

Maguk az osztályok lehetnek mind diszkrét ismértértékek (erre volt példa a korábbi alfejezetben a panellakásban lévő háztartások száma taglétszám szerint. Abban a példában 50 megfigyelés volt, az ismértértékek 6 különböző értéket vettek fel, ekkor az egyes osztályok ezeknek a különböző értékeknek felelnek meg), mind intervallumok (tartományok), amelyeket **osztályközöknek** nevezünk. A tartományokat mindig ismértérték szerint növekvő sorrendben adjuk meg.

Osztályközök készítésekor először azt kellene eldöntenünk, hogy hány darab, illetve milyen széles osztályokat használjunk. Erre számos megoldási javaslat ismert a gyakorlatban, melyek közül négyet emelnék ki:

1. Véletlen hasra ütésszerűen jelöljük ki az osztályokat.
2. Azonos szélességű osztályokat készítünk. Ekkor az osztályközök k száma az a legkisebb egész szám lesz, melyre először teljesül a $2^k > N$ összefüggés. Azaz például 30 elemű sokaság esetén $k=5$ ($2^5=32$) osztályt érdemes választani. Ez a szabály nem szentírás, csupán ajánlás. Ezek után az osztályok h hossza könnyen adódik az alábbi összefüggésből, melyben az osztályközök száma mellett a legkisebb, illetve a legnagyobb ismértérték szerepel:

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

3. Olyan osztályokat készítünk, melyek mindegyikébe ugyanannyi elem esik. Ezeket a sorokat kvantiliseloszlásoknak nevezzük, és a középértékeknél fogjuk tárgyalni.
4. Változó osztályköz hosszokat alakítunk ki úgy, hogy ahol jobban sűrűsödnek az ismértértékek, ott rövidebb, ahol kevésbé sűrűsödnek, ott hosszabb osztályköz hosszt használunk.

Vigyázzunk, bármelyik technikát is választjuk, e típusú osztályozás információvesztéssel jár, így alaposan meggondolandó, hogy melyik módszert alkalmazzuk.

Osztályközök kialakításakor fokozottan ügyelnünk kell az osztályozás kritériumaira. A teljesség miatt először az ismértértékek tartományát hézagmentesen intervallumokra osztjuk fel. Ezeknek a határait **valódi határoknak** hívjuk. Például a vízállás esetében egy lehetséges felosztás: 0-5m; 5-10m; 10-15m; 15-25m. A tartományok, osztályok hossza lehet azonos, illetve különböző. A valódi határok problémája az átfedés mentesség nem teljesülése: például az 5 méteres vízállás melyik osztályba sorolandó az elsőbe vagy a másodikba? Ennek kiküszöbölésére a táblázatokban nem a valódi, hanem az úgynevezett **közölt határok** szerepelnek, azonban **számolni a valódi határokkal szoktunk**. A közölt határok az adatok pontosságának segítségével kerülnek kialakításra. Például, ha a vízállást 0,1 méter pontossággal mérjük, akkor egy lehetséges felosztás: 0-5,0 m; 5,1-10,0 m; 10,1-15,0 m; 15,1-25,0 m¹. Az i -edik osztály valódi alsó határát $X_{i,0}$; felső határát pedig $X_{i,1}$ szimbólummal jelöljük. Ekkor az i -edik osztály hossza a

$$h_i = X_{i,1} - X_{i,0}$$

összefüggésből adódik. Ha mindegyik osztály ugyanolyan széles, akkor az osztályköz hosszát egyszerűen csak h szimbólummal jelöljük.

¹ Szabályszerűen ezekhez a közölt határokhöz 0-5,05 m; 5,05-10,05 m; 10,05-15,05 m; 15,05-25,0 m valódi határok tartoznak.

Adatok közlésekor előfordulhat úgynevezett nyitott osztályköz használata. Ez azt jelenti, hogy osztályköz alsó vagy felső határa hiányzik. Ez nyilvánvalóan csak az első és az utolsó osztályköz esetében fordulhat elő. Ebben az esetben a számításokhoz az osztályközt zárttá kell tenni. Ez vagy elméleti megfontolásokon alapszik vagy pedig az osztályközt egyszerűen olyan hosszúnak tekintjük, mint a szomszédos osztályközt. Mi ezt a második technikát fogjuk alkalmazni.

Jogosan merül fel a kérdés, hogy osztályközös sorokból hogyan lehet értékösszeget és relatív értékösszeget számolni. A probléma az, hogy osztályközök esetében csak azt tudjuk, hogy egy-egy osztályban hány elem helyezkedik el, de azt nem tudjuk, hogy ezek értéke pontosan mekkora, azaz az x_i értékeket nem ismerjük. Ilyenkor minden osztályból kiveszünk egy reprezentáns értéket, ami nem más, mint az

$$X_i = \frac{X_{i,0} + X_{i,1}}{2}$$

osztályközép. Számításoknál, az ismértértékeket ezekkel az osztályközepekkel helyettesítjük. Az eljárásból következően osztályközös sorok esetében mindent csak becsülni tudunk. Ezt a képletekben úgy fejezzük ki, hogy az adott mennyiség tetejére egy kis kalapot rajzolunk. Ekkor az értékösszeg és a relatív értékösszeg az alábbi képletekkel becsülhető.

$$\hat{S}_i = f_i \cdot X_i$$

$$\hat{Z}_i = \frac{\hat{S}_i}{\hat{S}} = \frac{f_i X_i}{\sum_{i=1}^k f_i X_i} = \frac{g_i X_i}{\sum_{i=1}^k g_i X_i}$$

Példa

Először az alsó osztályközt zárttá alakítjuk, azaz olyan hosszúnak vesszük, mint a másodikat. Majd kiszámítjuk az osztályközepeket az osztályközök valódi határai alapján.

Egy ABC forgalmára vonatkozó adatok

Vásárlások értéke (Ft)	Vásárlások száma f_i	X_i (osztályközép)	$\hat{S}_i = f_i \cdot X_i$	$\hat{Z}_i = \frac{\hat{S}_i}{\hat{S}} (\%)$
– 400	56	200 $= (0+400)/2$	11 200,00 $(= 56 \cdot 200)$	0,74 $(= 11200/1520200 = 0,0074 \rightarrow 0,74)$
401 – 800	89	600 $= (400+800)/2$	53 400,00	3,51
801 – 1200	250	1000	250 000,00	16,45
1201 – 2000	260	1600	416 000,00	27,36
2001 – 2800	210	2400	504 000,00	33,15
2801 – 4000	84	3400	285 600,00	18,79
Összesen	949	–	1 520 200,00	100,00

A második adatok értelmezéséből megállapíthatjuk, hogy a 400 és 800 forint közötti összegért vásárlók 53400 forinttal járultak hozzá az ABC forgalmához. Ez az összes forgalomnak 3,51 százaléka.

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.

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

SZÉCHENYI 2020



Európai Unió
Európai Szociális
Alap



BEFEKTETÉS A JÖVŐBE