4 Tilaston hajontaluvut

Jakauman keskikohdan selvittämisen lisäksi on usein tärkeä tietää, miten ja kuinka laajalle keskikohdan ympärille havaintoarvot ovat levittäytyneet. Tilaston tällaista ominaisuutta mittaamaan tai kuvaamaan on asetettu useitakin hajontalukuja, joista tärkein on **keskihajonta**.

Keskihajontaa määritettäessä lasketaan aina ensin keskiarvo. Tämä siis edellyttää mittaamisessa vähintään välimatka-asteikon tasoa. Keskiarvon määrittämisen jälkeen lasketaan jokaisen havaintoarvon poikkeama keskiarvosta ja korotetaan tämä poikkeama neliöön. Poikkeamien neliöt lasketaan yhteen, ja tämä summa jaetaan havaintoarvojen lukumäärällä. Lopuksi otetaan saadusta osamäärästä neliöjuuri. Tällainen menettely palauttaa mitattavan suureen keskihajonnalle saman dimension (laadun) kuin itse suureellakin on.

****

Määritelmä 1 Keskihajonta on keskiarvosta laskettujen poikkeamien neliöiden keskiarvon neliöjuuri:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}$$

****

Esim. 1 Laske lukujoukon 1, 2, 3, 4, 5, 6 ja 7 keskihajonta.

$$\text{Keskiarvo } \bar{x} = \frac{1+2+3+4+5+6+7}{7} = \frac{28}{7} = 4$$

x	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	-3	9
2	-2	4
3	-1	1
4	0	0
5	1	1
6	2	4
7	3	9
		28

$$\sigma = \sqrt{\frac{28}{7}} = \sqrt{4} = 2$$

Esim. 2 Erään oppilasjoukon matematiikan (x) ja äidinkielen (y) arvosanojen keskiarvot olivat molemmat tasan 7, ja muutoin oheisen taulukon mukaiset:

Taulukko 4. Matematiikkaa ja äidinkieltä

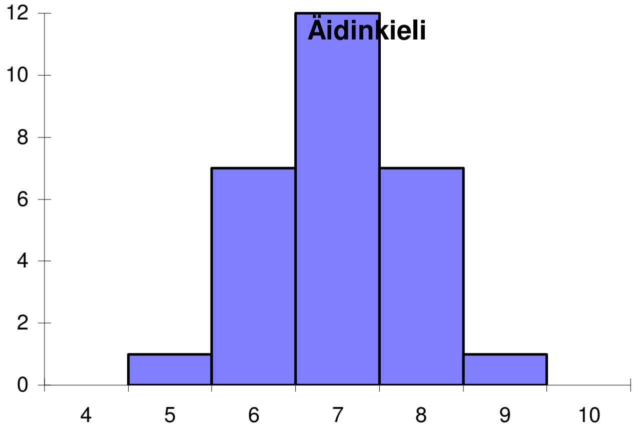
Arvosana	f_x	$x - \bar{x}$	$(x - \bar{x})^2$	$f_x \cdot (x - \bar{x})^2$	f_y	$y - \bar{y}$	$(y - \bar{y})^2$	$f_y \cdot (y - \bar{y})^2$
4	4	-3	9	36				
5	4	-2	4	16	1	-2	4	4
6	2	-1	1	2	7	-1	1	7
7	3	0	0	0	12	0	0	0
8	6	1	1	6	7	1	1	7
9	5	2	4	20	1	2	4	4
10	2	3	9	18				
Summat	26			98,00	26			22

$$\sigma_x = \sqrt{\frac{98}{26}} = 1.941... \approx 1.94 \qquad \sigma y = \sqrt{\frac{22}{26}} = 0.919... \approx 0.92$$

Matematiikan arvosanojen keskihajonta on yli kaksi kertaa niin suuri kuin äidinkielen arvosanojen. Tämä näkyy myös arvosanajakauman kuvaajassa.

Arvosana						
4	4	-3	9	36		
5	4	-2	4	16	1	-2
6	2	-1	1	2	7	-1
7	3	0	0	0	12	0
8	6	1	1	6	7	1
9	5	2	4	20	1	
10	2	3	9	18		
Summat	26			98,00	26	

Matematiikka



Kuva 4. Keskihajonta näkyy jo kuvasta.

Jos keskihajonnan kaavassa neliöjuuren alla olevan lausekkeen nimittäjä n korvataan (n – 1):llä, saadaan ns. **otoskeskihajonta**. Kun tehdään jokin tutkimus, ei ole aina mahdollista tutkia koko perusjoukkoa, jos populaation koko on hyvin suuri, esim. Tampereen kaupungin asukkaat tai Päijänteen ahvenet. Tutkimustulosta, ehkä luotettavaakin, saadaan kyllä aikaan, jos perusjoukosta valitaan **edustava otos** = joukko tilastoyksiköitä, jotka kuvaavat perusjoukkoa mahdollisimman hyvin. Mikäli otos todella on edustava, sitä koskevat tulokset voidaan yleistää kokemaan koko perusjoukkoa. Otoskeskihajontakaavan antama tulos eroaa varsinaisen keskihajonnan kaavan antamasta tuloksesta sitä vähemmän, mitä suurempi n on. Ero on merkityksetön, jos n alkaa olla luvun 30 tietämissä tai sen yli.