Gradus Vol 10, No 2 (2023)
ISSN 2064-8014

ÁTLAGOK A GYAKORLATBAN

AVERAGES IN PRACTICE

Ország Adrienn[0009-0000-7269-6287 1*], *Sugár András*[0009-0005-6390-9419 1]

[1] Alkalmazott Kvantitatív Módszertan Tanszék, Pénzügyi és Számviteli Kar, Budapesti Gazdasági Egyetem, Magyarország

https://doi.org/10.47833/2023.2.CSC.018

Kulcsszavak:
Átlagok fajtái
Átlagok gyakorlati alkalmazása
Súlyozás

Keywords:
Types of averages
Averages in practice
Weighting

Cikktörténet:
Beérkezett 2023. szeptember 1.
Átdolgozva 2023. november 3.
Elfogadva 2023. november 15.

Összefoglalás
A cikk röviden összefoglalja a gyakorlatban használt átlagokat és azok általánosításait, kitér a gyakorlati alkalmazások feltételeire, valamint egy újfajta súlyozási módot is bemutat.

Abstract
The article briefly summarizes the averages used in practice and their generalizations, discusses the conditions of practical applications, and also presents a new weighting method.

1. Bevezetés

Az átlagok a leggyakrabban használt statisztikai mutatók/módszerek közé tartoznak. A tankönyvek is eléggé hasonló módon, szerkezetben tárgyalják őket. Ebben a cikkben arra vállalkozunk, hogy röviden összefoglaljuk, mit is értünk átlagon, milyen fajtái vannak, melyek az átlagok tulajdonságai a szakirodalom (Pl. a gazdasági jellegű magyar felsőoktatás tananyagai, amelyek közül kiemeljük Köves-Párniczky (1981), Hunyadi-Vita (2003), Kerékgyártóné et al. (2009), Sándorné et al (2013) tankönyveket), illetve Statisztikai Szemle cikkek alapján. Másrészt néhány újszerű szemponttal bővítjük az átlagokról szóló diskurzust, főleg az átlagok fajtái, azok gyakorlati használata, valamint a leggyakrabban használt számtani átlag általánosításai, variációi kapcsán.

2. Az átlag fogalma, fajtái

Az átlag a középérték egy fajtája. A középérték az információsűrítés egy szélsőséges (de nagyon hasznos) eszköze, egyetlen értékben testesítjük meg, tömörítjük az információt, amit fontosnak tartunk. (Ennek természetesen az az ára, hogy sok információt, pl. a szóródásra, vagy az eloszlás alakjára vonatkozókat elveszítjük.) Az átlag fogalma, számítása szorosan kapcsolódik a valószínűségszámítás egyes területeihez is, de ezek tárgyalása meghaladja cikkünk lehetőségeit, pedig ez magyarázhatja azt is, hogy más matematikai területekhez képest az átlagok számítása történetileg nagyon későn került be a gazdasági/társadalmi gyakorlatba. L. erről Kendall (1956), Dusek (2022).

A középértékektől elvárjuk, hogy a) közepes helyzetűek, b) tipikusak, c) egyértelműen meghatározhatók, d) lehetőleg könnyen értelmezhetők legyenek. A gyakorlatban használt

* Kapcsolattartó szerző
E-mail cím: Orszag.Gaborne@uni-bge.hu

középértékek nem egyforma mértékben tesznek eleget ezeknek a követelményeknek, az átlagok az a) és c) követelményeknek minden esetben megfelelnek, b) és d) követelményeknek nem feltétlenül.

A középértékek két fajtája a helyzeti és a számított középérték. A helyzeti középértékek valamilyen jellegzetes helyzetüknél fogva (pl. szó szerint felezik a sokaságot, vagy tipikusak) jellemzik a sokaságot, a gyakorlatban is gyakran használjuk a mediánt és a móduszt, de ebben az áttekintésben a továbbiakban csak a másik középértéktípussal, a számított középértékekkel foglalkozunk. Mint az elnevezésben is benne van, a számított középértékek valamilyen számítás eredményei, ezeket a mutatókat szokás átlagoknak nevezni.

Az átlagok esetén a statisztika tankönyvek sok évtizede négyfajta átlagot szoktak ismertetni, a harmonikus, mértani, számtani és négyzetes átlagot. Az alábbi táblázatban összefoglaljuk a négyféle átlagra vonatkozó legfontosabb jellemzőket, alapvetően Köves (1961) alapján.* A képletekben az egyszerűség kedvéért elhagyjuk az összegzésnél és szorzatnál használatos futóindexek határait.

1. Táblázat. Átlagok fajtái és tulajdonságaik

Átlagfajta	*Súlyozatlan forma*	*Súlyozott forma*	*Visszavezetés a számtani átlagra*
Harmonikus ($\bar{X}_h$)	$\frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\dots+\frac{1}{x_n}}=\frac{n}{\sum\frac{1}{x_i}}$	$\frac{\sum f_j}{\sum\frac{f_j}{x_j}}$	$\frac{1}{\bar{X}_h}=\frac{\sum\frac{1}{x_i}}{n}$
Mértani ($\bar{X}_g$)	$\sqrt[n]{x_1\cdot x_2\cdot\dots\cdot x_n}=\sqrt[n]{\prod x_i}$	$\sqrt[\sum f_j]{\prod x_j^{f_j}}$	$\ln(\bar{X}_g)=\frac{\sum ln(x_i)}{n}$
Számtani ($\bar{X}_a$)	$\frac{x_1+x_2+\dots+x_n}{n}=\frac{\sum x_i}{n}$	$\frac{\sum f_j\cdot x_j}{\sum f_j}$	$\bar{X}_a=\frac{\sum x_i}{n}$
Négyzetes ($\bar{X}_q$)	$\sqrt{\frac{x_1^2+x_2^2+\dots+x_n^2}{n}}=\sqrt{\frac{\sum x_i^2}{n}}$	$\sqrt{\frac{\sum f_j\cdot x_j^2}{\sum f_j}}$	$\bar{X}_q^2=\frac{\sum x_i^2}{n}$

A leggyakrabban használt átlagfajta a számtani átlag, gyakran amikor átlagot mondunk, automatikusan a számtani átlagra utalunk. A számtani átlag definíció szerint az a szám, amit ha a számok helyébe írunk, akkor az összeg ugyanaz marad. Úgy is lehet interpretálni, ha az adatok összegét (az ún. értékösszeget) egyenletesen szétosztjuk, akkor egy megfigyelésre a számtani átlagnak megfelelő érték jut. Ez a megfogalmazás sugallja, hogy a számtani átlag akkor értelmes, ha az adatok logikailag összeadódnak, tehát értelmes az értékösszeg. Ez a legtöbb gyakorlati alkalmazás esetén igaz is (pl. átlagbérek esetén a bérek összege a kifizetett bértömeg, vagy egy termelékenységi mutatót említve az egy alkalmazott által előállított átlagos termékszám esetén az összes legyártott termék mennyisége), de nem mindig. Van, amikor ugyan összeadódnak logikailag az értékek, de a tényleges értékösszegnek csak erőltetett értelmet lehetne tulajdonítani (pl. az átlagos testmagasság vagy átlagos életkor esetén).

Több olyan általánosítása létezik a számtani átlagnak, amelyekbe a többi átlag is belefoglalható, ezek a gyakorlatban erőltetettek is lehetnek, de van, amikor további alkalmazásokra adnak alkalmat. Gyakori, hogy az ún. momentumot definiálják, illetve ennek további általánosításait, ezek közül a centrális momentum a legelterjedtebb. Az (1) képletek mutatják a sima és a centrális momentumok (súlyozatlan) számítását. A k. sima momentum (M_k) x_i k. hatványainak számtani átlaga, ahol k természetes szám. Ha k=1, akkor a számtani, átlagot kapjuk, ha k=2, akkor a négyzetes átlag négyzetét. A k. centrális momentum ($M_k(\bar{X})$) a számtani átlagtól vett eltérések k. hatványainak számtani átlaga. Ha k=1, akkor 0-t kapunk (hiszen az átlagtól vett eltérések kiegyenlítik egymást, összegük 0), a második centrális momentum a variancia, ami a szóródás mértékének egyik mutatószáma (értelmezni a gyökét, a szórást szoktuk). A centrális momentumok esetén k magasabb

* Köves Pál 60 évvel ezelőtti cikke a mai napig releváns, jó összefoglaló, de természetesen azóta is sok cikk született az átlagokról, azok fajtáiról, alkalmazásairól, egy viszonylag friss áttekintést lehet olvasni, pl. de Carvalho (2016)-ban.

értékeire számolt mutatókat is használjuk a gyakorlatban, k=3 esetén az aszimmetria irányát és intenzitását mérő mutatóhoz jutunk, k=4 esetén pedig a kurtózis (csúcsosság-lapultság) mértékét tudjuk számszerűsíteni. A statisztikai szoftverek által számolt aszimmetria- és kurtózis- mutatók szinte minden esetben a centrális momentumokon alapuló mutatók.

$$M_k = \frac{\sum x_i^k}{n} \qquad M_k(\bar{X}) = \frac{\sum (x_i - \bar{X})^k}{n} \tag{1}$$

Még egy formula (általánosítás) említése lehet hasznos (vannak továbbiak is, de – főleg a gazdasági/társadalomtudományi gyakorlatban – számunkra nem tűnnek fontosnak). A (2) képlet mutatja az ún. hatványközepet (H_α)), ami két területen is továbblép a momentumhoz képest.

$$H_\alpha = \left(\frac{\sum x_i^\alpha}{n}\right)^{\frac{1}{\alpha}}, \; x \epsilon \mathbb{R} \setminus \{0\}. \tag{2}$$

Egyrészt a k természetes számok mellett bármilyen valós (nem nulla) szám lehet a kitevőben, másrészt a hatványozás hatását a végén visszatranszformáljuk. Az 1. táblázatban felsorolt – a gyakorlatban használt 4 átlag felfogható a hatványközép 4 speciális eseteként, α=1 adja a számtani, α=2 a négyzetes átlagot, α=-1 esetén kapjuk a harmonikus átlagot, és a mértani átlag formulája is megkapható, ha képezzük a $\lim_{\alpha \to 0} H_\alpha$ értéket. Ezek közül egyedül az utolsó nem látszik ránézésre, de annak bizonyítása is viszonylag egyszerű, l. pl. Sykora (2009).

Itt érdemes kiemelni, milyen gyakorlati esetben használjuk igazán a harmonikus átlagot (ezt a területet szokták a rész- és összetett viszonyszámok vagy rész- és főátlagok közötti összefüggésnek nevezni), amennyiben egy V=A/B viszonyszámot részsokaságokra és a sokaság egészére is számolunk, akkor az ún. aggregát forma mellett az összetett viszonyszám a részviszonyszámok B_j-kkel súlyozott számtani vagy A_j-kkel súlyozott harmonikus átlagaként adódik. Ezt mutatják a (3) képletei. Ennek speciális esete a rész- és főátlagok közötti összefüggés, amit a (4) képletekben lehet nyomon követni.

$$V_j = \frac{A_j}{B_j} \qquad \bar{V} = \frac{\sum A_j}{\sum B_j} = \frac{\sum B_j \cdot V_j}{\sum B_j} = \frac{\sum A_j}{\sum \frac{A_j}{V_j}} \tag{3}$$

$$A_j = N_j \cdot \bar{X}_j = S_j \qquad B_j = N_j \qquad V_j = \bar{X}_j \qquad \bar{X} = \frac{\sum S_j}{\sum N_j} = \frac{\sum N_j \cdot \bar{X}_j}{\sum N_j} = \frac{\sum S_j}{\sum \frac{S_j}{\bar{X}_j}} \tag{4}$$

ahol S_j a j. részsokasághoz tartozó részletösszeg, N_j j. részsokaság mérete, $\bar{X}_j$ a j. részsokasághoz tartozó átlag. Egy technikai (de a gyakorlatban fontos) elem, hogy a súlyozáshoz elég a gyakoriságok vagy értékösszegek megoszlása is, ahogy az (5) képletek mutatják.

$$g_j = \frac{N_j}{\sum N_j} \qquad Z_j = \frac{S_j}{\sum S_j} \qquad \bar{X} = \sum g_j \cdot \bar{X}_j = \frac{1}{\sum \frac{Z_j}{\bar{X}_j}} \tag{5}$$

Végezetül érdemes iderakni a középértékek között nagyságviszonyt is, amennyiben a megfigyelések közül legalább 2 különbözik egymástól, akkor teljesül a (6) összefüggés.

$$x_{min} < \bar{X}_h < \bar{X}_g < \bar{X}_a < \bar{X}_q < x_{max} \tag{6}$$

Az átlagok közötti választás kérdésében érdemes a gyakorlati szempontokat figyelembe venni. Vegyük végig, melyik átlagfajta használata milyen körülmények esetén merül fel a gyakorlatban.

- A leggyakoribb számtani átlag használatáról már szó esett, minden olyan esetben, amikor az adatok tartalmilag összeadódnak, ennek az átlagnak a használata a logikus. Egy szempontra érdemes felhívni a figyelmet, a számtani átlag használata esetén legalább különbségi szintű mérést feltételezünk (a mértani és harmonikus átlag csak arányskálán értelmes), ugyanakkor a gyakorlatban számos esetben számítanak átlagokat sorrendi skálák esetében is (Likert-skálák értékeinek átlaga, érdemjegyek átlaga stb.) Ezek a számítások ugyan nem megfelelőek, de kiirthatatlanok, csak legyünk tisztában a hátulütőivel.

- A négyzetes átlag gyakorlati helye viszonylag egyértelmű. Ezt az átlagfajtát általában akkor használjuk, ha el akarunk tekinteni az adatok előjelétől, mert pl. azok összege nulla. Ez a helyzet a szóródás számszerűsítése esetén, amikor az átlagtól vett eltérések átlagát számoljuk (ezek számtani átlaga nulla), vagy az olyan modelleknél, ahol a hiba átlagos nagysága a kérdés, de a hibák kiegyenlítik egymást, az összegük szintén nulla. Az előjeltől eltekintés két módja az abszolút érték és a négyzetre emelés, mindkét logikára épülnek mutatószámok, de főleg következtető statisztikai megfontolások miatt az abszolút értékre alapozott mutatók kiszorultak a gyakorlatból. A két leggyakrabban használt négyzetes átlag forma a szórás (átlagtól vett eltérések négyzetes átlaga), ami a szóródás legfontosabb mutatója, és a reziduális szórás.
- A mértani átlag logikája mutatja, hogy akkor használatos, ha az adatok logikailag nem összeadódnak, hanem összeszorzódnak, az adatok szorzatának van tárgyi (gazdasági vagy egyéb) értelme. Éppen ezért a leggyakoribb felhasználása idősorok esetén a láncviszonyszámok átlagos értékének számolása (átlagos növekedési ütem). Több tankönyv kizárólag ezt a példát hozza az alkalmazásra. Azért használata máshol is megjelenik, főleg az indexszámítás területén (egyedi indexeknek esetleg nem számtani, hanem mértani átlagát veszik, erről l. részletesebben Köves (1957), vagy keresztezett indexformulák esetén használják (l. pl. Törnquist-index, vagy a legismertebb Fisher-index).
- A legellentmondásosabb átlagforma a harmonikus átlag. Definíciószerűen a harmonikus átlag akkor értelmes, ha az adatok reciprokösszegének van tárgyi értelme. A harmonikus átlag reciprokát írva az adatok helyébe, így lesz változatlan az eredeti adatok reciprok összege. Egyes esetekben említik, hogy fordított intenzitási viszonyszámok esetén értelmes. Egyedi adatok esetén nehézkes jó példát adni ennek használatára, több tankönyv meg is jegyzi ezt. Ugyanakkor láttuk, hogy a harmonikus átlag használata mennyire gyakran és logikusan adódik, amikor részviszonyszámok (vagy részátlagok) számlálóval súlyozott harmonikus átlagformáját kell használnunk az összetett viszonyszám (főátlag) számításához. Ez a harmonikus átlag igazi terepe. Érdemes megnézni 2 olyan példát, amit statisztika tankönyvek előszeretettel mutatnak a harmonikus átlag súlyozatlan számítására, hogy lássuk ennek a kérdéskörnek a problémáit.

Számos esetben a pénz vásárlóerejének (egységnyi pénzért kapható árumennyiség) példáját említik. Ebben az esetben a vásárlóerő a forint/darab egységár reciproka, közvetlenül nem összegezhető. Legyen számszerű példaként az alábbi: egy piacon 100 forintért egy helyen 1, egy másik helyen 2, a harmadik helyen 3 tojás vehető. Mekkora a piacon 100 forint vásárlóértéke tojásban? (Átlagosan a 3 standon hány tojást lehet venni 100 forintért?) A számtani átlagforma ebben az esetben hibás eredményre vezetne, hiszen a 3 standon nem ugyanannyiért kapjuk a tojás darabját. A helyes számolás:

$$\frac{100+100+100}{\frac{100}{1}+\frac{100}{2}+\frac{100}{3}}=\frac{3}{\frac{1}{1}+\frac{1}{2}+\frac{2}{3}}=\frac{3}{1+0{,}5+0{,}33}=1{,}64\,\frac{db}{100Ft}$$

azaz 100 forintért átlagosan 1,64 tojást lehet venni a 3 standon (és nem 2 darabot, amit a súlyozatlan számtani átlagforma eredményezne.

Ebben az esetben a harmonikus átlag azért írható fel súlyozatlan (pontosabban azonos súlyú) átlag formájában, mert eleve 100-100-100 forint vásárlóerejét számoltuk. De természetesen előfordulhat, hogy az egyik standon 100, a másikon 200, a harmadikon 400 forintért vennénk tojást, és akkor a forma már súlyozott. De az előző számítás úgy is felírható, hogy az egyes standokon 100 forintért vehető tojások számát (tehát a vásárlóerő reciprokát, esetünkben az egyenes mutatót) átlagoljuk számtani átlag formában, de ebben az esetben már súlyoznunk kell. A 3 standon 100 forintért rendre 1, 2 és 3 tojás vehető (db/100 forint), ezért az ehhez tartozó árak különböznek egymástól, rendre 100, 50 és 33,33 Ft/db. A számtani átlagformában való számolás:

$$\frac{100\cdot 1+50\cdot 2+33{,}33\cdot 3}{100+50+33{,}33}=1{,}64\,\frac{db}{100Ft}$$

Tartalmilag mindkét esetben ugyanazt számoljuk, a 183,33 forintból, amit a 3 standon elköltünk, hány tojást lehet megvenni. A harmonikus átlagforma azért tűnik súlyozatlannak, mert mind a 3 standon azonos költési összeget feltételezünk.

Talán még tisztábban látszik a fenti probléma a másik gyakori példán, amikor átlagsebességet számolunk. Amennyiben súlyozatlan harmonikus átlagformára szeretnénk példát mutatni, egyenlő hosszúságú utakat tételezünk fel. Legyen egy egyszerű példa, egy kerékpáros felteker egy hegyre 4 km/h átlagsebességgel, majd lejön 16 km/h átlagsebességgel. Ebben az esetben az átlagsebesség a két szakaszon mért átlagsebesség súlyozatlan harmonikus átlagaként is számolható, de csak azért lehet súlyozatlan formát használni, mert a két megtett út (fel a hegyre, illetve le a hegyről) nyilvánvalóan azonos (S) hosszúságú.

$$\bar{X}_h = \frac{S+S}{\frac{S}{4}+\frac{S}{16}} = \frac{2}{\frac{1}{4}+\frac{1}{16}} = 6{,}4\frac{km}{h}$$

Ugyanez az eredmény adódik számtani átlag formában is, ebben az esetben azonban nem egyforma a súly, hiszen lefelé 4-szer olyan gyorsan megy a kerékpáros, mint felfelé, azaz negyed annyi idő alatt ér le, mint fel, tehát az összes kerékpározásra fordított idő 20-80%-os arányban oszlik meg.

$$\bar{X}_a = 0{,}2 \cdot 16 + 0{,}8 \cdot 4 = 6{,}4\frac{km}{h}$$

A harmonikus átlag ezek alapján a gyakorlatban súlyozott formában jelenik meg, és tipikusan akkor használjuk, ha részviszonyszámok (vagy részátlagok) számláló megoszlással súlyozott harmonikus átlagformáját számoljuk. Olyan speciális esetben, amikor a számláló egyenletesen oszlik meg a részek között, ez felfogható súlyozatlan átlagnak.

3. A számtani átlag és gyakorlati variációi

A leggyakrabban tehát a számtani átlagot használjuk. Érdemes a számtani átlag tulajdonságait felsorolni:

- A számtani átlag (mint minden átlag) minden megfigyelési értéktől függ, egyértelműen meghatározható.
- Az átlag érzékeny a kilógó (outlier) értékekre, a mutató nem robusztus (nem úgy, mint a helyzeti középértékek).
- Amennyiben a megfigyelt értékeken ugyanazon lineáris transzformációt hajtjuk végre, a transzformált értékek átlaga is örökli ezt, azaz $\overline{A + B \cdot x_i} = A + B \cdot \bar{X}$
- A számtani átlag 3 szempontból is közepes helyzetű, a) a legkisebb és legnagyobb megfigyelt érték közé esik, b) a tőle vett eltérések kiegyenlítik egymást, c) minimálja az elemek négyzetes eltérésösszegét. Formálisan a (7) képletek mutatják az összefüggéseket.

$$a)\ x_{min} < \bar{X} < x_{max}\ \ b)\ \sum(x_i - \bar{X}) = 0\ \ c)\ \sum(x_i - A)^2 \to \min\ esetén\ A = \bar{X} \quad (7)$$

Többek között egy-két fenti tulajdonságra is reagálnak a számtani átlag egyes variációi, ezek közül 4-et sorolunk itt fel, és közülük a 4. variációt újszerűsége miatt alaposabban is megnézzük.

1. Nyesett átlag ($\bar{X}_\beta$): a nyesett átlag éppen arra reagál, hogy az outlierek befolyásolhatják az átlag értékét, elhúzhatják azt lefelé vagy felfelé. Ezért egy variáció, hogy elhagyjuk a legkisebb és legnagyobb β%-át a megfigyeléseknek, és ezek nélkül számítunk számtani átlagot.

2. Kronologikus átlag ($\bar{X}_k$): Ha egy időközönként megfigyelt állapotidősor átlagos szintje a kérdés, pl. készletek nagysága, valamilyen népességszám, stb., akkor az átlagos szint jellemzése kronologikus átlaggal történik ($\bar{X}_k = \left(\frac{1}{2} \cdot x_0 + x_1 + \cdots + x_{n-1} + \frac{1}{2} \cdot x_n\right)/n$). A kronologikus átlag tulajdonképpen az átmenet az álló jelenségről az időbelire (mozgóra).
3. Mozgóátlag $\left(\hat{X}_t\right)$: használatára akkor kerül sor, ha egy idősorban van tendencia, és ezt a tendenciát (trendet) átlagolással mutatjuk be. Ilyenkor az átlag értékét az idősor közepéhez igazítjuk. Ha páros a tagszám, akkor formailag egy kronologikus átlagot számolunk, de középre (konkrét időszakhoz) igazítva. A mozgóátlag egyszerűbb (súlyozatlan) esetei mellett számos súlyozott formát is használnak.
4. Végül egy olyan súlyozott számtani átlagfajtát mutatunk, ami nem magától értetődő, de éppen gyakorlati szempontból figyelemreméltó. Induljunk ki a részviszonyszámok (vagy részátlagok) súlyozott számtani átlag formájából. Mint ismeretes (és szó volt róla itt az előbbiekben is) a $\bar{V} = \frac{\sum B_j \cdot V_j}{\sum B_j}$ összetett viszonyszám a részviszonyszámok nevezőben levő értékekkel súlyozott számtani átlaga. Felmerül, hogy számoljuk a részviszonyszámok számlálóval súlyozott számtani átlagát, azaz marad az átlagforma típusa, csak más a súly. A számolás módja tehát: $\bar{V}_é = \frac{\sum A_j \cdot V_j}{\sum A_j}$. A $\bar{V}_é$ jelölés az érzékelt átlag elnevezésre utal, az érzékelt jelző majd világos lesz az alkalmazási példa ismertetése után.

Ez első ránézésre szokatlan, annyira elterjedt, hogy vagy a nevezővel súlyozott számtani, vagy a számlálóval súlyozott harmonikus átlagot számoljuk. Utánanéztünk, van-e előzménye ennek az átlagfajtának, de egyetlen cikket találtunk, ahol ez szerepel, Ertl (1983) leírását. Az alapötlet és a példa, amivel illusztráljuk a problémát, a cikkből származik, a szokásos számtani átlaggal való összehasonlítás, és az eltérés okainak elemzése a mi hozzájárulásunk a témához. Az érzékelt átlag (számlálóval való súlyozás) akkor merülhet fel, ha más az, ami egy jelenségből ténylegesen megvalósul, és amit a felhasználó ebből észlel/érzékel. (Az érzékelhető jelenség szerepelhet a számlálóban.) A legplasztikusabb példa erre a járatok utaskihasználtsága lehet (magáról a témáról összefoglalót ad pl. Kocsis-Nagy (2023)), de sok egyéb példa is hozható (népsűrűség, egy oktatóra jutó tanuló, stb.). Ha vonatok, tömegközlekedési eszközök esetén számszerűsítjük a járatok kihasználtságát (a viszonyszám alapvetően a tényleges utasok száma/kapacitás), akkor „belefuthatunk" abba a jelenségbe, hogy átlagosan nem túl magas átlagos kihasználtság mellett egyes utasok egész mást érzékelhetnek. Emögött a kihasználtság szóródása áll, minél nagyobb a szóródás mértéke, annál inkább lesznek tömött járatok (ahol sokan érzékelnek zsúfoltságot), és kevesen laza üzemmenetet, hiszen olyankor kevesen ülnek a járművön. Ilyen esetekben pl. az érzékelt átlagnak lehet létjogosultsága. Szélsőséges esetben legyen egy oda-vissza járat 100 fős kapacitással, ahol odafelé tele van a járat, visszafelé senki nem utazik. Az átlagos kihasználtság a hagyományos súlyozással (100*1+100*0)/200=0,5, azaz 50%. Ugyanakkor ebben a konkrét szélsőséges esetben az érzékelt zsúfoltság 100%-os lesz, hiszen ezt odafelé mindenki így érzi, visszafelé pedig nincs érzékelő személy, mert senki nem utazik a járaton. Az érzékelt átlag ezt jelzi is, (100*1+0*0)/100=1, azaz 100%.

Vezessük le, mi az összefüggés a hagyományos átlag és az érzékelt átlag között. Az egyszerűség kedvéért feltételezzük, hogy a nevezőben levő értékek megegyeznek egymással (példánkban minden járat azonos kapacitású, legyen ez B). Az eredeti átlag a (8) alatt látható módon írható át (n a megfigyelések száma)

$$\bar{V} = \frac{\sum B \cdot V_j}{\sum B} = \frac{\sum A_j}{n \cdot B} = \frac{\bar{A}}{B} \qquad (8)$$

Azt kaptuk tehát logikus módon, hogy az eredeti átlag (példánkban az átlagos kapacitáskihasználtság) a megfigyelésenkénti átlagos (utas)szám és a konstans kapacitás hányadosa. Az érzékelt átlagot alakítsuk át olyan módon, hogy lássuk, hogyan függ az értéke a tényleges átlagtól. (S az A_j-k összege, azaz az értékösszeg, a relatív értékek négyzetösszege a Herfindahl-index (HI), ami a koncentráció fokát méri, V az Aj-k relatív szórása).

$$\bar{V}_{é} = \frac{\sum A_j^2/B}{\sum A_j} = \frac{S}{B} \cdot \sum Z_j^2 = \frac{n \cdot \bar{A}}{B} \cdot HI = \bar{V} \cdot n \cdot HI = \bar{V} \cdot (V^2 + 1) \qquad (9)$$

A képlet mutatja, hogy egyrészt az érzékelt átlag nagyobb vagy egyenlő, mint a tényleges, és egyenlőség akkor és csak akkor áll fenn, ha az A_j adatok szórása 0, azaz minden adat megegyezik egymással (azaz minden járaton ugyanannyian utaznak), másrészt az is látszik, annál inkább meghaladja az érzékelt átlag a ténylegest, minél erősebb a relatív koncentráció mértéke.

4. Összefoglalás

Cikkünkben a tankönyvekben használt átlagfajtákat tekintettük át. Bemutattuk, hogy elméletileg többféle módon is lehet az átlagokat általánosítani, de a gyakorlatban a számtani átlag mellett a többi átlagot csak speciális esetekben használjuk. Különösen a harmonikus átlag használata esetén jelenik meg a gyakorlati alkalmazás szempontjának elsődlegessége az elméleti rendszerezéssel szemben. A leggyakrabban használt számtani átlag tulajdonságait részletesebben is szemügyre vettük. Bemutattunk egy ritkán használt, de gyakorlati szempontból hasznos átlagfajtát, az ún. érzékelt átlagot, valamint viszonyát a hagyományos számtani átlaghoz.

Irodalomjegyzék

[1] de Carvalho, M.: "Mean, what do you Mean?", The American Statistician, 2016, 70 (3): 764–776. DOI: 10.1080/00031305.2016.1148632

[2] Dusek, T.: A valószínűségszámítás ókori és középkori hiányának okai, Statisztikai Szemle, 2022, 100. évf. 4. sz. pp. 408–420. DOI: 10.20311/stat2022.4.hu0408

[3] Ertl, I.: Az intenzitási viszonyszámokkal mért átlagokról, Statisztikai Szemle, 1983, 61. évf.3. sz. pp.280-293

[4] Hunyadi, L.-Vita, L.: Statisztika közgazdászoknak, 2003, KSH Budapest

[5] Kendall, M. G.: Studies in the history of probability and statistics: II. The beginnings of a probability calculus. Biometrika, 1956, Vol. 43. Nos. 1–2. pp. 1–14. DOI: 10.1093/biomet/43.1-2.1

[6] Kerékgyártó, Gy. - L. Balogh, I.- Sugár, A. – Szarvas, B.: 2009 Statisztikai módszerek a gazdasági és társadalmi elemzésekben. 2009, Aula kiadó Budapest

[7] Kocsis, T.– Nagy, D.: Hálózati hányadosok Budapest kerületeinek közösségi közlekedéssel feltárt belső hálózataira, Területi Statisztika, 2023, 63(4): 99.466–491; DOI: 10.15196/TS630403

[8] Köves, P.: A mértani átlag statisztikai alkalmazásai, Statisztikai Szemle 1957, 35. évf. 4-5. sz.pp. 303–332

[9] Köves, P.: Az átlagszámítás helye a statisztika elméletében Statisztikai Szemle, 1961, 39. évf. 1. sz. pp. 3-30

[10] Köves, P.- Párniczky G.: Általános statisztika I., 1981. KJK Budapest

[11] Sýkora, SS.: Mathematical means and averages: basic properties , 2009, Vol. 3. Stan's Library: Castano Primo

[12] Sándorné Kriszt É., – Csesznák A.,- Ország G.: Statisztika I., 2013. Tankönyvkiadó Budapest