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

Lecture - 10
Spatial Entropy - II

So, now we have the task of evaluating the mean distance R bar for a given probability distribution, right? That is kind of straightforward we have an interpretation. The interpretation is that there is a linear city. It starts as starts at the origin. It goes out till a finite radius capital R, right and we just want to understand what is the mean distance right. So, what is the mean distance you know at which a marker location marker will be found going out from you know the origin, right?

So, we want the first moment R bar, right? So, we have a probability p(r) that we know we are going to be able to see a location marker observe a location marker and this is given as K e to the power minus lambda r, and K we have now evaluated is lambda over 1 minus exponential to the power minus lambda capital R so, times e to the power minus lambda small r.

So, ok let us move forward. So, we have the R bar which is given as something we are very well aware of 0 to R small r p(r) dr, right and I have I am going to try to write down I am going to write down r lambda e to the power minus lambda small r over 1 minus e to the power minus lambda capital R dr. I am going to take everything that is not a function of small r which is variable or which we are integrating outside the integration function, right? So, let us do that.

So, I have lambda over 1 minus e to the power minus lambda capital R careful that you know in the denominator we have capital R; capital R is fixed. So, the radius of the city is fixed right. So, integration 0 to R r times e to the power minus lambda r dr. This is something we did in the last lecture. So, we're going to do it slightly quicker. So, you can look at the last lecture and follow through with this integration. This is done by applying integration by parts, right? So, we designate a 1st and 2nd function.

So, we rank them and you know we know that integration by parts is nothing but some if you have an integral with a function of r times a function of g you know a function f of r function g of r multiplied with each other and you must integrate them with respect to r what you have is f of r g(r) dr minus integral del f by del r multiplied by integral g(r) d r whole thing integrated with respect to r, ok. So, I am going to simply apply this formula here ok.

So, I have lambda over 1 minus exponential minus lambda r sitting outside. So, I have r sitting as is e to the power minus lambda r integrated with respect to r. So, I have the following evaluated between 0 and R minus integral del r by del r is just 1 times e to the power minus lambda r over minus lambda dr, ok. So, this is the second component and I am going to just keep solving this ok. So, I am going to take it to the next page as I am going to try my best. So, we have ok.

So, we have the R bar equals ok I am going to go to the previous page a little bit and follow through. So, I have 1 minus sorry, I did and I made a mistake here. So, I think this r was supposed to be inside here. So, this r is inside here. So, you have a capital R. So, you have a capital R here is 0 sitting out here. So, this becomes simpler sorry, about the fumble.

So, you have lambda over 1 minus e to the power minus lambda R times R times e to the power minus lambda R over minus lambda minus just a 0. So, that is the first component minus the second components look like going to look like the following e to the power minus lambda R over lambda squared minus e to the power minus lambda into 0 over lambda squared ok.

So, I have the R bar given as lambda over 1 minus exponential (Refer Time: 06:47) lambda r times we are going to have 1 minus exponential minus lambda R 1 plus lambda R over lambda which can be further simplified. So, the R bar can be further simplified as 1 over lambda minus R e to the power minus lambda capital R divided by 1 minus e to the power minus lambda R.

And, this representation is particularly useful because if you recall from the previous lecture for an exponential distribution with an infinite radius when R was capital you know capital R was infinity. The exponential distribution would have given me a mean value that is exactly equal to 1 over lambda this is something that we saw last time.

And, now what happens when R is a positive you know R is a finite positive value I have 1 over lambda minus a and entity which is given as what you see on your screen and this entity is going to be a positive number. So, what I am saying is that the mean distance with a finite radius city is going to be smaller than the mean distance that we see with infinite for our infinite radius city which makes sense; that means, that you know when I have an infinite radius city I am going to have a larger mean distance than if I were to have you know slightly lower you know a sort of finite sort of city which is only as big enough right.

So, of course, you know the exercise that naturally follows is to evaluate or analyze the relationship between the R bar and this finite radius of the city R, ok. So, now, I am going to sort of ask you to stop for 5 minutes and evaluate this relationship ok. So, I want you to stop for 5 minutes ok. Stop for 5 minutes.

You can pause this video for 5 minutes and evaluate this value and then will come back and then you know we will resolve the matter, ok. Welcome back from the break we had given ourselves an exercise to evaluate the R bar which is the mean distance first moment of our spatial exponential distribution interpretation. We wanted to see how this R bar would depend on the finiteness of the city. So, this capital R if it goes to infinity, well that is an infinite city, that is an infinitely large linear city.

And, as soon as you restrict it to let us say 100 kilometers a 150 kilometers I mean you have a more realistic situation, but from the perspective or point of view of spatial distribution, it is a finite city that we are trying to model. So, we wanted to understand how would R bar depend on capital R. What you see on your screen is a plot of the R bar value.

As we change the value R and that one thing that we have discussed there are two variables in this distribution one is lambda which is the rate of occurrence of location markers in the city; the second is the extent of the city itself which is capital R. So, what do I do? I can fix the value of lambda; I can vary R, right? So, I can fix lambda, let us say lambda is fixed at 0.1, right?

Then I can vary the value of capital R, let us say we can go from 10 kilometers to 20 kilometers to 30 kilometers to 40 to 50 to 60 and all the way let us say we go to 200. So, we have these 20 unique values of capital R and what we can do is we can simply sort of you know throw in or fill in the R bar given the formula that we derived before the break, right?

So, we can do that. So, we can get the R bar from the formula that we derived in the previous you know on the previous slide and then we can plot them, right? I mean we have an R bar on the y-axis we have R on the x-axis and we can see the relationship ok. That is one way of doing it the other way is that I have a mathematical relationship between R bar and capital R.

So, I have R bar as a function of R lambda and you know it is an exponential form and everything that we saw in the previous on the previous slide, but you know we can simply you know we can say alternatively we can simply derive del R bar by del R which automatically by definition you know restricts d lambda equal to 0. So, lambda is kept fixed which I was doing manually when I was doing it on excel numerically drawing the plot.

But, I can do it using you know the differentiation operator and then I can sort of understand whether I find del R bar by del R to be a positive entity, a negative entity or you know maybe even 0, there is no change when our value changes. What do we see in the graph? We see that you know when we fix lambda as at 0.1. So, there is no change in the lambda value as we change the capital R-value on the x-axis lambda is kept fixed at 0.1.

R bar seems to be increasing according to the yellow line which looks like an exponential know you know the inverse of an exponential increase, right? So, the increase there is R bar is increasing in R, but at a decreasing rate ok and it is clear as you move to a higher value of lambda let us say we move to lambda equals 0.5 from lambda equals 0.1 to lambda equals 0.5. You see that the R bar still increases as R goes up, but after a certain level let us say around 11 kilometers the change in the R bar tends to be almost 0, right?

So, it attains a maximum value of 2, right and then it sort of you know goes to 0. So, it does not matter whether my city is you know 11 kilometers big or you know has a radius of 20 kilometers or 200 kilometers you know the average distance you know at which a location marker may be found or a distance at which I will on average you know I can expect to see a location marker will be 2 kilometers.

That is the physical interpretation of this exponential distribution-based relationship between R bar and R. So, what have we done till now we have adapted the exponential distribution which models time between you know independent and regularly occurring events to a space a spatial domain, right? So, now, our adapted interpretation of this exponential distribution is to observe location markers away from the origin for a linear city of a finite radius capital R, right? So, we have first adopted that.

Second, in the definition of this exponential distribution, it is slightly different from what it was in the traditional case of time you know time between events interpretation. We had a new parameter called k we evaluated what this k was we analyzed what k was. And, as the third step as a third step what we have done is we have evaluated the first moment of you know of this spatial you know enter spatial exponential distribution.

So, what is the next step? The next step is obvious, the second moment, but now we will study the second moment which is the variation through entropy.

So, once we have a spatially delineated you know distance based exponential distribution we are going to now as the next step which I am going to call step 3 figure out what is spatial entropy for this example of a linear city having radius capital R and the probability of observing location markers at distance small r from origin given as p(r) equals lambda over 1 minus exponential minus lambda capital R times exponential minus lambda r, ok.

Now, I have a spatial exponential distribution spatially delineated where distance, space right we study distance is the most fundamental entity of understanding you know or characterizing phenomena over space. So, now, we have a distribution which is you know characterized by that distance and distance goes from the origin and we know that this radius this distance r can only go as far as the maximum radius of this linear city under consideration.

So, then you know what would be spatial entropy? Well, it will be capital E equals minus integration goes from 0 to R p(r) times log of p (r) this definition we have studied in the previous lecture I am just writing it again right and alright. So, once we know that we can then expand it by putting the value of p(r) that we have figured out in this lecture.

So, we have 0 to R with the negative sign right k which is lambda over 1 minus exponential minus lambda capital R times exponential minus lambda small r multiplied by the log of this entire entity. So, log lambda 1 minus exponential lambda R dr, ok. So, now, what is this entropy a function of? So, first of all, it again depends on lambda, but lambda is the rate of occurrence. It is a distribution parameter, but the other more critical thing that it depends on is the size of the city and the maximum radius that we provide for studying this city.

So, we say this entropy is a function of R we call it ER, ok. And, we can show this ER has a little bit of you know a complicated mathematical formulation, but it has a very interesting physical formulation. So, that is something that we will see now. It is 1 plus log 1 over lambda plus log 1 minus the power minus lambda R minus lambda R e to the power minus lambda r over 1 minus lambda R.

Now, pay attention to these components with a little bit of cosh, and with a little bit of attention so, let us see. So, in the green square bracket I have a term that is quite complicated right it is the log of 1 minus exponential minus lambda R minus lambda R exponential minus lambda R over a complicated denominator, but what is outside this green bracket is 1 over log 1 over lambda. So, 1 over lambda if you remember is exactly equal to the standard deviation of an exponential distribution when R could go to infinity.

So, if capital R were infinity what we are looking at here is that the first term is equal to 1 plus log of SD which is nothing but ER capital R is set to infinity. So, what I am looking at is the entropy of you know a linear city which can be written as the entropy of an infinitely large city plus an entity that is a function again of lambda and R, right?

So, every time I mean there is a recurring pattern here that you know the functional understanding of how different moments for the spatial and you know distribution depends is critically this radius of the monocentric city that we are looking at. So, I am just going to say that you know if you are if you feel confused about this E infinity, well you can always go back to the previous lecture and so, we can say we showed this in the previous lecture for an exponential distribution, ok.

And, it turns out as a pointer as a note I am just going to say that it turns out that E infinity which is the entropy of an infinitely large linear city is less than or equal to this finitely sized monocentric city. The center is you know it; obviously, a has an origin and it has a finite radius capital R ok.

So, the entropy the variation of how you know these location markers would be observed over space would be higher you know in a finite space than if my space in my then if my constraint of the size of the city was relaxed and it was allowed to expand infinitely, ok. So, the entropy of a smaller city with sort of equal content you are an equal number of markers to appear over space, and the variation at which you are going to see these markers will be realized at a higher level than if the city was larger ok.

So, we right away get an interpretation of what to expect for things like population density, for real estate prices in a large city versus a small city given that lambda is held constant, right? So, here lambda you can think of is a control parameter of you know how frequently can you see a marker. So, how what is the interpretation in the real world? Well, you can think of it as the level of economic activity, right as the total population that I you know I expect in the city.

So, if it is a city of a 2 million population, and if we change the size of the city, how will the entropy change? So, it is kind of logical to think I am going to fit a lot of stuff sort of content in a smaller space I probably will see a lot of you know kind of higher variation, alright. So, as a next step, we will define so, so till now we have before we go to the next step I mean we have ER we have a formulation for ER we know that ER is E infinity plus some stuff, ok and this some stuff is a function of lambda and R, alright we understand that now ok. So, now, you know what the next step is.

I am going to call it to step 4 we are going to define the difference between entropy levels at finite rate radius R and with R approaching infinity. So, I am just saying E difference E diff equals ER minus E infinity which clearly from the first equation that I have written is nothing but some stuff right this sum stuff which is a function of lambda and R.

So, if I were to go on to the previous slide I can write down the formulation of this remainder you know which is a function of lambda and R and it will come out to be lambda e to the power minus lambda R times R divided by 1 minus lambda R minus log 1 minus lambda R, ok. What is this difference capturing? It is the distance of variation, ok.

Now, the second moment how different is the variation captured by an infinitely large city and a city which has a finite radius R? So, it is the difference in the second moment measure for a finite city relative to an infinitely large city and I had said that E infinity is usually lesser than ER

and we also sort of covered you know we just covered a and a physical interpretation of the same.
You can show that del Ediff by del R. So, whenever I do this whenever I use the differential operator what I am doing is I am fixing the other parameter which is lambda, right? So, I am trying to you know create an understanding of what is the impact of this capital radius R and it turns out that this is less than 0. So, that means, as I increase the size of my city I get closer and closer to the entropy levels of an infinitely large city, ok.
And, what is going to be interesting for us as a next step, right? So, I am going to say step 5 and I am going to cover it in 2 minutes, but what is going to be interesting is the proportion of variation or variability captured by a finite city finitely sized city relative to an infinitely sized city infinite radius city. So, what I am going to be interested in going forward is to figure out this proportion ER minus E infinity divided by ER ok.
So, I am normalizing the level of ER. So, you know I need to sort of create a percentage or a proportion to be able to normalize the level of entropy that I am studying. So, if I have a city that is 50 kilometers you know wide from the origin, then I want to sort of make sure that I am when I am comparing that with a 100-kilometer city, I am not you know I am not comparing apples and oranges. I am creating a proportion to normalize you know this difference by you know the level of entropy that I began with ok.
So, I am trying to see what is the extent or how what is the distance between the entropy that I would get if I had all the space I wanted versus if I have a smaller space, right? What would be a good optimal size or appropriate size of a city if I wanted most of the stuff most of the distribution in population the way the population is spread out the prices are spread out and all that if I had no space constrained whatsoever, right?
So, this sort of comes closer to some kind of a regional policy or regional planning exercise where policymakers might want to understand where should they draw a boundary of a city, right, where should we stop the national capital territory of Delhi, where should we stop you know a coastal city like Mumbai which can only expand in one direction, right.
So, this idea of spatial entropy and this metric Ediff by ER specifically provide a policy instrument to do that. Before I move forward in that direction what I am going to do is I am going to look at Ediff this difference and ER a little bit more closely. So, let us go and do that. So, first of all, I am looking at ER. So, I am sorry about the notational goof up these are all E sub-R plotted as a function of capital R.
So, this is the entropy captured as I increase the size of the city. So, the entropy captured goes up as we know increases the size of the city whereas, this different sort of comes down. So, del Ediff by del R as I said will be a negative entity, right? So, this is nothing but del EDIFF by del R less than 0 entity, right?
So, what it shows me is that as I increase the size of the city the difference between what would be the variation in an infinitely large city and how would it start to fall down and of course, you again see you know either an exponential functions shape when you see these curves or you see

the inverse of it, right. So, the shape is all determined by the analytical functional form of the distribution that we began with.

So, if we begin with a different distribution let us say we began with a normal distribution, a truncated normal distribution, a Poisson distribution right, or a beta distribution, then all of these results will automatically change according to how you know different parameters of those distributions behave or lend to a behavior of the distribution itself, ok.

So, having said that you know we have understood you know E sub-R, E difference now we are going to we wanted to understand a little bit more about this proportion Hdiff or Ediff over ER to be able to understand you know an appropriate size of a city. So, it turns out that I have a functional form for you know EDIFF which is a difference between a finitely sized city and an infinite size in the city, and the entropies between the two normalized by the level of entropy for this finitely sized city and I had said you know we have both of these the numerator as well as nominator.

They are both functions of lambda and R, right? We want to understand how much how close a finitely sized city is to an infinitely sized city in terms of the variation of a given interest you know metric of interest if you know we change the radius of the city itself. So, if we expand the city how much closer am I going to get to you know to an infinitely large city as well?

So, I am going to call this step 5 and in this step 5 I am going to set this ratio equal to 0.05 or 5 percent. So, this allows for the difference between the infinitely large city and the finite city to be only 5 percent. So, that means that I am looking at a situation in a setting where a linear city with radius capital R captures 95 percent variation relative to an infinitely large monocentric city, ok.

So, what can I do going forward ok? So, I have a function of lambda and R and I am setting it to 0.05. What I can do is I can set lambda to let me set lambda equals 0.1 and then I can take an inverse of this function and I can figure out the value of R which is the radius I can back out the radius of a city. I can back out the radius of a city that will capture 95 percent of variation as far as the entropy is concerned relative to an infinitely large city ok. So, if I can attain this radius given the lambda value I can achieve all the variation I want if I had no constraint of space. So, the constraint of space by itself can be realistically you know given some kind of boundedness, right? So, it may not be a very large constraint.

So, if R 0.95 or R 95 percent turns out to be a realistic value let us say within 100 kilometers, 200 kilometers well, we can realistically build such a city which we think will encompass all you know economic activity living standards as far as you know if they are considered as random variables to be able to be contained within that city, right. So, this will be nothing but l inverse of 0.05 being with lambda set as 0.1 Ok.

And, we can you know we can figure that out with a graph. So, here again, I have Ediff by ER I am going to focus on the case where I fix lambda equals 0.1 Ok.

So, let me fix lambda to be equal to 0.1 and I am going to look at the yellow line what I see here is if you look at Ediff by E sub-R Ediff by ER it is actually falling beyond a given distance from the origin and this value keeps falling and comes to this point of 0.05 which is the value of my interest at a distance of around 33 kilometers. So, R 95 percent or 0.95 turns out to be 33 kilometers.

What this means is that given lambda equals 0.1 the distance from origin equals 33 kilometers is considered an appropriate regional boundary for a monocentric city, right what is the definition of the appropriateness of this regional boundary that it is able to capture 95 percent of the variation, right?

Of course, with lambda values being different if lambda goes up you see that R 0.95 is drastically smaller, right? So, a lot can be done you know with smaller cities as well. So, that it is not all is lost with smaller towns and cities, right, but if lambda is smaller you are going to expect you know you can expect a higher you know sort of a larger distance or larger requirement of a larger you know spatial scope of a city to be able to you know encompass the variation that one would like, right.

So, this is the case, but probably lambda will be probably much less than 0.1, right? So, the idea is that you know schematically if you see I start at the origin my R stops here the R could go till infinity, but the idea is that you know there comes a value of capital R at 33 kilometers where you know ER is 0.95 which is 95 percent of E infinity, ok.

Now, as an exercise you know I am going to sort of ask you to you know to figure out the radius of a monocentric city that captures 99 percent of the variation. So, I am going to make it stricter. So, I am going to say figure out you know R you know 0.99 such that EDIFF by ER equals 0.01 you can set lambda equals 0.1, right? And, you should figure out what our 0.99 is you know in the next 3 minutes ok.

So, I will tell you R after I discuss this little schematic in front of us. So, what we are looking at here are different cities you know a spatial map of different cities including Jakarta, Paris, Moscow, Shanghai, Berlin, New York, and London, and what you see interestingly is that many of these cities if not all have something like a center.

And, around this center the city sort of distributes itself in a way that could be explained as an exponential decline in let us say population density or trade you know some kind of economic activity may be groundwater depletion right you might see a lot of depletion around the center and then that might actually become better as you move away from the city into the heartlands, right.

So, you see a similar shape for Shanghai you know, of course, the axis at which you want to analyze these things can be different, but the model that we are studying here can be useful to understand or characterize the spatial spread of real-world cities. Coming back to the question, well, the value of R 0.99 where if you wanted to capture 99 percent entropy which I had set for you a minute ago well, it is going to be 53 kilometers.

So, interestingly you know while R 0.95 was 33 kilometers, you would have to go out 20 kilometers more which is a lot of space if you think about it on land if you were to go from 95 percent to 99 percent you know a variation of a city. So, that is about it. I hope you enjoyed this exposition of spatial entropy, the use of spatial entropy for regional planning, to understand

characterized cities around us.
Thank you very much.