

# Taldata 1. Chancer gennem eksperimenter

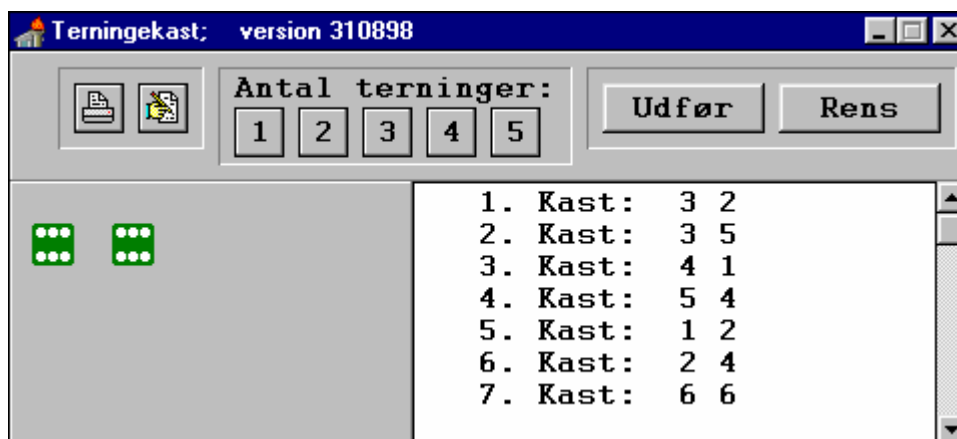
## Indhold

1. Kast med to terninger
2. Et pindediagram
3. Sumtabel
4. Median og kvartiler
5. Et trappediagram
6. Gennemsnit
7. En statistik
8. Anvendelse af edb
9. Opgaver

## 1. Kast med to terninger

*Du skal kaste med to terninger. Du skal kaste indtil de to terninger viser samme øjental, altså indtil de fx begge viser 2 eller begge viser 5. Du skal tælle op hvor mange kast du måtte udføre for at få samme øjental på de to terninger.*

*Når du skal udføre terningkast, kan du vælge mellem to muligheder: Du kan kaste med rigtige terninger, eller du kan bruge et edb-program der hedder KAST.*



## Hjælpeprogrammet KAST

Du kan hurtigt sætte dig ind i brugen af KAST. Prøv fx at vælge kast med to terninger. Klik derefter på knappen for *Udfør*. På skærmen kommer nu et billede af to terninger som kastes. Resultatet af kastet overføres til en liste, så du har en oversigt over hvad terningerne viste i de enkelte kast.

Klik på *Udfør* igen indtil du har et kast hvor de to terninger viser samme øjental.

Med *Rens* kan du nu gøre klar til en ny serie af kast. Prøv at se hvor mange kast der nu skal til, før du har et kast hvor terningerne igen viser sammen øjental.

## Sannes datasæt

Inden du går i gang, ser vi på de data Sanne har indsamlet.

Sanne deltager i et spil hvor der kastes med to terninger. For at komme ind i spillet gælder det om at få et terningkast hvor de to terninger viser sammen øjental.

Sanne er godt klar over at somme tider får man to ens allerede i første kast, men det sker også at man må udføre mange kast før de to terninger viser det samme. Sanne er derfor interesseret i at undersøge hvor svært det er at få to ens i et kast med to terninger.

Sanne vælger at udføre 50 eksperimenter. I hvert eksperiment kaster hun de to terninger indtil hun opnår to ens øjental.

Hun fører regnskab over hvor mange kast hun udfører i hvert af de 50 eksperimenter.

Her er hendes resultater:

```
2 - 1 - 12 - 7 - 1 - 2 - 9 - 1 - 2 - 3
3 - 4 - 1 - 14 - 6 - 1 - 8 - 3 - 5 - 10
6 - 3 - 3 - 1 - 7 - 4 - 4 - 11 - 2 - 7
15 - 3 - 5 - 1 - 7 - 12 - 6 - 2 - 4 - 5
5 - 2 - 10 - 3 - 14 - 1 - 1 - 6 - 8 - 4
```

I det første eksperiment skulle Sanne blot bruge 2 kast for at få to ens øjental. I det andet eksperiment kunne hun klare sig med kun 1 kast, men så måtte hun i det tredje eksperiment ud i 12 kast før det lykkedes hende at få to ens.

Af oversigten kan vi se at det tit forekommer at Sanne kan klare sig med 1 kast. Det sker i 9 af de 50 eksperimenter.

## En grundtabel

Sanne vil nu skaffe sig overblik over resultaterne. Hun sorterer sine 50 data og opstiller en tabel over dem:

Antal kast	Hyppighed	Frekvens
1	9	18%
2	6	12%
3	7	14%
4	5	10%
5	4	8%
6	4	8%
7	4	8%
8	2	4%
9	1	2%
10	2	4%
11	1	2%
12	2	4%

14	2	4%
15	1	2%
I alt: 50		100%

En sådan tabel over vore data vil vi kalde *en grundtabel* eller *en frekvenstabel*.

Tabellen giver hyppigheder og frekvenser for de data vi har indsamlet:

Under *hyppigheder* ser vi hvor mange gange de enkelte kasteantal er forekommet i serien af eksperimenter. For eksempel kan vi se at Sanne i 7 eksperimenter måtte bruge 3 kast for at opnå to ens øjental. Summen af alle hyppigheder i tabellen skal her være 50, der er jo udført 50 eksperimenter.

Under *frekvenser* i tabellens sidste kolonne kan vi se at de 7 eksperimenter svarer til 14% af de 50 eksperimenter der blev udført. Den sidste kolonne indeholder et procenttal for hvor tit de enkelte kastetal forekommer. Summen af procenterne skal selvfølgelig være 100%.

### Typetal

Af tabellen ser vi at den observation der forekommer flest gange, er 1. Den forekommer jo 9 gange, og ingen af de andre observationer når så højt op. En observation der har større hyppighed end de øvrige, kaldes *typetallet* i datasættet. I Sannes datasæt er typetallet altså 1.

Frekvenstabellen er grundstenen i Sannes statistiske undersøgelse. Den giver en god oversigt over hendes observationer.

Vi kan af tabellen se at den største observation er 15. Det største antal kast der blev brugt for i et af de udførte eksperimenter var altså 15. At den mindste observation ved eksperimentet er 1, har vi allerede set. Den største og den mindste observation fortæller os hvor bredt et område resultaterne i eksperimenterne spreder sig over.

Afstanden mellem de to yderste observationer kaldes *variationsbredden*. I Sannes datasæt med de 50 observationer er variationsbredden  $15 - 1$ , altså 14.

### Prøv selv!

1. Du skal udføre 50 eksperimenter hvor du kaster med to terninger. Du holder regnskab med hvor mange kast du må udføre for at få to ens øjental.

Hvis du arbejder med rigtige terninger, kan arbejdet godt tage nogen tid, men du kan måske slå dig sammen med en kammerat så I deler arbejdet imellem jer.

Hvis du bruger edb-programmet KAST, kan du sagtens selv klare at indsamle de data der skal til.

*Vigtigt:* Husk at du selv skal notere hvor mange kast der blev kastet i hver serie for at få to ens øjental. Når du først har brugt *Rens*, er listen forsvundet.

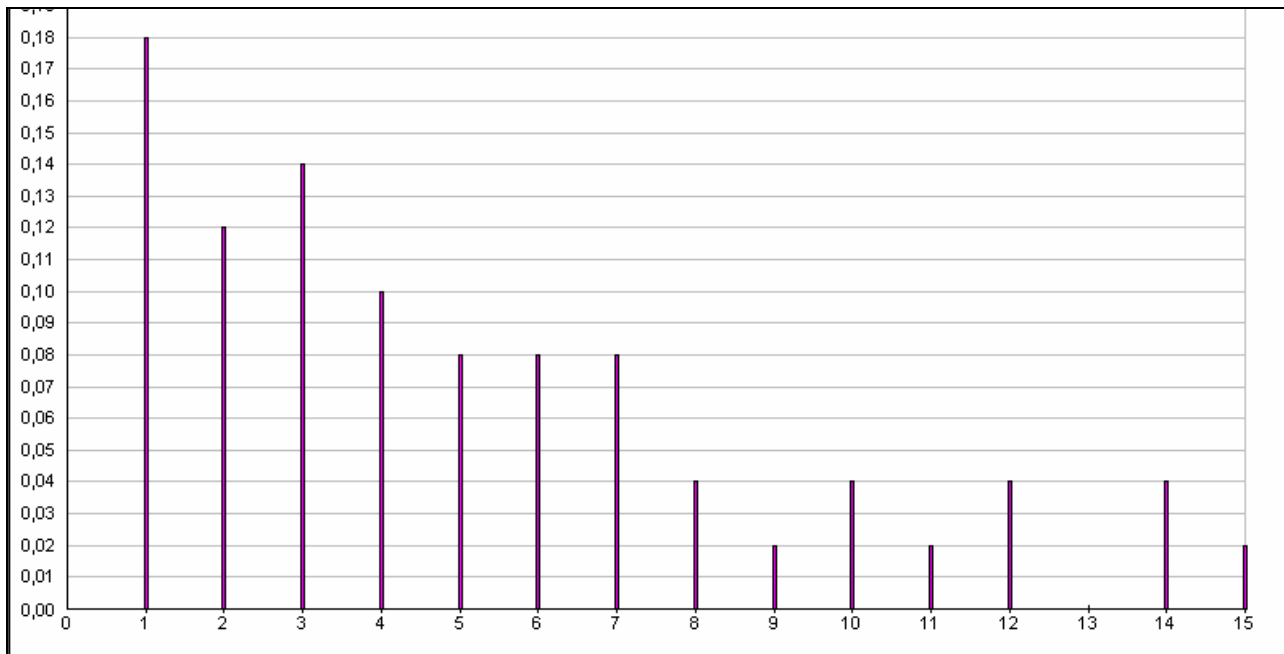
2. Opstil en grundtabel over dine 50 resultater.
3. Sammenlign din tabel med Sannes. Hvilke forskelle er der?  
Synes du, forskellene er væsentlige, eller er der kun tale om småting?
4. Hvad er typetallet i din tabel? Og hvad er variationsbredden?

*Nye begreber.* Lav en liste over de nye begreber du har arbejdet med.

## 2. Et pindediagram

Du har nu opstillet en grundtabel over de data der er indsamlet i de 50 eksperimenter. Af tabellen kan du få oplysninger om kasteantallene i eksperimenterne.

Men vi vil også godt have en tegning, et grafisk billede, hvor vi i ét blik kan få en god oversigt over resultaterne. Vi laver derfor et *pindediagram* over Sannes datasæt.



Pindenes højde viser hvor tit de forskellige kasteantal forekom. På den lodrette skala til venstre i diagrammet kan du aflæse frekvenserne.

### Hvad kan tabel og diagram bruges til?

Grundtabellen og pindediagrammet er en god hjælp for Sanne når hun vil sige noget om vanskeligheden ved at få to ens øjental.

Sanne vil mene at chancen for at få to ens allerede i første kast er ca. 18%. Hun påstår ikke at chancen er lige præcis 18%, for hun er godt klar over at hun ikke kan sige noget sikkert ud fra de 50 eksperimenter. De kan kun give hende en fornemmelse af hvad de ukendte chancer er. De kan hjælpe hende til at opstille nogle gæt som ikke er grebet ud af luften, men som er begrundet i data der er indsamlet gennem eksperimenter.

Sanne gætter altså på at chancen for at få to ens i første kast er ca. 18%, og hvis hun blot fortæller at hun gætter, og at hendes gæt er foretaget på grundlag af 50 eksperimenter, er alt i orden. Så er der ingen fare for at vi skal lægge mere i hendes udtalelser end de kan bære.

Når du selv udfører de 50 eksperimenter og kommer til andre resultater, må du blot konstatere at dine gæt om de ukendte chancer afviger fra Sannes. Nye eksperimenter kan så måske skaffe mere sikre oplysninger, og lange serier af eksperimenter er at foretrække frem for korte serier. Men selv nok så lange serier af eksperimenter kan ikke fortælle os

hvad "den rigtige chance" er. Heldigvis har man i praksis aldrig behov for at få besvaret det spørgsmål. Her er man tilfreds med et godt bud på hvad den ukendte chance er.

I praktiske anvendelser behøver man således ofte blot at vide at chancen fx er ca. 18%, eller at den er lidt under 20%, eller at den er 15-20%. Derfor kan en serie af eksperimenter i mange tilfælde give os alle de oplysninger vi har behov for.

### **Prøv selv!**

1. Lav et pindediagram over dine egne data.
2. Hvad viser dine data om chancen for at få to ens i første kast?

*Nye begreber.* Tilføj de nye begreber til din liste.

### 3. Sumtabel

Grundtabellen og pindediagrammet kan være en hjælp for Sanne når hun vil skaffe sig et billede af chancen for de enkelte kasteantal. Altså når hun fx vil vurdere hvad chancen er for at der skal bruges 5 kast. Men i praksis er det ikke så meget chancen for de enkelte kasteantal, man kigger på. Her er man mere interesseret i at få kendskab til hvad chancen fx er for at man kan klare sig med *højst* 2 kast, eller hvad chancen er for at der skal bruges *højst* 5 kast eller *mindst* 10 kast.

Vi indfører nu et vigtigt hjælpemiddel i arbejdet med taldata og statistik. Det er en tabel, en *sumtabel*, som vil være til stor nytte når vi skal have overblik over indsamlede data.

Vi ser på en sumtabel for Sannes datasæt:

Antal kast	Hyppighed	Summeret hyppighed	Summeret frekvens
1	9	9	18%
2	6	15	30%
3	7	22	44%
4	5	27	54%
5	4	31	62%
6	4	35	70%
7	4	39	78%
8	2	41	82%
9	1	42	84%
10	2	44	88%
11	1	45	90%
12	2	47	94%
14	2	49	98%
15	1	50	100%

Denne tabel kan give svar på Sannes spørgsmål. Ud for 2 kast ser vi at der er anført en "Summeret hyppighed" på 15. Det er de to hyppigheder for 1 kast og 2 kast der her er lagt sammen:  $9+6 = 15$ . At den summerede hyppighed er 15, betyder altså at i 15 af de 50 eksperimenter skulle man bruge *højst* 2 kast for at slå to ens. Chancen for at man kan klare sig med 2 kast er derfor ca.  $15/50$ , dvs. ca. 30%. Det er det tal vi finder i tabellens højre kolonne.

Af sumtabellen kan Sanne også få en ide om hvad chancen er for at man skal bruge op til 5 kast, dvs. *højst* 5 kast. Den summerede hyppighed er her:  $9+6+7+5+4 = 31$ . Ved 31 af de 50 eksperimenter er der altså blevet brugt *højst* 5 kast. Chancen for at man kan klare sig med *højst* 5 kast vil Sanne derfor sætte til ca.  $31/50$ , dvs. 62%.

#### Hvad kan sumtabellen bruges til?

Sumtabellen er et godt hjælpemiddel når Sanne vil udtale sig om vanskeligheden ved at få to ens øjental. Hun kan benytte tabellen på flere måder. Hun kan kigge i tabellen *før*

hun begynder at udføre et kasteeksperiment, og hun kan da få oplysning om hvad chancen er for at hun kan klare sig med et bestemt antal kast. Fx kan hun se i tabellen at hun har ca. 44% chance for at skulle bruge højst 3 kast for at få to ens.

Hun kan også se i tabellen at der er ca. 70% chance for at hun kan klare sig med 6 kast. Det vil sige at der er en chance (eller risiko) på ca. 30% for at hun skal bruge mere end 6 kast for at slå to ens.

Men hun kan også bruge sumtabellen *efter* at hun har udført sine kast og opnået to ens. Hvis hun for eksempel har fået to ens ved at bruge 2 kast, kan hun i tabellen se at dette resultat hører til blandt "*de bedste 30%*", eller "*de heldigste 30%*". Af tabellen fremgår jo at i 30% af Sannes 50 eksperimenter kunne man klare sig med 2 kast.

Har hun i en anden situation måttet bruge mere end 11 kast for at få to ens, kan hun se i sumtabellen at dette resultat hører til blandt "*de dårligste 10%*" eller "*de uheldigste 10%*". Af tabellen fremgår jo at man i 90% af Sannes eksperimenter kunne klare sig med højst 11 kast, så kun i 10% af tilfældene måtte der bruges mere end 11 kast.

Når Sanne benytter sumtabellen efter at hun har udført de kast der var nødvendige for at slå to ens, kan hun få oplysninger om hvor heldig eller uheldig hun har været med sine kast. Har hun kun brugt 2 kast vil hun nok synes at hun har været heldig. Har hun måttet bruge over 11 kast vil hun sikkert mene at hun har været uheldig.

### **Prøv selv!**

1. Opstil en sumtabel over dine data.
2. Hvad viser din sumtabel om chancen for at bruge højst 2 kast?
3. Hvad er chancen for at bruge højst 5 kast?
4. Hvad er chancen for at bruge mere end 10 kast?
5. Hvilke kasteantal hører til de bedste 30% i din tabel?

*Nye begreber.* Tilføj de nye begreber til din liste.



## 4. Median og Kvartiler

Når en spiller skal afgøre om han har været heldig eller uheldig med sit kasteantal, kan han se i Sannes sumtabel eller i din sumtabel hvordan hans resultat er placeret der. Men ofte er han slet ikke interesseret i hele sumtabellen. Han har måske blot brug for at vide om hans resultat hører til *den bedste halvdel*, eller til *den dårligste halvdel*. Eller måske om det hører til *den bedste fjerdedel* eller til *den dårligste fjerdedel*. Der kan derfor være brug for at dele datasættet op i halvdele eller måske i fjerdedele, så godt det nu lader sig gøre. Lad os se på Sannes datasæt.

Vi gør igen brug af sumtabellen. Af tabellen fremgår at den heldigste halvdel strækker sig til et kasteantal på 4. Først ved et kasteantal på 4 når den summerede frekvens jo op på 50%. Ved kasteantallet 3 er vi endnu kun nået op på 44%, men ved kasteantallet 4 når vi op på 54%. De 50% når vi altså først ved kasteantallet 4. Man siger at 4 er *median* i Sannes datasæt (Median betyder midtpunkt).

Medianen angiver hvor langt den nederste halvdel af resultaterne rækker. Hvis vi stiller alle Sannes 50 data op i størrelsesorden med de mindste forrest, vil medianen være den observation der markerer at nu er vi halvvejs i rækken, altså den observation der står som nr. 25 i rækkefølgen. Men vi behøver ikke at lave denne rangliste over Sannes observationer. Af sumtabellen kan vi jo let se hvor medianen i datasættet er.

At medianen i Sannes datasæt er 4, fortæller os at et kasteresultat på under 4 hører til den bedste halvdel af sumtabellen, og at et kasteresultat på over 4 hører til den dårligste halvdel.

Lad os nu se på opdelingen i *fjerdedele*. Af sumtabellen ser vi at vi når op til en summeret frekvens på 25% ved kasteantallet 2. Ved kasteantallet 1 er vi endnu kun nået op på 18%, men ved kasteantallet 2 når vi op på 30%. De 25% når vi altså ved kasteantallet 2. Dette antal kalder man *nedre kvartil* (eller *første kvartil*) i datasættet. Nedre kvartil angiver hvortil den nederste fjerdedel af observationerne rækker.

Ved *øvre kvartil* (eller *tredje kvartil*) forstår vi den observation hvortil de nederste 75% af observationerne rækker. Af sumtabellen ser vi at i Sannes observationssæt er 7 øvre kvartil. Vi har dermed følgende oplysninger om Sannes observationssæt:

Nedre kvartil: 2                  Median: 4                  Øvre kvartil: 7

Medianen kaldes også *midterste kvartil* (eller anden kvartil). De tre værdier for nedre kvartil, median og øvre kvartil kaldes datasættets **kvartilsæt**. Kvartilsættet for Sannes datasæt er altså: 2-4-7.

### Hvad kvartilsættet fortæller

Af kvartilsættet kan aflæses en række oplysninger. Kvartilsættet for Sannes datasæt fortæller således at observationer under 2 hører til den nederste fjerdedel af observationerne, og at observationer under 4 hører til den nederste halvdel af observationerne. Vi kan af kvartilsættet også se at observationer over 4 hører til den øverste halvdel af observationerne, og at observationer over 7 hører til den øverste fjerdedel.

Kvartilsættet er derfor et fint hjælpemiddel når man hurtigt skal vurdere resultater og chancer i nye eksperimenter. Nogle af de vigtigste oplysninger fra sumtabellen er indeholdt i kvartilsættet. Man kan faktisk sige at kvartilsættet er sumtabellen "kogt ned til" tre tal.

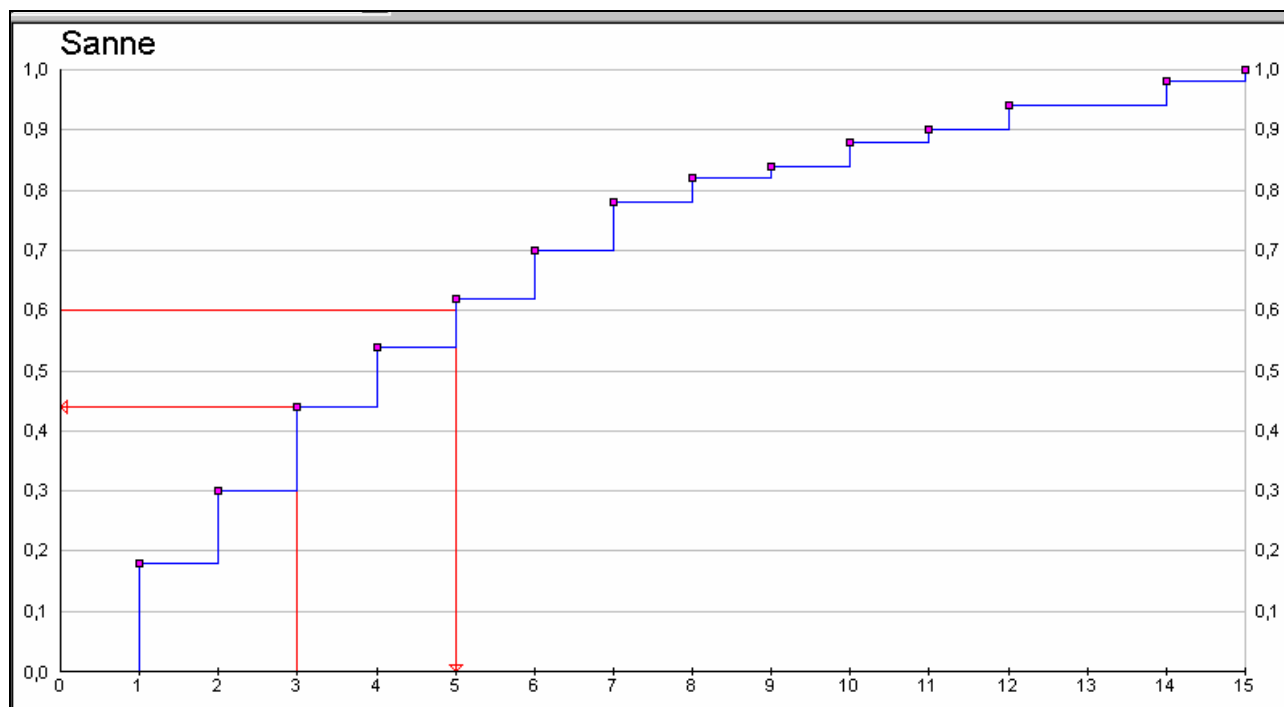
### **Prøv selv!**

1. Hvad er medianen for dine data?
2. Hvad er nedre kvartil i dit datasæt?
3. Og hvad er øvre kvartil?
4. Sammenlign dit kvartilsæt med Sannes. Er der forskelle?

*Nye begreber.* Tilføj de nye begreber til din liste.

## 5. Et trappediagram

Sumtabellen kan også beskrives ved et grafisk billede, et *trappediagram*. Diagrammet tegnes let ud fra de tal der er anført i sumtabellen under "Summeret frekvens". Vi opstiller her et trappediagram for Sannes data.



Trappediagrammet fortæller i et grafisk billede hvordan den summerede frekvens vokser fra 0% op til 100%. Den vokser ikke jævnt, men i spring af forskellig størrelse.

Det første spring forekommer ved et kasteantal på 1, her er springet i den summerede frekvens på 18%. Det sidste spring forekommer ved kasteantallet 15, her er springet i den summerede frekvens kun 2%.

Ved hver af de observationer der forekommer i Sannes datasæt, tager den summerede frekvens et spring opefter. Og springets størrelse viser hvor stor en frekvens den pågældende observation har i datasættet.

### At aflæse på diagrammet

På trappediagrammet kan man aflæse "nedefra og op": På figuren er vist hvordan der aflæses for et kasteantal på 3. Man kan også aflæse "oppefra og ned": For en summeret frekvens på 60% aflæses et kasteantal på 5. På denne måde kan medianen altså let aflæses på trappediagrammet.

Når man ved aflæsning nedefra rammer et lodret liniestykke på diagrammet, skal man gå op til det vandrette trin før man aflæser på skalaen til venstre. (Se aflæsningen for et kasteantal på 3). - Ved aflæsning oppefra skal der aflæses der hvor man rammer trappen. (Prøv fx at aflæse for en summeret frekvens på 70%).

## Fraktiler

Som du har set, kan man på trappediagrammet let aflæse datasættets median. Man begynder blot ved 50%-mærket på den lodrette akse, går ind til trappediagrammet, og aflæser herefter den søgte værdi på den vandrette akse. Tilsvarende begynder man ved 25%-mærket og 75%-mærket når den nedre kvartil og den øvre kvartil skal bestemmes.

Man kan selvfølgelig også foretage aflæsninger af denne art hvor man begynder ved andre mærker på den lodrette akse. Begynder man fx ved 35%-mærket på trappediagrammet over Sannes datasæt, aflæses på den vandrette akse et kasteantal på 3.

Denne værdi kalder man *35%-fraktilen* i datasættet (*fraktil* betyder andel eller brøkdel). De tidligere fundne kvartiler er blot nogle specielle fraktiler. Nedre kvartil er det samme som *25%-fraktilen*, medianen er det samme som *50%-fraktilen*, og øvre kvartil er det samme som *75%-fraktilen*.

## Prøv selv!

1. Tegn på millimeterpapir et trappediagram over dine data.
  1. Vær omhyggelig med tegningen.
2. Aflæs på trappediagrammet median, nedre kvartil og øvre kvartil.
3. Aflæs følgende fraktiler: 10%-fraktilen og 90%-fraktilen.
4. Hvordan kan man på trappediagrammet aflæse typetallet?
5. Hvordan kan man aflæse variationsbredden?

*Nye begreber.* Tilføj de nye begreber til din liste.

## 6. Gennemsnit

I det spil Sanne deltager i, gælder det at der skal betales 1 krone for hvert kast der bruges til at opnå de to ens øjental på terningerne.

Sanne er nu interesseret i at vide hvor dyrt det vil være at opnå to ens. Her kan hun igen få hjælp fra den undersøgelse hun har foretaget.

Af sumtabellen har vi allerede set at medianen i Sannes observationssæt er 4. Der er altså ca. 50% chance for at der ikke skal bruges mere end 4 kast for at slå to ens. Med andre ord: *Der er ca. 50% chance for at to ens øjental højst koster 4 kr.*

Men vi ved også at det godt kan blive noget dyrere. Sannes eksperimenter viser jo at det kan komme til at koste helt op til 15 kast, dvs. 15 kr. Det kan dog berolige Sanne lidt at øvre kvartil i hendes datasæt er 7. Det fortæller jo at der højst er en risiko på 25% for at to ens kommer til koste mere end 7 kr.

### Hvor dyrt bliver det?

I et spil kan der mange gange undervejs blive brug for at et kast med to ens øjental. Sanne er derfor interesseret i at få oplysninger om hvor meget man kan forvente at komme til at betale. Og det er ikke nok for hende at vide at der i hver situation er 50% chance for at hun kan slippe med at betale højst 4 kr., og at der kun er en risiko på 25% for at han skal betale mere end 7 kr. I stedet for vil hun være interesseret i at vide hvad hun *i gennemsnit* må regne med at skulle betale for at få to ens øjental.

Også her kan Sannes statistik give et svar. Sanne har jo udført en undersøgelse der omfatter 50 eksperimenter, og hun kan beregne hvor mange kast der i gennemsnit blev brugt i hvert eksperiment. For at kunne beregne det, må Sanne først tælle op hvor mange kast der i alt blev anvendt i de 50 eksperimenter. Hun kunne her gå tilbage til listen over de enkelte observationer og lægge alle 50 tal sammen:

$$2 + 1 + 12 + 7 + \dots + 6 + 8 + 4.$$

Men hun kan også bruge den grundtabel hun har opstillet:

Antal kast	1	2	3	4	5	6	7	8	9	10	11	12	14	15
Hyppighed	9	6	7	5	4	4	4	2	1	2	1	2	2	1

Den viser jo at der blandt de 50 tal der skal lægges sammen findes 9 1-taller, 6 2-taller, 7 3-taller osv. Summen af de 50 tal kan derfor let beregnes ud fra grundtabellen:

$$9 \cdot 1 + 6 \cdot 2 + 7 \cdot 3 + \dots + 2 \cdot 14 + 1 \cdot 15 = 257$$

I de 50 eksperimenter er der altså brugt i alt 257 kast. I gennemsnit er der derfor brugt

$$257:50 = 5.14$$

kast pr. eksperiment. Vi har hermed beregnet *gennemsnittet*, eller *middeltallet*, for Sannes datasæt. Middeltallet er altså 5.14.

Sanne har nu en anelse om hvad det vil koste at få to ens øjental. Hun må i gennemsnit regne med at skulle bruge over 5 kast for at opnå to ens. Den gennemsnitlige udgift til at slå to ens bliver derfor mere end 5 kr.

Sanne er godt klar over at hendes 50 eksperimenter kun kan give et fingerpeg om hvad det gennemsnitlige antal af kast i fremtidige spil vil være. Havde hun udført andre 50 eksperimenter, ville hun nok have fået et andet middeltal end 5.14. Men resultatet kan give hende en fornemmelse af den gennemsnitlige omkostning ved at slå to ens øjental.

### **Prøv selv!**

1. Beregn middeltallet for dine data.
2. Hvor dyrt bliver det i gennemsnit at få to ens øjental?

*Nye begreber.* Tilføj de nye begreber til din liste.

## 7. En statistik

Du har nu gennemført en statistisk undersøgelse og opstillet *en statistik*: Du har udført en række eksperimenter og dermed skaffet dig et datasæt. Herefter har du opstillet *tabeller*, tegnet *diagrammer*, og beregnet *mærketal*. De statistiske mærketal er talværdier som fortæller noget om det foreliggende datasæt.

Af mærketal har du set på median, nedre og øvre kvartil, fraktiler og middeltal. Og undervejs gjorde du også brug af typetal, variationsbredde og største og mindste observation. Alle disse tal er statistiske mærketal som kan give dig oplysninger om datasættet.

Ved hjælp af tabellerne, diagrammerne og mærketallene har du skaffet dig oplysninger om det problem du har villet undersøge, og du har kunnet fremsætte nogle formodninger om de ukendte chancer og forventninger i spillet. Du ved godt at dine undersøgelser ikke giver helt præcise svar, så derfor bruger du dine resultater med forsigtighed. Men du har god tiltro til at de på væsentlige punkter fortæller noget rigtigt.

En statistisk undersøgelse har ikke ét rigtigt svar som man skal nå frem til. Gennem anvendelsen af statistik kan vi give et svar på de stillede spørgsmål, men svaret er ikke det endelige. *Nye undersøgelser kan revidere vort svar.*

Dette forhold er typisk for omgangen med data som opstår gennem eksperimenter og undersøgelser. *De kan give os en ide om hvad de ukendte svar kan være, men de vil aldrig give os 100% sikkerhed.*

### En oversigt

I den statistiske undersøgelse gjorde du brug af en række hjælpemidler i form af tabeller og diagrammer:

*Grundtabel, Sumtabel, Pindediagram, Trappediagram*

Endvidere benyttede du følgende mærketal:

*Typetal, Median, Kvartiler, Variationsbredde, Middeltal*

Tabeller, diagrammer og mærketal kan være dig til hjælp når du skal i gang med at udføre en statistisk undersøgelse. Du behøver dog ikke altid at være lige grundig. I nogle tilfælde kan du sikkert nøjes med at opstille en tabel over de fundne hyppigheder og at beregne nogle af de mærketal der hører til denne tabel. I andre tilfælde vil du måske skulle gå lidt videre og se på frekvensfordelingen for dit datasæt. Du må da opstille en grundtabel eller tegne et pindediagram.

Den bedste oversigt får du gennem *sumtabellen* og *trappediagrammet*. Så i mange tilfælde vil du nok skulle tilrettelægge din undersøgelse sådan at du når frem til at opstille en sumtabel. Denne tabel indeholder jo også alt det du kan aflæse i grundtabellen.

### Prøv selv med et større datasæt

I denne opgave skal du prøve på at få et overblik over de nye begreber der er indført undervejs i arbejdet med taldata.

1. Slå dine data og Sannes data sammen til ét datasæt. Det indeholder altså 100 data.
2. Opstil en sumtabel for det store datasæt.
3. Hvad er her chancen for at man kan få to ens i højst to kast?
4. Hvad er risikoen for at man må bruge mere end 10 kast?
5. Find middeltallet for det store datasæt.
6. Angiv median og nedre og øvre kvartil for datasættet.



## 8. Anvendelse af edb

Ved anvendelse af edb kan du slippe lettere om ved arbejdet med at udføre beregninger og at opstille tabeller og fremstille diagrammer. Så snart du er fortrolig med de statistiske begreber, kan du derfor lade et edb-program overtage det besværlige arbejde med den statistiske behandling af dine data.

Det er dog dig der skal bestemme hvilke data edb-programmet skal anvendes på, og det er dig der bestemmer hvilke statistiske hjælpemidler du vil gøre brug af. Og ikke mindst: *Det bliver stadig din opgave at fortolke undersøgelsens resultater og at beslutte hvordan de anvendes ved løsning af de stillede problemer.* Her kan edb-programmet ikke hjælpe dig.

### Edb-programmet Visistat

Dette program kan du anvende i arbejdet med taldata. Det kan opstille tabeller, beregne mærketal og tegne diagrammer.

Du kan nu bruge programmet på de data du har indsamlet i din statistiske undersøgelse.

**Data.** Gå ind i Visistat og gå til menupunktet *Datatype*. Her vælger du *Observationer* og indtaster dine 50 data fra kasteeksperimenterne. Du indtaster i datasøjlen A.

**Gem.** Når de 50 data er indtastet, kan det være en god ide at gå til *Gem* under menupunktet *Filer*. Her skal du give dit datasæt et navn, det kan fx være TALDATA1. Klik derefter på *Gem*, og dit datasæt vil nu være gemt i en datafil. Det kan herefter tages frem igen ved nye anvendelser uden at du igen skal indtaste de 50 tal.

**Beskriv.** Klik på feltet A så dine data bliver markeret. Og klik på knappen "Redigering og beskrivelse". I det højre vindue vil du nu få adgang til en rullemenu. Her kan du vælge mellem følgende beskrivelser:

- Tabel (sumtabel)
- Mærketal
- Datapinde (en grafisk oversigt over dine data)
- Hyppighedsgraf (pindediagram)
- Frekvensgraf (pindediagram)
- Sumfrekvensgraf (trappediagram)
- Cirkeldiagram

Afprøv alle disse muligheder med datasættet TALDATA1, og sammenlign tabeller, diagrammer og mærketal med dem du har fremstillet ved håndkraft. Lad programmet udprinte tabellen og nogle af diagrammerne.

**Stort billede.** Lad programmet vise trappediagrammet og klik på knappen "Beskrivelsesvindue". Du får nu et billede som benytter hele skærmen.

**Beregning af fraktiler.** Ved siden af overskriften "Sumfrekvensgraf" findes en knap med to pile. Den giver dig adgang til at beregne samhørende værdier for x og y på trappediagrammet.

**Varians og standardafvigelse.** Bemærk at der under mærketal beregnes to værdier som vi ikke har omtalt: *Varians* og *Standardafvigelse*. Disse mærketal er mål for hvor stor en spredning der er i datasættet. De anvendes i videregående statistiske undersøgelser, og vi skal ikke tage dem op her.

Du kan læse mere om anvendelsen af Visistat i den brugervejledning der hører til programmet.

## 9. Opgaver

Du skal nu prøve selv at arbejde med indsamling af data og opstilling af en statistik. Her er nogle opgaver som du kan vælge imellem. I arbejdet kan du bruge Visistat når du skal opstille tabeller, tegne diagrammer og beregne mærketal. Og du kan bruge KAST hvis der skal udføres terningkast.

### Opgave 1. Hvor svært er det at slå seksere?

I Ludo får man en brik i spil hvis man kan slå en sekser med en terning.

Lav en statistisk undersøgelse som viser hvor svært det er at slå en sekser. Du skal udføre mindst 50 kasteserier.

Opstil tabeller, tegn diagrammer og beregn mærketal.

Svar dernæst på følgende:

1. Hvad er chancen for at du kan klare dig med 2 kast?
2. Hvad er chancen for at du kan klare dig med 4 kast?
3. Hvad er risikoen for at du skal bruge over 10 kast?
4. Hvor mange kast skal du i gennemsnit bruge for at få en sekser?

### Opgave 2. Nye Ludoregler

For at få lidt mere fart over Ludospillet foreslår Sanne at reglerne ændres så man ikke kun får en brik i spil hvis man slår en sekser, men også hvis man slår en femmer. Lav en statistisk undersøgelse som viser hvad denne ændring betyder for vanskeligheden med at få en brik i spil.

Giv dernæst dine svar på følgende spørgsmål:

1. Hvad er chancen for at man kan klare sig med to kast?
2. Hvad er chancen for at man kan klare sig med fire kast?
3. Hvad er risikoen for at man skal bruge mere end fem kast?
4. Hvilke kasteantal hører til "den heldigste halvdel"?
5. Hvilke kasteantal hører til "den uheldigste fjerdedel"?
6. Hvor mange kast skal der i gennemsnit til for at få en brik i spil?

### Opgave 3. To ens på en anden måde

Lav en statistisk undersøgelse af følgende spil: Du kaster med én terning. Terningen kastes indtil der opnås et øjental som tidligere er forekommet i kasterækken. Kasterækken omfatter altså mindst to kast og højst syv kast.

Giv på grundlag af din undersøgelse et svar på følgende spørgsmål vedrørende spillet:

1. Hvad er chancen for at man klarer sig med to kast?
2. Hvad er chancen for at man kan klare sig med tre kast?
3. Hvad er risikoen for at man skal bruge mere end fire kast?
4. Hvilket kasteantal vil man hyppigst skulle bruge i spillet?
5. Hvor mange kast skal der i gennemsnit bruges i spillet?
- 6.

#### **Opgave 4. To ens i træk**

Spillet fra opgave 3 ændres til at kasterækken fortsættes indtil to på hinanden følgende kast har givet samme øjental.

Lav en statistisk undersøgelse af dette spil, og giv svar på følgende spørgsmål:

1. Hvad er chancen for at man klarer sig med tre kast?
2. Hvad er risikoen for at man skal bruge mere end 10 kast?
3. Hvor mange kast skal der i gennemsnit bruges i spillet?

Opstil yderligere nogle spørgsmål, og besvar dem ved hjælp af resultaterne fra din statistiske undersøgelse.

#### **Opgave 5. Kvik-Ludo**

I Kvik-Ludo kastes med to terninger, og man får en brik i spil hvis en af terningerne (eller begge) viser 6.

Besvar følgende spørgsmål på grundlag af data fra en statistisk undersøgelse:

1. Hvad er chancen for at der højst skal udføres 3 kast for at få en brik i spil.
2. Hvad er chancen for at man kan klare sig med højst 5 kast?
3. Hvad er risikoen for at der skal udføres mere end 8 kast?
4. Hvor mange kast skal der i gennemsnit udføres?

#### **Opgave 6. Alle seks**

Sanne har fundet på følgende spil: En terning kastes indtil alle seks mulige øjental er forekommet i kasterækken.

(Kasterækken omfatter altså mindst seks kast.)

Sanne tilbyder dig at deltage i spillet på følgende vilkår:

Hvis du kan slå "alle seks" med højst 12 kast får du udbetalt 10 kr. Hvis du bruger mere end 12 kast skal du betale 10 kr. til Sanne.

Lav en statistisk undersøgelse af spillet, og afgør om du vil modtage Sannes tilbud.

### **Opgave 7. Alle seks. Et nyt tilbud**

Sanne tilbyder dig nu at deltage i spillet "Alle seks" fra forrige opgave på følgende vilkår:

Du betaler 25 øre pr. kast. Til gengæld får du udbetalt 3 kr. når du har slået "alle seks".

Afgør på grundlag af din undersøgelse fra opgave 8 om du vil modtage dette tilbud.

### **Opgave 8. Hvem rammer plet?**

Sanne foreslår Malene en konkurrence: Vi skal gætte på hvor mange kast der skal bruges for at slå en sekser. Du gætter på et antal, og jeg gætter på et andet antal. Derefter kastes terningen indtil vi opnår en sekser.

Den af os der har gættet det rigtige antal af kast får 10 kr. fra den anden. Hvis ingen har gættet det rigtige antal udføres en ny kasteserie.

Hvilket antal bør Malene gætte på?

### **Opgave 9. Hvem rammer nærmest?**

Malene foreslår Sanne at de ændrer reglerne fra opgave 8:

Sanne gætter først på et antal, derefter gætter Malene på et andet antal.

Terningen kastes indtil der er opnået en sekser. Den der har gættet nærmest ved det rigtige antal kast får 10 kr. af den anden.

Hvis der er dødt løb udføres en ny kasteserie.

Efter lidt betænkningstid modtager Sanne udfordringen, idet hun mener at have fundet det gæt som vil give hende størst chance for at vinde.

Hvilket antal gætter Sanne på?

### **Opgave 10. Observationssæt slås sammen**

Sanne og Malene har hver for sig udført samme statistiske undersøgelse.

De har begge benyttet et observationssæt som indeholder 25 observationer.

På grundlag af deres undersøgelse har de beregnet følgende statistiske mærketal:

	Sanne	Malene
Typetal	1	2
Median	3	5
Middeltal	5.60	6.20

De slår nu deres observationer sammen til ét observationssæt bestående af 50 observationer. Hvad kan du mon ud fra de oplysninger der er givet i tabellen sige om typetal og median for deres samlede observationssæt? Svar på følgende:

Kan typetallet være 1? 2? 3?

Kan medianen være 3? 5? 4? 6?

Og hvad er middeltallet for det samlede observationssæt?

Hvad ville du kunne sige om de tre mærketal for det samlede observationssæt hvis de to observationssæt havde forskellig størrelse: Sannes 25 observationer, Malenes 50 observationer?

### Opgave 11. Nogle chancesituationer

Hvis du har lyst til at arbejde med flere opgaver, er her et udvalg. I arbejdet med opgaverne kan du bruge KAST.

Undersøg en eller flere af følgende chancesituationer ved hjælp af en statistisk undersøgelse. Angiv først nogle spørgsmål som du ønsker at besvare gennem din undersøgelse, og opstil derefter en statistik der kan hjælpe dig til at besvare dine spørgsmål. Ved udarbejdelsen af statistikken kan du bruge Visistat.

1. Du kaster med tre terninger indtil du opnår et kast med mindst to seksere.
2. Du kaster med fire terninger indtil du opnår et kast med mindst to seksere.
3. Du kaster med tre terninger. I hvert kast ganger du de tre opnåede øjental sammen. Du kaster indtil du opnår et kast hvor resultatet af gangestykket er 100 eller derover.
4. Du kaster fem terninger indtil du opnår et kast med mindst tre terninger der viser samme øjental.
5. Du kaster fem terninger indtil du opnår et kast med "to par", dvs. to terninger der viser ét øjental, to terninger der viser et andet øjental, og en terning som viser et tredje øjental. (Eksempel: 23532).
6. Du kaster med tre terninger indtil du opnår et kast hvor øjentallet 6 forekommer, men ikke øjentallet 1.
7. Du kaster med to terninger indtil du opnår et kast hvor enten øjentallet 5 forekommer, eller hvor summen af øjentalene er 5.

8. Du kaster med tre terninger indtil du opnår et kast med tre på hinanden følgende øjental: 123, 234,...,456. (Rækkefølgen på terningerne betyder ikke noget, 312 er det samme som 123).