Maximization of Entropy is More General Than Reaction Balance

Francesco R. Ruggeri Hanwell, N.B. June 2, 2026

In the Maxwell-Boltzmann (MB) case, one may obtain the probability distribution by maximizing Shannon's entropy - Sum over i p(ei) ln(p(ei)) subject to the constraints Sum over i ei p(ei) = Eave and Sum over i p(ei)=1, or by reaction balance: ei+ej = ek+el with p(ei)p(ej) = p(ek)p(el). This might lead to the notion that the two approaches are essentially equivalent, i.e. reaction balance is maximizaiton of entropy. For the particular case of ln(p(ei)) = -ei/T, they are, but in general, they are not, we argue.

As noted, reaction balance suggests that ei is the only information in the problem (as T is linked to Eave). Certainly, writing p(ei)p(ej) as the probability for an interaction to occur between ei and ej is not the most general expression. It is an expression based on the least amount of information, namely that if one has an ei and ej present, they will interact. This allows one to write ln(p(ei)) = -ei/T. It is not just the case that only ei appears on the RHS, it is also the case that it only appears on the LHS, i.e. there is no other parameter. What would happen if another parameter existed, i.e there was extra information? We suggest that ln(p(ei)) may be generalized to F(p(ei,k)), where k is an extra parameter. F is the inverse of p, so in principle if p contains k, F may also contain k in such a way that F(p) removes p. If one uses F = ln(p), however, then ln does not constain k and hence cannot remove this extra parameter. This is the case for which there cannot be any more information than -ei/T.

The general process of maximization of entropy subject to constraints seems to be the following, we argue. One first approximates a physical system by average moments of ei, namely ei power 1 and ei power 0. This, however, gives no information on what distribution should appear because one may have many that satisfy Sum over i p(ei) = 1 and Sum over i ei p(ei) = Eave. In this approach there are two pieces of information, p(ei) and the moments used. One might suggest that one might generalize and write an average function Sum over i p(ei) F(p(ei)).

This function must describe the system and so F(p(ei)) is a function of ei and should only represent the same moments as the constraints, i..e F(p(ei)) = a ei + b. F is the inverse of p(ei) and so something interesting occurs. One may have information beyond ei contained in F(p(ei)), i.e in F and p separately as this information may cancel out. If one wishes to minimize information, one may argue that d/dp (p F(p)) = d/dp (a1 ei p(ei)) + a2 p(ei)). This leads to a second condition: p dF/dp = 1. Again if an extra parameter appears in the form of a power law, then p dF/dp can cancel the effects of this extra parameter (extra information) and if F=1/k p power k, the p dF/dp shows no trace of the parameter. Thus, one may create a special function F(p) which has more information than simply ei information (unlike the Maxwell-Boltzmann case) and still describe a physical system in terms of average moments of ei. By writing p(ei) F(p(ei) and minimizing information, one may obtain explicit forms for two solutions of p(ei), namely the MB one and a power law (i.e. the Tsallis case). The Tsallis case, however, cannot be represented by reaction balance because p(ei) p(ej) = p(ek) p(el) ((1)) for ei+ej = ek+el assumes no other information than ei and this is not the case for a Tsallis power law which involves an extra piece of information, namely the power law parameter. One cannot have the power law parameter disappear from ((1)). Formally, if one considers ((1)) with p(ei,k), then ln(p(ei,k) =

-ei/T and this means that k must disappear because it is not defined by this equation. Having any allowable k is the same as saying one may have many solutions p(ei,k) for k1,k2, k3 etc. This is no different than picking multiple solutions which satisfy Sum over i p(ei) =1 and Sum over i ei p(ei) = Eave.

Reaction Balance p(ei)p(ej)=p(ek)p(el) and Minimal Information

In the Maxwell-Boltzmann case, reaction balance:

p(ei)p(ej) = p(ek)p(el) for ei+ej = ek+el ((1))

yields p(ei) = C exp(-ei/T) just as maximization of Shannon's entropy:

S = - Sum over i p(ei) ln(p(ei) ((2a)) subject to Sum over i p(ei)=1 and Sum over i ei p(ei) = Eave ((2b))

does.

One might think that the two approaches are identical, i.e. reaction balance is maximizatin of entropy. We suggest, however, that the MB case is a special instance and try to analyze the situation in more detail in terms of information.

We argue that ((1)) implies that there is no other information present than ei. ((1)) is associated with the constraints ((2b)), but there are many p(ei) solutions which satisfy these constraints. ((1)) is a special kind of minimal information approach which yields a very specific p(ei) solution for which there is only ei information. As stated, this approach is equivalent to a formal maximization of Shannon's entropy subject to constraints. This begs the question: Is ((1)) the only condition which may exist in order to have maximum entropy? To answer this, we argue that one must consider the notion of entropy in more detail.

The Notion of Entropy

In the previous section, we saw that ((1)) is sufficient for obtaining the Maxwell-Boltzmann distribution. Even though one may find an equivalent solution using maximization of entropy, there is no need for the notion of entropy to find the MB distribution.Certainly, ((1)) is not the most general expression for an interaction because it states that given an ei and ej, they will interact in the same way. Physically , there may be other situations. Does this mean that they cannot be described by a maximization of some kind of entropy function?

We begin with the constraints ((2b)). They allow for multiple solutions for p(ei), suggesting that nature has many probability distributions. Experiment, however, shows that this is not the case. For many years, it was believed that the Maxwell-Boltzmann distribution was the key minimal information one. We ask: Is it possible to have minimal information with more information than that contained in ((1))?

Given the constraints ((2b)), one describes a physical system in terms of averages of moments ei power 0 and ei power 1. This leads to many solutions. One wishes to have a unique solution, but this must be based on some principle. We suggest a minimal information solution. The constraints ((2b)) contain two pieces of information, p(ei) and the ei moments. We suggest that one may write a single term (instead of a sum of two constraints), i.e.

p(ei) F(p(ei)) ((3))

Given that ((3)) describes the same physical system and that F(p(ei)) is ultimately a function of ei,

F(p(ei)) = aei + b ((4))

Mathematically, F and p(ei) are inverses, but the catch is that one may introduce a parameter into p(ei), i.e.

p(ei, k) ((5))

Siuch that F removes this parameter. This then represents a system with more information than that contained in ((1)). This also means that ((1)) does not apply to such a system. Can this system still maximize entropy given by ((3)) subject to ((2b)?

For that to be the case, one requires:

p dF/dp = 1 ((6))

A power law, however, is a solution to ((6)), i.e. p = (a1 - a2 ei/T) power 1/(1-q)

In particular, one may use: p(ei) = (1 - (1-q) ei/T) power 1/(1-q) and

F = (1+ p power 1-q) / (1-q) ((7))

F becomes ln in the q-> 1 limit and ((6)) is satisfied. One may have a minimziation of entropy scheme with information formally given by F(p(ei,k)) and minus average informjaiton given by p(ei) F(p(ei).

The power law case is the Tsallis one. In general, one may use p(ei) power q in the ei constraint, i.e.

Sum over i ei p(ei) power q = Eave and Sum over i p(ei)=1 and define entropy usings

-Sum over i p(ei) power q F(p(ei,k) with F(p(ei,k)) = aei + b ((8))

The point is that one may create a single function of information average over the probability used for the ei moment and then minimize it subject to constraints. It is possible to bring in an extra parameter linked to a power law (e.g. q) in this manner and still describe the system in terms of average moments of ei/T. This is an alternative to the simpler Maxwerll-Boltzmann solution which contains no such extra power law parameter information.

Conclusion

In conclusion, we suggest that the idea of entropy is really the idea of having an information function F(p(ei,k) such that its average

Sum over i p(ei) F(p(ei,k)

descrbes the same avearge moments of ei, i.e. F(p(ei,k)) = aei + b, but at the same time may be minimized subject to the ei moment constraints to yield a unique solution. In the simplest case, this approach is equivalent to reaction balance p(ei)p(ej) = p(ek)p(el) with ei+ej = ek+el, but given that F and p(ei,k) are inverses, it is mathematically possible to include an extra parameter k in p which is removed through F.. Furthermore, in the minimization procedure, the parameter k also vanishes from p dF/dp = 1. As a result, one may have a more complicated (system with more information) than the reaction balance MB case. In fact, one may have a power law solution which is the Tsallis case.