

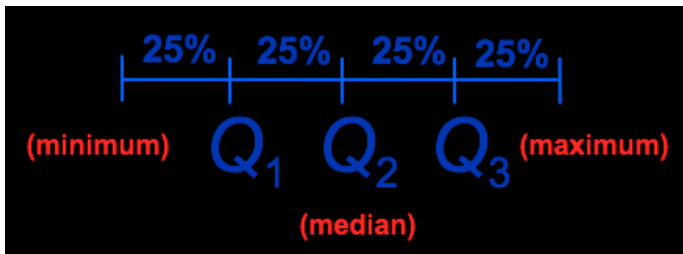
Kvartiler

Kvartiler

Sorter dataene.

- Nedre kvartil Q_1 , skiller de nedre 25% fra de øvre 75%
- Midterste kvartil Q_2 , skiller de nedre 50% fra de øvre 50% (Medianen)
- Øvre kvartil Q_3 , skiller de nedre 75% fra de øvre 25%

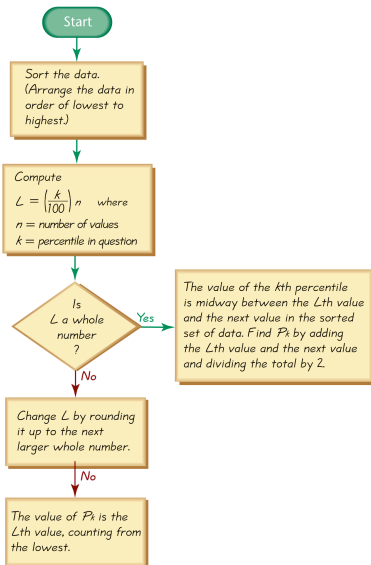
Kvartilene deler de sorterte dataene i fire like store deler.



Persentiler

Persentil

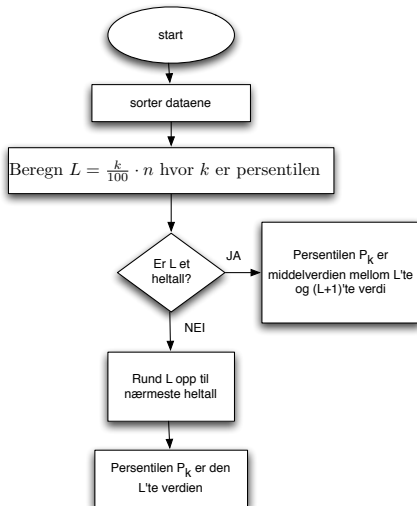
- Kvartiler gir firdeling
- Med persentiler deler man opp i 100 like store deler
- P_1, P_2, \dots, P_{99}
- $Q_1 = P_{25}$ og $Q_3 = P_{75}$
- Bruk Figur 3-6



Persentiler

Persentil

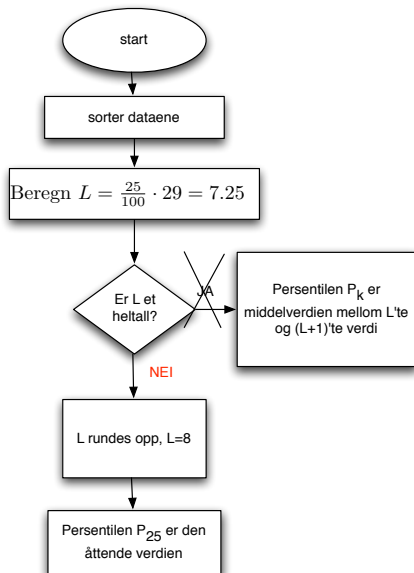
- Mange måter å beregne kvartiler/persentiler på
- SPSS, JMP, Excel er uenige
- Vi bruker Figur 3-5 i boka



Kvartil eksempel

Example

- Vi ville finne nedre kvartil
 $Q_1 = P_{25}$
- Stikkprøven har størrelse
 $n = 29$
- Da er $k = 25$ i Figur 3-6
- Konklusjon: Nedre kvartil er
8. verdi fra bunn



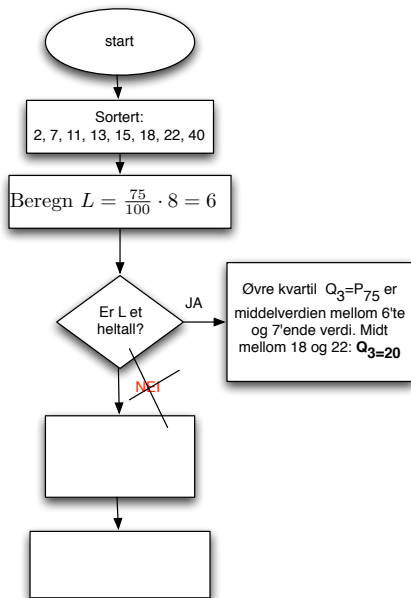
Kvartil eksempel

Example

Stikkprøven:

13 22 15 18 11 7 2 40

- Finn øvre kvartil $Q_3 = P_{75}$
- Stikkprøven har størrelse $n = 8$
- Da er $k = 75$ i Figur 3-6
- Øvre kvartil ligger midt mellom 6. og 7. verdi
- Konklusjon: $Q_3 = 20$



En uteligger (ekstrem observasjon) kan ha drastisk effekt på

- Gjennomsnittet
- Standardavviket
- På skalaen til histogrammet, slik at det ikke gir et godt bilde av fordelingen

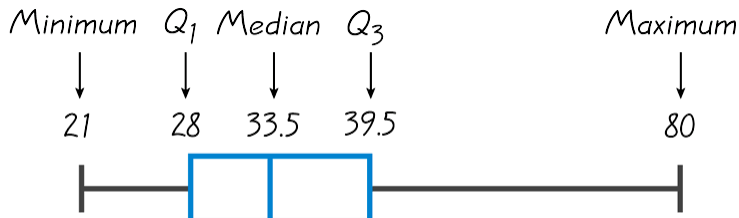
5-talls oppsummering og boksplott

5-talls oppsummering

Minimumsverdi, Nedre kvartil, Median, Øvre kvartil og maksimumsverdien

Boksplott

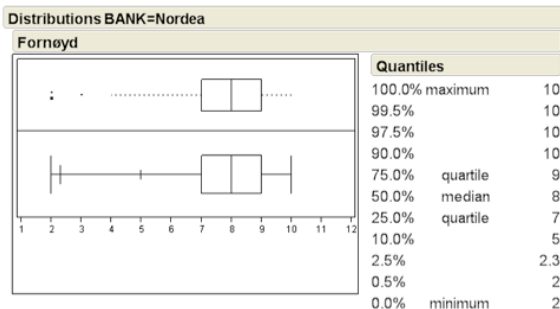
Er en graf som viser 5-talls oppsummeringen.



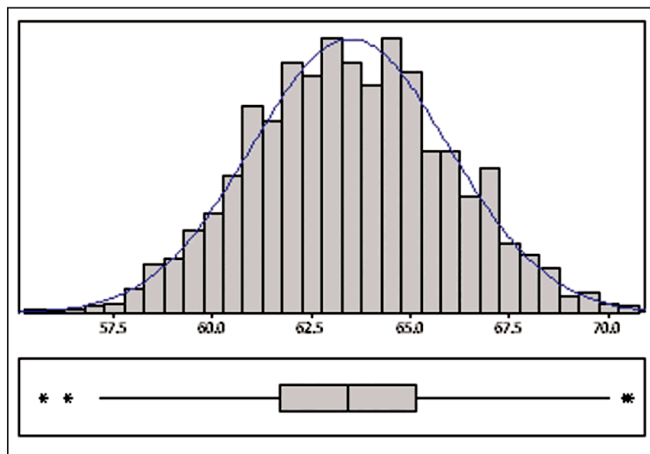
Boksplott i JMP

JMP har to boksplott

- JMPs outlier box plot er det som boka (section 3-5) kaller *Modified box plot*. Det brukes til å oppdage uteliggere.
- JMPs quantile box plot er det som boka kaller *box plot* og som ble vist på forrige slide.

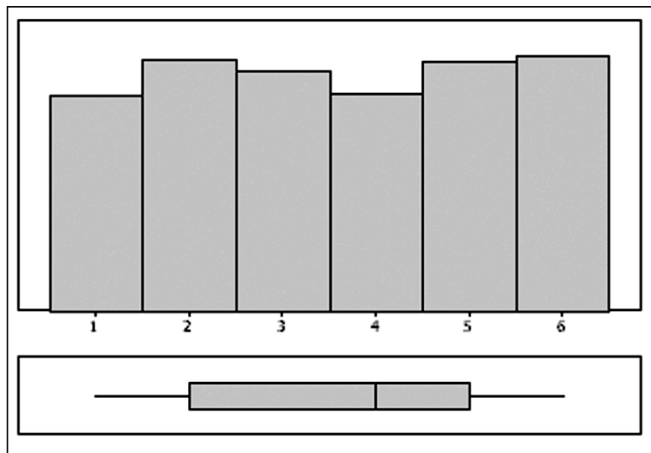


Boksplott for normalfordeling



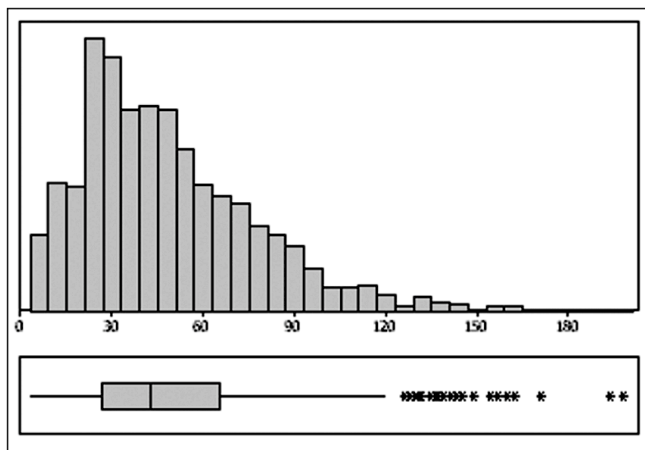
(a) Høyde (inch) for 10000 kvinner. Normalfordelt

Boksplott for uniform fordeling



(b) 1000 terningkast. Uniform fordeling

Boksplott for høyreskjev fordeling



(c) Månedslønn til 1000 markedsførere. Høyreskjev.

Outlier box plot

BoksploTT for å oppdage uteliggere

Disse boksploTTene markerer en verdi som uteligger.

- Kvartilbredden er definert som $IQ = Q_3 - Q_1$
- Markert som uteligger dersom større enn $Q_3 + 1.5 \cdot IQ$
- Eller mindre enn $Q_1 - 1.5 \cdot IQ$

