# Sokaságok típusai

**1. Diszkrét vagy folytonos:**

* **Diszkrét**: Egyértelműen elkülöníthető egységek, például:
  + Magyar népesség 2022. január 1-én: 9 689 010 fő.
  + Magyarországra érkezett külföldiek száma 2021-ben: 36 688 fő.
* **Folytonos**: Összefüggő egységek, például:
  + 2021-es teljes búzatermés: 5 316 074 tonna.
  + Belföldön közúton szállított áruk mennyisége 2021-ben: 184 218 tonna.

**2. Álló vagy mozgó:**

* **Álló**: Valamely konkrét időpontra vonatkozik (stock), például:
  + Magyar népesség 2022. január 1-én.
  + Beiratkozott hallgatók száma az IK karon 2022. szeptember 5-én.
* **Mozgó**: Időtartamra vonatkozik (flow), például:
  + 2021-es teljes búzatermés.
  + Hallgatók által elfogyasztott sör mennyisége a 2021/22-es tanév második félévében.

**3. Véges vagy végtelen:**

* **Véges**: Meghatározott számú egységgel rendelkezik, például:
  + Magyar népesség 2022. január 1-én.
* **Végtelen**: Elméletileg végtelen, például:
  + 2022 lehetséges búzatermés eredményei (eloszlással megadott fiktív sokaság).

# Aggregált sokaság

**Definíció:** Az aggregált sokaság különböző termékekből vagy szolgáltatásokból származó összesített értékeket tartalmaz.

**Példák:**

1. **Magyarország teljes exportja 2021-ben:**
   * Összérték: **42 781,5 milliárd forint**.
2. **IK hallgatói által a 2021/22-es tanév 2. félévében elfogyasztott alkoholtartalmú italok összértéke:**

# Ismérvek

**Definíció:** Az ismérvek olyan vizsgálati szempontok, amelyek alapján egy sokaság részekre bontható. Az ismérvek lehetővé teszik, hogy a sokaság egységeit adott szempontok szerint csoportosítsuk.

**Ismérvek fajtái:**

* **Területi:**
  + Például lakhely, születési hely.
* **Időbeli:**
  + Például születési idő, munkába állás időpontja.
* **Minőségi:**
  + Például nem, foglalkozás.
* **Mennyiségi:**
  + Például életkor, testmagasság, testtömeg, tanulmányi átlag.

# Mérési skálák

Az ismérvváltozatok átfordíthatók számokká, és csak olyan műveletek végezhetők, amelyek az eredeti változatokkal is összhangban vannak. A mérési szintek négy fő típusa a következő:

**1. Nominális skála**

* **Leírás:** Csak az vizsgálható, hogy két érték egyenlő-e vagy sem.
* **Példák:** Név, lakhely, foglalkozás.
* **Mértékegység:** Nincs.

**2. Ordinális skála**

* **Leírás:** Csak az értékek sorrendje számít, a közöttük lévő távolság nem.
* **Példák:** Vizsgajegyek, végzettség.
* **Mértékegység:** Nincs.

**3. Különbségi skála**

* **Leírás:** Az értékek közötti különbségek is információt hordoznak, de az arányuk nem értelmezhető.
* **Példák:** Hőmérséklet (Celsius, Kelvin, Fahrenheit).
* **Kezdőpont:** Önkéntes (választható).
* **Mértékegység:** Van (pl. fokok).

**4. Arány skála**

* **Leírás:** A kezdőpont egyértelműen adott, az arányok is értelmezhetők.
* **Példák:** Havi jövedelem, testmagasság.
* **Kezdőpont:** Egyértelmű (abszolút nulla).
* **Mértékegység:** Van.

**Kontigencia táblázat:** Olyan statisztikai eszköz, amely két vagy több kategorizált változó közötti kapcsolatot vizsgálja és ábrázolja. Ezzel a módszerrel könnyen áttekinthetővé válnak az egyes változók lehetséges kombinációi és azok előfordulási gyakoriságai. Gyakran használják a **khi-négyzet teszttel** együtt, hogy megvizsgálják, van-e statisztikailag szignifikáns kapcsolat a változók között.

# Statisztikai alapműveletek

A statisztikai alapműveletek három fő kategóriába sorolhatók, amelyek segítenek a sokaságok jellemzésében, összehasonlításában és osztályozásában.

**I. A sokaság jellemzése**

* **Leírás:** A sokaság egészének jellemzése megfelelő adatokkal vagy mutatószámokkal.
* **Példák:**
  + A sokaság nagysága.
  + Átlagos értékek (pl. átlag életkor, átlag jövedelem).
  + Várható értékek.

**II. Összehasonlítás**

* **Leírás:** A statisztikai adatok összehasonlítása különböző szempontok alapján, mint például időbeli alakulás, területi eltérések vagy egymáshoz kapcsolódó jelenségek viszonyai.
* **Fontosság:** Az adatok összehasonlíthatósága kulcsfontosságú; csak akkor vonhatunk le következtetéseket, ha az adatok azonos alapokon állnak.
* **Módszerek:** Különbségek és hányadosok képzése az összehasonlításhoz.

**III. Osztályozás**

* **Leírás:** A sokaság tagolása egy vagy több ismérv szerint, amely során a csoportok homogénebbek, mint az egész sokaság.
* **Osztályok:** Az osztályozás során kapott csoportok, amelyek az adott ismérvek alapján alakulnak ki.
* **Csoportképző ismérvek:** Az osztályok elhatárolására szolgáló ismérvek.
* **Példa:** Az évfolyam osztályozása a Mikroökonómia jegyei alapján.

# Viszonyszámok

**Definíció:** A viszonyszám két adat hányadosa, amely a következőképpen alakul:

V = A / B

* **V:** viszonyszám
* **A:** a viszonyítást tárgyát képező adat
* **B:** a viszonyítást alapját képező adat

**Típusai**

1. **Dinamikus viszonyszámok:**
   * **Leírás:** Olyan viszonyszámok, amelyeket idősorok adataiból számítanak.
   * **Példa:** Egy adott év gazdasági növekedésének arányosítása az előző évhez képest.
2. **Intenzitási viszonyszámok:**
   * **Leírás:** Két egymással kapcsolatban lévő, de nem feltétlenül azonos fajta egységekből álló sokaság nagyságából képzett hányadosok.
   * **Példa:** A születések számának és a női népesség számának aránya egy adott évben.
3. **Megoszlási viszonyszámok:**
   * **Leírás:** Valamely sokaság részének az egészhez viszonyított nagyságát mutatják.
   * **Példa:** Egy adott korcsoport aránya a teljes népességen belül.

# Grafikus ábrázolás

A statisztikai adatok grafikus ábrázolása különböző diagramok segítségével történik, amelyek megkönnyítik az adatok vizualizálását és értelmezését. Íme a leggyakoribb ábrázolási formák:

1. **Idősorok ábrázolása: Vonaldiagram**
   * **Leírás:** Az időbeli változások, trendek bemutatására szolgál.
   * **Alkalmazás:** Például a havi vagy éves értékesítési adatok, gazdasági mutatók alakulásának szemléltetése.
2. **Mennyiségi ismérvek kapcsolata: Pontdiagram**
   * **Leírás:** Két mennyiségi adat közötti kapcsolat ábrázolására használják.
   * **Alkalmazás:** Például a testmagasság és a testsúly közötti összefüggés bemutatása.
3. **Szerkezeti megoszlás ábrázolása: Osztott kör-, oszlop- vagy szalagdiagram**
   * **Leírás:** A sokaság szerkezeti összetevőit szemlélteti, mutatva az egyes részek arányát az egészhez képest.
   * **Alkalmazás:** Például a népesség korcsoportok szerkezete, vagy egy termék összetevőinek megoszlása.
4. **Mennyiségi ismérv eloszlásának ábrázolása: Hisztogram**
   * **Leírás:** A mennyiségi adatok eloszlásának vizuális reprezentációja, amely oszlopok formájában mutatja be az értékek gyakoriságát.
   * **Alkalmazás:** Például a diákok vizsgaeredményeinek eloszlása.

# Mennyiségi sorok

A mennyiségi sorok a statisztikai elemzések egyik alapvető eszköze, amely lehetővé teszi a mennyiségi ismérvek rendszerezését és a sokaság elemzését.

**Alapfogalmak**

* **Y:** Mennyiségi ismérv.
* **N:** A sokaság elemszáma.
* **Y1, Y2, ..., YN:** Az Y ismérv változatai, amelyek különbségi vagy arány skálán mért számértékek.

**Típusok**

1. **Diszkrét ismérv:**
   * **Leírás:** Olyan értékeket vehet fel, amelyek csak megszámlálható számosságúak. Az értékek pontosan megadhatók.
   * **Példák:**
     + Háztartások nagysága.
     + Családban lévő gépjárművek száma.
   * **Jellemző:** Az értékek egyértelműen és pontosan definiáltak.
2. **Folytonos ismérv:**
   * **Leírás:** Olyan értékeket vehet fel, amelyek kontinuum számosságúak, azaz bármilyen mérés eredménye lehet. Az értékek csak bizonyos pontosságra kerekítve adhatók meg.
   * **Példák:**
     + Háztartás éves jövedelme.
     + Családban lévő gépjárművek összértéke.
   * **Jellemző:** Az értékek a mérés pontosságának függvényében változnak.
3. **Folytonos kezelés diszkrét esetben:**
   * Ha a diszkrét ismérv nagyon sok értéket vehet fel, kezelhetjük folytonosként is.
   * **Példa:** Nagy városok népessége, ahol a létszámot gyakran folytonos adatokként kezeljük a statisztikai elemzések során.

**Rangsor**

* A megfigyelési egységekhez tartozó Yi ismérvértékek monoton nemcsökkenő sorrendben történő felsorolása.
* A rangsor i-edik tagját Yi∗-gal jelöljük.

# Kvantilisek

**Definíció**

Az Yi/k i-edik k-adrendű kvantilis az a szám, amelynél az összes előforduló ismérvérték legfeljebb i/k-ad része kisebb, és legfeljebb (1 – i/k)-ad része nagyobb.

**Rangsor**

Az adatsor rangsorba állítása a következőképpen történik:

* Az adatok rangsorba állítása: Y1∗,Y2∗,…,YN\*

**Kvantilisek kiszámítása**

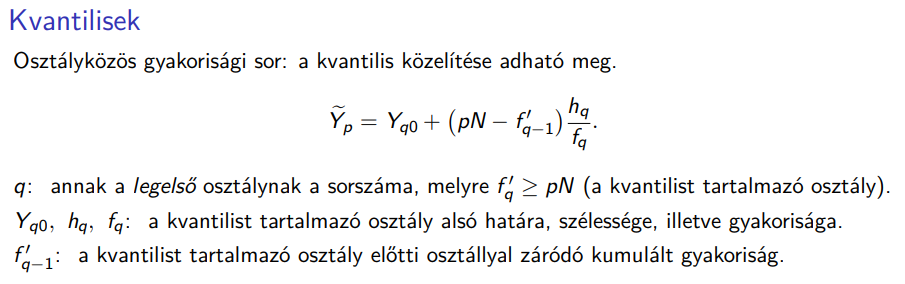
A kvantilisek kiszámítása a következőképpen történik:

1. **Számítsuk ki:** sp = p \* (N+1)
   * sp: kvantilis pozíció
   * p: a kívánt kvantilis aránya (pl. medián p=0.5)
2. **Két eset:**

A képen szöveg, Betűtípus, kézírás, kalligráfia látható

Automatikusan generált leírás

Kvantilisek kiszámítása osztályközös gyakorisági sor esetén:



# Gyakorisági eloszlások grafikus ábrázolása

**Gyakorisági eloszlások grafikus ábrázolása**

A gyakorisági eloszlások vizuális ábrázolása segíti az adatok elemzését, megértését és az eloszlás jellemzőinek gyors áttekintését. Az alábbiakban bemutatjuk a leggyakoribb grafikus ábrázolási formákat: a leveles ág ábrát, a doboz ábrát és a hisztogramot.

**1. Leveles Ág Ábra (Stem-and-Leaf Plot)**

A leveles ág ábra egy kombinált grafikus forma, amely lehetővé teszi a statisztikai adatok részletesebb bemutatását.

* **Felépítés:**
  + **Függőleges vonal:** A bal oldalon található, amely az ismérvértékek legelső helyi értékű számjegyeit (ágakat) mutatja.
  + **Levelek:** A vonal jobb oldalán az ágakhoz tartozó további számjegyek (levelek) szerepelnek, amelyek a számjegyeket szóközökkel vagy vesszőkkel választják el egymástól.

**2. Doboz Ábra (Box Plot, Box-and-Whiskers Plot)**

A doboz ábra a statisztikai adatok eloszlásának vizuális megjelenítésére szolgál, különösen a kvartilisek kiemelésére.

* **Felépítés:**
  + **Doboz:** A doboz a Q1 (első kvartilis) és a Q3 (harmadik kvartilis) közötti tartományt jelöli.
  + **Osztóvonal:** A doboz közepén található a Me (medián), amely az adatok középpontját mutatja.
  + **Bajuszok (Whiskers):** A doboz mellett a legnagyobb és legkisebb értékeket (outlier-eket) kapcsolják be.

**3. Hisztogram**

A hisztogram egy osztályközös gyakorisági sor ábrázolásának formája, amely oszlopokkal vizualizálja az adatokat.

* **Felépítés:**
  + Az osztályközök fől a vízszintes tengelyen helyezkednek el, míg a függőleges tengelyen az osztályok gyakoriságát (gyakoriság hisztogram) vagy relatív gyakoriságát (sűrűség hisztogram) tüntetik fel.
  + Minden oszlop magassága arányos az adott osztályban előforduló értékek számával.

# Helyzetmutatók (középértékek)

A képen szöveg, képernyőkép, Betűtípus, algebra látható

Automatikusan generált leírás

A képen szöveg, képernyőkép, Betűtípus látható

Automatikusan generált leírásA képen szöveg, Betűtípus, képernyőkép, kézírás látható

Automatikusan generált leírás

# Terjedelem mutatók

**Fogalom:** A terjedelem mutatók az adatok közötti távolságot mérik, azaz a legkisebb és legnagyobb érték közötti különbséget.

**Terjedelem:** A terjedelem a legnagyobb és a legkisebb érték közötti különbséget jelenti, amely megmutatja az adathalmaz szélességét.

**Interkvartilis távolság:** A harmadik kvartilis (Q3) és az első kvartilis (Q1) közötti távolságot méri. Ez az érték az adathalmaz középső 50%-ának szóródását jellemzi, így ellenállóbb a szélsőértékekkel szemben.

# Szóródási mutatók

**Tétel:** Arra szolgálnak, hogy megmutassák, mennyire szóródnak az értékek az átlag (vagy más középérték) körül.

A képen szöveg, Betűtípus, képernyőkép, sor látható

Automatikusan generált leírás

# Koncentráció

* **Fogalom:** A koncentráció azt jelenti, hogy a sokaságban lévő teljes értékösszeg jelentős része vagy egészének egy kis egységre való összpontosulása. Más szavakkal, a koncentráció azt tükrözi, hogy az adatok vagy értékek milyen mértékben csoportosulnak egy vagy néhány jellemző érték körül.

**Koncentrációs együttható**

* **Definíció:** A koncentrációs együttható a Lorenz görbe és a négyzet átlója által bezárt terület (koncentrációs terület) arányát méri a négyzet teljes területéhez viszonyítva.
* **Jele:** L

**Herfindahl index:** Egy olyan mérőszám, amely a piac koncentrációját méri, és segít értékelni, hogy egy iparág mennyire van monopolizálva vagy fragmentálva.

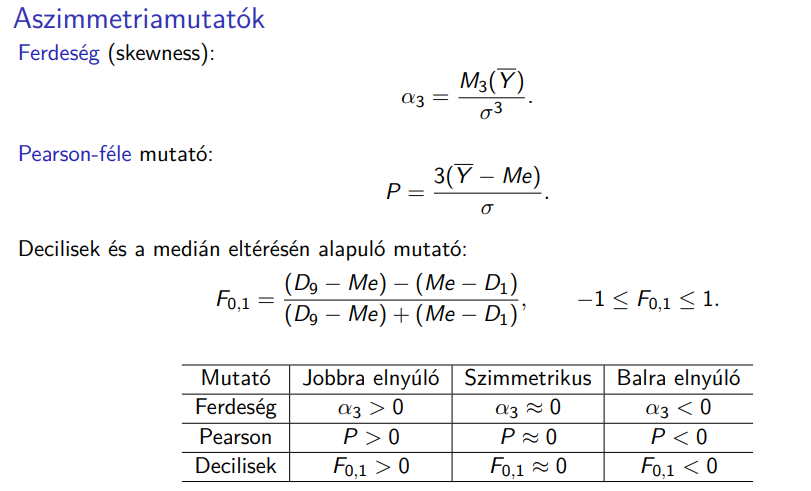
* Kiszámítása:
* HI = 1/N 🡪nincs koncentráció
* HI = 1 🡪 teljes a koncentráció

# Momentumok

A képen szöveg, Betűtípus, sor, képernyőkép látható

Automatikusan generált leírás

# Asszimetria mutatók

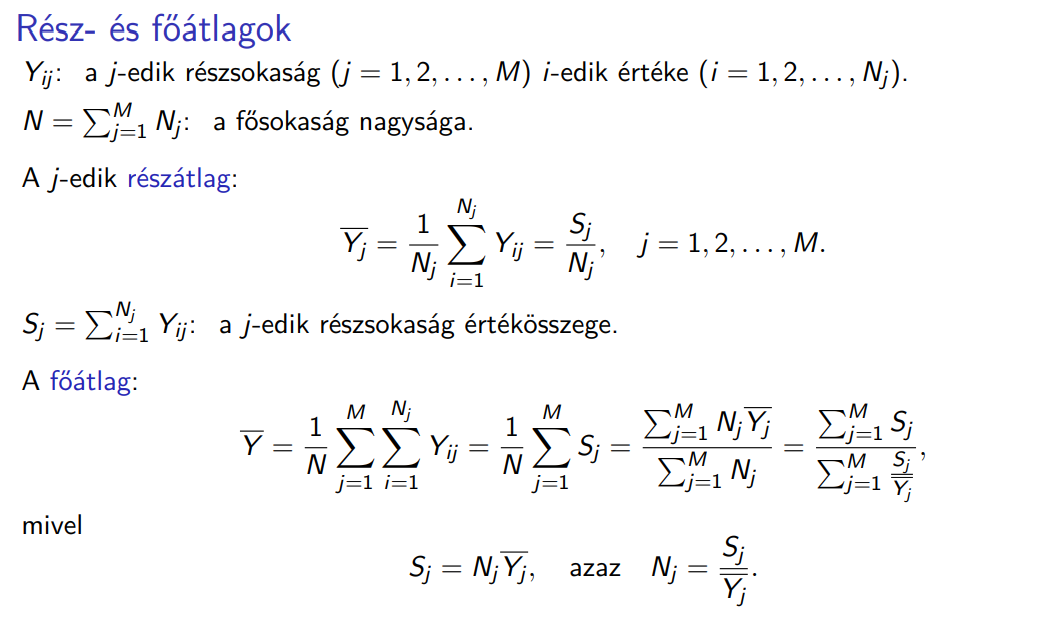


# Csúcsossági mutató

A képen szöveg, képernyőkép, Betűtípus, sor látható

Automatikusan generált leírás

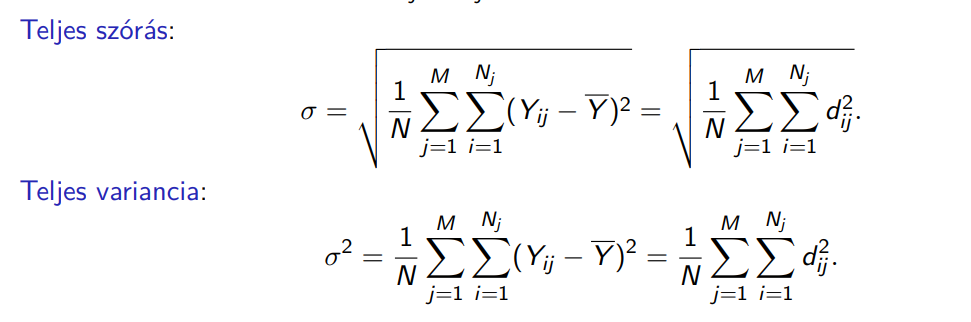
# Rész és főátlagok



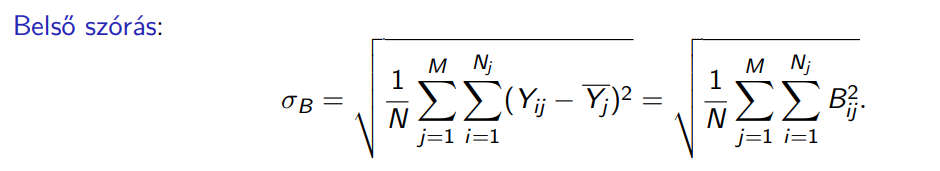
Tétel: a fősokaság átlaga a részsokaságok átlagainak súlyozott átlaga, ahol a súlyok az egyes részsokaságok elemszámai. Ha az egyes részsokaságok különböző nagyságúak, akkor azok a részátlagok, amelyek nagyobb elemszámú részsokaságra vonatkoznak, nagyobb hatással lesznek a főátlagra.

Harmonikus átlag: A harmonikus átlag egy speciális típusú átlag, amely akkor használatos, ha az értékek fordított arányban vannak egymással. Ha a vizsgált értékek valamilyen fordított kapcsolatban állnak egy mennyiséggel (például sebesség, ahol a nagyobb érték kisebb időt jelent egy adott távolság megtételére), a harmonikus átlag pontosabb képet ad, mint a számtani átlag.

# Teljes szórás és variancia



# Belső szórás és variancia

A képen szöveg, Betűtípus, fehér, képernyőkép látható

Automatikusan generált leírás

Tétel: A belső szórás azt mutatja meg, hogy a fősokaság egyes elemei mennyire térnek el a saját részátlaguktól. Ez azt méri, hogy a részsokaságokon belül mekkora eltérés van a megfigyelések között.

# Külső szórás és variancia

A képen szöveg, Betűtípus, sor, fehér látható

Automatikusan generált leírás

Tétel: A **külső szórás** azt méri, hogy az egyes részsokaságok átlagai (részátlagok) mennyire térnek el a fősokaság átlagától. Vagyis azt mutatja meg, hogy a részsokaságok közötti különbségek hogyan járulnak hozzá az egész sokaság szóródásához.

Tétel: A teljes variancia megegyezik a külső és a belső variancia összegével, kiszámítási módja: σ2 = σ2B + σ2K

# Négyzetösszegek között összefüggés

# Az ismérvek közötti kapcsolatokat az alábbi fő kategóriákba sorolhatjuk:

1. **Független kapcsolat**: Az ismérvek egymástól teljesen függetlenek, nincs köztük kapcsolat. Példa: hajszín és testmagasság.
2. **Sztochasztikus (valószínűségi) kapcsolat**: Az egyik ismérv (X) alapján következtetni lehet a másik ismérv (Y) értékére, de nem teljes bizonyossággal. Példa: hajszín és szemszín.
3. **Függvényszerű (determinisztikus) kapcsolat**: Az egyik ismérv teljesen meghatározza a másik ismérv értékét. Példa: ösztöndíjátlag és tanulmányi ösztöndíj.

# Az ismérvek kapcsolatai az ismérv típusai alapján:

* **Asszociáció**: Mindkét ismérv minőségi vagy területi (nominális skálán mért).
* **Vegyes kapcsolat**: Az egyik ismérv mennyiségi, a másik minőségi vagy területi.
* **Korreláció**: Mindkét ismérv mennyiségi (különbségi vagy arányskálán mért).
* **Rangkorreláció**: Mindkét ismérv sorrendi skálán mért.

# Tényleges és várt gyakoriság

Várt gyakoriság kiszámítása: f∗ij = fi \* fj / N

# A kapcsolat szorossága

A képen szöveg, képernyőkép, Betűtípus, szám látható

Automatikusan generált leírás

# PRE eljárás és a szorosság mérése

A **PRE (Proportional Reduction in Error)** eljárás a kapcsolat szorosságának mérésére szolgál, és azt méri, hogy mennyivel csökkenthetjük a hibatényezőt, ha ismerjük egy ismérv (X) alapján egy másik ismérv (Y) értékeit.

**PRE eljárás lépései:**

1. **Első hibamérték meghatározása** (E1): Megnézzük, mekkora hibát követnénk el, ha az Y ismérv szerinti hovatartozást csak az Y ismérv megoszlása alapján próbálnánk megjósolni (az X ismerete nélkül).
2. **Második hibamérték meghatározása** (E2): Megnézzük, mekkora lenne a hiba, ha ismernénk az X ismérv szerinti hovatartozást, és ez alapján próbálnánk meg jósolni az Y ismérvet.
3. **Relatív hibacsökkenés kiszámítása**: PRE = (E1 – E2) / E1  
    Ez azt mutatja meg, hogy mennyivel csökken a hibák aránya, ha az X ismérvet is figyelembe vesszük.
4. **PRE értelmezése:**
5. **PRE = 0**: Az X és Y ismérvek függetlenek egymástól, vagyis az X ismerete nem csökkenti a hibát.
6. **PRE = 1**: Az X és Y ismérvek között függvényszerű, determinisztikus kapcsolat van, azaz X ismerete teljesen meghatározza az Y értékeit.

# Vegyes kapcsolat A képen szöveg, képernyőkép, Betűtípus, szám látható Automatikusan generált leírás

Tétel: A PRE mutató értéke mindig 0 és 1 között mozog, és minél nagyobb, annál szorosabb az X és Y közötti kapcsolat. Ha PRE = 0 akkor X és Y függetlenek, ha PRE = 1 akkor X és Y között függvényszerű kapcsolat van.

# Variancia hányados

# Empirikus regressziófüggvény

Fogalom: Az empirikus regressziófüggvény a statisztikában egy olyan modell, amely a két mennyiségi ismérv, X és Y közötti kapcsolatot írja le, a megfigyelések (tapasztalati adatok) alapján. Lényegében azt vizsgálja, hogy hogyan változik Y értéke X változásával.

**Alapfogalmak:**

* **X és Y:** Két mennyiségi ismérv, amelyek szerepe felcserélhető.
* **X:** Csoportképző ismérv, amely alapján az osztályokat rangsorolhatjuk.

**Kapcsolat iránya:**

* **Pozitív kapcsolat:** Ha X növekedésével Y is növekszik.
* **Negatív kapcsolat:** Ha X növekedésével Y csökken.

**Empirikus regressziófüggvény:**

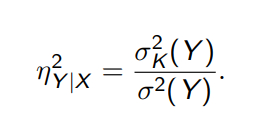
* Az X szerinti csoportosított részpopulációkhoz hozzárendelt Y részátlagok sorozata, amely az Y változó X-re vonatkozó empirikus regressziófüggvényét jelenti.

**Grafikus ábrázolás:**

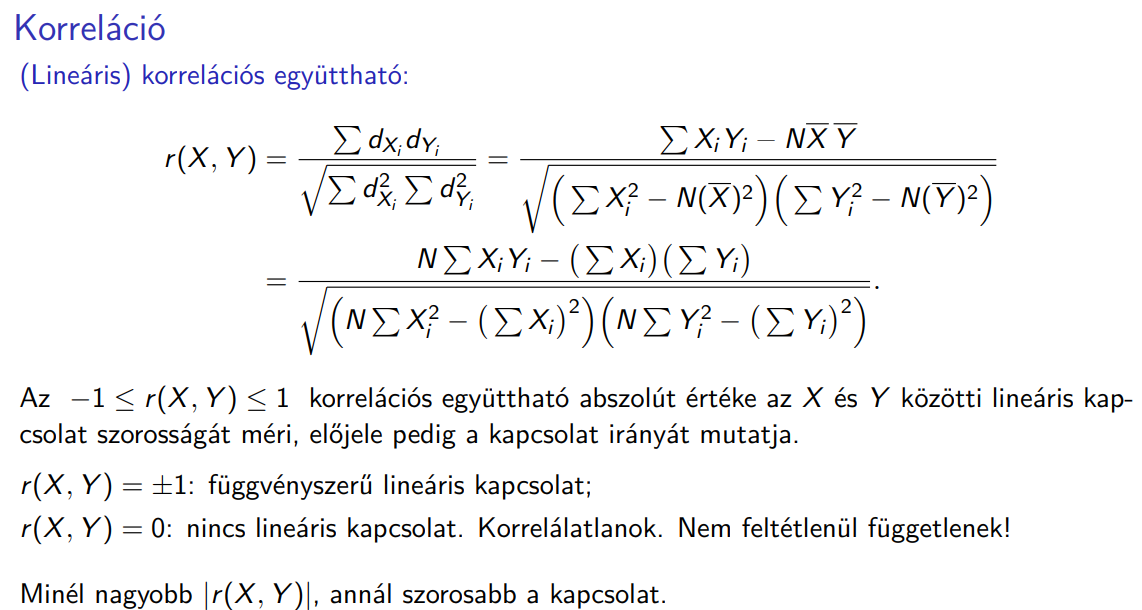
* Az (Xi, Yi) pontok összekötésével vonaldiagramon jelenítjük meg az adatokat.

**Determinációs hányados:**

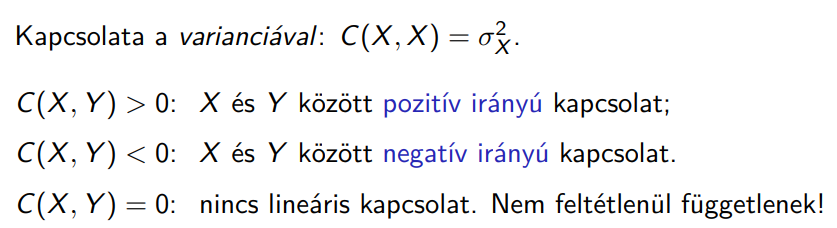
* Y-nak X-re vonatkozó determinációs hányadosa:



# Korreláció

Kovariancia A képen szöveg, Betűtípus, képernyőkép, sor látható

Automatikusan generált leírás

Determinációs együttható A képen szöveg, Betűtípus, képernyőkép, sor látható

Automatikusan generált leírásRangkorreláció A képen szöveg, képernyőkép, Betűtípus látható

Automatikusan generált leírás

# Összetett intenzitási viszonyszámok összehasonlítása

A képen szöveg, Betűtípus, sor, képernyőkép látható

Automatikusan generált leírás

V0​ és V1​ eltérésének okai:  
Eltérőek lehetnek a két sokaság ugyanazon részeire számított V0​i​ és V1i​ részviszonyszámok, és/vagy eltérő lehet a két sokaság szerkezete (összetétele).

Részviszonyszám különbségek:  
kj = V1j − V0j ​ahol V1j és V0j az adott sokaság egy részére vonatkozó viszonyszámok két különböző időpontban vagy állapotban.

Részviszonyszám hányadosok:  
ij = V1j / V0j ami a két időpont viszonyszámainak arányát adja meg az adott sokaságrészre vonatkozóan.

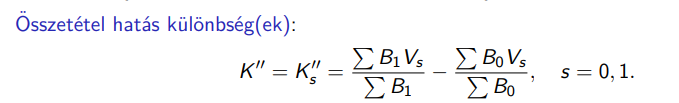
Összetett viszonyszám különbségek:  
K = V1 – V0 ami a két összetett viszonyszám közötti abszolút különbséget jelenti.

Összetett viszonyszám hányadosok:  
I = V1 / VO ami az összetett viszonyszámok arányát mutatja meg.

# Különbségfelbontás

A képen szöveg, Betűtípus, képernyőkép, sor látható

Automatikusan generált leírás



A képen szöveg, Betűtípus, képernyőkép, sor látható

Automatikusan generált leírásHányadosfelbontás A képen szöveg, képernyőkép, Betűtípus, szám látható

Automatikusan generált leírásAggregátumok összehasonlításaA képen szöveg, képernyőkép, Betűtípus látható

Automatikusan generált leírás

# Egyedi indexekA képen szöveg, képernyőkép, Betűtípus, szám látható Automatikusan generált leírás

# Érték és volumenindexA képen szöveg, képernyőkép, Betűtípus látható Automatikusan generált leírás

# ÁrindexA képen szöveg, képernyőkép, Betűtípus, algebra látható Automatikusan generált leírás

# Legfontosabb volumen és árindexformulákA képen szöveg, Betűtípus, képernyőkép, kézírás látható Automatikusan generált leírás

# Indexek átlagformáiA képen szöveg, képernyőkép, Betűtípus, sor látható Automatikusan generált leírás

# ÖsszefüggésekA képen szöveg, Betűtípus, sor, képernyőkép látható Automatikusan generált leírásA képen szöveg, Betűtípus, sor, kézírás látható Automatikusan generált leírás

# Laspeyres és a Paache féle indexek eltéréseiA képen szöveg, Betűtípus, képernyőkép, kézírás látható Automatikusan generált leírás

# Árollók

Két egymással valamilyen kapcsolatban álló csoport indexeit hasonlítjuk össze, leggyakrabban árindexeket. Az **árolló** két árindex hányadosát jelenti, és az alábbi esetekben használatos:

* **Agrárolló**: a mezőgazdasági termékek termelői árindexének és a mezőgazdasági ráfordítások árindexének hányadosa. Ez azt mutatja meg, hogy a mezőgazdasági termelők jövedelmezősége hogyan változik a termékek eladási ára és a termelési költségek árindexei viszonyában.
* **Cserearányindex** (*terms of trade*): az exportált és importált termékek árindexének hányadosa. Ez azt jelzi, hogy az exportált termékekért mennyi importált termék kapható az adott időszakban a bázisidőszakhoz képest.

Bortkiewicz formulaA képen szöveg, képernyőkép, Betűtípus látható

Automatikusan generált leírás

# **Cserearányindex**

A képen szöveg, képernyőkép, Betűtípus, sor látható

Automatikusan generált leírásAggregátummátrixA képen szöveg, képernyőkép, Betűtípus, sor látható

Automatikusan generált leírás

# Indexsorok

1. **Bázisindexsorok**: Ha a főátlóban vagy egy adott sorban/oszlopban lévő minden aggregátumot egy bázisidőszak aggregátumával osztunk, akkor bázisindexsorokat kapunk.  
   **Láncindexsorok**: Ha viszont minden aggregátumot az adott sorban/oszlopban közvetlenül előtte található aggregátummal osztunk, akkor láncindexsorokhoz jutunk.
2. **Állandó súlyú indexsorok**: Ha a volumen- és árindexsorok minden egyes tagját ugyanazon oszlop vagy sor aggregátumaiból számítjuk, akkor állandó súlyú indexsorokat kapunk.  
   **Változó súlyú indexsorok**: Ha az indexsor minden tagját más-más oszlopban/sorban található két aggregátum hányadosaként számítjuk, akkor változó súlyú indexsorokat kapunk.
3. **Laspeyres- vagy Paasche-féle indexek**: A változó súlyú láncvolumen- vagy láncárindexek tagjairól mindig egyértelműen eldönthető, hogy Laspeyres- vagy Paasche-féle indexek-e. Ezt a súlyozás alapján lehet megállapítani.
4. **Láncindexek és bázisindexek kapcsolata**: Egymást követő láncindexek szorzata (érték vagy állandó súlyú volumen, illetve árindexek esetén) egy bázisindexet ad.

# Több időszak közötti összehasonlításA képen szöveg, képernyőkép, Betűtípus látható Automatikusan generált leírás

# Területi indexek

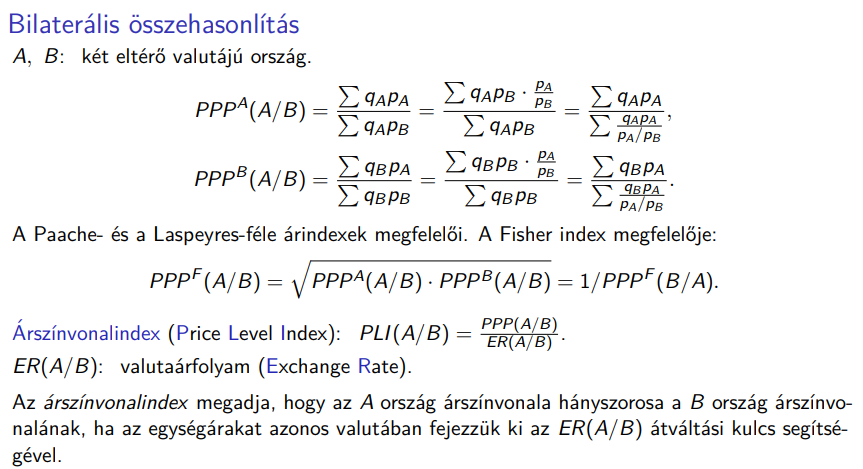
A területi indexek esetében az időszakok helyett területi egységek kerülnek összehasonlításra, és az ezekhez kapcsolódó aggregátumokat vetjük össze. A területi indexek sajátosságai a következők:

* **Területi egységek sorrendje**: A területi egységek között nincs sorrend, de fontos, hogy az azonos relációjú összehasonlítások azonos eredményhez vezessenek.
* **Alkalmazásuk**: Adott ország régiói vagy azonos valutájú országok esetén a területi indexek jelentése megegyezik a hagyományos indexekével. Azonban eltérő valutájú országok esetében a területi indexeket **nemzetközi indexeknek** nevezzük, ahol az egyes országok aggregátumai különböző valutákban vannak kifejezve.
* **Vásárlóerő-paritás (PPP)**: A nemzetközi árindexeket vásárlóerő-paritással (PPP - *Purchasing Power Parity*) mérjük.
  + A **PPP(A/B)** jelölésű nemzetközi árindex azt mutatja meg, hogy **B ország** egy valutaegysége mennyi **A ország** valutaegységével egyenértékű, ha a vizsgált termékek megvásárlására fordítjuk.

**Vásárlóerő-paritás (PPP)**

* Az **egyedi vásárlóerő-paritás** az árak arányát jelenti két ország között, és azt mutatja meg, hogy egy ország valutaegysége mennyit ér a másik ország valutájában. Matematikailag így fejezhetjük ki:
* PPP = PA / PB
* ahol PA és PB az A és B országban egy adott termék ára. Ez azt mutatja, hogy B ország egy valutaegysége hány A valutát ér, ha egy azonos terméket vásárolunk.

# Bilaterális összehasonlítás



# Sokaságok megadása

A sokaság egy adott jelenség vagy folyamat minden egyes megfigyelését tartalmazza. Az adott sokaság megadásának többféle módja van, attól függően, hogy véges vagy végtelen elemekből áll, illetve hogy az adatok diszkrétek vagy folytonosak.

* **Véges, N elemű sokaság**: Ha a sokaság elemeinek száma véges (például 10 ember magassága), akkor az egyes elemeket felsoroljuk. Ez így néz ki: Y1​,Y2​,…,YN

**Végtelen sokaság**: Ha a sokaság végtelen számú elemből áll (például az emberek magasságának eloszlása), akkor a diszkrét és folytonos ismérvek szerint adhatjuk meg az eloszlást

# Minta

A minta egy véges számú (általában n elemű) adatból álló részhalmaza a sokaságnak. Az egyes elemek a sokaság elemeiből kerülnek ki.

* **Minta**: A minta nem más mint az egyes mintaelemek. Mivel a minta elemei a sokaság elemeiből származnak, minden egyes mintaelem valószínűségi változó, amely a sokaság eloszlását tükrözi.

**2. Véletlen mintavétel:**

A véletlen mintavétel során minden egyes sokasági elem egy **előre meghatározott valószínűséggel** kerül a mintába. Azaz a mintavétel probabilisztikus jellegű: minden egyes elemet véletlenszerűen választunk ki a sokaságból, és a kiválasztás valószínűsége a sokaságban lévő elemek eloszlásán alapul.

* **Véletlen mintavétel**: A mintavétel során a választott elemek közül mindegyik az adott eloszlás szerint kerül kiválasztásra, tehát a mintavétel minden elemre egy valószínűséget rendel.

**3. Mintaelemek és a realizációk:**

Miután a mintavétel megtörtént, a mintában szereplő elemek konkrét számértékeket tartalmaznak. A **mintavétel előtt a mintaelemek véletlen változók**, amelyek az adott eloszlásból származnak, a **mintavétel után pedig egy konkrét "realizációt**" kapunk, azaz egy valós adatot, amely a minta egyes elemeit tartalmazza.

* **Valószínűségi változók**: A mintavétel előtt minden mintaelem egy valószínűségi változó, amely az adott sokaság eloszlásának megfelelően alakul.
* **Realizáció**: A mintavétel után a minta egy konkrét eredményt ad, amely az egyes mintaelemek valódi, konkrét értékeit tartalmazza (pl. konkrét számadatok).

**4. Mintajellemzők:**

A minta jellemzői, mint például az **átlag**, **szórás**, **kvantilisek**, stb., statisztikai mutatók, amelyek a minta alapján jellemzik az adatokat. Mivel ezek a mutatók a minta alapján kerülnek kiszámításra, így minden mintajellemző valószínűségi változóként kezelhető, amely az egyes mintákból származó adatok eloszlásának függvénye.

* **Mintajellemzők**: Az adatok elemzésére szolgáló mérőszámok, mint például az átlag, a szórás, a medián, a kvartilisek. Ezek mind statisztikai mutatók, amelyek a minta adatai alapján számíthatók.

**5. Mintavételi eloszlás:**

A mintavételi eloszlás a mintajellemzők eloszlását jelenti, amely azt mutatja meg, hogyan oszlanak el a különböző mintákból származó jellemzők. Ha például egy mintaátlagot számítunk, akkor annak különböző mintákból történő meghatározása különböző eloszlásokkal rendelkezhet.

* **Mintavételi eloszlás**: A mintavételi eloszlás a mintajellemzők eloszlását jelenti, amely azt mutatja meg, hogy a mintaátlag (vagy más jellemző) hogyan változik különböző minták esetén, ha a minták véletlenszerűen kerülnek kiválasztásra a sokaságból.

# FAE minta (független, azonos eloszlású minta)

A FAE minta az a minta, amelyet **független** és **azonos eloszlású** valószínűségi változókból veszünk. Ennek lényege, hogy:

* A mintaelemek **függetlenek** egymástól, azaz egy mintaelem kiválasztása nem befolyásolja a többi mintaelem kiválasztását.
* A mintaelemek **azonos eloszlásúak**, tehát mindegyik azonos valószínűségi eloszlásból származik (például mindegyik a sokaság átlagát és szórását követi).

A FAE minta két fő esetben keletkezhet:

* **Véges, homogén sokaság esetén**: Ha a sokaság véges és homogén (minden elem ugyanazon jellemzők szerint van meghatározva), akkor minden egyes elem azonos valószínűséggel kerül kiválasztásra, például visszatevéses mintavétel alkalmazásával.
* **Nagy sokaság esetén**: Ha a sokaság nagyon nagy (szinte végtelen), akkor még a visszatevés nélküli mintavétel is közelítőleg FAE mintát ad, mivel a kiválasztott elemek közötti kölcsönhatás elhanyagolható.

**FAE minta eloszlása:**

* **Eloszlással megadott sokaság esetén**: Ha a sokaság eloszlásával rendelkezünk, akkor az FAE minta elemei függetlenek és az adott eloszlás szerint véletlenszerűen választódnak ki. A kiválasztott minták mindegyike tehát ugyanazzal az eloszlással rendelkező valószínűségi változókat tartalmaz.

**Várható értéke, szórásnégyzete, standard hibája:**

* várható érték: E(y)=(1/n) \* ​(n⋅μ)=μ
* szórásnégyzete: σ2​ / n
* standard hiba: σy​= σ / gyök(n)​

# Egyszerű véletlen minta (EV minta)

Az **egyszerű véletlen minta** (EV minta) egy olyan mintavételi technika, amelyet **homogén**, **véges elemszámú** sokaság esetén alkalmazunk. Az EV mintavételnél az egyes mintaelemek **visszatevés nélküli** kiválasztása történik, és minden lehetséges minta azonos valószínűséggel kerül kiválasztásra. Ez azt jelenti, hogy minden egyes elem a sokaságból ugyanakkora eséllyel kerül a mintába, és minden egyes minta kifejezetten véletlenszerű.

**A folyamat:**

* **Homogén sokaság**: Azaz a sokaság minden eleme azonos jellemzőkkel rendelkezik, és a mintavétel nem befolyásolja az egyes elemek kiválasztását.
* **Visszatevés nélküli mintavétel**: Az EV minta esetén minden egyes mintaelem kiválasztása után az adott elem nem kerül vissza a sokaságba, tehát a következő kiválasztásnál már nem szerepelhet újra.

**Várható érték és szórás:**

Az EV minta várható értéke és szórása hasonló a független, azonos eloszlású mintánál, de figyelembe kell venni a **mintavételi hibát**, amely a minta elemszámának és a sokaság elemszámának arányából adódik.

1. **Várható érték (E(yEV))**: Az EV minta várható értéke megegyezik a sokaság várható értékével, tehát: E(yEV​) = μ  
   Ez azt jelenti, hogy a minta várható értéke nem torzul a kiválasztás módszere miatt, és az EV minta az átlagos sokasági jellemzőket tükrözi.
2. **Szórás (Var(yEV))**: Az EV minta szórása függ a minta elemszámától és a sokaság elemszámától. A következő képlettel számítható ki:   
   Var(yEV​)= (σ2​ / n) \* [(N - n) \* (N-1)] ~ (σ2​ / n) \* (1 – n / N)

# A rétegezett minta fogalma és eljárás

* **Heterogén sokaság**: A sokaság, amely különböző, egymástól eltérő csoportokból (rétegekből) áll. Például, ha a lakosság jövedelmi viszonyait vagy iskolázottságát szeretnénk vizsgálni, akkor a sokaságot rétegezhetjük például **lakóhely** alapján (pl. **Budapestre**, **megyeszékhelyekre**, **vidéki településekre**), vagy **jövedelmi szintek** szerint.
* **Rétegezett mintavétel lépései**:
  1. **A sokaság rétegekre bontása**: Először a sokaságot különböző, egymástól jól elkülöníthető **rétegekre** soroljuk. A rétegeket olyan ismérvek szerint képezzük, amelyek jelentős hatással lehetnek a vizsgált jellemzőre. A rétegeknek **átfedésmentesen kell lefedniük** a teljes sokaságot, tehát minden egyes egyed a sokaságban pontosan egy réteghez tartozik.
  2. **Rétegeken belüli mintavétel**: Miután meghatároztuk a rétegeket, az egyes rétegeken belül **egyszerű véletlen mintavételt** (EV) alkalmazunk. Ritkábban előfordulhat, hogy **független, azonos eloszlású (FAE) mintavételt** használunk egy-egy rétegen belül. A rétegeken belüli mintavétel során a rétegekhez tartozó elemek egymástól **függetlenek**, és minden rétegen belül minden elem egyenlő eséllyel kerül a mintába.

# Egyenletes eloszlás

* **Leírás**: Az egyenletes elosztás során minden rétegből **azonos elemszámú** mintát veszünk, függetlenül a réteg elemszámától. Ez azt jelenti, hogy minden rétegből ugyanannyi mintát választunk ki

**Előnyök**:

* **Egyszerű**: Könnyen számolható, mivel az egyes rétegekből ugyanakkora számú minta kerül kiválasztásra.
* Az egyes rétegek jellemzői könnyen meghatározhatók a megfelelő számadatokból.

**Hátrányok**:

* **Nem optimális**, ha a rétegek között nagy eltérés van a varianciákban, mivel a kisebb elemszámú rétegek aránytalanul nagy befolyást gyakorolhatnak a becslésre.

# Arányos eloszlás

* **Leírás**: Az arányos elosztás során a minták elemszámai **arányosak** a rétegek elemszámaival. Azaz az egyes rétegekből vett minta elemszáma úgy van meghatározva, hogy az arányban álljon a réteg elemszámával az összes réteg elemszámához viszonyítva

**Előnyök**:

* **Egyszerű** és könnyen alkalmazható, mivel az egyes rétegek súlya arányos a réteg elemszámával a sokaságban.
* Jobb reprezentációt ad a rétegek arányának megfelelően, így a mintában minden réteg megfelelően lesz jelen.

**Hátrányok**:

* Ha a rétegek közötti különbségek (pl. varianciák) nagyok, az arányos mintavétel nem biztosít optimális becslést, mivel a kisebb elemszámú rétegek varianciája torzíthatja a mintát.

# Neyman-féle optimális eloszlás

* **Leírás**: A Neyman-féle optimális elosztás célja, hogy a nagyobb **szórású** rétegekből **nagyobb számú mintát** vegyünk. A mintavételi elemszámot az alábbiak szerint számítjuk ki

**Előnyök**:

* Az optimális eloszlás csökkenti a mintavételi hibát, mivel a szórás alapján súlyozza a minták elemszámát. Ha egy réteg szórása nagy, akkor abból több mintát veszünk, így az ilyen rétegekre vonatkozó becslés pontosabb lesz.

**Hátrányok**:

* **Nehezebb alkalmazni**, mivel a rétegek szórása általában nem ismert előre. Emiatt a pontos eloszlás megvalósítása gyakran problémás, mivel a szórások ismerete nélkül a mintavételi elemszámok nem számíthatók ki pontosan.

# Egylépcsős (1L) minta

* **Leírás**: Az egylépcsős csoportos mintavétel során a sokaságot **csoportokra** (más néven **klaszterekre**) osztjuk, és ezekből a csoportokból választunk ki véletlenszerűen mintát. Miután a csoportokat kiválasztottuk, azokat **teljeskörűen megfigyeljük**, azaz minden egyes elemét vizsgáljuk, amely a csoportba tartozik.
* **Jellemzők**:
  + A kiválasztott csoportokon belül **minden elemre** adatokat gyűjtünk.
  + **Homogén csoportok** esetén az eljárás hatékony, mivel ha a csoportok belső heterogenitása alacsony, akkor a csoportok egészének vizsgálata gyorsabb és egyszerűbb lehet.
  + **Heterogén csoportok** esetén az eljárás nem lesz olyan hatékony, mivel a csoporton belüli nagy különbségek a becsléseket torzíthatják.
* **Előnyök**:
  + Könnyen alkalmazható, ha a csoportok jól meghatározottak.
  + Alkalmas akkor, amikor az összes egyes elem megfigyelése túl költséges vagy nem lehetséges, de a csoportok adatai hozzáférhetők.
* **Hátrányok**:
  + Ha a csoportok nem homogének, az egyes minták megbízhatósága csökkenhet, mivel a csoportok belső varianciája nagy hatással lehet a becslések pontosságára.

# 2. Többlépcsős (TL) minta

* **Leírás**: A többlépcsős mintavétel egy továbbfejlesztett változata az egylépcsős mintavételnek. Itt a mintavétel több **lépcsőben történik**. Az első lépcsőben csoportokat (klasztereket) választunk ki, majd az egyes csoportokból újabb mintákat veszünk.
* **Jellemzők**:
  + Az első lépcsőben egy nagyobb csoportot választunk ki, majd a második lépcsőben az egyes csoportokból további mintákat választunk.
  + Például egy kétlépcsős mintavétel esetén először kiválasztunk néhány iskolát (csoportot), majd az iskolákból további diákokat (elemeket) választunk véletlenszerűen.
* **Előnyök**:
  + **Kevesebb redundancia**: Mivel az egyes csoportok kiválasztása után csak azokból választunk, így a második lépcsőben már nem ismétlődnek a minták, és kevesebb redundáns adat keletkezik, mint egy egyszerű, egylépcsős mintavételnél.
  + Lehetővé teszi, hogy az adatgyűjtést költséghatékony módon végezzük, különösen nagy elemszámú sokaságok esetén.
* **Hátrányok**:
  + **Komplexebb megvalósítás**: Többlépcsős mintavételnél több lépést kell végrehajtani, ami bonyolultabb adatfeldolgozást igényel.
  + A minták kiválasztása különböző lépcsőkben történik, így a becslések pontossága függhet az egyes lépcsőkben alkalmazott mintavételi módszerektől.

# Becslőfüggvény definíciója

A becslőfüggvény egy statisztika, ami a minta elemeiből kiszámított értéket ad. Ezt a statisztikai eljárást a minta egy vagy több eleme alapján végezzük, és az eredmény egy **valószínűségi változó**, amely közelíti a sokaság paraméterét.

* A becslőfüggvény egy adott **sokasági jellemző** (θ) értékének becslésére szolgál. A becslőfüggvényt a minta adatai alapján számítjuk ki.
* **Korrigált tapasztalati szórásnégyzet**: A korrigált szórásnégyzet egy pontosabb becslést ad a sokaság szórásnégyzetére, ha **visszatevés nélküli mintavételt** alkalmazunk. A képen szöveg, Betűtípus, kézírás, képernyőkép látható

  Automatikusan generált leírás

# **Torzítatlanság **definíciója****

A becslőfüggvény torzítatlanságát akkor mondhatjuk el, ha annak várható értéke pontosan a becsülni kívánt paraméter értékével egyezik meg. Ez azt jelenti, hogy ha ismételten végeznénk a mintavételt és minden esetben kiszámítanánk a becslőfüggvényt, akkor az átlagos becslés mindig a sokaság valódi paraméterét adná. Mind FAE, mind EV minta esetén: E(µ) = E(y) = μ

# Mintavételi szórásnégyzet

A becslőfüggvény szórásnégyzetét **mintavételi szórásnégyzetnek** nevezzük, és azt jelenti, hogy milyen mértékben változik a becslés, ha különböző mintákat veszünk a sokaságból. A mintavételi szórásnégyzet a becslőfüggvény szórásának négyzete, és az alábbi képlettel számolható ki: Var(θ) = mintavételi szórásnégyzet

Standard hiba

A becslőfüggvény **standard hibája** (standard error, SE) a becslés szórásának a négyzete

# Minimum Variance Unbiased Estimator (MVUE)

**Minimális Szórásnégyzetű Torzítatlan Becslő:** Egy **MVUE** olyan torzítatlan becslés, amely nemcsak torzítatlan, hanem a legalacsonyabb szórásnégyzettel rendelkezik a torzítatlan becslők között. Más szóval, ha több torzítatlan becslő létezik, akkor az MVUE az, amelynek a szórásnégyzete a legkisebb.

**Hatékonyság:**

A becslőfüggvény hatékonysága azt jelenti, hogy egy adott becslés mennyire optimális a paraméter becslésére. A hatékonyságot a szórásnégyzet és a variancia minimalizálásával mérhetjük.

* Ha egy becslőfüggvény **MVUE**, akkor az a legjobb torzítatlan becslő, amelyet választhatunk a variancia minimalizálására.
* A **MVUE** becslések a **leghatékonyabbak**, mivel biztosítják, hogy a becslés a legkisebb szórásnégyzettel rendelkezzen a torzítatlan becslők között.

# MSE (Mean Squared Error)

Az **MSE (Mean Squared Error)** vagy **átlagos négyzetes hiba** egy olyan mérőszám, amely a becslés és a valós érték közötti eltérést méri, és segít összehasonlítani a különböző becslőfüggvényeket. Az MSE a becslés **pontosságát** és **torzítottságát** egyaránt figyelembe veszi, tehát nemcsak a becslés átlagos eltérését, hanem annak szórását is. Az a becslőfüggvény a "jobb", amelyiknek az **átlagos négyzetes hibája (MSE)** kisebb, mivel az MSE a becslés pontosságát és megbízhatóságát méri.

**Aszimptotikus torzítatlanság:**

Az **aszimptotikus torzítatlanság** arra utal, hogy egy becslőfüggvény torzítása a mintaelemek számának növekedésével csökkenhet, és végül a minta méretének növekedésével a becslés torzítása nulla lesz, ha a becslőfüggvény megfelelő.

A **becslőfüggvény konzisztenciája** azt jelenti, hogy ahogy a minta mérete növekszik, a becslés egyre közelebb kerül a valódi paraméter értékéhez, és végül a becslés megbízhatóan közelíti azt.  
Momentumok módszere

A **momentumok módszere** egy statisztikai eljárás, amelyet a paraméterek becslésére használnak, különösen, ha az eloszlás konkrét formája nem ismert, de a momentumszámítások elérhetők. A módszer célja, hogy egy eloszlás paramétereit olyan módon becsülje meg, hogy a minta momentumai (vagyis a statisztikai jellemzők, mint az átlag, szórás, harmadik és negyedik rendű momentumaik) megegyezzenek az elméleti eloszlás momentumaival.

# 1. Exponenciális eloszlás1. Exponenciális eloszlásEloszlások diagrammja

# Egyenletes eloszlás | Dr. Csallner András Erik, Vincze Nándor: Bevezetés a valószínűség-számításba és a matematikai statisztikábaLikehood becslés

**Likelihood függvény**: A *likelihood* (valószínűségi) függvény azt fejezi ki, hogy a minta hogyan jön létre egy adott paraméterérték mellett. Diszkrét eloszlás esetén a minta együttes valószínűségét, folytonos eloszlás esetén pedig a minta együttes sűrűségét adja meg.

**Log-likelihood függvény**: A *log-likelihood* a likelihood függvény logaritmusát jelenti, ami gyakran matematikai értelemben kényelmesebb a becsléshez. A logaritmus alkalmazása nem változtat a maximális helyen, de leegyszerűsíti a számításokat.

**ML becslés**: Az **ML becslés** a paraméterek legjobb becslését adja meg, amely maximális értéket ad a likelihood vagy log-likelihood függvény számára. Más szóval, a paramétereket úgy választjuk, hogy azok a legnagyobb valószínűséget biztosítsák a megfigyelt adatok számára.

**Maximum és minimum helyek:**

1. **Likelihood függvény maximuma**: A ML becslés során azt a paramétert keressük, amelyik a *likelihood* vagy *log-likelihood* függvényt maximális értékre hozza. Azaz, az adott adat mellett a paraméterek közül azt választjuk, amely a legnagyobb valószínűséggel magyarázza a megfigyelt adatokat.
2. **Függvény maximuma**: A maximális érték megtalálásához a likelihood függvény első deriváltját (gradienst) kell megvizsgálni. A maximum akkor érhető el, amikor a derivált nulla, és a második derivált negatív, jelezve, hogy az adott pont valóban egy maximum.
3. **Függvény minimuma**: Habár a maximum keresése az alapvető cél, a minimumok is megjelenhetnek a függvényekben, különösen, ha a függvény nem csak egy globális maximumot tartalmaz. Ha a függvény egy paraméterre vonatkozóan nem rendelkezik világos maximumhelyekkel, akkor a minimum helyét is figyelembe kell venni, hogy elkerüljük a nem releváns megoldásokat.

# Markov-egyenlőtlenség

A **Markov-egyenlőtlenség** egy általános egyenlőtlenség, amely bármilyen nem-negatív valószínűségi változóra alkalmazható. Az egyenlőtlenség azt mondja, hogy a valószínűsége, hogy egy nem-negatív valószínűségi változó **Y** nagyobb vagy egyenlő legyen egy adott δ pozitív számnál, nem lehet nagyobb, mint az Y várható értéke (E(Y)) osztva δ-val.

# Csebisev-egyenlőtlenség

A **Csebisev-egyenlőtlenség** egy erősebb eszköz, amely a **szórás** vagy **variancia** figyelembevételével korlátozza, hogy egy valószínűségi változó mennyire térhet el a várható értékétől.

**Összegzés egyszerűbben:**

* **Markov-egyenlőtlenség**: Ha egy várható értéket tudsz, akkor a nagyobb értékek elérésének valószínűsége nem lehet nagyobb, mint a várható érték osztva a kérdéses nagyobb értékkel. Pl. ha a nyeremény átlagos értéke 10 forint, akkor a 20 forintnyi nyeremény elérésének valószínűsége legfeljebb 50%.
* **Csebisev-egyenlőtlenség**: Ha tudod, hogy egy érték átlagosan mekkora, és hogy mennyit szóródik körülötte, akkor a nagy eltérések valószínűsége is korlátozott. Pl. ha az átlagos magasság 5 láb, és a szórás 4 hüvelyk, akkor annak a valószínűsége, hogy valaki legalább 3 hüvelykkel eltér az átlagtól, legfeljebb 44,44%.

# Nagy Számok Gyenge Törvénye (NSZT)

A Nagy Számok Gyenge Törvénye arról szól, hogy ha van egy sorozat **független, azonos eloszlású** valószínűségi változókból (Y1, Y2, ..., Yn) és mindegyiknek van egy véges várható értéke (E(Y1) = µ) és szórása, akkor:

* Az átlaguk (Sn / n, ahol Sn = Y1 + Y2 + ... + Yn) **sztochasztikusan konvergál** a várható értékükhöz, vagyis ahogy a minta mérete nő, az átlag egyre inkább közelíteni fog a valódi várható értékhez (µ).

**Példa a gyenge törvényre**

Képzeld el, hogy egy kísérletet végzel, például egy pénzérme feldobását, ahol a **várható érték** (µ) a két kimenet (fej vagy írás) arányát jelenti, tehát ha ideális pénzérmét dobsz, akkor az eredmények várhatóan 50-50% arányban fej és írás lesznek.

* Ha egy vagy két érmét dobsz, lehet, hogy a fej és írás aránya nem pontosan 50%. De ahogy egyre több pénzérmét dobsz, az arány egyre közelebb fog kerülni a valódi 50%-hoz.
* Tehát a nagy számok gyenge törvénye szerint, minél több érmét dobsz (vagyis minél több mintát veszel), annál jobban fog közelíteni az átlagos arány (például a fej-írás arány) a valódi várható értékhez (ami itt 50%).

# Központi Határeloszlás Tétel

Azt mondja, hogy független és azonos eloszlású valószínűségi változók sorozata egyre inkább normális eloszlást fog követni, ahogy a minta mérete nő. Ennek alapvető következménye, hogy a nagy minták esetén a mintaátlagok eloszlása jól közelíthető normál eloszlással, és a középértékekre vonatkozó becslések egyre pontosabbak lesznek. Nagy n esetén Y = Sn / Y ​​ eloszlása hozzávetőlegesen normális μ várható értékkel és σ2 / n szórásnégyzettel.

# Khi-négyzet eloszlás

A **khi-négyzet eloszlás (chi-square distribution)** egy folyamatos valószínűségi eloszlás, amely a statisztikában gyakran előfordul. Ha van egy **X₁, X₂, ..., Xn** független standard normális eloszlású valószínűségi változónk, akkor az ezek négyzetének összege



**Várható érték:** Az **n** szabadságfokú khi-négyzet eloszlás várható értéke **n**.  
**Szórás négyzete:** A khi-négyzet eloszlás szórásának négyzete **2n**. A képen szöveg, képernyőkép, Betűtípus, tervezés látható

Automatikusan generált leírás

# t-eloszlás

A **t-eloszlás** (Student-féle t-eloszlás) egy fontos statisztikai eloszlás, amelyet a standard normális eloszlásból vezetünk le, és elsősorban kis minták elemzésére használunk.

A képen szöveg, képernyőkép, Betűtípus látható

Automatikusan generált leírás

A képen szöveg, képernyőkép, Betűtípus látható

Automatikusan generált leírás

# F-eloszlás

A képen szöveg, képernyőkép, Betűtípus látható

Automatikusan generált leírásAz **F-eloszlás** (Fisher-féle eloszlás) egy fontos statisztikai eloszlás, amelyet független khi-négyzet eloszlású változókból vezetünk le

A képen szöveg, képernyőkép, Betűtípus látható

Automatikusan generált leírás

A képen szöveg, képernyőkép, Betűtípus, tervezés látható

Automatikusan generált leírás

A képen szöveg, képernyőkép, Betűtípus látható

Automatikusan generált leírás

# Az intervallumbecslés alapjai

Az intervallumbecslés egy statisztikai módszer, amelynek célja, hogy egy minta alapján olyan intervallumot határozzunk meg, amely egy előre megadott valószínűséggel tartalmazza a sokaság valamely ismeretlen jellemzőjét (θ). Az így meghatározott intervallumot **konfidencia intervallumnak** nevezzük.

**Főbb fogalmak és definíciók:**

1. **Konfidencia intervallum:**
   * Egy adott megbízhatósági szinthez (1−α) tartozó intervallum, amely az ismeretlen paraméter (θ\thetaθ) becslésére szolgál.
   * Matematikailag: A képen szöveg, képernyőkép, Betűtípus látható

     Automatikusan generált leírás

2. **Konfidenciaszint (1 − α):**

* Az az arány, amellyel a konfidencia intervallum a paramétert „elméletben” tartalmazza. Például, ha 95%-os konfidenciaszintet használunk, az azt jelenti, hogy az ilyen módon számított intervallumok 95%-a tartalmazza az ismeretlen jellemzőt.

3. **Szignifikanciaszint (α):**

* Az a valószínűség, hogy az intervallum nem tartalmazza az ismeretlen paramétert. Például, ha α = 0.05, akkor az intervallumok 5%-a „téves” lesz, azaz nem tartalmazza a θ paramétert.

A képen szöveg, képernyőkép, Betűtípus, szám látható

Automatikusan generált leírásA képen szöveg, Betűtípus, képernyőkép, kézírás látható

Automatikusan generált leírásA képen szöveg, Betűtípus, sor, képernyőkép látható

Automatikusan generált leírás

A képen szöveg, képernyőkép, Betűtípus, algebra látható

Automatikusan generált leírás

A képen szöveg, Betűtípus, kézírás, képernyőkép látható

Automatikusan generált leírás

A képen szöveg, képernyőkép, Betűtípus, szám látható

Automatikusan generált leírás