

**O`ZBEKISTON RESPUBLIKASI OLIY VA
O`RTA MAXSUS TA`LIM VAZIRLIGI**

TOSHKENT MOLIYA INSTITUTI

“STATISTIKA” KAFEDRASI

STATISTIKA FANIDAN

REFERAT

Mavzu: Statistik jamlash va guruhlash

BAJARDI: KBI-64-guruh talabasi
Maxmasoliyeva Sh.

TEKSHIRDI: o`q.Utanov B.

TOSHKENT-2014

STATISTIK JAMLASH VA GURUHLASH

1. Statistikada jamlash: turlari va mohiyati

O'rganilayotgan hodisa bo'yicha statistik kuzatish o'tkazish natijasida u haqida ko'pdan-ko'p va turli-tuman tarqoq ma'lumotlar to'planadi. Bu ma'lumotlar asosida hali hech qanday fikr yuritib bo'lmaydi, chunki ular tarqoq va har xildir. Shuning uchun ham navbatdagi vazifa to'plangan ma'lumotlarni bir tizimga keltirish, tartibga solish, umumlashtirishdan iboratdir. Bu muammo statistikada jamlash (svodkalash) metodini qo'llash bilan hal etiladi. Jamlashdan statistik tekshirishning ikkinchi bosqichi boshlanadi.

Statistik jamlash (svodkalash) deganda har kuni radio va televizorda beriladigan informatsiyani tushunmaslik kerak. Yuqorida ta'kidlaganimizdek informatsiya ma'lum bir ishning bajarilishi to'g'risidagi operativ ma'lumotdir.

Statistik jamlash deganda to'plangan ma'lumotlarni ilmiy tekshirishdan ko'zlangan maqsad va vazifalar nuqtai-nazaridan qayta ishlash tushuniladi.

Statistik jamlash hisoblash texnologiyasiga qarab oddiy va murakkab, tashkil qilinishiga qarab esa markazlashgan va markazlashmagan jamlashlarga bo'linadi.

Oddiy jamlash deganda olingan ma'lumotlarni umumiy yakunlarini hisoblash tushuniladi, murakkabda esa to'plam birliklari guruh va guruhchalarga ajratilib, ularning har biri va umumiy bo'yicha jamlar chiqariladi.

Markazlashgan jamlashda barcha ma'lumotlar bir erga to'planib, o'sha erda qayta ishlanadi. Masalan, aholi ro'yxati ma'lumotlarini qayta ishlash. Markazlashmagan jamlashda ish pog'onama-pog'ona amalga oshiriladi, masalan, dastlab tumanda, keyin viloyatda va h. k.

Statistik jamlash bir qancha bosqichlarda amalga oshiriladi va quyidagi elementlarni o'z ichiga oladi:

1. To'plangan ma'lumotlarni qayta ishlash rejasi va dasturi tuziladi;
2. To'plam birliklari o'rganilayotgan belgi bo'yicha guruhlariga ajratiladi.
3. Har bir guruh va umumiy to'plam bo'yicha jamlar chiqariladi.
4. Natijalar statistik jadvallarga joylashtiriladi va grafiklarda tasvirlanadi.

2. Guruhlash metodi: mohiyati, ahamiyati va turlari

Statistik kuzatish ma'lumotlarni qayta ishlash va tahlil qilishning keng tarqalgan metodlaridan biri guruhlashdir. Statistikada guruhlash deb to'plam birliklarini eng muhim belgilari bo'yicha bir xil guruh va guruhchalarga ajratib o'rganishga aytiladi.

Guruhlash metodi statistikada ko'p yillardan (XVIII asr) beri qo'llanib kelinmoqda. Bu metodni qo'llashdan maqsad, to'plam birliklarini qanday bir bo'laklarga bo'lish emas, balki faqat shu hodisaga xos xususiyatlarni ochib berish, undagi mavjud tendentsiya va qonuniyatni baholash, miqdor o'zgarishlardan sifat o'zgarishlarga, sifat o'zgarishlardan miqdor o'zgarishlarga o'tish jarayonlarini aniqlash va baholashdir.

Guruhlash metodi oldida o'rganilayotgan to'plam birliklarini tiplarga ajratish, hodisalar o'rtasidagi bog'lanishlarni va to'plam tuzilishini o'rganish vazifalari ham turadi. Bu vazifalar guruhlashning uch (tipologik, analitik, tuzilmaviy) turidan foydalanish orqali hal qilinadi.

Aholini jinsi bo'yicha guruhlariga ajratish, sinflarga bo'lish, mulkni davlat va shaxsiy mulkka bo'lish va boshqalar tipologik guruhlashga misol bo'la oladi. Bu misollardan ko'rinib turibdiki, o'rganilayotgan to'plamning turli xildagi birliklari tipologik guruhlash yordamida sifat jihatdan bir xil guruhlariga, ya'ni tiplarga ajratiladi.

Analitik guruhlashdan maqsad hodisalar o'rtasidagi bog'liqlikni o'rganishdir. Masalan, chakana savdo do'konlarida tovar oboroti hajmi va rentabellik darajasi o'rtasidagi bog'liqlikni o'rganish uchun do'konlarni tovar oboroti hajmi bo'yicha guruhlariga ajratib har bir guruh uchun rentabellik darajasini hisoblaymiz. Agarda tovar oboroti oshib borishi bilan rentabellik darajasi oshib borsa, ular o'rtasida to'g'ri bog'lanish mavjud, aksincha bo'lsa, teskari bog'lanish mavjud bo'ladi.

Amaliyotda shunday qiyin va murakkab hodisalar uchraydiki, ularni chuqur o'rganish uchun ikki va undan ortiq belgilar bo'yicha guruhlariga ajratish maqsadga muvofiq bo'ladi. Statistika bunday guruhlash kombinatsion guruhlash metodi deb yuritiladi.

Tuzilmaviy guruhlashga taqsimot qatorlari, variatsion qatorlar yorqin misol bo'ladi (Ushbu savollar kelgusi mavzularda ko'riladi). Tuzish usuliga qarab guruhlash quyidagi turlarga bo'linadi:

1. Umumdavlat klassifikatorlari;
2. Matematik va ekspert-analitik metodlar asosida guruhlash;
3. Bir o'lchovli yoki ko'p o'lchovli guruhlash

Guruhlashni amalga oshirish uchun dastlab guruhlash belgisini aniqlab olish muhimdir.

Iqtisodiy fanlardagi nazariy holatlar va aniq tekshirishning vazifalariga tayangan holda, guruhlashni amalga oshirishda mavjud belgilardan eng hal qiluvchisi tanlanadi. Hal qiluvchi belgi bo'lib, o'rganilayotgan ob'ektni har tomonlama va aniq xarakterlovchi, uning o'ziga xos holat va xususiyatlarini tanlashda yordam beruvchi belgilarga aytiladi. Masalan, savdo korxonalari, savdo zalining maydoni, xodimlar soni va hokazolar bilan tasniflanishi mumkin. Lekin savdo korxonalari faoliyati (ish hajmi)ni baholashda tovar oboroti hajmi hal qiluvchi belgi hisoblanadi, chunki savdoning asosiy maqsadi tovar sotishdan iboratdir.

Shunday qilib, guruhlash belgisini tanlashda, statistika quyidagi shartlarga e'tibor berishni taklif etiladi: guruhlar negiziga doimo hodisani to'la-to'kis tavsiflab beruvchi muhim belgilarni asos qilib olish zarur; guruhlash belgisini tanlashda uning aniq vaqt va joy sharoitini, o'sha davrning mohiyatini ifodalovchi, zamonaviy masalalarni yoritib oladigan belgilarga e'tibor berish zarur; hodisalarni guruhlashda belgilar soni etarli bo'lgani ma'qul.

Guruhlash belgisi ifodalanishiga qarab atributiv va miqdoriy belgilarga, ta'siriga qarab omil va natijaviy belgilarga, ko'zlangan maqsad va vazifalarga qarab muhim va muhim bo'lmagan belgilarga bo'linadi.

Guruhlash belgisi to'g'ri tanlangandan so'ng, eng muhim masalalardan biri, to'plam birliklarini guruhlariga ajratishdir. Bu erda guruhlar soni va oralig'i (intervali)ni aniqlash muammosi paydo bo'ladi. Bu muammo miqdoriy belgilar bo'yicha guruhlashga tegishlidir, chunki tipologik guruhlashda, odatda, guruhlar soni tekshirishning vazifasi bilan aniqlanadi. Masalan, aholini jinsi bo'yicha guruhlash vazifasi qo'yilsa, bu erda guruhlar soni doimo aniq-ikkita. Banklarni kapital miqdori bo'yicha guruhlariga ajratishdan maqsad, ularni katta, o'rtacha va kichik bank ekanligini aniqlash bo'lsa, guruhlar soni bu erda ham aniq -uchta.

Xo'sh guruhlar soni qanday aniqlanadi? Tayyor retsept yo'q. Odatda, hodisaning taqsimlanish xarakterini aniqlash uchun guruhlar soni ko'proq olinadi. Bu erda o'rganilayotgan belgining tebranishi (o'zgarishi) e'tiborga olinadi, ya'ni tebranish qancha katta bo'lsa, odatda, guruhlar soni ko'proq bo'ladi. Yana bir narsa. O'rganilayotgan to'plamdagi birliklar soni ham muhimdir. Agarda, ularni soni juda ko'p bo'lmasa, guruhlar sonini ko'paytirish maqsadga muvofiq emas, chunki guruhlarda to'plam birliklari soni kam bo'ladi va ularni ifodasi etarli tipik bo'lmasligi mumkin. Lekin bu qoidani mahkam ushlab olish ham noto'g'ridir. Bizga ma'lumki, ayniqsa bozor iqtisodiyotiga o'tish davrida, yangi, ilg'or, zamonaviy hodisalar kamroq bo'ladi (chunki ular hali ommaviy emas, to'g'risini aytish kerak ularga xalq qo'rqibroq yondashadi). Guruhlash metodining vazifasi bu faktlarni ular qanchalik kam sonli bo'lishiga qaramasdan aniqlash va chuqur o'rganishdir.

Guruhlar soni aniqlangandan keyingi muhim masala, guruhlar oralig'ini (intervali)ni aniqlashdir. Oraliq (interval)deb, guruhdagi eng maksimal va minimal variantlarning farqi tushuniladi.

O'rganilayotgan belgi birliklarining taqsimlanish xarakteriga qarab guruh oralig'i teng va teng bo'lmagan holda hamda ochiq va yopiq, maxsus ko'rinishlarda bo'lishi mumkin.

Agar belgining variatsiyasi juda tor chegarada namoyon bo'lib, taqsimlanish bir muncha tekis bo'lsa, oraliqlar teng qilib belgilanadi. Teng oraliq deyilganda barcha guruhlar uchun bir xil bo'lgan oraliq tushuniladi. U quyidagicha hisoblanadi:

$$i = \frac{X_{\max} - X_{\min}}{n} \quad \text{ёku} \quad i = \frac{X_{\max} - X_{\min}}{1 + 3,322 \cdot \lg \cdot N}$$

bu erda: i-oraliq kattaligi; $X_{\max}$ -belgining eng katta qiymati; $X_{\min}$ -belgining eng kichik qiymati; N-to'plamdagi birliklar soni; n-guruhlar soni. Agarda guruhlar soni aniq bo'lmasa, ularning optimal sonini Sterdjess formulasi bilan aniqlaymiz:

$$n = 1 + 3,322 \lg \cdot N$$

Faraz qilaylik, savdo korxonalarining oylik tovar oboroti 50 mln. so'mdan 80 mln. so'mgacha bo'lsa, ularni 6 ga teng intervalli guruhga ajratsak, u holda interval miqdori quyidagicha bo'ladi:

$$i = \frac{80 - 50}{6} = \frac{30}{6} = 5 \text{ млн. сум}$$

Guruhlash belgisining eng kichik qiymatiga 5mln. so'mni qo'shsak, birinchi guruhning chegarasi kelib chiqadi: 50+5=55mln. so'm. Demak, birinchi guruhga 50-55mln. so'mgacha tovar oborotiga ega bo'lgan korxonalar kiradi. Qolgan guruhlar: 55-60; 60-65; 65-70; 70-75; 75-80mln. so'm.

Teng bo'lmagan oraliq deganda guruhdan guruhga o'zgarib boruvchi interval tushuniladi. Bunday oraliqlar, odatda, to'plam birliklari juda katta tarqoqlikka ega bo'lgan hollarda qo'llaniladi. Masalan, do'konlar kichik, o'rta va yirik do'konlardan tashkil topgan bo'lsa, ularni tovar oboroti bo'yicha bir xil intervallarda guruhlarga ajratsak ayrim guruhlarga bitta ham do'kon tushmasligi mumkin.

Guruhlar chegarasini aniqlashda ayrim qoidalarga rioya qilishga to'g'ri keladi. Birinchidan, "gacha" deganda yoki "unda yuqori" so'zlari yozilganda qanday yo'l to'tish kerak. Masalan, do'kon xodimlarining mehnat unumdorligi bo'yicha 5 ta guruh tuzilgan: 90ming so'mgacha; 90-120; 120-150; 150-180; 180 dan yuqori. Bunday misollarni bir necha xilini keltirish mumkin. Bu erda masalan, 120 ming so'mlik mehnat unumdorligiga ega bo'lgan xodimni qaysi guruhga kiritish tekshiruvchining o'ziga va aniqrog'i uning yozuviga bog'liq. Statistikada ikkita tamoyil (qo'shilgan va qo'shilmagan holda) mavjud. Agarda "qo'shilgan holda" tamoyilini qo'llasak 90 ming so'm unumdorlikka ega bo'lgan savdo xodimi birinchi guruhga, "qo'shilmagan holda" tamoyili qo'llansa-ikkinchi guruhga qo'shiladi. Endi yozuvga e'tibor bering, oxirgi guruhda "180 dan yuqori" deb yozilgan. Demak, 180 ming so'mlik unumdorlikka ega bo'lgan xodim oxiridan oldingi guruhga kiradi. Agarda "180 va yuqori" so'zi yozilgan bo'lsa edi, bu xodimni oxirgi guruhga kiritgan bo'lar edik.

Oraliqlarni o'rtachasini aniqlash ham muhim ishlardan biri. Bu ish quyidagicha bajariladi. Intervalni quyi va yuqori chegarasi qo'shib ikkiga bo'linadi. Ikkinchi guruh uchun $u = (90+120):2 = 105$ ming so'mga teng. Uchinchi guruh uchun $(120+150):2 = 135$ ming so'm. Bu natijani ikkinchi guruh intervali o'rtachasi 105 ming so'mga interval farqini qo'shish orqali ham olish mumkin ($105+30=135$) Ochiq intervalli guruhlarda birinchi va oxirgi guruhlarining o'rtacha darajasini aniqlash quyidagicha amalga oshiriladi. Teng oraliqli guruhlarda, birinchi guruh o'rtachasini aniqlash uchun ikkinchi guruh o'rtachasidan oraliq hajmi ajratiladi ($105-30=75$), oxirgi - o'zidan oldingi guruh o'rtachasiga oraliq hajmi qo'shiladi ($165+30=195$). Teng bo'lmagan oraliqlarda birinchi guruh oralig'i ikkinchi guruh oralig'iga teng, oxirgi guruh oralig'i o'zidan oldingi guruh oralig'iga teng deb qabul qilinadi.

Ikkilamchi guruhlash. Guruhlashning xususiy turi bo'lib ikkilamchi guruhlash hisoblanadi. Ikkilamchi guruhlash deb oldingi tuzilgan guruhlar asosida yangi guruhlar tuzish operatsiyasiga aytiladi. Agarda birlamchi guruhlashda statistik kuzatishning boshlang'ich ma'lumotlari asosida guruhlar tuzilsa, ikkilamchi guruhlash dastlabki guruhlash oraliqlarini yiriklashtirish va

oraliqlarning nisbatiga asosanib yangi guruhlarini hosil qilish usullarida amalga oshiriladi.

Faraz qilaylik, Chilonzor tumanida 100 ta do'kon bor. Ular inkassatsiya qilish summolari bo'yicha 10 guruhga ajratilgan: 100 ming so'mgacha; 100-200; 200-300; 300-400; 400-500; 500-600; 600-700; 700-800; 800-900; 900 va yuqori. Bu intervallar oralig'ini ikki baravarga yiriklashtirib quyidagi guruhlarini hosil qilish mumkin: 200 ming so'mgacha; 200-400; 400-600; 600-800; 800 va undan yuqori. Ikkilamchi guruhlashning boshqa usullari ham qo'llanilishi mumkin. Bu qo'yilgan maqsad va vazifaga bog'liq.

Ko'p o'lchamli guruhlash (klaster-tahlil). Keyingi paytlarda guruhlash bir vaqtning o'zida bir necha belgi orqali amalga oshirilmoqda. Buning o'zi guruhlash metodini ko'p o'lchamli tahlilga aylanib borishidan darak beradi. Ma'lumki, ko'p o'lchamli guruhlashda yoki klaster – tahlilida kuzatish ob'ektlarini xohlangan belgilar soni bo'yicha bir jinsli guruhlariga birlashtirish mumkin. Shunisi qiziqki kuzatilayotgan ob'ekt sifatida iqtisodiy birliklar-korxonalar yoki belgilarning o'zi qatnashishi mumkin.

Klaster-tahlil algoritmlari ikki asos bo'ladigan paytni hisobga olgan holda ishlab chiqiladi:

1. Bir turlilikni, yoki "o'xshamas" ob'ektlarni ifodalovchi belgilarning geometrik maydonda juda ko'p nuqtalarni tiqis to'plamini ko'rsatib berish sharoitlarini.

2. Geometrik maydonda ikki turli ob'ektlar bir-biridan bir muncha uzoqlikda joylashgan va ular orasidagi masofa qancha uzoqlashsa, ular shuncha o'xshamas va qancha yaqinlashsa ularning o'xshashligi shunga ortadi; nollik variant hamma vaqt qandaydir bir ob'ektdan o'zigacha, bu erda to'liq o'xshashlik.

Aniq algoritmnini tanlashga qaramasdan, klaster-tahlil quyidagi qadamlarni birin-ketinlik bilan bajarish sharoitida amalga oshiriladi:

a) "X"- boshlang'ich ma'lumotlarni $n \times m$ razmerdagi matritsalarini tuzish, bu erda n – kuzatish ob'ektlari soni; m – guruhlashtiruvchi belgilar soni;

b) boshlang'ich ma'lumotlar matritsalaridan normalashtirilgan ma'lumotlar matritsalariga o'tish (z). Bu masalani echilishi bilan o'z tabiati bo'yicha turli bo'lgan belgilar bitta asosga keltiriladi. O'tish har bir qiymatni qayta hisoblash X_{ij} va Z_{ij} quyidagi variantlar orqali amalga oshadi.

$$1. Z_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_i}; \quad 2. Z_{ij} = \frac{x_{ij}}{x_j}; \quad 3. Z_{ij} = \frac{x_{ij}}{x_j(\text{эталон})};$$

$$4. Z_{ij} = x_{ij}(\text{max}); \quad 5. Z_{ij} = \frac{x_{ij} - \bar{x}_j}{x_{j(\text{max})} - x_{j(\text{min})}}.$$

v) barcha juft ob'ektlar orasidagi masofani aniqlash (d_{ij}) va dastlabki matritsalar masofasini tuzish (D_0). Kuzatish ob'ektlari o'rtasidagi masofani hisoblash uchun bir qancha metrikalar (i_1 – norma; Minkovski; Evklidovo masofa; Maxalanobis) mavjud. Ularni qaysi birini qo'llash tekshiruvchining xohishiga bog'liq.

g) klaster-tahlilni aniq protsedurasi tanlanadi va matritsa (D_0) ma'lumotlari bo'yicha birin-ketinlik bilan bir turli guruhlar ajratiladi. Eslatib o'tmoqchimiz, hozirgi kunda klasterlashni 200 dan ortiq turli xil protseduralari mavjud. Ularni quyidagi 6 ta guruhga bo'lish mumkin: ierarxik klaster-tahlil; guruhlashning iterativ metodlari; zichlikning model qiymatini izlash metodlari; omiliy metodlar; quyuqlashishni izlash metodlari; graflar nazariyasini qo'llovchi metodlar.

Yuqorida keltirilgan qadamlar faqat miqdoriy o'lchovga ega bo'lgan belgilar tahlil qilingan paytda qo'llaniladi. Agarda tahlilda tartibli (ranglar) va boshqa sifat ko'rsatkichlar qatnashsa, keltirilgan algoritmdan oldin nomiqdoriy ma'lumotlarni oshifrovkalash etaplari birma-bir bajariladi.

Asosiy tayanch iboralar: statistik jamlash (svodkalash), statistik guruhlash, guruhlash belgisi, guruh intervali (oralig'i), ikkilamchi guruhlash, ko'p o'lchamli guruhlash.

Xaritalarning yana bir turi **xaritadiagrammalardir**. U geografik xaritalardagi hududlar hodisalarini diagrammalarda tasvirlash orqali hosil qilinadi. Bunda tasvirlovchi belgi(diagramma)lar bo'lib nuqta, chiziq, ustun, lenta, kvadrat, doira, har xil figuralar xizmat qiladi. Bu belgilar hududlarga tegishli hodisalarning darajasiga yoki xususiyatiga ko'ra turli o'lchovda geografik xaritalarga joylashtiriladi.

Siyosiy dunyo xaritasidagi shaharlarni ifodalovchi belgi – nuqtalar (doira) xaritadiagrammaga misol bo'ladi. Bunda nuqtaning kattaligi u shaharda aholi sonini ifodalaydi.

Xaritadiagramma xaritagrammalarga nisbatan geografik holatlarni o'rganishda qulayroq hisoblanadi. Unda murakkab statistik-geografik ma'lumotlarni ko'rinarli va tushunarliroq holda tasvirlash mumkin.

Shunday qilib, xulosa qilib aytganda, statistik grafiklar ijtimoiy-iqtisodiy va siyosiy hodisalarni o'rganishda muhim statistik qurollardan biri bo'lib, ularda o'rganilayotgan hodisa va jarayonlarning tashqi tasviri yorqin namoyon bo'ladi. Ularni insoniyat hayotida uchraydigan hodisa va jarayonlarning oynasi deb atasak xato bo'lmasa kerak.

Asosiy tayanch iboralar: statistik jadval, jadval maketi, statistik grafik, grafik tasvir, grafik maydon, maydon o'lchovi, masshtab (miqyos)shkala, grafik talqin, dinamik grafiklar, variatsion grafiklar, geografik grafiklar, analitik grafiklar, nuqtali grafiklar, chiziqli grafiklar, yassi grafiklar, hajmli grafiklar, diagramma, xaritogramma, gistogramma, poligon, kumulyata, ogiva, taqqoslovchi grafiklar