

Statisztika gyakorlat

1.

Gazdasági agrármérnök szak

II. évfolyam

2007.02.13.

Félévi követelményrendszer

- **Heti óraszám:** 2+2
- **Aláírás feltételei:**
 - az előadásokon való részvétel nem kötelező, de AJÁNLOTT!
 - a gyakorlatokon való részvétel kötelező (max. 3 hiányzás)
 - 1 db ZH elégséges teljesítése
 - 2007.03.26. 9:40 – 11:10 E2
 - elméleti és gyakorlati rész min. 50-50%-os teljesítése
 - pótZH:
 - 2007.04.02. 9:40 – 11:10 E2
- **Beszámoló nem jár automatikusan!!!**
- **Vizsga:**
 - írásban történik
 - elméleti és gyakorlati rész min. 50-50%-os teljesítése

Órai munka

- Gyakorlatokon való részvétel **kötelező!**
- **Ajánlott felszerelés:**
 - Számológép!!!
 - Nagy Mónika Zita – Barna Katalin – Dr. Molnár Tamás: **Egyszerűen statisztika 1.**
 - Barna Katalin – Nagy Mónika Zita – Dr. Molnár Tamás: **Egyszerűen statisztika 2.**
 - Barna Katalin – Kovács Bernadett – Dr. Molnár Tamás: **Statisztikai képletek és táblázatok gyűjteménye**
- **Aktív órai munka:**
 - a gyakorlatokon való aktív részvétel beleszámít a ZH-ba

Előadás összefoglalása 1.

- **Mi a statisztika?**
 1. Adatok halmaza
 2. Tudományos módszertan
 - Adatelemzés egyszerű és bonyolultabb módszerekkel
 3. Hivatalos statisztikai szervezet
 4. Gyakorlati tevékenység
 - Statisztikai program készítése
 - Adatgyűjtés
 - Adatok rendezése
 - **Elemzés**
 - Közzététel
 - Tárolás

Előadás összefoglalása 2.

- Alapvető fogalmak és jelölések:
 - Alapsokaság (jele: N)
 - Mintasokaság (jele: n)
 - Sokaság egyedei (x_i)
 - Elemek gyakorisága (f_i)

Előadás összefoglalása 3.

- **Adatrendezés módszere:**
 - csoportosítás ismérvek alapján → statisztikai sorok → statisztikai tábla
- **Csoportosító ismérvek:**
 - **Minőségi:**
 - pl.: hajszín, nem, autó márka
 - **Mennyiségi:**
 - *Diszkrét* (pl.: gépkocsik kor szerinti megoszlása 2004 év végén a Dél-Dunántúli régióban)
 - *Folytonos* vagy *osztályközös* (pl.: hallgatók-, alkalmazottak száma)
 - **Idő:**
 - *Állapot idősorok* (pl.: kisvállalkozások száma Magyarországon)
 - *Tartam idősorok* (pl.: árbevétel alakulása adott vállalkozásnál)
 - **Területi:**
 - pl.: egy főre jutó GDP alakulása régióként

Előadás összefoglalása 4.

- Statisztikai sorok csoportosítása:

	Azonos fajta adatokból állnak		Különböző fajta adatokból állnak
Keletkezés szerint	Csoportosító	Összehasonlító	Leíró
Az ismérv fajtája szerint	Minőségi Mennyiségi	Idő Területi	

- Sokaság részsokaságok összegére bontható, így összegzésnek van értelme
- Összegzésnek nem mindig van értelme

Előadás összefoglalása 5.

- Csoportosítás szabályai:

- Fő szabályok:
 - Teljesség
 - Átfedés mentesség
 - Homogenitás
- Kiegészítő szabályok:
 - Osztályok határai lehetőleg kerek számok legyenek
 - Osztályszélességek azonosak legyenek
 - Nyílt osztályközös gyakorisági sort hozzunk létre
 - Egy elemnek ne „nyissunk” külön osztályt
 - Osztályközök száma legyen optimális

Előadás összefoglalása 6.

- Csoportosítás módszerei

- Matematikai módszerekkel:

$$k = 1 + 3,3 \cdot \lg n$$

- ahol
 - k = az osztályközök száma,
 - n = a csoportosítani kívánt sokaság elemszáma

$$h = \frac{x_{\max} - x_{\min}}{k}$$

- ahol
 - h = osztályköz hossza,
 - x_{\max} = a csoportosítandó adatsor legnagyobb eleme,
 - x_{\min} = a csoportosítandó adatsor legkisebb eleme.

- Empirikus (tapasztalati) úton

Előadás összefoglalása 7.

- Statisztikai táblák:

- A statisztikai sorok összefüggő rendszerét statisztikai tábláknak hívjuk.
- Jellemzői:
 - A táblának minden esetben *címmel* kell rendelkeznie,
 - a címben *utalni kell arra, hogy a táblázat mit tartalmaz*, milyen időpont adatai találhatóak a táblában,
 - meg kell jelölni az adatok *forrását*,
 - ahol szükséges ott a *magyarázó szövegeket* is fel kell tüntetni, és
 - fontos megadni a táblázatban szereplő adatok *mértékegységét* is.
- Típusai:
 - Egyszerű
 - Csoportosító
 - Kombinatív

Osztóértékek meghatározása

- A folytonos mennyiségi ismérveket nem csak csoportosítással, hanem osztóértékekkel is rendezhetjük.
- Az osztóértékek (kvantilis) a **nagyság szerint sorrendbe rendezett** adatsort **egyenlő nagyságú** (egyenlő gyakoriságú) részekre osztják.
- A kvantiliseknek nagyon sok fajtája van, melyek közül a leggyakrabban a következőket használjuk:

Osztóértékek fajtái

Osztóérték		Egyenlő részek száma	Egy rész az egész mekkora hányada, %	Osztóértékek száma az adott típusból
neve	jele			
Felező (medián)	Me	2	50	1
Harmadoló (tercilis)	T	3	33,3	2
Negyedelő (kvartilis)	Q	4	25	3
Ötödölő (kvintilis)	K	5	20	4
Tizedelő (decilis)	D	10	10	9
Szazadoló (percentilis)	P	100	1	99

Osztóértékek meghatározása 1.

- **Eredeti adatsorból:**

1. Adatok nagyság szerinti sorba rendezése
2. Osztóérték sorszáma meghatározása

$$S_j = j \cdot \frac{n+1}{k}$$

- S_j = az osztóérték sorszáma
 - j = az adott típusú osztóérték sorszáma \rightarrow hányadikat keressük az adott osztóértékből
 - n = az adatok száma
 - k = az osztóérték típusa \rightarrow hány részre osztja fel az adatsort az adott osztóérték
3. Osztóérték konkrét értékének megállapítása
- Ha a sorszáma egész szám:
 - Egyszerűen leolvassuk a konkrét adatot
 - Ha nem egész szám az osztóérték sorszáma:
 - A sorszámot két részre bontjuk (egész rész és tört rész)

Osztóértékek meghatározása 2.

- **Osztályközös gyakorisági sorból:**

1. Adatok nagyság szerinti sorba rendezése
 - Gyakorlatilag a csoportosítással (táblázat elkészítése) megtörténik
2. Osztóérték sorszáma meghatározása

$$S_j = j \cdot \frac{n+1}{k}$$

3. Osztóértéket tartalmazó osztályköz meghatározása
 - *Kumulálás alulról felfelé*
 - az alsó osztályköztől a felsőig történő kumulálás
 - az irány nem a vertikálitást jelenti!
 - Addig halmozzuk (addig az osztályközökig) a gyakoriságokat, amíg a kumulatív gyakoriságok éppen meghaladják az osztóérték sorszámat.

Osztóértékek meghatározása 3.

4. Osztóérték meghatározása

- Nyers:
 - Az osztóértéket tartalmazó osztály osztályközepe (egyszerű számtani átlaga)

- Becsült:

$$K_j = X_{j0} + \frac{S_j - \sum f_i}{f_j} \cdot h$$

- K_j = a számítandó osztóérték
- X_{j0} = az osztóértéket tartalmazó osztály alsó határa
- S_j = az adott osztóérték sorszáma
- $\sum f_i$ = az osztóértéket tartalmazó osztályt megelőző osztályok kumulatív gyakorisága
- f_j = az osztóértéket tartalmazó osztály gyakorisága
- h = az osztóértéket tartalmazó osztály szélessége (terjedelme)

Medián (Me)

- A nagyság szerint sorba rendezett adatokat **két egyenlő részre** osztja.
- A negyedelő (kvartilis) értékek közül megegyezik a másodikkal ($Me = Q_2$)
- **Számítása történhet:**
 - Eredeti adatsorból
 - Osztályközös gyakorisági sorból

Tercilis (T)

- A nagyság szerint sorba rendezett adatokat **három egyenlő részre** osztja.
- **Számítása történhet:**
 - Eredeti adatsorból
 - Osztályközös gyakorisági sorból

Kvartilis (Q)

- A nagyság szerint sorba rendezett adatokat **négy egyenlő részre** osztja.
- A negyedelő (kvartilis) értékek közül a második (vagy középső) egyenlő a mediánnal ($Q_2 = Me$)
- **Számítása történhet:**
 - Eredeti adatsorból
 - Osztályközös gyakorisági sorból

Kvintilis (K)

- A nagyság szerint sorba rendezett adatokat **öt egyenlő részre** osztja.
- Számítása történhet:
 - Eredeti adatsorból
 - Osztályközös gyakorisági sorból

Feladatok

1. feladat

Egy kisvállalkozásnál dolgozók jövedelmének alakulása a következő. Az adatok ezerFt-ban vannak megadva.

120, 115, 148, 126, 156, 132, 155

- Határozza meg a mediánt (Me) és értelmezze az eredményt!
- Határozza meg az alsó negyedelő (Q_1) értékét és értelmezze az eredményt!
- Határozza meg a középső negyedelő (Q_2) értékét és értelmezze az eredményt!
- Határozza meg a felső negyedelő (Q_3) értékét és értelmezze az eredményt!
- Határozza meg a harmadik kvintilis (K_3) értékét és értelmezze az eredményt!

a) Határozza meg a mediánt (Me).

1. Az adatok sorba rendezése

115, 120, 126, **132**, 148, 155, 156

2. Medián sorszámának meghatározása

$$S_{Me} = 1 \cdot \frac{7+1}{2} = 4$$

3. A medián értékének megállapítása:

Jelen esetben, gyakorlatilag csak le kell olvasnunk a sorszám melletti értéket. Nálunk a 4. elem most a 132 ezerFt-os kereset.

Válasz!!!

Az alkalmazottak 50%-a kevesebbet, 50%-a többet keres, mint 132 eFt/fő.

b) Határozza meg Q_1 értékét.

1. Az adatok sorba rendezése

115, **120**, 126, 132, 148, 155, 156

2. Az alsó negyedelő sorszámának meghatározása

$$S_{Q_1} = 1 \cdot \frac{7+1}{4} = 2$$

Ez azt jelenti, hogy a sorba rendezett adataink közül a 2. elem az alsó negyedelő (Q_1), ami alatt a keresetek 25%-a, felette a keresetek 75%-a helyezkedik el.

3. A negyedelő értékének megállapítása:

Jelen esetben, szintén csak le kell olvasnunk a sorszám melletti értéket. Nálunk a 2. elem most a 120 ezerFt-os kereset.

Válasz!!!

Az adott vállalkozásnál, a dolgozók 25%-a 120 ezerFt/fő-nél kevesebbet, míg 75%-a többet keres.

c) Határozza meg Q_2 értékét.

1. Az adatok sorba rendezése

115, 120, 126, **132**, 148, 155, 156

2. Az középső negyedelő sorszámának meghatározása

$$S_{Q_2} = 2 \cdot \frac{7+1}{4} = 4$$

Ez azt jelenti, hogy a sorba rendezett adataink közül a 4. elem a középső negyedelő (Q_2), ami alatt a keresetek 50%-a, felette szintén a keresetek 50%-a helyezkedik el.

3. A negyedelő értékének megállapítása:

Jelen esetben, szintén csak le kell olvasnunk a sorszám melletti értéket. Nálunk a 4. elem most a 132 ezerFt-os kereset.

Válasz!!!

A középső negyedelő értéke 132 ezerFt/fő. Ez alatt a keresetek 50%-a, felette pedig a keresetek 50%-a található.

d) Határozza meg Q_3 értékét.

1. Az adatok sorba rendezése

115, 120, 126, 132, 148, **155**, 156

2. Az felső negyedelő sorszámának meghatározása

$$S_{Q_3} = 3 \cdot \frac{7+1}{4} = 6$$

Ez azt jelenti, hogy a sorba rendezett adataink közül a 6. elem a felső negyedelő (Q_3), ami alatt a keresetek 75%-a, felette a keresetek 25%-a helyezkedik el.

3. A negyedelő értékének megállapítása:

Jelen esetben, szintén csak le kell olvasnunk a sorszám melletti értéket. Nálunk a 6. elem most a 155 ezerFt-os kereset.

Válasz!!!

Az felső negyedelő értéke 155 ezerFt/fő. Ez alatt a keresetek 75%-a, felette pedig a keresetek 25%-a található.

e) Határozza meg K_3 értékét.

1. Az adatok sorba rendezése

115, 120, 126, *132*, *148*, 155, 156

2. Az alsó negyedelő sorszámának meghatározása

$$S_{K_3} = 3 \cdot \frac{7+1}{5} = 4,8$$

3. Az ötödölő értékének megállapítása:

Mivel az kiszámított osztóérték sorszámja nem egész szám, így konkrét értékének meghatározása a következő:

$$K_{K_3} = 132 + [(148 - 132) \cdot 0,8] = 144,8 \text{ eFt} / \text{fő}$$

Válasz!!!

A harmadik kvintilis értéke 144,8 ezerFt/fő. Ez alatt a keresetek 60%-a, felette pedig a keresetek 40%-a található.

2. feladat

Egy vállalkozásnál dolgozók jövedelmének alakulása a következő.

Bércsoportok (ezerFt / fő)	Létszám (fő)
- 100	8
100 - 150	7
150 - 200	4
200 - 250	3
250 -	2
Összesen	24

- Határozza meg a mediánt (Me) és értelmezze az eredményt!
- Határozza meg a második tercilis (T_2) értékét és értelmezze az eredményt!
- Határozza meg a középső negyedelő (Q_2) értékét és értelmezze az eredményt!
- Határozza meg a felső negyedelő (Q_3) értékét és értelmezze az eredményt!
- Határozza meg az alsó negyedelő (Q_1) értékét és értelmezze az eredményt!

a) Határozza meg a mediánt (Me) 1.

1. Az adatok sorba rendezése

2. Medián sorszámának meghatározása

$$S_{Me} = 1 \cdot \frac{24+1}{2} = 12,5$$

3. A mediánt tartalmazó osztályköz meghatározása

- Alulról felfelé történő kumulálással történik

Bércsoportok (ezerFt / fő)	Létszám (fő)	Kumulatív gyakoriság
- 100	8	8
100 - 150	7	8+7=15
150 - 200	4	8+7+4=19
200 - 250	3	8+7+4+3=22
250 -	2	8+7+4+3+2=24
Összesen	24	-

a) Határozza meg a mediánt (Me) 2.

4. A medián értékének megállapítása:

1. Nyers medián meghatározása:

- A mediánt tartalmazó osztály középe: **125 eFt/fő**

2. Becsült medián meghatározása:

- Figyelembe veszi az osztályközök gyakoriságát is.

$$K_{M_i} = X_{j_0} + \frac{S_j - \sum f_i}{f_j} \cdot h = 100 + \frac{12,5 - 8}{7} \cdot 50 = 132,14 \text{ eFt} / \text{fő}$$

Válasz!!!

A nyers medián értéke 125 eFt/fő, a becsült medián értéke **132,14 eFt/fő**, ami azt jelenti, hogy ezek alatt és felett egyaránt a keresetek 50-50%-a található.

b) Határozza meg a T_2 értékét 1.

1. Az adatok sorba rendezése

2. Tercilis sorszámának meghatározása

$$S_{T_2} = 2 \cdot \frac{24+1}{3} = 16,67$$

3. A tercilist tartalmazó osztályköz meghatározása

- Alulról felfelé történő kumulálással történik

Bércsoportok (ezerFt / fő)	Létszám (fő)	Kumulatív gyakoriság
- 100	8	8
100 - 150	7	8+7=15
150 - 200	4	8+7+4=19
200 - 250	3	8+7+4+3=22
250 -	2	8+7+4+3+2=24
Összesen	24	-

b) Határozza meg a T_2 értékét 2.

4. A tercilis értékének megállapítása:

$$K_{T_2} = X_{j_0} + \frac{S_j - \sum f_i}{f_j} \cdot h = 150 + \frac{16,67 - 15}{4} \cdot 50 = 170,88 \text{ eFt / fő}$$

Válasz!!!

A második tercilis értéke **170,88 eFt/fő**, ami azt jelenti, hogy ez alatt a keresetek 66%-a, míg felette a keresetek 33%-a található.

c) Határozza meg a Q_2 értékét 1.

1. Az adatok sorba rendezése
2. Kvartilis sorszámanak meghatározása

$$S_{Q_2} = 2 \cdot \frac{24+1}{4} = 12,5$$

3. A kvartilist tartalmazó osztályköz meghatározása
- Alulról felfelé történő kumulálással történik

Bércsoportok (ezerFt / fő)	Létszám (fő)	Kumulatív gyakoriság
- 100	8	8
100 - 150	7	8+7=15
150 - 200	4	8+7+4=19
200 - 250	3	8+7+4+3=22
250 -	2	8+7+4+3+2=24
Összesen	24	-

c) Határozza meg a Q_2 értékét 2.

4. A kvartilis értékének megállapítása:

$$K_{Q_2} = X_{j_0} + \frac{S_j - \sum f_i}{f_j} \cdot h = 100 + \frac{12,5 - 8}{7} \cdot 50 = 132,14 \text{ eFt / fő}$$

Válasz!!!

A második kvartilis értéke **132,14 eFt/fő**, ami azt jelenti, hogy ez alatt és felett a keresetek 50-50%-a található.

d) Határozza meg a Q_3 értékét 1.

1. Az adatok sorba rendezése
2. Kvartilis sorszámanak meghatározása

$$S_{Q_3} = 3 \cdot \frac{24+1}{4} = 18,75$$

3. A kvartilist tartalmazó osztályköz meghatározása
- Alulról felfelé történő kumulálással történik

Bércsoportok (ezerFt / fő)	Létszám (fő)	Kumulatív gyakoriság
- 100	8	8
100 - 150	7	8+7=15
150 - 200	4	8+7+4=19
200 - 250	3	8+7+4+3=22
250 -	2	8+7+4+3+2=24
Összesen	24	-

d) Határozza meg a Q_3 értékét 2.

4. A kvartilis értékének megállapítása:

$$K_{Q_3} = X_{j_0} + \frac{S_j - \sum f_i}{f_j} \cdot h = 150 + \frac{18,75 - 15}{4} \cdot 50 = 196,88 \text{ eFt / fő}$$

Válasz!!!

A harmadik negyedelő értéke **196,88 eFt/fő**, ami azt jelenti, hogy ez alatt a keresetek 75%-a, míg felette a keresetek 25%-a található.

e) Határozza meg a Q_1 értékét 1.

1. Az adatok sorba rendezése
2. Kvartilis sorszámanak meghatározása

$$S_{Q_1} = 1 \cdot \frac{24+1}{4} = 6,25$$

3. A kvartilist tartalmazó osztályköz meghatározása
- Alulról felfelé történő kumulálással történik

Bércsoportok (ezerFt / fő)	Létszám (fő)	Kumulatív gyakoriság
50 - 100	8	8
100 - 150	7	8+7=15
150 - 200	4	8+7+4=19
200 - 250	3	8+7+4+3=22
250 - 300	2	8+7+4+3+2=24
Összesen	24	-

e) Határozza meg a Q_1 értékét 2.

4. A kvartilis értékének megállapítása:

$$K_{0.25} = X_{j_0} + \frac{S_j - \sum_{f_j} f_j}{f_j} \cdot h = 50 + \frac{6,25 - 0}{8} \cdot 50 = 89,06 \text{ eFt / fő}$$

Válasz!!!

Az alsó negyedelő értéke **89,06 eFt/fő**, ami azt jelenti, hogy ez alatt a keresetek 25%-a, míg felett a keresetek 75%-a található.

Köszönöm a figyelmet!