ON MAXIMUM-ENTROPY AND RELATED PRINCIPLES IN STATISTICAL EQUILIBRIUM

C.G. CHAKRABARTI and V. MUKHOPADHYAY
Department of Applied Mathematics
University of Calcutta
Calcutta - 700009, India
(Received May 25, 1988)

ABSTRACT. The paper aims to study the importance and equivalence of the principles of maximum-entropy, and sufficient and stable inferences in the statistical characterization of the thermal equilibrium of a closed system.

KEY WORDS AND PHRASES. Maximum-entropy, stable and sufficient inferences, $\sigma$-algebra, statistical equilibrium.

1980 AMS SUBJECT CLASSIFICATION CODE. 82A05.

1. INTRODUCTION.

The maximum-entropy principle is a well-known procedure of estimating an unknown probability distribution on the basis of partial information or data. This was first formulated by Jaynes [1] in the subjective formalism of statistical mechanics and it has found wide applications in different branches of science [2]. There are different estimation principles, but they are not so well known in physics, as the method of maximum entropy principle [3]. Tikochinsky et al [4] proposed three other alternative approaches which lead to the same operational procedure as the maximum-entropy principle.

The object of the present paper is to study the importance of the principles of maximum-entropy, and sufficient and stable inferences in the statistical characterization of thermal equilibrium of a closed system. It also studies the equivalence of these principles in the context of statistical equilibrium of a closed system.

2. CLOSED SYSTEM: ENTROPY.

Let us consider a classical system whose state-space is the phase-space $\Sigma$. The points of the phase-space represent the microstates of the system. Let $p(\omega)$ be the probability density of microstates for any non-equilibrium state of the system and $p^0(\omega)$ be that for the stationary equilibrium state of the closed system. The entropy of the system at the non-equilibrium state $p(\omega)$ relative to the equilibrium state $p^0(\omega)$ is defined by the functional [5]
where the integration is with respect to the Liouville's measure $m$ of the phase-space. According to Kullback [6] $S(\rho|\rho^0)$ gives a measure of directed divergence between the equilibrium and non-equilibrium states. The relative entropy $S(\rho|\rho^0)$ is the basic starting concept and for the case of closed system it is more appropriate than that of the simple entropy [5]. The macroscopic description of the system is realized by the countable set of measurable non-overlapping phase-cells or partition $\pi = \{\Gamma_a\}$ which corresponds to the different observational values of the energy $X(\omega)$ of the system. Let $\mathcal{A}$ be the $\sigma$-algebra formed by the partition $\{\Gamma_a\}$ including the null set $\varnothing$. The coarse-grained probability density is then defined by

$$P_{\alpha} = \frac{1}{m_a} \int_{\Gamma_a} p(\omega) dm(\omega) = \frac{\rho_a}{m_a}, \quad m_a = m(\Gamma_a)$$

where $\rho_a$ are the probability distributions of the partition $\Gamma_a$. Then the relative entropy for the coarse-grained distribution is defined by

$$\tilde{S}(\rho|\rho^0) = -k \int_{\Gamma} \frac{\rho}{\rho^0} dm(\omega) = -k \int_{\Gamma} p(\omega) \ln \frac{\rho}{\rho^0} dm(\omega)$$

or

$$\tilde{S}(\rho|\rho^0) = -k \sum_{\alpha} \left( \int_{\Gamma_a} p(\omega) dm(\omega) \right) \ln \frac{\rho_a}{\rho^0} = -k \sum_{\alpha} P_a \ln \frac{\rho_a}{\rho^0}$$

3. MAXIMUM-ENTROPY AND SUFFICIENT PRINCIPLES.

Let us consider the statistical equilibrium of the closed system with the environment. The statistical equilibrium is completely characterized by the average energy $\langle X(\omega) \rangle = E_o$ of the system. According to the generalization of the maximum-entropy principle [6], the probability distribution of statistical equilibrium will correspond to the maximum of the relative entropy $S(\rho|\rho^0)$ subject to the condition (or information) $\langle X(\omega) \rangle = E_o$.

Let

$$S^*(\lambda) = \text{Max} S(\rho|\rho^0) \{ p(\omega) | \langle X(\omega) \rangle = E_o \}$$

In the process of maximization, we have [6]

$$S(\rho|\rho^0) < \tilde{S}(\rho|\rho^0) < S^*(\lambda)$$

The equality between the first two members of (3.2) holds if and only if [6]

$$\frac{\rho(\omega)}{\rho} = \frac{\rho^0(\omega)}{\rho^0} \quad \text{or} \quad \frac{p(\omega)}{p_a} = \frac{\rho^0(\omega)}{\rho^0_a}, \quad \text{for all} \ \Gamma_a \in \Gamma$$

which shows the constancy of the conditional probability density of microstates given the partition $\{\Gamma_a\}$. From the statistical point of view this implies the sufficiency of the $\sigma$-algebra $\mathcal{A}$ for the macroscopic states of the system. Under this criteria of
sufficiency, the conditional expectation of the energy \( X(\omega) \) given the partition
\( \{ \Gamma_\alpha \} \) or \( \sigma \)-algebra \( \mathcal{A} \), remains invariant, that is,
\[
\langle X(\omega) | \mathcal{A} \rangle \rho(\omega) = \frac{1}{\mathcal{A}} \int_{\Gamma_\mathcal{A}} \rho(\omega) X(\omega) \, d\omega
\]
\[
= \frac{1}{\mathcal{A}} \int_{\mathcal{A}} \rho^0(\omega) X(\omega) \, d\omega
\]
\[
= \langle X(\omega) | \mathcal{A} \rangle \rho^0(\omega) \quad .
\] (3.4)

For the statistical equilibrium of the system, \( S(\rho|\rho^0) \) must achieve the upper bound \( \mathcal{S}^*(\lambda) \) that is, we must have the equality \( S(\rho|\rho^0) = \mathcal{S}^*(\lambda) \). This corresponds to the to the minimal (or least) sufficiency of the \( \sigma \)-algebra \( \mathcal{A} \) that is, to the \( \sigma \)-algebra \( \mathcal{A}_{\min} = \{ \varnothing, \mathcal{A} \} \) where \( \varnothing \) is the null set and \( \mathcal{A} \) is the width of variation of \( X(\omega) \).

For the minimal sufficient \( \sigma \)-algebra \( \mathcal{A}_{\min} = \{ \varnothing, \mathcal{A} \} \), we have for \( P(\mathcal{A}) = 1 \),
\[
\langle X(\omega) | \mathcal{A}_{\min} \rangle \rho(\omega) = \int_{\mathcal{A}} \rho(\omega) X(\omega) \, d\omega
\]
\[
= \langle X(\omega) | \mathcal{A}_{\min} \rangle \rho^0
\]
\[
= \langle X(\omega) \rangle = E_0
\] (3.5)

which implies the conservation of the average energy \( \langle X(\omega) \rangle = E_0 \). The statistical equilibrium of the closed system is thus equivalent to the minimal (or least) sufficiency of the \( \sigma \)-algebra \( \mathcal{A}_{\min} = \{ \varnothing, \mathcal{A} \} \) induced by the energy \( X(\omega) \) of the system. This is of great significance giving a statistical criteria of the coarse-grained interpretation of thermal equilibrium of a closed system. This analysis also leads to another important result namely, the equivalence of the principles of maximum-entropy and sufficient inferences for the study of the statistical equilibrium of a closed system.

4. MAXIMUM-ENTROPY AND STABLE PRINCIPLES.

Let us study the stable character of the maximum-entropy estimate. The probability distribution \( P^0_\alpha \) for statistical equilibrium is completely characterized by the average energy
\[
\langle X \rangle = \sum_\alpha P_\alpha X_\alpha
\]
where \( X_\alpha \) is the value of the energy \( X(\omega) \) in the macroscopic state \( \Gamma_\alpha \). So the coarse-grained probability distribution \( P^0_\alpha \) for statistical equilibrium can be written in the form \( P^0_\alpha = P_\alpha(\langle X \rangle) \). Let \( \Delta \langle x \rangle \) be the error in the average value \( \langle x \rangle \), then we have
\[
\tilde{S}(P|P^0) = - \sum_\alpha P_\alpha(\langle X \rangle + \Delta \langle x \rangle) \ln \frac{P_\alpha(\langle X \rangle + \Delta \langle x \rangle)}{P^0_\alpha(\langle X \rangle)}
\] (4.1)
Expanding in powers of $\Delta \langle X \rangle$ and neglecting higher powers of $\Delta \langle X \rangle$, we have [6]

$$
\delta S(P|\beta) = -\frac{1}{2} k \sum_a \left( \frac{\partial \ln P}{\partial \langle X \rangle} \right)^2 \left( \Delta \langle X \rangle \right)^2 = \frac{1}{2} k \sum_a \frac{\delta P_a}{P_a} \left( \Delta \langle X \rangle \right)^2
$$

(4.2)

where the quantity $\left( \frac{\delta P_a}{P_a} \right)^2$ represents the measure of the sensitivity of the distribution $P_a$ subject to an error $\Delta \langle X \rangle$ in the average value $\langle X \rangle$ [4]. The maximization of $S(\beta|P_0)$ then corresponds to the lower-bound of $\left( \frac{\partial \ln P_a}{\partial \langle X \rangle} \right)^2$ or equivalently $\left( \frac{\delta P_a}{P_a} \right)^2$.

By the Cramer-Rao inequality [7], we have

$$
\left( \frac{\partial \ln P_a}{\partial \langle X \rangle} \right)^2 > \frac{1}{\left( \Delta \langle X \rangle \right)^2}
$$

(4.3)

where $(\Delta \langle X \rangle)^2 = \langle (X-\langle X \rangle)^2 \rangle$ is the variance of $X(\omega)$. The lower-bound of the left-hand member of (4.3) corresponds to the equality in (4.3) and this holds if and only if [7]

$$
\frac{\partial \ln P_a}{\partial \langle X \rangle} = \beta \langle X \rangle (X_a - \langle X \rangle)
$$

(4.4)

Integrating (4.4), we get

$$
P_0 = P_a (\langle X \rangle) = h(X_a) \exp \{ \lambda (\langle X \rangle) X_a \}/Z(\lambda (\langle X \rangle))
$$

(4.5)

which is the maximum-entropy estimate of the probability distribution $P_0$ and it is the canonical distribution of energy with $\lambda = -1/kT$, $k$ being the Boltzmann's constant and $T$ being the absolute temperature. The probability distribution (4.5) according to (4.2), is the most stable probability in the sense of its least sensitiveness to the statistical error in the average value $\langle X \rangle$. The procedure of obtaining the probability distribution as a consequence of the criteria of stability in the sense of its least sensitiveness is known as the procedure of stable inference [4]. The above analysis also leads to the important result of equivalence of the principles of maximum-entropy and stable inferences.

5. STATISTICAL EQUILIBRIUM: CHARACTERIZATION.

Let us now draw from section (3) and (4) some conclusions about the statistical characteristics of thermal equilibrium. From section (3), we have obtained an important statistical criteria of thermal equilibrium. The thermal equilibrium of the closed system corresponds to the minimal (or least) sufficiency of the energy $X(\omega)$ or to the minimal sufficiency of the partition or the $\sigma$-algebra $\alpha$ generated by the partition. In section (4) we have proved the stable character of the probability distribution of thermal equilibrium. From the equality of (4.3) we see that the spread $\Delta X$ of $X(\omega)$ in statistical equilibrium is given by the lower-bound of the root
mean-square of $X(\omega)$ that is,

$$\Delta X = [\text{Min } \langle (\Delta X)^2 \rangle]^{1/2}.$$  \hspace{1cm} (5.1)

For the width of variation $\Delta X$ of $X(\omega)$, we have in general $\Delta X < X$. However, $X(\omega)$ being minimal sufficient for statistical equilibrium we have $\Delta X = \delta X$ [7] which is an important characteristic of thermal equilibrium. This leads to the important result that the canonical distribution (4.5) is valid over average of width $\delta X$ which is of the same order of magnitude as the spread $\Delta X$ [8] and this, therefore, corresponds to the stable region of equilibrium derived from the criteria of stable probability and the criteria of minimal sufficiency of the energy $X(\omega)$ for thermal equilibrium of the closed system.

The principles of maximum-entropy, and stable and sufficient inferences have led to some important statistical characteristics of thermal equilibrium and have in fact, a greater significance in the study of the equivalence of the different estimation principles in general [9, 10].

REFERENCES
