

Signifikans och felkällor

Exempel Har svenska elever blivit bättre i matematik?
För att svara på frågan tar vi hjälp av diagrammet till höger. Det visar svenska elevers resultat vid en internationell undersökning av 15-åringars matematikkunskaper (PISA).

Diagrammet visar att resultaten har förbättrats från år 2012 till år 2018.

I diagrammet ser vi att resultatet år 2009 var 494 poäng.

Mätningarna var stickprovsundersökningar, så vi kan inte vara helt säkra på att resultaten gäller för alla svenska 15-åringar.

felmargin Man kan med statistiska beräkningar (som inte ingår i denna matematikurs) visa att resultat från år 2009 hade en *felmargin* på ± 6 poäng.

Det betyder att om alla 15-åringar i Sverige hade deltagit, hade resultatet med 95 % säkerhet varit mellan 494 ± 6 poäng dvs. mellan 488 och 500 poäng.

konfidensintervall Detta intervall kallas ett *konfidensintervall* på 95 %-nivån.

I diagrammet till höger visas konfidensintervallen med lodräta streck.

signifikant Eftersom intervallen för år 2012 och 2015 inte överlappar varandra är ökningen mellan dessa år statistiskt säkerställd. Man kan också säga att ökningen är statistiskt *signifikant*.

Detsamma gäller ökningen från år 2012 till år 2018.

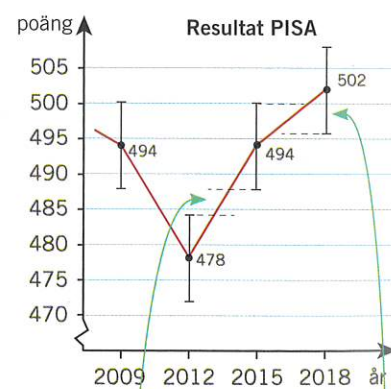
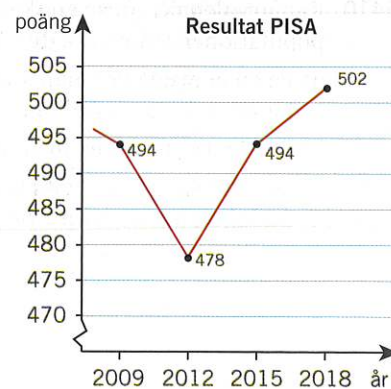
Däremot kan vi inte säga att ökningen är statistiskt säkerställd från år 2015 till 2018 eftersom intervallen överlappar varandra.

Värdet för hela populationen kan t.ex. ha varit 498 poäng både år 2015 och 2018.

Vårt svar på den inledande frågan kan då vara:

"Enligt PISA-undersökningen har svenska 15-åringars matematikkunskaper ökat både mellan år 2012 och 2015 och mellan år 2015 och 2018.

Den första ökningen är statistiskt *signifikant*."



Intervallen överlappar inte varandra.

Intervallen överlappar varandra.

Felkällor

Vid statistiska undersökningar kan det finnas många felkällor. Vi beskriver några av dem här:

- urvalsfel** Om personerna eller mätvärdena i ett stickprov inte är representativa för hela populationen ger detta ett *urvalsfel*.
- stickprovstorlek** Om *stickprovets storlek* är för liten kan slumpen göra att resultatet inte är representativt för hela populationen. Ett större stickprov innebär att *felmargin*alen minskar.
- bortfall** *Bortfall* är ett stort problem vid många undersökningar. Det kan handla om en enkät som är ofullständigt ifylld eller en utvald person som inte vill eller kan vara med i en undersökning.
- mätfel** En undersökning kan ge felaktiga data på grund av svårtolkade frågor i en enkät. En undersökning kan även ge felaktiga data på grund av att mätinstrument visar felaktiga värden eller avläses fel. Dessa typer av felkällor kallas *mätfel*.
- tolkningsfel** Personen som *tolkar* resultaten kan dra felaktiga slutsatser.

5415

Vid en undersökning ställde man följande fråga till 1 230 slumpmässigt valda personer i ett samhälle.
"Kan du tänka dig ett vindkraftverk nära din bostad?"

1 000 personer svarade:
Ja: 380 Nej: 410 Vet ej: 210

Hur stor andel svarade Ja om vi

a) bortser från bortfallet

b) antar att svaren i bortfallet har fördelningen
Ja: 30 % Nej: 30 % Vet ej: 40 %

a) Andelen Ja = $\frac{380}{1000} = 38\%$

Svar: 38 % svarade Ja.

b) Bortfallets storlek = $1230 - 1000 = 230$
30 % av 230 = $0,3 \cdot 230 = 69$

Antalet som skulle svara Ja om vi tar med bortfallet:
 $380 + 69 = 449$

Andelen Ja = $\frac{449}{1230} = 0,365 \approx 36,5\%$

Svar: 36,5 % antas svara Ja.

