

5. lecke: Aszimmetria, koncentráció

Gyakorlati feladatsor

1. feladat

Egy uszodában az alábbi távokat úszták az emberek:

Leúszott táv, méter	Úszók száma, fő
1-500	20
501-1000	45
1001-2000	60
2001-3000	12

Számítsa ki és értelmezze az aszimmetria P és a koncentráció Herfindahl és normált Herfindahl mérőszámát és értelmezze is azokat!

2. feladat

Az alábbi táblázat egy iparágban működő szervezetek bevételeit tartalmazza.

Szervezet	Éves nettó árbevétel (ezer Ft)
A	17 034 140
B	14 353 202
C	12 340 400
D	11 234 564
E	10 000 120
F	7 546 200
G	6 534 002
H	4 153 421
I	2 453 450
J	873 421
Összesen	86 522 920

Számítsa ki és értelmezze a koncentráció CR_4 , Herfindahl és normált Herfindahl mérőszámát és az aszimmetria P és F mutatóját! Értelmezze a kapott eredményt!

Ábrázolja az adatok megoszlását Boxploton, rajzolja be a medián vonalát: a doboz melyik oldalához lesz közelebb?



Excel feladat

Egy bank alkalmazottjaira vonatkozóan talál adatokat a bank2.xls állományban. Számítsa ki és értelmezze a bank dolgozóinak jelenlegi fizetésére vonatkozóan

- az aszimmetria P mutatóját,
- a koncentráció Herfindahl mérőszámát és a normált Herfindahl-indexet,
- készítsen Lorenz-görbét az adatok megoszlására vonatkozóan!

Oldja meg a feladatot Excel segítségével! Készítsen részletes szöveges elemzést!

Megoldás

Egy uszodában az alábbi távokat úszták az emberek:

Leúszott táv, méter	Úszók száma, fő	Osztályközép, méter	f_i'
1-500	20	250	20
501-1000	45	750	65
1001-2000	60	1500	125
2001-3000	12	2500	137
Összesen	137		

Számítsa ki és értelmezze az aszimmetria P és F mérőszámait és értelmezze is azokat! Ábrázolja az adatok megoszlását Boxploton, rajzolja be a medián vonalát: a doboz melyik oldalához lesz közelebb?

A P mutató kiszámításához 3 adatra van szükségünk: a leúszott távok átlagára, mediánjára és szórására. A feladat is ismerős lehet a Szóródás leckéből, ezért a keresett 3 számot már könnyedén ki kell tudjuk számolni a korábban tanultak felelevenítésével:

Az átlag kiszámításához elsőként határozzuk meg a leúszott táv osztályközök osztályközépeit, majd ebből számoljuk ki az átlagosan leúszott távot súlyozott számtani átlagformát használva:

$$\bar{x} = \frac{\sum_{i=1}^k f_i * x_i}{\sum_{i=1}^k f_i} = \frac{20 * 250 + 45 * 750 + 60 * 1500 + 12 * 2500}{20 + 45 + 60 + 12} = 1158,759 \text{ méter}$$

Az egyes úszók által leúszott táv átlagosan 1158,759 méter // Az uszodában az emberek átlagosan 1158,759 métert úsztak.

A medián a sorbarendezett adatsor középső eleme. Mivel páratlan elemszámú ismérvértékünk van, 137 darab, így a keresett elemet az $(N+1)/2$ képlettel tudjuk kiszámolni $\rightarrow 138/2=69$, tehát a medián a 69. elem lesz. A leúszott távok kumulált gyakoriságait kiszámítva megállapíthatjuk, hogy a 69. elem az 1001-2000 osztályközben van, amelynek osztályközépe 1500, így ez lesz a nyers medián értéke is, tehát

$Me=1500$

A szóráshoz elsőként számítsuk ki a teljes eltérés négyzetösszeget:

$$\begin{aligned} SST &= \sum_{i=1}^k f_i * (x_i - \bar{x})^2 \\ &= 20(250 - 1158,759)^2 + 45(750 - 1158,759)^2 \\ &\quad + 60(1500 - 1158,759)^2 + 12(2500 - 1158,759)^2 = 52609489,05 \end{aligned}$$

A teljes eltérés négyzetösszeget elosztva az elemszámmal kapjuk meg a szórásnégyzetet, azaz varianciát:

$$\sigma^2 = \frac{SST}{N} = \frac{SST}{\sum_{i=1}^k f_i} = \frac{52609489,05}{20 + 45 + 60 + 12} = 384010,869$$

A szórásnégyzetből gyököt vonva számíthatjuk ki a szórást:

$$\sigma = \sqrt{\sigma^2} = \sqrt{384010,869} = 619,686 \text{ méter}$$

Az egyes emberek által leúszott távok átlagosan 619,686 méterrel térnek el az átlagosan leúszott távtól.

Megvan tehát minden információnk, amire a P kiszámításához szükségünk van:

$\bar{x}=1158,759 \text{ m}$

$Me=1500 \text{ m}$

$$\sigma = 619,686 \text{ m}$$

Helyettesítsünk be a képletbe:

$$P = 3 * \frac{\bar{x} - Me}{\sigma} = 3 * \frac{1158,759 - 1500}{619,686} = -1,652$$

Mivel a P mutató negatív előjelű, így balra elnyúló jobboldali aszimmetriáról beszélünk, azaz az ismértékek többsége átlag feletti.

A Herfindahl-indexet kétféleképpen számíthatjuk ki: a relatív értékösszegek és minta relatív szórása segítségével. Mi most ez utóbbi módszerrel számolunk. A minta relatív szórását a szórás és az átlag hányadosaként tudjuk kiszámolni, amelyeket már az előző részfeladatban megkaptunk.

$$v = \frac{\sigma}{\bar{x}} = \frac{619,686}{1158,759} = 0,53479 \sim 53,48\%$$

Az egyes emberek által leúszott távok átlagosan 53,48%-kal térnek el az összes úszó által átlagosan leúszott távolságtól.

Helyettesítsünk be a Herfindahl-index képletébe:

$$HI = \frac{v^2 + 1}{N} = \frac{0,53479^2 + 1}{137} = 0,009386818$$

A Herfindahl-index értéke 0 és 1 közé eshet, minél nagyobb, annál nagyobb az ismértékek koncentrációja. A kapott érték alapján elmondhatjuk, hogy az egyes úszók által leúszott távolságok koncentrációja alacsony.

Mivel a Herfindahl-index értéke 1 és $1/N$ közé eshet, így nem alkalmas arra, hogy több adatsor koncentrációját is összehasonlíthassuk vele. Ehhez számoljuk ki a normált Herfindahl-index értékét, amely 0 és 1 közé esik. Ez alapján a 0,1 alatti értékeket alacsony, míg a 0,18 feletti értékeket magas koncentrációnak tekinthetjük. Helyettesítsünk be a képletbe:

$$HI^* = \frac{HI - \frac{1}{N}}{1 - \frac{1}{N}} = \frac{0,009386818 - \frac{1}{137}}{1 - \frac{1}{137}} = 0,00210$$

A kapott eredményünk alapján tehát továbbra is elmondhatjuk, hogy az egyes úszók által leúszott távok koncentrációja alacsony.

2. feladat

Az alábbi táblázat egy iparágban működő szervezetek bevételeit tartalmazza.

Szervezet	Éves nettó árbevétel (ezer Ft)	Relatív értékösszeg, Z_i
A	17 034 140	$17034140/86522920=0,19687 \rightarrow 19,687\%$
B	14 353 202	$0,16589 \rightarrow 16,59\%$
C	12 340 400	$0,14262 \rightarrow 14,26\%$
D	11 234 564	$0,12984 \rightarrow 12,98\%$
E	10 000 120	0,115578
F	7 546 200	0,087216
G	6 534 002	0,075518
H	4 153 421	0,048004
I	2 453 450	0,028356
J	873 421	0,010095
Összesen	86 522 920	1

Számítsa ki és értelmezze a koncentráció CR_4 , Herfindahl és normált Herfindahl mérőszámát és az aszimmetria P és F mutatóját! Értelmezze a kapott eredményt!

Ábrázolja az adatok megoszlását Boxploton, rajzolja be a medián vonalát: a doboz melyik oldalához lesz közelebb?



A CR_4 koncentrációs mérőszám megadja az iparágban tevékenykedő négy legnagyobb vállalat piaci részesedését. Ezt úgy számolhatjuk ki, hogy összeadjuk az adott vállalatok relatív értékösszegeit. A relatív értékösszeg jelen esetben azt mutatja meg, hogy az adott vállalat éves nettó árbevétele hányadrésze az adott iparág teljes éves nettó árbevételének. Ezt úgy számolhatjuk ki, hogy elosztjuk a vállalatok éves nettó árbevételeit az iparági összes éves nettó árbevétellel.

A 4 legnagyobb vállalkozás koncentrációjához adjuk össze tehát relatív értékösszegeiket:

$$CR_4 = \sum_{i=1}^4 Z_i = 0,19687 + 0,16598 + 0,14262 + 0,12984 = 0,63531 \sim 63,531\%$$

A kapott érték azt jelenti tehát, hogy a 4 legnagyobb vállalkozás piaci részesedése az adott iparágban 63,531%, amely közepesen koncentrált piacot jelent.

A Herfindahl-indexet az előző feladatban látottaktól eltérő módon is ki lehet számolni. Ehhez adjuk össze a vállalatok relatív értékösszegeinek négyzeteit, a képletbe behelyettesítve:

$$HI = \sum_{i=1}^N Z_i^2 = 0,19687^2 + 0,16598^2 + 0,14262^2 + 0,12984^2 + 0,11558^2 + 0,08721^2 + 0,07551^2 + 0,04800^2 + 0,02835^2 + 0,01009^2 = 0,133382$$

A Herfindahl-index értéke $1/N$ (azaz ez esetben $1/8=0,125$) és 1 közé esik, mivel a Herfindahl-index értéke közelebb van az $1/8$ -hoz, ezért elmondhatjuk, hogy:

A vállalkozások koncentrációja az iparágban viszonylag alacsony.

Normálva a Herfindahl-index értékét összehasonlíthatóvá válnak különböző adatsorok, összehasonlíthatjuk tehát, melyik esetben beszélhetünk nagyobb koncentrációról, ebben a feladatban, vagy az első feladat esetén. Helyettesítsünk be a képletbe:

$$HI^* = \frac{HI - \frac{1}{N}}{1 - \frac{1}{N}} = \frac{0,133382 - \frac{1}{8}}{1 - \frac{1}{8}} = 0,0095794$$

Mivel a kapott normált Herfindahl-index közelebb van a 0-hoz, mint az 1-hez, így igazoltuk, hogy az iparágban a vállalatok koncentrációja alacsony. Összehasonlítva az előző példában kapott 0,002-s értékkel, láthatjuk viszont, hogy bár nincs értelme összehasonlítani egy uszodát egy iparággal, az iparágban a vállalatok koncentrációja magasabb, mint az uszodában az egyes személyek által leúszott távok.

A koncentráció P mutatójának kiszámításához szükségünk van az adatsor átlagára és mediánjára, amelyek számításával már megismerkedtünk a Középtértékek leckében. Mivel ismerjük, mekkora volt az iparág összes nettó bevétele az adott évben, így azt elosztva az ismérvtékek számával, kiszámolhatjuk az iparági éves átlagos nettó bevételt:

$$\bar{x} = \frac{S}{N} = \frac{86522920}{10} = 8652292 \text{ Ft}$$

A kapott eredményünk értelmezése: az iparágban az éves nettó árbevétel átlagosan 8 652 292 Ft.

A medián páratlan számú ismerv esetén a sorba rendezett adatsor középső eleme, páros számú ismerv esetén a két középső elem átlaga. Mivel most 10 vállalkozásunk van, így számoljuk ki a két középső elem –vagyis az E és F vállalkozás éves nettó árbevétele- átlagát:

$$Me = \frac{10000120 + 7546200}{2} = 8773160 \text{ Ft}$$

A vállalkozások felének éves árbevétele tehát legalább vagy legfeljebb 8773160 Ft.

A P mutató kiszámításához szükségünk van még az ismérvtékek szórására is, amelyhez minden információnk meg van, amelyhez csak be kell helyettesítenünk minden ismérvtéket elsőként az SST képletébe:

$$\begin{aligned} SST &= \sum_{i=1}^k (x_i - \bar{x})^2 \\ &= (17034140 - 8652292)^2 + (14353202 - 8652292)^2 + \dots \\ &\quad + (873421 - 8652292)^2 = 2,4973 * 10^{15} \end{aligned}$$

Behelyettesítve a variancia képletébe:

$$\sigma^2 = \frac{SST}{N} = \frac{2,4973 * 10^{14}}{10} = 2,4973 * 10^{13}$$

Ebből gyököt vonva kapjuk meg a szórást:

$$\sigma = \sqrt{\sigma^2} = \sqrt{2,4973 * 10^{13}} = 4997294,957 \text{ Ft}$$

A kapott eredményt értelmezve:

Az iparágban az egyes vállalatok éves nettó árbevételei átlagosan 4997294,96 Ft-tal térnek el az iparági éves átlagos nettó árbevételtől.

Minden adatunk megvan tehát, hogy behelyettesítsünk a P -mutató képletébe:

$$P = 3 * \frac{\bar{x} - Me}{\sigma} = 3 * \frac{8652292 - 87731160}{4997294,957} = -0,072560056$$

A P -mutató értéke negatív és pozitív is lehet, megmutatja, hogy az ismérvtékek többsége átlagon felül vagy alul helyezkedik el. A P -mutató negatív értéke alapján elmondhatjuk, hogy:

Az iparágban a vállalkozók éves nettó árbevételeinek többsége átlag feletti.

Az F -mutató hasonlóan a P -mutatóhoz megmutatja, hol helyezkednek el az ismérvtékek: kiszámításához a medián mellett az első és a harmadik kvartilisre van szükségünk.

Az első és harmadik kvartilis kiszámításához először számoljuk ki hányadik ismérvtékeket keressük, az alábbi képletbe történő behelyettesítéssel:

$$\begin{aligned} s_{\frac{1}{k}} &= \frac{1}{k} (N + 1) \\ s_{\frac{1}{4}} &= \frac{1}{4} (10 + 1) = 2,75 \end{aligned}$$

$$s_{\frac{3}{4}} = \frac{3}{4}(10 + 1) = 8,25$$

Tudjuk, hogy a 2,75. és 8.25. elemeket keressük, ennek ismeretében helyettesítsünk bele a képletbe, amivel kiszámíthatjuk a kvartilisok értékét.

$$x_{i/k} = x_{[i/k]} + \{s_{i/k}\} * (x_{[i/k]+1} - x_{[i/k]})$$

A szögletes zárójelben levő értékek egész ismérveket jelölnek, a kapcsos zárójelben levő értékek törtrészeket. Behelyettesítve a képletbe valahogy így lehet őket értelmezni:

$$x_{2,75} = x_2 + 0,75 * (x_3 - x_2) = 2453450 + 0,75(4153421 - 2452450) = 3729178,25 \text{ Ft}$$

$$x_{8,28} = x_8 + 0,25 * (x_9 - x_8) = 12340400 + 0,25(14353202 - 12340400) = 12843600,5 \text{ Ft}$$

Fontos megemlíteni ennél a lépésnél, hogy bár az adattáblában csökkenő sorrendben láthatók az éves nettó árbevételek, az ismérvértékek sorszáma növekvő sorrendben értelmezendő, az 1. érték tehát a legkisebb, a 10. pedig ebben az esetben a legnagyobb ismérvérték, így a kvartilisek kiszámításakor a növekvő sorrendbe rendezett ismérvértékekkel dolgozunk!

Q_1 : Az iparág vállalkozásainak negyedének éves nettó árbevétele legfeljebb 3729178,25 Ft/ Az iparág vállalkozásainak háromnegyedének éves nettó árbevétele legalább 3729178,25 Ft/

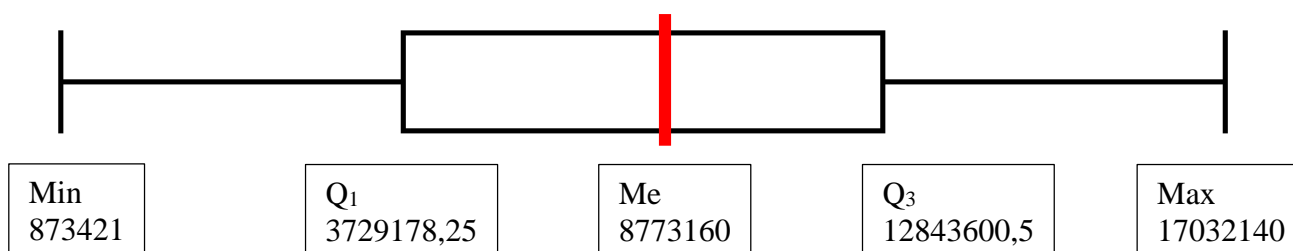
Q_3 : Az iparág vállalkozásainak háromnegyedének éves nettó árbevétele legfeljebb 12843600,5 Ft/ Az iparág vállalkozásainak negyedének éves nettó árbevétele legalább 12843600,5 Ft

Minden információnk megvan tehát az F-mutató kiszámításához:

$$F = \frac{(Q_3 - Me) - (Me - Q_1)}{(Q_3 - Me) + (Me - Q_1)} = \frac{(12843600,5 - 8773160) - (8773160 - 3729178,25)}{(12843600,5 - 8773160) + (8773160 - 3729178,25)} = -0,106813271$$

Az F-mutató hasonló eredményt hozott, mint a P-mutató, mivel értéke ebben az esetben is negatív: Az iparágban a vállalkozók éves nettó árbevételeinek többsége átlag feletti.

A boxplotra mostmár fel tudjuk rajzolni a szükséges értékeket; a boxplot két végpontja a minimum és a maximum, a doboz két oldala pedig az első és a harmadik kvartilis:



Kérdés volt, hogy a doboz „nyelve”, vagyis a medián melyik oldalához lesz közelebb a doboznak, ebből ugyanis sejthetjük, hogy az ismérvértékek többsége átlag feletti vagy alatti lesz. Mivel a harmadik kvartilis és a medián távolsága kisebb (Q_3 -ból kivonva a mediánt kettejük távolsága 4070440,5), mint az első kvartilis és a medián távolsága (a mediánból kivonva a Q_1 értékét 5043981,75-öt kapunk), így a medián a boxplotban a harmadik kvartilishez lesz közelebb, ismét alátámasztva, hogy az iparágban a vállalkozások többségének éves nettó árbevétele átlag feletti.

SZEGEDI TUDOMÁNYEGYETEM
GAZDASÁGTUDOMÁNYI KAR
KÖZGAZDÁSZ KÉPZÉS
TÁVOKTATÁSI TAGOZAT
LECKESOROZAT
COPYRIGHT © SZTE GTK 2017/2018

A LECKE TARTALMA, ILLETVE ALKOTÓ ELEMEI ELŐZETES,
ÍRÁSBELI ENGEDÉLY MELLETT HASZNÁLHATÓK FEL.

JELÉN TANANYAG
A SZEGEDI TUDOMÁNYEGYETEMEN KÉSZÜLT
AZ EURÓPAI UNIÓ TÁMOGATÁSÁVAL.
PROJEKT AZONOSÍTÓ: EFOP-3.4.3-16-2016-00014

SZÉCHENYI 2020



MAGYARORSZÁG
KORMÁNYA

Európai Unió
Európai Szociális
Alap



BEFEKTETÉS A JÖVŐBE