Spatial Statistics and Spatial Econometrics
Prof. Gaurav Arora
Department of Social Science and Humanities
Indraprastha Institute of Information Technology, Delhi

Lecture - 9
Spatial Entropy

Hello everyone, welcome back to Spatial Statistics and Spatial Econometrics; we are at the 7th lecture today and we will be looking at this topic called Spatial Entropy. In the previous lecture, we extensively covered the concept of entropy and in summary what we said was that, entropy provides a measure a measure of dispersion, a measure of dispersion for a given random process. In that spirit, entropy was characterized by us as an alternative, alternative measure to variance and standard deviation, ok.

And of course, you know when we are talking about variance and standard deviation, what we are really talking about is the second moment, second moment of a random process, right. And also as part of the previous lecture, we looked at the exponential distribution at length.

So, today what we are going to do is, we are going to sort of enhance or expand this idea of entropy in space. So, we are going to delineate what we studied as exponential distribution, which was to model a random process that characterizes time between regularly occurring events, regularly and independently occurring events to a spatial domain, right.

So, we will you know look at exponential distribution and its entropy in space, ok. So, let us move forward.

So, here I have an example of a monocentric city. So, we are talking about space already, we are talking about a city. So, a city is spread in space, right. So, for example, you can think about you know National Capital Territory of Delhi, you can think about Kanpur, you can think about Jaipur, you can think about Chennai and any other city alike.

So, what we say is let us that, let us consider a linear city with plausible location markers given by r in between 0 and a capital R units from the region, ok. So, I have used a few terms here, first of all I am saying I am talking about a linear city; I am talking about plausible location marker. So, it is not that we will locate markers everywhere in this linear city; we are going to have some plausible locations, where we can find these or locate these markers. What is a location marker? Well you can look at figure 1 to look to find a location marker, right.

A location marker is really when you are driving through a highway, when you see a kilometer marker which tells you ok in this example Haridwar is 24 kilometers away, that is a location marker rise, right. And when I look at location markers away from the origin, I can get a sense of how far out I have come from the city center, right.

And there is a probability attached, there is a probability attached to finding a location marker, which is you know given for every r distance away from the origin as p r equals K exponential minus lambda r where r goes from 0 to R. And K is a normalizing coefficient which just ensures that the sum of probabilities or the integral of probabilities in throughout this city must equal to 1, right.

So, probability in a defined domain d must equal to 1 for all the events that can occur, that is all the markers that we locate, ok. So, let us begin solving this problem, so ok. So, we are given what? We are given, first we are given location markers placed along a linear city right with and the city has an origin and a length or a radius R, ok.

So, my city starts at the origin; it is a linear city, so it goes out till the radius or length R, ok. So, I am talking about a spatial domain d, right. So, in previous lectures, we have talked about geostatistical data and we said that look you know the data could appear at any point in space.

So, here we have a is a 1 dimensional, a 1 dimensional spatial domain that we are working with right and D here is nothing, but real space in the positive direction. So, I do not have negative values of R, right. So, what I am saying is that if I go out a given distance from the origin; let us say I go out R units or R kilometers out from this origin, so let us think about kilometers being the units of measurement here. So, if I go R kilometers out, I am going to find a location marker I am going to find a location marker with probability p r, ok.

So, the probability of observing a location marker at an arbitrary distance from origin distance r from origin is p r, which is written as K times exponential minus lambda r right, such that r can be between 0 and R, right. So, that is reasonable, I mean r how far can I go from the origin at max that is capital R right; I can at max go till capital R or you know I can just stop at 0. So, I can have this arbitrary location or distance from the origin going from 0 to capital R.

So, what is K? We can say where K is a normalizing coefficient; I know all of this is already written, but I want to just ensure clarity, such that integral 0 to R p r d r equals 1, ok. So, before we move forward what we are going to do is, we are going to do an aside where we create this analogy between the exponential distribution that we had used in the previous lecture to model time to then model it own space. So, because we want to bring that idea of a exponential distribution, which we understand traditionally, right.

It is a where the random variable is basically time between two different independently occurring events; two space right, which is a city now, right. So, now, I am talking about modeling you know a random process of observing markers with some probability at a given distance from the origin as an exponential distribution.

So, I am going to now as a next step, I am going to create an aside an aside and I am going to sort of create this analogy for you.

So, in the previous lecture in the previous lecture, what we said was that we have a random variable x which is distributed exponential with parameter lambda and that implied and that implied that the distribution the PDF, the frequency of observing x you know a given value of x will be given as lambda e to the power minus lambda x, right.

And in that case we had said x could be greater than or equal to 0. So, there we had allowed for x to go between 0 and infinity, right. So, we just wanted x to be greater than 0 and after that we were ok even if it became very very large to the extent that it is not it is not defined, ok.

So, of course you know x was modeling, this was time between events that occur independently and regularly at or with an average rate average rate lambda, right. So, if we if we are to sort of adapt it to the notation that we have, you know we can simply think we can just replace, we can just replace x with r and we have our example right, to create an analogy with our example of a monocentric city, right. So, now why we do that, the first thing that we have to note is the. So, we have to note a couple of things, right.

We note a couple of things when we sort of bring this idea from time to distance right, from temporal dimension to the spatial dimension. Well, we will say note first of all we have reinterpreted time between events to what? To distance between two location markers right or distance between origin and a location markers, to distance between origin and location markers. So, that is the first sort of distinction for in what we have done in moving from entropy to spatial entropy.

So, this is entropy, time between events and what is distance between location markers origin and location markers, this is spatial entropy, ok. And one thing you realize straight away is that the second pointer that I want to note is that, if R were to be infinity in our example, right. So, what is R? Well, you know we know r can small r can go from 0 to capital R; but if r were infinity, we would be exactly mapping to the to the time dimension problem, then we will simply get K equals lambda.

If you want to see it, we can see it. So, the condition that we have is 0 to R d r. So, we have K, such that the probability integral through the linear city of a finite radius is equal to 1, right. And here if I were to solve this what I am going to get is 0 to R K times e to the power minus lambda r d r equals 1, now we are setting R equals infinity, right. So, R approaches infinity right more, if you want to be more correct and exactly correct, right.

So, if we integrate the left hand side I am going to have K times e to the power minus lambda r over minus lambda, this goes from 0 to infinity and the whole thing on the left hand side equals 1, which will give me simply K over lambda equals 1, which implies k equals lambda, ok. So, we have sort of as a second pointer what we have done is that, we have reinterpreted an infinite or unbounded time sequence to a finite linear city in space in the over the real number line, right. So, on the positive you know quadrant, right.

So, we have this implies that we have reinterpreted, we have reinterpreted an infinite or unbounded time sequence in case of what I am going to call as the traditional exponential distribution ok to a finite linear city of given radius capital R, right. So, in other words if I had a infinite linear city that is R goes to infinity; I could simply say that you know the probability of observing a marker at a distance small r from the origin is simply lambda exponential to minus lambda r, right.

So, I am just going to say this very quickly; in other words in other words if I had an infinite

infinitely large linear city, right. So, a very very large linear city which would just not end over space, that is I can say R capital R goes to infinity mathematically that is our expression for that; I could I can say that probability of observing a marker at a distance small r from the origin is simply given as lambda e to the power minus lambda r such that r goes r lies between 0 to infinity as on your screen, ok.
So, now we have completed this exercise of adapting of adapting what we had sort of seen as a spatial interpretation of the exponential distribution that we covered in the last class to understand entropy to understand spatial entropy over space, right.
So, now, we are going to go step by step use this you know probability distribution of location markers over a linear city with the schematics that are there on your screen, right. And then sort of go step by step and get to the point, where we understand entropy of a city and how do we interpret it in the physical world, ok. So, let us do that.

So, I am going to do that in steps I am say, I am going to say step 1. So, let us see what is step 1. So, the first step is to let us find K right and K you know if R is infinity K is simply lambda. So, if R is not infinity and R is a fixed capital R radius of the city K is likely to depend on capital R, ok. So, let us look at that.
So, let us find K with or let us say for a linear city for a linear city of radius capital R, ok. So, my condition is already sort of pretty clear 0 to R p r d r equals 1; this implies K times exponential minus lambda r d r, where the integral integrand ranges from 0 to R.
So, I am going to simply equate this to 1, ok. And if I sort of now sort of follow through the integration, K itself is a constant with respect to r in this integration exercise. So, I have K integral 0 to R e to the power minus lambda r d r equals 1; that would mean K e to the power minus lambda r this integration is straight forward 0 to R equals 1, this implies K e to the power minus lambda R by minus lambda minus e to the power minus lambda 0 over minus lambda, right. So, this is lambda time 0 right, I hope this is this is clear.
So, I am sorry about that. So, minus lambda into 0 right and this whole thing is equal to 1, ok. So, we have K e to the power minus lambda R over minus lambda minus. So, this entity the numerator is simply equal to 1, because e to the power 0 is 1, right. So, 1 over minus lambda with a negative sign in front, so I have 1 over lambda equals 1; this implies that I can write K equals lambda over 1 minus minus lambda R, ok.
So, now let us analyze this value K, right. So, first of all I have a the denominator; if I look at if I focus on the denominator what I see here is that; it is 1 minus some entity. So, what is this some entity? This entity is e to the power minus lambda R, R by itself is a positive number, lambda is positive right; it is the rate of occurrence, so average rate of occurrence of events that follow exponential distribution.
So, the rate has to be a positive rate likely between 0 and 1, because it is the rate of occurrence right, that is the physical interpretation. So, what I have is exponential to the power in negative entity.

So, what I have is 1 over exponential to a positive power; that means what I have is 1 minus an entity that will be vary between 0 and 1 right, it can go to 0 and 1 with close brackets. That means, I have a denominator which will be slightly or somewhat lower than 1, but still positive. So, my K which will imply that, my K will be at least as large as lambda right. So, my K will be at least as large as lambda. So, if I want to do some analysis, if I want to do some analysis; I can say that ok I want to know as R increases what is the impact on K and as you know lambda increases what is the impact on K.

So, for the case of lambda for the case of lambda; we can how do we evaluate it? Well, the way to evaluate it mathematically is to first do for case of lambda, we would have to figure out what is del K by del lambda and for R we will have to figure out what is del K by del R, ok. So, let me evaluate del K by del R for you and then you can evaluate del K by del lambda at your own time, ok.

So, del K by del R by itself is going to be given as, ok. So, I have a lambda times minus 1 which is the power on 1 minus exponential minus lambda R. So, I am going to apply, I am applying the chain rule to conduct this differentiation. So, I am now going to say minus to the e to the power minus lambda R multiplied by minus lambda; whole thing divided by exponential, sorry it is going to be 1 minus exponential 1 minus exponential to the power minus lambda R the whole squared.

And if you look at it, it is going to be less than 0, ok. So, as R goes up, K will decrease and what you can find is that as lambda goes up, the rate of occurrence goes up, K will increase by evaluating this differential, ok.

So, what we really understand is that, you know we can then sort of evaluate, we can then evaluate you know we. So, we know one thing. So, if we go back to the expression of K, it depends on two variables say; one is lambda and second is R. So, if I fix lambda, if I fix lambda right and I increase R; my understanding is that K will go down, ok.

So, let us go and check. So, if I fix lambda, if I fix lambda and sorry about that. So, this is the PDF. So, first the first graph that I am showing you is basically p of r given as e to the power minus lambda r times over 1 minus e to the power minus lambda R, ok.

So, just a little fumble there, sorry about that. So, I have e to the power minus lambda r over what I have is K is 1 minus e to the power minus lambda capital R. So, just be careful that on the numerator I have e to the power minus lambda small r and in the denominator I have 1 minus e to the power lambda times capital R.

So, by the way there is also a lambda sitting here at the numerator, because K is equal to lambda over 1 minus e to the power minus lambda R, right. So, having understood that, we want to understand how does p vary by the two variables in this distribution; one is capital R and the other is the lambda. So, 0 to R, right. So, you can see that as capital R increases, the propensity or the probability of observing a location marker dips down exponentially. So the curve the curve, the shape of the curve comes from the exponential function right.

So, the probability is very very high to observe a location marker near the origin; but as soon as I leave the origin and I start going out, this probability comes down steeply with an exponential rate you know and eventually becomes 0 at capital R. So, if I were to set R as 20, I am going to have this yellow line touch 20, right. If it were going up to 40, then of course you know it will keep going down till it sort of touches 40, ok.
So, that is the; that is an understanding. And what happens with lambda? What you see is that you know as you know as if you fix a particular radius, let us say you fix a radius of 6 and you vary lambda; as the rate of occurrence increases, the probability at that given radius goes down, right. So, I am going to have a lower p r value, if I were to look at different values of lambda and as lambda goes up, p r comes down. So, that is about p r, right.
Now, we want to understand about K. So, that is the next step. So, let us look at it.

So, this is K R over lambda, right. So, and lambda is fixed at different values. And what I am showing you here is that as R increases K of R is going to fall down right and it is happening at a exponential rate, which is not surprising; because if you look at the formulation of K, it has R as part of an exponential function, right.
So, that is an understanding or analysis of this normalizing coefficient K. So, I am looking at both the density function, I am plotting it just to give you a flavor of how it looks like and what is the shape and all that, so that does not remain a you know a mathematical functional form you understand that look; when I am going out from the origin 2 kilometers this and I can draw this vertical line and figure out what is the probability of observing a location marker given a fixed lambda.
And then I can go out 6 kilometers and I can see that you know probability of observing the location marker is much lower and this fall in probability is exponential in nature, which is where the exponential distribution comes into picture. And finally, you know to analyze how does K behave as a function of R, you know we plot K as a function of R by fixing lambda at a given value. So, I am going to look at just lambda equals 0.1 and I am going to look at K R values and see what it looks like, ok.
So, that is step 1 in our movement in our journey from going to a exponential distribution interpreted in in the space and then finally, we are going to go towards entropy, ok. So, now, we have understood the distribution by itself how does it look like, how do we interpret it, what is K and so on and so forth.

And as step 2, as step 2 you know we are going to evaluate the first moment; the first moment which is the mean distance from origin, means distance from origin called as R bar right for our linear city having radius capital R and consequently you know consequently K as lambda over 1 minus exponential to the power minus lambda R right, which we have just established.
So, we are going to look at this step 2 as the you know as the second sort of module of this lecture. So, we will join you shortly.