

GII-07 3 dalis: Erdvinė statistika

2 užduotis: Vietinė erdvinio išsidėstymo analizė

Atlikimo terminas: šiai užduočiai atlikti reikia maždaug savaitės.

Praktinio darbo vertinimas: šis laboratorinis darbas vertinamas iki 37 balų, jo vertė sudaro 6 proc. viso kurso bendro pažymio.

Aprašymas ir uždaviniai

Šiame darbe pagrindinis dėmesys skiriamas erdvinės statistikos metodų taikymui geografiniams duomenims vertinant nusikalstamumą Lietuvoje. Studentai išmėgins šiuos erdvinės statistikos metodus:

- 1) vietinių autokoreliacijos erdvėje rodiklių skaičiavimą ir interpretavimą;
- 2) vietinio erdvinio išsidėstymo analizę;
- 3) vietinių rezultatų palyginimą su ankstesniais globalios erdvinės statistikos rezultatais.

Atsiskaitymo reikalavimai

Surašykite atsakymus į visus klausimus atskirame dokumente (*MS Word*) ir nepamirškite įtraukti žemėlapių ar lentelių, kurias parengėte atsakydami į užduoties klausimus. Visos iliustracijos turi būti įdėtos į *MS Word* dokumentą, o visos lentelės turi būti parengtos *MS Word* lentelių *Table* funkcijomis. Visi skaičiai turi būti surašyti trijų skaitmenų po kablelio tikslumu.

Pasiruošimas

Darbui reikia šio duomenų failo: Lithuania_2005.SHP.

Šiame SHAPE faile yra tokie kintamieji, kurie gali būti panaudoti atliekant 2005 metų nusikalstamumo duomenis nagrinėjančią užduotį:

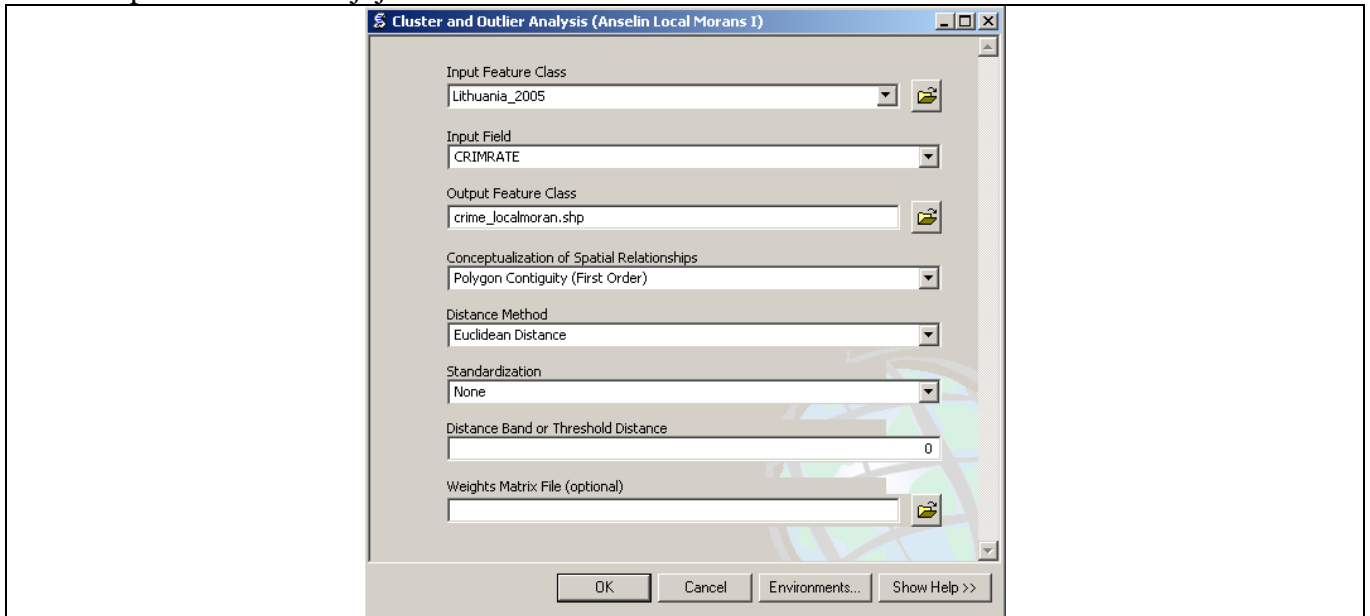
UNIQUE_ID	Unikalus kiekvienos savivaldybės identifikacinis numeris
COUNTY	Unikalus kiekvienos apskrities identifikatorius
VARDAS	Savivaldybės pavadinimas
APSKRITIS	Apskrities pavadinimas
POPULATION	Savivaldybės gyventojų skaičius
CRIMES	Nusikaltimų skaičius savivaldybėje
THEFT	Vagysčių skaičius savivaldybėje
BNE	Įsilaužimų į butus skaičius
AUTO	Pavogtų automobilių skaičius
JUVDEL	Nepilnamečių padarytų nusikaltimų skaičius
CRIMRATE	Nusikalstamumo lygis savivaldybėje 10 000 gyventojų
THFTRATE	Vagysčių lygis savivaldybėje, skaičiuojant 10 000 gyventojų
BNERATE	Įsilaužimų į butus lygis savivaldybėje, skaičiuojant 10 000 gyventojų
AUTORATE	Automobilių vagysčių lygis savivaldybėje, skaičiuojant 10 000 gyventojų
JDELRATE	Nepilnamečių nusikalstamumo lygis savivaldybėje, skaičiuojant 10 000 gyventojų

Visi duomenys šiems skaičiavimams paimti iš Lietuvos statistikos departamento:
<http://www.stat.gov.lt/en/>.

A dalis. Vietinė autokoreliacija erdvėje

Atlikdami šį pratimą pasinaudosime Lietuvos savivaldybių ribomis, kad išnagrinėtume vietinio nusikalstamos veiklos išsidėstymo erdvėje (jei toks yra) ypatumus. Pagal keletą kintamųjų apskaičiuosime vietinę autokoreliaciją erdvėje.

1. Paleiskite *ArcMap* ir pridėkite failą *Lithuania_2005.SHP*.
2. Paleiskite *ArcToolbox*.
3. Spragtelėkite *Spatial Statistics Tools* → *Mapping Clusters* → *Cluster and Outlier Analysis (Anselin Local Morans I)* (Erdvinės statistikos įrankiai → Klasterių išdėstymas → Klasterių ir riktų analizė (Anselino vietinė Morano I)).
4. Klasterių ir riktų analizės (*Cluster and Outlier Analysis*) lange įveskite:
5. Įvesties elementų klasė (*Input Feature Class*) yra *Lithuania_2005.SHP*; įvesties laukas – *CRIMRATE*; išvesties elementų klasė (*Output Feature Class*) – *crime_localmoran*; o erdvinių sąryšių konceptualizavimas (*Conceptualization of Spatial Relationships*) – pirmos eilės poligonų gretimumas (*Polygon Contiguity (First Order)*); palikite likusių laukų numatytojas reikšmes; spragtelėkite *OK*.
6. Tai suformuos dar vieną *SHAPE* failą, turintį du kintamuosius: *LMiContig* ir *LMzContig*, kaip aptarta dalies teorijoje.



Pakeiskite šio naujo *SHAPE* failo simboliką, kad *LMiContig* būtų parodytas kaip kiekybinė spalvų paletė, nuo žalios iki raudonos, ir patvirtinkite numatytuosius nustatymus (5 natūralios pakopos).

Iš ankstesnio darbo prisiminkite, kad taikant globalią Morano I erdvinę statistiką nė vienas iš nusikalstamumo santykių neparodė autokoreliacijos erdvėje.

1 klausimas: ką tik žemėlapyje pavaizdavote vietinius autokoreliacijos erdvėje įverčius. Aprašykite vietinės Morano I pasiskirstymo žemėlapyje ypatumus CRIMRATE kintamajam. (5 balai)

Pakartoję šiuos veiksmus sukurkite naujus *THFTRATE* ir *JDELRATE* kintamųjų *SHAPE* failus.

2 klausimas: sukurkite ir darbo ataskaitoje pateikite THFTRATE ir JDELRATE žemėlapius. Aprašykite vietinės autokoreliacijos erdvėje išsidėstymo ypatumus kiekviename šių žemėlapių. (8 balai)

3 klausimas: pažvelkite į *LMiContig* kintamojo reikšmes. Ar šis vietinis autokoreliacijos erdvėje matas jums pasako daugiau nei praeitoje užduotyje nagrinėtas globalusis autokoreliacijos erdvėje rodiklis? Paaiškinkite. (4 balai)

B dalis. Vietinė klasterių analizė

Vietinės autokoreliacijos erdvėje rodiklių vaizdavimas žemėlapyje yra įtikinamas, ir parodo vietinės analizės naudą. Tačiau, kaip ir globalaus rodiklio atveju, reikia išnagrinėti statistinio patikimumo reikšmę. Šioje užduoties dalyje vietiniai autokoreliacijos erdvėje matai ir atitinkamos *Z* reikšmės bus panaudotos vietiniams klasteriams nustatyti.

1. Atidarykite CRIMRATE vietinių Morano I rezultatų atributų lentelę.
2. Dešiniuoju pelės klavišu spragtelėkite LMzContig ir pasirinkite *Sort Ascending* (rikiuoti didėjančiai).
3. Dabar pasirinkite visus stebėjimus, kurių LMzContig (*Z* reikšmė) absoliučia reikšme didesnė už 1,28: $LMzContig < -1.28$ or $LMzContig > 1.28$
4. *ArcMap* turinyje dešiniu klavišu spragtelėkite šio SHAPE failo pavadinimą.
5. Spragtelėkite *Data* → *Export Data...* (Duomenys → Eksportuoti duomenis).
6. Atsidarys *Export Data...* (Duomenų eksporto) langas.
7. Eksportuokite tik pasirinktus elementus ir naudokite tą pačią koordinačių sistemą, kaip šio sluoksnio pradiniai duomenys.
8. Išsaugokite išvesties SHAPE failą kaip *crime_localmoran_significant.shp* ar panašiu pavadinimu.
9. Pridėkite šį sluoksnį į žemėlapi.
10. Tai sluoksnis, kuriame rodomi vietiniai autokoreliacijos erdvėje klasteriai.
11. Ištrinkite pasirinktus elementus iš CRIMRATE vietinio Morano sluoksnio.

Šiame naujame sluoksnyje matomos tik savivaldybės, kurių vietinės Morano I reikšmės statistiškai reikšmingos. Paprastas metodas parodyti visą Lietuvą taip, kad ji būtų susijusi su jūsų žemėlapiu, yra atlikti tai:

1. Atidarykite Lithuania_2005.SHP *Symbolology* (simbolikos) langą.
2. Pasirinkite *Features* → *Single Symbol* (Elementai → Vienas simbolis).
3. Spragtelėkite simbolį ir vietoje spalvos pasirinkite *Hollow* (tuščias).
4. Padarykite šį sluoksnį aktyvų, bet palikite jį po sluoksniu, rodančiu statistiškai reikšmingas vietines Morano I statistikas.
5. Dabar rodomos tik savivaldybių ribos...
6. ... ir tos savivaldybės, kurių vietinės Morano I reikšmės statistiškai reikšmingos, matomos visos Lietuvos kontekste.

Dabar, norint pateikti šiuos rezultatus, kyla klausimas, kaip šią informaciją prasmingu būdu pavaizduoti žemėlapyje. Kurso dalies teorijoje tai buvo aprašyta taip: D–D ir M–M kategorijų savivaldybių vietinė autokoreliacija erdvėje yra teigiama (vietinės Morano I reikšmės panašios), D–M ir M–D savivaldybių – neigiama (vietinės Morano I reikšmės nepanašios).

Dabar tos savivaldybės, kurių vietinės Morano I reikšmės statistiškai reikšmingos, turi būti suklasifikuotos, kad patektų į vieną iš šių keturių klasių. Vidutinis nusikalstamumas savivaldybėse yra 183 nusikaltimai 10 000 gyventojų. Ši statistika gali būti panaudota suklasifikuoti savivaldybėms į didelio arba mažo nusikalstamumo. Naudodami tiek Lithuania_2005.SHP, tiek SHAPE failą, kuriame yra reikšminga vietinė Morano I statistiką, suklasifikuokime kiekvieną statistiškai reikšmingą savivaldybę.

Pavyzdžiui, jei statistiškai reikšminga savivaldybė turi teigiamą vietinę Morano I reikšmę, ir jos nusikalstamumo lygis didesnis už vidutinį, ji bus D–D. Kitu atveju, jei statistiškai reikšminga savivaldybė turi neigiamą vietinio Morano I reikšmę, ir jos nusikalstamumo lygis mažesnis už vidutinį, ji bus M–D.

Kad išdėstytumėte šiuos rezultatus žemėlapyje, turėsite sukurti naują kintamąjį.

1. Atidarykite statistiškai reikšmingų vietinės Morano I statistikų atributų lentelę.
2. Spustelėkite *Options* (parinkčių) mygtuką ir pasirinkite *Add Field...* (pridėti lauką).
3. Nustatykite pavadinimą: *Cluster*; tipą *Type* kaip tekstinį *Text*; o ilgį *Length* – 50.
4. Paspauskite *Start* ir pradėkite redagavimo seansą *Edit Session*.
5. Išanalizuokite kiekvienos statistiškai reikšmingą vietinę Morano I reikšmę turinčios savivaldybės aplinką nusikalstamumo santykio požiūriu. Įveskite klasterio tipą (D–D, M–M, D–M arba M–D) kiekvienai statistiškai reikšmingai vietinei Morano I reikšmei.

Sudarykite šį rezultatą vaizduojantį žemėlapi, naudodami tokias spalvas: M–M (žalia), D–D (raudona), M–D (mėlyna), D–M (rausva).

4 klausimas: pateikite CRIMRATE klasterių tipus vaizduojantį žemėlapi. Aptarkite šį žemėlapi, nurodykite, kiek, kur ir koks klasterio tipas. (7 balai)

5 klausimas: kokios informacijos, kurios neturi globalūs autokoreliacijos erdvėje (iš ankstesnės užduoties) matai, pateikia autokoreliacijos erdvėje vietinių klasterių analizė? (5 balai)

C dalis. Vietinė klasterių analizė ir kintamo plotinio vieneto problema

Kaip ir ankstesnėje šios dalies užduotyje, gauti rezultatai gali būti tiesiog pasirinktų analizuoti erdviųjų vienetų padarinys. Kad patikrintume prieš tai rastų vietinių klasterių jautrumą, dabar ketiname atlikti tokią pat analizę naudodami Lietuvos apskritis.

Pridėkite atliekant ankstesnę užduotį sugeneruotą Lietuvos apskričių sluoksnį, arba atlikite tos užduoties veiksmus, kad sugeneruotumėte tą sluoksnį iš naujo. Skaičiuoti CRIMRATE jums reikės tik tuo atveju, jei pametėte ar ištrynėte ankstesnį apskričių sluoksnį.

Naudodamiesi anksčiau nusakytais veiksmais, sudarykite reikšmingų vietinių Lietuvos apskričių nusikalstamumo lygio klasterių žemėlapi. Gautajame žemėlapyje rastiems klasteriams vaizduoti turėtų būti naudojama anksčiau aprašyta simbolika.

6 klausimas: pateikite CRIMRATE klasterių tipus apskrityse vaizduojantį žemėlapi apskritims. Aptarkite šį žemėlapi, nurodykite kiek, kur ir koks klasterio tipas. (3 balai)

7 klausimas: kiek šie rezultatai skiriasi nuo globalios Lietuvos apskričių nusikalstamumo santykio autokoreliacijos erdvėje? Kas analizėje pasikeitė naudojant kitą erdvinę skiriamąją gebą? Ar tai turi reikšmės? (5 balai)