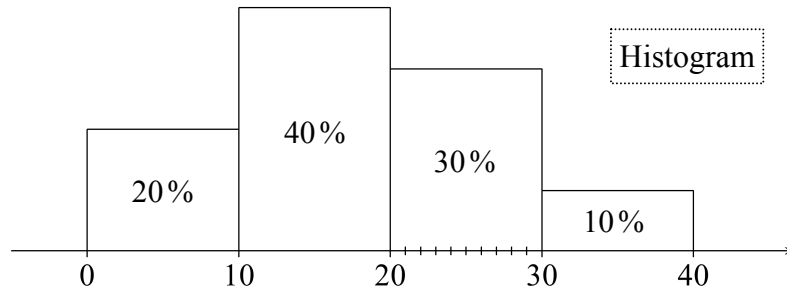


Sumkurve og lineær sammenhæng

Histogrammet viser et grupperet observationssæt:



Intervalleret 20-30 deler vi op i 10 lige store dele (se figur).

Hver af disse små intervaller må indeholde en tiendedel af hele intervallets observationer, dvs. de indeholder hver 3 % af samtlige observationer.

(x, y) er et punkt på sumkurven, dvs.

y er den procentdel af observationerne der har størrelse x eller derunder.

Af histogrammet ovenfor ser vi:

$$\text{Når } x = 20 \text{ er } y = 0,20 + 0,40 = 0,60$$

$$\text{Når } x = 21 \text{ er } y = 0,60 + 0,03 = 0,63$$

$$\text{Når } x = 22 \text{ er } y = 0,63 + 0,03 = 0,66$$

Hver gang x bliver 1 større, vil y blive 0,03 enheder større, så y vokser lineært i intervallet fra $x = 20$ til $x = 30$.

Derfor er grafen en ret linje i dette interval, og ligningen er

$$y = 0,03x + b.$$

Vi udregner b :

$$\text{Når } x = 20 \text{ er } y = 0,60 \text{ så}$$

$$0,60 = 0,03 \cdot 20 + b.$$

Heraf ser vi at $b = 0$, så ligningen er

$$y = 0,03x.$$

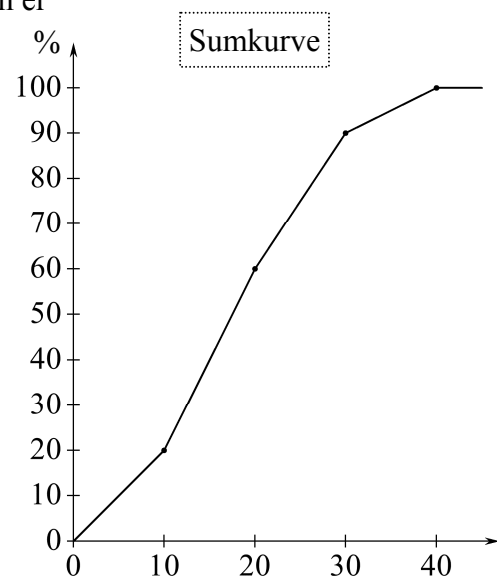
For de fire intervaller er ligningerne:

$$0-10: \quad y = 0,02x$$

$$10-20: \quad y = 0,04x - 0,2$$

$$20-30: \quad y = 0,03x$$

$$30-40: \quad y = 0,01x + 0,6$$



Hvor mange procent af observationerne har størrelse 27 eller derunder?

Vi ser at vi skal bruge ligningen fra tredje interval:

$$y = 0,03 \cdot 27 = 0,81$$

dvs. 81 % af observationerne er 27 eller derunder.

Hvor stor er nedre kvartil?

Vi skal gå ud fra 25 % på y-aksen. Vi ser at vi skal bruge ligningen fra andet interval:

$$0,25 = 0,04x - 0,2.$$

Vi løser denne ligning mht. x og får 11,25,

dvs. nedre kvartil er 11,25.