

**Majsko srečanje '89 Sekcije za klasifikacije in nomenklature  
Zveze statističnih društev Jugoslavije  
v Mostarju (poročilo)**

**Anton Kramberger**

Na že tradicionalnem srečanju članov Sekcije za klasifikacije in nomenklature ZSD Jugoslavije, ki je potekalo v prijetnih prostorih vile "Neretva" v Mostarju, pod pokroviteljstvom Zveznega Zavoda za statistiko in skrbno organizacijo Ekonomske fakultete "Džemal Bijedić" iz Mostarja, so se letos prepletale teme teoretične statistike, aplikacije klasifikacijskih metod, predstavitve izkušenj v delu z računalniki in programskimi paketi ter dileme klasifikacij, ki jih kot standarde uporablja statistična služba. V juliju je Institut za statistiko Zveznega zavoda za statistiko izdal zbornik 18 prispevkov s tega srečanja, ki jih v tem besedilu na kratko predstavljamo.

Največ prispevkov je bilo s področja teoretične statistike. Ti avtorji so z različnih vidikov skušali odgovoriti na temeljno vprašanje sodobne teorije klasifikacije (taksonomije, razvrščanja ipd.): koliko je dobljena rešitev (razvrstitev) dobra, objektivna, če jo primerjamo z celo družino dopustnih razvrstitev nad isto množico enot ali pa celo z realnostjo? Tovrstnim raziskovalnim zamislim sledijo nekoliko bolj aplikativne raziskovalne pobude. Naj navedemo nekaj glavnih rezultatov predstavljenih raziskovanj:

**B. Ivanović:** je v svojem prispevku "Razdalja med dvema klasifikacijama iste statistične množice" obravnaval naslednji problem: če razvrščamo enote neke množice po različnih kriterijih, nas zanima kakšna je razvrstitev teh različnih razvrščanj. V ta namen avtor definira in numerično določi razdaljo med dvema razvrstitvama, dokaže, da algebrska struktura, ki jo sestavljajo vse možne razvrstitve

in operacija razdalja, tvori metrični prostor. Tako lahko dobi matriko razdalj med klasifikacijami in z eno od znanih metod, recimo hierarhičnim združevanjem, določi drevo združevanja vseh klasifikacij. Teoretična rešitev je bila ponazorjena s praktičnim primerom vpliva metropol na nekdanje kolonije.

**M. Siudak in M. Trocki:** poljska avtorja sta v svojem članku uredila postopke za merjenje podobnosti particij (razvrstitev) izvorne množice v dve najpogostejši družini: v prvi družini primerov se primerjajo rezultati razvrstitve množice po dveh kriterijih, v drugi družini primerov pa se primerja stabilnost klasifikacije istih enot v dveh časovnih točkah. Avtorja sta ponudila enostavno metodo za merjenje podobnosti razvrstitev in jo ilustrirala s primerom.

**A. Ferligoj:** avtorica je v prispevku "Metoda prestavljanj za večkriterijsko razvrščanje v skupine" predstavila rezultate svojega kontinuiranega teoretičnega dela na večkriterijskem razvrščanju, v katerem skuša najti najboljšo rešitev, s kar največjim upoštevanjem uporabljenih kriterijev. Pri tem je uporabila pristop, ki se naslanja na multikriterijsko optimizacijsko teorijo, vendar se ne približuje rešitvi preko lokalnih optimizacij, ampak direktno, preko optimizacije celotnega multikriterijske razvrščanja. Ker ni na voljo eksatne rešitve problema, je potrebno rešitve najti s približnimi procedurami, pri čemer je avtorica predstavila učinkovit računalniški algoritem, ki počiva na relokacijski metodi.

**V. B. Kudrjavcev in K. Momirović:** avtorja sta z delom "O enem pristopu problema prepoznavanja" prikazala zelo splošno rešitev družine problemov prepoznavanja vzorcev, ki po vsebini lahko predstavlja tudi teorijo, pristop h klasifikaciji enot (po podobnosti). V ta namen sta definirala splošno matriko objektov (z nekimi vnaprej podanimi lastnostmi), posplošila sta pojem podobnosti in

uvedla projekтивne operatorje nad objekti in razredi objektov, ki predstavljajo podmatrke splošne matrice. Dokazala sta nekaj teoremov, ki jamčijo obstoj rešitev zastavljenih problemov, tudi v primeru nepopolnih in celo napačnih podatkov.

**K. Momirović:** v prispevku "Kvazikanonična diskriminativna analiza v Pearsonovi diskriminativni metriki" nam avtor predlaga eno varianto kanonske diskriminativne analize, namenjene analizi takšnih podatkov, ki ne zadoščajo znanim predpostavkami za multivariatno analizo (večrazsežna normalna porazdelitev, odsotnost ekstremnih vrednosti, dobra determiniranost matrice kovarianc). Rešitev, ki je naslonjena na kanonično analizo kovarianc reparametriziranih spremenljivk, je tako možna, vendar daje robustnejše rešitve.

**M. Jerman:** mladi srednješolec, ki je v Sloveniji za predstavljeno delo "Vloga koeficientov v Lance-Williamsovi formuli pri razvrščanju v skupine" dobil najvišjo nagrado kot mladi raziskovalec, se je ukvarjal s posebno družino problemov v okviru hierarhičnih metod združevanja; skušal je poiskati obetavno metodo združevanja enote in skupine, pri čemer naj bi imela ta metoda v splošni Lance-Williamsovi formuli konstantne koeficiente. Pri empiričnem raziskovanju je variiral koeficiente, naletel na zanimivo metodo, jo izločil in podrobneje obravnaval ter prikazal njene lastnosti.

**S. Mičković:** avtor je v referatu z naslovom "Metoda združevanja, zasnovana na meri različnosti, specifični za časovne serije" podal posebno metodo razvrščanja za časovne serije. Mere različnosti upoštevajo specifične funkcije časovnih serij (recimo avtokorelacijo). Avtor je prikazal posebni algoritem, ki pripelje do voditeljev skupin, metodo pa je se raziskal z Monte Carlo eksperimenti. Zaključki so zaenkrat veljavni za dva modela časovnih serij in za dva

voditelja. Na kraju je avtor podal primerjavo z metodo voditeljev, ki operira z evklidsko razdaljo.

**F. Prot:** za učinkovito uporabo metod taksonomske analize je predpogoj objektivnost v postopkih razvrščanja. To izhodišče je avtorju služilo pri izdelavi algoritma in programa, ki ju je predstavil v referatu z naslovom "Algoritem in program za prepoznavanje skupin igralcev v športnih igrah". Osrednja zamisel algoritma je v simultanem izvajanju objektivnega in konfirmativnega (kriterijskega) tipa razvrščanja. S primerjavo obeh rešitev lahko ocenimo uspeh objektivnega razvrščanja in prepoznamo spremenljivke, ki so dominantno določile položaj objektov v taksonomskem prostoru. Algoritem je splošne narave: z njim lahko razvrstimo objekte, če imamo podano množico objektivnih meritev in še podatek o pripadnosti objektov k eni distinktivni podmnožici.

**K. Bosnar:** avtorica se v referatu "Uporaba klasifikacijskega postopka nad parametri regresijske analize" loteva zanimivega praktičnega problema: ko raziskovalec/znanstvenik skuša interpretirati rezultate svojih analiz, mu obstoječi izpisi iz različnih programskih orodij tega ne olajšujejo, saj so podrejeni algoritmom izračunavanja, ne pa algoritmom interpretacije. Izpeljevanje zaključkov pa je kognitivna funkcija, ki se sicer lahko do neke mere naslanja na paralelno procesiranje sklepanja, pri preveliki količini informacij pa se mora opreti tudi na serijsko procesiranje informacij. Avtorica s primerom objektivne klasifikacije nad rezultati regresijske analize skuša pokazati, da je možno subjektivno interpretacijo tudi objektivno prepoznati in določiti numerično mero ujemanja subjektivne in objektivne sodbe.

**D. Acketa in N. Carić:** avtorja sta v prispevku "Postopek za grupiranje organizacijskih enot v podjetju" podala algoritem za razvrščanje organizacijskih enot v podjetju v večje organizacijske enote, ki operira nad vhodno matriko, v

kateri so podane kooperativne vezi med delovnimi mesti in organizacijskimi enotami. Ko se preko različnih transformacijskih funkcij dobi matrika različnosti, se izvede še hierarhično združevanje, ki pa ga je v samem programu mogoče interaktivno usmerjati.

**N. Viskić-Štalec in R. Bujanović-Pastuović:** avtorici zelo zanimivega referata "Evolucija devijantnega in delikventnega vedenja" sta na vzorcu 860 ljudi iz splošne populacije, starih od 18 do 27 let, ki so izpovedali lastno delikventno obnašanje, prišli do naslednjih zaključkov: mlade osebe, ki jim lahko dobo zorenja razdelimo v dve poddobi, prvo, od 14. do 18. leta starosti, in drugo, ki objema starost nad 18 let, se lahko znajdejo primežu enega od treh faktorjev (splošni faktor persistence devijantnega in delikventnega vedenja, faktor anarhoidnega vedenja in faktor transformacije od lažje k teji delikvenci). Med tendencami, ki so prisotne v adolescenci, in tendencami, ki prevladujejo v mlajši polnoletni dobi, so ujemanja, potrjena tudi s kongruenco nekaterih kanoničnih in kvazikanoničnih faktorjev.

**M. Gredelj, A. Hošek in K. Momirović:** avtorji metodološko in vsebinsko zanimive študije "Taksonomska analiza kvalitativnih sprememb v strukturi kriminalnosti" so predstavili razlike v pogostosti kaznivih postopkov proti čivljenju in telesu in proti imetju v hrvaških občinah v 1979. in 1987. letu. Metodološko je bil uporabljen algoritem kanonične kovariančne analize, ki preiščuje kvalitativne spremembe objektov, opisanih v dveh časovnih točkah, in ki je definiran nad množico kvalitativnih spremenljivk absolutnega tipa. V vsebinskem smislu se je algoritem pokazal kot učinkovit mehanizem za prepoznavanje strukturnih razlik v kriminalnosti po občinah Hrvaške v dveh časovnih točkah.

Sledi strokovni referat, nekoliko pedagoško orientiran, ki povzema izkušnje pri računalniškem razvrščanju z različnimi paketi:

**J. Jug:** avtor v prispevku z naslovom "Programi za razvrščanje v skupine v programskih paketih CLUSE, SPSS in SAS" primerja podprograme za razvrščanje v skupine v omenjenih programskih paketih, ki se razlikujejo med seboj po obsegu, zmogljivosti in namenu. Od vseh treh je najbolj specializiran paket CLUSE, druga dva pa se med seboj razlikujeta po tem, da je SPSS zaprt proceduralni jezik, SAS pa odprt proceduralni jezik - oba sta namenjena splošni statistični analizi, ki vsebuje tudi postopke za razvrščanje. Avtor podrobneje obdela naslednje vsebinke sklope po paketih: tip vhodne matrike, način obravnavanja manjkajočih vrednosti, mere različnosti in podobnosti, metode združevanja, algoritme za razvrščanje velikega števila enot.

Naslednja referata sta izrazito aplikativne narave:

**Z. Jevtić in D. Djurić-Bijelić:** avtorja v članku "Določanje položaja Jugoslavije v skupini izbranih dežel na osnovi nekaterih kazalnikov razvoja" razvrščata 26 držav, med njimi Jugoslavijo, v dveh obdobjih: prvo obsega čas med 1965. in 1980. letom, drugo pa od 1980. do 1986. leta. Najprej uporabljata faktorsko analizo: za prvo obdobje so značilni faktorji dinamika rasti družbenega proizvoda, zunajetrovinska menjava in osebna potrošnja; za drugo obdobje pa so značilni uvoz in izvoz, investicije in splošna potrošnja. Nato države razvrstita s hierarhičnim združevanjem: za prvo obdobje dobita tri skupine, v drugem obdobju pa šest skupin, ki v glavnem nastanejo zaradi integracijskih procesov v evropski skupnosti oziroma zaradi povezav v EFTI.

**S. Stević:** Avtor v članku "Klasifikacija občin severovzhodne Bosne glede na delež zaposlenih v dejavnostih družebnega sektorja" najprej na osnovi dveh tridimenzionalnih nominalnih spremenljivk (pretežna sektorska dejavnost in tip pretežne dejavnosti) razvrsti vse občine obravnavne regije v 9 skupin, nato pa si pomaga z nazorno metodo grafike trikotnika, kjer je lepo razvidnih nekaj taksonomskih značilnosti prevladujočih skupin.

Na kraju naj povzamemo še nekaj odprtih dilem, ki izhajajo iz pragmatičnih skrbi v sistemu statističnih raziskovanj; letos so trije avtorski prispevki najavili morebitno novo vsebino vsakoletnih mostarjih srečanj:

**S. Misović:** avtorica v izčrpnem poročilu z naslovom "Klasifikacije in nomenklature proizvodov in storitev v statističnem sistemu" podaja poglobljena odprta vprašanja, ki se tičejo omenjenih standardov statističnega sistema. Kritizira pomanjkljivosti in nekonsistentnost kriterijev za razvrščanje v uradnih klasifikacijah, hkrati pa oceno omili z navedbo, da so te neljube lastnosti statističnih standardov prisotne tudi v mednarodni statistiki.

**M. Nikolić:** z dvema prispevkoma, ki se oba dotikata nekih specifičnosti klasifikacijskih postopkov v uradni statistiki, avtor odpira vrsto vprašanj, ki bi se lahko reševala tudi bolj teoretično. V prvem prispevku, "O postopku klasificiranja glede na opisne značilnosti" navaja dobro znane principe v induktivnih in deduktivnih procedurah, ki jim sledi uradna statistika. Predlaga uporabo metod statistične analize za strožje ocenjevanje klasifikacijske kvalitete. V drugem prispevku z naslovom "K razpravi o klasifikacijah uradne statistike" pa splošno razpravo nekoliko natančneje opredeli in meni, da bi Statistično društvo z večjim zanimanjem za opisne značilnosti lahko mnogo pripomoglo k izboljšavi uradnih klasifikacij: nomenklaturi poblicev, klasifikaciji gospodinjstev, klasifikaciji naravnih nesreč, klasifikaciji

dejavnosti in nomenklaturi proizvodov. Avtor predlaga, da bi te teme lahko postale samostojne točke dnevnega reda naslednjih mostarskih srečanj.