



Soviet-era science, translated into English

M. S. PINSKER

1960

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196001.38423>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

MATHEMATICS

M. S. PINSKER

ENTROPY, THE RATE OF CREATION OF ENTROPY, AND ENTROPIC STABILITY OF GAUSSIAN RANDOM VARIABLES AND PROCESSES

(Presented by Academician A. N. Kolmogorov, 7 III 1960)

1° Entropy. Let $\mathcal{P}_1\{\cdot\}$ and $\mathcal{P}_2\{\cdot\}$ be probability measures defined on one and the same space (X, S_X) . If the measure $\mathcal{P}_1\{\cdot\}$ is absolutely continuous with respect to the measure $\mathcal{P}_2\{\cdot\}$, then put

$$h_{\mathcal{P}_2}(x, \mathcal{P}_1) = \ln a(x) = \ln \frac{d\mathcal{P}_1\{\cdot\}}{d\mathcal{P}_2\{\cdot\}}, \quad H_{\mathcal{P}_2}(\mathcal{P}_1) = \int_X h_{\mathcal{P}_2}(x, \mathcal{P}_1) \mathcal{P}_1\{dx\}. \quad (1)$$

If, however, the measure $\mathcal{P}_1\{\cdot\}$ is not absolutely continuous with respect to the measure $\mathcal{P}_2\{\cdot\}$, then put $H_{\mathcal{P}_2}(\mathcal{P}_1) = \infty$, and in the case of singularity of the measure $\mathcal{P}_1\{\cdot\}$ with respect to the measure $\mathcal{P}_2\{\cdot\}$, $h_{\mathcal{P}_2}(x, \mathcal{P}_1) = \infty$, $h_{\mathcal{P}_2}(x, \mathcal{P}_1)/H_{\mathcal{P}_2}(\mathcal{P}_1) = 1$.

We shall call ⁽¹⁻³⁾ $H_{\mathcal{P}_2}(\mathcal{P}_1)$, $h_{\mathcal{P}_2}(x, \mathcal{P}_1)$, respectively, the **entropy** and the **entropy density** of the measure $\mathcal{P}_1\{\cdot\}$ with respect to the measure $\mathcal{P}_2\{\cdot\}$. (These notions, under other names, have previously occurred in works on mathematical statistics ⁽⁴⁾.) In the case where $\mathcal{P}_1\{\cdot\}$, $\mathcal{P}_2\{\cdot\}$ are measures determined by the distributions $P_\xi(E) = \mathcal{P}\{\xi \in E\} = \mathcal{P}_1\{E\}$, $P_\eta(E) = \mathcal{P}\{\eta \in E\} = \mathcal{P}_2\{E\}$, $E \in S_x$, of random variables ξ and η taking values in one and the same measurable space (X, S_X) , we shall adopt the notation

$$H_{\mathcal{P}_2}(\mathcal{P}_1) = H_{P_\eta}(P_\xi) = H_\eta(\xi), \quad h_{\mathcal{P}_2}(x, \mathcal{P}_1) = h_{P_\eta}(x, P_\xi) = h_\eta(x, \xi).$$

In this case the expression $h_\eta(\xi, \xi) = h_\eta(\xi)$ is a random variable and $H_\eta(\xi) = Mh_\eta(\xi)$.

A random variable $\xi = \{\xi_t\}$, $t \in N$ (N is some set), formed from one-dimensional random variables ξ_t , is called **Gaussian** if, for any finite set $t_1, \dots, t_n \in N$ of values of the parameter t , the joint distribution of the random variables $\xi_{t_1}, \dots, \xi_{t_n}$ is normal.

Let $\xi = (\xi_1, \dots, \xi_n)$ and $\eta = (\eta_1, \dots, \eta_n)$ be n -dimensional Gaussian random variables. J. Hájek proposed the following method for computing $H_\eta(\xi)$. In ⁽⁵⁾ a nondegenerate linear transformation of the random variables $\xi_1, \dots, \xi_n; \eta_2, \dots, \eta_n$ was considered,

$$\xi'_j = \sum_{k=1}^n c_{jk} \xi_k,$$

$$\eta'_j = \sum_{k=1}^n c_{jk} \eta_k$$

such that the Gaussian random variables $\xi' = (\xi'_1, \dots, \xi'_n)$, $\eta' = (\eta'_1, \dots, \eta'_n)$ consist of mutually independent random variables $\xi'_1, \dots, \xi'_n; \eta'_1, \dots, \eta'_n$. Then, obviously,

$$H_\eta(\xi) = H_{\eta'}(\xi') = \sum_{j=1}^n H_{\eta'_j}(\xi'_j) = \frac{1}{2} \sum_{j=1}^n \left[\frac{\sigma_{\xi'_j}^2 + (m_{\xi'_j} - m_{\eta'_j})^2}{\sigma_{\eta'_j}^2} - 1 - \ln \frac{\sigma_{\xi'_j}^2}{\sigma_{\eta'_j}^2} \right],$$

where $\sigma_{\xi'_j}^2 = D\xi'_j$, $\sigma_{\eta'_j}^2 = D\eta'_j$.

* Here and below D denotes the variance of a random variable.

The method considered for computing entropy can be generalized to arbitrary Gaussian random variables.

Associate with the random variable $\beta = \sum_{j=1}^n c_j \eta_{t_j}$ the random variable $\theta(\beta) = \sum_{j=1}^n c_j \xi_{t_j}$. The entropy $H_\eta(\xi)$ is finite if and only if the following set of conditions is satisfied:

- 1) The correspondence $\theta(\beta)$ can be extended to a linear operator carrying out a one-to-one mapping of the unitary space B_ξ —the closed linear hull of the random variables ξ_t , $t \in N$, onto the unitary space B_η —the closed linear hull of the random variables η_t , $t \in N$.
- 2) There exists a finite or countable sequence of independent random variables $\eta'_1, \eta'_2, \dots \in B_\eta$ such that the sequence $\xi'_1 = \theta(\eta'_1)$, $\xi'_2 = \theta(\eta'_2)$, ... is a sequence of independent random variables and

$$H_\eta(\xi) = \sum_j H_{\eta'_j}(\xi'_j), \quad h_\eta(\xi) = \sum_j h_{\eta'_j}(\xi'_j). \quad (2)$$

The entropies $H_\eta(\xi)$ and $H_\xi(\eta)$ are simultaneously finite or infinite.

2°. **Rate of creation of entropy.** Let $\xi = \{\xi(\cdot)\}$ and $\eta = \{\eta(\cdot)\}$ be random processes, generalized or non-generalized; let ξ_0^T, η_0^T be segments of these processes formed from the random variables $\xi(t), \eta(t)$, $0 \leq t \leq T$, if ξ and η are non-generalized processes, and from $\xi(\varphi), \eta(\varphi)$, $\varphi(t) \in \Phi$, $\varphi(t) = 0$ for $t \notin [0, T]$, if ξ and η are generalized processes. The quantity

$$\bar{H}_\eta(\xi) = \lim_{T \rightarrow \infty} \frac{1}{T} H_{\eta_0^T}(\xi_0^T), \quad (3)$$

defined only in the case when the right-hand side of this formula exists, will be called the **rate of creation of entropy of the process** ξ with respect to the process η .

Let $\xi = \{\xi(\cdot)\}$ and $\eta = \{\eta(\cdot)\}$ be one-dimensional stationary in the broad sense random processes; put

$$\mathcal{E}_{\xi\eta} = \frac{1}{2\pi} \int \left(\frac{f_{\xi\xi}(\lambda)}{f_{\eta\eta}(\lambda)} - 1 - \ln \frac{f_{\xi\xi}(\lambda)}{f_{\eta\eta}(\lambda)} \right) d\lambda, \quad (4)$$

where the integral is taken over the limits from 0 to π , if ξ and η are random processes with discrete argument, and over the limits from 0 to ∞ , if ξ and η are random processes with continuous argument or generalized random processes; $f_{\xi\xi}(\lambda) = F'_{\xi\xi}(\lambda)$, $f_{\eta\eta}(\lambda) = F'_{\eta\eta}(\lambda)$, and $F_{\xi\xi}(\lambda)$ and $F_{\eta\eta}(\lambda)$ are the spectral functions of the processes ξ and η .

If $\xi(\cdot) = \eta(\cdot) + \nu(\cdot)$, where $\nu = \{\nu(\cdot)\}$ is a Gaussian stationary random process independent of η , then formula (4) takes the form

$$\mathcal{E}_{\xi\eta} = \frac{1}{2\pi} \int \left(\frac{f_{\nu\nu}(\lambda)}{f_{\eta\eta}(\lambda)} - \ln \left(1 + \frac{f_{\nu\nu}(\lambda)}{f_{\eta\eta}(\lambda)} \right) \right) d\lambda. \quad (5)$$

Theorem 1. *Let ξ and η be one-dimensional Gaussian stationary random processes, generalized or non-generalized, and $M\xi(\cdot) = M\eta(\cdot) = 0$. Then:*

1)

$$\bar{H}_\eta(\xi) = \mathcal{E}_{\xi\eta} \quad (6)$$

in the following cases: a) $\mathcal{E}_{\xi\eta} = \infty$; b) ξ and η are regular random processes with discrete argument and the function $f_{\xi\xi}(\lambda)/f_{\eta\eta}(\lambda)$ is bounded above; c) ξ and η are random processes with continuous argument or generalized random processes with rational spectral densities.

2) If η is a random process of continuous argument or a generalized random process with rational spectral densities and $f_{\eta\eta}(\lambda) \neq 0$, then from the

equality $\mathcal{E}_{\xi\eta} = \infty$ it follows that, for any $T > 0$, the measures $P_{\xi_0}^T$ and $P_{\eta_0}^T$ are singular with respect to one another.

In the case when ξ and η have rational spectral densities, the assertion of item 2) of the theorem follows in an obvious way from the corresponding result of Slepian (6).

In writing out the formula generalizing (6) to multidimensional random processes, we shall use the result of the following lemma, which is of known independent interest.

Lemma. Let $\xi = (\xi_1, \dots, \xi_n)$, $\eta = (\eta_1, \dots, \eta_n)$ be n -dimensional Gaussian stationary random processes, generalized or non-generalized. Then there exist n -dimensional Gaussian stationary random processes $\xi' = (\xi'_1, \dots, \xi'_n)$, $\eta' = (\eta'_1, \dots, \eta'_n)$ such that:

- 1) The random processes $\xi'_1, \dots, \xi'_n; \eta'_1, \dots, \eta'_n$ are mutually independent.
- 2) ξ and ξ' , η and η' are stationary-related and $B_\xi = B_{\xi'}$, $B_\eta = B_{\eta'}$.
- 3)

$$F_{\xi'_j \xi_k}(\lambda) = \sum_{l=1}^n \psi_{jl}(\lambda) F_{\xi_l \xi_k}(\lambda), \quad F_{\eta'_j \eta_k}(\lambda) = \sum_{l=1}^n \psi_{jl}(\lambda) F_{\eta_l \eta_k}(\lambda),$$

where $F_{\xi'_j \xi_k}(\lambda)$, $F_{\eta'_j \eta_k}(\lambda)$ are the mutual spectral functions of the processes ξ'_j and ξ_k , η'_j and η_k ; $\psi_{jl}(\lambda)$, $j, l = 1, \dots, n$, are measurable functions.

Put

$$\mathcal{E}_{\xi\eta} = \mathcal{E}_{\xi'\eta'} = \sum_{j=1}^n \mathcal{E}_{\xi'_j \eta'_j} = \frac{1}{2\pi} \int \sum_{j=1}^n \left(\frac{f_{\xi'_j \xi'_j}(\lambda)}{f_{\eta'_j \eta'_j}(\lambda)} - 1 - \ln \frac{f_{\xi'_j \xi'_j}(\lambda)}{f_{\eta'_j \eta'_j}(\lambda)} \right) d\lambda, \quad (7)$$

where the limits of integration are determined exactly as in formula (4).

For those values of λ for which

$$\det \|f_{\eta_j \eta_k}(\lambda)\| \neq 0, \quad \det \|f_{\xi_j \xi_k}(\lambda)\| \neq 0, \quad j, k = 1, \dots, n,$$

the integrand in (7) is equal to

$$\sum_{j,k=1}^n f_{\xi_j \xi_k}(\lambda) f_{\eta_k \eta_j}^{(-1)}(\lambda) - n - \ln \frac{\det \|f_{\xi_j \xi_k}(\lambda)\|}{\det \|f_{\eta_j \eta_k}(\lambda)\|}, \quad (8)$$

where $f_{\eta_k \eta_j}^{(-1)}(\lambda)$ are the elements of the matrix inverse to the matrix $\|f_{\eta_j \eta_k}(\lambda)\|$.

In the case $n = 1$, formula (7) passes into formula (4).

Theorem 2. Let $\xi = (\xi_1, \dots, \xi_n)$, $\eta = (\eta_1, \dots, \eta_n)$ be n -dimensional Gaussian stationary random processes, generalized or non-generalized, for which $M\xi_j(\cdot) = M\eta_j(\cdot) = 0$.

Then:

- 1) In the following cases formula (6) holds: a) $\mathcal{E}_{\xi\eta} = \infty$; b) ξ and η are regular random processes of discrete argument of rank n , and the function

$$\sum_{j,k=1}^n f_{\xi_j\xi_k}(\lambda) f_{\eta_k\eta_j}^{(-1)}(\lambda)$$

is bounded above; c) ξ and η are random processes of continuous argument or generalized random processes with rational spectral and mutual spectral functions.

- 2) If η is a random process of continuous argument or a generalized random process with rational spectral and mutual densities, and each minor of the matrix $\|f_{\eta_j\eta_k}\|_{j,k=1,\dots,n}$ is either identically equal to zero or everywhere different from zero, then from the equality $\mathcal{E}_{\xi\eta} = \infty$ it follows that, for any $T > 0$, the measures $P_{\xi_0}^T$ and $P_{\eta_0}^T$ are singular with respect to one another.

Remark. It is easy to show that always

$$\underline{\lim} \frac{1}{T} H_{\xi_0^T}(\eta_0^T) \geq \mathcal{E}_{\xi\eta}.$$

It would be desirable to find the most general conditions imposed on the spectral functions of the processes ξ and η under which Theorems 1 and 2 remain valid.

3°. Entropic stability. Let ξ^t and η^t be random variables taking values in one and the same measurable space (X^t, S_{X^t}) . We shall call a sequence of random variables $\{\xi^t\}$, $t = t_1, t_2, \dots$, $\lim_{n \rightarrow \infty} t_n = \infty$, **entropically stable with respect to the sequence** $\{\eta^t\}$, $t = t_1, t_2, \dots$, if, in probability,

$$\lim_{t \rightarrow \infty} (h_{\eta^t}(\xi^t)/H_{\eta^t}(\xi^t)) = 1.$$

Theorem 3. *In order that the sequence of Gaussian random variables $\{\xi^t\}$, $t = t_1, t_2, \dots$, be entropically stable with respect to the sequence of random variables $\{\eta^t\}$, $t = t_1, t_2, \dots$, it is necessary and sufficient that the following collection of conditions be satisfied:*

$$\lim_{t \rightarrow \infty} H_{\eta^t}(\xi^t) = \infty, \quad \lim_{t \rightarrow \infty} (\mathbf{D}h_{\eta^t}(\xi^t)/(H_{\eta^t}(\xi^t))^2) = 0. \quad (9)$$

Moreover, if

$$\lim_{t \rightarrow \infty} \mathbf{D}h_{\eta^t}(\xi^t) = \infty, \quad (10)$$

then as $t \rightarrow \infty$ the information density $h_{\eta^t}(\xi^t)$ is asymptotically normal with the usual normalization.

We shall call a random process $\xi = \{\xi(\cdot)\}$ **entropically stable with respect to the random process** $\eta = \{\eta(\cdot)\}$, if either $\bar{H}_\eta(\xi) = 0$, or every sequence of random variables $\xi_0^{t_1}, \xi_0^{t_2}, \dots, \lim t_n = \infty$, is entropically stable with respect to the sequence of random variables $\eta_0^{t_1}, \eta_0^{t_2}, \dots$.

Theorem 4. *Under the conditions of items 1) and 2) of Theorems 1 and 2, the random process $\xi = \{\xi(\cdot)\}$ is entropically stable with respect to the random process $\eta = \{\eta(\cdot)\}$. Moreover, if $\mathcal{E}_{\xi\eta} < \infty$, then, when the condition of item 1) of Theorems 1 and 2 is fulfilled,*

$$\lim_{t \rightarrow \infty} \frac{1}{t} \mathbf{D}h_{\eta_0^t}(\xi_0^t) = \frac{1}{2\pi} \sum_{j=1}^n \int \left(\frac{f_{\xi_j \xi_j}^{(\lambda)}}{f_{\eta_j \eta_j}^{(\lambda)}} - 1 \right)^2 d\lambda, \quad (11)$$

where the limits of integration are determined exactly as in formula (4), and as $t \rightarrow \infty$ the information density $h_{\eta_0^t}(\xi_0^t)$ is asymptotically normal with the usual normalization.

Remark. The results of Theorems 1, 2, and 3 can be generalized without particular difficulty to a much broader class of random processes, in particular to processes of the form $\xi(t) = \xi'(t) + b(t)$, where $\xi' = \{\xi'(t)\}$ is a Gaussian stationary random process, and $b(t)$ is a nonrandom function.

Received
7 III 1960

References

1. A. Rényi, Trans. First Prague Conference of Information Theory, Statistical Decision Function, Random Processes, Prague, 1958, p. 183.
2. A. Perez, Theory Probab. and Its Appl., 4, no. 1, 105 (1959).
3. M. Rosenblatt-Roth, DAN, 112, no. 1, 16 (1957).
4. S. Kullback, R. A. Leibler, Ann. Math. Stat., 22, 79 (1951).
5. Ya. Gaek, Czechoslovak Math. Journal, 8 (83), 610 (1958).
6. D. Slepian, Trans. I.R.E., Sect. Inf. Theory, 4, no. 2, 65 (1958).

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.