JIRSS (2010)
Vol. 9, No. 2, pp 115-126

ADK Entropy and ADK Entropy Rate in Irreducible- Aperiodic Markov Chain and Gaussian Processes

Morteza Khodabin

Department of Mathematics, Karaj Branch, Islamic Azad University, Karaj, Iran. (m-khodabin@kiau.ac.ir)

Abstract. In this paper, the two parameter ADK entropy, as a generalized of Re'nyi entropy, is considered and some properties of it, are investigated. We will see that the ADK entropy for continuous random variables is invariant under a location and is not invariant under a scale transformation of the random variable. Furthermore, the joint ADK entropy, conditional ADK entropy, and chain rule of this entropy is discussed. The ADK entropy rate is defined and is used for deriving the entropy rate of stationary Gaussian processes and an irreducible- aperiodic Markov chain.

1 Introduction

The concept of entropy or (more accurately) entropy rate whether of a stochastic process, information source or dynamical system has proved to be instrumental in many fields of science. Shannon [16] in-

Key words and phrases: ADK entropy, ADK entropy rate, gaussian process, Markov chain, Renyi entropy, Renyi entropy rate, Shannon entropy, Shannon entropy rate, spectral density function, stationary process.

troduced his entropy and axiomatic characterization of it. But since then a number of entropy-like quantities have appeared in the scientific literature. All these new quantities share some but not all properties with the Shannon entropy. The most important examples are the Renyi entropy and the Tsallis entropy. There are so many other definitions of entropy-like quantities that Arndt [2] was able to write a whole book entitled "Information Measures ". The Renyi entropy was introduced by Renyi [13, 14] and soon after it found application in graph theory. The original reason for Renyi to introduce his new entropy is said to be that he planned to use it in an information theoretic proof of the central Limit Theorem. Renyi entropies of order greater than 2 are also known to be related to search problems, see [6, 18]. Aczel and Daroczy [1], Kapur [8], have introduced a general form of entropy,which will henceforth be called ADK entropy, and of which Re'nyi entropy is a special case. The focus in this paper will be on some properties of this entropy. the entropy rate was defined for stochastic processes. The Shannon entropy rate is extensively studied for stochastic processes, especially for stationary processes with discrete or continuous time (see [5] and references therein). For example, the rate of Shannon entropy for a stationary Gaussian process was obtained by Kolmogorov [9]. The rate of Renyi entropy for stochastic processes was obtained by Rached et al. [12]. He also obtained an operational characteristic for the Renyi entropy rate in coding theory. Obviously this paper can only present a few results on this huge topic and give some pointers to the literature. The fact is that most of the alternative definitions of entropy may be useful in very special situations or may fulfill modified axiom systems but they do not have operational definitions. Some of the alternative and definitions will find applications and operational definitions in the future, but most of them have already been more or less forgotten. This was what happened with the ADK entropy.

This paper is organized as follows:

In Section 2, the ADK entropy for discrete random variables are introduced and the chain rule of this entropy is proved. In Section 3, the ADK entropy for continuous random variables is presented and the chain rule of this entropy is obtained. In section 4 a relation for obtaining the rate of ADK entropy is obtained. Then, using this relation, the ADK entropy rate for an irreducible-aperiodic Markov chain is derived. Furthermore, the ADK entropy rate is calculated for Gaussian stationary processes. Also in this section its shows that the ADK entropy rate for two family (AR(p) and MA(q)) time series is same. Finally, the conclusions are drawn in section 5.

2 ADK entropy for discrete random variables

In information theory, the ADK entropy, as a generalization of Renyi and Shannon entropy, is one of a family of functional for quantifying the diversity, uncertainty or randomness of a system. It is named ADK because Aczel, Daroczy and Kapur initial works. In this paper assumes that:

1. All integrals and sums exist.
2. $\alpha, \beta > 0 \quad , \quad \alpha + \beta > 1 \quad , \quad \alpha \neq 1.$
3. log is to the base 2 and entropy is expressed in bits.

Let $(S, \beta_S, P)_p \in \Delta_n$ be an statistical space, where $S = \{x_1, x_2, \dots, x_n\}$, $\Delta_n = \{P = (p_1, p_2, \dots, p_n), p_i \geq 0, \sum_{i=1}^{n} p_i = 1\}$ and β_S is the σ-field of all the subsets of S.

Definition 2.1. ADK entropy of any probability distribution $P = (p_1, p_2, \dots, p_n)$ of a random variable is defined as

$$H_{\alpha,\beta}(X) \equiv H_{\alpha,\beta}(P) = \frac{1}{1-\alpha} \log \frac{\sum_{i=1}^{n} p_i^{\alpha+\beta-1}}{\sum_{i=1}^{n} p_i^{\beta-1}} \tag{2.1}$$

Definition 2.2. (The joint ADK entropy):If $X_1, X_2, \dots, X_n$ are distributed according to $p(x_1, x_2, \dots, x_n)$ then

$$H_{\alpha,\beta}(X_1, X_2, \dots, X_n) = \frac{1}{1-\alpha} \log \frac{\sum_{x_1,x_2,\dots,x_n} p(x_1, x_2, \dots, x_n)^{\alpha+\beta-1}}{\sum_{x_1,x_2,\dots,x_n} p(x_1, x_2, \dots, x_n)^{\beta-1}} \tag{2.2}$$

Definition 2.3. (Conditional ADK entropy): Let $X_1, X_2, \dots, X_n$ are distributed according to $p(x_1, x_2, \dots, x_n)$. Then the conditional ADK entropy of random variable X_n , given $X_1, X_2, \dots, X_{n-1}$, is defined as

$$H_{\alpha,\beta}(X_n | X_1, X_2, \dots, X_{n-1}) = \frac{1}{1-\alpha} \log \tag{2.3}$$

$$\frac{\sum_{x_1,x_2,\dots,x_{n-1}} p(x_1, x_2, \dots, x_{n-1})^{\beta-1} \sum_{x_1,x_2,\dots,x_n} p(x_1, x_2, \dots, x_n)^{\alpha+\beta-1}}{\sum_{x_1,x_2,\dots,x_{n-1}} p(x_1, x_2, \dots, x_{n-1})^{\alpha+\beta-1} \sum_{x_1,x_2,\dots,x_n} p(x_1, x_2, \dots, x_n)^{\beta-1}}$$

It's easy to show that the ADK entropy of a collection of random variables is the sum of the conditional ADK entropies. i.e.

$$H_{\alpha,\beta}(X_1, X_2, \ldots, X_n) = \sum_{j=1}^{n} H_{\alpha,\beta}(X_j|X_{j-1}, \ldots, X_1) \tag{2.4}$$

3 ADK entropy for continuous random variables

Let X be a random variable having an absolutely continuous distribution with density function $f(x)$. The ADK entropy of order α and β is defined as

$$H_{\alpha,\beta}(X) = \frac{1}{1-\alpha} \log \frac{\int_{-\infty}^{+\infty} f^{\alpha+\beta-1}(x)dx}{\int_{-\infty}^{+\infty} f^{\beta}(x)dx}. \tag{3.1}$$

Two particular cases of this family are :

$$H_{\alpha,1}(X) = \frac{1}{1-\alpha} \log \int_{-\infty}^{+\infty} f^{\alpha}(x)dx, \tag{3.2}$$

which is the Renyi entropy of order α, and

$$H(X) = \lim_{\alpha\to 1} H_{\alpha,1}(X) = -\int_{-\infty}^{+\infty} f(x) \log f(x)dx, \tag{3.3}$$

is the Shannon entropy.

Theorem 3.1. *Let $X \sim N(\mu, \sigma^2)$ then*

$$H_{\alpha,\beta}(X) = \log \sigma\sqrt{2\pi} - \frac{\log(\alpha+\beta-1)}{2(1-\alpha)} + \frac{\log \beta}{2(1-\alpha)} \tag{3.4}$$

Proof.

$$\begin{aligned}
\int_{-\infty}^{+\infty} f^{\alpha+\beta-1}(x)dx &= \int_{-\infty}^{+\infty} \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{(\alpha+\beta-1)} e^{\frac{-(\alpha+\beta-1)(x-\mu)^2}{2\sigma^2}} dx \\
&= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{(\alpha+\beta-1)} \frac{\sigma}{\sqrt{\alpha+\beta-1}} \int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz \\
&= \frac{(\sigma\sqrt{2\pi})^{2-\alpha-\beta}}{\sqrt{\alpha+\beta-1}},
\end{aligned}$$

and similarly,

$$\int_{-\infty}^{+\infty} f^{\beta}(x)dx = \frac{(\sigma\sqrt{2\pi})^{1-\beta}}{\sqrt{\beta}}.$$

Then

$$\frac{\int_{-\infty}^{+\infty} f^{\alpha+\beta-1}(x)dx}{\int_{-\infty}^{+\infty} f^{\beta}(x)dx} = \frac{(\sigma\sqrt{2\pi})^{1-\alpha}(\alpha+\beta-1)^{-\frac{1}{2}}}{\sqrt{\beta}}.$$

By taking log and multiply in $\frac{1}{1-\alpha}$ the theorem is proved.

Corollary 3.2. *The Renyi and Shannon entropy for normal distribution are*

$$H_{\alpha,1}(X) = \log\sigma\sqrt{2\pi} - \frac{\log\alpha}{2(1-\alpha)},$$

and

$$H(X) = \lim_{\alpha\to 1} H_{\alpha,1}(X) = \log\sigma\sqrt{2\pi} + \frac{1}{2},$$

respectively.

Theorem 3.3. *The ADK entropy is invariant under a location and is not invariant under a scale transformation of the random variable i.e. , suppose that $a \neq 0$ then*

$$H_{\alpha,\beta}(aX+b) = \log|a| + H_{\alpha,\beta}(X) \tag{3.5}$$

Proof. Put $Y = aX + b$ then

$$\begin{aligned}\int_{-\infty}^{+\infty} f_Y^{\alpha+\beta-1}(y)dy &= (\frac{1}{a})^{(\alpha+\beta-1)}\int_{-\infty}^{+\infty} f_X^{\alpha+\beta-1}(\frac{y-b}{a})dy\\ &= a^{2-\alpha-\beta}\int_{-\infty}^{+\infty} f_X^{\alpha+\beta-1}(x)dx,\end{aligned}$$

and similarly,

$$\int_{-\infty}^{+\infty} f_Y^{\beta}(y)dy = a^{1-\beta}\int_{-\infty}^{+\infty} f_X^{\beta}(x)dx.$$

Then from (3.1) relation (3.5) is obtained.

Definition 3.4. The joint ADK entropy $H_{\alpha,\beta}(X_1, X_2)$ of a pair of continue random variables (X_1, X_2) with a joint distribution $f(x_1, x_2)$ is defined as

$$H_{\alpha,\beta}(X_1, X_2) = \frac{1}{1-\alpha}\log\frac{\int_{R^2} f^{\alpha+\beta-1}(x_1, x_2)dx_1dx_2}{\int_{R^2} f^{\beta}(x_1, x_2)dx_1dx_2}. \tag{3.6}$$

From this definition its obvious that if X_1 and X_2 are independent random variables then

$$H_{\alpha,\beta}(X_1, X_2) = H_{\alpha,\beta}(X_1) + H_{\alpha,\beta}(X_2).$$

Definition 3.5. If $(X_1, \ldots, X_n)$ be a random vector with density function $f(x_1, \ldots, x_n)$ then

$$H_{\alpha,\beta}(X_1, \ldots, X_n) = \frac{1}{1-\alpha} \log \frac{\int_{R^n} f^{\alpha+\beta-1}(x_1, \ldots, x_n) dx_1 \ldots dx_n}{\int_{R^n} f^{\beta}(x_1, \ldots, x_n) dx_1 \ldots dx_n}. \tag{3.7}$$

Definition 3.6. If $(X_1, X_2) \sim f(x_1, x_2)$ then the conditional ADK entropy of random variable X_2, given X_1, is defined as

$$H_{\alpha,\beta}(X_2|X_1) = \frac{1}{1-\alpha} \log \frac{\int_R f^{\beta}(x_1) dx_1 \int_{R^2} f^{\alpha+\beta-1}(x_1, x_2) dx_1 dx_2}{\int_R f^{\alpha+\beta-1}(x_1) dx_1 \int_{R^2} f^{\beta}(x_1, x_2) dx_1 dx_2}. \tag{3.8}$$

From this equation we get

$$H_{\alpha,\beta}(X_2|X_1) = H_{\alpha,\beta}(X_1, X_2) - H_{\alpha,\beta}(X_1).$$

For two independent random variables X_1 and X_2 we have

$$H_{\alpha,\beta}(X_2) - H_{\alpha,\beta}(X_2|X_1) = 0,$$

and we have

$$H_{\alpha,\beta}(X_2|X_1) = H_{\alpha,\beta}(X_2).$$

Definition 3.7. If $X_1, X_2, \ldots, X_n$ are distributed according $f(x_1, x_2, \ldots, x_n)$, then the conditional ADK entropy of random variable X_n, given $X_1, X_2, \ldots, X_{n-1}$, is defined as small

$$H_{\alpha,\beta}(X_n|X_1, X_2, \ldots, X_{n-1}) = \frac{1}{1-\alpha} \log$$
$$\frac{\int_{R^{n-1}} f^{\beta}(x_1, \ldots, x_{n-1}) dx_1 \ldots dx_{n-1} \int_{R^n} f^{\alpha+\beta-1}(x_1, \ldots, x_n) dx_1 \ldots dx_n}{\int_{R^{n-1}} f^{\alpha+\beta-1}(x_1, \ldots, x_{n-1}) dx_1 \ldots dx_{n-1} \int_{R^n} f^{\beta}(x_1, \ldots, x_n) dx_1 \ldots dx_n}. \tag{3.9}$$

Theorem 3.8. ***(Chain rule)*** *Let $(X_1, \ldots, X_n)$ be a random vector with density function $f(x_1, \ldots, x_n)$ and have finite ADK entropy for every n, then*

$$H_{\alpha,\beta}(X_1, \ldots, X_n) = \sum_{i=1}^{n} H_{\alpha,\beta}(X_i|X_{i-1}, \ldots, X_1). \tag{3.10}$$

Proof. We can write

$$\frac{\int_{R^n} f^{\alpha+\beta-1}(x_1,\dots,x_n)dx_1\dots dx_n}{\int_{R^n} f^{\beta}(x_1,\dots,x_n)dx_1\dots dx_n}$$

$$= \frac{\int_R f^{\alpha+\beta-1}(x_1)dx_1}{\int_R f^{\beta}(x_1)dx_1}$$

$$\times \frac{\int_R f^{\beta}(x_1)dx_1 \int_{R^2} f^{\alpha+\beta-1}(x_1,x_2)dx_1dx_2}{\int_R f^{\alpha+\beta-1}(x_1)dx_1 \int_{R^2} f^{\beta}(x_1,x_2)dx_1dx_2}$$

$$\times \frac{\int_{R^2} f^{\beta}(x_1,x_2)dx_1dx_2 \int_{R^3} f^{\alpha+\beta-1}(x_1,x_2,x_3)dx_1dx_2dx_3}{\int_{R^2} f^{\alpha+\beta-1}(x_1,x_2)dx_1dx_2 \int_{R^3} f^{\beta}(x_1,x_2,x_3)dx_1dx_2dx_3}$$

$$\times \quad \dots \times$$

$$\frac{\int_{R^{n-1}} f^{\beta}(x_1,\dots,x_{n-1})dx_1\dots dx_{n-1} \int_{R^n} f^{\alpha+\beta-1}(x_1,\dots,x_n)dx_1\dots dx_n}{\int_{R^{n-1}} f^{\alpha+\beta-1}(x_1,\dots,x_{n-1})dx_1\dots dx_{n-1} \int_{R^n} f^{\beta}(x_1,\dots,x_n)dx_1\dots dx_n}$$

Now, multiplying both sides of the above relation through $\frac{1}{1-\alpha}$ and taking the log, the desired result is obtained.

4 ADK entropy rate

Roughly speaking, the entropy rate quantifies the limiting average of uncertainty, disorder or irregularity generated by a process or system per time unit.

Definition 4.1. The regular relation for ADK entropy rate of a discrete-time stochastic process $\{X_n\}_{n\in\mathbb{Z}}$ is defined as

$$\overline{H}_{\alpha,\beta} = \lim_{n\to\infty} \frac{H_{\alpha,\beta}(X_1,\dots,X_n)}{n}, \tag{4.1}$$

when the limit exists. Furthermore, by (2.4) we have

$$\overline{H}_{\alpha,\beta} = \lim_{n\to\infty} \frac{\sum_{i=1}^{n} H_{\alpha,\beta}(X_i|X_{i-1},\dots,X_1)}{n}, \tag{4.2}$$

when the limit exists.

Remark 4.2. If the limits exists, then by Cesaro mean

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} H_{\alpha,\beta}(X_i|X_{i-1},\dots,X_1)}{n} = \lim_{n\to\infty} H_{\alpha,\beta}(X_n|X_{n-1},\dots,X_1) \tag{4.3}$$

A discrete-time stochastic process is said to be stationary if the joint distribution of any subset of the sequence of random variables is invariant with respect to shifts in the time index, i.e., if the distribution of $(X_{n_1+h}, \ldots, X_{n_k+h})$ is independent of h for any positive integer k and $n_1, \ldots, n_k \in \mathbb{Z}$. Hence we can write

$$H_{\alpha,\beta}(X_{n_k}|X_{n_1}, \ldots, X_{n_{k-1}}) = H_{\alpha,\beta}(X_{n_k+h}|X_{n_1+h}, \ldots, X_{n_{k-1+h}}).$$

Then the ADK entropy rate of a stationary process is equal to

$$\overline{H}_{\alpha,\beta} = \lim_{n\to\infty} H_{\alpha,\beta}(X_1|X_0, \ldots, X_{2-n}) = H_{\alpha,\beta}(X_1|X_0, \ldots) \quad (4.4)$$

Theorem 4.3. *The ADK entropy rate of an ergodic Markov chain with a infinite state space is*

$$\overline{H}_{\alpha,\beta} = \frac{1}{1-\alpha}(\log R^{-1} + \log \tilde{R}^{-1}), \quad (4.5)$$

where, R is the convergence radius of the matrix $M = [p_{x_i x_j}^{\alpha+\beta-1}]_{x_i,x_j\in S}$ and $\tilde{R}$ is the convergence radius of the matrix $\tilde{M} = [p_{x_i x_j}^{\beta}]_{x_i,x_j\in S}$.

Proof. From Markov chain properties and (2.3) we have

$$H_{\alpha,\beta}(X_n|X_1, X_2, \ldots, X_{n-1}) = \frac{1}{1-\alpha}\log \quad (4.6)$$
$$\frac{\sum_{x_1,x_2,\ldots,x_{n-1}} p_{x_1}^{\beta} p_{x_1x_2}^{\beta} \ldots p_{x_{n-2}x_{n-1}}^{\beta} \sum_{x_1,x_2,\ldots,x_n} p_{x_1}^{\alpha+\beta-1} p_{x_1x_2}^{\alpha+\beta-1} p_{x_{n-1}x_n}^{\alpha+\beta-1}}{\sum_{x_1,x_2,\ldots,x_{n-1}} p_{x_1}^{\alpha+\beta-1} p_{x_1x_2}^{\alpha+\beta-1} \ldots p_{x_{n-2}x_{n-1}}^{\alpha+\beta-1} \sum_{x_1,x_2,\ldots,x_n} p_{x_1}^{\beta} p_{x_1x_2}^{\beta} p_{x_{n-1}x_n}^{\beta}}.$$

By defining two row vectors $U = [p_{x_i}^{\alpha+\beta-1}]_{x_i\in S}$ and $\tilde{U} = [p_{x_i}^{\beta}]_{x_i\in S}$ a column vector $\mathbf{1}$ we have (see [10, 15])

$$H_{\alpha,\beta}(X_n|X_1, X_2, \ldots, X_{n-1}) = \frac{1}{1-\alpha}\log[\frac{\mathbf{UM^{n-1}1}}{\mathbf{UM^{n-2}1}} \times \frac{\mathbf{\tilde{U}\tilde{M}^{n-2}1}}{\mathbf{\tilde{U}\tilde{M}^{n-1}1}}]. \quad (4.7)$$

Suppose that $R_{ij} = Sup_{z\geq 0}\{z : \sum_{k=0}^{\infty} M_{ij}^k z^k < \infty\}$ and $\tilde{R}_{ij} = Sup_{z\geq 0}\{z : \sum_{k=0}^{\infty} \tilde{M}_{ij}^k z^k < \infty\}$ are convergence radius for generating functions of $\mathbf{M}$ and $\tilde{\mathbf{M}}$ respectively. Since, this chain is irreducible, the matrices $\mathbf{M}$ and $\tilde{\mathbf{M}}$ are also irreducible and by Theorem1 of the chapter 6 of Seneta's book [15], these matrices have common convergence radiuses R and $\tilde{R}$ respectively, where $0 < R, \tilde{R} < 1$. Now by using this and taking the limit as $n \to \infty$ and making use of the fact that all of the assumptions of Theorem 5 of chapter 6 of Seneta's

book [15] hold, relation (4.5) is obtained.
Suppose that for all n, $E(X_n) = 0$. Consider autocovariance function $\gamma(k) = E(X_n X_{n+k})$ and the spectral density function $f(w)$ as

$$f(w) = \frac{1}{2\pi}\Sigma_{k=-\infty}^{\infty}\gamma(k)e^{-iwk}, \quad -\pi \leq w \leq \pi. \tag{4.8}$$

For stationary Gaussian processes we have the following representation [3, 7]

$$X_n = \Sigma_{j=0}^{\infty}\varphi_j Z_{n-j} \;\;, \quad n \in \mathbb{Z}, \tag{4.9}$$

where φ_j with $j \geq 0$ are constant, $\Sigma_{j=0}^{\infty}\varphi_j^2 < \infty$, and $\{Z_n\}$ is a sequence of independent Gaussian random variables with identical distribution $N(0,1)$. This expression is known as the moving average representation of the process.

Theorem 4.4. *For stationary Gaussian processes, the ADK entropy rate is equal to*

$$\overline{H}_{\alpha,\beta} = \log\sigma\sqrt{2\pi} - \frac{\log(\alpha+\beta-1)}{2(1-\alpha)} + \frac{\log\beta}{2(1-\alpha)} + \frac{1}{4\pi}\int_{-\pi}^{\pi}\log 2\pi f(w)dw. \tag{4.10}$$

Proof. By relations (4.4) and (4.9) we have

$$\overline{H}_{\alpha,\beta} = H_{\alpha,\beta}(\Sigma_{j=0}^{\infty}\varphi_j Z_{1-j}|X_0,\ldots) = H_{\alpha,\beta}(\varphi_0 Z_1 + \Sigma_{j=1}^{\infty}\varphi_j Z_{1-j}|X_0,\ldots). \tag{4.11}$$

We know that

$$\sigma\{Z_k,\ k \leq 1\} = \sigma\{X_k,\ k \leq 1\},$$

and since Z_1 is independent of $\{Z_k,\ k \leq 0\}$, then Z_1 is independent of $\{X_k,\ k \leq 0\}$. Since $\Sigma_{j=1}^{\infty}\varphi_j Z_{1-j}$ is measurable with respect to $\sigma\{X_k,\ k \leq 0\}$, is constant, and by ((3.5) we have

$$\overline{H}_{\alpha,\beta} = H_{\alpha,\beta}(\varphi_0 Z_1) = H_{\alpha,\beta}(Z_1) + \log|\varphi_0|.$$

Finally we can use (3.4) to get

$$\overline{H}_{\alpha,\beta} = \log\sigma\sqrt{2\pi} - \frac{\log(\alpha+\beta-1)}{2(1-\alpha)} + \frac{\log\beta}{2(1-\alpha)} + \frac{1}{2}\log\varphi_0^2, \tag{4.12}$$

where φ_0 is constant. On one hand, from [7] we have

$$\varphi_0^2 = 2\pi 2^{\frac{1}{2\pi}\int_{-\pi}^{\pi}\log f(w)dw}. \tag{4.13}$$

Consequently, by (4.12) and (4.13) we get (4.10).

Corollary 4.5. *Let $\{X_n\}$ be an autoregressive process of order p, i.e.,*

$$X_n = \Sigma_{j=0}^{p} \pi_j X_{n-j} + Z_n,$$

or, a moving average process of order q with $\phi_0 = 1$, i.e.,

$$X_n = \Sigma_{j=0}^{q} \phi_j Z_{n-j},$$

for these processes [7], $\int_{-\pi}^{\pi} \log 2\pi f(w)dw = 0$. Then from (4.10) the ADK entropy rate for these Gaussian stationary processes is equal to ADK entropy rate for normal distribution, i.e.

$$\overline{H}_{\alpha,\beta} = \log \sigma\sqrt{2\pi} - \frac{\log(\alpha+\beta-1)}{2(1-\alpha)} + \frac{\log \beta}{2(1-\alpha)}. \tag{4.14}$$

Its reminded that [7] in time series analysis we often use $X_n = \Sigma_{j=0}^{\infty} \varphi_j Z_{n-j}$ form with $\varphi_0 = 1$ which, also has ADK entropy rate as normal distribution.

5 Conclusion

In this paper, I considered a definition for ADK entropy and some correlated concepts. Here, demonstrated that the chain rule holds for this definition. Furthermore, two relations for the rate of ADK entropy and stationary Gaussian processes was obtained. We used this relation to obtain the rate of ADK entropy for an irreducible-aperiodic Markov chain. Also we showed that the rate of stationary Gaussian processes depends on the spectral density function of the processes. In other words, whatever carried out here is an extension of [10, 6].

References

[1] Aczel J. and Daroczy, Z. (1963), Characterisierung der entropien positiver ordnung und der Shannons chen entropie. Act. Math. Acad. Sci. Hunger, **14**, 95-121.

[2] Arndt, C. (2001), Information Measures. Berlin: Springer.

[3] Ash, R. B. and Gardner, M. F. (1975), Topics in Stochastic Processes. Academic Press, Inc.

[4] Cover, T. and Thomas, J. A. (2006), Elements of Information Theory, New York: Wiley.

[5] Girardin, V. and Limnios, N. (2004), Entropy rate and maximum entropy methods for countable Semi-Markov chains. Communications in Statistics: Theory and Methods, **33**, 609622.

[6] Golshani, L., Pasha, E., and Yari, G. (2009), Some properties of Rnyi entropy and Rnyi entropy rate. Information Sciences **179** 2426-2433.

[7] Ihara, S. (1993), Information Theory for Continuous Systems. Singapore: World Scientific.

[8] Kapur, J. N. (1967), Generalized entropy of order and type. The Math. Seminar, **4**, 78-82.

[9] Kolmogorov, A. N. (1958), A new invariant for transitive dynamical systems. Dokl. An. SSR, **119**, 861864.

[10] Pasha, E. and Golshani, L., The Rnyi entropy rate for Markov chains with countable state space. Bulletin of the Iranian Mathematical Society (submitted for publication).

[11] Pfister, C. E. and Sullivan, W. G., (2004), Renyi entropy, guesswork moments, and large deviations. IEEE Trans. Inform. Theory, **50**(11) 2794-2800.

[12] Rached, Z., Fady, A., and Campbell, L. L. (1999), Rnyi's entropy rate for discrete Markov sources. In: Proceedings of the CISS'99, Baltimore, MD, March, 17- 19.

[13] Renyi, A. (1961), On measures of entropy and information. In: Proceedings of the Fourth Berkeley Symposium on Mathematics, Statistics and Probability, **1**, Berkeley: University California Press, 547-561.

[14] Renyi, A. (1965), On the foundation of information theory. Rev. Inst. Int. Stat., **33**, 1-14.

[15] Seneta, E. (1981), Nonnegative Matrix and Markov Chains. New York: Springer-Verlag.

[16] Shannon, C. E. (1984), A mathematical theory of communication. Bell Syst. Tech. J., **27**, 379-423, 623-656.

[17] Wei, W. W. S. (1990), Time Series Analysis Univariate and Multivariate Methods. Addison-Wesley.

[18] Wimmer, H. K. (1974), Spectral radius and radius convergence. American Mathematical Monthly, **81**, 625-627.