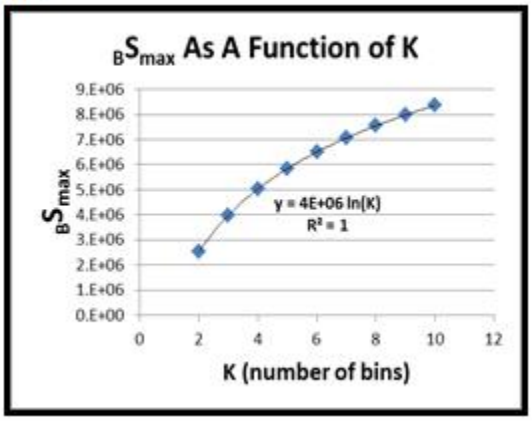
NOTE TO FILE:
Garvin H Boyle
Dated: R2:180119

Entropy in a Histogram, and in ABMs

Frontispiece

Figure 01 – Table of values – $_BS_{max}$ as K varies

A	K	α	$_BS_{max}$
3628800	1	3628800	0.00000
3628800	2	1814400	2515284.62976
3628800	3	1209600	3986628.81662
3628800	4	907200	5030562.09360
3628800	5	725760	5840298.11141
3628800	6	604800	6501899.37618
3628800	7	518400	7061274.24634
3628800	8	453600	7545825.91877
3628800	9	403200	7973230.01610
3628800	10	362880	8355555.32853
	3628800		



Substituting equations [12] and [15] into equation [14] I get:

$$ {}_BS_{index} = \frac{{}_BS}{{}_BS_{max}} = \frac{\left[\ln(A!) - \sum_{i=1}^{K}\left(ln(a_i!)\right)\right]}{A\ln(K)} \qquad [16] $$

If I wish, in resolving the ln/ln division, I can write this as:

$$ {}_SS_{index} = -\sum_{i=1}^{K}(p_i \log_K(p_i)) \qquad [28] $$

Table of Contents

1 - References 1
2 - Background 1
3 - Purpose 3
4 - Discussion 3
4.1 - S_{max}? 3
4.2 - ${}_{B}S_{max}$ – Boltzmann Regime 5
4.3 - S_{index}? 7
4.4 - ${}_{B}S_{index}$ – Boltzmann Regime 8
4.5 - ${}_{S}S_{max}$ – Shannon Regime 9
4.6 - ${}_{S}S_{index}$ – Shannon Regime 10
5 - Summary 11

1 - References

A. A. Drăgulescu and V.M. Yakovenko (2000) "Statistical mechanics of money", Eur. Phys. J. B 17, 723-729.
B. 140218 Email from Yakovenko R1.pdf
C. 140409 NTF Discussion With Dr Yakovenko R1.pdf
D. 150527 PPR - Definition of EI R17.pdf
E. 180118 NTF Custom Functions in Excel R3.pdf
F. 180210 NTF Shannon Vs Boltzmann R3.pdf
G. 180214 NTF Entropy and Units of Measure R5.pdf
H. 180214 NTF FactLn() and GammaLn() R1.pdf
I. V.M. Yakovenko (2010) "Statistical mechanics approach to the probability distribution of money", arXiv:1007.5074v1 [q-fin.ST] 28 July 2010.
J. V.M. Yakovenko (2010) "Statistical mechanics of money, debt, and energy consumption", arXiv: 1008.2179v1 [q-fin.ST] 12 Aug 2010.
K. V.M. Yakovenko (2012) "Applications of statistical mechanics to economics: Entropic origins of the probability distributions of money, income, and energy consumption", arXiv: 1204.6483v1 [q-fin.ST] 29 Apr 2012.
L. 180115 XLS Entropy In a Histogram R2.xlsx
M.

2 - Background

This series of diary notes is a rework of a set of notes partially completed in 2013-2014. In 2000 Dr Victor Yakovenko and his student (Drăgulescu) published a set of eight capital exchange models which have come to be known as the BDY model (for Benatti-Drăgulescu-Yakovenko). Later, on his website, Dr Yakovenko had produced a demonstration of rising entropy in his BDY models, and I decided to do the same in my own agent-based models (ABMs). This was all with the goal of understanding the role of the Maximum Entropy Principle (MEP) and Maximum Entropy Production Principle (MEPP) in ABMs such as ModEco or PSoup. That study and the associated diary notes were set aside for a while as I studied Odum's MPP. Now, in 2018, I want to review, update, and complete this set of diary notes.

Ref A is the paper in which I first saw the BDY models described, and Ref B is a series of email in which I discuss entropy with a friend, with comments from Dr Yakovenko. I had opportunity in April 2014 to meet with Dr Yakovenko in Toronto, and Ref C captures my discussion with him, and a set of thoughts that occurred to me over the four hours as I drove home again.

Ref D is a draft paper in which I develop a measure of entropy as exhibited in operating ABMs similar to the BDY models. This series of NTFs is being (re-)written in support of that document.

Ref E is a technical diary note describing how to write a custom function in MS Excel. This skill is needed to pursue a study of entropy in ABMs.

of measure for entropy are dimensionless numbers; and (2) that the formula I used for calculating entropy in an ABM is most closely associated with nats (units of measure from information theory) when the number of agents is very large, but diverges from nats when the number of agents is of more practical size. I therefore name my units of measure hnats, and the prepended 'h' stands for 'histogram'.

In Ref H I try to closely examine the two functions ln(x!) and GammaLn(x) that are intimately connected to my definition of entropy in histograms, and in agent-based models.

All of this was preparatory work needed to undertake the work on this diary note, and it is all background to the Ref D draft paper.

Ref L is an MS Excel spreadsheet associated with this diary note.

3 - Purpose

The purpose of this diary note is to present a consolidated description of how to calculate the entropic index of some conserved quantity within an agent-based model. I am going to do this using the two different base equations for entropy in each of the two "regimes" of thought explored at Ref F xxx.

4 - Discussion

In this diary note I am going to develop the equations for S_{max} in each of the two regimes of thought explored in the Ref F diary note, and, following that, develop the equations for S_{index} in each of the two regimes. I do this noting that it is my intention to only use the Boltzmann regime in the future, but I want to explore the connection between the two at the same time. Why? Because, when (1) the number of agents in a model is large and (2) when the histograms are not sparsely populated, then the distortion introduced by Stirling's approximation disappears and the two regimes of thought should be identical in practical outcome.

4.1 - S_{max}?

In my various readings about entropy, I admit that I have only come across the concept of "maximum entropy" in reference to closed systems. With that in mind, I am uncertain how these ideas apply to open systems. So, all of the discussion in this diary note is meant explicitly to apply to closed systems, but its applicability to open systems has yet to be determined.

In a closed system, in which some quantity is conserved in all transactions and merely transferred from agent to agent, then a history of the collective ownership of that quantity can be constructed as a time series of histograms, and there is a maximum possible value of entropy associated with each such time series. An economic version of the second law of thermodynamics would tell us that the entropy of such a time series should rise towards that maximum value and hover in that vicinity. But, there are several design characteristics of an ABM and of the entropy measurement regime that might affect the actual value of entropy calculated for any histogram, or for any time series of histograms. Using "money" as the conserved quantity, as an example, these design characteristics include:

Cancelling the As I get:

$$_{S}S_{index} = \frac{[-\sum_{i=1}^{K}(p_i \ln(p_i))]}{\ln(K)}$$	[27]

If I wish, in resolving the ln/ln division, I can write this as:

$$\boldsymbol{_{S}S_{index} = -\sum_{i=1}^{K}(p_i \log_K(p_i))}$$	**[28]**

5 - Summary

In either regime of calculations, I get a surprisingly compact formula for the entropic index. (See equations [16] and [28].)

In both cases, the maximum entropy possible is A, and the index is the entropy divided by A.

In both cases, the result is independent of the number of bins in the histogram, as long as the number of agents in the model does not change, and as long as the bin widths do not change.

The assumption that K is a constant has been relaxed, as it is of no consequence if it changes. However, the assumptions that A and W are constant has not been relaxed.