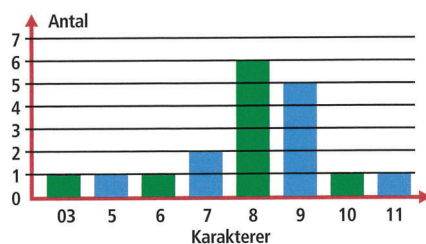


- 1 Find gennemsnittet i hvert af følgende talsæt.

- a 2, 5, 4, 1, 1, 6, 4, 5, 3, 4
b 63, 49, 50, 62, 48, 33, 55, 48, 51, 46, 43
c 8, 4, -2, 5, 12, -1, 0, 0, -6, 8, 4, 7

- 2 Find medianen i hvert af følgende talsæt.

- a 5, 5, 7, 8, 9, 10, 11
b 5, 5, 7, 8, 9, 10
c 8, 4, -2, 5, 12, -1, 0, 0, -6, 8, 4, 7



- 3 a Hvad viser diagrammet?
b Hvilke oplysninger var nødvendige for at fremstille diagrammet?
c Hvor mange elever deltog?
a Beskriv andre tal, der er "skjult" i diagrammet.



- 4 I 9.z ville man finde ud af, hvilken farve der var den enkelte elevs yndlingsfarve. Eleverne kunne vælge mellem, rød, gul, blå, grøn, brun og sort.

Svarene var: rød, blå, rød, gul, rød, sort, gul, rød, brun, grøn, blå, rød, sort, gul, blå, sort, rød, blå, rød og sort.

- a Foretag en optælling.
b Hvilken farve foretrækkes af flest?
c Hvilken farve foretrækkes af færrest?
d Vis optællingen grafisk.
e Hvad er gennemsnitsfarven? Begrund dit svar.

- 5 I 9.x har eleverne noteret, hvor mange musikanlæg de har derhjemme:

5, 6, 2, 3, 5, 4, 7, 4, 2, 5, 8, 4, 1, 7, 6, 5, 3, 4, 3, 2

- a Foretag en optælling.
b Hvilke oplysninger giver denne optælling?
c Hvad er det gennemsnitlige antal? Begrund dit svar.
d Vis optællingen grafisk.
e Beskriv undersøgelsens resultater.



- 6 I 9.y er der 24 elever, som også har foretaget en optælling af musikanlæg. Efter at have behandlet disse data statistisk præsenterer eleverne bl.a. følgende resultater:

- 25% af eleverne har 4 musikanlæg. Lige så mange elever har 5.
- Alle elever har mindst 2, og halvdelen har mere end 4 anlæg.
- Det højeste antal musikanlæg er 7 hos 2 elever.
- Der er dobbelt så mange elever der har 3 musikanlæg, end der er elever, der har 2 musikanlæg.

Kan du regne dig "tilbage" til klassens optælling?



- 7 På Pærefabrikken Luma Spot ønsker man at tjekke kvaliteten af de producerede pærer. Da man jo ikke kan afprøve alle, har man taget én pære fra hver af produktionens 30 kasser, som hver indeholder 250 pærer.

- Opstil en tabel over hyppigheden af de enkelte observationer og find mindsteværdi, størsteværdi, typetal, gennemsnit og median. Vurder pærernes kvalitet på det grundlag.
- Hvad er forskellen på observationerne i denne undersøgelse sammenlignet med dem i 9.x og 9.y?
- Hvordan ville du foretage optællingen for Luma Spot?



Brændetid i timer:

1051, 1069, 1049,
1099, 1038, 1031,
1097, 1085, 1042,
1034, 1110, 1026,
1097, 1010, 1108,
1031, 1108, 1052,
1109, 1010, 1075,
1058, 1037, 1056,
1022, 1008, 1017,
1036, 988, 1065

GRUPPEREDE OBSERVATIONER

Når man skal beskrive observationssæt med mange forskellige observationer, kan det ofte gøre observationssættet mere overskueligt at gruppere observationerne i intervaller. I den forbindelse må man tage stilling til nogle forhold.

$a;b$	$h(a;b)$
$[0;5[$	8
$[5;10[$	20
$[10;15[$	15
$[15;20[$	7
$[20;25[$	5
SUM:	55

$a;b$	$im(a;b)$
$[0;5[$	2,5
$[5;10[$	7,5

$$\begin{array}{r}
 h(a;b) \cdot im(a;b) \\
 20 \\
 150 \\
 187,5 \\
 122,5 \\
 112,5 \\
 \hline
 \text{Sum: } 592,5
 \end{array}$$

$$\begin{array}{r}
 \text{Modelleret} \\
 \text{gennemsnit} \\
 592,5 \\
 55 \\
 \hline
 \approx 10,8
 \end{array}$$

Antallet af intervaller er et skøn. Det afhænger af variationsbredden og hyppigheden af de enkelte observationer. Med få intervaller mister man meget information om observationernes beliggenhed, med mange intervaller mister grupperingen sin overskuelighed.

Intervalbredden er afstanden mellem de to intervalyderpunkter, a og b .

Intervalyderpunkter

Man kan vælge, at a skal være med, og b ikke skal være med: Det skrives $[a;b[$ $a \leq x < b$ (x er en observation i intervallet).

eller

Man kan vælge, at a ikke skal være med og b skal være med: Det skrives $]a;b]$ $a < x \leq b$ (x er en observation i intervallet).

BESKRIVELSE AF GRUPPEREDE OBSERVATIONER

Ved grupperede observationer kender man kun antallet af observationer i hvert interval. Man ved ikke hvor i et interval, den enkelte observation befinder sig – den information ses der bevidst bort fra for at gøre observationssættet mere overskueligt.

Gennemsnit

Gennemsnittet af et sæt enkeltobservationer findes ved at dividere summen af alle observationer med antallet af observationer.

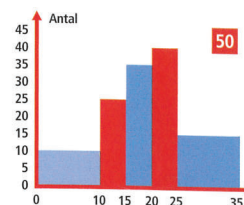
Ved grupperede observationer anvender man intervalmidtpunkterne, $im(a;b)$ til at få et bud på summen af observationsværdierne i hvert interval. Hvis man antager, at observationerne i hvert interval er jævnt fordelt omkring midtpunktet, kan denne sum beregnes som intervalhyppighed gange intervalmidtpunkt. Ved at gøre dette for hvert interval, og lægge disse tal sammen, får vi et bud på summen af alle observationsværdierne.

Hvis denne sum divideres med det samlede antal observationer, får vi det modellerede gennemsnit.

Diagrammer

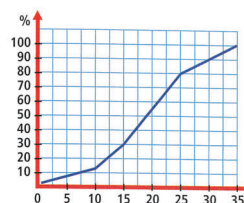
Histogram. Et histogram viser frekvensen eller hyppigheden i hvert interval. I et histogram er det arealet af rektanglet over hvert interval, der bruges som mål for frekvensen eller hyppigheden og ikke højden af rektanglerne. Enheden angives derfor ved et stykke areal, ikke ved valg af enhed på andenaksen, som bare er en talakse.

Da arealet svarer til den frekvens eller hyppighed, man vil angive, og bredden er intervallets bredde, findes højden af hvert rektangel som frekvensen eller hyppigheden divideret med intervalbredden.



Sumkurve. En sumkurve viser den summerede frekvens eller hyppighed.

Ved grupperede observationer kender man kun den summerede frekvens eller hyppighed ved hvert intervals størsteværdi. For at kunne aflæse et bud på den summerede frekvens, $F(x)$, eller hyppighed, $H(x)$ overalt, forbinder man de punkter, der svarer til størsteværdierne, med rette linjer. Man antager således, at observationerne er jævnt fordelt i hvert interval.



Den sumkurve, man på denne måde kan tegne for grupperede observationer, er en matematisk model – de virkelige summerede frekvenser eller hyppigheder kan man ikke beregne, når først observationerne er grupperede.

Deskriptorer ved grupperede observationer

x: Mulig observation.

$h(a;b)$: Hyppigheden af observationer i intervallet $a;b$.

$f(a;b)$: Frekvensen af observationer i intervallet $a;b$. Findes ved at dividere $h(a;b)$ med det samlede antal observationer.

Typeinterval: Det interval der indeholder flest observationer.

Størsteværdi: Slutværdien (b) i sidste interval.

Mindsteværdi: Startværdien (a) i første interval.

Variationsbredde: Forskellen mellem størsteværdi og mindsteværdi.

Modelleret gennemsnit: For hvert interval ganges $h(a;b)$ med $im(a;b)$, produkterne lægges sammen, og summen divideres med det samlede antal observationer.

$\bar{H}(x)$: Modelleret summeret hyppighed – Antallet af observationer mindre end eller lig med x ifølge model-sumkurven.

$\bar{F}(x)$: Modelleret summeret frekvens – Andelen af observationer mindre end eller lig med x ifølge model-sumkurven. Findes ved at dividere $\bar{H}(x)$ med det samlede antal observationer.

1. kvartil: Den mindste observation hvor $\bar{F}(x) \geq 0,25$

2. kvartil (Median): Den mindste observation hvor $\bar{F}(x) \geq 0,50$

3. kvartil: Den mindste observation hvor $\bar{F}(x) \geq 0,75$

Kvartil kommer af "kvart", lige som "et kvarter" ($\frac{1}{4}$ af 1 time).