Standardavvikelse

Variationsbredd och lådagram är två exempel på spridningsmått, dvs. två sätt att beskriva spridning. Ett tredje spridningsmått är standardavvikelse.

Hur mycket mätvärdena i en datamängd i genomsnitt avviker från datamängdens medelvärde kallas **standardavvikelse** och anges med ett tal.

Ju större standardavvikelse desto större spridning. Ju mindre standardavvikelse desto mindre spridning. Om vi inte har någon spridning alls är standardavvikelsen 0.

Standardavvikelsen beräknas på två olika sätt, beroende på om vi har en totalundersökning (alla i populationen undersöks) eller en stickprovsundersökning (en del av populationen undersöks). Vid en totalundersökning betecknas standardavvikelsen σ (lilla sigma) och vid en stickprovsundersökning (en del av populationen undersöks) betecknas standardavvikelsen *s*.

I den här kursen räcker det att vi kan beräkna standardavvikelsen med ett digitalt verktyg, se exemplet längre fram.

För förståelsens skull tittar vi först på ett enkelt exempel utifrån två datamängder.
Datamängd 1: 16, 17, 17, 17, 18 (liten spridning)
Datamängd 2: 2, 4, 12, 24, 43 (stor spridning)

Medelvärdet för båda dessa datamängder är 17.

Medelvärde för datamängd 1 = $\frac{16+17+17+17+18}{5}$ = 17

Medelvärde för datamängd 2 = $\frac{2+4+12+24+43}{5}$ = 17

Avvikelserna från medelvärdet fås genom att ta varje mätvärde minus medelvärdet.
Medelvärdet skrivs ibland $\overline{x}$. Så $\overline{x}$ = 17.

Avvikelser datamängd 1	Avvikelser datamängd 2
(16 − 17) = **−1**	(2 − 17) = **−15**
(17 − 17) = **0**	(4 − 17) = **−13**
(17 − 17) = **0**	(12 − 17) = **−5**
(17 − 17) = **0**	(24 − 17) = **7**
(18 − 17) = **1**	(43 − 17) = **26**

Vi ser att avvikelserna för datamängd 1 är små. De skiljer sig som **minst 0** och som **mest 1**. Så standardavvikelsen (medelvärdet för avvikelserna) för datamängd 1 kommer att bli ett tal mellan **0** och **1**.

Avvikelserna för datamängd 2 är större. De skiljer sig som **minst 5** och som **mest 26**. Så standardavvikelsen för datamängd 2 kommer att bli ett tal någonstans mellan **5** och **26**, alltså större än för datamängd 1.

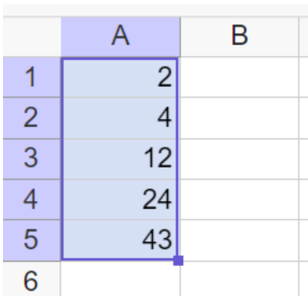
Kuriosa – Standardavvikelse
Förklaringen att standardavvikelsen är medelvärdet av avvikelserna är något förenklad. För att få standardavvikelsen ska avvikelserna dessutom kvadreras och när de har summerats och dividerats med antalet värden ska kvadratroten ur kvoten dras. Detta <u>ingår inte i kursen</u>, men det kan ändå vara intressant att veta vad det är som våra digitala hjälpmedel gör.

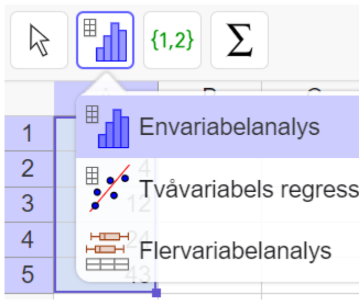
Standardavvikelse för totalundersökning: $$\sigma = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n}}$$

Standardavvikelse för stickprov: $$s = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n - 1}}$$

Ex. Beräkna standardavvikelsen för mätvärdena 2, 4, 12, 24 och 43.

Vi startar Geogebra, klickar på knappen "Växla till kalkylblad" i den översta menyraden och skriver in värdena i den vänstra kolumnen. Vi markerar därefter alla värden, klickar på menyn "blå staplar" och väljer *Envariabelanalys*.





Klicka på knappen "Visa statistik" för att få fram statistiska data.

Statistik

n	5
Medel	17
σ	15.12614
s	16.91153
Σx	85
Σx²	2589
Min	2
Q1	3
Median	12
Q3	33.5
Max	43

Eftersom det inte framgår i uppgiften att det är en stickprovsundersökning, så förutsätter vi att det är en totalundersökning. Vi avläser standardavvikelsen vid σ och avrundar.

Svar: 15,1
(Om det hade varit ett stickprov hade vi avläst standardavvikelsen vid s och svarat 16,9.)