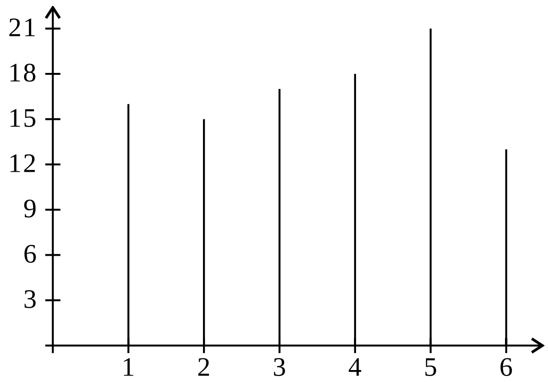
2.2 Stolpediagram/søylediagram

En frekvenstabell som vi har i tabell 2.1, kan anskueliggjøres i et *stolpediagram*. Utfallsrommet til variabelen (antall øyne terningen viser) avsettes langs førsteaksen. Frekvensen (antall observasjoner) avsettes langs andreaksen. Med dataene fra tabell 2.1 får vi følgende stolpediagram.
Som vi ser, framkommer stolpediagrammet ved at vi reiser en stolpe for hver verdi variabelen tar. Høyden på stolpen er lik frekvensen for utfallet. Merk at en stolpe minner om et linjestykke i det den ikke har noen bredde.



Figur 2.1: Stolpediagram

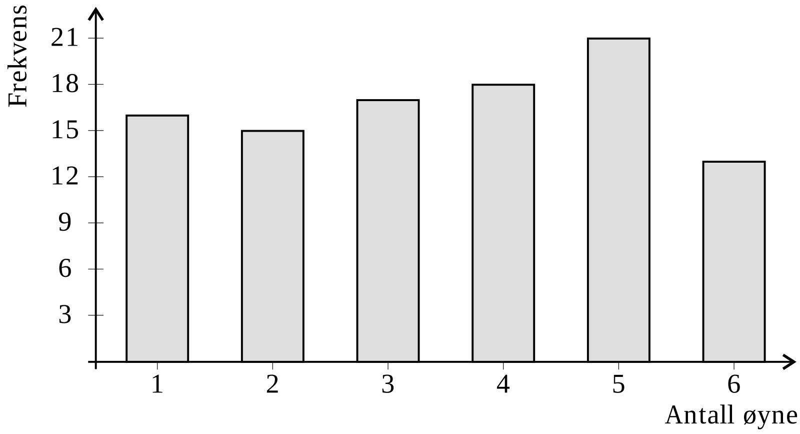
Fordelen med et slikt stolpediagram er at en får en visuelt bedre oversikt over hvordan observasjonene i et forsøk fordeler seg. Og da ikke bare hvilke verdier frekvensene har, men også hvor høye de er i forhold til hverandre. Vi sier gjerne at vi har framstilt tabell 2.1 *grafisk*. Eller: Stolpediagrammet er en grafisk framstilling av tabell 2.1.

Søylediagrammet er nær beslektet med stolpediagrammet, så nær at de ofte blir brukt og omtalt om hverandre. Grunnen er at det bare er en nyanseforskjell mellom dem. Forskjellen er kun en definisjonssak. Vi kan godt si at et stolpediagram er et søylediagram hvor søylene ikke har noen bredde. Eller like gjerne omvendt: Et søylediagram er et stolpediagram hvor vi tillater stolpene å ha bredde. Figur 2.2 viser et søylediagram som er reist på grunnlag av dataene i tabell 2.1. Dette kan for eksempel utføres i et regneark.

Mens stolpediagrammet kan virke noe spinkelt, vil søylediagrammet gi et mer robust inntrykk, og følgelig ofte være det mest hensiktsmessige diagrammet å benytte av de to. Spesielt kan presentasjonen av et datamateriale virke mer iøynefallende dersom vi har anledning til å fargelegge søylene. Enkelte hevder at en benytter stolpediagram når verdiene på førsteaksen er tall, mens en ellers benytter søylediagram. Men noen absolutt regel er ikke dette.

matemania

Denne teksten er hentet fra *Matematiske sammenhenger: Statistikk og sannsynlighetsregning*
Skrevet av Knut Ole Lysø
Utgitt av Caspar forlag as – www.caspar.no



Figur 2.2 Søylediagram

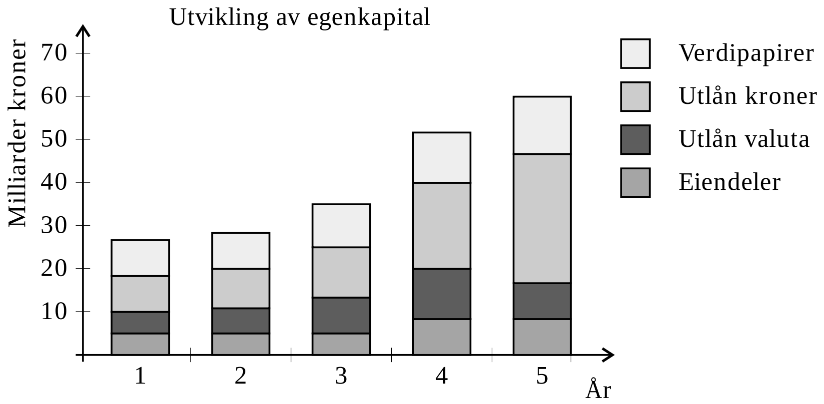
Prinsippet for konstruksjonen av et søylediagram, er at høyden på søylene er lik frekvensen for hendelsen. Søylenes bredde er en smaksak, men skal framstillingen ha noen verdi, må bredden i hver søyle være den samme.

Både stolpediagrammet og søylediagrammet avslører at det av og til kan være vanskelig å avlese helt nøyaktige verdier i et diagram. Hvis en har behov for disse, kan en finne fram det opprinnelige datamaterialet som opptegningen av diagrammet er basert på. Det viktigste med slike grafiske framstillinger er å få fram hvordan frekvensene varierer, og om det er noen verdier som skiller seg ut. Særlig nyttig kan dette være om en ønsker å sammenligne ulike grupper med hverandre eller når en ønsker å beskrive en utvikling over tid (jfr. oppgave 6 og 7).

En stor fordel med søylediagrammet er at en lett kan inndele hver søyle i ulike underområder. Vi får da et såkalt *stabeldiagram*, se eksempel 2.2.

Eksempel 2.2
I en norsk banks årsberetning får en presentert utviklingen av bankens egenkapital de siste 5 årene gjennom en grafisk framstilling (figur 2.3).

Dette stabeldiagrammet inneholder mye informasjon. For det første vil høyden av søylene angi årets totale verdi av eiendelene. Videre ser vi at hver søyle er inndelt i fire underområder som til sammen utgjør bankens totale egenkapital. Størrelsen av underområdene varierer også fra år til år. Av diagrammet går det klart fram at bankens totale eiendeler har steget jevnt disse årene. Faktisk representerer verdien i år 5 mer enn en fordobling i forhold til situasjonen i år 1. Det vesentligste av denne økningen ser vi forklares av økte utlån i kroner.



Figur 2.3: Stabeldiagram

La oss dvele litt ved figur 2.3. Hva kan grunnen til disse svingningene (variasjonene) være? Kan en nedgang i et underområde skyldes svingninger i markedet eller uheldige investeringer? Kan utlån i kroner skyldes lavere rente eller mindre konkurranse fra andre aktører på markedet? Spørsmålene kan være mange, men disse kan vi selvsagt ikke besvare på bakgrunn av dette diagrammet alene. Det må videre undersøkelser til for å avdekke mulige årsaker til variasjonene. Kanskje vil en finne at variasjonene skyldes den ene av disse faktorene eller en kombinasjon av dem. Det er fullt mulig at svingningene ikke kan forklares av noen av disse faktorene, men av helt andre faktorer. En siste mulighet er at en ikke kan finne noen bestemte årsaker til variasjonene. I statistikken sier en da gjerne at svingningene trolig skyldes *tilfeldig variasjon*. I så fall er variasjonene i dette eksemplet i prinsippet de samme som vi så i forbindelse med terningkastene i eksempel 1.1. De har ingen bestemt årsak, men er en konsekvens av de tilfeldigheter som kjennetegner de typer forsøk som behandles i statistikken.

Å avgjøre om en har forhold som kan forklares ved tilfeldig variasjon eller ikke, er et viktig område i statistikken; et område som omfattes av metodelæren (analytisk statistikk, jfr. kapittel 1.1), og som vi kommer inn på i kapittel 12.

matemania

Denne teksten er hentet fra *Matematiske sammenhenger: Statistikk og sannsynlighetsregning*
Skrevet av Knut Ole Lysø
Utgitt av Caspar forlag as – www.caspar.no