„Tudásfejlesztés és – hasznosítás a Nyíregyházi Egyetemen” **- EFOP-3.4.3-16-2016-00018**
A.4. Kárpát-medencei oktatási tér kialakítása

STATISZTIKA I.

A tananyag elkészítését a **„Tudásfejlesztés és – hasznosítás a Nyíregyházi Egyetemen”** című, **EFOP-3.4.3-16-2016-00018** számú projekt támogatta. A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.

Szerző: Makszim Györgyné dr. Nagy Tímea

Lektor: Vargáné dr. Bosnyák Ildikó

Kézirat lezárva: 2018. december

ISBN:

TARTALOMJEGYZÉK

ELŐSZÓ

Napjaink felgyorsult világában az egyik legfontosabb erőforrás az információ. Az információ előállításában, megszerzésében, feldolgozásában és elemzésében kitüntetett szerepe van a statisztikának. Éppen ezért a Statisztika I. jegyzet célja módszertani ismereteket nyújtani a társadalmi-gazdasági jelenségek és folyamatok megfigyeléséhez, számszaki összehasonlításához, összefüggések felismeréséhez és feltárásához, következtetések levonásához. Középpontba helyezni a vállalati igényeknek megfelelő, gyakorlatorientált feladatok, problémák megoldását, továbbá a lényeglátó gondolkodásmód fejlesztését.

E jegyzet elsősorban a jövő szakembereit jelentő egyetemi hallgatóknak készült, akik a különböző képzési területeken találkoznak a statisztikával, mint tantárggyal.

A legfontosabb feladat, olyan szemlélet kialakítása a hallgatóban, hogy képes legyen összetett feladatok rugalmas megoldására, önálló adatgyűjtésekre, adatfeldolgozásra, illetve komplex statisztikai elemzésekre. Az elemzések eredményei alapján tudjon felelős következtetéseket levonni és javaslatokat tenni a problémák megoldására.

A Statisztika című tantárgy oktatásának lényege a gyakorlati elemzésekhez szükséges legfontosabb módszerek elsajátíttatása.

Ezen módszerek két részre oszthatók:

1) a leíró statisztika módszerei,
2) a következtető statisztika módszerei.

Ebben a jegyzetben, a továbbiakban a leíró statisztika módszereinek bemutatására kerül sor. Ide tartoznak a :

- viszonyszámok,
- középértékek,
- az adatok változékonyságának vizsgálati módszerei,
- a gyakorisági sorok empirikus eloszlás vizsgálata,

- a koncentráció elemzése,
- az értékindexkör indexei, továbbá
- a főátlagindexkör indexei.

A következtető statisztika módszertani bemutatására a Statisztika II. jegyzetben kerül sor.

Kívánom, hogy ezt a törzsanyagot forgassa hasznosan, és segítse Önt a statisztikai szempontú problémák megoldásában.

A Szerző

1 Statisztikai alapfogalmak

Jelen fejezetben bevezető jelleggel betekintést szeretnék nyújtani a statisztika alapfogalmi rendszerébe, a magyar hivatalos statisztikai szolgálatba, továbbá ismertetni szeretném a statisztikai információs rendszer alapjait, a statisztikai munka folyamatát, valamint a mérési skálák típusait.

1.1 A statisztika fogalma

Statisztikán három különböző fogalmat szokás érteni. Statisztikának nevezik ugyanis:

- azt a **gyakorlati tevékenységet**, amelynek keretében statisztikai adatokhoz jutunk,
- e tevékenység eredményeképpen kapott **adatok összességét**, végül
- statisztikán egy bizonyos **módszertant** is értenek.

A statisztika tehát egy olyan tudományos módszertan, illetve gyakorlati tevékenység, amely a tömegesen előforduló jelenségek egyedeire vonatkozó információk gyűjtésével, feldolgozásával, elemzésével, valamint a vizsgált jelenség egészének tömör, számszerű jellemzésével foglalkozik.

A statisztika célja:

- új információ, új ismeret nyújtása,
- határozatlanság, bizonytalanság csökkentése,
- optimális döntési helyzetek felismerése,
- több lehetőség közötti választás elősegítése.

Vizsgálatának tárgya: a társadalomi, gazdasági élet tömegesen (nagy számban) előforduló jelenségei.

A statisztika helye a tudományok rendszerében:

A statisztika a tudományágak rendszerében a társadalomtudományok közé tartozik, ezen belül a **közgazdaságtudományok egyik ága**.

A statisztika nem önálló tudomány, hanem **módszertudomány**, vagy ún. **tudományos módszertan**, mivel a statisztika módszereit szinte minden tudományterületen alkalmazzák. Miközben a statisztika más tudományok módszeréül szolgál, maga is sokrétű kapcsolatban áll a szaktudományokkal (matematika, könyvvitel, vállalati gazdaságtan stb.), sőt a különböző tudományterületek kapcsolódó területein ún. interdiszciplináris tudományterületek kialakulásához (pl. biometria, ökonometria stb.) vezetett.

A statisztika, mint módszertan részei:

- **Általános statisztika:** általános elméleti kérdésekkel, elméletekkel foglalkozik.
- **Szakstatisztika:** a társadalmi, gazdasági élet egy adott területének vizsgálatát tekinti céljának. (A szakstatisztika ugyanakkor az általános statisztika által kidolgozott eredményeket alkalmazza!)

A statisztikai tevékenységnek többféle ágát különböztetjük meg:

- Népességstatisztika;
- Gazdaságstatisztika;
- Ágazati statisztikák;
- Vállalati, üzemi statisztika;
- Társadalomstatisztika.

1.2 A statisztikai szolgálat jogi keretei

Az egyes országokban, így Magyarországon is működik hivatalos statisztikai szolgálat, amelynek tevékenységét általában törvény szabályozza. A jelenleg hatályos törvény a 2016. évi CLV. törvény a hivatalos statisztikáról. E törvény többek között meghatározza azokat a szerveket, amelyek a hivatalos statisztikai szolgálatot alkotják. A Hivatalos Statisztikai Szolgálat tagja az a szervezet lehet, amely közfeladatának részeként hoz nyilvánosságra hivatalos statisztikai adatokat, és amelynek szervezete és működése megfelel a Nemzeti Statisztika Gyakorlati Kódexében és az Európai Statisztikai Rendeletben foglaltaknak. Ez alapján a statisztikai szolgálat tagja:

- Központi Statisztikai Hivatal (KSH): A KSH elnöke látja el Magyarország főstatisztikusi feladatait.
- Magyar Nemzeti Bank (MNB);
- Minisztériumok;
- Országos Bírósági Hivatal (OBH);
- Az igazságszolgáltatás legfőbb szervei;
- Gazdasági Versenyhivatal;
- Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal;
- Egyéb szervek.

A statisztikáról szóló törvényen kívül persze számos egyéb jogforrás is hatással van a statisztika tevékenységi rendszerére, melyek közül fontos kiemelni az információs önrendelkezési jogról és az információszabadságról szóló 2011. évi CXII. törvényt.

A nemzetközi statisztikai tevékenységet az ENSZ Statisztikai Bizottsága koordinálja. Magát az ENSZ-ben folyó statisztikai tevékenységet a Titkárság Statisztikai Osztálya végzi. Emellett az ENSZ-nek vannak regionális bizottságai (pl. a Genfben székelő Európai Gazdasági Bizottság – ECE), valamint szakosított szervei (pl. a Nemzetközi Munkaügyi Szervezet (ILO), az ENSZ Nevelésügyi Tudományos és Kulturális Szervezete (UNESCO), az Egészségügyi Világszervezet (WHO), a Kereskedelmi Világszervezet (WTO) stb).

Az Európai Unió is rendelkezik statisztikai szolgálattal, mely nem más, mint a luxemburgi székhelyű EUROSTAT, amelynek tevékenysége az ENSZ statisztikájával harmonizál.

A statisztikai információk közül a legfontosabbakat Magyarországon az 1884 óta évenként megjelenő Statisztikai Évkönyv teszi közzé. Ugyancsak évenként jelenik meg a Magyar Statisztikai Zsebkönyv, havonta a Statisztikai Havi Közlemények, és rendszeresen vagy időszakonként a különböző gazdasági és társadalmi jelenségeket és folyamatokat ismertető szakmai, valamint területi évkönyvek és időszaki kiadványok. Pl.: Mezőgazdasági Évkönyv, Beruházási Évkönyv, stb. Mindezek mellett igen kiterjed statisztikai információk állnak rendelkezésünkre a KSH honlapján is.

1.3 A statisztika információs rendszere

A **statisztikai információs rendszer** a statisztikai adatok, mutatószámok egymással kapcsolatban álló, logikailag összefüggő rendszere.

A statisztikai információs rendszer azonban nemcsak **tartalmi oldalról** közelíthető meg, hanem **működési** és **szervezeti** oldalról is. A statisztikai információs rendszer működési oldalát azok a műveletek és eljárások jelentik, melyek a rendszert alkotó statisztikai információk keletkezését, regisztrálását, tárolását, szükség szerinti átalakítását és mozgatását biztosítják. A szervezeti oldal pedig nem más, mint a működést lehetővé tevő szervezeti egységek, illetve intézmények összessége, valamint azok egymás közötti kapcsolatai.

A statisztikai információs rendszer tárgyát a társadalmi-, gazdasági állapotok és folyamatok alkotják. Az információs rendszer elsősorban a döntéshozók érdekeit szolgálja, emellett a lakosság tájékoztatását, valamint a tudományos célú felhasználást.

A magyar statisztikai információs rendszerben vannak *központi* és *területi* rendszerek, amelyek együtt egységes integrált rendszert alkotnak.

A statisztikai információk feldolgozásának fázisai a következők:

- statisztikai információs rendszer tervezése,
- metaadatok kezelése,
- regiszterek karbantartása,
- adatgyűjtés szervezése,
- adatok előkészítése,
- adatok feldolgozásának módja és rendszere,
- adatbázis-lekérdezés módja,
- tájékoztatás és elemzés.

A statisztikai információs rendszer kialakításakor lényeges a statisztikai adatok iránti igények megismerése és időben történő elemzése. Ezt csakis a felhasználókkal való rendszeres kapcsolat alapján lehet megvalósítani. Éppen ezért célszerű előzetesen informálódni a felhasználók igényeiről, illetve meghatározott időszakonként értékelést kérni a felhasználóktól.

Az adattartalom meghatározása is igen fontos, mert ez nemcsak a megfigyelni kívánt mutatószámokat és a megfigyelés egységeit tartalmazza, hanem a felvétel keretrendszerét, a megfigyelés gyakoriságát, az alapmutatókból képzett mutatók körét és tartalmát, valamint a megfigyelés módszerét is.

A megfigyelés módszere lehet:

- teljes körű: a sokaság minden elemére kiterjed;
- részleges: csak bizonyos körre vonatkozik.

Minden adatgyűjtés célja a tájékoztatás és az elemzés, ezért az adatgyűjtések, feldolgozások megszervezését megelőzi a kiadványok tervezése. A kiadvány természetesen nemcsak a hagyományos nyomtatott formát jelenti, hanem a CD-n, interneten való közlés módját is.

A hazai jogrend szempontjából különösen fontos az *adatvédelem* és az *adatok hozzáférhetősége*. Az ezekkel kapcsolatos jogok védelmének és elősegítésének legfontosabb szerve a Nemzteti Adatvédelmi és Információszabadság Hatóság (NAIH).

1.4 Alapfogalmak

A statisztika szűk értelembe vette alapfogalmi bázisát a SOKSÁG-EGYED-ISMÉRV, illetve az ADAT-MUTATÓSZÁM-STATISZTIKAI MODELL fogalmai képezik.

STATISZTIKAI SOKASÁG: a statisztikai megfigyelés tárgyát képező egyedek összessége, halmaza. Sokaságot mindig a Mi? Hol? Mikor? kérdésekre történő válaszadással adunk meg.

Pl.: A Nyíregyházi Egyetem hallgatóinak száma 2018. szeptember 1-jén.

A sokaság tipizálása:

a) Az egységek számossága alapján a sokaság lehet:

- **véges:** pl. a népesség adott időben, térben.
- **végtelen:** azonos körülmények közt tetszőlegesen sokszor megismételhető kísérlet eredményei.

b) A múló időhöz való viszonyulás alapján a sokaság lehet:

- **álló:** állapot, időpont (stock) jellegű, pl. a MÁV személyszállító vonatainak száma 2018. január 1-jén.
- **mozgó:** folyamat, időtartam (flow) jellegű, pl. a MÁV által 2018 januárjában szállított utasok száma.

c) A vizsgálati halmaz nagysága alapján a sokaság lehet:

- **fő:** a Nyíregyházi Egyetem hallgatóinak száma 2018. szeptember 1-jén.
- **rész:** a Nyíregyházi Egyetem Gazdálkodástudományi Intézetének hallgatói létszáma 2018. szeptember 1-jén.

A vizsgálati halmaz nagysága alapján létezik egy másik csoportosítás is:

- **teljes:** a Nyíregyházi Egyetem hallgatóinak száma 2018. szeptember 1-jén.
- **minta:** a Nyíregyházi Egyetem egyszerű véletlen mintavétellel kiválasztott 500 hallgatója 2018. szeptember 1-jén.

d) A létező világhoz való viszonyulás alapján a sokaság lehet:

- **valóságos:** létező elemekből áll.
- **elméleti:** csak elképzelt, fiktív elemekből áll.

EGYED vagy **EGYSÉG**: a sokaságot alkotó egyedeket a sokaság egységeinek nevezzük.

Pl.: A Nyíregyházi Egyetem egy konkrét hallgatója.

Az egység tipizálása:

Az egység lehet:

- valóságos: ilyenek az álló sokaság egyedei.
- esemény: ilynek a mozgó sokaság egyedei.

ISMÉRV: azokat a kritériumokat, amelyek szerint a sokaság egységeit jellemezzük, ismérveknek nevezzük.

Ismérvváltozat: egy adott ismérv szerint a sokaság egységei többféle tulajdonsággal rendelkeznek. Ezek a tulajdonságok az ismérv változatai, variációi.

Pl.: ***közös ismérv***: érettségi bizonyítvánnyal rendelkeznek,

megkülönböztető ismérv: nem → ismérv változat: férfi – nő

Az ismérvek tipizálása:

a) Az ismérv lehet:

- **Közös:** adott tulajdonság szerint egyforma sokasághoz tartoznak. Éppen ezért a sokaság megadása közös ismérvek megadásával történik.
- **Megkülönböztető:** egy adott ismérv többféle változattal rendelkezhet. A sokaság jellemzése a megkülönböztető ismérvek szerinti megoszlás vizsgálatából áll._Pl.: 70% nő – 30% férfi (Az elemzéshez szolgáltatnak alapot.)

b) Az ismérv lehet:

- **Mennyiségi:** (számmal) életkor, testmagasság, lábméret;
- **Minőségi:** nem, foglalkozás, hajszín;

Tárgyi ismérvek: a mennyiségi és a minőségi ismérvek közös elnevezése.

- **Időbeli:** időpont vagy időtartam megnevezéséből áll;
- **Területi:** földrajzi megjelölés, pl. ország, megye, város, község.

A mennyiségi ismérv lehet: diszkrét és folytonos.

A **diszkrét** mennyiségi ismérv csak véges vagy megszámlálhatóan sok, egymástól jól elkülönülő értéket vehet fel.

A **folytonos** mennyiségi ismérv egy adott intervallumon belül bármilyen, tehát kontinuum számosságú értéket felvehet.

A diszkrét ismérv értékei mindig valamilyen számlálás eredményét adják meg, míg a folytonos ismérv értékei mérés eredményeképpen keletkeznek. Így például a háztartásban élők száma diszkrét ismérv, míg a háztartások havi jövedelme folytonos ismérv. A háztartásban élők száma ugyanis csak valamilyen pozitív egész szám lehet, míg a másik ismérv egy adott intervallumban bármilyen értéket felvehet. A diszkrét mennyiségi ismérv értékei pontosan, a folytonos mennyiségi ismérv értékei mindig csak bizonyos pontosságra kerekítve adhatók meg.

c) Az ismérv lehet:

- **Két változattal rendelkező, vagy alternatív:** a tulajdonság meglétét vagy hiányát is kifejezi, pl.: nem;
- **Több változattal rendelkező:** pl. hajszín.

STATISZTIKAI ADAT: valamely statisztikai sokaság tagjainak száma, vagy a sokaság valamilyen számszerű jellemzője. Az adat mindig tartalmaz fogalmi jegyeket (adatazonosítókat, mennyiségi, minőségi, időbeli, és területi azonosítókat), egy számértéket, illetve mértékegységet.

A statisztikai adatok (eredetük szerint) lehetnek:

- **Abszolút adatok:** közvetlen mérés, számlálás útján jönnek létre (pl.: a lakosság száma a népszámlálás adott időpontjában).
- **Leszármaztatott adatok:** különböző matematikai műveletek elvégzésével jutunk származtatott adatokhoz (pl.: a lakosságon belül a férfiak aránya százalékban kifejezve).

  A leszármaztatott számok lehetnek:

  - viszonyszámok,
  - átlagok,
  - indexek, stb.

Mind az abszolút, mind pedig a leszármaztatott adatok lehetnek:

- **Elsődleges** (primer) forrásból származó adatok: statisztikai adatok nyerésének céljából végzett felmérések eredményei. Saját kutatás eredményeképpen születnek.
- **Másodlagos** (szekunder) forrásból származó adatok: adminisztratív tevékenységek melléktermékeként létrejött nyilvántartásokból átvett adatok. Nem saját kutatás eredményei, hanem mások által már korábban összegyűjtött (és publikált) adatokról van szó.

MUTATÓSZÁM: mutatószámoknak nevezzük azokat az abszolút, illetve leszármaztatott statisztikai adatokat, és adatkategóriákat, amelyekkel valamilyen rendszeresen megismétlődő társadalmi, gazdasági jelenséget statisztikailag szoktunk. A szakstatisztika feladata a mutatószámok kidolgozása. Pl.: termelékenységi, hatékonysági, hitelképességi, munka színvonalát jellemző mutatók.

STATISZTIKAI MODELL: a valóság lényegi összefüggéseit tömören jellemző logikai, matematikai, statisztikai konstrukciók.

1.5 A statisztikai munka és szakaszai

Statisztikai munkának a meghatározott cél érdekében végzett statisztikai tevékenységek rendszerét nevezzük.

A statisztikai munka szakaszai:

1. **Statisztikai programkészítés**, amely áll:
   - a célkitűzés megfogalmazásból,
   - az elemzés megtervezéséből,
   - adatfeldolgozási terv készítéséből,
   - az adatgyűjtés előkészítéséből,
   - szervezési feladatokból.

2. **Adatgyűjtés** (milyen adatokat, milyen csoportosításban, mely adatszolgáltatótól)

   Az adatgyűjtés módjai:
   - megfigyelés (résztvevő megfigyelés),
   - megkérdezés: interjú (strukturáltalan-strukturált), kérdőív,
   - önszámlálás.

   Az adatfelvétel köre:
   - teljes körű: a sokaság minden egyedét megfigyeljük, vagy
   - részleges: amikor a sokaságnak csak egy része képezi az adatfelvétel tárgyát.
     - *reprezentatív*: tükrözi a teljes populációt,
     - *kontrollált kísérlet* (mezőgazdaságban dolgozták ki),
     - *egyéb részleges megfigyelés*.

3. **Adatfeldolgozás:** a tömeges jelenség egészére vonatkozóan rendszerezettebb, áttekinthetőbb, tömörített adathalmazt eredményez. Legfontosabb statisztikai eszközei a statisztikai sorok és táblák.

4. **Elemzés, értékelés, közzététel:** amelynek során a tanult statisztikai módszerek kerülnek felhasználásra.

   Elemzési módszerek:

   - ***Leíró (egyszerű) módszerek:***
     - grafikus ábrázolás,
     - viszonyszámok,
     - középértékek,
     - szóródási mérőszámok,
     - statisztikai indexek.
   - ***Következtető (matematikai-statisztikai) módszerek:***
     - statisztikai becslés,
     - hipotézisvizsgálat,
     - összefüggés-vizsgálat (asszociációs-, vegyes-, korrelációs kapcsolat),
     - idősorok komponenseinek vizsgálata.

A statisztikai eljárásoknak alapvetően két csoportját különböztetjük meg: a leíró statisztikai eljárásokat és a matematikai statisztikai (következtető statisztikai) eljárásokat. A matematikai statisztikai eljárások feltételezik a leíró statisztikákat. A következő táblázat komplex szemléletben foglalja össze a statisztikai eljárásokat. A táblázat alapvetően elkülöníti egymástól a leíró és a matematikai statisztikai módszereket, majd külön csoportban tárgyalja az összefüggések feltárására, illetve a különbségek kimutatására szolgáló eljárásokat. További differenciáló tényezőt jelentenek az adatok fajtái és a vizsgált minták száma.

A statisztikai módszerek csoportjai

<table>
<tr><th colspan="4">LEÍRÓ STATISZTIKA</th></tr>
<tr><td>gyakoriságok</td><td>középértékek</td><td colspan="2">szóródási viszonyok és indexek</td></tr>
<tr><td>abszolút</td><td>átlagok</td><td colspan="2">szóródási terjedelem</td></tr>
<tr><td>relatív (százalékos)</td><td>módusz</td><td colspan="2">interkvartilis félterjedelem</td></tr>
<tr><td rowspan="7">kumulatív</td><td>medián</td><td colspan="2">átlagos eltérés</td></tr>
<tr><td rowspan="6">egyéb osztótényezők</td><td colspan="2">variancia</td></tr>
<tr><td colspan="2">szórás</td></tr>
<tr><td colspan="2">relatív szórás</td></tr>
<tr><td colspan="2">koncentráció-elemzés</td></tr>
<tr><td colspan="2">indexszámítás (érték-, és főátlagindexek)</td></tr>
<tr></tr>
<tr><th colspan="4">MATEMATIKAI STATISZTIKA – KÜLÖNBÖZŐSÉGVIZSGÁLATOK (Jelentős-e a különbség?)</th></tr>
<tr><td>adatfajták →
minták száma ↓</td><td>intervallum</td><td>ordinális</td><td>nominális</td></tr>
<tr><td>egy</td><td>egymintás t-próba</td><td>Wilcoxon-próba</td><td>kereszttábla-elemzés, λ^2 -próba</td></tr>
<tr><td>kettő</td><td>kétmintás t-próba és F-próba</td><td>Mann-Whitney-próba</td><td>kereszttábla-elemzés, λ^2 -próba</td></tr>
<tr><td>több</td><td>varianciaanalízis (ANOVA)</td><td>Kruskall-Wallis-próba</td><td>kereszttábla-elemzés, λ^2 -próba</td></tr>
<tr><th colspan="4">MATEMATIKAI STATISZTIKA – ÖSSZEFÜGGÉS-VIZSGÁLATOK (Van-e szoros összefüggés?)</th></tr>
<tr><td>adatfajták →
változók száma↓</td><td>intervallum</td><td>ordinális</td><td>nominális</td></tr>
</table>

kettő	korrelációszámítás	Rangkorrelá-ció-számítás	kereszttábla-elemzés, λ^2 -próba
kettő vagy több	regresszióanalízis	-	-
több, mint kettő	parciális korrelációszámítás	-	-
	faktoranalízis	-	-
	klaszteranalízis	-	-
IDŐSOROK ANALITIKUS VIZSGÁLATA			

Forrás: Saját szerkesztés

E tananyagban kizárólag a leíró statisztika fontosabb elemzési módszereivel kívánok foglalkozni. (A táblázatban a statisztikai módszereket az átfogó rendszerezés érdekében tüntettem fel.) A leíró statisztikai elemzés a kutatás folyamatában fontos szerepet tölt be. A leíró elemzésekkel hipotéziseket (feltételezéseket) ugyan nem lehet bizonyítani, de az előzetes eredmények, illetve azok szemléletes megjelenítése az első olyan információ, amit a kutató a vizsgált adatokkal kapcsolatban megismer. Ez sok esetben további ötleteket adhat a már meglévő hipotézisek kiegészítéséhez.

1.6 Csoportosítás és összehasonlítás

A **csoportosítás** a sokaság egységeinek egy vagy több tulajdonság szerinti felosztása. Azt a tulajdonságot, amely szerint a sokaságot csoportosítjuk, **csoportképző ismérv**nek nevezzük.
Szerepét betöltheti a statisztikai ismérvek mindegyike, így beszélhetünk:

- **Az ismérvek fajtái szerint:**
  - mennyiségi (pl. a csoport hallgatóinak magasság szerinti rendezése),
  - minőségi (pl. a csoport hallgatóinak rendezése nemek szerint),
  - területi (pl. a csoport hallgatóinak rendezése lakhely szerint),

- időbeli ismérv szerinti csoportosításról (pl. a csoport hallgatóinak rendezése születési év szerint).

- **Az ismérvek száma alapján beszélhetünk:**
  - egy ismérv szerinti: **egyszerű**,
  - több ismérv szerinti: **kombinatív** csoportosításról.

A csoportosítás követelményei:

- **teljesség**: a sokaság minden elemét érintse,
- **átfedésmentesség:** az elemek mindegyike besorolható legyen valamelyik, de csak egy kialakított osztályba,
- **homogén osztályokat eredményezzen:** azaz a sokaság azonos osztályba sorolt egységei jobban hasonlítsanak egymáshoz, mint a sokaság más osztály(ok)ba sorolt egységeihez.

Az azonos fajta, de térben és időben változó jelenségek egybevetését **összehasonlítás**nak nevezzük.

Módszerei:

- nagyságrendi viszonyok megállapítása,
- az adatok különbségének meghatározása (abszolút eltérés),
- az adatok hányadosának meghatározása (relatív eltérés).

Fontos, hogy a statisztikai elemzés célját csak akkor érheti el, ha az összehasonlított adatok valóban csak olyan okok miatt térnek el, melyeknek szerepét kutatjuk.

1.7 Mérési skálák

A **mérés** a számok meghatározott szabályok szerinti hozzárendelése jelenségekhez, illetve ezek bizonyos tulajdonságaihoz. A hozzárendelési szabályok alapján négy mérési skálát különböztetünk meg:

a) Névleges (nominális) skála;
b) Sorrendi skála;
c) Intervallum skála;
d) Arány skála.

a) A **névleges (nominális) skála** a számok kötetlen hozzárendelését jelenti. Területi és minőségi ismérvek szerinti megfigyeléseknél alkalmazzuk. E skálán való méréskor a számok csak a sokaság egyedeinek azonosítására szolgálnak. Pl.: rendszám, irányítószám, nemek, stb.

b) Amikor minőségi ismérv szerint vizsgálódunk általában nincs olyan ismérvváltozat, amely a sokaság egészét jellemzi. Ekkor nem lehet egy számszerű értéket a sokasághoz rendelni, de a sorrendet meg lehet határozni az adott jellemző alapján közöttük. Ezt nevezzük **rendezési vagy sorrendi skálán** történő mérésnek. Pl.: a hallgatók osztályzata, az országok hitelképességének sorrendje, a lóversenyen a lovak helyezése, stb.

c) Az **intervallumskála** zérus pontja önkényesen választható meg. Így a zérus pont nem jelenti azt, hogy az adott egyed nem rendelkezik az adott tulajdonsággal. Az előző két skálával ellentétben e skálának már valamilyen mértékegység is a szerves tartozékát képezi. E skála esetében nem értelmezhető két adat összege vagy aránya, de két adat különbsége, illetve a különbségek összege és arány értelmezhető. Pl.: Celsius-fok, Fahrenheit-fok, naptári idő, stb.

d) Az **arányskála** zérus pontja természetesen adódik. E skála esetében bármely két érték aránya független a mértékegységtől, és az értékek összege is értelmezhető. Pl.: hosszúság, jövedelem, tengerszint feletti magasság, vetésterület, stb.

A névleges mérési szintet a legalacsonyabb, az arányskála által képviselt mérési szintet pedig a legmagasabb mérési szintnek nevezhetjük. Egy ismérvet általában akkor szokás „valódi” mennyiségi ismérvnek tekinteni, ha az adott ismérv értékei különbségi vagy arányskálán mért értékek. A minőségi ismérvek változatai viszont többnyire egy névleges skála értékei. Ezek ismerete azért fontos, mert az ismérvek fajtája, illetve a mérés adott szintje mindig behatárolja az elemzés adott esetben szóba jöhető módszereit.

1.8 Mintafeladat

a) Definiáljon egy konkrét sokaságot és soroljon fel a definiált sokaságra vonatkozóan közös és megkülönböztető ismérveket minden ismérvfajtára!
b) Határozza meg, milyen típusú a sokaság!

Megoldás:

a) A Nyíregyházi Egyetemen az első Statisztika előadáson részt vevő gazdálkodási és menedzsment szakos hallgatók képezik a megfigyelt sokaságot 2018. szeptember 5-én, az A épület III. emeletének 352-es termében 14-16 óráig.

Ismérv	Közös	Megkülönböztető
Mennyiségi	Gyakorlatok száma	Magasságuk, életkoruk
Minőségi	Főiskolások	Hajszínük, nemük
Területi	Nyíregyházán tanulnak	Lakóhelyük
Időbeli	Gyakorlat időpontja	Születési idejük

b) Véges, álló, valóságos sokaság.

1.9 Ellenőrző kérdések

Milyen három dolgot értünk a statisztika fogalma alatt?

Fogalmazza meg egy mondatban, hogy mit jelent a statisztika?

Mi képezi a statisztika vizsgálatának tárgyát?

Mit jelent az, hogy a statisztika módszertudomány?

Jelenleg melyik törvény szabályozza Magyarországon a hivatalos statisztikai szolgálatot és annak tevékenységét?

Soroljon fel legalább öt olyan hivatalos szervet, amely a hivatalos statisztikai szolgálatot alkotja!

A nemzetközi statisztikai tevékenységet mely szerv koordinálja?

Az Európai Unió melyik szerve jelenti az EU hivatalos statisztikai szolgálatát?

Sorolja fel a statisztikai információk feldolgozásának fázisait?

Definiálja a statisztikai sokaságot!

Írjon/mondjon egy példát a statisztikai sokaságra!

Mit nevezünk egyednek?

Írjon/mondjon egy példát az egyedre!

Definiálja az ismérv fogalmát!

Írjon/mondjon egy példát a mennyiségi, a minőségi, a területi és az időbeli ismérvre!

Írjon/mondjon egy példát az álló sokaságra!

Írjon/mondjon egy példát a mozgó sokaságra!

Írjon/mondjon egy példát a két változattal, illetve a több változattal rendelkező ismérvre!

Mi az a három fontos jellemző, amivel egy statisztikai adatot megadunk?

Mit jelent az abszolút adat?

Mit jelent a leszármaztatott adat? Írjon/mondjon rá legalább egy példát!

Mikor beszélünk elsődleges (primer) adatról?

Mikor beszélünk másodlagos (szekunder) adatról?

Mit nevezünk mutatószámnak?

Írjon/mondjon példát mutatószámra!

Nevezze meg a statisztikai munka szakaszait!

Rendszerezze az adatfelvétel körét!

Milyen lehet a csoportosítás az ismérvek száma alapján?

Milyen követelményei vannak a csoportosításnak?

Milyen módszerei vannak az összehasonlításnak?

Mit nevezünk mérésnek!

Rendszerezze a mérési skálák típusait!

Jellemezze a nominális skálát! Mondjon rá példát!

Jellemezze a rendezési (sorrendi) skálát! Mondjon rá példát!

Jellemezze az intervallumskálát! Mondjon rá példát!

Jellemezze az arányskálát! Mondjon rá példát!

2 Statisztikai sorok, táblák, mérlegek, grafikus ábrázolás

A statisztikai adatok feldolgozásának, szemléltetésének, elemzésének fontos eszközeit jelentik a statisztikai sorok, táblák, a statisztikai mérlegek és a grafikonok.

2.1 Statisztikai sorok

A statisztikai adatoknak meghatározott összefüggésben, valamilyen meghatározott ismérv szerinti felsorolásait **statisztikai sornak** nevezzük.

A statisztikai sor abszolút számokból és leszármaztatott számokból is állhat.

A statisztikai sorok csoportosítása:

1. A sort alkotó adatok fajtái szerint:

- **Valódi sorok:** egynemű adatokból állnak,
- **Leíró sorok:** különnemű adatokból állnak.

2. A sorok keletkezési módja szerint:

- **Csoportosító sorok:** Egy fősokaság és a megfelelő részsokaságok nagyságát adják meg. Tartozéka összesen adat.
- **Összehasonlító sorok:** Ez is azonos fajta és mértékegységű adatokat tartalmaz, de azok mégsem adhatók össze, vagy összeadhatók, de az összeg elmaradhat.
- **Leíró sorok:** Általában különböző fajta (mértékegységű) adatok, amelyek mindegyike egy meghatározott jelenségre, társadalmi vagy gazdasági egységre vonatkozik több különböző szempontból. Összesen adatot nem tartalmaz.

3. A csoportosító és az összehasonlító sorok osztályozhatók az ismérv fajtája szerint:

- **Idősorok:** a jelenségek, folyamatok időbeli alakulását mutatják. Alfajai:
  - *Állapot idősorok:* álló sokaság időbeli alakulását mutatják.
  - *Tartam idősorok:* mozgó sokaság időbeli alakulását mutatják.
- **Területi sorok:** területi sorok esetén a csoportképző ismérv a terület, amely lehet egy ország vagy az országon belül egy régió, megye, város, stb.
- **Minőségi sorok:** betekintést nyújtanak a sokaság összetételébe, szerkezetébe.
- **Mennyiségi sorok:** mennyiségi ismérv szerinti csoportosítás eredményei.

A mennyiségi ismérveket **változóknak**, lehetséges kimeneteleiket **ismérvértékeknek** nevezzük. Az **ismérvértékek** intervallum- vagy arányskálán mért, valamilyen mértékegységgel bíró számértékek. Ha egy sokaságot valamilyen mennyiségi ismérv szerint vizsgálunk, akkor első lépésként általában az ismérvértékeket sorba rendezzük, ún. rangsort készítünk.

A **rangsor** a mennyiségi ismérv értékeinek monoton sorozata.

A mennyiségi sorok altípusai:

– **Gyakorisági sor:** a mennyiségi ismérv szerinti osztályozás eredményeként kapott speciális csoportosító sor.

Ha a mennyiségi ismérv diszkrét és kevés változattal rendelkezik (pl. a lakások szobaszáma), akkor a gyakorisági sorban minden ismérvértéket felsorolunk.

Ha a mennyiségi ismérv folytonos, vagy diszkrét ugyan, de az általa felvehető értékek sokfélék lehetnek (pl. az aktív keresőket havi keresetük szerint csoportosítjuk), akkor az ismérvértékek tartományát egymást át nem fedő intervallumokra, ún. **osztályközökre** bontjuk. Az így képzett sort **osztályközös gyakorisági sornak** nevezzük.

A **gyakoriság** (f_i) azt mutatja, hogy a mennyiségi ismérv szerint képzett egy-egy osztályba a sokaságnak hány egysége tartozik.

– **Relatív gyakorisági sor:** megmutatja, hogy a mennyiségi ismérv szerint képzett egy-egy osztályba (osztályközbe) a sokaságnak hányad része (hány százaléka) tartozik.

A **relatív gyakoriságok** (gi) nem mások, mint a gyakoriságokból számított megoszlási viszonyszámok:

$$g_i = \frac{f_i}{N}, \text{ ahol}$$

f_i: az i-edik osztályhoz (osztályközhöz) rendelt gyakoriság,

g_i: az i-edik osztályhoz (osztályközhöz) rendelt relatív gyakoriság,

N: a sokaság elemszáma.

– **Kumulált gyakorisági/relatív gyakorisági sorok:** a gyakorisági/relatív gyakorisági sorokban rejlő információk tovább bővíthetők a gyakoriságok/relatív gyakoriságok halmozott összeadásával, azaz kumulálásával. Megkülönböztetünk:

- *felfelé, illetve*
- *lefelé kumulált gyakorisági/relatív gyakorisági sorokat.*

A **felfelé kumulált** gyakoriságok (f_i'), illetve relatív gyakoriságok (g_i') adatai azt mutatják, hogy az adott osztályköz felső határának megfelelő és annál kisebb ismérvértékek hányszor (f_i'), illetve milyen arányban (g_i') fordulnak elő. A kumulált gyakorisági, illetve relatív gyakorisági sorokat úgy képezzük, hogy a gyakoriságokat, illetve relatív gyakoriságokat rendre halmozva összeadjuk felülről lefelé haladva.

A **lefelé kumulált** gyakoriságok (f_i''), illetve relatív gyakoriságok (g_i'') adatai azt mutatják, hogy az adott osztályköz alsó határánál nagyobb ismérvértékek hányszor (f_i”), illetve milyen arányban (g_i”) fordulnak elő.

– **Értékösszegsor:** a mennyiségi ismérv alapján kialakított osztályokhoz (osztályközökhöz) az azokba tartozó egységek ismérvértékeinek összegét rendeli.

A vizsgált mennyiségi ismérv értékeinek egyes osztályokon (osztályközökön) belüli összegeit **értékösszegeknek** (S_i) nevezzük.

Az egyes osztályokhoz tartozó értékösszegeket (S_i) az ismérvértékek (x_i) és a gyakoriságok (f_i) szorzataként kapjuk:

$$S_i = f_i \cdot x_i$$

A sokaság **teljes értékösszeg**e (S):

$$\sum_{i=1}^{k} S_i = S$$

Ha csak az osztályközös gyakorisági sor áll rendelkezésre, akkor az értékösszegeket (S_i) a gyakoriságok (f_i) és az osztályközepek (x_i) szorzataként becsüljük. Az i-edik osztályközép:

$$x_i = \frac{x_{ia} + x_{if}}{2}, \text{ ahol}$$

x_i: az i-edik osztályközép

x_{ia}: az i-edik osztályköz alsó határa

x_{if}: az i-edik osztályköz felső határa

– **Relatív értékösszegsor:** az értékösszegek megoszlását mutatja.

Relatív értékösszegen (Z_i) olyan megoszlási viszonyszámot értünk, amely az egyes osztályok értékösszegét (S_i) a teljes értékösszeghez (S) viszonyítja. Az i-edik osztály relatív értékösszege (Z_i):

$$Z_i = \frac{S_i}{S}$$

– **Kumulált értékösszeg/relatív értékösszegsor:** a gyakorisági sorokhoz hasonlóan az értékösszegsorból és a relatív értékösszegsorból is képezhetünk felfelé kumulált (S_i’, Z_i’), illetve lefelé kumulált (S_i”, Z_i”) sorokat.

2.2 Statisztikai táblák

A statisztikai sorok összefüggő rendszerét **statisztikai táblának** nevezzük. Azt a számot, amely azt jelzi, hogy a tábla egy-egy adata hány statisztikai sorhoz tartozik, a **tábla dimenziójának** nevezzük.

A statisztikai táblák csoportosítása:

1. **Rendeltetésük szerint (milyen célra használjuk):**
   - **Munkatábla:** azért készítjük, hogy belőle további számításokat végezzünk;
   - **Feldolgozói tábla:** az adatokat feldolgozás közben állítjuk össze;
   - **Közlési tábla:** a munka végső eredményeit foglalja össze.

2. **A táblázatba foglalt sorok jellege szerint:**
   - **Egyszerű tábla:** összehasonlító vagy leíró sorokat tartalmaz. Tartalmazhat két összehasonlító sort, egy összehasonlító és egy leíró sort, de két leíró sort nem.
   - **Csoportosító tábla:** tartalmaz egy csoportosító sort (de kettő nem lehet!), és egy összehasonlító vagy egy leíró sort.

- **Kombinációs tábla:** legalább két csoportosító sora, ennek megfelelően két vagy több összesen adata van.

A tábla részei formai szempontból:

- **Oszlop:** a tábla függőleges része;
- **Sor:** a tábla vízszintes része;
- **Rovat:** a sor és az oszlop találkozása.

Szöveget is tartalmazó rovatok:

- **Fejrovat:** a táblában felül helyezkednek el;
- **Oldalrovat:** a sorok előtt találhatók;
- **Összegrovat:** a sorok és az oszlopok adatainak összegzését tartalmazó rovatok.

A statisztikai táblának vannak formai és tartalmi követelményei.

1. **Formai követelmények:**
    - a tábla címe,
    - fej és oldalrovatok megnevezése,
    - hálózat készítés,
    - forrásokra történő hivatkozás,
    - közös mértékegységek feltüntetése,
    - adatok tartalmára vonatkozó megjegyzések.

2. **Tartalmi követelmények:**
   - A táblában üres sor, vagy oszlop nem szerepelhet, üres rovat nem maradhat.
   - Ha valamely rovatba szám nem kerül, akkor a következőképpen járunk el:
     - Nincs adat: a rovatot kihúzzuk (-);
     - Nem ismerjük az adatot: a rovatot kipontozzuk (…);
     - Ha a táblában szereplő mértékegységhez képest a rovatban túl kocsi az adat, akkor 0,0-t írunk.

2.3 Statisztikai mérlegek

A **statisztikai mérlegek** speciális statisztikai táblák.

Típusaik:

- **Könyvviteli típusú mérlegek:** állósokaságok közötti összefüggések, változások bemutatására szolgálnak (pl. adott év létszámmérlege).
- **Sakktáblaszerű mérlegek:** mozgósokaságok elszámolására szolgálnak. Input-output tábláknak is szoktuk nevezni. (Használatuk demográfiai jellemzők vizsgálatánál különösen jellemző: pl. adott év belső vándorlási mérlege.)

2.4 Grafikus ábrázolás

A statisztikai adatok megjelenítésének, szemléltetésének és népszerűsítésének közkedvelt eszközei a **grafikus ábrák**. A szemléltetésen és népszerűsítésen túl azonban elemzési eszközök is, amelyek mindig arányokat érzékeltetnek. Ábrák segítségével jobban felismerhető a vizsgált jelenség természete, sőt a jelenségek között fennálló összefüggésekre is ráirányítják a figyelmünket. A statisztikai adatokat leggyakrabban vonalakkal, pontokkal, körökkel vagy oszlopokkal, azaz távolságokkal, területekkel, esetenként köbtartalommal szemléltetjük.

A grafikus ábráknak több fajtája ismeretes:

1) mértani alakzatok felhasználásával készült ábrák:
   - **a)** koordináta rendszerben (derékszögű vagy poláris),
   - **b)** koordináta rendszeren kívül,

2) térképeken alapuló ábrák,

3) figurális ábrák.

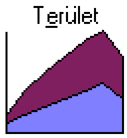
Az ábrák szerkesztése – a célnak legmegfelelőbb ábratípus kiválasztását követően – számítógépekkel már mindennapossá vált, ezért csak az ábraszerkesztés szempontjaira, illetve a típus megválasztásának okára kell felhívni a figyelmet. A statisztikai ábrának két fő része van. A szöveges magyarázó rész és az ábramezőben elhelyezett grafikus ábra. A szöveges magyarázó részhez tartozik az ábra pontos címe, amely sorszámot követően az ábra alatt kerül elhelyezésre. Tartozéka az egyértelmű jelmagyarázat, a mértékegység, valamint a forrásokra történő hivatkozás.

Diagramtípusok:

Az adatok egyértelmű és hatásos bemutatására több diagramtípus közül választhatunk, de készíthetők a korábbiak felhasználásával kombinált típusok is. Minden egyes diagramtípus több altípussal vagy változattal rendelkezik. A következőkben a főbb típusokat tekintjük át röviden.

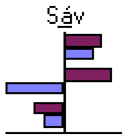
Területdiagram:

Értékek relatív fontosságát mutatja egy adott időszak során, a változás mértékét (az értékek nagyságát) emeli ki az idő és a változás sebessége helyett. Pl. forgalmi, termelési adatok esetén.



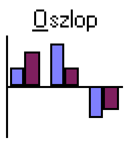
- *Sávdiagram:*

Tulajdonképpen elemek összehasonlítására szolgál, hasonlít az oszlopdiagramra, csak 90°-kal el van forgatva. A sávdiagramon a kategóriák függőlegesen, míg az értékek vízszintesen helyezkednek el, így a hangsúly inkább az összehasonlításon van. Statisztikai adatok szemléltetésére szívesen használják.



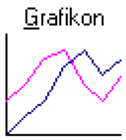
- *Oszlopdiagram:*

Habár hasonlít a sávdiagramhoz, az oszlopdiagram kategóriái vízszintesen, értékei pedig függőlegesen helyezkednek el. Felhasználási köre is hasonló, de a műszaki gyakorlatban is szívesen használják.



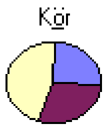
Grafikon:

Adott időszakot egyenlő nagyságú intervallumokra osztva az adatok alakulását, trendjét szemlélteti az időszakban. A változás mértéke helyett inkább az idő múlására és a változás mikéntjére, sebességére helyezi a hangsúlyt. Ha trendeket kell szemléltetünk, vagy ha az intervallumok nem egyenlő hosszúak, vagy ha osztályokba csoportosított intervallumokkal van dolgunk, akkor a *pontdiagram* általában megfelelőbb mint a grafikon. A műszaki alkalmazásokban ezért inkább ez utóbbit használjuk.



- *Kördiagram:*

A részeknek az egészhez való viszonyát, arányait szemlélteti. Ez a diagramtípus különösen akkor hasznos, ha egy lényeges elemet szeretnénk kiemelni. A kördiagram mindig egyetlen adatsort tartalmaz. Főleg statisztikai adatok szemléltetésére használják.

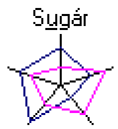


- *Perecdiagram:*

A kördiagramhoz hasonló. A fő különbség - a közepén lévő lyukon kívül - az, hogy a kördiagramtól eltérően perecdiagrammal több adatsor ábrázolása is lehetséges.

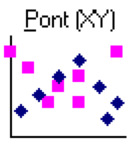
- *Sugárdiagram:*

  Adatsoroknak egy középponthoz illetve egymáshoz viszonyított változásait vagy gyakoriságát szemlélteti. Minden egyes kategória saját értéktengellyel rendelkezik, amelyek a középpontból sugárirányban ágaznak ki. Az azonos adatsorhoz tartozó adatjelölőket vonalak kötik össze.



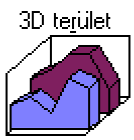
- *Pontdiagram:*

  A különböző adatsorokban lévő számértékek közötti összefüggést, az összefüggés fokát szemlélteti, illetőleg számok két csoportját, az x-y koordinátapárokat jeleníti meg. Pontdiagram esetén az adatok vagy adatcsoportok időben egyenetlen eloszlásúak is lehetnek, ezért műszaki-tudományos munkákban gyakran használjuk.



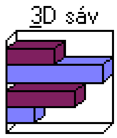
- *3D területdiagram:*

  A területdiagram háromdimenziós képét mutatja, amely a kirajzolt értékek összegét hangsúlyozza, és a diagram adatsorait külön sorokba választja szét, hogy szemléltesse az adatsorok közötti különbséget.



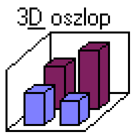
3D sávdiagram:

A sávdiagram háromdimenziós képét mutatja, amely az egyes elemeknek egy adott időpontban vett értékét hangsúlyozza, illetve összehasonlítja az elemeket.



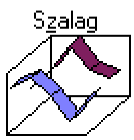
- *3D oszlopdiagram:*

Az oszlopdiagram háromdimenziós képét mutatja két változat valamelyikében. Az egyszerű 3D oszlopdiagram az x (kategória-) tengely mentén elhelyezkedő háromdimenziós oszlopjelölőket jeleníti meg, míg a 3D térhatású oszlopdiagram két tengely - az x (kategória-) tengely és az y (adatsor-) tengely - mentén elhelyezkedő adatpontokat hasonlítja össze. Az adatpontok mindkét változatnál a z tengely mentén helyezkednek el. Ez a diagramtípus lehetővé teszi, hogy egy adatsoron belül az adatokat könnyebben össze tudjuk hasonlítani, de az adatok továbbra is kategóriák szerint jelenjenek meg.



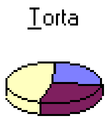
- *Szalagdiagram típus:*

Tulajdonképpen a grafikon vonalait mutatja háromdimenziós szalag formájában. Ezt a diagramtípust előadások, bemutatók során használjuk gyakran az adatok tetszetős megjelenítésére.



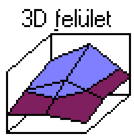
- *Tortadiagram:*

A kördiagram háromdimenziós változata. A tortadiagram csak egyetlen adatsort tud ábrázolni.



- *Felületdiagram:*

Azt mutatja meg, hogyan nézne ki, ha gumihártyát feszítenénk rá egy 3D oszlopdiagramra. A felületdiagram két adatkészlet optimális kombinációjának meghatározásában nyújthat segítséget. Ez a diagramtípus nagy mennyiségű adatok közötti olyan kapcsolatokat is képes feltárni, melyeket egyébként csak nehezen lehetne észrevenni. A domborzati térképekhez hasonlóan azonos színek vagy mintázatok nem adatsorokat jelölnek, hanem az azonos értéktartományt megjelenítő felületdarabokat jelölik. Készíthető „*drótkeret*” formátumban is, amely az adatokat színek nélkül jeleníti meg.



2.5 Mintafeladat

Egy vállalat részmunkaidőben foglalkoztatott dolgozóinak kereseti adatai 2018 novemberében:

Nettó kereset (eFt/fő)	Dolgozók száma (fő)
-70,00	12
70,01–90,00	20
90,01–110,00	34
110,01–130,00	25
130,01–150,00	9
Összesen	**100**

a) Állapítsa meg a statisztikai sor és az ismérv típusát!

b) Készítsen relatív gyakorisági sort, becsült értékösszegsort, illetve relatív értékösszegsort, továbbá kumulálja a gyakoriságokat, a relatív gyakoriságokat, az értékösszegeket és a relatív értékösszegeket!

c) Értelmezzen egy-egy adatot minden egyes statisztikai sorból!

<u>**Megoldás:**</u>

a) Osztályközös gyakorisági sor, diszkrét mennyiségi ismérv.

b)

Nettó kereset (eFt/fő)	f_i	g_i	u_i/x_i	S_i	Z_i	f_i'	g_i'	S_i'	Z_i'	f_i''	g_i''
-70,0	12	0,12	60	720	0,07	12	0,12	720	0,07	100	1,00
70,1–90,0	20	0,2	80	1600	0,16	32	0,32	2360	0,23	88	0,88
90,1–110,0	34	0,34	100	3400	0,34	66	0,66	5720	0,57	68	0,68
110,1–130,0	25	0,25	120	3000	0,30	91	0,91	8720	0,87	34	0,34
130,1–150,0	9	0,09	140	1260	0,13	100	1,00	9980	1,00	9	0,09
Össze-sen	100	1,00	-	9980	1,00	-	-	-	-	-	-

A mennyiségi sor típusa	**A mennyiségi sor értelmezése**
f_i=20	20 olyan dolgozó van, aki 70 és 90 eFt között keres.
g_i=0,25	A dolgozók 25%-a 110 és 130 eFt közötti összeget keres.
S_i=1600	Az a 20 dolgozó, akinek a fizetése70 és 90 eFt közé esik, összesen 1.600 eFt-ot keres.
Z_i=0,341	A 90 és 110 eFt között kereső 34 dolgozó összkeresete az összes dolgozó keresetének 34,1%-a.
f_i'=66	66 olyan dolgozó van, aki 110 eFt-nál kevesebb összeget keres.
g_i'=0,32	A 90 eFt-nál kevesebbet keresők aránya 32%.
S_i'=5720	A 110 eFt-nál kevesebbet keresők összes bére 5.720 eFt.
Z_i'=0,874	A 130 eFt-nál kevesebbet kereső dolgozók összkeresete az összes dolgozó keresetének 87,4%-a.
f_i"=34	34 olyan dolgozó van, aki110 eFt-nál többet keres.
g_i"=0,68	A 90 eFt-nál többet keresők aránya 68%.
S_i"=4260	A 110 eFt-nál többet keresők összes bére 4.260 eFt.
Z_i"=0,768	A 90 eFt-nál többet kereső dolgozók összkeresete az összes dolgozó keresetének 76,8%-a.

2.6 Ellenőrző kérdések

Mit nevezünk statisztikai sornak?

Csoportosítsa a statisztikai sorokat a sort alkotó adatok fajtája szerint!

Csoportosítsa a statisztikai sorokat a sorok keletkezési módja szerint!

Csoportosítsa a statisztikai sorokat az ismérvek típusa alapján!

Hogyan nevezzük a mennyiségi ismérveket?

Hogyan nevezzük a mennyiségi ismérvek lehetséges kimeneteleit?

Mit nevezünk rangsornak?

Sorolja fel a mennyiségi ismérvek altípusait!

Mit mutat meg a gyakoriság? Hogyan jelöljük?

Mit jelent a relatív gyakoriság? Hogyan jelöljük? Hogyan határozzuk meg?

Mit jelent a kumulálás?

Mit nevezünk értékösszegnek? Hogyan jelöljük?

Osztályközös gyakorisági sor esetén hogyan becsüljük meg az értékösszegeket?

Mit mutat meg a relatív értékösszeg? Hogyan jelöljük? Hogyan határozzuk meg?

Mit nevezünk statisztikai táblának?

Mit értünk a tábla dimenziója alatt?

Rendszerezze a statisztikai táblák típusait rendeltetésük szerint!

Rendszerezze a statisztikai táblákat a táblázatba foglalt sorok típusa szerint!

Sorolja fel a tábla részeit formai szempontból!

Sorolja fel a szöveget is tartalmazó rovatokat!

Sorolja fel a statisztikai tábla formai követelményeit!

Sorolja fel a statisztikai tábla tartalmi követelményeit!

Mi a statisztikai mérleg, nevezze meg a fajtáit!

Mit nevezünk diagramnak/grafikus ábrának?

Mi a fő tulajdonsága a grafikus ábrázolásnak?

Rendszerezze a grafikus eljárásokat!

Mondjon példát koordináta-rendszeren kívüli grafikus ábrázolásra!

Sorolja fel a koordináta-rendszeren alapuló grafikus ábrák/diagramok fő típusait!

3 Viszonyszámok

A statisztikai elemzés egyszerűbb eszközeit a leszármaztatott számok képzésének módszerei alkotják. Abszolút számokból egyszerű műveletekkel történik a meghatározásuk.

A leíró statisztika kiemelt területei:

- viszonyszámok,
- középértékek,
- szóródási mérőszámok,
- indexek.

Jelen fejezetben a viszonyszámok módszertani bemutatására kerül sor.

A **viszonyszám** két egymással valamilyen kapcsolatban lévő (ugyanazon statisztikai sorhoz tartozó) statisztikai adat hányadosa.

Amit viszonyítunk: viszonyított adat, a viszonyítás tárgya.

Amihez viszonyítunk: viszonyítási alap, a viszonyítás bázisa.

$$V(Viszonyszám) = \frac{A \ (Viszonyított \ adat)}{B \ (Viszonyítási \ alap)}$$

A **viszonyszámok típusai** a statisztikai sorok típusai alapján határozhatók meg.

3.1 Csoportosító sorból számított viszonyszámok

Csoportosító sorból kétféle viszonyszámot tudunk számolni: megoszlási és koordinációs viszonyszámot.

a) **Megoszlási viszonyszám:** a sokaság egyes részeinek a sokaság egészéhez viszonyított arányát fejezi ki. Meghatározása úgy történik, hogy a statisztikai sor adatait elosztjuk az összesen adattal. Leggyakrabban minőségi és mennyiségi sorokból számítjuk. A megoszlási viszonyszámok a jelenségek struktúráját jellemzik, a sokaság belső szerkezetét önmagában fejezik ki. Pl.: A hallgatók 60%-a férfi.

$$V_m = \frac{x_i}{\sum x_i}, \text{ vagy } g_i = \frac{f_i}{\sum f_i}$$

b) **Koordinációs viszonyszám:** valamilyen csoportosító sor egyik részadatának egy másik részadatához való viszonyítása. Pl.: 1000 férfira jutó nők száma.

3.2 Összehasonlító sorból számított viszonyszámok

Összehasonlító sorból **összehasonlító viszonyszám** számítható, amely azt mutatja meg, hogy a vizsgált jelenség időben vagy térben különböző adatai hányszorosát, illetve milyen hányadát teszik ki a bázisul választott adatnak. Leggyakrabban idő és területi sorokból számítható.

Az összehasonlító viszonyszámok altípusai:

a) **Dinamikus viszonyszám:** időbeli összehasonlító viszonyszám. Két időszak vagy időpont adatának, mégpedig az összehasonlítás tárgyát képező tárgyidőszak, valamint az összehasonlítás alapját képező bázisidőszak adatának hányadosa.

Fajtái:

- *Bázisviszonyszám* (V_b): állandó bázissal számított dinamikus viszonyszám. Számításának módja: a tárgyidőszak adatainak (y_i) és az állandó bázisnak (y_0) a hányadosa.

$$V_b = \frac{y_i}{y_0}$$

- *Láncviszonyszám* (V_l): kettőnél több tagból számított olyan „láncszerűen” egymáshoz kapcsolódó dinamikus viszonyszám, amely minden esetben két „szomszédos” időszakot hasonlít össze.

Számításának módja: a tárgyidőszak (y_i) és a megelőző időszak (y_{i-1}) (változó bázis) adatainak hányadosa.

$$V_l = \frac{y_i}{y_{i-1}}$$

Bizonyos **összefüggések** határozhatók meg **a bázis- és láncviszonyszámok között:**

- A legelső időszakra láncviszonyszámot nem tudunk számítani.
- Az állandó bázisul választott időszakban a bázisviszonyszám egyenlő 1, azaz 100%.
- Az állandó bázisidőszak utáni első tárgyidőszakban a bázis és láncviszonyszámok megegyeznek.
- Az állandó bázis utáni k láncviszonyszám szorzata egyenlő a k-adik bázisviszonyszámmal.

$$V_{l1} \cdot V_{l2} \cdot ..... \cdot V_{lk} = V_{bk}, \text{ ugyanis } \frac{y_1}{y_0} \cdot \frac{y_2}{y_1} \cdot ..... \frac{y_k}{y_{k-1}} = \frac{y_k}{y_0}$$

- A bázisviszonyszámokból úgy számíthatunk láncviszonyszámokat, mint az eredeti abszolút számokból. Tehát az egymást követő két bázisviszonyszám hányadosa megadja a k-adik láncviszonyszámot.

$$V_{lk} = \frac{V_{bk}}{V_{b(k-1)}}, \text{ mert } \frac{y_k}{y_0} \div \frac{y_{k-1}}{y_0} = \frac{y_k}{y_{k-1}}$$

- Az eredeti abszolút számok ismerete nélkül is átszámíthatjuk a bázisviszonyszámokat új bázisra úgy, mintha a bázisviszonyszámok abszolút számok lennének.

b) Területi összehasonlító viszonyszám: két területre vonatkozó adat hányadosaként határozható meg. A bázis kiválasztása problémát okozhat. Számítás a bázisviszonyszámhoz hasonlóan történik.

c) Teljesítmény viszonyszámok: speciális összehasonlító viszonyszámok, mert a számításukhoz legalább két statisztikai sorra van szükség, amely nem is mindig összehasonlító, hanem csoportosító sor is lehet.

Leginkább a vállalati tevékenység elemzéséhez használják, amikor pl. két egymást követő év tervezett és ténylegesen teljesített adatait vetik egybe. Alkalmat ad a struktúra változásának vizsgálatára is. Leggyakrabban a növekedési célkitűzés, a tényleges teljesítés, a tényleges növekedés és a teljesítés tervszerűségének az elemzéséhez számolunk viszonyszámot. Ennek megfelelően a teljesítményviszonyszámoknak három fajtáját különböztetjük meg: a tervteljesítési, a tervfeladat és a tervszerűségi viszonyszámot.

- *Tervteljesítési viszonyszám:* valamilyen ténylegesen elért eredményt ugyanazon jelenség optimálisnak tartott, norma vagy terv szerinti értékéhez viszonyítjuk.

$$V_{tervteljesítési} = \frac{Tény}{Terv}$$

Azt mutatja meg, hogy hány %-ban teljesítettük túl vagy alul a tervet.

- *Tervfeladat viszonyszám:* valamilyen optimálisnak tartott, norma vagy terv szerinti értéket viszonyítunk a bázisul választott adathoz.

$$V_{tervfeladat} = \frac{Terv}{Bázis}$$

Azt mutatja meg, hogy a bázishoz képest hány %-os növekedést irányoztak elő.

- *Tervszerűségi viszonyszám:*

$$V_{tervszerüség} = \frac{Tervszerüség}{Terv}$$

Azt mutatja meg, hogy hány %-ban voltunk tervszerűek.

3.3 Leíró sorból számított viszonyszámok

Leíró sorból **intenzitási viszonyszám**ot számolunk, amely megmutatja, hogy az egyik jelenségből átlagosan mennyi jut a másiknak egy egységére, azaz, hogy a vizsgált jelenség milyen intenzitással fordul elő valamilyen más jelenség környezetében.

Jellegzetes típusai:

- sűrűségmutatók, pl.: népsűrűség, orvossűrűség,
- arányszámok, pl.: születési, halálozási,
- koordinációs viszonyszámok
- átlag jellegű mutatók.

Bizonyos intenzitási viszonyszámok megfordíthatók, tehát a két adat szerepét a vizsgálat céljától függően felcserélhetjük. Az így képzett mutatót **egyenes mutató**nak, a reciprok értéket pedig **fordított mutató**nak nevezzük.

Az lesz az egyenes mutató, amelyik mutatónak a növekedése a kedvező jelenségre utal.

Pl.: Határozzuk meg az orvosellátottság egyenes és fordított mutatóját!

$\frac{Lakos}{Orvos}$: Azt fejezi ki, hogy egy orvosra hány lakos jut.

Ha nő az egy orvosra jutó lakosok száma, akkor romlik az egészségügyi ellátás színvonala. Mivel ez a jelenség kedvezőtlen, ezért ez lesz a fordított mutató.

$\frac{Orvos}{Lakos}$: Azt fejezi ki, hogy egy lakosra hány orvos jut.

Amennyiben nő az egy lakosra jutó orvosok száma, akkor javul az egészségügyi ellátás színvonala. Tehát ez utal a kedvező jelenségre, ezért ez lesz az egyenes mutató.

Egyes intenzitási viszonyszámok számíthatók:

- **nyers**, illetve
- **tisztított** formában.

Nyers mutatóról akkor beszélünk, ha a jelenséget a vele lazább kapcsolatban álló másik jelenséghez (teljes viszonyítási alaphoz) viszonyítjuk, míg ha a vele közvetlenebb, szorosabb kapcsolatban álló jelenséghez (kisebbített viszonyítási alaphoz) viszonyítjuk, a viszonyszám **tisztított**.

$$V = \frac{A}{B} = \frac{A}{b} \cdot \frac{b}{B}, \text{ ahol}$$

b: részhalmaza B-nek, és b szorosabb kapcsolatban van az A-val,

$\frac{A}{B}$: nyers intenzitási viszonyszám,

$\frac{A}{b}$: tisztított intenzitási viszonyszám,

$\frac{b}{B}$: tiszta rész aránya (megoszlási viszonyszám).

Tehát **a nyers viszonyszám egyenlő a tisztított viszonyszám és a tiszta rész arányának szorzatával.**

Pl.: Születési arányszám

$$V = \frac{A}{B} = \frac{A}{b} \cdot \frac{b}{B}$$

Szül.-i arányszám=

$$\frac{Születések\ száma}{Nőő\ száma} = \frac{Születések\ száma}{SzülőKépes\ korú\ nőő\ sz.} \cdot \frac{SzülőKépes\ korú\ nőő\ sz.}{Nőő\ száma}$$

Az eredeti viszonyítási alapból (nők) kivettünk egy részhalmazt (szülőképes korú nőket), amely szorosabb kapcsolatban van a vizsgált jelenséggel (születések számával).

3.4 Mintafeladat

Egy térségben a gazdasági társaságok számának alakulása 2011 és 2017 között:

Év	Gazdasági Társaságok		
	(db)	2017=100%	Előző év=100%
2011	165	100,00	-
2012	170	103,03	103,03
2013	150	90,91	88,24
2014	145	87,88	96,67
2015	168	101,82	115,86
2016	155	93,94	92,26
2017	160	96,97	103,23

a) Számítsa ki és értelmezze a gazdasági társaságok számának alakulását jelző bázisviszonyszámokat 2011. évi bázison!

b) Határozza meg és értelmezze a láncviszonyszámokat!

Megoldás:

a) A tárgyidőszak adatait elosztjuk az állandó bázis, vagyis a 2011-es év adatával (165).

Az állandó bázisul választott időszakban a bázisviszonyszám egyenlő 100%-kal.

$$V_{b2012}=\frac{170}{165}=1{,}0303\rightarrow 103{,}03\%;$$

$$V_{b2013}=\frac{150}{165}=0{,}9091\rightarrow 90{,}91\%, \text{ stb.}$$

Értelmezés:

V_{b2012}=103,03%: 2011-ről 2012-re 3,03%-kal nőtt a gazdasági társaságok száma.

V_{b2013}=90,91%: 2011-hez képest 2013-ra 9,09%-kal csökkent a gazdasági társaságok száma.

b) A legelső időszakra láncviszonyszámot nem tudunk számolni.

Az állandó bázisidőszak utáni első tárgyidőszakban a bázis és láncviszonyszámok megegyeznek (103,03%).

$$V_{l2013}=\frac{150}{170}=0{,}8824\rightarrow 88{,}24\%;$$

$$V_{l2014}=\frac{145}{150}=0{,}9667\rightarrow 96{,}67\%, \text{ stb.}$$

Értelmezés:

V_{l2013}=88,24%: 2012-ről 2013-ra 11,76%-kal csökkent a gazdasági társaságok száma.

V_{l2014}=103,23%: 2013-ról 2014-re 3,23%-kal nőtt a gazdasági társaságok száma.

3.5 Ellenőrző kérdések

Mit nevezünk viszonyszámnak!

Rendszerezze a viszonyszámokat a statisztikai sorok típusa alapján!

Mit nevezünk megoszlási viszonyszámnak!

Mi a koordinációs viszonyszám lényege! Mondjon példát koordinációs viszonyszámra!

Definiálja a dinamikus viszonyszámot!

Mit nevezünk bázisviszonyszámnak!

Mit nevezünk láncviszonyszámnak!

Milyen összefüggések vannak a bázis- és a láncviszonyszámok között?

Definiálja a tervteljesítési viszonyszámot!

Definiálja a tervfeladat viszonyszámot!

Definiálja a tervszerűségi viszonyszámot!

Mit nevezünk intenzitási viszonyszámnak!

Mikor beszélünk egyenes, illetve fordított intenzitási viszonyszámról?

Értelmezze a nyers és a tisztított intenzitási viszonyszámot!

4 Középértékek

Azonos fajta adatok tömegének közös jellemzői a **középértékek**, amelyek egyetlen adatba tömörítik a sokaság vizsgálat szempontjából lényeges tulajdonságait.

Ahhoz, hogy a középértékek valóban „jól”, tömören jellemezzék a sokaságot vagy mintát, fontos **követelményék**nek kell eleget tenniük:

- Egyértelműen és algebrailag könnyen számíthatók legyenek;
- Tipikus, jellemző értékek legyenek;
- Szemléletesen, jól lehessen őket értelmezni;
- Közepes helyzetet foglaljanak el, azaz a legkisebb és a legnagyobb elem közé essenek.

A középértékek típusai:

KÖZÉPÉRTÉK

számított (átlagok)	***helyzeti (kijelölt)***
- számtani (aritmetikai) átlag	- módusz
- mértani (geometriai) átlag	- medián
- négyzetes (kvadratikus) átlag	- kvantilisek
- harmonikus átlag	

4.1 Számított középértékek

Egyetlen adatba sűrítik a sokaság vizsgálat szempontjából lényeges tulajdonságait. Van egy algoritmusa, számítási szabálya, ami alapján meghatározzuk az értékét. Nagyságára minden részadatnak befolyásoló szerepe van.

4.1.1 Számtani vagy aritmetikai átlag

Olyan számított középérték, amelyet az átlagolandó értékek helyébe írva, azok **összege** változatlan marad. A sokaság vagy minta elemeiből általában akkor számítható számtani átlag, ha az elemek összegének van valamilyen tárgyi értelme.

- Ha minden elem (x_i) csak egyszer fordul elő→egyszerű számtani átlagot,
- Ha az adatok előfordulása (f_i) különböző→súlyozott számtani átlagot számolunk.

Számításuk:

– *Egyszerű számtani átlag:* az átlagolandó adatok összegének (Σx_i) és számának (n) a hányadosa:

$$\overline{x_a} = \frac{\sum_{i=1}^{n} x_i}{n}, \text{ ahol}$$

x_i: a sokaság/minta i-edik eleme,

n: a sokaság/minta elemszáma.

– *Súlyozott számtani átlag:* az átlagolandó adatok és a súlyok szorzatösszegének, valamint a súlyok összegének hányadosa:

$$\overline{x_a} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}, \text{ ahol}$$

f_i: gyakoriság.

– *Súlyozott számtani átlag osztályközös gyakorisági sorokból:* az osztályközepek (osztályköz alsó és felső értékének átlaga: u_i vagy x_i) és a megfelelő súlyok szorzatösszegének, valamint a súlyok összegének hányadosa:

$$\overline{x_a} = \frac{\sum_{i=1}^{k} f_i x_i}{\sum_{i=1}^{n} f_i}, \text{ ahol}$$

x_i/u_i: az i-edik osztályközép.

A számtani átlag tulajdonságai:

1. Az átlagolandó adatok átlagtól vett eltéréseinek összege nulla:

$$\sum_{i=1}^{n}\left(x_i - \overline{x}\right) = 0$$

2. Az átlagolandó adatok mindegyikéből egy tetszőleges „a" állandót levonva ezen eltérések négyzetösszege akkor lesz minimális, ha az „a" állandó éppen a számtani átlag. **Ezt nevezzük a számtani átlag négyzetes minimum tételének:**

$$\sum_{i=1}^{n}\left(x_i - a\right)^2 = \text{minimális, ha } a = \overline{x_a}$$

3. Ha az átlagolandó adatokból rendre kivonunk (hozzáadunk) egy állandó számot, az új értékek számtani átlaga éppen a levont (hozzáadott) számmal fog eltérni az eredeti adatok átlagától:

Ha x_1, x_2,….., x_n, átlaga $\overline{x_a}$, akkor x_1+a, x_2+a,…..,x_n+a átlaga $\overline{x_a}$ +a lesz.

4. Ha az átlagolandó adatokat rendre megszorozzuk (elosztjuk) egy állandó számmal, akkor az így kapott elemek átlaga éppen az állandó szám szorosa (hányada) lesz az eredeti elemek átlagának:

Ha x_1, x_2,….., x_n, átlaga $\overline{x_a}$, akkor bx_1, bx_2,…..,bx_n átlaga $b\overline{x_a}$ lesz.

5. Az utóbbi két tulajdonságot együttesen alkalmazva a számtani átlag meghatározásának új képletéhez jutunk.

Számítási alap: egyenlő szélességű osztályközös gyakorisági sor, az állandó levonandó érték az egyik osztályközép (a), az állandó osztó az osztályszélesség (i).

A műveletek eredménye:

$$d_i = \frac{u_i - a}{i}$$

A számtani átlag számításának menete:

- Intervallumegységben meghatározott átlag, amely az osztályközepek állandó értéktől vett eltérései súlyozott szorzatösszegének és a súlyok összegének a hányadosa:

$$\bar{d} = \frac{\sum f_i d_i}{\sum f_i}$$

- Az intervallumegységben meghatározott átlag és az állandó osztályszélesség értékének szorzata: $\bar{d} \cdot i$
- A levont állandó és a $d \cdot i$ szorzat összege:

$$\bar{x}_a = a + \bar{d} \cdot i = a + \frac{\sum f_i d_i}{\sum f_i} \cdot i$$

6. A súlyozott számtani átlag nagyságát az átlagolandó adatok mellett a súlyarányok határozzák meg. Ebből következően a súlyok helyett azok megoszlási viszonyszámai is használhatók, így a számtani átlag az átlagolandó adatok és a relatív súlyok szorzatösszege:

$$\bar{x}_a = \sum g_i \cdot x_i = \sum g_i \cdot u_i$$

4.1.2 Mértani (geometriai) átlag

Olyan számított középérték, amelyet az átlagolandó adatok helyére írva azok **szorzata** változatlan marad. Mértani átlagot akkor számolunk, ha az adatok szorzatának van valamilyen tárgyi értelme.

Számítása:

- *Egyszerű geometriai átlag:* az adatok szorzatának (π) n-edik gyöke:

$$\overline{x_g} = \sqrt[n]{\pi x_i}$$

- *Súlyozott geometriai átlag:* az átlagolandó adatok előfordulásaiknak megfelelő hatványaiból képzett szorzat, súlyösszegének megfelelő gyökeként:

$$\sqrt[\Sigma f_i]{\pi x_i^{f_i}}$$

A mértani átlagot idősorok adatainak átlagolására használjuk, amikor az időbeli változás átlagos ütemét kívánjuk meghatározni.

4.1.3 Négyzetes (kvadratikus) átlag

Olyan számított középérték, amellyel az átlagolandó adatokat helyettesítve, azok **négyzetösszege** változatlan marad.

Számítása:

- *Egyszerű négyzetes átlag:* az átlagolandó adatok négyzetösszegének és az adatok számának hányadosából számított négyzetgyök érték:

$$\overline{x}_q = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}}$$

– *Súlyozott négyzetes átlag:* az átlagolandó adatok négyzetei súlyokkal képzett szorzatösszegének és a súlyok összegének hányadosából vont négyzetgyök érték:

$$\bar{x}_q = \sqrt{\frac{\sum_{i=1}^{n} f_i \cdot x_i^2}{\sum_{i=1}^{n} f_i}}$$

Az eredeti súlyok helyett a *relatív súlyokkal* is számolhatunk:

$$\bar{x}_q = \sqrt{\sum_{i=1}^{n} g_i \cdot x_i^2}$$

Négyzetes átlag alkalmazására akkor kerül sor, ha az átlagolandó értékek között (+) és (-) előjelű értékek is vannak. A gyakorlatban elsősorban a szóródás mérőszámainak meghatározásánál használjuk.

4.1.4 Harmonikus átlag

Olyan számított középértékek, amelyet az egyes átlagolandó adatok helyére helyettesítve, azok **reciprokösszege** változatlan marad. Harmonikus átlagot akkor számíthatunk, ha az átlagolandó adatok reciprokának és a reciprokok összegének van valamilyen tárgyi értelme.

Számítása:

– *Egyszerű harmonikus átlag:* az adatok számának (n) és értékeik reciprokösszegének a hányadosa:

$$\bar{x}_h = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

– *Súlyozott harmonikus átlag:* a súlyok összegének és az átlagolandó adatok reciprokai súlyokkal képzett szorzatösszegének a hányadosa:

$$\bar{x}_h = \frac{n}{\sum_{i=1}^{n} \frac{f_i}{x_i}}$$

Az eredeti súlyok helyett a *relatív súlyokkal* számolva:

$$\bar{x}_h = \frac{1}{\sum_{i=1}^{n} g_i \frac{1}{x_i}}$$

Ugyanazon adatsorból, ha mind a 4 féle számított középértéket meghatározzuk, eltérő nagyságrendű átlagokat kapunk:

$$\bar{x}_h \langle \bar{x}_g \langle \bar{x}_a \langle \bar{x}_q$$

Éppen ezért lényeges a helyes módszer megválasztása!

4.2 Idősorok elemzése átlagokkal

Az átlagolás célja egyrészt az idősor átlagos értékének meghatározása, másrészt az idősorban végbemenő átlagos változások kimutatása lehet.

- **Az idősorok átlagos értékének meghatározása:**

  Az állapot- és tartamidősor adatait eltérő módon átlagoljuk:

  – A **tartamidősor** adatai összegezhetők, ezért átlagolásukra a **számtani átlagot** használjuk:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Az így kiszámított átlag a megfigyelt jelenség egy időszakra jutó átlagos értékét mutatja.

– Az **állapotidősorok** adatai egy-egy időpontra vonatkoznak, összegüknek nincs tárgyi értelme. Így tehát állapotidősorokból a számtani átlag speciális változatát, az ún. **kronologikus átlagot** számítjuk, amely kétszeres számtani átlag.

Számítása (ha az időpontok távolságai megegyeznek): az első és utolsó adat felének, valamint a közbeeső adatok teljes értékének összegét, az adatok számától eggyel kevesebb értékkel osztjuk el:

$$\bar{x}_k = \frac{\frac{x_1 + x_n}{2} + \sum_{i=2}^{n-1} x_i}{n-1}$$

- <u>**Idősorok átlagos változásának vizsgálata:**</u>

Az átlagos változást két mutatóval mérhetjük: a fejlődés átlagos mértéke és a fejlődés átlagos üteme.

– **A fejlődés átlagos mértéke** ($\bar{d}$) az időszakról időszakra bekövetkező átlagos abszolút változást mutatja a vizsgált jelenség mértékegységében:

$$\bar{d} = \frac{x_n - x_0}{n}$$

– **A fejlődés átlagos üteme** ($\bar{l}$) az időszakról időszakra bekövetkező átlagos relatív változást mutatja. A mutató a láncviszonyszámok mértani átlaga:

$$\bar{l} = \sqrt[n]{V_{l1} \cdot V_{l2} \cdot \ldots\ldots \cdot V_{l_n}} = \sqrt[n]{\prod_{i=1}^{n} V_{li}}$$

Mivel $\prod_{i=1}^{n} V_{li} = b_{bn}$, ezért a fejlődés átlagos üteme az n-edik bázisviszonyszámból is kiszámítható: $V_l = \sqrt[n]{V_{bn}}$.

Továbbá $V_{bn}=\frac{x_n}{x_0}$ miatt a fejlődés átlagos üteme az idősor első és utolsó adatából az $V_l=\sqrt[n]{\frac{x_n}{x_0}}$ alakban is számítható.

4.3 Helyzeti középértékek

A statisztikai adatok rangsorban elfoglalt helyzetük alapján töltik be ezt a szerepkörüket. Meghatározásuk kijelöléssel történik, de eltérő módon egyszerű és súlyozott formában.

- **Medián** (Me): a rangsorba rendezett adatok közül a középső elemet mediánnak nevezzük. A medián tehát az az érték, amitől az adatok fele kisebb, mások fele pedig nagyobb.

  Sorszáma:

  - **páratlan tagszám** esetén a középső adat: $S_{Me}=\frac{n+1}{2}$
  - **páros tagszám** esetén a két középső adat egyszerű számtani átlaga.

A medián meghatározása osztályközös (egyenlő szélességű) gyakorisági sorok esetén:

- meghatározzuk a sorszámot: $S_{Me}=\frac{n}{2}$
- megkeressük azt az osztályközt, amelyben az $\frac{n}{2}$ sorszámú adat található $\rightarrow$ mediánt tartalmazó osztály (r-edik osztály);
- az osztályköz alsó értékéhez (x_{r_0}) hozzáadjuk az osztályköz arányos terjedelmét, amely: a sorszám és a mediánt megelőző osztály kumulált

gyakoriságának a különbsége, osztva a mediánt tartalmazó osztály gyakoriságával, szorozva az osztályszélességgel:

$$Me = x_{r0} + \frac{S_{Me} - f'_{r-1}}{f_r} \cdot i$$

- **Módusz** (Mo): a leggyakrabban előforduló elemet jelenti.

Osztályközös gyakorisági sorokban a leggyakoribb osztályt modális osztálynak nevezzük (r-edik), és azon belül keressük a móduszt, mégpedig úgy, hogy az osztályköz alsó értékéhez hozzáadjuk az osztályköz arányos terjedelmét. Ennek értéke: a modális és a megelőző osztály gyakoriságának a különbsége, valamint ugyanez, plusz a modális és követő osztály gyakorisága különbségének hányadosa, az osztályszélességgel szorozva:

$$M_O = x_{r0} + \frac{f_r - f_{r-1}}{(f_r - f_{r-1}) + (f_r - f_{r+1})} \cdot i$$

- **Kvantilisek** (egyéb osztóértékek): Legyen $0<q<1$. Ha a rangsorba rendezett sokaságot egy ismérvérték q:(1-q) arányban osztja ketté, akkor ezt az ismérvértéket q-ad rendű vagy q-adik kvantilisnek nevezzük.

A **leggyakrabban előforduló kvantiliseke**t külön névvel és jelöléssel is illetjük:

- Tercilisek: T_1(alsó tercilis), T_2(felső tercilis),
- Kvartilisek: Q_1(alsó kvartilis, Q_2=Me(medián), Q_3(felső kvartilis),
- Kvintilisek:K_1, K_2, K_3, K_4,
- Decilisek:D_1, D_2,….., D_9,
- Percentilisek: P_1, P_2,….., P_{99}.

A kvantilisek rangsorból való meghatározásának és osztályközös gyakorisági (relatív gyakorisági) sorból történő becslésének menete azonos a mediánnál ismertetett eljárással.

Kvartilisek fajtái: Q_1 alsó sorszáma: $\frac{n}{4}$

Q_2 középső (Me) sorszáma: $\frac{n}{2}$

Q_3 felső sorszáma: $\frac{3n}{4}$

$$Q_1 = x_{r0} + \frac{S_{q1} - f'_{r-1}}{f_r} \cdot i$$

$$Q_3 = x_{r0} + \frac{S_{q3} - f'_{r-1}}{f_r} \cdot i$$

4.4 Mintafeladat

Egy hétköznap egy egyetemi büfében megvizsgálták a vásárlások értékét. Ötven vásárlót kérdeztek meg, hogy ki mennyit költött:

A vásárlások értéke (Ft)	Vásárlók száma (fő)
0–500	7
501–1000	10
1001-1500	18
1501–2000	12
2001 felett	3
Σ	**50**

Jellemezze középértékekkel a vásárlási értékek eloszlását!

Megoldás: Becsült számtani átlagot, móduszt, mediánt, illetve alsó- és felső kvartilist számolunk.

x_i	f_i	u_i / x_i	$f_i \cdot x_i$	d_i	$f_i d_i$	f_i'
0–500	7	250	1750	-2	-14	7
501–1000	10	750	7500	-1	-10	17
1001-1500	18	1250	22500	0	0	35
1501–2000	12	1750	21000	1	12	47
2001 felett	3	2250	6750	2	6	50
Σ	**50**	**-**	**59500**	**-**	**4**	**-**

Becsült számtani átlagot kétféle módon is számolhatunk:

$$\bar{x}_a = \frac{\sum f \cdot x_i}{\sum f_i} = \frac{59500}{50} = 1190\text{Ft}$$

$$\bar{x} = a + \frac{\sum f_i \cdot d_i}{\sum f_i} \cdot i = 1250 + \frac{4}{50} \cdot 500 = 1190\,\text{Ft}$$

Egy vásárló átlagosan 1190 Ft-ot költött.

$$M_o = x_{ro} + \frac{f_r - f_{r-1}}{(f_r - f_{r-1}) + (f_r - f_{r+1})} \cdot i = 1000 + \frac{18-10}{(18-10)+(18-12)} \cdot 500 = 1285{,}7$$

Ft

A legtöbb vásárló 1286 Ft-ot költött, vagy a legtipikusabb vásárlási érték 1286 Ft volt.

$$M_e = x_{r0} + \frac{S_{Me} - \sum f'_{i-1}}{f_r} \cdot i = 1000 + \frac{25-17}{18} \cdot 500 = 1222{,}2\,\text{Ft}$$

A medián sorszáma: $S_{Me} = \frac{n}{2} = \frac{50}{2} = 25$

A vásárlók fele 1222 Ft-nál kevesebbet, a másik fele pedig 1222 Ft-nál többet költött.

$$Q_1 = x_{r0} + \frac{S_{Q1} - \sum f_{i-1}^{'}}{f_r} \cdot i = 500 + \frac{12{,}5 - 7}{10} \cdot 500 = 775 \text{Ft}$$

A felső kvartilis sorszáma: $S_{Q1} = \frac{n}{4} = \frac{50}{4} = 12{,}5$

A vásárlók negyede 775 Ft-nál kevesebbet, háromnegyede pedig 775 Ft-nál többet költött.

$$Q_3 = x_{r0} + \frac{S_{Q3} - \sum f_{i-1}^{'}}{f_r} \cdot i = 1500 + \frac{37{,}5 - 35}{12} \cdot 500 = 1604{,}2 \text{ Ft}$$

Az alsó kvartilis sorszáma: $S_{Q3} = \frac{3 \cdot n}{2} = \frac{3 \cdot 50}{2} = 37{,}5$

A vásárlók háromnegyede 1604 Ft-nál kevesebbet, negyede pedig 1604 Ft-nál többet költött.

4.5 Ellenőrző kérdések

Ismertesse a középértékek fogalmát és jellemzőit!

Csoportosítsa a középértékeket!

Határozza meg a számtani átlag fogalmát!

Hogyan számolunk számtani átlagot egyszerű és súlyozott formában?

Ismertesse a számtani átlag tulajdonságait!

Határozza meg a mértani átlag fogalmát!

Hogyan számolunk mértani átlagot egyszerű és súlyozott formában?

Határozza meg a négyzetes átlag fogalmát!

Hogyan számolunk négyzetes átlagot egyszerű és súlyozott formában?

Határozza meg a harmonikus átlag fogalmát!

Hogyan számolunk harmonikus átlagot egyszerű és súlyozott formában?

Hogyan átlagoljuk a tartamidősorok adatait?

Hogyan átlagoljuk az állapotidősorok adatait?

Mit értünk a fejlődés átlagos mértéke alatt?

Mit értünk a fejlődés átlagos üteme alatt?

Értelmezze a medián fogalmát!

Hogyan határozzuk meg a mediánt páros, illetve páratlan számú adatok esetén?

Mit nevezünk módusznak?

Mik a kvantilisek? Nevezze meg a kvantilisek főbb fajtáit!

5 Szóródási viszonyok elemzése

Szóródáson azonos fajta számszerű értékek különbözőségét értjük. A szóródás, vagyis az értékek különbözősége egyrészt az értékek egymástól való különbözőségében, másrészt valamely középértéktől való eltérésben fejeződik ki.

SZÓRÓDÁSI MÉRŐSZÁMOK

Közelítő értékek

-A szóródás terjedelme: $R = x_{\max} - x_{\min}$

-Kvartilis eltérés: $Q_e = \dfrac{Q_3 - Q_1}{2}$

-Abszolút átlageltérés: $\delta = \dfrac{\sum \left| x_i - \overline{x_a} \right|}{n}$

-Középeltérés: $K_e = \dfrac{\sum \left| x_i - M_e \right|}{n}$

-Átlagos különbség: $G = \dfrac{\sum_{j=1}^{n} \sum_{i=1}^{n} \left| x_i - x_j \right|}{n^2}$

Egzakt mutatók

-Szórás: $\sigma = \sqrt{\dfrac{\sum \left(x_i - \overline{x_a} \right)^2}{n}}$

$$\sigma = \sqrt{\frac{\sum f_i \left(x_i - \overline{x_a} \right)^2}{\sum f_i}}$$

$$\sigma = i \cdot \sqrt{\frac{\sum f_i d_i^2}{\sum f_i} - \left(\frac{\sum f_i d_i}{\sum f_i} \right)^2}$$

-Szórásnégyzet: σ^2

-Relatív szór.: $V = \dfrac{\sigma}{\overline{x_a}} \cdot 100$

Aszimmetria mérőszámok:

-Pearson-féle mérőszám: $A_p = \dfrac{\overline{x_a} - M_o}{\sigma}$

-Bowley-féle mérőszám: $F = \dfrac{\left(Q_3 - M_e \right) - \left(M_e - Q_1 \right)}{\left(Q_3 - M_e \right) + \left(M_e - Q_1 \right)} = \dfrac{Q_3 - 2M_e + Q_1}{Q_3 - Q_1}$

-Yule-Pearson féle mutató: $A_y = \frac{3\left(\overline{x_a} - M_e\right)}{\sigma}$

Koncentráció számítása (koncentrációs együtthatóval):

-Boldrini-féle közelítés módszere: $K \approx 1 - \sum_{i=1}^{k} g_i\left(z_i^{'} + z_{i-1}^{'}\right)$

-Gini-féle közelítés módszere: $K = \frac{G}{2 \cdot \overline{\overline{x_a}}}$

5.1 Közelítő szóródási mérőszámok

A közelítő szóródási mérőszámok kiegészítő információt nyújtanak az adatok változékonyságáról.

- **A szóródás terjedelme (R):** az előforduló elemek közül a legnagyobb és legkisebb érték különbsége:

  $$R = x_{\max} - x_{\min}$$

  A mutatószám kifejezi, hogy mekkora értékközben ingadoznak az ismérvértékek. Gyakorlatban kevéssé használatos, mert csupán a két szélső értékre támaszkodik. Osztályközös gyakorisági sorból nem is számítható, hiszen szélső helyzetben előfordulhat, hogy az osztályközök határai csak jelzésértékűek, vagy nincsenek megadva, azaz nyitott osztályköz áll rendelkezésre. A gyakorlatban az elemzés során jobban használható az interkvartilis terjedelem.

- **Interkvartilis terjedelem (IQR):** a kvartilis értékek közötti távolság, ami a rangsorba rendezett elemek középső ötven százalékának elhelyezkedését mutatja:

$$IQR = Q_3 - Q_1$$

Kvartilis eltérés (Q_e): az interkvartilis félterjedelem:

$$Q_e = \frac{Q_3 - Q_1}{2}$$

- **Középeltérés (K_e):** a mediántól számított eltérések abszolút értékeinek a számtani átlaga:

$$K_e = \frac{\sum |x_i - Me|}{n}$$

- **Abszolút átlageltérés (δ):** az egyedi értékeknek a számtani átlagtól mért átlagos abszolút eltérését mutatja:

$$\delta = \frac{|x_i - \bar{x}_a|}{n}$$

Súlyozott forma:

$$\delta = \frac{\sum_{i=1}^{k} f_i |x_i - \bar{x}_a|}{\sum_{i=1}^{k} f_i}$$

- **Átlagos különbség (G):** a változékonyságot a statisztikai adatok egymástól való abszolút eltérései alapján jelzi.

  Számítása:

  – egyszerű formában: valamennyi adatnak valamennyi adattól számított abszolút különbségei összegének és az eltérések számának ($n \cdot n = n^2$) a hányadosa:

$$G = \frac{\sum_{j=1}^{n} \sum_{i=1}^{n} \left| x_i - x_j \right|}{n^2}$$

– súlyozott formában: ugyanezen eltérések súlyszorzatokkal súlyozott összege és a súlyösszegek szorzatainak ($\sum f_i \cdot \sum f_j = n^2$) a hányadosa:

$$G = \frac{\sum_{i=1}^{k} \sum_{j=1}^{k} f_i \cdot f_j \cdot \left| x_i - x_j \right|}{n^2}$$

Az átlagos különbség mérőszámát a koncentráció mértékének jellemzéséhez használjuk fel.

5.2 Egzakt mutatók

- **Szórás (σ):** az egyedi értékek átlagtól való eltéréseinek a négyzetes átlaga, vagy az átlagtól mért átlagos négyzetes eltérés. Megmutatja, hogy a vizsgált adatok közül az egyes értékek átlagosan mennyivel térnek el a vizsgált adatok számtani átlagától.

  Meghatározása:

  – egyszerű formában: a számtani átlagtól vett eltérések négyzetösszegének és az adatok számának hányadosából vont négyzetgyök érték:

$$\sigma = \sqrt{\frac{\sum \left(x_i - \overline{x_a} \right)^2}{n}}$$

  – súlyozott formában: az eltérések súlyozott négyzetösszegének és a súlyok összegének a hányadosából vont négyzetgyök érték:

$$\sigma = \sqrt{\frac{\sum f_i \left(x_i - \overline{x_a} \right)^2}{\sum f_i}}$$

  – a relatív súlyokkal számolva:

$$\sigma = \sqrt{\sum g_i \left(x_i - \bar{x}_a\right)^2}$$

A szórás tulajdonságai:

1. Ha az x_i értékekhez egy állandó „a” számot hozzáadunk (vagy levonjuk x_i-ből ezt a számot), akkor a szórás értéke nem változik:

$$\sigma_z = \sigma_{a+x} = \sigma_x$$

2. Ha az egyes x_i értékeket egy állandó számmal (b) szorozzuk (vagy osztjuk), az így kapott $z_i = b \cdot x_i$ értékek szórása megegyezik az eredeti értékek szórásának
b-szeresével (vagy b-ed részével):

$$\sigma_z = \sigma_{b \cdot x} = |b| \cdot \sigma_x$$

3. A számtani átlag négyzetes minimum tulajdonságából következik, hogy egy adott „a” értéktől számított eltérésnégyzetek számtani átlagának, illetve az eltérések négyzetes átlagának minimuma a szórásnégyzet, illetve a szórás:

$$\frac{\sum (x_i - a)^2}{n} = \text{minimum, ha } a = \bar{x} \text{, illetve}$$

$$\sqrt{\frac{\sum (x_i - a)^2}{n}} = \text{minimum, ha } a = \bar{x}.$$

Ha az állandó érték nulla (a=0), akkor a szórásnégyzet (σ^2) a négyzetes átlag négyzetének ($\bar{x}_q$) és a számtani átlag négyzetének ($\bar{x}_a^{\,2}$) a különbsége.

A szórás tehát a négyzetes átlag és a számtani átlag négyzetei különbségéből vont négyzetgyök érték:

$$\sigma = \sqrt{\bar{x}_q^{\,2} - \bar{x}_a^{\,2}}$$, amiből a számtani átlag 5. tulajdonságát is felhasználva:

$$\sigma = i \cdot \sqrt{\frac{\sum f_i d_i^2}{\sum f_i} - \left(\frac{\sum f_i d_i}{\sum f_i}\right)^2}$$

- **Relatív szórás (V):** a számtani átlaghoz viszonyított (%-osan megjelenített) szórás, amely kifejezi, hogy az egyedi értékek átlagosan hány %-kal térnek el az átlagos értéktől. Számítása: a szórás és a számtani átlag hányadosaként történik:

$$V = \frac{\sigma}{\overline{x}_a} \cdot 100$$

- **Szórásnégyzet:** σ^2

5.3 Aszimmetria viszonyok mérése

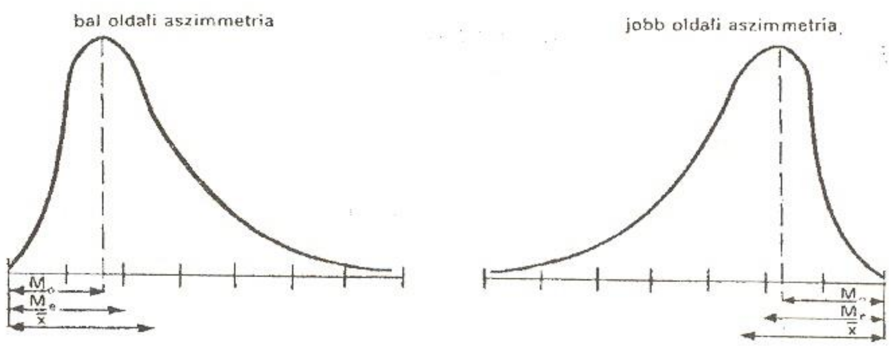
A gyakorisági sorokat ábrázolva megállapítható, hogy a görbék igen változatosak lehetnek, de nagy többségük bizonyos szabályszerűséget mutat, így besorolható néhány jellegzetes típusba.

Az eloszlás lehet:

a) egymóduszú (unimodális) eloszlás:
   - szimmetrikus ($\overline{x}_a = M_e = M_o$),
   - aszimmetrikus ($\overline{x}_a \neq M_e \neq M_o$),

b) többmóduszú (bi-, illetve polimodális) eloszlás.

Szemléltetését és az aszimmetria irányait az alábbi ábra tartalmazza:



Az aszimmetrikus eloszlások esetén vizsgáljuk az aszimmetria irányát és mértékét. A célra kiválasztott mérőszám előjele az irányt, számértéke a mértéket fejezi ki.

Az aszimmetria mérőszámai:

a) Leggyakrabban alkalmazott a Pearson-féle mérőszám (A_p):

Meghatározása: $A_p = \dfrac{\overline{x_a} - M_o}{\sigma}$

- Ha $\overline{x}_a - M_o = 0 \rightarrow$ az eloszlás szimmetrikus;
- Ha $\overline{x}_a - M_o > 0 \rightarrow$ az eloszlás baloldali;
- Ha $\overline{x}_a - M_o < 0 \rightarrow$ az eloszlás jobboldali.

Az $\overline{x}_a - M_o$ különbségének a nagyságrendjét a szórás is befolyásolja, amitől függetleníteni kell, így az aszimmetria mérőszáma a számtani átlag és a módusz különbségének a szóráshoz viszonyított aránya:

0,5-től nagyobb érték már erős aszimmetriát mutat.

b) Bowley-féle mutató (F):

$$F = \frac{(Q_3 - M_e) - (M_e - Q_1)}{(Q_3 - M_e) + (M_e - Q_1)} = \frac{Q_3 - 2M_e + Q_1}{Q_3 - Q_1}$$

c) Yule-Pearson féle mutató (A_y):

$$A_y = \frac{3(\bar{x}_a - M_e)}{\sigma}$$

5.4 Koncentráció

Koncentráción általában tömörülést, összpontosulást értünk.

Statisztikai szempontból **koncentrációnak** (K) nevezzük azt a jelenséget, amikor az ismérvértékek különbözősége következtében a kisebb ismérvértékekkel rendelkező egységekhez az értékösszeg kisebb hányada tartozik, mint amilyen ezen egységeknek a sokaság egészében elfoglalt részaránya, a sokaság nagyobb ismérvértékekkel rendelkező egységeinél pedig fordított a helyzet, azaz a sokasághoz tartozó teljes értékösszeg jelentős része a sokaság kevés egységére összpontosul.

A szakirodalom különbséget tesz **abszolút és relatív koncentráció** között. **Abszolút koncentrációról** beszélünk (a sokaság nagysága eleve kicsi), ha a sokasági értékösszeg a szó igazi értelmében is kevés egységre összpontosul. Amennyiben a sokaság nagy, akkor a kevés vagy sok egységet relatív módon (a sokaság teljes nagyságához viszonyítva) értelmezzük. **Relatív koncentráció** esetén az értékösszeg egyenlőtlenül oszlik el a sokaság egységei között. Az abszolút koncentráció az egységek nagyságát, a relatív koncentráció pedig az egységek nagyságának változékonyságát jelenti.

A koncentrációt tehát különböző mérőszámokkal jellemezhetjük.

Az abszolút koncentrációnak az alsó határa nem definiálható, míg a felső határát az jelenti, ha a sokaság egy egyedből áll, és a teljes értékösszeg hozzá tartozik.

A relatív koncentrációnak van alsó határa, mégpedig az, amikor a sokaság egyes egyedeihez az értékösszegek azonos hányada tartozik (pl. a dolgozók 20%-a a bértömegből 20%-ban részesedik, 50%-a 50%-ban, 75%-a 75%-ban stb). Felső határát az egyedek maximális különbözősége adja, pl. a sokaság egy eleme rendelkezik az értékösszeg egészével.

Az abszolút koncentráció jellemzésére a sokaság tagszámát (N) és a számtani átlagot használhatjuk. Az előző azt méri, hogy az értékösszeg hány egységhez tartozik, az átlag pedig az egységek átlagos nagyságát fejezi ki.

A relatív koncentráció mérésére használhatjuk a **Lorenz-görbét**. A Lorenz-görbe a kumulált relatív értékösszegeket a kumulált relatív gyakoriságok függvényében ábrázolja, egy olyan derékszögű koordináta rendszerben, amelynek mindkét tengelybeosztása 0-100%-ig, vagy 0-1-ig terjed. A koncentráció hiánya esetén a görbe egybeesik az átlóval. Minél távolabb esik a görbe az átlótól, annál nagyobb fokú a koncentráció.

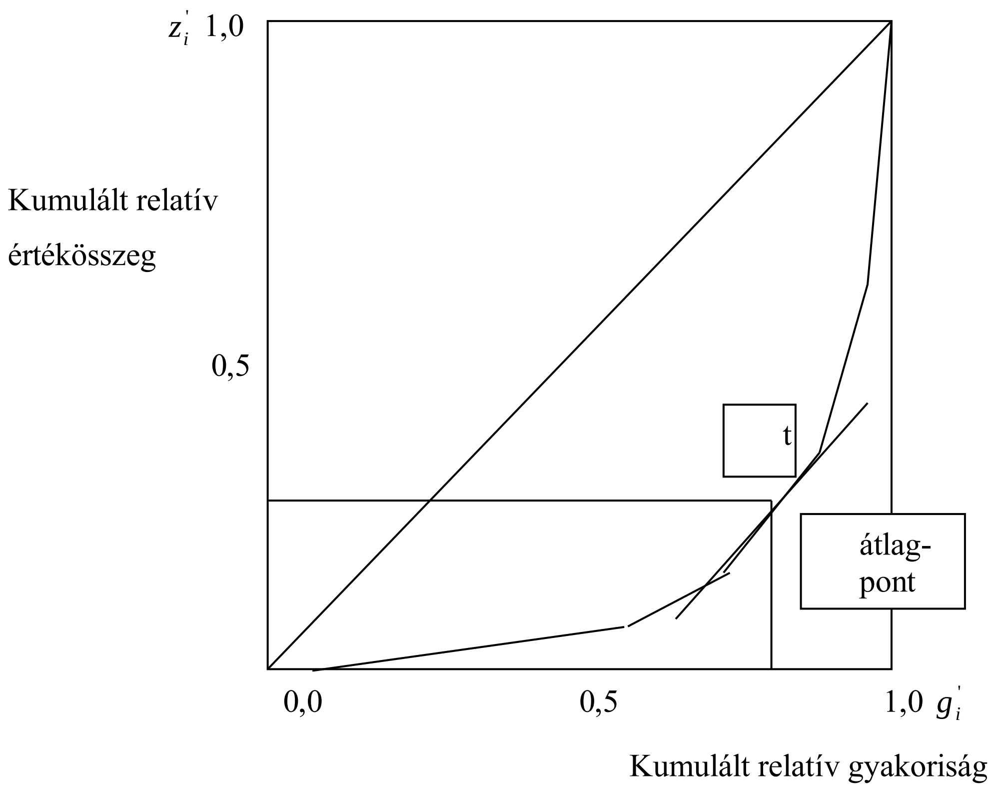
A koncentráció mértékét a 45°-os átló és az ábrázolt pontokat összekötő görbe által bezárt terület jelzi, amit **koncentrációs területnek** nevezünk. Ha a koncentrációs területet az átló és a tengelyek által bezárt háromszög területéhez viszonyítjuk, akkor e hányados alapján következtetni tudunk a koncentráció mértékére. A relatív koncentrációt tehát azáltal mérjük, hogy a koncentrációs terület nagyságát a négyzet felét képező háromszög területéhez viszonyítjuk. Az így kapott koncentrációs együttható (K) értéke 0 és 1 közé eső érték, amely azt is jelentheti, hogy a koncentrációs terület hány %-a a háromszög területének, a vizsgált jelenség koncentrációja hogyan közelíti meg a teljes koncentrációt.

Ennek meghatározására szolgál a **Boldrini-féle koncentrációs együttható**, amely közelítő érték, de ugyanazon adatokból számítható, amelyek a Lorenz-görbe szerkesztéséhez is szükségesek. Számításának képlete:

$$K \approx 1 - \sum_{i=1}^{k} g_i \left(z_i^{'} + z_{i-1}^{'} \right)$$

Közelítő pontossággal a Boldrini-féle együttható segítségével:

- osztályközönként képezzük az adott osztály és a megelőző osztály kumulált relatív értékösszeg adatainak az összegét, amit megszorzunk az adott osztályhoz tartozó relatív gyakorisági értékkel,
- az előzőekben meghatározott összegszorzatokat összesítjük és kivonjuk egyből:



A Lorenz görbe

A görbe annál nagyobb fokú koncentrációt jelez, minél nagyobb koncentrációs területet határol. Amennyiben a görbe az átlóval egybeesik (bázisvonal) nem beszélünk koncentrációról, amennyiben a tengelyekkel párhuzamossal azonos, teljes koncentrációról beszélünk. A görbe és az átlóval párhuzamos egyenes érintési pontja az átlagpont, amelynek vízszintes tengelyvetülete az átlagnál kisebb egységek sokaságon belüli aránya, a függőleges tengelyvetülete pedig az értékösszeg egészének ezen egységekhez tartozó hányada.

Egzakt mérőszám a Gini-féle koncentrációs együttható, amelynek formulája:

$$K = \frac{G}{2 \cdot \bar{x}_a}$$

A Gini-féle koncentrációs együttható az átlagos különbség és a számtani átlag kétszeresének a hányadosa:

Vannak olyan mutatószámok, amelyekkel mind az abszolút, mind pedig a relatív koncentrációt elemezhetjük. Ezek az ún. általános mutatók, amelyek közül legelterjedtebbek a:

- **CR koncentrációs arányszám** és a
- **Herfindahl-index** (HI).

A **koncentrációs arányszám** azt mutatja, hogy a sokaság néhány legnagyobb egysége (általában 3-20 egysége) hogyan részesedik a termelésből. Gyakran használatos a 3-5 legnagyobb egység kiemelése, és a relatív értékösszegből való részesedési arány megállapítása. Pl. A országban az egyik iparágban 5 vállalat működik, amelyek terelésből való részesedési aránya 8, 16, 20, 26 és 30%, a B országban ugyanezen iparágban 7 vállalat működik, amelyek termelésből való részesedése 6, 8, 10, 14, 17, 20 és 25%.

A három legnagyobb vállalat termelése alapján a koncentráció helyzete:

$$CR_A = 30 + 26 + 20 = 76\%$$

$$CR_B = 25 + 20 + 17 = 62\%$$

Tehát az A országban koncentráltabb a termelés. Arra vonatkozóan, hogy hány egységet emeljünk ki, nincs megkötés. A CR maximális értékét akkor veszi fel, ha a kiemelt egységekhez tartozik az értékösszeg egésze, minimális értéket pedig teljesen egyenletes megoszlás mellett mutat. Különösen kedvelt elemzési területe a jövedelem különbségek vizsgálata, pl. a felső tízezer fő részesedése a lakossági összjövedelemből.

A Herfindahl index az egységek értékösszegből való részesedési arányainak négyzeteit összegzi. Ennek megfelelően:

$$HI = \sum \left(\frac{x_i}{\sum x_i} \right)^2 = \sum z_i^2$$

$$\text{értéke } \frac{1}{N} \leq H \leq 1$$

$\frac{1}{N}$ az érték, ha minden egység relatív értékösszege megegyező, míg a maximális 1 érték a teljes koncentráció esetének felel meg.

Példánkban:

$$HI_A = 0{,}08^2 + 0{,}16^2 + 0{,}20^2 + 0{,}26^2 + 0{,}30^2 + = 0{,}2296$$

$$HI_B = 0{,}06^2 + 0{,}08^2 + 0{,}10^2 + 0{,}14^2 + 0{,}17^2 + 0{,}20^2 + 0{,}25^2 = 0{,}1710$$

Ezen mutató alapján is az A országban nagyobb a koncentráció. A HI nagyságát az egyedek száma (N) mellett az értékösszegek szóródása is befolyásolja, ezért a HI-nek másik számítási módja:

$$HI = \frac{V^2 + 1}{N}$$

ahol V a relatív szórás mérőszáma.

Amennyiben egy termék gyártásával 7 vállalat foglalkozik és termelésük adott hónapban: 50, 70, 180, 500, 800, 1700 és 3000 db, akkor a HI mutatója:

$$HI = \sum z_i^2 = 0{,}008^2 + 0{,}011^2 + 0{,}029^2 + 0{,}079^2 + 0{,}127^2 + 0{,}270^2 + 0{,}476^2 = 0{,}3228$$

$$\bar{x}_a = 900 \qquad \sigma = 1010 \qquad V = 1{,}122222$$

$$HI = \frac{1{,}122222^2 + 1}{7} = 0{,}3228$$

A két számítási mód azonos mutatóértéket eredményezett, amely közepes koncentráltságot mutat.

5.5 Szórás egyéb alkalmazási területei

A szórás egyéb alkalmazási területei közül fontosnak tartom kiemelni az **alternatív ismérvek** és a **heterogén sokaságok** szórásának módszertani bemutatását. Mindkét módszertan alapul szolgál a következtető statisztika egyes elemzési területeihez.

5.5.1 Alternatív ismérvek szórása

Az **alternatív ismérvek** fontos tulajdonsága, hogy két változattal rendelkeznek.

1. Először át kell alakítani az eredetileg minőségi ismérveket mennyiségi ismérvekké. Az alternatív ismérvek átalakítása mennyiségi ismérvekké úgy történik, hogy az ismérv egyik változatának 1, a másiknak 0 értéket adunk.
   - tulajdonság megléte 1
   - tulajdonság hiánya 0

Ezek után a definíció szerint már számítható az átlag és a szórás:

2. **Alternatív ismérvek átlaga:** az 1-gyel jelzett ismérvváltozat relatív gyakorisága.

$$\bar{x}_a = \frac{f_0 \cdot 0 + f_1 \cdot 1}{f_0 + f_1} = \frac{f_1}{f_0 + f_1} = g_1$$

3. A relatív súlyok felhasználásával történő szórásnégyzet meghatározása (tanult módon, lásd. szórás!):

$\sigma^2 = g_0(0-g_1)^2 + g_1(1-g_1)^2 = g_1 g_0$, ahonnan

Az alternatív ismérvek szórása: a relatív gyakoriságok mértani átlaga.

$$\sigma = \sqrt{g_1 \cdot g_0}$$

5.5.2 Heterogén sokaságok szórása

A **heterogén sokaságok** jellemzője, hogy homogén részekre bontható fel.
Nem elég egyetlen átlaggal jellemezni:

- főátlagot és
- a homogén részek átlagait is meg kell határozni.

Ha az n tagú sokaság m részre osztható, a sokaság értékei x_{ij} (i = 1, 2, ..., n; j = 1, 2, ..., m).

- **A fősokaság átlaga:** az összes előforduló érték összegének és számának a hányadosa:

$$\bar{\bar{x}} = \frac{\sum_{j=1}^{m} \sum_{i=1}^{n} x_{ij}}{\sum_{j=1}^{m} \sum_{i=1}^{n} f_{ij}}, \text{ vagy}$$

- **A fősokaság átlaga:** a részsokaságok átlagainak a részek nagyságaival súlyozott átlaga:

- részátlag: $\bar{x}_j = \dfrac{\sum_{j=1}^{m} x_j}{n_j}$
- főátlag: $\bar{\bar{x}} = \dfrac{\sum_{j=1}^{m} n_j \cdot \bar{x}_j}{\sum_{j=1}^{m} n_j}$

– **Az átlagtól vett eltérések többféleképpen értelmezhetők:**

- Belső eltérés (B): az egyedi értékek és a részátlag különbsége: $x_{ij} - \bar{x}_j$
- Külső eltérés (K): a részátlagok és a főátlag különbsége: $\bar{x}_j - \bar{\bar{x}}$
- Teljes eltérés (D): az egyedi értékek és a főátlag különbsége: $x_{ij} - \bar{\bar{x}}$

A teljes eltérés egyenlő a belső és a külső eltérés összegével. Ez az összefüggés igaz az eltérésnégyzetek összegeire, sőt azok szórásnégyzeteire is. Így:

$$\sigma^2 = \sigma_B{}^2 + \sigma_K{}^2$$

- **A belső szórásnégyzet meghatározása:**

  1. A homogén részek szórásnégyzeteinek meghatározása a tanult módon:

$$\sigma_j^2 = \frac{\sum_{i=1}^{nj} \left(x_{ij} - \bar{x}_j\right)^2}{n_j}$$

  2. A homogén részek szórásnégyzeteiből súlyozott átlagot számítunk:

$$\sigma_B^2 = \frac{\sum_{j=1}^{m} n_j \sigma_j^2}{\sum_{j=1}^{m} n_j}$$

- **A külső szórásnégyzet meghatározása:** a részátlagoknak a főátlagoktól számított eltérésnégyzetei súlyozott átlagaként:

$$\sigma_K^2 = \frac{\sum_{j=1}^{m} n_j \left(\overline{x_j} - \overline{\overline{x}}\right)^2}{\sum_{j=1}^{m} n_j}$$

- **A heterogén sokaság szórása:** $\sigma = \sqrt{\sigma_B^2 + \sigma_K^2}$

A külső, illetve belső szórásnégyzet arányát megoszlási viszonyszám formájában határozhatjuk meg, amely:

$$1 = \frac{\sigma_\mathrm{B}^2}{\sigma^2} + \frac{\sigma_\mathrm{K}^2}{\sigma^2}$$

Az összefüggésből a külső szórásnégyzet arányát szórásnégyzet-hányadosnak nevezzük, értékét pedig felhasználjuk a vegyes jellegű kapcsolatszorossági mérőszám meghatározásához. A szórásnégyzet-hányados tehát:

$$H^2 = \frac{\sigma_\mathrm{K}^2}{\sigma^2}, \text{ vagy } H^2 = 1 - \frac{\sigma_\mathrm{B}^2}{\sigma^2}$$

A szórásnégyzet-hányadosból vont négyzetgyök érték a **szóráshányados**, amely a vegyes kapcsolatok szorosságának mérőszáma:

$$H = \sqrt{\frac{\sigma_\mathrm{K}^2}{\sigma^2}}$$

$$0 \leq \mathrm{H} \leq 1$$

0-hoz közel eső érték laza jellegű kapcsolatra utal, míg 1-hez közelítve erősödik az összefüggés. Ha a két ismérv között nincs kapcsolat, a mutató értéke 0. Ez azt is jelenti, hogy a részátlagok nem szóródnak. Ebben az esetben a külső szórás 0, a belső szórás pedig egyenlő a teljes szórással. Az ellenkező szélső esetben az egyes csoportokon belül egyeznek meg az értékek, vagyis a belső szórás értéke 0, illetve a külső szórás a teljes szórással egyezik meg.

5.6 Mintafeladat

Egy vállalkozás részmunkaidőben foglalkoztatott dolgozóinak nettó kereset szerinti megoszlása 2018 októberében:

Nettó jövedelem (eFt/fő)	**Létszám (fő)**
-60,0	25
60,1–80,0	37
80,1–100,0	55
100,1–120,0	28
120,1-	10
Összesen	**155**

a) Állapítsa meg a csoportképző ismérv fajtáját!

b) Nevezze meg a statisztikai sor típusát!

c) Számítsa ki és értelmezze:

- a szóródási mérőszámokat,
- az aszimmetria mérőszámait és
- a bérkoncentrációt!

Megoldás:

a) Diszkrét mennyiségi ismérv

b) Osztályközös gyakorisági sor

c) A szóródási mérőszámok közül meghatározzuk az egzakt mutatókat, tehát szórást és relatív szórást számolunk. A gyakorisági sor egyenlő osztályszélességű.

Az aszimmetria mérőszámokhoz meg kell határozni a móduszt, a mediánt és a számtani átlagot is.

Végül a koncentrációt a Boldrini-féle együtthatóval jellemezzük.

x_i	f_i	d_i	$f_i d_i$	$f_i d_i^2$	f_i'
-60,0	25	-2	-50	100	25
60,1–80,0	37	-1	-37	37	62
80,1–100,0	55	0	0	0	117
100,1–120,0	28	1	28	28	145
120,1-	10	2	20	40	155
Összesen	**155**	**-**	**-39**	**205**	**-**

Szóródási mérőszámok meghatározása:

A szórás számításának előfeltétele a számtani átlag meghatározása.

$$\bar{x} = a + \frac{\sum f_i \cdot d_i}{\sum f_i} \cdot i = 90 + \frac{-39}{155} \cdot 20 = 84{,}97 \text{ eFt}$$

A vállalatnál egy dolgozó átlagosan 84,97 eFt-ot keres.

$$\sigma = i \cdot \sqrt{\frac{\sum f_i \cdot d_i^2}{\sum f_i} - \left(\frac{\sum f_i \cdot d_i}{\sum f_i}\right)^2} = 20 \cdot \sqrt{\frac{205}{155} - \left(\frac{-39}{155}\right)^2} = 21{,}73 \text{ eFt}$$

A dolgozók keresete átlagosan 21,73 eFt-tal tér el az átlagtól.

$$V = \frac{\sigma}{\bar{x}_a} \cdot 100 = \frac{21{,}73}{84{,}97} \cdot 100 = 25{,}57\%$$

A dolgozók keresete átlagosan 25,57%-kal tér el az átlagtól.

Aszimmetria mérőszámok meghatározása:

-Pearson-féle mérőszám: $A_P = \dfrac{\bar{x} - M_0}{\sigma}$

$$M_o = x_{ro} + \frac{f_r - f_{r-1}}{(f_r - f_{r-1}) + (f_r - f_{r+1})} \cdot i = 80 + \frac{55-37}{(55-37)+(55-28)} \cdot 20 = 88$$

$$A_P = \frac{84{,}97 - 88}{25{,}57} = -0{,}12$$

Gyenge jobboldali az aszimmetria jellemzi a keresetek eloszlását.

-Bowley-féle mérőszám:

$$F = \frac{(Q_3 - M_e) - (M_e - Q_1)}{(Q_3 - M_e) + (M_e - Q_1)} = \frac{Q_3 - 2M_e + Q_1}{Q_3 - Q_1} = \frac{99{,}73 - 2 \cdot 85{,}64 + 67{,}43}{99{,}73 - 67{,}43} = -0{,}13$$

Gyenge jobboldali az aszimmetria jellemző.

$$M_e = x_{r0} + \frac{S_{Me} - \sum f'_{i-1}}{f_r} \cdot i = 80 + \frac{77{,}5 - 62}{55} \cdot 20 = 85{,}64 \text{ eFt}$$

A medián sorszáma: $S_{Me} = \dfrac{n}{2} = \dfrac{155}{2} = 77{,}5$

$$Q_1 = x_{r0} + \frac{S_{Q1} - \sum f'_{i-1}}{f_r} \cdot i = 60 + \frac{38{,}75 - 25}{37} \cdot 20 = 67{,}43 \text{ eFt}$$

A felső kvartilis sorszáma: $S_{Q1} = \dfrac{n}{4} = \dfrac{155}{4} = 38{,}75$

$$Q_3 = x_{r0} + \frac{S_{Q3} - \sum f'_{i-1}}{f_r} \cdot i = 80 + \frac{116{,}25 - 62}{55} \cdot 20 = 99{,}73 \text{ eFt}$$

Az alsó kvartilis sorszáma: $S_{Q3} = \frac{3 \cdot n}{2} = \frac{3 \cdot 155}{2} = 116{,}25$

-Yule-Pearson féle mutató: $A_y = \frac{3\left(\bar{x}_a - M_e\right)}{\sigma} = \frac{3(84{,}97 - 85{,}64)}{25{,}57} = -0{,}08$

Az aszimmetria a Yule-Person mutató szerint is gyenge jobboldali.

A koncentráció meghatározása a Boldrini-féle együttható segítségével:

u_i: osztályközép,

$g_i = \frac{f_i}{\sum f_i}$,

$Z_i = \frac{f_i \cdot u_i}{\sum f_i \cdot u_i}$

Z_i': kumulált relatív értékösszeg

u_i	$f_i\ u_i$	g_i	z_i	$z_i^{'}$	$g_i(z_i^{'} + z_{i-1})$
70	1750	0,161	0,108	0,108	0,0174
90	3330	0,239	0,205	0,313	0,1006
110	6050	0,355	0,372	0,685	0,3543
130	3640	0,181	0,224	0,909	0,2885
150	1500	0,065	0,092	1,000	0,1241
$\sum$	**16270**	**1,000**	**1,000**	**-**	**0,8849**

Osztályközönként képezzük az adott osztály és a megelőző osztály kumulált relatív értékösszeg adatainak az összegét, amit megszorzunk az adott osztályhoz tartozó relatív gyakorisági értékkel. Majd az előzőekben meghatározott összegszorzatokat összesítjük és kivonjuk egyből:

$$K \approx 1-\sum_{i=1}^{k} g_i\left(z_i^{'}+z_{i-1}^{'}\right)=1-0{,}8849=0{,}1151$$

Alacsony fokú a koncentráció.

Mintafeladat: heterogén sokaság szórásának elemzésére egyedi adatok és osztályközös gyakorisági sor esetén

Egy szerelőüzemben dolgozó 17 főből, 7 fő segédmunkás és 10 fő szakmunkás. Egyéni termelékenységi adataik:

x_{1i}= 5, 7, 9, 11, 6, 4, 8 db/műszak

x_{2i}=7, 8, 11, 13, 10, 8, 9, 12, 13, 12 db/műszak.

$$\bar{x}_1 = 7{,}1\text{db}$$

$$\bar{x}_2 = 10{,}3 \text{ db}$$

$$\bar{\bar{x}} = 9 \text{ db}$$

A szakképzettségi csoportonkénti szórásnégyzet és szórás:

$$\sigma_1^2 = \frac{34{,}86}{7} = 4{,}98 \qquad \sigma_1 = 2{,}23 \text{ db}$$

$$\sigma_2^2 = \frac{44{,}1}{10} = 4{,}41 \qquad \sigma_2 = 2{,}1 \text{ db}$$

A belső szórásnégyzet és szórás:

$$\sigma_B^2 = \frac{7\cdot 2{,}23^2 + 10\cdot 2{,}1^2}{17} = 4{,}64 \qquad \sigma_B = 2{,}15 \text{ db}$$

A külső szórásnégyzet és szórás:

$$\sigma_K^2 = \frac{7\cdot(7{,}1-9{,}0)^2 + 10\cdot(10{,}3-9{,}0)^2}{17} = 2{,}48 \qquad \sigma_K = 1{,}57 \text{ db}$$

A teljes szórásnégyzet és szórás.

$$\sigma^2 = 4{,}64 + 2{,}48 = 7{,}12 \qquad \sigma = 2{,}67 \text{ db}$$

Egy vállalatnál a fizikai és szellemi dolgozók kereset szerinti megoszlása alapján elemezzük az átlagkeresetek alakulását:

Kereseti kategória (eFt/fő)	Dolg. Létszám (fő)		d_i	$f_i \cdot d_i$		$f_i d_i^2$	
	fizikai	szellemi		fizikai	szellemi	fizikai	szellemi
60,1-100,0	52	0	0	0	0	0	0
100,1-140,0	31	26	1	31	26	31	26
140,1-180,0	28	44	2	56	88	112	176
180,1-220,0	19	32	3	57	96	171	288
220,1-260,0	7	18	4	28	72	112	288
260,1-300,0	3	5	5	15	25	75	125
Összesen	**140**	**125**	**-**	**187**	**307**	**501**	**903**

A fizikaiak átlagkereset:

$$\bar{x}_1 = 80 + \frac{187}{140} \cdot 40 = 133{,}43 \text{ eFt/fő},$$

A szellemiek átlagkeresete:

$$\bar{x}_2 = 80 + \frac{307}{125} \cdot 40 = 178{,}24 \text{ eFt/fő},$$

Az együttes átlagkereset:

$$\bar{\bar{x}} = \frac{140 \cdot 133{,}43 + 125 \cdot 178{,}24}{265} = 154{,}57 \text{ eFt/fő}.$$

A fizikai és szellemi dolgozók bérkülönbségeinek vizsgálatához a táblázati részletszámításokat a $\sum f_i d_i^2$ oszloppal kell kiegészíteni, amely a fizikaiaknál: 0+31+112+171+112+75=501 a szellemieknél: 0+26+176+288+288+125=903.

A fizikaiak szórása:

$$\sigma_1 = 40 \cdot \sqrt{\frac{501}{140} - \left(\frac{187}{140}\right)^2} = 53{,}58 \text{ eFt}$$

A szellemiek szórása:

$$\sigma_2 = 40 \cdot \sqrt{\frac{903}{125} - \left(\frac{307}{125}\right)^2} = 43{,}67 \text{ eFt}$$

A belső szórásnégyzet:

$$\sigma_B^2 = \frac{140 \cdot 53{,}58^2 + 125 \cdot 43{,}67^2}{265} = 2416{,}2$$

A belső szórás: $\sigma_B = \sqrt{2416{,}2} = 49{,}15 \text{ eFt}$

A külső szórásnégyzet:

$$\sigma_K^2 = \frac{140 \cdot (133{,}43 - 154{,}57)^2 + 125 \cdot (178{,}24 - 154{,}57)^2}{265} = 500{,}38$$

A külső szórás: $\sigma_K = \sqrt{500{,}38} = 22{,}37 \text{ eFt}$

A teljes szórásnégyzet:

$$\sigma^2 = 2356{,}25 + 500{,}38 = 2916{,}58$$

A teljes szórás:

$$\sigma = \sqrt{2916{,}58} = 54 \text{ eFt}$$

Ez utóbbi példa osztályközös gyakorisági sorban mutatta be a szórásnégyzet felbontás módszerét.

5.7 Ellenőrző kérdések

Mit értünk szóródáson?

Rendszerezze a szóródási mérőszámokat!

Rendszerezze az aszimmetria mutatókat!

Rendszerezze a koncentráció elemzési módszereit!

Mit értünk a szóródás terjedelme alatt?

Mit értünk interkvartilis terjedelem alatt?

Mit értünk középeltérés alatt?

Mit mutat meg az abszolút átlageltérés?

Mit értünk átlagos különbség alatt?

Definiálja a szórás fogalmát!

Hogyan határozzuk meg a szórást egyszerű és súlyozott formában?

Mit értünk relatív szórás alatt?

Ismertesse a szórás tulajdonságait!

Rendszerezze a gyakorisági eloszlásokat!

Mikor beszélünk baloldali aszimmetriáról?

Mikor beszélünk jobboldali aszimmetriáról?

Mikor beszélünk koncentrációról?

Mit jelent az abszolút koncentráció?

Mit jelent a relatív koncentráció?

Ismertesse a Lorenz-görbe lényegét!

Mit értünk koncentrációs terület alatt?

Hogyan határozzuk meg az alternatív ismérvek átlagát?

Hogyan határozzuk meg az alternatív ismérvek szórását?

Mit nevezünk heterogén sokaságnak?

Mit értünk a fősokaság átlaga alatt?

Mit értünk belső, külső és teljes eltérés alatt heterogén sokaság esetén?

Hogy határozzuk meg a belső szórásnégyzetet?

Hogy határozzuk meg a külső szórásnégyzetet?

Hogy határozzuk meg a heterogén sokaság szórását?

Értelmezze a szóráshányados (H) mutatót!

6 Indexszámítás

Az **indexszám** két vagy több, valamilyen szempontból együvé tartozó, de az adatok jellegét tekintve különnemű, közvetlenül nem összesíthető statisztikai adat együttes átlagos változását kifejező összetett összehasonlító viszonyszám.

Az indexek egyszerre:

- összetett viszonyszámok és
- átlagok, amelyek időbeli vagy területi változást tükrözhetnek.

Tárgyalási szempontból két nagy körre oszthatók:

1. Abszolút számokból számított indexek: **értékindexkör,**
2. Leszármaztatott számokból számított indexek: **főátlagindexkör**.

6.1 Az értékindex-kör indexei

A különnemű adatok kezelésének elsődleges problémája a közös mértékegységben történő számbavétel. Legkézenfekvőbb a pénzbeli kifejezés, amikor a mennyiség és az ár együttes alkalmazásával az értékhez jutunk.

Az érték (v) tehát a mennyiség (q) és az ár (p) szorzataként határozható meg.

A mennyiség,- illetve áradatokat értelmezhetjük:

- Bázis: q_0, p_0, illetve
- Tárgyidőszakra: q_1, p_1.

Az indexszámítás keretén belül az egyes termékekre, cikkekre vonatkozó viszonyszámokat **egyedi indexek**nek nevezzük:

- Egyedi értékindex (i_v): $i_v = \frac{q_1 \cdot p_1}{q_0 \cdot p_0} = \frac{v_1}{v_0}$
- Egyedi árindex (i_p): $i_p = \frac{p_1}{p_0}$
- Egyedi volumenindex (i_q): $i_q = \frac{q_1}{q_0}$

Az indexszámításban az alapformát **aggregát** formának nevezzük. Az egymással szorosan összefüggő érték-, ár- és volumenindexből álló együttest **értékindexkörnek** nevezzük.

Az értékindexkör indexei:

- Értékindex (I_v),
- Árindex (I_p) és
- Volumenindex (I_q).

6.1.1 Értékindex

Az **aggregát értékindex** (I_v) a termékek, cikkek összességére nézve a termelési (eladási stb.) érték együttes, átlagos változását mutatja. Számítása két különböző időszak aggregátumának a hányadosaként történik, de az értékindex az egyedi indexek v_0-lal súlyozott számtani, illetve v_1-gyel súlyozott harmonikus átlagaként is kiszámítható:

$$I_v = \frac{\sum q_1 \cdot p_1}{\sum q_0 \cdot p_0} = \frac{\sum v_1}{\sum v_0} = \frac{\sum v_0 \cdot i_v}{\sum v_0} = \frac{\sum v_1}{\sum \frac{v_1}{i_v}}$$

A formulák az értékváltozást tükrözik. Nagyságára két tényező, a volumen és az árváltozás van hatással, amelynek mértékét meg akarjuk ismerni. Ezt mutatják a következő indexek:

6.1.2 Árindex

Az **aggregát árindex** (I_p): a különböző termékek, árucikkek árainak együttes, átlagos változását, röviden árszínvonal változását mutatja. A meghatározás során vagy a tárgy, vagy a bázisidőszakban fiktív volumeneket tételezünk fel, és így képezünk aggregátumokat. Az aggregát forma mellett azonban az egyedi árindexek súlyozott számtani, illetve harmonikus átlagaként is kiszámítható az árindex. Ennek megfelelően:

Bázisidőszaki súlyozású

$$I_p^0 = \frac{\sum q_0 \cdot p_1}{\sum q_0 \cdot p_0} = \frac{\sum v_0 \cdot i_p}{\sum v_0} = \frac{\sum q_0 \cdot p_1}{\sum \frac{q_0 \cdot p_1}{i_p}}$$

Tárgyidőszaki súlyozású

$$I_p^1 = \frac{\sum q_1 \cdot p_1}{\sum q_1 \cdot p_0} = \frac{\sum q_1 \cdot p_0 \cdot i_p}{\sum q_1 \cdot p_0} = \frac{\sum v_1}{\sum \frac{v_1}{i_p}}$$

6.1.3 Volumenindex

Az **aggregát volumenindex** (I_q): a különböző termékekből termelt, eladott, forgalmazott vagy fogyasztott mennyiségek együttes átlagos változását mutatja. Meghatározása úgy történik, hogy a tárgy vagy a bázisidőszakban fiktív árakat tételezünk fel és így képezünk aggregátumokat. De hasonlóan az érték,- illetve árindexhez, kiszámítható az egyedi volumenindexek súlyozott számtani, illetve harmonikus átlagaként is. Ennek megfelelően:

Bázisidőszaki súlyozású

$$I_q^0 = \frac{\sum q_1 \cdot p_0}{\sum q_0 \cdot p_0} = \frac{\sum v_0 \cdot i_q}{\sum v_0} = \frac{\sum q_1 \cdot p_0}{\sum \frac{q_1 \cdot p_0}{i_q}}$$

Tárgyidőszaki súlyozású

$$I_q^1 = \frac{\sum q_1 \cdot p_1}{\sum q_0 \cdot p_1} = \frac{\sum q_0 \cdot p_1 \cdot i_q}{\sum q_0 \cdot p_1} = \frac{\sum v_1}{\sum \frac{v_1}{i_q}}$$

A kétféle súlyozású ár- és volumenindex számszerű értéke nem azonos, elnevezésük is különböző:

- bázisidőszaki súlyozású: **Laspeyres-formula**,
- tárgyidőszaki súlyozású: **Paasche-formula**.

6.1.4 Index-összefüggések

Az egyes cikkekre vonatkozó $v = q \cdot p$ szorzatszerű összefüggés a megfelelő egyedi indexek között is fennáll:

$$i_v = i_q \cdot i_p$$

A megfelelő indexek is hasonló módon kapcsolódnak egymáshoz, de annak feltétele a volumen és árindexek eltérő súlyozása:

$$I_v = I_q^0 \cdot I_p^1, \text{ vagy } I_v = I_q^1 \cdot I_p^0$$

Ezen összefüggések ismeretének az a gyakorlati jelentősége, hogy bármely két index ismeretében kiszámítható a harmadik:

$$I_v = I_q^0 \cdot I_p^1, \text{ ebből } I_q^0 = \frac{I_v}{I_p^1}, \text{ illetve } I_p^1 = \frac{I_v}{I_q^0}, \text{ vagy}$$

$$I_v = I_q^1 \cdot I_p^0, \text{ ebből } I_q^1 = \frac{I_v}{I_p^0}, \text{ illetve } I_p^0 = \frac{I_v}{I_q^1}$$

Az eltérő súlyozások különböző eredményei keresztezett indexformulák születéséhez vezettek. **Fisher** a mértani átlag formulát alkalmazva kreálta az:

$I_p^F = \sqrt{I_{p0} \cdot I_{p1}}$, illetve az $I_q^F = \sqrt{I_{q0} \cdot I_{q1}}$ indexeket.

Így $I_v = I_q^F \cdot I_p^F$, ebből $I_q^F = \frac{I_v}{I_p^F}$, illetve $I_p^F = \frac{I_v}{I_q^F}$

A gyakorlati elemző munkában a hányadosok helyett az aggregátumok különbségeinek is konkrét tartalmat tulajdonítunk. Ennek megfelelően az:

- Értékváltozás: $K_v = \sum q_1 \cdot p_1 - \sum q_0 \cdot p_0$
- Árváltozások okozta értékkülönbségek:

$$K_p^0 = \sum q_0 \cdot p_1 - \sum q_0 \cdot p_0$$

$$K_p^1 = \sum q_1 \cdot p_1 - \sum q_1 \cdot p_0$$

$$K_p^F = \frac{K_p^0 + K_p^1}{2}$$

- Volumenváltozások okozta értékkülönbségek:

$$K_q^0 \sum q_1 \cdot p_0 - \sum q_0 \cdot p_0$$

$$K_q^1 = \sum q_1 \cdot p_1 - \sum q_0 \cdot p_1$$

$$K_q^F = \frac{K_q^0 + K_q^1}{2}$$

- A különbségek közötti összefüggések:

$$K_v = K_q^0 + K_p^1$$

$$K_v = K_q^1 + K_p^0$$

$$K_v = K_p^F + K_q^F$$

6.1.5 Indexsorok

Az indexszámok eddigi tárgyalása során mindig két időszak adatait hasonlítottuk össze. Az összehasonlítandó időszakok száma azonban természetesen kettőnél több is lehet. Ilyenkor nem egy indexet kell számítani, hanem az indexeknek egész összefüggő sorát, indexsort kell készíteni. Ezeket három szempont szerint lehet csoportosítani:

1. Az indexsorok abból a szempontból, hogy milyen jelenség változását mutatják, lehetnek:
   - érték,
   - ár és
   - volumenindexsorok.
2. Az időszakok összehasonlításának rendje szempontjából megkülönböztetünk:
   - bázis és
   - láncindexsorokat.
3. A súlyozás módja szempontjából beszélünk:
   - állandó súlyú és
   - változó súlyú indexsorokról.

Az értékindexsor:

- bázisindexei:

$$\frac{\sum q_{1i}\cdot p_{1i}}{\sum q_{0i}\cdot p_{0i}};\quad \frac{\sum q_{2i}\cdot p_{2i}}{\sum q_{0i}\cdot p_{0i}}\cdots\frac{\sum q_{ki}\cdot p_{ki}}{\sum q_{0i}\cdot p_{0i}}\cdots\frac{\sum q_{ni}\cdot p_{ni}}{\sum q_{0i}\cdot p_{0i}}$$

- láncindexei:

$$\frac{\sum q_{1i}\cdot p_{1i}}{\sum q_{0i}\cdot p_{0i}};\quad \frac{\sum q_{2i}\cdot p_{2i}}{\sum q_{1i}\cdot p_{1i}}\cdots\frac{\sum q_{ki}\cdot p_{ki}}{\sum q_{(k-1)i}\cdot p_{(k-1)i}}\qquad\frac{\sum q_{ni}\cdot p_{ni}}{\sum q_{(n-1)i}\cdot p_{(n-1)i}}$$

A súlyozás módja szerint állandó súlyú és változó súlyú indexsorokat különböztetünk meg. Ez a megkülönböztetés a súlyadatok rendjét, vagyis volumenindexek esetében az ár, árindexek esetében pedig a volumenadatok rendjét jelenti.

Az állandó súlyú indexsorok esetében a súlyok mindvégig változatlanok, a változó súlyú indexsorban viszont az egyes tagok más-más időszakok súlyadatait kapják. Ez utóbbi megkülönböztetés csak a volumen-és árindexek esetében lehetséges.

Meghatározásukhoz az indexekről és a dinamikus viszonyszámokról tanultak kerülnek együttes alkalmazásra.

6.2 A főátlagindexkör indexei

Kiinduló alap a heterogén sokaságok elemzésének azon sajátossága, hogy a homogén részeit külön kell jellemeznünk átlaggal, és ezek súlyozott átlagai alapján képezhetjük a heterogén sokaság főátlagát. A főátlag időbeli változását összetett dinamikus viszonyszámokkal elemezhetjük, a tanultak szerint:

$$I_{\bar{x}} = \frac{\dfrac{\sum f_1 \cdot x_1}{\sum f_1}}{\dfrac{\sum f_0 \cdot x_0}{\sum f_0}} = \frac{\bar{x}_1}{\bar{x}_0}$$

A formulából látható, hogy a főátlag értékét két tényező határozza meg:

1. a részátlagok nagyság (x_1, x_2)
2. a fősokaság összetétele (f_0, f_1)

Ebből következik, hogy az összehasonlított főátlagok hányadosát is ez a két tényező befolyásolja.

Az egyes tényezők szerepének pontos, számszerű jellemzésére a standardizálás módszere szolgál.

Standardizálás segítségével a főátlagindexet ($I_{\bar{\bar{x}}}$) két részre bontjuk:

1. Részátlagindexre ($I_{\bar{x}}$) és
2. Összetételhatás indexre ($I_{ö}$).

Az $I_{\bar{\bar{x}}}$, $I_{\bar{x}}$, $I_{ö}$, indexek együttesét **főátlagindexkörnek** nevezzük.

A **standardizálás** lényege, hogy a főátlagokat a részátlagok súlyozott átlagaként kiszámítjuk oly módon, hogy a két tényező valamelyike szempontjából összehasonlíthatóvá tesszük azokat, majd az így kapott standardizált főátlagokat hasonlítjuk össze. Összehasonlíthatóvá úgy tesszük a főátlagokat, hogy az éppen vizsgált komponens eredeti adatsorai mellett a másik összetevő tényleges adatsorai helyett standard, „állandó", mindkét főátlag számításánál rendre megegyező adatsort használunk fel.

6.2.1 Főátlagindex

A **főátlagindex** a heterogén sokaság átlagos színvonaljellemzőjének dinamikus változását mutatja. Meghatározása a tényleges főátlagok hányadosaként történik:

$$I_{\bar{\bar{x}}} = \frac{\bar{\bar{x}}_1}{\bar{\bar{x}}_0} = \frac{\dfrac{\sum f_1 \cdot x_1}{\sum f_1}}{\dfrac{\sum f_0 \cdot x_0}{\sum f_0}} = \frac{\sum f_1 \cdot x_1}{\sum \dfrac{f_1 \cdot x_1}{x_1}} \div \frac{\sum f_0 \cdot x_0}{\sum \dfrac{f_0 \cdot x_0}{x_0}}$$

6.2.2 Részátlagindex

A **részátlagindex** a részátlagok megváltozásának a főátlag változására gyakorolt hatását fejezi ki. Megmutatja, hogyan változott volna a főátlag, ha a változás kizárólag a részátlagok megváltozásából adódott volna. A részátlagindexet tehát változó részátlagokkal és állandó összetétellel számított főátlagok hányadosaként számítjuk ki:

$$I^{1}_{\bar{x}} = \frac{\dfrac{\sum f_1 \cdot x_1}{\sum f_1}}{\dfrac{\sum f_1 x_0}{\sum f_1}} = \frac{\sum f_1 \cdot x_1}{\sum f_1 x_0} = \frac{\sum f_1 \cdot x_1}{\sum \dfrac{f_1 \cdot x_1}{\dfrac{x_1}{x_0}}}$$

6.2.3 Összetételhatás index

Az **összetételhatás index** a fősokaság összetételében bekövetkezett változásnak a főátlag változására gyakorolt hatását fejezi ki. Arra a kérdésre felel, hogy hogyan változott volna a főátlag, ha e változás kizárólag az összetétel megváltozásából adódott volna. Az összetételhatás index tehát olyan főátlagok hányadosaként adódik, amelyek kiszámításánál az összetétel változik, a részátlagok pedig változatlanok:

$$I^{0}_{ö} = \frac{\dfrac{\sum f_1 \cdot x_0}{\sum f_1}}{\dfrac{\sum f_0 \cdot x_0}{\sum f_0}}$$

Az értékindexkörnél ismertetett szorzatszerű összefüggés itt is fennáll:

$$I_{\bar{\bar{x}}} = I^{1}_{\bar{x}} \cdot I^{0}_{ö}, \text{ ebből } I^{1}_{\bar{x}} = \frac{I_{\bar{\bar{x}}}}{I^{0}_{ö}}, \text{ illetve } I^{0}_{ö} = \frac{I_{\bar{\bar{x}}}}{I^{1}_{\bar{x}}}$$

$$I_{\bar{\bar{x}}} = I^{0}_{\bar{x}} \cdot I^{1}_{ö}$$

Vannak olyan esetek, amikor az elemzési céljaink megkívánják, és a vizsgált jelenség természete meg is engedi, hogy a két indexkör indexeit együttesen alkalmazzuk a következtetéseink alátámasztására. Ilyenkor feltétel a volumenek (súlyok) összesíthetősége, ami 5 féle index meghatározásának teremti meg a lehetőségeit. Azért nem 6 féle, mert észre kell vennünk, hogy a részátlagindex képlete analóg az árindex képletével, így a kettő azonos eredményhez vezet.

6.3 Mintafeladat

Értékindexkör:

Egy vállalat három termékének értékesítési adatai a következők:

Termék	Értékesített mennyiség (ezer db)		Eladási ár (Ft/db)	
	2018 jan. q0	2018 júl. q1	2018 jan. p0	2018 júl. p1
A	135	120	60	70
B	120	115	50	55
C	65	55	110	105

a.) Számítsa ki és értelmezze a vállalat értékesítési bevételének ár-, volumen- és értékindexeit a három termékre vonatkozóan!

b.) Igazolja az értékindex-kör indexei közötti összefüggések meglétét!

c.) Az értékesítés bevételének változását bontsa fel az ár és a volumenváltozás hatására!

Megoldás:
a)

Termék	q0p0	q1p1	q0p1	q1p0
A	8100	8400	9450	7200
B	6000	6325	6600	5750
C	7150	5775	6825	6050
Összesen	**21250**	**20500**	**22875**	**19000**

$$I_p^0 = \frac{\sum q_0 p_1}{\sum q_0 p_0} = \frac{22875}{21250} = 1{,}0765 \rightarrow 107{,}65\%$$

Az eladási egységárak átlagosan 7,65%-kal nőttek 2018 januárjáról 2018 júliusára bázisidőszaki volumenadatok figyelembe vételével.

$$I_p^1 = \frac{\sum q_1 p_1}{\sum q_1 p_0} = \frac{20500}{19000} = 1{,}0789 \rightarrow 107{,}89\%$$

Az eladási egységárak átlagosan 7,89%-kal nőttek 2018 januárjáról 2018 júliusára tárgyidőszaki volumenadatok figyelembe vételével.

$$I_p^F = \sqrt{I_p^0 \cdot I_p^1} = \sqrt{1{,}0765 \cdot 1{,}0789} = 1{,}0777 \rightarrow 107{,}77\%$$

Az eladási egységárak átlagosan 7,77%-kal nőttek 2018 januárjáról 2018 júliusára mindkét időszak volumenadatait figyelembe véve.

$$I_q^0 = \frac{\sum q_1 p_0}{\sum q_0 p_0} = \frac{19000}{21250} = 0{,}8941 \rightarrow 89{,}41\%$$

Az értékesítés volumene átlagosan 10,59%-kal csökkent 2018 januárjáról 2018 júliusára bázisidőszaki áradatok figyelembe vételével.

$$I_q^1 = \frac{\sum q_1 p_1}{\sum q_0 p_1} = \frac{20500}{22875} = 0{,}8962 \rightarrow 89{,}62\%$$

Az értékesítés volumene átlagosan 10,38%-kal csökkent 2018 januárjáról 2018 júliusára tárgyidőszaki áradatok figyelembe vételével.

$$I_q^F = \sqrt{I_q^0 \cdot I_q^1} = \sqrt{0{,}8941 \cdot 0{,}8962} = 0{,}8951 \rightarrow 89{,}51\%$$

Az értékesítés volumene átlagosan 10,49%-kal csökkent 2018 januárjáról 2018 júliusára mindkét időszak áradatait figyelembe véve.

$$I_v = \frac{\sum q_1 p_1}{\sum q_0 p_0} = \frac{20500}{21250} = 0{,}9647 \rightarrow 96{,}47\%$$

A vállalat értékesítési bevétele 3,53%-kal csökkent 2018 januárjáról 2018 júliusára.

b)

$$I_v = I_q^0 \cdot I_p^1 = 0{,}8941 \cdot 1{,}0789 = 0{,}9647 \rightarrow 96{,}47\%$$

$$I_v = I_q^1 \cdot I_p^0 = 0{,}8962 \cdot 1{,}0765 = 0{,}9647 \rightarrow 96{,}47\%$$

$$I_v = I_q^F \cdot I_p^F = 0{,}8951 \cdot 1{,}0777 = 0{,}9647 \rightarrow 96{,}47\%$$

c)

Értékváltozás: $K_v = \sum q_1 \cdot p_1 - \sum q_0 \cdot p_0 = 20500 - 21250 = -750$

A vállalat bevétele 2018 januárjáról 2018 júliusára 750 eFt-tal csökkent.

Árváltozások okozta értékkülönbségek:

$$K_p^0 = \sum q_0 \cdot p_1 - \sum q_0 \cdot p_0 = 22875 - 21250 = 1625$$

Az árváltozás hatására a vállalat bevétele 2018 januárjáról 2018 júliusára 1625 eFt-tal nőtt bázisidőszaki súlyozás mellett.

$$K_p^1 = \sum q_1 \cdot p_1 - \sum q_1 \cdot p_0 = 20500 - 19000 = 1500$$

Az árváltozás hatására a vállalat bevétele 2018 januárjáról 2018 júliusára 1500 eFt-tal nőtt tárgyidőszaki súlyozás mellett.

$$K_p^F = \frac{K_p^0 + K_p^1}{2} = \frac{1625+1500}{2} = 1562{,}5$$

Az árváltozás hatására a vállalat bevétele 2018 januárjáról 2018 júliusára 1562,5 eFt-tal nőtt mindkét időszak súlyadatait figyelembe véve.

Volumenváltozások okozta értékkülönbségek:

$$K_q^0 \sum q_1 \cdot p_0 - \sum q_0 \cdot p_0 = 19000 - 21250 = -2250$$

A volumenváltozás hatására a vállalat bevétele 2018 januárjáról 2018 júliusára 2250 eFt-tal csökkent bázisidőszaki súlyozás mellett.

$$K_q^1 = \sum q_1 \cdot p_1 - \sum q_0 \cdot p_1 = 20500 - 22875 = -2375$$

A volumenváltozás hatására a vállalat bevétele 2018 januárjáról 2018 júliusára 2375 eFt-tal csökkent bázisidőszaki súlyozás mellett.

$$K_q^F = \frac{K_q^0 + K_q^1}{2} = \frac{-2250 - 2375}{2} = -2312{,}5$$

A volumenváltozás hatására a vállalat bevétele 2018 januárjáról 2018 júliusára 2312,5 eFt-tal csökkent mindkét időszak súlyadatait figyelembe véve.

A különbségek közötti összefüggések:

$$K_v = K_q^0 + K_p^1 = \text{-}2250 + 1500 = \text{-}750$$

$$K_v = K_q^1 + K_p^0 = \text{-}2375 + 1625 = \text{-}750$$

$$K_v = K_q^F + K_p^F = \text{-}2312{,}5 + 1562{,}5 = \text{-}750$$

Főátlagindexkör:

Egy vállalkozás részmunkaidőben foglalkoztatott dolgozóinak iskolai végzettség szerinti megoszlásáról és kereseti viszonyainak alakulásáról a következő információink vannak:

Iskolai végzettség	2017.		2018.	
	Létszám (Fő) f0	Átlagkereset (eFt/fő) x0	Létszám (fő) f1	Átlagkereset (eFt/fő) x1
Alapfok	30	76,0	35	83,0
Középfok	82	90,0	85	95,0
Felsőfok	28	120,0	20	135,0
Összesen	140	…	140	…

Állapítsa meg az átlagos kereseti színvonalváltozás indexét és elemezze a változás okait!

Megoldás:

Az átlagkereset 2017-ben:

$$\bar{x}_0 = \frac{\sum f_0 \cdot x_0}{\sum f_0} = \frac{30 \cdot 76 + 82 \cdot 90 + 28 \cdot 120}{140} = \frac{13020}{140} = 93 eFt / fő$$

Az átlagkereset 2018-ban:

$$\bar{x}_1 = \frac{\sum f_1 \cdot x_1}{\sum f_1} = \frac{35 \cdot 83 + 85 \cdot 95 + 20 \cdot 135}{140} = \frac{13680}{140} = 97{,}71 eFt / fő$$

Főátlagindex:

$$I_{\bar{\bar{x}}} = \frac{\bar{x}_1}{\bar{x}_0} = \frac{\dfrac{\sum f_1 \cdot x_1}{\sum f_1}}{\dfrac{\sum f_0 \cdot x_0}{\sum f_0}} = \frac{\dfrac{35\cdot 83 + 85\cdot 95 + 20\cdot 135}{140}}{\dfrac{30\cdot 76 + 82\cdot 90 + 28\cdot 120}{140}} = \frac{13680}{13020} = 1{,}0507 \rightarrow 105{,}07\%$$

Részátlagindex:

$$I^{1}_{\bar{x}} = \frac{\dfrac{\sum f_1 \cdot x_1}{\sum f_1}}{\dfrac{\sum f_1 x_0}{\sum f_1}} = \frac{\dfrac{35\cdot 83 + 85\cdot 95 + 20\cdot 135}{140}}{\dfrac{35\cdot 76 + 85\cdot 90 + 20\cdot 120}{140}} = \frac{13680}{12710} = 1{,}0763 \rightarrow 107{,}63\%$$

Összetételhatás index:

$$I^{0}_{ö} = \frac{\dfrac{\sum f_1 \cdot x_0}{\sum f_1}}{\dfrac{\sum f_0 \cdot x_0}{\sum f_0}} = \frac{\dfrac{35\cdot 76 + 85\cdot 90 + 20\cdot 120}{140}}{\dfrac{30\cdot 76 + 82\cdot 90 + 28\cdot 120}{140}} = \frac{12710}{13020} = 0{,}9762 \rightarrow 97{,}62\%$$

Elemzés:

A vállalkozás dolgozóinak átlagos kereseti színvonala 2017-ről 2018-ra 5,07%-kal nőtt. Mindhárom állománycsoport átlagbére növekedett, így változatlan létszámösszetételt feltételezve a vállalati átlagkereset 7,63%-kal nőtt. A létszám arányok a tárgyidőszakra megváltoztak, ez a vállalati átlagkereset 2,38%-os csökkenését okozta. Az első hatás (a részátlagok változásának hatása) volt erősebb a kereseti színvonal megváltozása terén.

6.4 Ellenőrző kérdések

Mit értünk indexszám alatt?

Mit értünk értékindexkör, illetve főátlagindexkör alatt?

Mit nevezünk egyedi indexnek?

Mit értünk aggregát forma alatt?

Nevezze meg az értékindexkör indexeit!

Értelmezze az értékindexet!

Értelmezze az árindexet!

Értelmezze a volumenindexet!

Ismertesse a különböző indexek közötti összefüggéseket!

Mit értünk indexsorok alatt? Rendszerezze az indexsorokat!

Mi a standardizálás lényege?

Nevezze meg a főátlagindexkör indexeit!

Értelmezze a főátlagindexet!

Értelmezze a részátlagindexet!

Értelmezze az összetételhatás indexet!

Milyen összefüggés mutatható ki a főátlagindexkör indexei között?

IRODALOMJEGYZÉK

Galó M. – Makszim Gy-né (2012): Statisztika I. Nyíregyházi Főiskola, Gazdasági és Társadalomtudományi Kar, Bessenyei Könyvkiadó, Nyíregyháza

Falus I. – Ollé J. (2008): Az empirikus kutatások gyakorlata. Nemzeti Tankönyvkiadó, Budapest

Hunyadi L.-Vita L.(2008): Statisztika I. Aula Kiadó, Budapest;

Kerékgyártó Gy-né L. Balogh I.-Sugár A.- Szarvas B. (2009): Statisztikai módszerek és alkalmazásuk a gazdasági és társadalmi elemzésekben. Aula Kiadó, Budapest