Conditional maximum entropy and superstatistics

Sergio Davis

Comisión Chilena de Energía Nuclear, Casilla 188-D, Santiago, Chile
Departamento de Física, Facultad de Ciencias Exactas, Universidad Andres Bello.
Sazié 2212, piso 7, 8370136, Santiago, Chile
E-mail: sergio.davis@cchen.cl

Received 5 May 2020, revised 4 September 2020
Accepted for publication 9 September 2020
Published 12 October 2020

Abstract

Superstatistics describes nonequilibrium steady states as superpositions of canonical ensembles with a probability distribution of temperatures. Rather than assume a certain distribution of temperature, recently [2020 J. Phys. A: Math. Theor. 53 045004] we have discussed general conditions under which a system in contact with a finite environment can be described by superstatistics together with a physically interpretable, microscopic definition of temperature. In this work, we present a new interpretation of this result in terms of the standard maximum entropy principle using conditional expectation constraints, and provide an example model where this framework can be tested.

Keywords: superstatistics, maximum entropy, conditional expectation

1. Introduction

The maximum entropy principle (MaxEnt) [1] provides a powerful explanation, deeply rooted in information theory, for the wide success of Boltzmann–Gibbs (BG) statistical mechanics in equilibrium systems. Despite all this success, there are systems for which non-canonical steady states are the norm, mainly systems with long-range interactions [2, 3] such as collisionless plasmas [4–6], self-gravitating systems [7, 8], but also systems in high-energy physics [9–11], complex systems [12], finite-size systems in biophysics and nanotechnology [13, 14], systems presenting anomalous diffusion [15], financial markets [16], earthquakes [17] among many others. This fact has proved difficult to reconcile with BG statistics, and has led to alternative frameworks such as nonextensive (Tsallis) statistics [18] and superstatistics [19, 20], among others. While Tsallis statistics departs from the BG entropy by proposing a generalization of the entropic functional, superstatistics augments the MaxEnt by providing a mechanism that produces non-canonical distributions without the need to abandon the foundations of BG.
statistics. This is achieved by postulating a superposition of canonical ensembles, governed by a continuous or discrete distribution of inverse temperature $\beta = 1/k_B T$. More precisely, superstatistics produces ensembles of the form

$$P(x|S) = \rho(H(x); S),$$  

where $x$ represents a possible microstate of the system, $S$ denotes the state of knowledge under which the model is constructed, $H$ is the Hamiltonian function, and $\rho(E; S)$ is the ensemble function (also referred to as the generalized Boltzmann factor), given by

$$\rho(E; S) = \int_0^\infty d\beta f(\beta; S) \exp(-\beta E).$$  

Note that the state of knowledge is written in abstract terms as $S$, but will be replaced with the actual parameters of the model where appropriate. In the following we may write the functions $\rho(E; S)$ and $f(\beta; S)$ simply as $\rho(E)$ and $f(\beta)$ respectively, omitting their implicit dependence on $S$ for clearer notation.

The traditional interpretation of superstatistics invokes models with a fluctuating parameter $\beta$ that follows some predefined stochastic dynamics [21–24]. Besides this, a Bayesian interpretation [25–27] of superstatistics is possible in which the superposition of temperatures can be understood as a manifestation of the uncertainty in the value of the mean energy used as constraint in a maximum entropy setting. In this case, equations (1) and (2) are written, using the marginalization rule of probability [28, 29], as

$$P(x|S) = \int d\beta P(\beta|S) P(x|\beta)$$

under the equivalence $f(\beta) = P(\beta|S)/Z(\beta)$.

Recently [30], a mechanism leading to superstatistics has been explored where the original (target) system with microstates $x$ is in contact with an environment with microstates $y$ and Hamiltonian $G(y)$, in such a way that equation (3) is recovered by integration over the microstates of the environment, that is,

$$P(x|S) = \int d\beta P(\beta; S) P(x|\beta)$$

and where a function $B = B(x, y)$ can be defined such that the expected values

$$\langle g(\beta) \rangle_S = \int_0^\infty d\beta P(\beta|S) g(\beta),$$

$$\langle g(B) \rangle_S = \int dx dy P(x, y|S) g(B(x, y)),$$

are equal for any function $g$. Here and in the rest of the paper we adopt the Bayesian notation of expectation values with a subscript indicating the state of knowledge. The meaning of the equality of $\langle g(\beta) \rangle_S$ and $\langle g(B) \rangle_S$ is that $\beta$, originally a parameter in superstatistics, can be observed or measured using the function $B$. Accordingly, we will refer to $B$ as the microscopic inverse temperature, in the sense that it is a function of the microscopic degrees of freedom that in principle can be computed using a single configuration or ‘snapshot’ of the system. This is in
contrast with the superstatistical inverse temperature $\beta$, which is a global, statistical parameter that describes the possible macroscopic equilibrium states in the canonical ensemble.

In reference [30] we prove that equations (3)–(5) are simultaneously true if and only if the conditional distribution of the system given a fixed environment is canonical, with a temperature which is exclusively a property of the environment, i.e., if and only if

$$P(x|y,S) = \left[ \frac{\exp(-\beta H(x))}{Z(\beta)} \right]_{\beta = B(G(y))}. \quad (6)$$

Here the microscopic inverse temperature $B = B(G)$ is a function of the environment energy only. Using equation (5) with the choice $g(\beta) = \delta(\beta - \beta_0)$ with $\beta_0 \in [0, \infty)$ an arbitrary value of inverse temperature implies

$$\langle \delta(\beta - \beta_0) \rangle_S = \langle \delta(B - \beta_0) \rangle_S, \quad (7)$$

but because $P(A = a|I) = \langle \delta(A - a) \rangle_I$ for any quantity $A$ and value $a$, it follows that

$$P(\beta = \beta_0|S) = P(B = \beta_0|S).$$

In other words, the sampling distribution of $B$ (fraction of microscopic states with $B = \beta_0$) coincides with the superstatistical distribution of $\beta$, $P(\beta = \beta_0|S)$ [27]. This result restores some aspects of the frequentist intuition behind superstatistics, but at the same time it seems to lose some of the direct connection with the statistical inference frameworks, either Bayesian probability or MaxEnt. Furthermore, the condition in equation (6) is strongly suggestive of an underlying maximum entropy formulation. This suggests the need in this formalism for some extra elements.

In this work we show that equation (6), the necessary and sufficient condition for superstatistics with a microscopic definition of environment temperature, is the natural consequence of a maximum entropy analysis using conditional expectations.

This paper is organized as follows. First, in section 2, we present the use of the standard MaxEnt in the case where the ratio of two expectations is given as a constraint. In section 3 we use this result to explore the effect of a conditional expectation of energy on the MaxEnt, recovering equation (6). Section 4 discusses an application of this formalism that can be solved exactly. Finally we close with some concluding remarks in section 5.

2. Maximum entropy with ratio constraints

Before starting our main analysis, let us consider an important case in the maximum entropy formalism that is not commonly presented in the literature, but simple enough to solve. Consider a constraint over the ratio of two expectations,

$$\frac{\langle f \rangle}{\langle g \rangle} = R, \quad (8)$$

where $f = f(x)$ and $g = g(x)$. The individual values of the expectations are unknown, only their ratio $R$ is assumed known. How do we determine the most unbiased model for the probability of $x$ given the value of $R$? In Bayesian notation, this distribution will be denoted as $P(x|R)$, and according to the MaxEnt, $P(x|R)$ is the function that maximizes the BG entropy functional

$$S[p] = -\int dx \ p(x) \ln \ p(x) \quad (9)$$
under the constraint on the ratio in equation (8) and normalization. On first inspection, this is not a linear constraint, as it is the ratio of two linear functionals. However, it can simply be rewritten as

$$\langle f - R \cdot g \rangle = 0$$

in order to make it linear. The application of the principle of maximum entropy [28] over this rewritten constraint directly yields as a solution

$$P(x|R) = \frac{1}{Z(\zeta)} \exp(-\zeta [f(x) - R \cdot g(x)]) ,$$

where the constraint equation that fixes the value of the Lagrange multiplier $\zeta$ is

$$-\frac{\partial}{\partial \zeta} \ln Z(\zeta) = \langle f - R \cdot g \rangle_\zeta = 0 .$$

A simple proof that the constraints in equations (8) and (10) are equivalent for the maximum entropy formalism is given in the appendix A.

A direct consequence of the solution in equations (11) and (12) of the constraint in equation (8) is the following general result. Let $A(x)$ and $B(x)$ be functions of $x$ and consider the conditional expectation constraint

$$\langle A \rangle_{B=B_0,I} = a(B_0),$$

that is, the expected value of $A$ over the states with fixed $B(x) = B_0$ under the state of knowledge $I$ is known to be the function $a(B_0)$. By using Bayes’ theorem [28, 29] we can write this conditional expectation as

$$\langle A \rangle_{B=B_0,I} = \int dx \ P(x|B=B_0,I) A(x)$$

$$= \int dx \ \left[ \frac{P(x|I) \cdot P(B = B_0|I,x)}{P(B = B_0|I)} \right] A(x)$$

$$= \frac{1}{P(B = B_0|I)} \int dx \ P(x|I) \delta(B(x) - B_0) A(x)$$

$$= \frac{\langle A \cdot \delta(B - B_0) \rangle}{P(B = B_0|I)} ,$$

where we have used the straightforward fact that knowledge of $x$ determines $B(x)$ completely, hence the probability of $B$ given $x$ is a Dirac delta,

$$P(B = B_0|I,x) = \delta(B(x) - B_0).$$

Choosing $A = 1$ in equation (14) leads to $P(B = B_0|I) = \langle \delta(B - B_0) \rangle_I$, thus finally

$$\langle A \rangle_{B=B_0,I} = \frac{\langle A \cdot \delta(B - B_0) \rangle}{\langle \delta(B - B_0) \rangle_I} = a(B_0).$$

From equation (11) we directly obtain the maximum entropy solution under the constraint in equation (13) as

$$P(x|\zeta) = \frac{1}{Z(\zeta)} \exp(-\zeta [A(x) - a(B_0)] \delta(B(x) - B_0)) ,$$

with $\zeta$ a function of $a(B_0)$, therefore a function of $B_0$. 

4
Having the solution in equations (11) and (12) to the ratio constraint in equation (8), we are now equipped to present a generalization of the textbook maximum entropy problem that leads to the canonical ensemble. Instead of the usual constraint on the mean energy,

\[ \langle H \rangle_S = \bar{E}, \] (17)

where \( \bar{E} \) is a known constant, and which leads to a single value of the inverse temperature \( \beta = \beta(\bar{E}) \), we will consider that our system is placed in contact with an environment with Hamiltonian \( G \) and we know the conditional mean value of the energy, \( \bar{E}(G_0) \) given that the environment is ‘frozen’ at energy \( G = G_0 \), that is,

\[ \langle H \rangle_{S,G=G_0} = \bar{E}(G_0), \] (18)

for every value of \( G_0 \in [G_{\text{min}}, G_{\text{max}}] \). This is a generalization of the constraint in equation (17). Just as in the canonical ensemble the value of \( \bar{E} \) leads to a unique value of inverse temperature \( \beta \), the function \( \bar{E}(G) \) will lead to a distribution of inverse temperatures, as we will see shortly.

Using the result in equation (15), we can rewrite the constraint in equation (18) as a ratio constraint,

\[ \frac{\langle H \rangle_{S,G=G_0}}{\langle G - G_0 \rangle_S} = \bar{E}(G_0), \forall G_0. \] (19)

Hence by using equation (16) the maximum entropy solution is given by

\[ P(x, y|S) = \frac{1}{\eta[\zeta, \bar{E}] \exp \left( -\int_{G_{\text{min}}}^{G_{\text{max}}} dG_0 \zeta(G_0) \delta(G_0 - G_0) \left[ H(x) - \bar{E}(G(y)) \right] \right)} \]

\[ = \frac{1}{\eta[\zeta, \bar{E}] \exp \left( -\zeta(G(y)) \left[ H(x) - \bar{E}(G(y)) \right] \right)} , \] (20)

where \( \eta[\zeta, \bar{E}] \) is a partition functional (normalization constant), given by

\[ \eta[\zeta, \bar{E}] = \int dx \, dy \, \exp \left( -\zeta(G(y)) \left[ H(x) - \bar{E}(G(y)) \right] \right) . \] (21)

Now we show that this solution is compatible with equation (6). We obtain first the marginal distribution of \( y \) by integrating over \( x \),

\[ P(y|S) = \int dx \, P(x, y|S) \]

\[ = \frac{1}{\eta[\zeta, \bar{E}] \exp \left( \zeta(G(y))\bar{E}(G(y)) \right)} \times \left[ \int dx \, \exp(\zeta(G(y))H(x)) \right] \]

\[ = \frac{1}{\eta[\zeta, \bar{E}] \exp \left( \zeta(G(y))\bar{E}(G(y)) \right)} \cdot Z(\zeta(G(y))), \] (22)

where \( Z(\beta) \) is the standard partition function of the target system,
By dividing \( P(x, y | S) \) by \( P(y | S) \) we clearly obtain
\[
P(x | y, S) = \frac{P(x, y | S)}{P(y | S)} = \frac{\exp(-\zeta(G(y))H(x))}{Z(\zeta(G(y)))},
\]
but this is precisely equation (6) if we identify \( \zeta(G) \) with \( B(G) \),
\[
P(x | y, S) = \frac{\exp(-\beta H(x))}{Z(\beta)} \bigg|_{\beta = B(G)}.
\]

We determine the Lagrange multiplier function \( B(G) \) by imposing the constraint in equation (18),
\[
\bar{E}(G) = \langle H \rangle_{\beta | G} = \int dx \ P(x | G, S) H(x)
\]
\[
= \int dx \ H(x) \int dy \ P(x | y, S) P(y | G, S)
\]
\[
= \int dx \ H(x) \int dy \ \frac{\exp(-\beta H(x))}{Z(\beta)} \bigg|_{\beta = B(G)} \times \left[ \frac{\delta(G(y) - G)}{\Omega_G(G)} \right]
\]
\[
= \int dx \ H(x) \left[ \frac{\exp(-\beta H(x))}{Z(\beta)} \right] \bigg|_{\beta = B(G)} \times \left[ \int dy \ \frac{\delta(G(y) - G)}{\Omega_G(G)} \right]
\]
\[
= \int dx \ \frac{\exp(-B(G)H(x)H(x))}{Z(B(G))},
\]
where \( \Omega_G(G) \) is the density of states of the environment, given by
\[
\Omega_G(G) := \int dy \ \delta(G(y) - G),
\]
and where in the third and fourth lines we have used equation (25).

Defining the canonical caloric curve \( \varepsilon(\beta) \) as
\[
\varepsilon(\beta) := \langle H \rangle_{\beta} = -\frac{\partial}{\partial \beta} \ln Z(\beta).
\]
we can directly write \( \bar{E}(G) \) as
\[
\bar{E}(G) = \left[ -\frac{\partial}{\partial \beta} \ln Z(\beta) \right] \bigg|_{\beta = B(G)} = \varepsilon(B(G)).
\]

Therefore, if \( \varepsilon(\beta) \) is invertible, we have
\[
B(G) = \varepsilon^{-1}(\bar{E}(G)).
\]

The physical interpretation of the connection between \( B(G) \) and \( \varepsilon(\beta) \) is that the microscopic inverse temperature function \( B(G) \) is the one that, through the canonical caloric curve, yields the known mean energy
\[
E(G) = \langle H \rangle_{\beta = B(G)}
\]
for every admissible value of \( G \). Thus, in the target system, \( \beta \) follows the fluctuations of \( G \) mapped through the canonical caloric curve.
The marginal distribution $P(x|S)$ resulting from integration over $y$ of equation (20) will be described by superstatistics with $P(\beta|S) = P(B = \beta|S)$, which we can write in terms of our original quantity $\bar{E}(G)$. This gives

$$P(\beta|S) = \langle \delta(B(G) - \beta) \rangle_S$$

$$= \int dy \left[ \frac{1}{\eta} Z(B(G(y))) \exp \left( B(G(y))\bar{E}(G(y)) \right) \delta(B(G(y)) - \beta) \right]$$

$$= \frac{1}{\eta} Z(\beta) \int dy \exp(\beta \bar{E}(G(y))) \delta(B(G(y)) - \beta)$$

$$= \frac{1}{\eta} Z(\beta) \int dG \Omega_G(G) \exp(\beta \bar{E}(G)) \delta(B(G) - \beta).$$  \hspace{1cm} (31)

By using the following property of the Dirac delta function,

$$\delta(B(G) - \beta) = \sum_{G^*} \frac{\delta(G^* - G)}{|B'(G)|},$$  \hspace{1cm} (32)

where $G^*$ are the solutions of $B(G^*) = \beta$, we derive the explicit formula

$$P(\beta|S) = \frac{1}{\eta} Z(\beta) \sum_{G^*} \frac{\Omega_G(G^*)}{|B'(G^*)|} \exp(\beta \bar{E}(G^*)).$$  \hspace{1cm} (33)

In order to determine the values of $G^*$, we use equation (29), which gives us the connection between $\bar{E}$ and $B$,

$$\bar{E}(G) = \epsilon(B(G)).$$

In the case of $G = G^*$, this becomes $\bar{E}(G^*) = \epsilon(B(G^*)) = \epsilon(\beta)$ and so, if the relation $\bar{E}(G)$ is invertible, we have a single value of $G^*$, namely

$$G^*(\beta) = E^{-1}(\epsilon(\beta)).$$  \hspace{1cm} (34)

and after replacing equations (34) and (29) into equation (33), we arrive at our main result,

$$P(\beta|S) = \frac{1}{\eta} Z(\beta) \Omega_G(G^*(\beta)) \left| \frac{\epsilon'(\beta)}{E'(G^*(\beta))} \right| \exp(\beta \epsilon(\beta)),$$  \hspace{1cm} (35)

where we have also used the derivative of equation (29) with respect to $G$ to replace $B'(G)$,

$$E'(G) = \frac{dE(G)}{dG} = \frac{d\epsilon(\beta)}{d\beta} \bigg|_{\beta=B(G)} \cdot \frac{dB(G)}{dG} = \epsilon'(B(G))B'(G).$$  \hspace{1cm} (36)

Note that we can also write equation (35) as an application of Bayes’ theorem with

$$P(\beta|S) \propto P(\beta|I_0) \times \frac{\Omega_G(G^*(\beta))}{|B'(G^*(\beta))|}$$  \hspace{1cm} (37)

with a prior probability

$$P(\beta|I_0) = \frac{1}{\eta_0} \exp(\beta \epsilon(\beta) + \ln Z(\beta)) = \frac{1}{\eta_0} \exp(s(\beta)),$$  \hspace{1cm} (38)
where
\[ s(\beta) := \langle -\ln P(x|\beta) \rangle_{\beta} = \beta \varepsilon(\beta) + \ln Z(\beta) \]
is the canonical entropy of the target system. The prior \( P(\beta|I_0) \) coincides with the maximum entropy distribution of inverse temperature derived by Abe et al \[31\], and the extra factor in equation (37) contains the new information related to the constraint on \( E(G) \). We can write our result in terms of the usual notation of superstatistics, as
\[ f(\beta) = \frac{P(\beta|S)}{Z(\beta)} = \frac{1}{\eta} \Omega G(\beta') \left| \frac{\varepsilon'(\beta)'}{E'(G'(\beta'))} \right| \exp(\beta \varepsilon'(\beta')), \quad (39) \]
and the corresponding ensemble function \( \rho(E) \) for the target system is
\[ \rho(E) = \frac{1}{\eta} \int_0^\infty d\beta \Omega G(\beta') \left| \frac{\varepsilon'(\beta)'}{E'(G'(\beta'))} \right| \exp(-\beta[E - \varepsilon(\beta)]). \quad (40) \]

4. Example

Let us demonstrate this formalism with a concrete example. Consider the target and environment in such a state that
\[ \langle H + G \rangle_{S,G} = \langle H \rangle_{S,G} + G_0 = E_0, \quad (41) \]
with \( E_0 \) the fixed total energy of the system. Furthermore, let us assume that the target has constant specific heat \( C > 0 \), therefore \( E(T) = CT \) and the microcanonical inverse temperature is
\[ \beta_{\Omega}(E) = \frac{1}{k_B T(E)} = \frac{\partial}{\partial E} \ln \Omega(E) = \frac{C}{k_B E}, \quad (42) \]
so that, by integration in \( E \) we have
\[ \Omega(E) = c_0 E^\alpha, \quad \alpha := \frac{C}{k_B} > 0, \quad (43) \]
with \( c_0 \) an unimportant integration constant. Because the specific heat is an extensive quantity, \( \alpha \) must be proportional to the size of the target system. For the environment, we assume a constant microcanonical inverse temperature \( \beta_0 \), so we have
\[ \beta_{\Omega}(G)^{env} := \frac{\partial}{\partial G} \ln \Omega G(G) = \beta_0, \quad (44) \]
and again by integration it follows that
\[ \Omega G(G) = c_1 \exp(\beta_0 G), \quad (45) \]
with \( c_1 \) an unimportant integration constant. From the density of states of the target (equation (43)) we can obtain its partition function,
\[ Z(\beta) = \int_0^\infty dE \Omega(E) \exp(-\beta E) = c_0 \Gamma(\alpha + 1) \beta^{-(\alpha + 1)}, \quad (46) \]
as well as its canonical caloric curve,
\[ \varepsilon(\beta) = -\frac{\partial}{\partial \beta} \ln Z(\beta) = \frac{\alpha + 1}{\beta}. \]  

From equation (41) we can obtain \( \bar{E}(G) \) as
\[ \bar{E}(G) = \langle H \rangle_{S,G} = E_0 - G, \]  
and replacing equations (47) and (48) into equation (29), we obtain
\[ \bar{E}(G) = E_0 - G = \varepsilon(B(G)) = \frac{\alpha + 1}{B(G)}. \]  

from which we read the microscopic inverse temperature
\[ B(G) = \frac{\alpha + 1}{E_0 - G} \]  
that the environment imposes over the system. Because the energy \( G \) of the environment is such that \( G \in [0, E_0] \), we have \( B \in [\beta_{\text{min}}, \infty) \), where we have defined
\[ \beta_{\text{min}} := \frac{\alpha + 1}{E_0}. \]  
This means the microscopic temperature \( T(G) = 1/(k_B B(G)) \) is bounded from above,
\[ 0 \leq T(G) \leq \frac{E_0}{k_B(\alpha + 1)}. \]  
Moreover, because the function \( \bar{E}(G) \) is invertible, we have
\[ G^*(\beta) = E_0 - \frac{\alpha + 1}{\beta} = E_0 \left(1 - \frac{\beta_{\text{min}}}{\beta}\right). \]  

In order to make the statistical treatment smoother, let us consider the limit where \( E_0 \) is large enough that we can take \( \beta \in [0, \infty) \) and \( E \in [0, \infty) \). Now we have everything to compute the statistical properties. First, we compute the joint energy-inverse temperature distribution,
\[ P(E, \beta|S) = P(E|\beta) \times P(\beta|S) \]
\[ = \exp(-\beta E)\Omega(E)f(\beta) \]
\[ = \frac{[\beta_0(\alpha + 1)]^{\alpha+2}}{\Gamma(\alpha + 1)\Gamma(\alpha + 2)} \exp(-\beta_0(\alpha + 1)/\beta)\beta^{-2} \exp(-\beta E)E^\alpha. \]  

At this point we identify the state of knowledge \( S \) with the parameters \((\alpha, \beta_0)\). Note that the constant \( c_0 \) has canceled because both \( \Omega(E) \) and \( Z(\beta) \) are proportional to it, and \( f(\beta) = P(\beta|S)/Z(\beta) \).

The marginal distribution of \( \beta \), equivalent to equation (35), yields
\[ P(\beta|\alpha, \beta_0) = \frac{[\beta_0(\alpha + 1)]^{\alpha+2}}{\Gamma(\alpha + 1)\Gamma(\alpha + 2)}\beta^{-\alpha-3} \exp(-\beta_0(\alpha + 1)/\beta), \]  
that is, an inverse gamma distribution, which means the temperature \( T = 1/(k_B \beta) \) that the target system ‘sees’ follows a gamma distribution. This distribution belongs to one of the
three main universality classes of superstatistics, namely gamma ($\chi^2$), inverse gamma (inverse $\chi^2$) and lognormal superstatistics [32], and has found applications in several contexts from nonequilibrium physics [33] and complex systems [34] to cancer survival times [35] and concentration of air pollutants [36]. Incidentally, the inverse gamma distribution arises naturally from Crooks’ hyperensemble formalism [37] under the same density of states of equation (43) without additional constraints. Our result in equation (54) agrees with Crooks’ distribution (equation (12) of reference [37])

$$P(T) \propto \left(\frac{T}{T^*}\right)^{\lambda-1} \exp(-c\lambda T/T^*), \quad (55)$$

if we set Crooks’ parameter $\lambda$ equal to 1, because his target partition function is $Z \propto \beta^{-c}$, so his $c$ parameter corresponds to $\alpha + 1$.

From this distribution we can compute the mean inverse temperature

$$\langle \beta \rangle_{\alpha, \beta_0} = \beta_0 \quad (56)$$

and its normalized variance,

$$\frac{\langle (\delta \beta)^2 \rangle_{\alpha, \beta_0}}{\langle \beta \rangle_{\alpha, \beta_0}^2} = \frac{1}{\alpha} \quad (57)$$

and we see that $\alpha \to \infty$ keeping $\beta_0$ fixed leads to $\langle (\delta \beta)^2 \rangle_{\alpha, \beta_0} \to 0$, recovering the canonical ensemble with $\beta = \beta_0$. The ensemble function $\rho(E; \alpha, \beta_0)$ of the target system is

$$\rho(E; \alpha, \beta_0) = \frac{2}{c_0 \Gamma(\alpha + 1) \Gamma(\alpha + 2)} [\beta_0(\alpha + 1)]^{\alpha+1} \sqrt{\beta_0(\alpha + 1)\sqrt[2]{\beta_0(\alpha + 1)E} K_1(2 \sqrt[2]{\beta_0(\alpha + 1)E})}, \quad (58)$$

where $K_1(z)$ is the modified Bessel function of the second kind, and the corresponding energy distribution of the target is

$$P(E|\alpha, \beta_0) = \rho(E)\Omega(E) = \frac{2\beta_0(\alpha + 1)}{\Gamma(\alpha + 1) \Gamma(\alpha + 2)} [\beta_0(\alpha + 1)E]^{\alpha+\frac{1}{2}} K_1(2 \sqrt[2]{\beta_0(\alpha + 1)E}). \quad (59)$$

The mean target energy is

$$\langle E \rangle_{\alpha, \beta_0} = \frac{\alpha + 2}{\beta_0}, \quad (60)$$

while its normalized variance is

$$\frac{\langle (\delta E)^2 \rangle_{\alpha, \beta_0}}{\langle E \rangle_{\alpha, \beta_0}^2} = \frac{2}{\alpha + 1} \quad (61)$$

In this superstatistical ensemble described by equation (53), the fluctuations of (inverse) temperature and energy are well defined, and we can explore what their joint behavior is. As we can see from equations (57) and (61), when $\alpha$ increases with fixed $\beta_0$, the variance of $\beta$ tends to zero as $1/\alpha$, while the variance of $E$ increases linearly with $\alpha$, so their product remains constant for large enough $\alpha$. 

10
In fact, it is possible to construct a thermodynamic uncertainty relation for the unnormalized variances of $\beta$ and $E$ in the target system, as have been proposed and studied in previous works [38, 39], namely

$$\langle (\delta \beta)^2 \rangle_{\alpha, \beta_0} \cdot \langle (\delta E)^2 \rangle_{\alpha, \beta_0} = 2 \left[ \frac{(\alpha + 2)^2}{\alpha(\alpha + 1)} \right] \geq 2. \quad (62)$$

Here the variance of the intensive parameter $\beta$ can in fact be interpreted as fluctuations of a physical quantity $B$, which decrease when the fluctuations of energy increase. On the other hand, we can compare this uncertainty relation with the correlation between $\beta$ and $E$ that we can obtain from the joint distribution,

$$\langle \delta \beta \cdot \delta E \rangle_{\alpha, \beta_0} = \langle \beta E \rangle_{\alpha, \beta_0} - \langle \beta \rangle_{\alpha, \beta_0} \langle E \rangle_{\alpha, \beta_0} = (\alpha + 1) - (\alpha + 2) = -1, \quad (63)$$

in direct agreement with Schwarz inequality,

$$\langle (\delta \beta)^2 \rangle_S \cdot \langle (\delta E)^2 \rangle_S \geq \langle \delta \beta \cdot \delta E \rangle_S^2$$

for an arbitrary state of knowledge $S$.

5. Concluding remarks

We have complemented the framework initiated in reference [30], by connecting it with the MaxEnt under a conditional energy expectation constraint, $\langle H \rangle_{S, G} = \bar{E}(G)$. In this formalism, knowledge of $\bar{E}(G)$, the canonical caloric curve $\varepsilon(\beta)$ and the density of states $\Omega_G$ of the environment completely determines the most unbiased form of superstatistics, with $f(\beta)$ or $\rho(E)$ given by equations (34), (39) and (40). The frequentist interpretation of $P(\beta|S)$ as a sampling distribution of an observable $B$ of the environment remains valid. We have explored a simple model of target and environment for which one can obtain fluctuations of the intensive parameter $\beta$ in a manner consistent with a frequentist interpretation of superstatistics, and also consistent with the Bayesian/MaxEnt framework, only requiring the standard elements of traditional MaxEnt and the laws of probability. In this example, it is clear that in the thermodynamic limit for the target system the statistics become canonical, as was shown also for the case of an infinite environment in reference [30]. Superstatistics in the thermodynamic limit may still be observed in other situations, such as nonequilibrium driven systems, where oscillations of $B(G)$ may induce a probability distribution $P(\beta|S)$.

Acknowledgments

This work is supported by the Anillo ACT-172101 grant. I thankfully acknowledge the careful and critical reading of this work by four anonymous reviewers.

Appendix A. Use of known ratio of expectation values as a constraint in maximum entropy inference

Consider the constraint

$$\frac{\langle f \rangle}{\langle g \rangle} = R, \quad (A.1)$$
where \( f = f(x) \) and \( g = g(x) \), and where only the value of \( R \) is given. We seek to verify that this constraint is equivalent to the constraint
\[
\langle f - R \cdot g \rangle = 0,
\]
when used in maximum entropy inference. In order to do this, we need to prove that the extremum equations
\[
\frac{\delta}{\delta p(x)} \left( S[p] + \mu \langle 1 \rangle + \lambda \frac{\langle f \rangle}{\langle g \rangle} \right) = 0 \tag{A.3}
\]
and
\[
\frac{\delta}{\delta p(x)} \left( S[p] + \mu \langle 1 \rangle + \zeta \langle f - R \cdot g \rangle \right) = 0 \tag{A.4}
\]
have the same solution \( p^*(x) \). Here \( S[p] \) represents the Gibbs entropy functional
\[
S[p] = -\int dx \ p(x) \ln p(x),
\]
and the term \( \mu \langle 1 \rangle \) comes from the normalization constraint. We should have, then,
\[
\lambda \frac{\delta}{\delta p(x)} \left( \frac{\langle f \rangle}{\langle g \rangle} \right) \bigg|_{p=p^*} = \zeta \frac{\delta}{\delta p(x)} \left( \langle f - R \cdot g \rangle \right) \bigg|_{p=p^*} = \zeta \langle f(x) - R \cdot g(x) \rangle. \tag{A.5}
\]
Rewriting the left-hand side,
\[
\lambda \frac{\delta}{\delta p(x)} \left( \frac{\langle f \rangle}{\langle g \rangle} \right) = \frac{\lambda}{\langle g \rangle} f(x) - \frac{\lambda \langle f \rangle}{\langle g \rangle^2} g(x) = \frac{\lambda}{\langle g \rangle} (f(x) - R \cdot g(x)), \tag{A.6}
\]
we in fact see that the constraints are equivalent, provided that we define \( \zeta \) as
\[
\zeta = \frac{\lambda}{\langle g \rangle}.
\]
Finally we have the maximum entropy solution
\[
P(x|R) = P(x|\zeta) = \frac{1}{Z(\zeta)} \exp \left( -\zeta \left[ f(x) - R \cdot g(x) \right] \right), \tag{A.7}
\]
where we have replaced the dependence on \( R \) by a dependence on the Lagrange multiplier \( \zeta \), whose value is determined from the constraint in equation (A.2),
\[
\langle f - R \cdot g \rangle_\zeta = \int dx \ P(x|\zeta) \left[ f(x) - R \cdot g(x) \right] = \frac{1}{Z(\zeta)} \int dx \ \exp \left( -\zeta \left[ f(x) - R \cdot g(x) \right] \right) \left[ f(x) - R \cdot g(x) \right] = -\frac{\partial}{\partial \zeta} \ln Z(\zeta) = 0. \tag{A.8}
\]
ORCID IDs

Sergio Davis 🐳 https://orcid.org/0000-0003-2757-332X

References

[1] Jaynes E T 1957 Information theory and statistical mechanics Phys. Rev. 106 620–30
[2] Tsallis C, Rapisarda A, Pluchino A and Borges E P 2007 On the non-Boltzmannian nature of quasi-stationary states in long-range interacting systems Physica A 381 143–7
[3] Ourabah K, Gougam L A and Tribeche M 2015 Nonthermal and suprathermal distributions as a consequence of superstatistics Phys. Rev. E 91 121133
[4] Aravindakshan H, Yoon P H, Kakad A and Kakad B 2020 Theory of ion holes in space and astrophysical plasmas Mon. Not. R. Astron. Soc. 497 69–75
[5] Davis S 2020 On the possible distributions of temperature in nonequilibrium steady states J. Phys. A: Math. Theor. 53 045004

[6] Abe S, Beck C and Cohen E G D 2007 Superstatistics, thermodinamics and fluctuations Phys. Rev. E 76 31102
[32] Beck C 2009 Recent developments in superstatistics Braz. J. Phys. 39 357–63  
[33] Sattin F and Salasnich L 2002 Multiparameter generalization of nonextensive statistical mechanics Phys. Rev. E 65 5–8  
[34] Abul-Magd A Y, Dietz B, Friedrich T and Richter A 2008 Spectral fluctuations of billiards with mixed dynamics: from time series to superstatistics Phys. Rev. E 77 046202  
[35] Leon Chen L and Beck C 2008 A superstatistical model of metastasis and cancer survival Physica A 387 3162–72  
[36] Williams G, Schäfer B and Beck C 2020 Superstatistical approach to air pollution statistics Phys. Rev. Res. 2 013019  
[37] Crooks G E 2007 Beyond Boltzmann–Gibbs statistics: maximum entropy hyperensembles out of equilibrium Phys. Rev. E 75 41119  
[38] Uffink J and Van Lith J 1999 Thermodynamic uncertainty relations Found. Phys. 29 655–92  
[39] Velazquez L and Curilef S 2009 A thermodynamic fluctuation relation for temperature and energy J. Phys. A: Math. Theor. 42 095006