

különösen akkor lehet nagy hasznát venni, amikor a minta szélső elemei bizonytalanok.

Hasonló szellemben foglalkozik a szerző a szórásra adható becslésekkel is.

(Ism.: *Krekó Béla*)

Gautschi, Werner:

Néhány megjegyzés a szisztematikus mintavételről

(Some remarks on systematic sampling.) *The Annals of Mathematical Statistics*. 1957. No. 2. 385—394. p.

A szisztematikus (vagy más elnevezéssel mechanikus) reprezentatív mintavétel lényege a következő: tegyük fel, hogy az alapsokaság, amelyet vizsgálunk, N elemből áll, a mintába pedig n számú elemet kívánunk válogatni, továbbá

$$\frac{N}{n} = k$$

Ez esetben az alapsokaság egyedeit valamilyen módon sorba rendezzük és az első k elem közül véletlenszerűen választunk egyet, majd ezt követően minden k -adikat. Így nyerjük az n elemű mintát, amelyből számított átlag az alapsokaság megfelelő értékének torzítatlan becslése.

A fenti eljárás *J. Tukey* által bevezetett általánosítása a több véletlen kezdőpontból kiinduló szisztematikus mintavétel (systematic sampling with multiple random starts). A kiválasztás lebonyolítása eszerint úgy történik, hogy az első k elemből ismétlés nélküli egyszerű véletlen kiválasztással egy s nagyságú mintát veszünk, majd e minta elemeiből kiindulva, a sorba rendezett elemekből minden k -adikat választjuk ki. Az így kapott minta terjedelme $n \cdot s$ és ugyancsak torzítatlan becslést tesz lehetővé.

A szisztematikus mintavétel pontosságát eddig elsősorban az egyszerű véletlen kiválasztással és a rétegzett kiválasztással összehasonlítva vizsgálták. Ezzel kapcsolatban a cikk *W. G. Cochran* eredményeit ismerteti.

Szerző viszont a közönséges szisztematikus kiválasztás és az imént vázolt, több kezdőpontot felhasználó eljárás eredményességének viszonylagos mértékét vizsgálja. Azonos nagyságú, és pedig ns elemből álló mintát véve alapul, kiszámítja a

mintaátlag szórásnégyzetének várható értékét.

A kétféle mintavétel pontosságának aránya az alapsokaság elemei *sorrendjének jellegétől* függ. Ettől függően a következő eseteket különböztethetjük meg:

a) Az elemek sorrendje véletlenszerű. Ebben az esetben a szisztematikus mintavétel és a több kezdőponttól kiinduló mintavétel standard hibája egyenlő.

b) A sorrendben lineáris trend érvényesül. Ilyenkor a közönséges szisztematikus mintavétel jobb eredményt ad.

c) A sorrendben „szériális korreláció” érvényesül, vagyis bizonyos távolságra álló elemek — a vizsgált ismérv szempontjából — nem függetlenek egymástól. A kétféle mintavétel viszonylagos hatékonysága ez esetben a sorozatok között fennálló korreláció típusán és szorosságán múlik. Exponenciális típus esetén a szisztematikus mintavétel standard hibája kisebb, mint a több kezdőpontból kiinduló mintavételé. Ugyanakkor csak az utóbbi módszer használata révén lehet a sokaság sorrendjének jellegére való tekintet nélkül a mintából torzítatlan becslést adni a mintaátlag szórásnégyzetére.

Ennélfogva szisztematikus mintavétel alkalmazásánál a kétféle lehetőség közötti döntés előtt helves az elemek sorrendjének jellegét tanulmányozni.

(Ism.: *Párniczky Gábor*)

Jugenburg, Sz.:

A területi indexek kiszámításának néhány kérdése

(Nekotoriye voproszű iszcsiszlenija territorial' nüh indekszov.) — *Vesztnik Sztatisztiki*. 1958. No. 4. 56—65. p.

Az egyes rajonok (kerületek) gazdasági eredményeinek egybevetése gyakran szükségessé teszi az ún. statikus, vagyis egy meghatározott időpontbeli állapotot kifejező indexek alkalmazását. Az ilyen indexek kiszámításának módszertanával ugyanakkor eddig alig foglalkoztak: elterjedt az a nézet, hogy az indexek alkalmazásának területe csak a társadalmi ielenségek dinamikája lehet.

Igen fontos statikus indexek a területi indexek, melyekben valamely rajon adatait egy másik, bázisul vett rajon adataival való összehasonlítás alapján fejezzük

ki. Kiszámításuk gyakran nehézséggel jár, különösen az összefoglaló mutatószámok esetében, az összehasonlított sokaságok jelentős strukturális különbözősége miatt.

Ha a vetésterület összetétele szempontjából jelentősen eltérő rajonok gabona-termés-átlagait hasonlítjuk össze az indexben, az egyes kultúrák termésátlagainak különbözőségén kívül nyilván a vetésterületi struktúrák különbözősége is kifejeződik. Hogyan küszöbölhető ki a strukturális különbségek befolyása a területi index kiszámításakor?

Két út lehetséges. Kiszámítható az index a II. rajon termésátlaga és az I. rajon vetésterületi struktúrája, valamint az I. rajon termésátlaga és a II. rajon vetésterületi struktúrája alapján.

Általános nézet, hogy a munkatermelékenység, az önköltség, az ár, az átlagtermés stb. indexeit mindig azon objektum súlyai szerint kell mérlegelnünk, amelyekre a kiszámítandó index vonatkozik. Területi index esetén annak a rajonnak a vetésterületi struktúrája alapján, amelyeknek a termésátlagát egybevetjük valamely más, bázisul vett rajon átlagtermésével. Ha az I. rajon termésátlag indexének kiszámításáról van szó, akkor az I. rajon vetésterületi struktúrája és a II. rajon termésátlagai veendőek alapul, ha a II. rajonról, akkor megfordítva.

Ezt a megoldást azonban a szerző nem tartja kielégítőnek. Hasonló nehézségek mutatkoznak a termelés fizikai terjedelmének, a munkatermelékenység színvonalának stb. területi összehasonlítása esetén is, különösen ha nemcsak két vagy több rajon, hanem egyes országok összehasonlításáról van szó. Nemzetközi összehasonlítások esetén még nagyobbak lehetnek a struktúra-különbségekből eredő nehézségek. A struktúra-különbségek hatását a fenti módon kiküszöbölő módszert azért is helyteleníti, mert megfosztja az indexeket alapsajátosságuktól: az objektív viszonyok tükrözésétől.

Szerző véleménye szerint a területi indexek esetében a standardizált koefficiensek módszerét kellene alkalmazni. Tehát nem egyik vagy másik vizsgált rajon gazdasági struktúrája alapján kell súlyozni, hanem az összehasonlított rajonok vagy az egész ország gazdasági struktúrája alapján. Ez esetben a nevező és a

számláló különbsége megmutatná az elemzett tulajdonságnak a vizsgált tényezők egyikének megváltozása következtében beálló növekedését vagy csökkenését is. Ezután példákon mutatja be a standardizált koefficiensek módszerét a termelés volumene területi indexeinek kiszámításával kapcsolatban, két rajont hasonlítva össze.

A szerző úgy véli, hogy ipari területi indexek kiszámításakor az egész szovjet ipar struktúráját kell standardnak tekinteni, ami hosszabb-rövidebb ideig változatlannak is vehető.

(Ism.: *Valkovics Emil*)

Szameitat, Klaus — Koller, Siegfried:

A mintasokaság terjedelme és pontossága

(Über den Umfang und die Genauigkeit von Stichproben.) — *Wirtschaft und Statistik*. 1958. No. 1. 10—16. p.

A reprezentatív eljárás — klasszikus alkalmazási területeit (például minőségvizsgálat, termésbecslés, háztartásstatisztika stb.) túllépve — egyre gyakrabban kerül felhasználásra a hivatalos adatgyűjtések körében. A módszer alkalmazhatóságának feltételeire és korlátaira, valamint a módszerből adódó hibák becslésére vonatkozó ismeretekre egyre szélesebb körben van szükség.

A szerzők a módszerből adódó problémákat sokoldalúan világítják meg és jól használható nomogramot közölnek a relatív hiba becslésének leegyszerűsítésére, egyszerű véletlenszerű mintavétel esetén.

A felvétel gyorsasága, olcsósága, bizonyos mértékig maga a megbízhatóság is a minta terjedelmétől függ. Optimális az a minta, amely a lehető legkisebb költséggel és rövid feldolgozási idővel az eredményt megbízhatóan, a kívánt pontossággal adja meg. A minta terjedelme nincs eleve meghatározva, hanem több tényező figyelembevételével keressük meg a legjobb megoldást. Módszertanilag leghelyesebb először azt a pontosságot meghatározni, amellyel a táblákat kívánjuk elkészíteni, ezután megkeresni azt a mintavételi eljárást, amely ennek a követelménynek a legkisebb költséggel tesz eleget. A gyakorlatban fordított esetek is előfordulnak: eleve adott tényező a költség, amit ráfordíthatunk, vagy adva van