Prof. dr. sc. Dražen Barković
Ekonomski fakultet u Osijeku

UDK 331.16
Izvorni znanstveni članak

OTKRIVANJE PROSTORNE AUTOKORELACIJE POMOĆU LOKALNIH INDIKATORA

Sažetak
Prostorna autokorelacija mogla bi se definirati kao odnos između vrijednosti pojedine varijable koje se javljaju radi geografskog uređenja područja u kojemu se te varijable pojavljuju. Ona mjeri sličnost objekata unutar jednog područja, stupanj po kojemu je neki prostorni fenomen u korelaciji sam sa sobom u prostoru, razinu nezavisnosti između varijabli, prirodu i jačinu međuovisnosti, tj. prostorna autokorelacija je procjena korelacije neke varijable u odnosu na prostornu lokaciju varijable. Drugim riječima, treba procijeniti jesu li vrijednosti međusobno povezane i ako jesu postoji li prostorni uzorak korelacije, odnosno postoji li prostorna autokorelacija. Sredstva prostorne autokorelacije testiraju jesu li vrijednosti varijable na jednoj lokaciji nezavisne o vrijednosti varijable koja je u susjednoj lokaciji. Prostorna autokorelacija mogla bi se klasificirati kao pozitivna ili negativna. Pozitivna prostorna autokorelacija ima sve slične vrijednosti koje se zajedno pojavljuju, dok negativna prostorna autokorelacija ima različite vrijednosti koje se pojavljuju u susjednim lokacijama.
Prostorna autokorelacija može se mjeriti na lokalnoj i globalnoj razini. U studiji su pokazane obje najvažnije mjere i ilustrirane su na jednome praktičnom primjeru.
Ključne riječi: prostorna autokorelacija, lokalni Moranov koeficijent, globalni Moranov koeficijent, Getis-Ord statistike, „hot spots" i „cold spots"

1. Uvod

Vrlo se često postavljaju pitanja koja se tiču regionalnoga gospodarstva odnosno kakav je međusobni utjecaj regija i njihova jakost, postoji li međuregionalni utjecaj kod pojedinih varijabli.

Takva prostorna autokorelacija javlja se također u regresijskim modelima. Posljedica je toga da su regresijski koeficijenti iskrivljeni ili test signifikantnosti gubi svoju valjanost.[1]

Prostorna autokorelcija može se ispitati globalnim ili lokalnim mjerama. Neke lokalne mjere posjeduju svojstvo da su proporcionalne s prosjekom ili sumom jednoga globalnog pokazatelja. Na engleskom je taj pojam lokalnih mjera poznat pod nazivom LISA (Local Indicators of Spatial Assotiation). Pokazatelji iz modela LISA mjere za svaku pojedinu regiju susjedske odnose. Široka primjena tih mjera opravdava se time što prostorne zavisnosti najčešće nisu u svim regijama jednako jake (homogene), nego se razlikuju lokalno.

Prostorna autokorelacija mogla bi se definirati kao odnos između vrijednosti pojedinih varijabli koji se javlja radi geografskog uređenja područja u kojemu se te varijable pojavljuju. Ona mjeri sličnost objekata unutar jednog područja, stupanj po kojemu je neki prostorni fenomen u korelaciji sam sa sobom u prostoru (Cliff i Ord 1973., 1981.), razinu nezavisnosti između varijabli, prirodu i jačinu međuovisnosti, tj. prostorna autokorelacija je procjena korelacije neke varijable u odnosu na prostornu lokaciju varijable. Drugim riječima, treba procijeniti jesu li vrijednosti međusobno povezane i ako jesu, postoji li prostorni uzorak korelacije, odnosno postoji li prostorna autokorelacija.

Sredstva prostorne autokorelacije testiraju jesu li vrijednosti varijable na jednoj lokaciji nezavisne od vrijednosti varijable koja je u susjednoj lokaciji. Prostorna autokorelacija mogla bi se klasificirati kao pozitivna ili negativna. Pozitivna prostorna autoko-

1 Anselin, I.: Spatial Econometrics, Methods and Models, Dordrecht, Boston, London, 1988., str. 57.

relacija ima sve slične vrijednosti koje se zajedno pojavljuju dok negativna prostorna autokorelacija ima različite vrijednosti koje se pojavljuju u susjednim lokacijama.

Pozitivna prostorna autokorelacija se odnosi na uzorak mape na kojoj geografsko obličje sličnih vrijednosti teži grupiranju na mapi dok negativna prostorna autokrelacija indicira uzorak mape na kojoj su geografske jedinice slične vrijednosti razasute u mapi. Ako ne postoji statistička signifikantnost prostorne autokorelacije, uzorak je prostorne distribucije slučajan. Postoje mnogi indikatori prostorne autokorelacije (Cliff i Ord 1973., 1981.; Goodchild, 1986.; Haining, 1990.; i Chou, 1997.):

1. Globalni indikatori prostorne asocijacije, zajednička statistika; Moranov I (Moran, 1948) i Geary-ev c (Geary, 1954.); nulta hipoteza nealternativna hipoteza; opća statistika unakrsnog produkta; normalni, slučajni i permutacijski pristup; i prostorni korelogram.
2. Lokalni indikatori prostorne asocijacije: LISA,Gi i Gi* statistika; LISA statistika; lokalni Moranov koeficijent.
3. Pristup variograma prostornim asocijacijama geostatističke perspektive; variogram i polu-variogram; variogram vs. korelogram; izglađivanje modela variograma ; robusna procjena variograma; interpretacija variograma; variogram i prostorni uzorak.

U ovom radu istražujemo praktičnu primjenu lokalnih indikatora (LISA).[2] Pokazatelji LISA mjere za svaku pojedinu regiju odnose sa susjedstvom. U međuvremenu je primjena tog pokazatelja pokazala da prostorne zavisnosti nisu najčešće jednake za sve regije nego su lokalno različito izražene. U tom slučaju se govori o prostornoj nestacioniranosti. Često postoje, jedna pored druge, grupe regija s natprosječnim ili ispotprosječnim vrijednostima nekih varijabli (prostorna koncentracija) iste ili suprotne orijentacije (pozitivna ili negativna prostorna autokorelacja). To se može identificirati pomoću izračunavanja u modelu LISA. Na jednome jednostavnom primjeru će se pokazati i objasniti obje najvažnije mjere LISA.

2 Kosfeld, R.,Eckey, F.H., Turck, M.A.: Lisa (Local Indicators of Spatial Association, Wist, br.3, str.157-162.

2. Lokalni Moranov-koeficijent

Lokalni Moranov koeficijent razvio je Anselin[3]. Taj mjerni broj polazi od matrice susjedstva **W.** U najjednostavnijem slučaju **W** pokazuje da li dvije regije „i" i „j" međusobno graniče:

$$w^*_{ij} = \begin{cases} 1, \text{ ako „i" i „j" imaju zajedničku granicu} \\ i \neq j \\ 0, \text{ suprotno} \end{cases} \qquad (1)$$

Tumačenje lokalnog Moranova koeficijenta poboljšat će se ako se matricu susjedstva $\mathbf{W}^*$ tako normira da suma reda bude jednaka jedan. Zato se dijeli svaki element sa sumom reda:

$$w_{ij} = \frac{w^*_{ij}}{\sum_{j=1}^{n} w^*_{ij}} \qquad (2)$$

Time se dobiva standardizirana matrica vaganja **W.** Slično kao kod izračunavanja koeficijenata korelacije po Bravais i Pearsonu za izračunavanje lokalnih Moranovih koeficijenata uzimaju se unakrsni produkti. Ovdje se ne primjenjuju odstupanja između dvije različite varijable X i Y od njihove srednje vrijednosti,

$$(x_i-\bar{x}) \cdot (y_i-\bar{y}) \qquad (3)$$

nego pojedinih varijabli. Umjesto druge varijable Y stavlja se aritmetička sredina graničnih područja:

$$\sum_{j=1}^{n} w_{ij} \cdot (x_j-\bar{x}) \qquad (4)$$

Mjerit će se produkt odstupanja od srednje vrijednosti u i-toj regiji sa srednjom vrijednosti svih odstupanja u susjednim regijama:

$$(x_i-\bar{x}) \sum_{j=1}^{n} w_{ij} \cdot (x_j-\bar{x}) \qquad (5)$$

3 Anselin, I.: Local indicators of Spatial Association-LISA u; Geographical Analysis, Vol. 27 1995, str. 98.

Kada objašnjavamo taj unakrsni produkt iz formule (5) koji određuje predznak, a time i interpretaciju lokalnh Moranovih koeficijenata, ističemo:

a) Unakrsni koeficijent u (5) uzima visoke vrijednosti kada se u i-toj regiji kao i u graničnim regijama pojavljuju iznatprosječne ili u obrnutom slučaju ispotprosječne x-vrijednosti.
b) U skladu s tim je unakrsni produkt negativan kada x-vrijednost u i-toj regiji kao i njegova aritmetička srednja vrijednost odstupaju od srednje vrijednosti u susjednim regijama po suprotnom predznaku.
c) Unakrsni produkt se javlja u blizini nule kada x-vrijednost u i-toj regiji ili prosjek x-vrijednosti u graničnim regijama odgovaraju približno ukupnoj sredini.

Taj unakrsni produkt mora se još normirati. To se postiže, ako se uzme u obzir prosječna suma kvadrata odstupanja u odnosu na X. Formula po kojoj se izračunavaju lokalni Moranovi koeficijenti i-te regije glasi[4]

$$I_i = \frac{(x_i-\bar{x}) \sum_{j=1}^{n} w_{ij} \cdot (x_j-\bar{x})}{\sum_{j=1}^{n} (x_j-\bar{x})^2 / n} \qquad (6)$$

U mnogim regijama kod kojih manjka autokorelacija očekivat će se nula za vrijednost od I_i. Kod manjeg broja regija može očekivana vrijednost koja je zadana standardiziranom vaganom matricom

$$E(I_i) = \frac{\sum_{j=1}^{n} w_{ij}}{n-1} = \frac{1}{n-1} \qquad (7)$$

značajno odstupati od nule u negativnom području.

Emprijske studije pokazuju da „spillover" efekti jako opadaju s prostornom udaljenosti [5]. Time se obrazlaže česta primjena binarne matrice susjedstva regija koje izravno graniče (matrica susjedstva prvog reda). U empirijskim istraživanjima rijetko se poseže za matricom susjedstva drugog reda koja mjeri efekte između regija koje razdvaja neka treća regija.

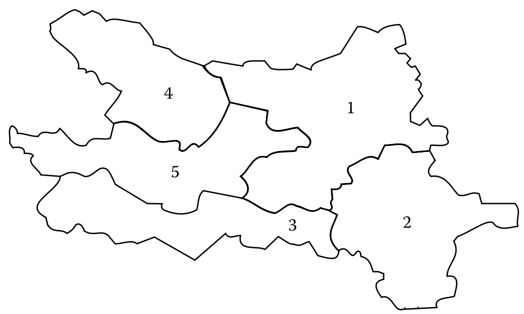
Kako se očekivane vrijednosti lokalnih i globalnih Moranovih koeficijenata poklapaju može se globalni Moranov koeficijent prikazati kao aritmetička sredina lokalnih pokazatelja:

$$I = \frac{1}{n} \sum_{j=1}^{n} I_i \qquad (8)$$

Globalni Moranov-koeficijent I,

$$I = \frac{\sum_{j=1}^{n} (x_i-\bar{x}) \cdot \sum_{j=1}^{n} w_{ij} \cdot (x_j-\bar{x})}{\sum_{j=1}^{n} (x_j-\bar{x})^2} \qquad (9)$$

leži kod primjene standardizirane vagane matrice skoro uvijek u intervalu između minus i plus jedan. On omogućava spoznaju, je li lokalna autokorelacija izražena srednjom vrijednošću, natprosječnom ili ispotprosječnom. Točnija su istraživanja koja se vrše preko lokalnih pokazatelja. Kada se lokalni Moranovi koeficijenti unesu na zemljopisnu kartu, mogu se identificirati lokalne nestacionarnosti. Ovdje je riječ o povezanim regijama koje u usporedbi s globalnim Moranovim koeficijentima pokazuju jako stvaranje klastera ili pak jednu suprotstavljenu prostornu autokorelaciju (onih koji strše izvan serije podataka).



4 Fotheringham, A.S.,Brunsdon, C., Charlton, M.E., Quantitative Geographie. Perspectives on Spatial Data Analysis, London New Delhi, 2000, str.102.

5Audretsch, D.B.,Feldman, M.P.:Knowlidge Spillovers and the Geography of Innovation u Handbook of Regional and Urban Economics, vol. 4 Amsterdam 2004., str. 2713-2742. .

Slika 1: Grafički prikaz položaja regija

Pojasnimo izračunavanje lokalnih Moranovih koeficijenata na jednom primjeru. Polazimo od slike 1.

Za pet regija mora se najprije uspostaviti matrica susjedstva **W***. Regija 1 graniči s regijom 2, zato je za elemente (1,2) i (2,1) unesena jedinica. U skladu s tim na sličan način su povezane i preostale regije. Na glavnoj dijagonali stavljene su nule. Matrica susjedstva za pet regija ima ovaj izgled:

$$W^* = \begin{array}{c@{}c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} & \\ \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix} & \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \end{array}$$

Osim toga moraju se izračunati sume redova. Suma prvog reda i suma drugog reda iznosi na primjer

$$\sum_{j=1}^{5} w^*_{ij} = 0+1+1+1+1=4$$

$$\sum_{j=1}^{5} w^*_{ij} = 1+0+1+0+0=2$$

Svaki element prvog reda mora se podijeliti s četiri, drugog s dva. Provedu li se ta računanja za sve redove, dobije se standardizirana vagana matric

$$W^* = \begin{array}{c@{}c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} & \\ \begin{bmatrix} 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1/3 & 1/3 & 0 & 0 & 1/3 \\ 1/2 & 0 & 0 & 0 & 1/2 \\ 1/3 & 0 & 1/3 & 1/3 & 0 \end{bmatrix} & \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \end{array}$$

Pored standardizirane vagane matrice W potrebne su nam x-vrijednosti za regije. Zanima nas pitanje koliko cijena potrošačke košare u jednoj regiji zavisi o odgovarajućim vrijednostima u susjednim (graničnim) područjima. Cijena u regiji 1 iznosi 5680 kuna, u regiji 2 i 3 iznosi 5042 i 5620 kuna itd. Cijene za svih 5 regija dane su vektorom x:

$$x = \begin{bmatrix} 5680 & 5042 & 5620 & 5843 & 5300 \end{bmatrix}'$$

Aritmetička sredina cijena u pet regija je

$$\bar{x} = \frac{1}{5} \cdot \sum_{i=1}^{5} x_i$$

$$= \frac{1}{5} \cdot (5680 + 5042 + 5620 + 5843 + 5300)$$

$$= \frac{1}{5} \cdot 27485 = 5497$$

Time se može odrediti suma kvadrata odstupanja od aritmetičke sredine za nazivnik u formuli (6):

$$\sum (x_j - \bar{x})^2 =$$

$$(5680 - 5497)^2 + (5042 - 5497)^2 + (5620 - 5497)^2 +$$

$$(5843 - 5497)^2 + (5300 - 5497)^2 = 414168$$

Kako tada glasi lokalni Moranov koeficijent za prvu regiju? Aritmetička je sredina odstupanja susjednih regija dana s

$$w_{1J} (x_j - \bar{x})^2 =$$

$$0 \cdot (5680 - 5497) + 1/4 \cdot (5042 - 5497)$$

$$- 1/4 \cdot (5620 - 5497) + 1/4 \cdot (5843 - 5497)$$

$$+ 1/4 \cdot (5300 - 5497) = -46$$

Regije koje su u susjedstvu, obilježene su u prvom redu matrice susjedstva s jedinicom, ta se područja uzimaju u obzir. Tako se dobiva sljedeći lokalni Moranov koeficijent

$$I_1 = \frac{(x_i - \bar{x}) \sum_{j=1}^{5} w_{1j} \cdot (x_j - \bar{x})}{\sum_{j=1}^{5} (x_j - \bar{x})^2 / 5}$$

$$= \frac{(5680 - 5497) \cdot (-46)}{414168 / 5} = -0.102$$

Regija 1 u usporedbi sa susjednim regijama ne pokazuje otklon u istom smjeru. Točnije, cijena košare od 5680 kuna u regiji 1 kao i 5620 kuna u susjednoj regiji 3 i 5843 u regiji 4 iznad su prosjeka, dok je u graničnoj regiji 2 cijena košare od 5042 kuna ispod prosjeka kao i u regiji 5. Imamo dakle jedan negativan doprinos globalnoj autokorelaciji koja se preslikava preko globalnih Moranovih koeficijenata.

Koja bi se vrijednost mogla očekivati u primjeru za prve lokalne Moranove koeficijente kada nedostaje autokorelacija? Tu treba primjeniti formulu (7):

$$E(I_1) = \frac{\sum_{j=1}^{5} w_{1j}}{5-1} = \frac{0 + 0.25 + 0.25 + 0.25 + 0.25}{5-1}$$

$$= -\frac{1}{4} = -1.025$$

Lokalni Moranov koeficijent može se odrediti i za drugu regiju po istom postupku kao i za prvu regiju. Ovaj put dobivamo vrijednost

U skladu s tim određuju se ostali lokalni Mora-

$$I_2 = \frac{(x_2 - \bar{x}) \sum_{j=1}^{5} w_{2j} \cdot (x_j - \bar{x})}{\sum_{j=1}^{5} (x_j - \bar{x})^2 / 5}$$

$$= \frac{(5042 - 5497) \cdot 153}{82833.6} = -0.840$$

novi koeficijenti:

$$I_3 = -0.233,\ I_4 = -0.029\ i\ I_5 = -0.5167$$

Globalni Moranov koeficijent dobiva se kao prosjek lokalnih mjera:

$$I = \frac{1}{5} \cdot \sum_{i=1}^{5} I_i$$

$$= \frac{1}{5} \cdot (-0.01 - 0.84 - 0.23 - 0.30 - 0.52)$$

$$= -0.32$$

Riječ je o negativnoj autokorelaciji. Regije s natprosječnom cijenom košarice proizvoda područno graniče s regijama u kojima u cijene ispod prosjeka i vice versa. Lokalni Moranovi koeficijenti pokazuju da regije 2 i 5 daju najveći doprinos negativnoj vrijednosti Moranovog koeficijenta.

Prostorno klasteriranje se identificira oko regija 1, 4 i nešto manje kod regije 3 jer njihovi lokalni Moranovi koeficijenti premašuju globalni Moran koeficijent i očekivane vrijednosti.

$I_1 = -0.102 > E(I_1) = -0.25$
→ pozitivna lokalna prostorna autokorelacija

$I_1 = -0.102 > I = -0.32$
→ prostorno klasteriranje (visoke ili niske vrijednosti klasteriranja oko regije 1.

$I_3 = -0.233 > E(I_3) = -0.25$
→ pozitivna lokalna prostorna autokorelacija

$I_3 = -0.233 > I = -0.32$
→ prostorno klasteriranje (visoke ili niske vrijednosti klasteriranja oko regije 3.

$I_4 = -0.002 > E(I_4) = -0.25$
→ pozitivna lokalna prostorna autokorelacija

$I_4 = -0.002 > I = -0.32$
→ prostorno klasteriranje (visoke ili niske vrijednosti klasteriranja oko regije 4.

Regije 2 i 5 se analiziraju ovako:

$I_2 = -0.840 < E(I_1) = -0.25$
→ negativna prostorna autokorelacija

$I_2 = -0.840 < I = -0.32$
→ nema prostornog klasteriranja oko regije 2

$I_5 = -0.517 < E(I_1) = -0.25$
→ negativna prostorna autokorelacija

$I_5 = -0.517 < I = -0.32$
→ nema prostornog klasteriranja oko regije 5

3. Getis – Ord G – statistike

Je li riječ o lokalnom klasteru „hot spot“ (koncentracija viših x-vrijednosti), ili o „cold spots“ (koncentracija nižih x-vrijednosti) ne može se vidjeti na temelju lokalnih Moranovih koeficijenata jer se u oba slučaja javlja regionalno ograničena autokorelacija. O tom problemu su Getis i Ord [6] dali izvještaj u svojim Getis-Ord statistikama. Ti su lokalni pokazatelji primjenjivi kod pozitivnih x-vrijednosti koje imaju prirodnu nul-točku (barem skalu odnosa ili omjera). Oni polaze od udaljenosti između regija koja se mjeri obično udaljenostima ili vremenom vožnje između regionalnih centara. Ostajemo kod pet regija (slika 1). Udaljenost između centara regije 1 i regije 2 je na primjer 6 km. U skladu s tim daju se i preostale udaljenosti koje su zabilježene u tablici 1.

Udaljenosti regionalnih središta (Osijek, Vukovar, Slavonski Brod, Virovitica, Požega) unosimo u matricu udaljenosti D. Na glavnoj dijagonali D stoje nule jer je udaljenost između središta jedne te iste regije nula:

$$D = \begin{bmatrix} 0 & 36 & 93 & 126 & 96 \\ 36 & 0 & 100 & 173 & 132 \\ 93 & 100 & 0 & 165 & 42 \\ 126 & 173 & 165 & 0 & 105 \\ 96 & 132 & 42 & 105 & 0 \end{bmatrix}$$

Potrebna je kritična vrijednost udaljenosti d da bi se matrica udaljenosti transformirala u binarnu matricu (matrica s nulama i jedinicama). Biramo aritmetičku sredinu svih udaljenosti različitih od nule iz matrice udaljenosti

6 Getis, A., Ord. J.K.: The Analysis of Spatial Associatiion by Use of Distance Statistics u Geographical Analysis, Vol. 24, 1992., str. 189 – 206.

Prof. dr. sc. Dražen Barković: Otkrivanje prostorne autokorelacije pomoću lokalnih indikatora

Tablica 1: Udaljenosti izeđu regionalnih centara

Regija	1	2	3	4	5
1	0	36	93	126	96
2	36	0	100	173	132
3	93	100	0	165	42
4	126	173	165	0	105
5	96	132	42	105	0

1. Osječko-baranjska županija
2. Vukovarsko-srijemska županija
3. Brodsko-posavska županija
4. Virovitičko-podravska županija
5. Požeško-slavonska županija

koja leži kod

Sada se određuje binarna matrica tako da se

$$d = \frac{1}{10} \cdot (36 + 93 + 126 + 96 + 100 + 173$$
$$+ 132 + 165 + 42 + 105) = 106.8$$

za udaljenosti koje podbacuju kritične vrijednosti stavljaju jedinice. Svi preostali elementi bilježe se nulama:

$$w_{ij} = \begin{cases} 1, \text{ ako } d_{ij} < d \\ 0, \text{ suprotno} \end{cases}$$

U primjeru iznosi udaljenost između prve dvije regije 36 < 106.8, zbog čega su elementi (1,2) i (2,1) u binarnoj matrici jedinica. Potpuna binarna matrica ima sljedeći izgled:

U binarnoj matrici stoje jedinice, kada dvije regije leže relativno blizu jedna pored druge, dakle kada

$$W(d = 116.8) = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

njihova udaljenost podbacuje kritičnu udaljenost d. Kod izračunavanja jednostavnih Getis-Ordovih statistika G_i, suma x-vrijednosti graničnih regija u odnosu je na sumu x-vrijednosti svih drugih regija:

$$G_i = \frac{\sum_{j \neq i} w_{ij}(d) \cdot x_j}{\sum_{j \neq i} x_j}$$

Indeks koji se pojavljuje kod sumiranja je i ≠ j. Time se naznačava da se sumira preko svih regija izuzimajući i-tu regiju. G_i daje udjel suma obilježja svih preostalih regija u odnosu na regije koje nisu daleko udaljene. Očekivana vrijednost statistike G_i iznosi

$$E(G_i) = \frac{\sum_{j \neq i} w_{ij}(d)}{n-1}$$

Premaši li G_i znatno očekivanu vrijednost tada se radi o pozitivnoj autokorelaciji s regionalnom koncentracijom visokih x-vrijednosti („hot spots“). Slično tome je dana pozitivna autokorelacija kod lokalne koncentracije nižih x-vrijednosti („cold spots“), kada G_i jako podbacuje svoju očekivanu vrijednost.

Pored toga, postoji još jedna varijacija te formule kada se vlastita regija uvede u brojnik i nazivnk[7]

$$G^*_i = \frac{\sum_{j=1}^{n} w_{ij}(d) \cdot x_j}{\sum_{j=1}^{n} x_j}$$

G^*_i se može tumačiti kao udio sume obilježja svih regija koji se odnosi na vlastitu regiju i područja koja nisu od nje jako udaljena. Očekivana vrijednost G^*_i

$$E(G^*_i) = \frac{\sum_{j=1}^{n} w_{ij}(d)}{n}$$

I kod Getis-Ord statistike postoji odgovarajući globalni pokazatelj

7Haining, R.: Spatial Data Analysis, Theory and Practice, Cambridge. 2004, str. 253.

$$G = \frac{\sum_{j=1}^{n} \sum_{j \neq i} w_{ij}(d) \cdot x_i \cdot x_j}{\sum_{j=1}^{n} \sum_{j \neq i} x_i \cdot x_j}$$

koji uzima vrijednosni interval između nula i jedan i ima očekivanu vrijednost

$$G = \frac{W}{n \cdot (n-1)} \quad \text{gdje je}$$

$$W = \sum_{j=1}^{n} \sum_{j \neq i} w_{ij}(d)$$

Suprotno od lokalnih Moranovih koeficijenata, G se ne može izraziti preko G_i. Umjesto toga mogu se opet uzeti u obzir globalni koeficijenti i pomoću lokalnih mjera G_i na karti identificirati „hot spots“ i „cold spots“.

Izračunavanje ćemo ponovno ilustrirati na primjeru cijena u pet regija. Počnimo s G_i statistikom prema formuli (11). x-vrijednosti (cijene) glase:

$$x = \begin{bmatrix} 5680 & 5042 & 5620 & 5843 & 5300 \end{bmatrix}'$$

Izračunajmo za prvu regiju brojčane veličine prema (11). Množe se elementi prvog reda iz **W**(d=106.8) s x-vrijednostima. Sumira se dakle preko regija 2 do 5. Prvu regiju za koju se treba izračunati Getis-Ord statistika G_i ne uzimamo u obzir:

$$\sum_{j \neq i} w_{1j}(d = 106.8)\, x_j = 1 \cdot 5042$$
$$+ 1 \cdot 5620 + 0 \cdot 5843 + 1 \cdot 5300 = 15962$$

Od regije 1 nisu daleko udaljene regija 2 , regija 3 i regija 5. Zbog toga pokazuje binarna matrica W(d=106.8) za te elemente jedinice i samo cijene te dvije regije se uvrštavaju u brojnik.

U nazivniku stoji suma svih x-vrijednosti, pri čemu se prva regija ponovno ne uzima u obzir:

Getis-Ord statistika iznosi

Približno 73 % agregiranih cijena košarica svih preostalih regija odnose se na regije koje nisu daleko

$$\sum_{j \neq i} w_j = 5042 + 5620 + 5843 + 5300 = 21805$$

udaljene od regije 1. Vrijednost 0.732 ne premašuje

$$G_1 = \frac{\sum_{j \neq 1} w_{1j}(d = 106.8) \cdot x_j}{\sum_{j \neq 1} x_j} = \frac{15962}{21805} = 0.732$$

vrijednost koja bi se očekivala za prvu regiju kada manjka lokalna autokorelacija:

Oko regije 1 leži snop regija (=klaster) s relativino niskim cijenama proizvoda („cold spots“). Za drugu regiju dobivamo na primjer Getis-Ord

$$E(G_1) = \frac{\sum_{j \neq 1} w_{ij}(d = 106.8)}{5-1} = \frac{1 + 1 + 0 + 1}{4} = 0,75$$

statistiku G_i

Slično tome može se odrediti druga varijanta G_i^*. Ovdje se uvrštava i vlastita regija:

$$G_2 = \frac{\sum_{j \neq 2} w_{2j}(d = 106.8) \cdot x_j}{\sum_{j \neq 2} x_j} = \frac{11300}{22443} = 0,50$$

Usporedba s očekivanom vrijednosti

Ponovno se pokazuje stvaranje klastera s niskim

$$G^*_1 = \frac{\sum_{j=1}^{5} w_{ij}(d = 106.8) \cdot x_j}{\sum_{j=1}^{5} x_j}$$

$$= \frac{1 \cdot 5680 + 1 \cdot 5042 + 1 \cdot 5620 + 0 \cdot 5843 + 1 \cdot 5300}{5680 + 5042 + 5620 + 5843 + 5300}$$

$$= \frac{21642}{27485} = 0,787$$

atributskim vezama (x-vrijednosti) oko regije 1.

$$E(G^*_1) = \frac{\sum_{j=1}^{5} w_{ij}(d = 106.8)}{n}$$

$$= \frac{1 + 1 + 1 + 0 + 1}{5} = \frac{4}{5} = 0,8$$

Kako proizlazi iz tablice 3, nailazimo na slična poklapanja lokalnih Getis-Ord pokazatelja.

Globalna Getis-Ord statistika

$$G = \frac{\sum_{i=1}^{5} \sum_{j \neq i} w_{ij}(d = 106.8) \cdot x_i \cdot x_j}{\sum_{i=1}^{5} \sum_{j \neq i} x_i \cdot x_j}$$

Tablica 2 Getis-Ord statistike

Regija	1	2	3	4	5
Gi	0.73	0.50	0.73	0.24	0.77
E(Gi)	0.75	0.50	0.75	0.25	0.75
Gi*	0.78	0.59	0.78	0.40	0.80
E(Gi*)	0.80	0.60	0.80	0.40	0.80

$$= \frac{1 \cdot 5680 \cdot 5042 + 1 \cdot 5680 \cdot 5620 + 1 \cdot 5680 \cdot 5300 + 1 \cdot 5042 \cdot 5680 + 1 \cdot 5042 \cdot 5620 + 1 \cdot 5620 \cdot 5680 + 1 \cdot 5620 \cdot 5042 + 1 \cdot 5620 \cdot 5300 \ldots + 1 \cdot 5300 \cdot 5843}{5680 \cdot 5042 + 5680 \cdot 5620 + 5680 \cdot 5843 + 5680 \cdot 5300 + 5042 \cdot 5680 + 5042 \cdot 5620 + 5042 \cdot 5843 + 5042 \cdot 5300 + \ldots + 5300 \cdot 5843}$$

$$= \frac{348}{603926012} = 0,595$$

globalno daje evidenciju prostorne koncentracije visokih x-vrijednosti koje jasno ne premašuju svoju očekivanu vrijednost oko prve tri regije.

$$E(G) = \frac{\sum_{i=1}^{5} \sum_{j \neq i} w_{ij}(d = 106.8)}{5 \cdot (5 - 1)} = \frac{12}{20} = 0,6$$

Lokalne mjere iz tablice 2 pokazuju da se klaster („hot spots) koncentrira oko regije 5.

Zaključak

Globalna se autokorelacija često javlja kod regionalnih ekonomskih istraživanja. Ona ne mora u svim regijama biti jednaka, nego može biti lokalno suprotno usmjerena. Osim toga kod suprotnih prostornih autokorelacija mogu se lokalni klasteri viših vrijednosti („hot spots“) i nižih vrijednosti („cold spots“) atributivnih varijabli smjestiti jedni pored drugih. U daljnjim istraživanjima mogu se za te klastere procjenjivati posebni modeli.

U ovom su radu predstavljene obje najupotrebljivije LISA mjere – lokalni Moranov koeficijent i lokalne Getis-Ord statistike. Getis-Ord statistike transfomiraju matricu udaljenosti u binarnu matricu čime je skopčan značajan informacijski gubitak. Ako se pojave negativne vrijednosti lokalnih Moranovih koeficijenata kod pozitivnih globalnih autokorelacija, pojavljuju se prostorne nestacionarnosti u obliku onih koji strše izvan serije podataka. Izvanredno visoki pozitivni lokalni Moranovi koeficijenti upućuju na regionalni klaster visokih („hot spots“) ili niskih („cold spots“) vrijednosti atributivne varijable. O kojem je od obje vrste klastera riječ, ne može se raspoznati na osnovu Moranovih koeficijenata. Tu se moraju uključiti Getis-Ord statistike na osnovi kojih se mogu identificirati „hot spots“ i „cold spots. „hot spot “ se pokazuje na temelju jako natprosječnih G_i ili G_i^* vrijednosti dok se „cold spot“ prepoznaje na temelju jasne ispotprosječne vrijednosti G_i ili G_i^*.

Za lokalne pokazatelje postoji također test signifikantnosti koji je integriran u postojeći software za GIS (Geoinformacijski sustav). Izračunavanjne tih veličina je djelomično dosta kompleksno. Posebno se mora voditi računa o tome da se više signifikacijskih testova mora provesti istovremeno.

Prema statistici G_i:

Tendencija prostorne koncentracije viših vrijednosti: regija 5 i okolne regije.

Tendencija prostorne koncentracije niskih vrijednosti: regija 1, 3, 4 i okolne regije.

Nema prostorne koncentracije: u okolini regije 2.

Prema statistici G_i^*:

Tendencija prostorne koncentracije niskih vrijednosti: regija 1, 2, 3 i okolne regije.

Nema prostorne koncentracije: regije 4, 5 i okolne regije.

Literatura

1. Anselin, I.: Spatial Econometrics, Methods and Models, Dordrecht, Boston, London, 1988.
2. Anselin, I.: Local Indicators of Spatial Association-LISA u; Geographical Analysis, Vol.27 1995.
3. Audretsch, D.B.,Feldman, M.P.:Knowlidge Spillovers and the Geography of Innovation u Handbook of Regional and Urban Economics, vol. 4 Amsterdam 2004., str. 2713-2742.
4. Cliff,A.D, Ord,J.K.: Spatial Autocorrelation, Pion, London, 1973.
5. Cliff,A.D, Ord,J.K.:Spatial Processes. Models and Applications, London, Pion, 1981.
6. Fotheringham, A.S.,Brunsdon, C., Charlton, M.E., Quantitative Geographie. Perspectives on Spatial Data Analysis, London, New Delhi, 2000.
7. Getis, A., Ord. J.K.: The Analysis of Spatial Associatiion by Use of Distance Statistics u Geographical Analysis, Vol. 24, 1992.
8. Goodchild,M.F.: Spatial Autocorrelation CATTIOG 47, Geo Books Norwich, 1986.
9. Goodchild,M.F., Haining, R., Wise, B.: Integrating gis and spatial dana analysis: probklems and possibilities. International Journal of Geographical Information Systems, 6, str. 407-423.
10. Haining, R.: Spatial Data Analysis, Theory and Practice, Cambridge, 2004.
11. Haining, R.P. : Spatial data analysis in the social and environmental sciences, Cambridge University Press, Cambridge 1990.
12. Moran, P.A.P.:The interpretation of statistical maps, Journal of the Royal Statistical Society, B 10, 1948.
13. Thompson, P.W., Huggert, R.J.: Modelling in Geography:A Mathematical Approach, Rowman,Littlefield.

Prof. dr. sc. Dražen Barković
Faculty of Economics, Osijek

REVEALING A SPATIAL AUTOCORRELATION WITH LOCAL INDICATORS

Summary

Spatial autocorrelation may be defined as the relationship among values of a single variable that comes from the geographic arrangement of the areas in which these values occur. It measures the similarity of objects within an area, the degree to which a spatial phenomenon is correlated to itself in space, the level of interdependence between the variables, the nature and strength of the interdependence, i.e. spatial autocorrelation is an assessment of the correlation of a variable in reference to spatial location of the variable. Assess if the values are interrelated, and if so is there a spatial pattern to the correlation, i.e. is there spatial autocorrelation.
Spatial autocorrelation tools test whether the observed value of a variable at one locality is independent of values of the variable at neighboring localities. Spatial autocorrelation may be classified as either positive or negative. Positive spatial autocorrelation has all similar values appearing together, while negative spatial autocorrelation has dissimilar values appearing in close association. map. When no statistically significant spatial autocorrelation exists, the pattern of spatial distribution is considered random.
Spatial autocorrelation can be measured on local and global level. This study presents both of these measures and ilustrates them on a practical example.
KEY WORDS: spatial autocorrelation, local Moran coefficient, global Moran coefficient, Getis-Ord statistics, „hot spots“ i „cold spots“.