Universitetet i Stavanger

DET TEKNISK-NATURVITENSKAPELIGE FAKULTET

BACHELOROPPGAVE

Studieprogram/spesialisering: Lektorutdanning for 8-13. trinn. Realfag, matematikk.	Vårsemesteret, 2021 Åpen
Forfatter: Ine Sæbø Kaltveit	Ine Sæbø Kaltveit …………………………………… (signatur forfatter)
Fagansvarlig: Anders Tranberg Veileder: Jan Terje Kvaløy	
Tittel på bacheloroppgaven: Hvordan avsløre løgn med statistikk? Engelsk tittel: How to expose lies with statistics?	
Studiepoeng: 10	
Emneord: Bacheloroppgave i matematikk – lektor LMABAC-1	Sidetall: 31 inkludert forside og referanser Stavanger, 15.05.2021 dato/år

Forord

Denne oppgaven er et resultat av mitt arbeid med bacheloroppgaven i matematikk - lektor ved Universitetet i Stavanger våren 2021.

Arbeidet har vært lærerikt og interessant.

Jeg vil rette en stor takk til min veileder, professor Jan Terje Kvaløy. Takk for særdeles god veiledning og oppfølgning. Jeg setter pris på dine kloke ord, din hjelp og din tilgjengelighet. Du er inspirerende, både som professor og veileder.

Stavanger 12. mai 2021

Ine Sæbø Kaltveit

Sammendrag

Hensikten med denne oppgaven er å beskrive vanlige feilkilder/fallgruver i sannsynlighetsregning og statistikk. De ulike feilkildene/fallgruvene er beskrevet både skriftlig og matematisk. Videre kommer forklaringer på hvordan man kan unngå og/eller avsløre disse. Oppgaven inneholder utvalgte eksempler som skal gi leseren et godt bilde på hvordan feilkildene/fallgruvene kommer til uttrykk. Noen av eksemplene har jeg utarbeidet selv, mens andre er hentet fra ulike kilder. I det siste kapitlet får man et innblikk i hvordan denne problematikken kan brukes i undervisningsituasjoner. Oppgaven rundes av med avsluttende ord.

Innhold

1 Introduksjon

Vi lever i en stor og komplisert verden, og for å forstå den trenger vi tall [1]. Hva er status på den pandemien vi nå er inne i? Hva mener norske studenter om sitt studieprogram? Er det en sammenheng mellom ensomhet og lønn? Disse spørsmålene krever at vi beveger oss ut fra våre personlige opplevelser og inntrykk. Vi trenger kvantitative data, altså statistikk, til å veilede oss. Men vi kan ikke stole blindt på all statistikk vi møter. Følgelig er noen av tallene vi møter nøyaktige, mens andre kan være ren gjetting [1]. Vi må ha kunnskap for å vite forskjellen mellom god og dårlig statistikk. Vi bør møte statistikk med kritiske øyne, og være klar over at det vi ser ikke nødvendigvis er troverdig. Men hvordan vet vi hva vi skal se etter? Nettopp dette skal vi gå igjennom i denne oppgaven; Hvordan avsløre løgn med statistikk?

Denne oppgaven inneholder informasjon om hvordan man kan avsløre løgn med statistikk. Oppgaven tar utgangspunkt i syv vanlige fallgruver man kan møte i statistikken. Fallgruvene presentert i oppgaven omhandler misvisende grafer, absolutte kontra relative tall, ikke-representative utvalg, Simpsons paradoks, p-verdi, feil kobling mellom årsak og virkning og til slutt feiltolkning av sjeldne hendelser.

For å øke bevisstheten blant barn og unge, bør undervisningen i skolen rette lys mot denne problematikken. Når elever lærer om statistikk bør de få et innblikk i hvilke fallgruver de kan møte på, både for å unngå å gjøre samme feil selv, men også for å avsløre hvilke feil som har blitt begått. Denne oppgaven vil avslutningsvis diskutere hvordan fallgruver kan trekkes inn i undervisningssituasjoner.

2 Grafiske fremstillinger

Grafisk fremstilling er en måte å fremstille tallmateriale og informasjon på. Ved å ta i bruk grafiske fremstillinger blir informasjonen presentert på en visuell måte. Dersom fremstillingen er visuell, kan det være enklere å forstå og lese ulike oppsummeringer av tallmaterialer eller statistiske resultater. Det finnes en rekke type grafer som kan brukes for å presentere informasjon. Hvis man skal fordele individer eller enkeltobservasjoner i grupper etter visse egenskaper brukes ofte et histogram eller søylediagram. Skal man illustrere sammenhengen mellom to variable størrelser kan man benytte et spredningsplott [2].

Datarevolusjonen førte til at det ble enklere for blant annet journalister å lage grafer og iøyefallende fremstillinger av data. Noen ganger er resultatene av disse fremstillingene informative - andre ganger ikke. En stilig graf er ikke nødvendigvis en god graf. Det finnes flere feil man kan begå når man arbeider med grafiske fremstillinger. Og selv de mest kjente feilene blir begått av personer som burde visst bedre [1].

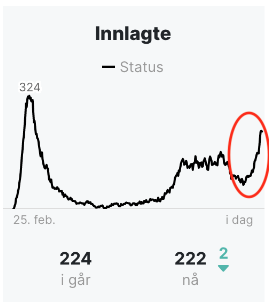
2.1 Utelate origo og akser

En av de mest kjente måtene data blir manipulert på i grafer, er å unngå å ha med origo, aksene eller begge deler. Når man arbeider med grafer kan det å utelate origo og akser være et bevisst valg, men også et ubevisst valg [3].

Et aktuelt tema i dagens nyhetsbilde er koronapandemien. Flere ulike norske nettaviser fremviser informasjon og tall knyttet til koronaviruset. En norsk nettavis har øverst på sin forside flere grafiske fremstillinger som viser dagens antall smittede, vaksinedoser, innlagte, positive tester og døde.
Denne nettavisen hadde på sin forside 17.03.21 denne grafen som presenterte utviklingen i antall innlagte de siste dagene.

Figur 1: Antall innlagte per 17.03.21. Hentet fra VG.

Grafen i figur 1 er det første man så på forsiden denne dagen. Leserne fikk mest sannsynlig et inntrykk av at antall innlagte de siste dagene har økt fra 0 til 200 personer på bakgrunn av at grafen mangler både origo og akser. Trykker man på selve grafen inne på nettsiden får man opp en ny graf vist i figur 2 under.



Figur 2: Antall innlagte fra 25.02.21 til 17.03.21. Hentet fra VG.

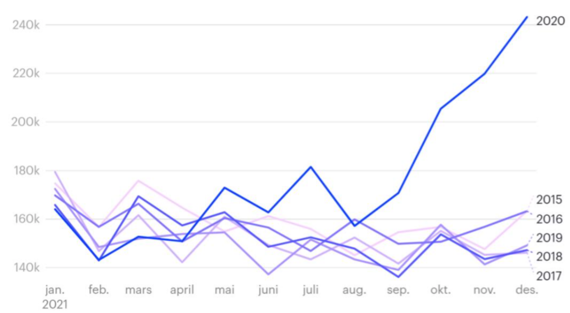
Grafen vist i figur 2 viser antall innlagte fra 25.02.20 til 17.03.21. Her fikk man som leser sett tallene for antall innlagte i et annet perspektiv ettersom grafen inneholdt en linje ved y=0. Grafen som blir fremvist i figur 1 er et utsnitt av grafen fremvist i figur 2. Utsnittet er markert med en rød sirkel på grafen i figur 2. Bunnpunktet inni den røde sirkelen viser at per 17.02.21 var det 69 innlagte personer. Her ser man tydelig at økningen ikke går fra 0 til 200 personer, men fra 69 til 200 personer. Dermed kan grafen i figur 1 være misvisende for mange dersom de ikke trykker seg videre og får se det hele bildet.
Dette eksempelet viser viktigheten med at y-aksen starter ved y=0 og hvor misvisende en graf kan være uten.

Grafen i figur 1 belyser også et annet problem som kan oppstå når man arbeider med grafiske fremstillinger - nemlig det å unngå å presentere vesentlig data. Grafen i figur 1 viser som nevnt kun et utsnitt av en hel graf.

2.2 Manipulere y-aksen

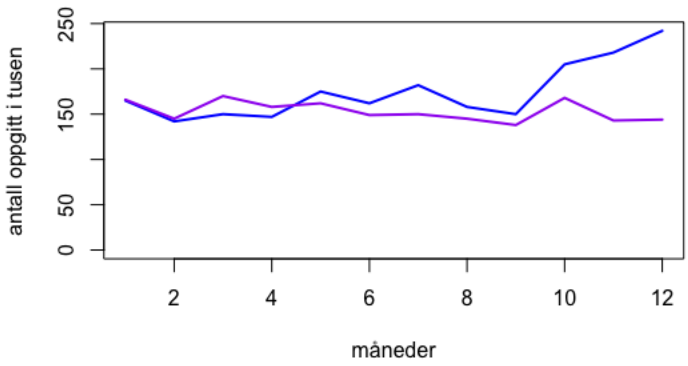
Manipulasjon av y-aksen kan føre til at en graf blir mer eller mindre dramatisk. Dette er en annen kjent feil som begås når man arbeider med grafer. Denne feilen kan begås bevisst eller ubevisst. Det vanligste problemet er å bare ta med deler av skalaen på y-aksen. Skalaen kan endres så mye at den mister kontekst [3].

Da smittetallene knyttet til koronaviruset økte i høst, ble det gradvis eller fullstendig nedstengning av de fleste landene i Europa. Russland valgte derimot å gå høsten i møte med åpne restauranter og barer. I 2020 hadde landet det høyeste dødstallet på 16 år. Grafen nedenfor viser dødsfall per måned i Russland de siste fem årene [4].



Figur 3: Manipulasjon av y-aksen. Hentet fra Aftenposten.

I figur 3 vises en graf med y-akse som starter på 140 tusen og slutter på 240 tusen. Dette fører til at grafen ser mer dramatisk ut enn det den egentlig er. Mellomrommet mellom den blå linjen som viser tall for 2020 og den blålilla linjen som viser tall for 2018 i måneden desember er stort. Tallene på y-aksen viser at differansen mellom de to linjene er fra rundt 140 000 tusen til 240 000 i denne måneden, med det visuelle inntrykket er en mangedobling.



Figur 4: Graf med y-akse som starter på 0.

Grafen i figur 4 inneholder deler av dataene oppgitt i figur 3. Den blå linjen representerer tallene for 2020 og den lilla representerer tallene for 2018. Denne grafen har en y-akse som starter på 0, og fremviser dataene på en mer realistisk måte. Mellomrommet mellom linjen for 2020 og linjen for 2018 i desember ser nå mindre dramatisk ut, og man ser at det tross alt ikke er snakk om en mangedobling.

2.3 Hvordan avsløre misvisende grafer

I kapittel 2 har jeg vist eksempler på hvordan grafer kan være misvisende. Det er viktig å se på grafiske fremstillinger med kritiske øyne, og være klar over hvilke fallgruver som kan ha oppstått. En graf bør vise akseinndelingen og ha med origo i plottet så sant det er rimelig [5]. Dersom en graf ikke inneholder disse elementene kan det fort oppstå misoppfatninger, som vist i figur 1. En annen fallgruve er manipulasjon av akser. Manipulasjon av aksene ved å for eksempel kutte deler av y-aksen vil gi en mer dramatisk fremstilling, som vist i figur 3.

3 Absolutte kontra relative tall

Man kan gjennomføre en statistisk undersøkelse ved enkel opptelling av antall hendelser. Tallene som blir brukt i undersøkelsen kan være absolutte tall eller relative tall, alt etter hva som er fokuset for undersøkelsen [5]. Absolutte tall vil si de tallene man får når man teller. Relative tall er tall satt i forhold til andre [6].

3.1 Eksempler

Eksempel 1: Anta at kvinner i en bedrift har en lønnsøkning på 6500 kroner i løpet av et år. Menn i denne bedriften har en lønnsøkning på 7000 kroner i løpet av det samme året. Det vil si at mennene har en absolutt lønnsøkning på 500 kroner mer enn kvinnene.
Eksempel 2: Kvinnene i samme bedrift som i eksempel 1 har en lønn på 325 000 kroner i året. Mennene har en årslønn på 450 000 kroner. I løpet av et år hadde kvinnene en lønnsøkning på 6500 kroner, og mennene en lønnsøkning på 7000 kroner. Det vil si at kvinnenes relative lønnsøkning var på 2 prosent, mens mennenes var på 1,56 prosent.
I eksempel 1 har det blitt brukt absolutte tall og i eksempel 2 har det blitt brukt relative tall. Det er en tydelig forskjell i disse to eksemplene. I eksempel 1 kan man konkludere med at mennene har en større lønnsøkning enn kvinnene. Men i eksempel 2 viser det seg at kvinnene har en større prosentvis økning enn mennene. Begge eksemplene viser korrekte resultater, men viser ulike aspekter.

3.2 Rot med absolutte kontra relative tall i praksis

I 2007 publiserte Aftenbladet en artikkel hvor en i fallskjermmiljøet kom med en påstand om at bilkjøring er farlige enn fallskjermhopping. Denne påstanden er basert på at det er flere ulykker ved bilkjøring enn fallskjermhopping [7].

Dagen etter denne avisartikkelen konkluderte en professor ved Universitetet i Stavanger at fallskjermhopping er farligere enn bilkjøring. Konklusjonen kommer av beregninger professoren har gjort. Beregningene viser at det er 0.1 prosent sanssynlighet for å omkomme i en fallskjermulykke i løpet av et år for en vanlig utøver, mens for bilkjøring er det tilsvarende tallet 0.01 prosent [8]. For tallene lå disse betingelsene til grunne:

For bilkjøring: Tall fra transportøkonomisk institutt viste at det ca var 20 omkomne per 100 millioner timer i bil. Per time vil det si 20 delt på 100 millioner. Dersom man bruker 500 timer i bil per år, vil sannsynligheten bli 500 ganger så høy. Utregningen blir da 1/10.000 - hvorav 0.01 prosent [8].

For fallskjermhopping: Professoren tok utgangspunkt i tall som viser én dødsulykke per 70.000 hopp. Hvis en fallskjermhopper hopper 70 hopp i året, vil utregningen bli 70/70.000 eller 1/1000 - hvorav 0.1 prosent [8].

Påstandene og konklusjonene gitt i begge avisartiklene er basert på absolutte eller relative tall. Teller man opp hvor mange ulykker som skjer i trafikken kontra hvor mange ulykker som skjer ved fallskjermhopp, vil man kunne konkludere med at antall ulykker er størst ved bilkjøring. Grunnen til dette er at det er langt flere som kjører bil enn som driver med fallskjermhopp. Ser man på ulykkestallene i forhold til andre tall, kan konklusjonen bli annerledes. Professoren ved Universitetet i Stavanger satte ulykkestallene i forhold til hvor mange timer man tilbringer i bil og hvor mange hopp en fallskjermhopper har, og konkluderte dermed med at det var større sannsynlighet for å omkomme i en fallskjermulykke enn i en bilulykke [8].

Oppsummert vil dette bety at dersom man undersøker absolutte tall, vil konklusjonen være at det skjer flere bilulykker enn fallskjermulykker. Undersøker man de relative tallene vil konklusjonen være at det er større sannsynlighet for å omkomme i en fallskjermulykke enn i en bilulykke. Men hva er da den farligste aktiviteten av disse to? Det kommer an på hvilket ståsted man har. Fra et samfunnsmessig ståsted kan det være av mest interesse å fokusere på de absolutte tallene. Den samfunnsmessige gevinsten kan kanskje være større ved å forsøke å få ned ulykkestallene ved en aktivitet som mange driver med og som har større totalt skadeantall, fremfor en aktivitet som få driver med [5].

3.3 Hvordan avsløre rot med absolutte kontra relative tall

I en statistisk undersøkelse hvor man gjør opptelling av antall hendelser må man vurdere om man skal bruke absolutte eller relative tall. Alt avhenger av hvilket fokus undersøkelsen retter seg mot. Både relative og absolutte tall kan være av interesse [5]. Dersom man for eksempel ønsker å undersøke skader ved ulike aktiviteter, vil det for den enkelte utøver være hensiktsmessig å vite noe om relative tall. Fra et samfunnsmessig perspektiv vil det kanskje være av mest interesse å vite noe om totalt antall ulykker, altså absolutte tall [5].

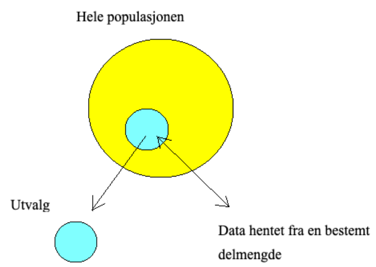
Er man klar over forskjellen på absolutte og relative tall kan man enkelt avsløre om en undersøkelse fremstiller feilaktige eller misvisende resultater. Det er også nødvendig å vite nøyaktig hva som er fokus i undersøkelsen, for å kunne avgjøre om det har blitt rot med absolutte kontra relative tall [5].

4 Ikke-representative utvalg

En viktig faktor for at resultatet i en statistisk undersøkelse skal være korrekt, er at de innsamlede dataene er hentet fra et representativt utvalg. For å få et representativt utvalg må man trekke et tilfeldig utvalg fra populasjonen. Et utvalg som er tilfeldig trukket fra populasjonen vil gjøre det mulig å trekke konklusjoner angående hele populasjonen [5]. Likevel er det flere undersøkelser som er basert på ikke-representative utvalg. Ikke-representative utvalg er utvalg fra en bestemt delmengde, og er ikke tilfeldige. Dersom et utvalg er ikke-representativt vil det være vanskelig å trekke konklusjoner for hele populasjonen [5]. I noen tilfeller kan det være utfordrende å se at utvalget er ikke-representativt.

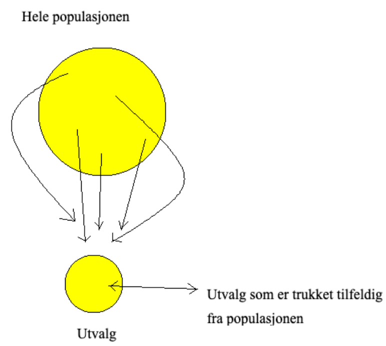
4.1 Eksempel og illustrasjoner

Anta at det skal gjennomføres en undersøkelse som undersøker hvordan Norges befolkning stiller seg til vindkraftutbygging. Dersom man spør 1000 personer som er bosatt i nærheten av det berørte området vil man ikke kunne bruke resultatet til å si noe om hva Norges befolkning generelt mener om temaet. Undersøkelsen har da et utvalg fra en bestemt delmengde av populasjonen, og det er ikke mulig å trekke konklusjoner for hele populasjonen basert på dataene i utvalget, se figur 5 [5].



Figur 5: Illustrasjon av ikke-representativ utvalg.

For at man skal kunne si noe om hvordan Norges befolkning stiller seg til vindkraftutbygging må man ha et representativt utvalg. Utvalget må være tilfeldig trukket fra populasjonen, se figur 6. Bare da kan man bruke de innsamlede dataene til å trekke konklusjoner for hele populasjonen [5].



Figur 6: Illustrasjon av representativt utvalg.

4.2 Reell undersøkelse: studiebarometeret

Studiebarometeret er en nasjonal undersøkelse som i teorien skal gjennomføres av 60.000 studenter ved høyskoler og universiteter hver høst. Den skal gjennomføres av 2. og 5. års studenter, og de aktuelle studentene får tilsendt en personlig lenke til undersøkelsen via e-post. Hensikten med undersøkelsen er å samle kunnskap om hva studentene mener om kvaliteten på studieprogrammene sine [9].

Merk at studentene får tilsendt en lenke på e-post som de må trykke seg inn på for å delta i undersøkelsen. Vil dette gi grunnlag for et representativt utvalg? Mest sannsynlig ikke. Det kan tenkes at flere studenter overser denne e-posten og dermed ikke svarer på undersøkelsen. Å overse e-posten kan skje bevisst eller ubevisst. Kanskje havner den i søppelpost og studentene ubevisst ikke får deltatt. Eller så kan studentene bevisst velge å ikke trykke seg inn på lenken. Dette innleder neste problem - nemlig at undersøkelsen er frivillig. Når en undersøkelse er frivillig, er det da tilfeldig hvem som svarer og ikke svarer? Sannsynligvis vil misfornøyde studenter se en anledning til å si ifra om det de er misfornøyde med og de som er forholdsvis fornøyde ikke gidde å svare på undersøkelsen - eller motsatt [5].

På studiebarometeret.no kan man søke på alle de representative studieprogrammene. Søker man på lektorutdanningen 8-13.trinn ved Universitetet i

Stavanger får man opp at dette studieprogrammet hadde en svarprosent på 26,4 prosent høsten 2020 [10]. Denne svarprosenten er svært lav, og vil ikke gi grunnlag for at Universitetet i Stavanger skal få innsikt om hva alle studentene på dette studieprogrammet mener.

Generelt er spørreundersøkelser hvor deltakerne må gjøre noe aktivt for å delta ikke til å stole på, ettersom man ikke får et representativt/tilfeldig utvalg [5].

4.3 Hvordan avsløre ikke-representative utvalg

For å kunne avsløre et ikke-representativt utvalg trenger man innsikt i hvordan de gitte dataene er samlet inn. Er de innsamlede dataene hentet tilfeldig fra populasjonen, eller er de hentet ut fra en delmengde? Det er også viktig å være klar over hvilke eventuelle begrensninger som ligger i måten innsamlingen er gjort på [5]. Dersom innsamlingen har blitt gjort gjennom en spørreundersøkelse hvor deltakerne aktivt må gjøre noe for å delta, vil innsamlingen mest sannsynlig ha begrensninger. Det er opp til hver enkelt person om de vil delta eller ikke, og det vil da være vanskelig å si noe konkret om hva hele populasjonen mener.

5 Simpsons paradoks

Simpsons paradoks er et fenomen beskrevet av Edward Simpson, som innebærer at en forskjell eller sammenheng kan forsvinne eller reverseres når grupper slås sammen. På motsatt vis kan en forskjell først komme til syne når man studerer ulike undergrupper [11].

5.1 Eksempel

Anta at et universitet hadde 8440 søkere. Resultatet av antall studenter tatt opp eller ikke tatt opp ved universitetet er fremvist i tabellen nedenfor. Dataene er delt inn i to grupper; kvinner og menn.

	Studenter tatt opp ved universitetet	Studenter ikke tatt opp ved universitetet	Total	Total prosentandel aksepterte søkere
Kvinner	2440	3300	5740	42,5%
Menn	850	1850	2700	32%
Total	3290	5150	8440	39%

Tabell 1: Resultat.

Tabell 1 viser at 42,5 prosent av de kvinnelige søkerne ved universitetet ble tatt opp, mens 32 prosent av de mannlige søkerne ble tatt opp. Tabellen antyder at en større andel kvinner enn menn blir tatt opp ved det gitte universitetet. Men er det nødvendigvis slik?

Ved å undersøke søkertallene nærmere kan man foreta enda en kategorisering. Anta at universitetet har kun to fakultet, et samfunnsvitenskapelig og et teknisk-naturvitenskapelig. De samme resultatene kan da presenteres på en annen måte med to undergrupper.

	Samfunns-vitenskapelig fakultet			
	Tatt opp	Ikke tatt opp	Total	Total prosentandel tatt opp
Kvinner	1550	400	1950	79%
Menn	150	10	160	94%
Total	1700	410	2110	80,5%

	Teknisk naturvitenskapelig fakultet			
	Tatt opp	Ikke tatt opp	Total	Total prosentandel tatt opp
Kvinner	890	2900	3790	23%
Menn	700	1840	2540	27,5%
Total	1590	4740	6330	25%

Tabell 2: Tabell med to undergrupper.

I tabell 2 ser man at bildet er snudd på hodet. I den første tabellen, tabell 1, ble data delt inn i to grupper: kvinner og menn som ble tatt opp eller ikke tatt opp ved universitetet. Ved å undersøke to nye undergrupper, altså to fakultet, kommer det frem at menn oftere blir tatt opp enn kvinner på begge fakultetene ved det gitte universitetet. Grunnen til at vi får motsatt resultat når vi ser på hele universitetet samlet er at menn i større grad søker på det teknisk naturvitenskapelige fakultetet enn det samfunnsvitenskapelige fakultetet. På det teknisk naturvitenskapelige fakultetet er det færre studieplasser og en lavere andel av studenter som blir tatt opp. På det samfunnsvitenskapelige fakultetet er det flere studieplasser og en høyere andel opptatte studenter.

Eksempelet vist i 5.1 belyser Simpsons paradoks ved at en forskjell først kom til syne når man studerte undergrupper. Hadde man ikke studert undergruppene ville konklusjonen vært at en større andel kvinner ble tatt opp ved universitetet, men egentlig viser det seg at en større andel menn ble tatt opp ved universitetet på begge fakultet.

Dette eksempelet er inspirert av en faktisk hendelse. Høsten 1973 ble opptakstallene til Universitetet i California, Berkeley, publisert. Opptakstallene ble oppfattet som kjønnsdiskriminerende ettersom det var mer sannsynlig at menn ble tatt opp ved universitetet enn kvinner [12]. Da undergrupper ble studert ble scenariet annerledes. Det viste seg at kvinner søkte på avdelinger hvor det var vanskelig å komme inn uavhengig av kjønn, mens menn søkte på avdelinger hvor det var flere studieplasser og lettere å komme inn. De samlede og korrigerte dataene viste en liten, men statistisk signifikant skjevhet til fordel for kvinner [12].

5.2 Matematisk fremstilling

Studenter tatt opp ved universitetet	
Kvinner	Menn
42,5%	32%

Fordelt på fakultet			
Samfunnsvitenskapelig fakultet		Teknisk naturvitenskapelig fakultet	
Kvinner	Menn	Kvinner	Menn
79%	83%	23%	27%

Tabell 3: Studenter tatt opp fordelt på fakultet.

Gitt A=tatt opp ved universitetet, B=kvinner og C=samfunnsvitenskapelig fakultet, da får man

$$P(A|B) = 0.425 > P(A|B^c) = 0.32 \tag{1}$$

$$P(A|B \cap C) = 0.79 < P(A|B^c \cap C) = 0.83 \tag{2}$$

$$P(A|B \cap C^c) = 0.23 < P(A|B^c \cap C^c) = 0.27 \tag{3}$$

Loven om total sannsynlighet gir:

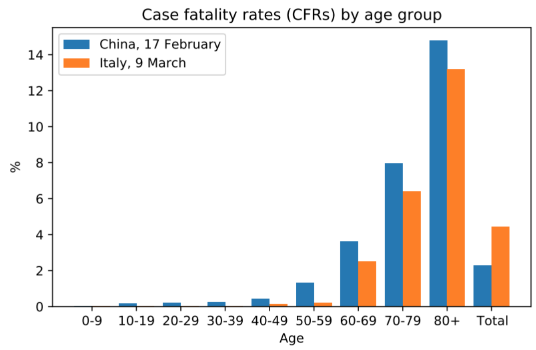
$$P(A|B) = P(A|B \cap C) * P(C|B) + P(A|B \cap C^c) * P(C^c|B) \tag{4}$$

$$P(A|B^c) = P(A|B^c \cap C) * P(C|B^c) + P(A|B^c \cap C^c) * P(C^c|B^c) \tag{5}$$

Fra likning (1) vet man at sannsynligheten for at en student blir tatt opp ved universitetet gitt at studenten er kvinne er større enn sannsynligheten for at en student blir tatt opp ved universitetet gitt at studenten er mann. Videre endrer ulikhetstegnene seg i likning (2) og (3), dette forteller at B og C må være ulike. Ut fra loven om total sannsynlighet ser man at paradokset kun vil skje dersom C og B er avhengige [13]. At C og B er avhengige vil si at sannsynligheten for hendelse C påvirkes av om hendelse B inntreffer. Dette kalles betinget sannsynlighet [14]. I dette tilfellet betyr det at sannsynligheten for at en student tas opp ved det samfunnsvitenskapelige fakultet påvirkes av om studenten er en kvinne eller en mann.

5.3 Simpsons paradoks i praksis

Koronapandemien har eksistert i verden i over et år. Italia og Kina var to av de hardest rammede landene i den tidlige fasen. Da dødelighetsdata fra disse to landene ble publisert, og man talte opp totalt antall døde, kom Italia dårligst ut [15]. Dødeligheten var dobbelt så stor i Italia som i Kina, sett i forhold til antall smittede. Men koronaviruset rammer befolkningen ulikt. Da forskerne tok alder med i betraktningen ble resultatet annerledes. For alle aldersgrupper var dødeligheten høyere i Kina enn i Italia dersom man så på aldersgruppene hver for seg. Altså - dødeligheten for alle innbyggerne sett under ett var høyest i Italia, men dødeligheten for aldersgrupper hver for seg var høyest i Kina [15].



Figur 7: Koronadødsfall kategorisert i aldergrupper. Hentet fra: https://arxiv.org/pdf/2005.07180v2.pdf

Dødeligheten av koronaviruset er størst blant eldre, og Italia har større andel eldre i befolkningen enn Kina. Samtidig har Italia flere smittede i de eldste aldersgruppene. Dette medfører at Italia kommer dårligst ut om man ukritisk ser på tallene for alle samlet [15].

5.4 Hvordan avdekke Simpsons paradoks

Man må være på vakt når man undersøker ulike størrelser og grupper. Statistiske konklusjoner må bære preg av sunn fornuft og god kunnskap om det fenomenet man undersøker. Å utelate en viktig variabel kan resultere i et helt feil bilde og gale konklusjoner [14]. Det er også viktig å bruke tid på å forstå hva undersøkelsen egentlig skal innebære [16].

For å kunne avdekke Simpsons paradoks kreves det at man har kunnskap om viktige grupperinger i dataene, og at man må ha informasjon om dataene i hver av disse gruppene. I eksempel 5.1 krevdes det kunnskap om de to fakultetene for å kunne avdekke paradokset. Det var flere studieplasser ved det samfunnsvitenskapelige fakultet kontra det teknisk naturvitenskapelige fakultet. Samtidig søkte flertallet av menn på det teknisk naturvitenskapelige fakultet hvor det var få studieplasser og en lav andel opptatte studenter.

6 P-verdi

Dersom man kaster en terning fem ganger, og får seksere på alle kast, har man skikkelig flaks. Det er en usannsynlig tilfeldighet. Likevel er det en teoretisk mulighet for at sekseren dukker opp uforholdsmessig ofte dersom man kaster terningen mange ganger [16]. Gjennomføres det en påfølgende dataanalyse kan denne vise at det er størst sjanse for å få seks på terningen. Men når det gjelder terninger, vet man allerede at det er like stor sannsynlighet for å få hvert av tallene på terningen. Det man trenger er et regneverktøy som kan beskrive usikkerheten tilfeldighetene skaper. Det mest brukte og kanskje mest omdiskuterte regneverktøyet er p-verdien [16]. En p-verdi sier noe om hvor sannsynlig et observert datasett er under en viss hypotese.

6.1 Hypotesetesting

Statistisk hypotesetesting er en metode for testing av hypoteser. Metoden innebærer at det er formulert en hypotese og at det finnes en motsatt hypotese, slik at man har to hypoteser å velge mellom. Disse kalles nullhypotese og alternativ hypotese [14]. En hypotesetest har fem trinn, disse er beskrevet i boken: Statistikk for universiteter og høgskoler, av Gunnar G.Løvås [14].

1. Velg passende sannsynlighetsmodell og formuler hypotesene.
2. Identifiser testobservator og bestem forkastningsområdet.
3. Velg hvor stor sannsynlighet for feilkonklusjon du kan godta.

4. Bestem forkastningsområdets kritiske grenseverdi. Vurder utvalgsstørrelsen og testens styrke.

5. Samler inn data og sammenligner observert verdi på testobservatoren med grenseverdien. Konkluderer til slutt.

6.2 Signifikanssannsynlighet/p-verdi

De siste trinnene i hypotesetestingen kan alternativt utføres ved å bruke p-verdi. Istedenfor å ta utgangspunkt i et valgt signifikansnivå kan man heller snu problemet på hodet og stille spørsmålet: Hvor sannsynlig er det å få de data som peker minst like mye i retning av alternativ hypotese som de man har observert, under antagelsen om at nullhypotesen er korrekt? Denne sannsynligheten kalles testens signifikanssansynlighet eller p-verdi [14]. P-verdien gir ikke indikasjoner på om hypotesen er sann eller ikke, men kan indikere om resultatet er påvirket av tilfeldigheter. P-verdien måler observasjonene opp mot den spesifikke nullhypotesen [17].

P-verdien oppgis mellom 0 og 1. Usikkerheten er mindre jo lavere p-verdien er [16]. En p-verdi på 0,05 eller lavere regnes ofte som statistisk signifikant. Dersom p-verdien er lav, betyr det at man kan forkaste nullhypotesen uten særlig risiko for å gjøre feil. Vanligvis kreves en p-verdi lavere enn 0.05 dersom man skal tørre å forkaste nullhypotesen [14].

6.2.1 Eksempel med p-verdi

Hvor sannsynlig er det å få tre seksere på rad under nullhypotesen at terningen er normal slik at P(sekser)=1/6.

$$P(\text{tre seksere}|\text{normal terning}) = 1/6^3 = 0.0046 \tag{6}$$

Sannsynligheten for å få tre seksere på tre kast av terningen, gitt at terningen er normal, er 0.0046. Dersom man skal teste nullhypotesen at terningen er normal, mot alternativt forhøyet sannsynlighet for sekser, vil man med dette resultatet forkaste nullhypotesen.

6.3 Feilbruk av p-verdier

Jakten på lave p-verdier har blitt et vanlig og alvorlig problem i vitenskapen [16]. Forskere i vår tid kan sitte på enorme mengder data, og når mengden variabler er stor nok vil man alltid kunne finne tilsynelatende sammenhenger når man begynner å lete [17].

Her kan man som en illustrasjon trekke inn terning-eksempelet fra 6.2.1. Anta at man utfører et forsøk med mange terninger. Man vil da teste om de har en forhøyet sannsynlighet for sekser, og forkaster nullhypotesen dersom man får tre seksere på rad. Utføres testen med 100 normale terninger er sannsynligheten for å få tre seksere på minst én av de

$$1 - (1 - 0.0046)^{100} = 0.369386 \tag{7}$$

Det vil si at det er ca. 37 prosent sannsynlighet for at minst én normal terning feilaktig blir erklært å ha forhøyet sannsynlighet for sekser.

I mange sammenhenger brukes

$$p - verdi < 0.05 \tag{8}$$

som grense for å forkaste nullhypotesen. Dersom man i terningeksempelet hadde valgt å forkaste nullhypotesen om normal terning dersom man kastet terningen to ganger og fikk to seksere, ville dette tilsvare en grense på p-verdi

$$P(\text{to seksere}|\text{normal terning} = 1/6^2 = 0.0278 \tag{9}$$

og for 100 terninger:

$$1 - (1 - 0.0278)^{100} = 0.94 \tag{10}$$

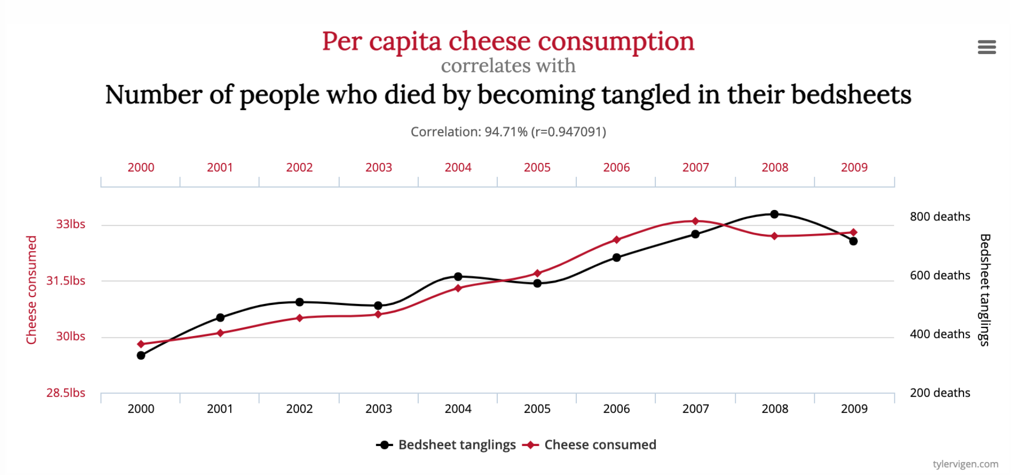
Ved kast av 100 terninger er det 94 prosent sannsynlighet for at minst én normal terning feilaktig blir erklært å ha forhøyet sannsynlighet for sekser.

6.4 Fisketur-problemet

Det å lete etter lave p-verdier, eller også kalt å dra på fisketur etter lave p-verdier, gir grunnlag for problemer. Det er stor fare for at det man finner ikke er reelle effekter når man leter rundt i dataene [5]. Når man gjennomfører mange tester vil det bli stor sannsynlighet for at minst én av dem gir en p-verdi på mindre enn for eksempel 0,05 uten at det er noen reell effekt [16].

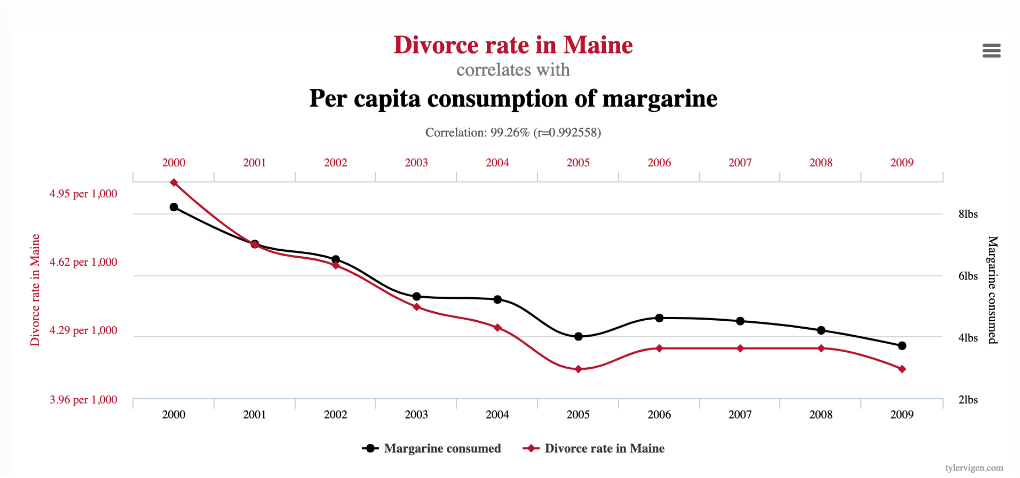
6.5 Feilbruk av p-verdier i praksis

Tyler Vigen har lett rundt i store databaser etter variabler som viser sammenheng, og viser så frem en del av de variablene som tilfeldigvis samvarierer på en nettside. Vigens nettside, tylervigen.com, inneholder grafiske fremstillinger av disse falske sammenhengene [18]. Nedenfor vises noen av dem.



Figur 8: Falsk sammenheng. Hentet fra tylervigen.com.

Grafen i figur 8 viser en falsk korrelasjon mellom antall mennesker som døde av å bli surret inn i sengetøyet og osteforbruk per innbygger [18].



Figur 9: Falsk sammenheng. Hentet fra tylervigen.com.

Også grafen i figur 9 viser en falsk korrelasjon. Her mellom skilsmisseraten i staten Maine i USA og margarinforbuk per innbygger.

Disse eksemplene er bevis på at hvis man leter rundt i store datamengder vil man finne ting som åpenbart ikke har noe å gjøre med hverandre [17].

6.6 Hvordan avdekke fisketurproblemet

Man må være klar over at konklusjoner som har blitt gjort gjennom hypotesetesting kan være et resultat av fisketur-problemet. Dette problemet oppstår når det har blitt gjort mange hypotesetester. For å avdekke om et funn ved en hypotesetest er et resultat av fisketur-problemet kan man undersøke funnet nærmere ved en oppfølgningsundersøkelse [5].

I situasjoner hvor man ønsker å teste mange ting bør man ideelt sett gå frem på denne måten:

1. På forhånd spesifiser hvilke hypoteser som skal testes
2. Bruk α/k som grense for hvor lav en p-verdi skal være før forkastning av nullhypotese. k=antall tester som skal gjennomføres, mens $\alpha = 0.05$.

Denne korreksjonen kalles en Bonferroni-korreksjon. Hvis man skal utføre 20 tester og normalt ville brukt 0.05 som en grense for p-verdien ved en enkel test, bør man heller bruke 0.05/20=0.0025 som en grense for forkastning av nullhypotesen [5]. Ved å bruke denne grensen er man i trygg grunn"i den forstand at den totale sannsynligheten for å gjøre feilaktig forkastning fremdeles er maksimalt 0.05. Uten denne Bonferroni-korreksjonen ville sannsynligheten for å feilaktig forkaste nullhypotesen i minst én av de 20 testene øke til inntil 64 prosent [5].

7 Feil kobling mellom årsak og virkning

Feil kobling mellom årsak og virkning er en vanlig fallgruve. Feilaktige konklusjoner om årsaksforhold er et vedvarende problem som finnes i vitenskapen [16]. En påvist statistisk sammenheng mellom to eller flere variable betyr ikke at det nødvendigvis er påvist et årsak-/virkningsforhold mellom dem [5]. For eksempel, er det påvist en statistisk sammenheng mellom lite søvn og hjertesykdom. Dette betyr ikke at lite søvn generelt gir hjertesykdom. Kanskje kan det hende at stress er den underliggende faktoren som påvirker både søvnen og risikoen for hjertesykdom. Feiltolkning av statistiske sammenhenger som årsakssammenhenger er velkjent, men er fremdeles noe man stadig ser [16].

7.1 Årsak/virkningsforhold eller andre underliggende årsaker?

En ny stor norsk studie har sett på hvordan ensomhet i ungdomsårene preger livet videre. Forsker Tilmann von Soest har i flere år undersøkt dette temaet, og høsten 2020 publiserte han en helt spesiell studie på ensomhet [19]. Langtidsstudien viser at de som har vært ensomme i ungdomsårene, tjener mindre når de kommer ut i jobblivet. Forskeren forteller videre at det å være ensom over lengre tid kan til slutt gjøre det vanskelig å å fungere sosialt på jobb. Thomas Severinius Nilsen som er seniorrådgiver ved Folkehelseinstituttet, forteller at en mulig sammenheng mellom ensomhet og lønn kan være at ensomhet kan påvirke kognitive funksjoner [19].

Studien viser at ensomhet påvirker hvor mye du tjener og at det er et funn som går igjen uansett hvordan tallene måles [19]. Men er denne påviste statistiske sammenhengen et årsak/virkningsforhold mellom ensomhet og lønn, eller skyldes sammenhengen andre underliggende årsaker? Generelt sett er dette vanskelig å konkludere sikkert om [5]. Det er mulig at det hadde vært mer hensiktsmessig å si at den underliggende årsaken er sosiale ferdigheter. Kanskje kan det være slik at svekkede sosiale ferdigheter både øker sjansen for å være ensom og for å tjene mindre. Det er ikke sikkert at det er rett å konkludere med at ensomhet automatisk fører til lavere lønn.

7.2 Hvordan avdekke/unngå feil kobling mellom årsak og virkning

Man må være oppmerksom på at en påvist statistisk sammenheng nødvendigvis ikke betyr at det er et årsak-/virkningsforhold. Det er vanskelig å finne en måte å avgjøre om en påvist statistisk sammenheng er et årsak-/virkningsforhold eller om sammenhengen skyldes andre underliggende årsaker. Ved å utføre en undersøkelse på den rette måten kan det være mulig å si noe om årsak/virkning, og andre ganger kan kunnskap om fenomenet man undersøker være til hjelp. Det er viktig å alltid stille spørsmålet: Finnes det en annen forklaring? [5].

8 Feiltolkning av sjeldne hendelser

Ingen dag er eksakt lik den foregående, selv om livet kan føles rutinepreget. Sjansen for at noe usannsynlig skal skje i løpet av en dag er kanskje ikke stor, men siden det er så mange dager i et liv bør man være klar over at man plutselig kan oppleve noe utenom det vanlige [20]. Fra en tid til annen hører vi om ”usannsynlige”hendelser som har inntruffet, f.eks at samme person vinner førstepremien i lotto to ganger. Selv kan man også oppleve slike "usannsynlige" hendelser. Kanskje har man tenkt på en person, og møter akkurat denne personen på gata dagen etter [5].

8.1 Den norske lotto-trekningen

Sannsynligheten for å vinne førstepremien i lotto-trekningen dersom man tipper en rekke er

$$1/5379616 = 1.86 * 10^{-7} \tag{11}$$

Tipper man flere rekker, er sannsynligheten for å vinne førstepremien fortsatt forsvinnende liten. Likevel er det førstepremie-vinnere nesten hver uke. Forklaringen på dette er at det vanligvis leveres inn rundt 20 millioner rekker [5].

Dette betyr at for den enkelte så er sjansen for å få 7 rette ekstremt lav, men på grunn av det store antallet innleverte rekker, vil det vanligvis være en eller få som får 7 rette [5]. Det blir feil å tenke at den eller de personene som vinner førstepremien er spesielle eller sjeldne, for ut fra hvor mange som tipper vil det som oftest være en tilfeldig deltaker som vinner uansett.

8.2 Feiltolkning av sjelden hendelse i rettsaker

Det faktum at en hendelse isolert sett har liten sannsynlighet for å oppstå ved tilfeldighet kan ikke nødvendigvis brukes som bevis for lovbrudd [13].

I England ble flere mødre dømt for drap basert på at to eller tre av deres barn døde i krybbedød. Bakgrunnen for dommen var at det er svært usannsynlig at en mor vil oppleve å miste to eller flere barn i plutselig krybbedød. Men i England er det flere mødre som opplever å få to eller flere barn, og over tid er ikke sjansen for at et fåtall av disse mødrene vil oppleve krybbedød veldig lav. Det vil si at det er svært lite sannsynlig at hver eneste mor i England vil oppleve dette, men veldig sannsynlig for samfunnet i helhet å se slike usannsynlige hendelser [13].

8.3 Hvordan unngå feiltolkning av sjeldne hendelser

Det å legge for mye tolkning i at det har skjedd sjeldne hendelser bør man unngå. Imidlertid kan man heller tenke gjennom hvor usannsynlig dette egentlig er ved å tenke over hvor mange ganger noe slikt har hatt muligheten til å inntreffe [5]. I juridiske sammenhenger er denne type tenkning også viktig. Hvis det eneste beviset mot en person er noe som kunne skjedd ved en tilfeldighet, må man legge til grunn en betraktning på hvor mange sjanser det har vært for at noe slikt kunne skjedd ved en tilfeldighet med hvem som helst og når som helst [5]. I krybbedødeksempelet vist i 8.2 burde tankegangen vært slik. Mødrene ble dømt for drap basert på at det hadde skjedd en sjelden hendelse. Det faktum at dette kunne skjedd ved en tilfeldighet med hvem som helst og når som helst ble ikke tatt i betraktning.

9 Bruk av fallgruver i undervisningssituasjoner

Matematikk er et sentralt fag i barn og unges utdanningsløp. Elevene møter faget allerede i første klasse, og noen av dem har matematikk i alle sine år på skolebenken. Matematikk er et bredt fag, med flere tema. Sannsynlighetsregning og statistikk er et av disse temaene. Jeg skal nå komme med eksempler på hvordan man kan trekke inn fallgruver i undervisningssituasjoner.

9.1 Sannsynlighetsregning og statistikk i barneskolen

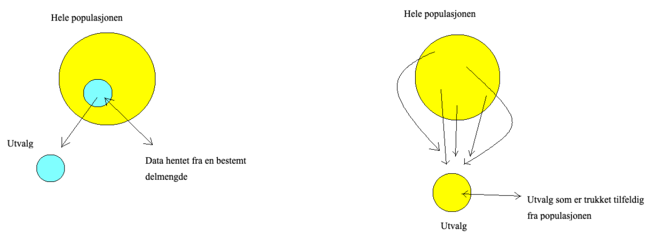
Elever i 7.klasse skal etter endt undervisning, i følge læreplanen, kunne logge, sortere, presentere og lese data i tabeller og diagrammer og begrunne valget av fremstilling [21].

La oss tenke at noen elever i en 7.klasse skal gjennomføre en undersøkelse hvor de vil finne ut hva som er klassens favorittbrus. Det første elevene må gjøre er å samle inn data. Når data er samlet inn må de sortere svarene og deretter velge en fremstilling som er egnet til denne undersøkelsen. Anta at Fanta var den brusen mesteparten av elevene likte best, og konklusjonen ble dermed at Fanta var klassens favorittbrus. Men, for at elevene skal kunne konkludere med at Fanta er klassens favorittbrus må de være sikre på at de innsamlede dataene er hentet fra et representativt utvalg.

9.2 Eksempel på hvordan ikke-representative utvalg kan vises i en undervisningssituasjon i 7.klasse

I brus-eksempelet vist i 9.1 vil utvalget være representativt dersom alle elevene i klassen var til stede og svarte på undersøkelsen. Bare da kan elevene trekke en konklusjon som gjelder for hele klassen. Dersom noen av elevene var borte fra skolen, eller ikke ville svare på undersøkelsen, kan ikke konklusjoner for hele klassen trekkes.

Jeg mener at det kan være lurt å la elevene gjennomføre en undersøkelse og finne et resultat, før de blir introdusert for representative og ikke-representative utvalg. Det kan være hensiktsmessig å først la elevene opparbeide en generell forståelse på hvordan man innheter og sorterer data. Deretter kan representative og ikke-representative utvalg trekkes inn. For at elevene skal forstå konseptet bedre kan illustrasjoner vises, se figur 10.



Figur 10: Illustrasjoner av representative og ikke-representative utvalg. Vist i kapittel 4.

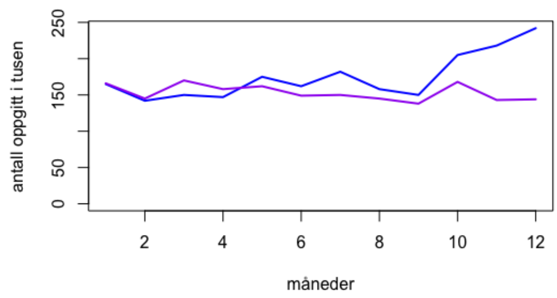
9.3 Sannsynlighetsregning og statistikk i ungdomsskolen

Tar man utgangspunkt i ungdomsskolen, møter elevene sannsynlighetsregning og statistikk på 9. trinn [21]. I læreplanen er det formulert ti kompetansemål, hvor store deler av dem omhandler sannsynlighetsregning og statistikk. Jeg skal ta utgangspunkt i to av disse kompetansemålene; Etter endt undervisning skal elevene blant annet kunne tolke og kritisk vurdere statistiske fremstillinger fra mediene, og de skal kunne utforske og argumentere for hvordan fremstillinger av tall og data kan fremme ulike synspunkter [21].

Disse to kompetansemålene innebærer at elevene har en viss kunnskap om hvordan data kan bli fremstilt. For å kunne tolke og kritisk vurdere ulike fremstillinger av data, er det i første omgang viktig at elevene vet hvordan man bruker statistikk riktig. Jeg mener at i en statistikkundervisning er det hensiktsmessig å først fokusere på gode måter å bruke statistikk på, før det eventuelt fokuseres på hvor enkelt det kan være å lyve med statistikk. Denne meningen kan begrunnes med det faktum at det ikke er statistikken det er noe galt med, men at det er folks kunnskap om statistikk som kan føre til misoppfatninger.

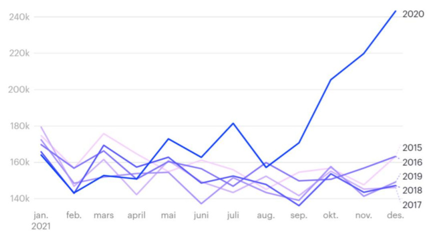
9.4 Eksempel på bruk av statistiske fremstillinger i en undervisningsituasjon på 9.trinn

Som nevnt ovenfor, ville jeg i en undervisningssituasjon gjerne begynt med eksempler som viste riktig fremstilling av data. Grafen i figur 11 inneholder akseinndeling og en y-akse som starter på 0. Denne grafen presenterer data riktig og er ikke misvisende på noen måte.



Figur 11: Riktig fremstilling av data. Vist i kapittel 2.

Etterhvert som elevene opparbeider seg kunnskap om riktig bruk av statistikk, tenker jeg at det vil bli enklere for dem å peke på feil. Da kan grafen nedenfor tas i bruk.



Figur 12: Feilaktig fremstilling av data. Vist i kapittel 2. Hentet fra Aftenposten.

Dersom elevene vet hvordan man lager riktige og troverdige grafer, vil de kanskje kunne oppdage feilen ved grafen i figur 12. Den åpenbare feilen, som nevnt i kapittel 2, er her at y-aksen starter på 140 tusen og ikke 0. Å manipulere y-aksen på denne måten vil gjøre grafen mer dramatisk.

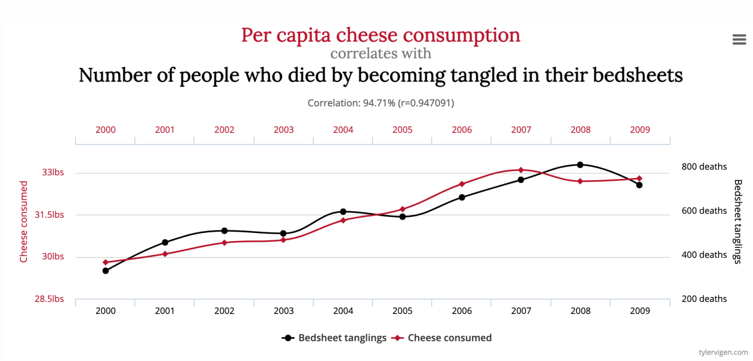
Ved å utforske disse to grafene vil elevene sannsynligvis være godt på vei til å nå kompetansemålene. Grafene kan bidra til økt bevissthet om at det kreves kunnskap for å forstå og tolke statistikk.

9.5 Sannsynlighetsregning og statistikk i videregående skole

I 3.klasse på videregående skole kan elevene velge å ta faget matematikk for samfunnsfag, S2. I dette faget skal elevene blant annet arbeide med hypotesetesting, og i følge læreplanen skal elevene kunne gjennomføre hypotesetesting i reelle datasett og kunne tolke resultatet [22]. I arbeid med hypotesetesting må elevene være klar over fisketurproblematikken. Dersom elevene gjennomfører mange tester kan de oppleve at det de finner ikke er reelle effekter.

9.6 Eksempler på hvordan feilbruk av p-verdier kan tematiseres i Vg.3

For å vise at mange hypotesetester kan gi grunnlag for funn uten reelle effekter ville jeg forklart terningeksempelet fra kapittel 6.3. Dersom man leter rundt i store databaser vil man kunne finne falske sammenhenger. For at elevene skal få en bedre forståelse for dette, ville jeg vist dem grafiske fremstillinger som viser falske sammenhenger. Et eksempel på en grafisk fremstilling som viser en falsk sammenheng er vist i figur 13 nedenfor.



Figur 13: Falsk sammenheng. Vist i kapittel 6.5. Hentet fra tylervigen.com.

10 Avsluttende ord

For at vi skal ha mest mulig nytte av statistikken omkring oss er det viktig at vi har tilstrekkelig kunnskap om fenomenet. Dette innebærer at vi må kunne skille mellom hva som er troverdig og ikke. Gjennom denne oppgaven har jeg presentert ulike fallgruver som man bør være obs på i møte med statistikk. Det er viktig å påpeke at det ikke er statistikken i seg selv som skaper problemer, men hvordan vi bruker og tolker den. Dersom vi er obs på disse fallgruvene vil det være enklere for oss å avsløre løgn med statistikk.

Referanser

[1] J. Best, *Stat-spotting: A field guide to identifying dubious data.* Univ of California Press, 2013.

[2] K. E. Aubert, *Grafisk fremstilling*, Store norske leksikon på snl.no, hentet 24.02.21, 2020. adresse: `https://snl.no/grafisk_fremstilling`.

[3] R. Mccready, *5 Ways Writers Use Misleading Graphs To Manipulate You*, hentet 24.02.21, 2020. adresse: `https://venngage.com/blog/misleading-graphs/`.

[4] H. Skjeggestad, *Hva ville det ha kostet om samfunnet fortsatte som normalt? Et land har forsøkt.* Aftenposten, 15.03.2021. adresse: `https://www.aftenposten.no/verden/i/yRm4yg/hva-ville-det-ha-kostet-om-samfunnet-fortsatte-som-normalt-et-land-ha`.

[5] J. T. Kvaløy, *Bruk statistikk riktig!* Upublisert notat, 2008. adresse: `https://www.ux.uis.no/~jtk/Rettbruk.pdf`.

[6] E. Nyberg, *Derfor må du kunne statistikk*, forskning.no, 2016. adresse: `https://forskning.no/data-statistikk-media/derfor-ma-du-kunne-statistikk/400798`.

[7] L. Lorentzen, *Bilkjøring er farligere enn fallskjermhopping*, Aftenbladet, 10.04.2007. adresse: `https://www.aftenbladet.no/lokalt/i/mvzb4/bilkjoering-er-farligere-enn-fallskjermhopping`.

[8] M. I. Line, *Fallskjerm mye farligere enn bil*, Aftenbladet, 11.04.2007. adresse: `https://www.aftenbladet.no/magasin/i/r957R/fallskjerm-mye-farligere-enn-bil`.

[9] Universitetet i Stavanger, 2020. adresse: `https://www.uis.no/nb/studiebarometeret-si-din-mening`.

[10] Kunnskapsdepartementet, 2020. adresse: `https://www.studiebarometeret.no/no/student/studieprogram/1160_m-lektor/`.

[11] J. E. Kristiansen, «Når tallene har noe å skjule,» *Samfunnsspeilet*, nr. 3, s. 42–45, 2010. adresse: `http://hdl.handle.net/11250/179892`.

[12] P. J. Bickel, E. A. Hammel og J. W. O'Connell, «Sex bias in graduate admissions: Data from Berkeley,» *Science*, årg. 187, nr. 4175, s. 398–404, 1975.

[13] J. T. Kvaløy, *Pitfalls and paradoxes in data analysis*, upublisert seminarnotat.

[14] G. G.Løvås, *Statistikk for universiteter og høgskoler.* Universitetsforlaget, 2013.

[15] J. Røislien og J. T. Kvaløy, *Simpsons paradoks - når pluss og pluss blir minus*, Tidsskriftet, Den Norske Legeforening, 2021. adresse: https://tidsskriftet.no/2021/03/medisin-og-tall/simpsons-paradoks-nar-pluss-og-pluss-blir-minus.

[16] K. S. Grønli, «Datamishandling - eller sju måter å feile med statistikk på,» *Forskningsetikk, Et magasin fra De nasjonale forskningsetiske komiteene*, nr. 4, s. 16–18, 2018.

[17] E. Torgersen og I. Kvittingen, *Hva er p-verdi og hva betyr statistisk signifikant?* forskning.no, 2019. adresse: https://forskning.no/matematikk-om-forskning-samfunn/hva-er-p-verdi-og-hva-betyr-statistisk-signifikant/1321080.

[18] T. Vigen, *Spurious Correlations*. adresse: https://tylervigen.com/spurious-correlations.

[19] E. N. N. Rydland, *Ung og ensom? Hold ut! Det blir bedre*, NRK, Viten, 25.04.2021. adresse: https://www.nrk.no/viten/ny-forskning-pa-ensomhet-blant-unge_-kan-pavirke-hvor-mye-du-tjener-1.15389790.

[20] J. Røislien og M. Nome, *Siffer*. Versal forlag, 2011.

[21] Utdanningsdirektoratet, *Læreplan i matematikk 1.-10.trinn*, 2020. adresse: https://data.udir.no/kl06/v201906/laereplaner-lk20/MAT01-05.pdf?lang=nob.

[22] Utdanningdirektoratet, *Læreplan i matematikk for samfunnsfag (matematikk S)*, 2021. adresse: https://data.udir.no/kl06/v201906/laereplaner-lk20/MAT04-02.pdf?lang=nob.