

for $m < 0, \alpha > 0,$

iv)

$$\alpha \cdot VT_{m,r}(G) \ge \alpha \cdot \alpha \cdot VT_{m,r}(P) + (1-\alpha)\alpha \cdot VT_{m,r}(Q) \quad (15)$$

for $m < 0, \alpha < 0.$

$$VT_{m,r,extreme} = \frac{m+r-2}{m-r}\frac{N^{m-1}-1}{m-1}$$

3.2 maximum value

It is well known that the Tsallis entropy attains an extreme value for all values of m when all p_i , $i = \overline{1,N}$ are equal, i. e. $p_i = \frac{1}{N}$ and this extreme value is

$$T_{m,extreme} = \frac{N^{m-1}-1}{m-1} \quad (16)$$

For $m > 0$ this extreme value is a maximum value and for $m < 0$ this extreme value is a minimum value. Considering the equations (10), (10') the extreme value for Varma-Tsallis entropy will be given by

$$VT_{m,r,extreme} = \frac{m+r-2}{m-r}\frac{N^{m-1}-1}{m-1} \quad (17)$$

4. THE PRINCIPLE OF MAXIMUM ENTROPY

Considering the following principles of ancient wisdom:

- “speak truth and nothing but truth
- make use of all the given information you are given and scrupulously avoid using the information not given to you
- make use of all the given and be maximally uncommitted to the missing information or be maximally uncertain about it” [9], E. T. Jaynes (1957) [1,2] formulated the principle of maximum entropy (POME), which states that *“one should choose the distribution that has the highest entropy, subject to the given information”*.

The implication here is that POME considers all of the given information and, at the same time, avoids consideration of any information that is not given. This is consistent with Laplace’s principle of insufficient reason (or principle of indifference), according to which all outcomes of an experiment should be considered equally likely unless there is information to the contrary.

Therefore, POME enables entropy theory to achieve the probability distribution of a given random variable [8].

To obtain the probability distribution of a given random variable by POME, it can be used the following procedure:

- fix the kind of entropy, in this case Varma-Tsallis entropy (9’)
- give the constraints
- maximize the entropy by POME
- obtain the probability distribution according to constraints
- determine the Lagrange multipliers
- determine the maximum entropy.

4.1 SPECIFICATION OF CONSTRAINTS

Given a sample of random variable *X*, $(x_1, x_2, ..., x_N)$, a type of restriction can be given by the following equations

$$\int_0^{\infty} x^k f(x)dx = \overline{x^k} \ , \ k = 0,1,2,3... \tag{18}$$

where $\overline{x^k}$, $k = 0,1,2,3...$ are empirical moments of random variable *X*.

Remark: In water engineering, empirical moments k=0,1,2,3 are considered.

The constraints (18) are not sufficient to determine $f(x)$ uniquely, because there may be many, even infinity of probability distributions satisfying (18).

4.2. Entropy maximization using lagrange multipliers

To determine $f(x)$ we should maximize the Varma-Tsallis entropy (9’) subject to (18) using the method of Lagrange multipliers.

There are two fortunate circumstances favoured the great success of the POME, since in all optimization problems the difficulties arise when we have to decide whether

- the extreme value found is a maximum or minimum
- the maximum obtained is local or global
- the non-negativity constraints are satisfied

namely: the Varma-Tsallis entropy function is a concave function and the pdf is always non-negative. The Lagrangian function *L* is given, in this case, by

$$\begin{aligned} L = \frac{1}{m-r}\left\{1 - \int_0^{\infty}\left(f(x)\right)^{m+r-1} dx\right\} - \lambda_0\left\{\int_0^{\infty} f(x)dx - 1\right\} - \\ - \sum_{i=1}^{k}\lambda_i\left\{\int_0^{\infty} x^i f(x)dx - \overline{x^i}\right\} \ , \ k = 1,2,3... \end{aligned} \tag{19}$$

where λ_i , $i = \overline{1,k}$ are the Lagrange multipliers.

Differentiating equation (19) with respect to $f(x)$ and equating the derivative to zero, we obtain:

$$\frac{\partial L}{\partial f(x)} = 0 = \frac{1}{m-r} - \frac{m+r-1}{m-r}\left[f(x)\right]^{m+r-2} - \sum_{i=0}^{k}\lambda_i x^i \ , \ k = 1,2,3... \tag{20}$$

Thus, the pdf of *X* is

$$f(x) = \left\{\frac{1}{m+r-1}\left[\lambda_0 - (m-r)\sum_{i=1}^{k}\lambda_i \cdot x^i\right]\right\}^{\frac{1}{m+r-2}} \ , \ k = 1,2,3... \tag{21}$$

Substituting equation (21) in equation (18) the result is, respectively:

$$\int_0^{\infty}\left\{\frac{1}{m+r-1}\left[\lambda_0 - (m-r)\sum_{i=0}^{k}\lambda_i \cdot x^i\right]\right\}^{\frac{1}{m+r-2}} dx = 1 \ , \ k = 1,2,3... \tag{22}$$

$$\int_0^{\infty} x^j \left\{\frac{1}{m+r-1}\left[\lambda_0 - (m-r)\sum_{i=0}^{k}\lambda_i \cdot x^i\right]\right\}^{\frac{1}{m+r-2}} dx = \overline{x^j}, \quad j = \overline{1,k}, \ k = 1,2,3... \tag{23}$$

The system given by equations (22)-(23) do not have generally an analytical solution but can be solved with numerical methods.

Substitution of equation (21) in equation (9’) leads to maximum Varma-Tsallis entropy

$$VT_{m,r} = \frac{1}{m-r}\left\{1-\int_0^\infty\left[\frac{1}{m+r-1}\left(\lambda_0-(m-r)\sum_{i=1}^{k}\lambda_i\cdot x^i\right)^{\frac{1}{m+r-2}}\right]^{m+r-1}dx\right\},\ k=1,2,3... \tag{24}$$

Last equation shows that the $VT_{m,r}$ of the distribution probability of X depends only on the constraints, since the Lagrange multipliers themselves depend on the same constraints.

5. APPLICATIONS

The design of the hydraulic structures like spillways, dykes or diversions is based on the maximum discharges corresponding to standard values of the annual probability of exceedance (usually in the range 1% - 0,1%). The length of the registered data rarely exceeds 50 years, which means that the empirical probabilities of exceedance of the maximum annual discharges are in the range 2-98%. The main problem is the real probability of exceedance of the outliers is not known, meaning that the values of the statistical parameters are influenced by the empirical probability which is assigned to the extreme values.

The method described above was used to determine the pdf for the maximum annual discharges of Prut River recorded at Radauti and the pdf of the data for the Somes River recorded at Satu Mare. The maximum annual discharges rates of the river Prut at Radauti gauge station between 1978 and 2015 (Fig.1) , X_i , $i=\overline{1,38}$, and of the river Somes at Satu Mare gauge station between 1928 and 1988 (Fig 2) , X_i , $i=\overline{1,64}$ are used to obtain the probability distributions of discharges in order to be able to make predictions of floods.

The measured discharge data are normalized

$$x_i = \frac{X_i}{\max(X_i)-\min(X_i)}\ ,\ i=\overline{1,38}\ ,\ respectively\ i=\overline{1,64}$$

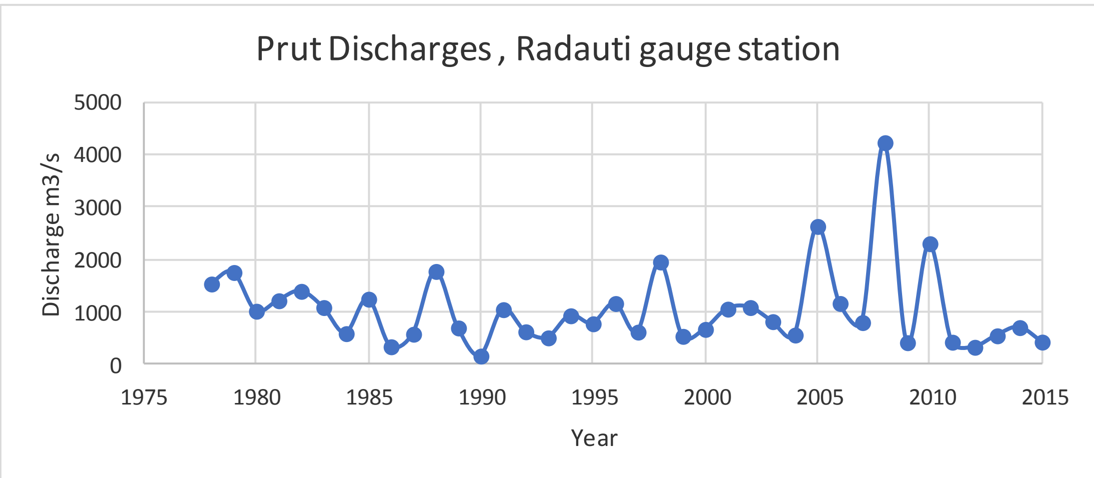

FIG.1. Maximum annual discharges of the Prut River recorded at Radauti station

$X\max=4240$, $X\min=163$, $w=X\max-X\min=4077$

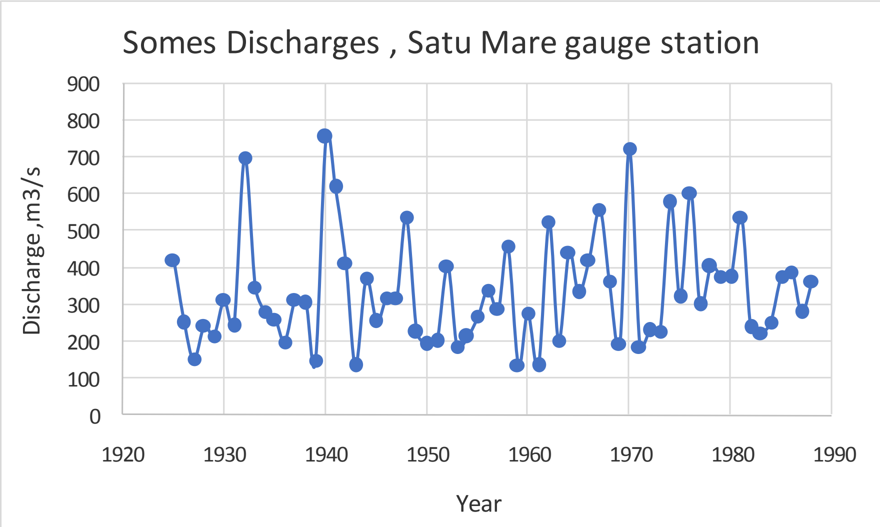

FIG. 2. Maximum annual discharges of the Somes River recorded at Satu Mare station
$X\max = 756$, $X\min = 134$, $w = X\max - X\min = 622$

The empirical moments of these normalized records make out the first data set:

$\overline{x^1} = 0.25729$, $\overline{x^2} = 0.10104$, $\overline{x^3} = 0.05829$

and for the second data set

$\overline{x^1} = 0.53851$, $\overline{x^2} = 0.34619$, $\overline{x^3} = 0.26046$

The non-linear equations systems for Lagrange's multiplier, considering $m = 2$, $r = 0.5$, $k = 2$ are solved for each case using a numerical method.

We obtain for the recorded discharges at Radauti

$\lambda_0 = -2.579$, $\lambda_1 = -2.671$, $\lambda_2 = 1.104$.

and for the recorded discharges at Satu Mare

$\lambda_0 = -2.002$, $\lambda_1 = -0.581$, $\lambda_2 = -0.251$.

Finally, the pdf for $(m = 2 \, , \, r = 0.5 \, , \, k = 2)$ is given by

$$f(x,\lambda,m,r,k) = \frac{1}{m+r-1}\left[\lambda_0 - (m-r)\sum_{i=1}^{k}\left(\lambda_i \cdot x^i\right)^{\frac{1}{m+r-2}}\right]$$

and cumulative distribution function

$$F(x,\lambda,m,r,k) = \frac{1}{m+r-1}\int_0^x\left[\lambda_0 - (m-r)\sum_{i=1}^{k}\left(\lambda_i \cdot t^i\right)^{\frac{1}{m+r-2}}\right]dt$$

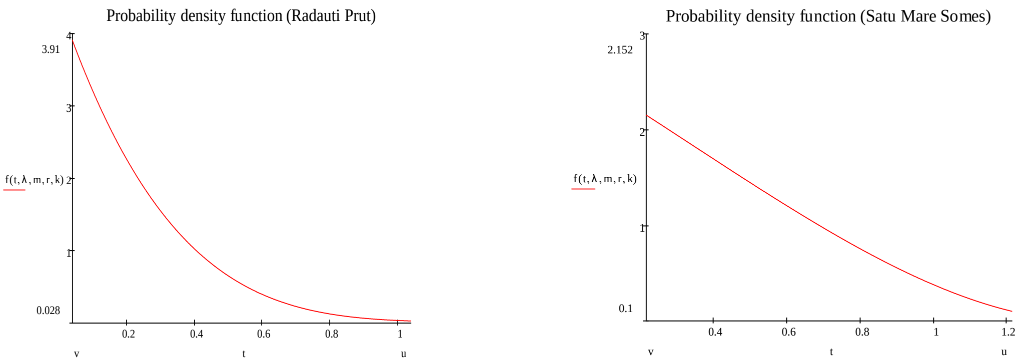

FIG. 3. Graphics of probability density functions for the two gauges

A comparison between the maximum annual discharge quantiles corresponding to different mean return intervals (periods) that are estimated using the obtained probability distributions and other reference probability distributions recommended in Statistical Hydrology is presented in Table 1.

Table 1. The maximum annual discharge quantiles corresponding to different mean return periods

Gauge station	Quantiles	Zα				
	Probabilities	α=0.90	α=0.95	α=0.98	α=0.99	α=0.995
	T (years)	**10**	**20**	**50**	**100**	**200**
Satu Mare, Somes	Varma-Tsallis	555	620	681	711	730
	Log-Pearson type III	532	627	756	857	963
	Lognormal 3-parameter	534	633	768	875	987
	Generalized Extreme Value	530	627	760	866	979
	Gumbel Max	529	612	720	801	882
Radauti, Prut	Varma-Tsallis	2138	2598	3127	3467	3742
	Log-Pearson type III	1953	2487	3272	3934	4662
	Lognormal 3-parameter	1924	2421	3135	3725	4361
	GEV (Generalized Extreme Value)	1900	2477	3415	4292	5348
	Gumbel Max	2055	2488	3048	3468	3886

5. CONCLUSIONS

In this paper we introduced a generalization of Tsallis entropy, called Varma-Tsallis entropy, highlighted some properties and showed how it can be used to determine a probability density function of a random variable. The method presented here was used for the recorded maximum annual discharges of the Prut River and the Somes River. The results we obtained show that the method is reliable.

REFERENCES

[1] Jaynes, E.T., *Information theory and statistical mechanics*, I. Phys. Rev. ,1957, 106, 620–630;

[2] Jaynes, E.T., *Information theory and statistical mechanics*, II. Phys. Rev. ,1957, 108, 171–190;

[3] Kapur, J.N., Generalized entropy of order α and type β . *The Math. Seminar*, 1967, 4, 78–94;.

[4] Madan Mohan Sati, Nitin Gupta, *On Partial Monotonic Behaviour of Varma Entropy and Its Application in Coding Theory*, Journal of the Indian Statistical Association, May 24, 2016;

[5] Papalexiou, S.M.; Koutsoyiannis, D., *Entropy based derivation of probability distributions: A case study to daily rainfall*. Adv. Water Resources, 2012, 45, 51–57;

[6] Renyi, A., *On measures of entropy and information*. In: Neyman, J. (Ed.), Proceedings of the Fourth Berkeley Symposium on Mathematics, Statistics and Probability,1961, Vol. I. University of California Press, Berkeley, CA, 547-56;

[7] Shannon, C.E., *A mathematical theory of communication*,1948, Bell System Technical Journal, 27, 379-423, 623-656;

[8] Singh V. P., Sivakumar, B., Cui H., *Tsallis Entropy Theory for Modeling in Water Engineering: A Review*, Entropy ,2017, 19, 641; doi:10.3390/e19120641, www.mdpi.com/journal/entropy;

[9] Singh, V. P., Fiorentino, M. (editors), *Entropy and Energy Dissipation in Water Resoources*, 3-20,1992, Kluwer Academic Publishers;

[10] Tsallis, C., *Possible generalization of Boltzmann-Gibbs statistics*. J. Stat. Phys. 1988, 52, 479–487;

[12] Varma, R.S., Generalization of Renyi’s entropy of order α, *Journal of Mathematical Sciences*, 1966,1, 34-48.