Lecture 6: 28th October 2011

Spatial Interaction: Scaling Across Space

Michael Batty

m.batty@ucl.ac.uk

twitter @jmichaelbatty

http://www.complexcity.info/

http://www.spatialcomplexity.info/

Outline of Lecture 6

Scaling across Space: Gravitation

Accessibility: In-degrees, Out-degrees & Potentials

The Family of Spatial Interaction Models

Entropy-Maximising

Defining Entropy

We will examine Symmetry and Related Concepts in the next lecture and complete our definitions of entropy …..

Scaling across Space: Gravitation

So far we have looked at scaling with respect to how an object changes in size – allometry – and how a set of objects which are all different sizes relate to one another – rank-size laws but space has been strangely absent from our treatment. It is time to redress this and this will involve us in what some have called 'The First Law of Geography' from a paper by Waldo Tobler (1970)

Tobler's Law states that: "Everything is related to everything else, but near things are more related than distant things" This is often expressed as a power law or rather an inverse power law where the strength of the relationship varies as

$$\boxed{\frac{1}{\text{distance}} \; or \; \frac{1}{d^2} \; or \text{ something like this}}$$

Tobler's first law of geography - Wikipedia, the free encyclopedia - Windows Internet Explorer

http://en.wikipedia.org/wiki/Tobler's_first_law_of_geography

Tobler's first law of geography

From Wikipedia, the free encyclopedia

The **first law of geography** according to Waldo Tobler is "Everything is related to everything else, but near things are more related than distant things."[1]

This observation is embedded in the gravity model of trip distribution. It is also related to the law of demand, in that interactions between places are inversely proportional to the cost of travel between them, which is much like the probability of purchasing a good is inversely proportional to the cost.

It is also related to the ideas of Isaac Newton's Law of universal gravitation and is essentially synonymous with the concept of spatial dependence that forms the foundation of spatial analysis.

The link structure of Wikipedia's collection of geolocated articles has been demonstrated to be consistent with Tobler's first law of geography.[2]

References

1. ^ Tobler W., (1970) "A computer movie simulating urban growth in the Detroit region". *Economic Geography*, 46(2): 234-240.
2. ^ Hecht, B., Moxley, E.: "Terabytes of Tobler: Evaluating the first law in a massive, domain-neutral representation of world knowledge". In Hornsby, K.S., Claramunt, C., Denis, M., Ligozat, G., eds.: Spatial Information Theory, 9th International Conference, COSIT 2009, Aber Wrac'h, France, September 21-25, 2009, Proceedings. Volume 5756 of Lecture Notes in Computer Science., Springer (2009) 88-105

This article about geography terminology is a stub. You can help Wikipedia by expanding it.

A basic model of forces between nodes in a graph was first articulated by Newton as his second law of motion – force is proportional to mass times acceleration. In more conventional terms we might write the force between two bodies as $F_{12} \sim M_1 M_2 / (d^2)_{12}$

There is a very long history of analogies between force and social interaction going back to Newton himself and I will add some refs to the bibliography. In fact in the last lecture we noted Ravenstein's contribution in the 1880s but apparently Carey used it for human systems before in 1850 and there is some sense in which the French speculated about such models in the late 17^{th} and 18^{th} centuries.

But let me immediately generalise this and say that we need to define many interactions – we break our system into areas or points which we define as origins and destinations i and j

And then we usually measure the distance not as distance per se but as travel cost or rather generalised cost c_{ij}

We also define the mass at the origins and destinations as O_i and D_j but first we define this as the populations P_i and P_j and we then write the conventional spatial interaction of gravity model as

$$T_{ij} \sim \frac{P_i P_j}{c_{ij}^2} = K \frac{P_i P_j}{c_{ij}^2}$$

where K is the gravitational constant

Now this is a flow model, not a network model, in terms of our previous discussion. In fact when we examine the indegrees and outdegrees, then this model lets us compute accessibilities not counts.

The equivalent model for counts which we assume are observed indegrees and outdegrees involve origin and destination activities O_i and D_j . We can now compare the models – the population gravity model & the generic model

$$T_{ij} = K\frac{P_i P_j}{c_{ij}^2} \quad \text{and} \quad T_{ij} = K\frac{O_i D_j}{c_{ij}^2}$$

These do look the same but note that the first is symmetric and the second is not – this lets us make some key distinctions

Accessibility: In-degrees, Out-degrees & Potentials

Now let us take the population gravity model and compute its indegrees and outdegrees as follows:

$$V_i = \sum_j T_{ij} = KP_i \sum_j P_j c_{ij}^{-2}$$

$$V_j = \sum_i T_{ij} = KP_j \sum_i P_i c_{ij}^{-2}$$

You can immediately see that the indegrees and outdegrees are identical, that is (if the cost matrix is symmetric of course).

$$V_i = V_j \text{ for all } i = j$$

This is defined as the potential – in fact the population potential by Stewart and Warntz (1958) in analogy to potential force in physics –the integration of force at a point.

A useful extension although it makes no difference to the symmetry of the model is to divide potential by population and produce a per capita measure as

$$\frac{V_i}{P_i} = \frac{V_j}{P_j} \text{ for all } i = j$$

In fact, let us do the same for the origin-destination model where the potentials are not symmetric; for completeness we just state the equations

$$V_i = \sum_j T_{ij} = KO_i \sum_j D_j c_{ij}^{-2} \quad \text{and} \quad V_j = \sum_i T_{ij} = KD_j \sum_i O_i c_{ij}^{-2}$$

$$V_i \neq V_j \text{ for all } i = j \quad \text{and} \quad \frac{V_i}{P_i} \neq \frac{V_j}{V_j} \text{ for all } i = j$$

Of course, it is not only the indegrees and outdegrees – in the case of a model – the accessibilities that are equal in the population gravity model but also the flows themselves as the underlying network is symmetric

That is

$$T_{ij} = T_{ji} = KP_i P_j c_{ij}^{-2} = KP_j P_i c_{ji}^{-2}$$

As long as the cost or distance matrix is symmetric which we can readily assume but only if the model is pitched at a level of aggregation where this is acceptable; as we said in the last lecture, if our models are of street systems, this is rarely ever true. It is usually only true of highly aggregate systems where

$$c_{ij} = c_{ji}$$

Ok we can define many kinds of potential or accessibility in this manner and to illustrate this let me turn to our land use transport model of London and show you how some of these play out. Here we have a variety of accessibility indicators

Help: Accessibilities Defined

Accessibility measures are computed with respect to the origin zone i which in this case is where the employment E_i is located, or the destination zone j which in this case is where the population P_j is located. A_i is the area and hence (E_i / A_i) and (P_j / A_j) are densities. c_{ij} is the travel cost from origin zone i to destination zone j. $\bar{C}$ is the mean travel cost with all these cost specific to each of the four modes. We show all these accessibility measures for the origin i zone.

Absolute Potential

$$V_i = \sum_j P_j c_{ij}^{-1}$$

Potential Density

$$V_i = \sum_j (P_j / A_j) c_{ij}^{-1}$$

Absolute Benefit

$$V_i = \sum_j P_j \exp(-c_{ij} / \bar{C})$$

Benefit Density

$$V_i = \sum_j (P_j / A_j) \exp(-c_{ij} / \bar{C})$$

These benefits are proportional to the log sum benefits which are the log of these

(Inverse) Absolute Travel Cost

$$V_i = (\sum_j c_{ij})^{-1}$$

(Inverse) Weighted Absolute Travel Cost

$$V_i = (\sum_j P_j c_{ij})^{-1}$$

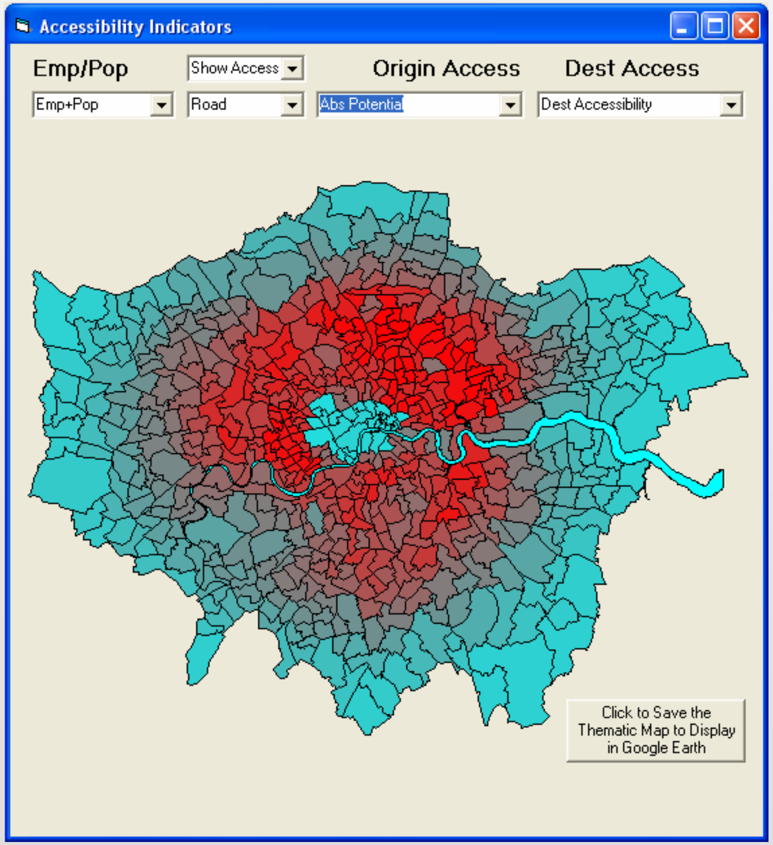
(Inverse) Weighted Absolute Travel Cost Density

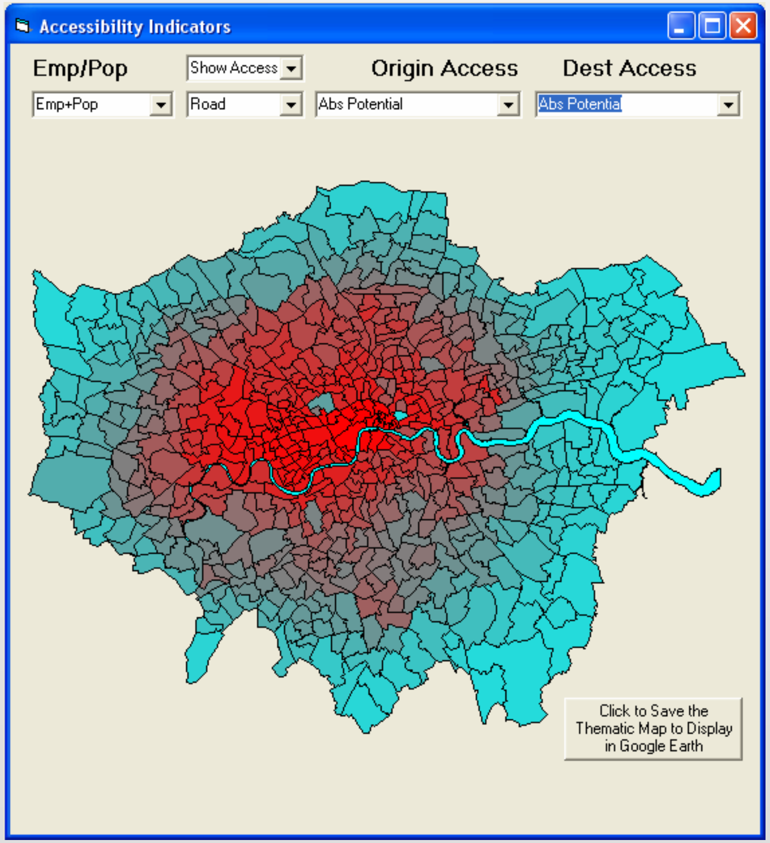
$$V_i = \sum_j (P_j / A_j) c_{ij}$$

Population within Mean Travel Cost

$$V_i = \sum_j P_j \ \mathit{for\ all}\ c_{ij} \leq \bar{C}$$

The simplest measure we compute is for origins and destinations over the road system where origins are employment and destinations are population, so they are not symmetric at all





Now I need to explain why these two maps are so different. First for employment accessibility – origins – there are large quantities of employment in the CBD and also congestion charging which makes the accessibility to employment higher outside the CBD. Hence the light blue zone In the centre has lower accessibility

For destination accessibility, there are only small volumes of population in the CBD but large ones just outside and this evens things out such that the effect of congestion charging relative to population in general in Greater London is as high in the CBD as outside it.

This is reflected in all our measures – in fact if I have time I will demonstrate all these measures in class by running the model. We will return to symmetry in the next lecture.

The Family of Spatial Interaction Models

We have already generalised the gravity model to capture the usual case where origins and destinations are not symmetric and the model we will work with from now on is

$$T_{ij} = K \frac{O_i D_j}{c_{ij}^2}$$

This is the model that has been used for years but in the 1960s and 1970s various researchers cast it in a wider framework – deriving the model by setting up a series of <u>constraints</u> on its form which showed how it might be solved and produced various <u>generating mechanisms</u> that could generate consistent models. We will first show how we produce models from constraints then from the generating methods.

The <u>constraints logic</u> led to consistent accounting

The <u>generative logic</u> lead to analogies between utility and entropy maximising and opened a door that has not been much exploited to date between entropy, energy, urban forms physical morphology and economic structure

In particular the economic logic rather than the energy-entropy logic was called choice theory, specifically discrete choice.

Let us redefine our terms. We will refer to the size of volume of origins and destinations not as population P any more but as O_i and D_j assuming they are different from one another. We will also assume that the inverse square law on distance or travel cost does not apply and that whenever c_{ij} appears it will be parameterised with a value that varies which we call λ

We will assume trips are as we have defined them as T_{ij} but we will also normalise trips by their total volume T to produce probabilities which we use in the generative mechanisms

$$p_{ij} = \frac{T_{ij}}{T} = \frac{T_{ij}}{\sum_{ij} T_{ij}}$$

Note that we use summation extensively in what follows

We must move quite quickly now so let me introduce the basic constraints on spatial interaction and then state various models. The constraints are usually specified as origin constraints and destination constraints as

$$O_i = \sum_j T_{ij} \qquad D_j = \sum_i T_{ij}$$

And we can take our basic gravity model and make it subject to either or both of these constraints or not at all

So what we get are four possible models which are key members of the family of such models

<u>Unconstrained</u> $T_{ij} = KO_i D_j c_{ij}^{-\lambda}$ subject to $\sum_{ij} T_{ij} = T$

<u>Singly (Origin) Constrained</u> so that the volume of trips at the origins is conserved $T_{ij} = A_i O_i D_j c_{ij}^{-\lambda}$ subject to $\sum_j T_{ij} = O_i$

<u>Singly (Destination) Constrained</u> so that the volume of trips at the destinations is conserved $T_{ij} = B_j O_i D_j c_{ij}^{-\lambda}$ subj to $\sum_i T_{ij} = D_j$

<u>Doubly Constrained</u> trip volumes at origins + destinations are conserved $T_{ij} = A_i B_j O_i D_j c_{ij}^{-\lambda}$ st $\sum_j T_{ij} = O_i$ and $\sum_i T_{ij} = D_j$

The first three are location models, the last a traffic model. Ok, so what are these parameters that enable the constraints to be met – well they can be easily produced by summing each model over the relevant subscripts –origins or destinations and then simply substituting and rearranging. I will do this but I will leave you to work through the algebra in your own time and many of you will know this anyway. Here are the factors which are sometimes called balancing factors

Unconstrained $K = T \Big/ \sum_i \sum_j O_i D_j c_{ij}^{-\lambda}$

Origin Constrained $A_i = 1 \Big/ \sum_j D_j c_{ij}^{-\lambda}$

Destination Constrained $B_j = 1 \Big/ \sum_i O_i c_{ij}^{-\lambda}$

Doubly Constrained $A_i = 1 \Big/ \sum_j B_j D_j c_{ij}^{-\lambda}$ $B_j = 1 \Big/ \sum_i A_i O_i c_{ij}^{-\lambda}$

Entropy-Maximising

Now we have only dealt with *constraints* through consistent accounting – we now need to deal with generative methods that lead to the same sort of accounting– entropy maximising, information-minimising, utility-maximising and random utility-maximising, and also various forms of nonlinear optimisation – in fact all these methods may be seen as a kind of optimisation of an objective function – entropy utility and so on – subject to constraints

We will define entropy maximising. First we define entropy as Shannon information and we convert all our equations and constraints to probabilities. Shannon entropy is

$$H = -\sum_i \sum_j p_{ij} \log p_{ij}$$

We maximise this entropy subject to the previous constraints – dependent on what kind of model we seek but noting now that we need another constraint on travel cost which is equivalent to energy so that we can derive a model

We thus set up the problem as

$$\max \quad H = -\sum_i \sum_j p_{ij} \log p_{ij}$$

subject to

$$\sum_j p_{ij} = p_i$$

$$\sum_i p_{ij} = p_j$$

$$\sum_i \sum_j p_{ij} c_{ij} = \hat{C}$$

But note that the probabilities always add to 1, that is

$$\sum_i \sum_j p_{ij} = \sum_i p_i = \sum_j p_j = 1$$

I am not going to work this through by setting up a Lagrangian and differentiating it and then getting the result. There is a lot of basic algebra involved and all I want to show is the result. You can find this in any standard text on spatial interaction

For this optimisation the model that we get can be written as

$$p_{ij} = \exp(-\lambda_i - \lambda_j - c_{ij}^{-\lambda})$$

or

$$T_{ij} = Tp_{ij} = A_i O_i B_j D_j \exp(-\lambda c_{ij})$$

Let us note many things

1. This is the doubly constrained model but with an exponential of travel cost replacing the inverse power
2. We can get any of the other constrained models in the family by dropping constraints and we can do this directly

3. We can begin to explore what entropy means by substituting the probability model into the entropy equation – paper to read on this relevant to this lecture is my GA 2010 (below).
4. We can think of this method as one in which the most likely model is generated given the information which is in the constraints
5. In terms of statistical physics, this model is essentially the Boltzmann-Gibbs distribution
6. Entropy can be seen as utility under certain circumstances
7. We solve the model – i.e. find its parameters by solving the entropy program which is equivalent to solving the maximum likelihood equations
8. We can then use this scheme to develop many different kinds of model – where we add more and more constraints and also disaggregate the equations to deal with groups

Residential Location, Modal Split

Let me illustrate in two ways how we can build models using this framework

First if we say that residential location depends on not only travel cost but also on money available for housing we can argue that

1. The model is singly constrained – we know where people work and we want to find out where they live – so origins are workplaces and destinations are housing areas
2. The model then lets us predict people in housing
3. We argue that people will trade off money for housing against transport cost

And we then set up the model as follows

It is

$$\sum_j T_{ij} = O_i$$

$$\sum_i \sum_j T_{ij} c_{ij} = C$$

$$\sum_i \sum_j T_{ij} R_j = R$$

leads to

$$T_{ij} = A_i O_i \exp(\vartheta R_j) \exp(-\lambda c_{ij})$$

Note that R_j is the average house price in j and R total housing costs in the system. We can find out from this location model how many people live in destination housing zones, so it is a distribution as well as a location model

$$P_j = \sum_i T_{ij}$$

If we want a modal split model, we can break the trips into different modes and then let the modes compete with locations for travellers

In this way we produce a combined modal split location model.

Sometimes we may want the modes to be constrained and in generating specific constraints on total travellers by mode, this is equivalent to adding parameters that distort the travel costs – in fact the generic equation can be seen as one where the travel cost or energy is modified by the volume constraints

That is

$$p_{ij} = \exp(-\lambda_i - \lambda_j - c_{ij}^{-\lambda})$$

$$or$$

$$T_{ij} = Tp_{ij} = A_i O_i B_j D_j \exp(-\lambda c_{ij})$$

Defining Entropy

At one level, you don't need to know what it is. You just need to be familiar that there is a technique of maximising a quantity subject to known information – constraints

You could think of this quantity as Accessibility or as Utility – in fact many people do.

Maximising utility is easy enough to understand

Now there are some very useful insights if we think of entropy as information. So we maximise information rather than entropy but there are some really interesting issues about entropy and thermodynamics that we don't have time to go into here. To give a taste of these, we need to look at the properties.

So this is a bit of digression to begin with but let us not forget that this mysterious quantity called entropy is not widely understood even by physicists, perhaps especially by physicists.

Von Neumann to Shannon in 1948 says it all:

"You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one really knows what entropy really is, so in a debate you will always have the advantage!"

Ok. Let me first state the formula for entropy as information which Shannon derived. It is the same as we have been using

$$H = -\sum_{i=1}^{n} p_i \log p_i$$

How do we get this? Now we can get it many ways but the easiest in my view is this. We define information from the probability of an event occurring p_i. If the probability is low and the event occurs, the information gain is high

and vice versa, so we define raw info as

$$\frac{1}{p_i}$$

But if an event occurs and another event occurs which is independent, then the raw info is

$$\frac{1}{p_i p_j}$$

Now information gained should be additive, we should be able to add the first info and the second info to get this but

$$\frac{1}{p_i p_j} \neq \frac{1}{p_i} + \frac{1}{p_j}$$

The only function to do this is the log of

$$\log \frac{1}{p_i}$$

And we thus write the information as follows

$$\left.\begin{aligned} F\left(\frac{1}{p_1 p_2}\right) &= F\left(\frac{1}{p_1}\right) + F\left(\frac{1}{p_2}\right) \\ -\log(p_1 p_2) &= -\log(p_1) - \log(p_2) \end{aligned}\right\}$$

And if we take the average or expected value of all these probabilities in the set, we multiply the info by the probability of each and sum

To get

$$H = -\sum_{i=1}^{n} p_i \log p_i$$

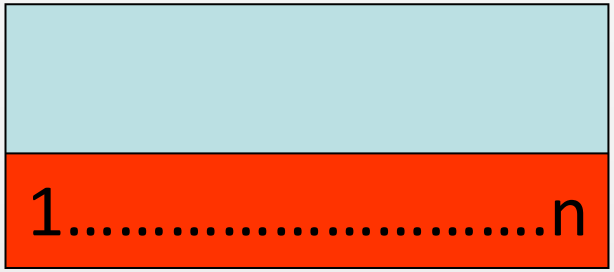
Now entropy or information is large – big – when all the probabilities are the same – uniform

$$p_i = \frac{1}{n}$$

And it is small – in fact 0 – when one probability is 1 and the rest are zero

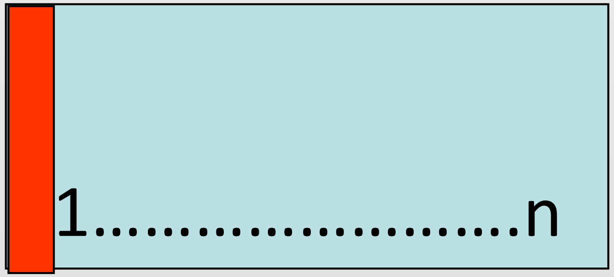
$$p_i = 1, \quad and\ the\ rest\ are\ p_j = 0, \forall j \neq i$$

We can draw a graph of all these probabilities as follows – first when there are all equal



p = 1/n and Entropy H = max

And then when only one is equal to 1



p = 1, the rest 0, Entropy H = min

In the first case there is extreme homogeneity and in the second extreme order. These profiles are like population density slices.

Essentially in E-M, we choose a probability distribution so that we let there be as much uncertainty as possible subject to what information we know which is certain

This is not the easiest point to grasp – why would we want to maximise this kind of uncertainty – well because if we didn't we would be assuming more than we knew – if we know there is some more info then we put it in as constraints. If we know p=1, we say so in the constraints. Let us review the process,

I will repeat it all in standard statistical mechanics terms:

Maximise $H = -\sum_{i=1}^{n} p_i \log p_i$

Subject to $\sum_i p_i = 1$ and $\sum_i p_i c_i = \overline{C}$

We can think of this as a one dimensional probablity density model where this might be population density

And we then get the classic negative exponential density function which can be written as

$$p_i = \frac{\exp(-\lambda c_i)}{\sum_i \exp(-\lambda c_i)} \qquad , \qquad \sum_i p_i = 1$$

Now we don't know this is a negative function, it might be positive – it depends on how we set up the problem but in working out probabilities wrt to costs, it implies the higher the cost, the lower the probability of location.

We can now show how we get a power law simply by using a log constraint on travel cost instead of the linear constraint.

We thus maximise entropy subject to a normalisation constraint on probabilities and now a logarithmic cost constraint of the form

Max $$H = -\sum_{i=1}^{n} p_i \log p_i$$

Subject to $$\sum_i p_i = 1 \qquad \sum_i p_i \boxed{\log c_i} = \bar{C}$$

Note the meaning of the log cost constraint. This is viewed as the fact that travellers perceive costs logarithmically according the Weber Fechner law and in some circumstances this is as it should be.

If we do all this we get the following model where we could simply put $\log c_i$ into the negative exponential getting

$$p_i = \frac{\exp(-\lambda \log c_i)}{\sum_i \exp(-\lambda \log c_i)} \quad \Rightarrow \quad p_i = \frac{c_i^{-\lambda}}{\sum_i c_i^{-\lambda}}$$

A power law. But this is not the rank size relation as in the sort of scaling we looked at last week. Let us see if we can get such a relation. But this will be next week. Before we go, let us give one reference to my GA 2010 paper

Space, Scale, and Scaling in Entropy Maximizing, *Geographical Analysis 42* (2010) 395–421 which is at http://www.complexcity.info/files/2011/06/batty-ga-2010.pdf

The blog will have more and more references as the course continues

Questions

www.complexity.info

www.spatialcomplexity.info