Arild Bjørndal og Dag Hofoss

STATISTIKK

for helse- og sosialfagene

Revidert utgave

GYLDENDAL
AKADEMISK

Arild Bjørndal og Dag Hofoss

Statistikk for helse- og sosialfagene

GYLDENDAL
AKADEMISK

© Gyldendal Norsk Forlag AS 2004

2. utgave

ISBN: 978-82-05-44467-6

Omslagsdesign: Merete Berg Toreg
Layout og sats: Laboremus Prepress AS
Brødtekst: Times 11,5/13,5 pkt

Alle henvendelser om boken kan rettes til
Gyldendal Akademisk
Postboks 6730 St. Olavs plass
0130 Oslo

www.gyldendal.no/akademisk
akademisk@gyldendal.no

Arild Bjørndal har mottatt støtte fra Det faglitterære fond.

Det må ikke kopieres fra denne boken i strid med åndsverkloven eller avtaler om kopiering inngått med KOPINOR, interesseorgan for rettighetshavere til åndsverk. Kopiering i strid med lov eller avtale kan medføre erstatningsansvar og inndragning, og kan straffes med bøter eller fengsel.

Forord

Mange synes statistikk er et vanskelig fag, og formler og tall gir ikke uten videre mening for alle. Samtidig har mange bruk for de hjelpemidler som statistikken tilbyr. Vi ønsker kanskje å redegjøre for et stort eller lite forskningsmateriale vi har samlet inn, eller vi vil være (passe) kritiske lesere av andres publikasjoner og ikke ta alt som serveres for god fisk. Vi vil også gjerne forstå noen av de grunnleggende trekkene ved den læringsformen vi kaller vitenskap. Da trenger vi en viss innsikt i tenkningen som statistikken er tuftet på.

Denne boken er blitt til fordi vi så det som en pedagogisk utfordring å presentere grunnleggende begreper, teknikker og tenkning på en slik måte at flere enn de med matematisk teft skulle ha en introduksjonsbok i statistikk på norsk. Boken retter seg mot dem som arbeider i helse- og sosialtjenesten og som trenger kunnskap i statistikk. De som ønsker en grunnleggende innføring, kan lese de første og det siste kapitlet. For andre er det bare å plukke fra de kapitlene som passer.

Vi håper at alle lesere sitter igjen med et håp («jeg skjønner det nesten, om ikke helt»), og at den svarte fortvilelsen kan holdes på avstand. Husk at de fleste av oss som liker statistikk har gått på mange kurs og lest flere bøker, og vi må ofte vende tilbake til bøkene. Statistikk er ikke et fag du tilegner deg ved å lese denne boka tvers igjennom én gang. Les litt, prøv deg på egne eller andres tall, les så litt til eller noe om igjen. En god definisjon av å lære er å kjenne igjen.

Forskning – i form av andres og egne små og store prosjekter – er spennende intellektuelt arbeid. I tillegg er det nyttig. Vi bringer verden fremover ved å skape mer innsikt og ved å dokumentere nytten av det ar-

beidet vi driver med i helse- og sosialtjenesten. Derfor er det vårt håp at mange av dere vil engasjere dere i forskning og bli ivrige brukere av forskningsbasert informasjon.

Lykke til!

Oslo 8. desember 2003
Arild Bjørndal og Dag Hofoss

Innhold

Kapittel 1 Introduksjon 13
1.1 Hvorfor er kunnskap i statistikk viktig? 13
1.2 Bokens siktemål 14
1.3 Statistikkens begrensninger 15
1.4 Om denne og andre lærebøker 15
1.5 Bokens oppbygging 16

Kapittel 2 Forskningsprosessen 18
2.1 Krav til forskning 18
2.2 Noen grunnleggende trekk ved forskningsprosessen 19
2.3 Grunnleggende om design 24
Eksperimenter, forsøk, andre effektvurderinger 24
Observerende design 25
Kunnskapsoppsummeringer 26
2.4 Spørsmålet avgjør metoden 26
2.5 Den typisk medisinske og den typisk samfunnsfaglige angrepsvinkelen 28
2.6 Problemstilling og modell i samfunnsfagene 29

Kapittel 3 Grunnleggende begreper 31
3.1 Verdi, variabel, enhet 31
3.2 Målenivå 32
Nominalt nivå 32
Ordinalt nivå 32
Intervallnivå 33
3.3 Populasjon og utvalg 34

3.4 Variasjon 35
Systematiske feil (bias) 36
Bakenforliggende faktorer 36
Tilfeldig variasjon 36
En valid sammenheng 38
3.5 Tilfeldighet og sannsynlighet 38

Kapittel 4 Univariat beskrivelse 41
4.1 Å fortette informasjon (fra rådata til frekvensfordelinger) . . . 41
4.2 Å beskrive det typiske 43
Gjennomsnitt 43
Median 43
Modus 44
4.3 Å beskrive variasjon 44
Variasjonsbredde 44
Prosentiler, desiler, n-tiler 45
Kvartiler, interkvartilbredden 45
Standardavvik 47
Variasjonskoeffisient 49
4.4 Å visualisere (grafisk framstilling) 50

Kapittel 5 Tilfeldig variasjon og konfidensintervall 53
5.1 Normalfordelingen 53
5.2 Hva er det en utnytter ved normalfordelingen? 56
5.3 Sentralgrenseteoremet 58
5.4 Hva avhenger utvalgsvariasjonen av? 60
5.5 Feilmarginer, konfidensintervall 62
5.6 Konfidensintervall omkring gjennomsnittet 63
5.7 Konfidensintervall omkring andeler 64
5.8 Noen poenger til 65
Alle størrelser beregnet ut fra utvalgsdata har en feilmargin 65
Statistikkprogrammer og konfidensintervallberegning 65
Den presise betydningen av konfidensintervallet 66
Normalfordeling og t-fordeling 66

Kapittel 6 Konfidensintervall for forskjeller 69
6.1 Konfidensintervallet omkring differansen mellom to andeler . 69
6.2 Hvorfor egen formel? 70
6.3 Konfidensintervallet omkring forskjellen mellom to gjennomsnitt 72

6.4 Hvor stort utvalg for å kunne fastslå en signifikant forskjell i andel? 73

Kapittel 7 Signifikanstesting, prinsippene 76
7.1 Signifikansnivå 76
7.2 Konfidensintervall og signifikans 76
7.3 Hvordan går en fram? 77
7.4 Tanken bak testen 79
7.5 Tolkningsproblemer 80
7.6 Tosidige og ensidige tester 82
7.7 Hvor stor er effekten? 83

Kapittel 8 Signifikanstesting, testene 84
8.1 Hvilken test? 84
8.2 To store utvalg: z-test 84
Et eksempel 85
Forutsetninger 86
8.3 To små utvalg: t-test 86
Eksempel 87
Forutsetninger 88
8.4 Pardata: t-test 89
8.5 Testing når dataene er på ordinalt målenivå (rangsumtester) . 90
8.6 Ordinale data fra to utvalg: Mann-Whitney-test 90
Eksempel 91
8.7 Ordinale pardata: Wilcoxon-test 93
Eksempel 93
8.8 Andre rangsumtester 94

Kapittel 9 Sammenheng i krysstabeller 96
9.1 Firefeltstabell og prosentdifferanse 96
9.2 Bivariate tabeller med mer enn fire ruter 96
9.3 Trevariate kryssfordelinger (å holde en tredje variabel konstant) 97
9.4 Firevariate kryssfordelinger (og vekting av prosentdifferanser) 99

Kapittel 10 Testing av sammenhenger i kryssfordelinger 106
10.1 Kji-kvadrattesten 106
10.2 Kji-kvadratfordelingen 108
10.3 Antall frihetsgrader 109
10.4 Forutsetninger 109

10.5 Begrensninger 109
10.6 Enda et eksempel 110
10.7 Kji-kvadrat for trend 113
10.8 Små utvalg: Fishers eksakte test for 2x2-tabeller 116
10.9 Hvilken test skal en velge når dataene er på nominalt eller ordinalt målenivå? 119

Kapittel 11 Korrelasjonskoeffisienter Statistisk testing av hypoteser om samvariasjon 120
11.1 Korrelasjon mellom nominalvariabler: kji-kvadrat (χ^2) og phi (ϕ) 120
11.2 Korrelasjon mellom grupperte variabler: korrelasjonskoeffisienten gamma (γ) 121
11.3 Korrelasjon mellom variabler på ordinalt nivå: korrelasjonskoeffisienten (Spearmans) rho (r) 123
11.4 Korrelasjon mellom intervallvariabler: produktmomentkorrelasjonskoeffisienten r (Pearsons r) 124
11.5 Statistisk testing av samvariasjon 127

Kapittel 12 Variansanalyse 131
12.1 Hva er variansanalyse? 131
12.2 Gjennomsnitt og varians 132
12.3 Dekomponering av varians: prinsippene 133
Hypotesetesting ved hjelp av F-brøken 134
12.4 Et talleksempel på variansanalyse 135
12.5 Multiple sammenlikninger 136
12.6 Et mål for grupperingsvariabelens effekt: eta^2 136
12.7 Forutsetninger for variansanalyse 137
12.8 Enveis, toveis og flerveis variansanalyse 137

Kapittel 13 Regresjonsanalyse 139
13.1 Fra korrelasjon til enkel lineær regresjon 139
Den rette linjen 141
13.2 Minste kvadraters metode 144
13.3 Å bestemme regresjonsstørrelsene a og b 145
13.4 Hva er hensikten med en regresjonslikning? 147
13.5 Ikke ekstrapolér utover observasjonsområdet 148
13.6 Forklart varians 149
13.7 Regresjon og utvalgstilfeldighet 150
13.8 Multippel regresjon 151
13.9 Standardiserte regresjonskoeffisienter 154

13.10 Regresjonsforutsetninger 156
Ikke-linearitet. 157
Interaksjon (ikke-additivitet) 158
Multikollinearitet 159
Heteroskedastisitet 161

Kapittel 14 Logistisk regresjon 164
14.1 Enkel logistisk regresjon (bare én forklaringsvariabel) 164
14.2 Multippel logistisk regresjon (det vil si flere enn én prediktor). 169
14.3 Å vurdere hvor god modellen er (goodness of fit) 173
Klassifikasjonstabellen. 173
Log likelihood (–2LL) 174

Kapittel 15 Overlevelsesanalyse 176
15.1 Overlevelsesanalyse og logistisk regresjon 176
15.2 Gjennomsnittsvarighet, problemet med sensurerte data. 176
15.3 Overlevelsesfunksjonen 178
15.4 Kaplan-Meiertabellen. 179
15.5 Partielle Meier-Kaplankurver, test av om kurvene er ulike . . 181
15.6 Cox-regresjon 184
Den avhengige variabelen i Cox-regresjon: hasardraten. 185
Prediktorene og deres koeffisienter: b og exp(b) 185
Modellers datatilpasning, «goodness of fit» 186
Forutsetninger for Cox-regresjon 186

Kapittel 16 Hvor stort må utvalget være? 189
16.1 Tilfeldige utvalg – og andre slags utvalg 189
Skjønnsmessige utvalg. 189
Bekvemmelighetsutvalg 189
Selvselekterte utvalg 190
16.2 Kravet til generaliseringstrygghet og presisjon bestemmer utvalgsstørrelse. 191
16.3 En likning for utregning av utvalgsstørrelse: prosentandel . . 192
16.4 En korreksjonsfaktor: utvalgsbrøkens betydning 194
16.5 Et eksempel med gjennomsnitt. 195
16.6 Betydningen av måten utvalget er trukket på. 196
16.7 Frafallsproblemet 197
16.8 En oppsummering i syv punkter. 199

Kapittel 17 Type 1- og type 2-feil
Statistisk styrke og dens betydning for utvalgsstørrelsen 201
17.1 Godtroenhet og overforsiktighet: type 1- og type 2-feil 201
17.2 Statistisk styrke 201
17.3 Et eksempel med differansen mellom to andeler. 202
17.4 Løsningen: undersøk et større utvalg. 206
17.5 Fastlegging av utvalgsstørrelse når en skal vurdere differansen mellom to utvalgsgjennomsnitt 208
17.6 Utvalgsstørrelsen når en skal vurdere differansen mellom to gjennomsnitt 212
17.7 Korreksjon av beregnet utvalgsstørrelse 212

Kapittel 18 Avslutning: «Back to basics» 213
18.1 Design 213
18.2 Organisering, analyse og tolkning 214

Appendiks 1 217

Appendiks 2 232

Appendiks 3 236

Stikkord 267

Kapittel 1

Introduksjon

1.1 Hvorfor er kunnskap i statistikk viktig?

De som arbeider i og med helse- og sosialtjenesten møter statistikken med forskjellig innstilling og ulike behov. Noen møter den med skepsis. «Statistics will prove anything, even the truth», er det sagt. Faget er preget av matematikk og formler, som mange synes er vanskelig.

Denne skepsisen bør overvinnes. Det hefter variasjon og usikkerhet ved all informasjon vi innhenter i helse- og sosialfagene. Når tall skal tolkes i lys av denne usikkerheten, trenger en den hjelpen statistikken tilbyr. Det gjelder både i den foredlingsprosessen vi kaller kunnskapsproduksjon (forskning), og når vi gjør en faglig vurdering i arbeidet med pasienter og klienter. Statistisk kunnskap hjelper en dessuten å kryste essensen ut av et datasett på en så overbevisende måte at selv den mest motstrebende tviler må akseptere budskapet. For så vidt er statistikk ikke bare en matematisk, men også en pedagogisk disiplin.

Kunnskap i statistikk gjør det lettere å planlegge, innsamle, organisere, analysere og tolke kvantitative data. Forskere må derfor kunne statistikk. De ønsker å bearbeide relevant og pålitelig informasjon som er samlet inn, og å videreformidle funnene. De vil konsentrere seg om det viktigste, utelate det som er uvesentlig, sammenlikne, generalisere, lete etter mønstre og vurdere sammenhengen mellom årsak og virkning. «Forskere» er for øvrig et begrep vi ikke vil innsnevre i denne boken; med forskere mener vi både studenter og professorer, både de med små og de med store prosjekter. De har alle en grunnleggende ting felles; i sine «prosjekter» prøver de å «vaske gull», det vil si finne fram til kunnskap som har gyldighet for andre.

Men også alle praktikere som ønsker å følge med i sitt fag, trenger et visst kjennskap til statistikkens grunnleggende begreper og tankegang. En leser uten statistisk innsikt blir forsvarsløs og ikke kritisk nok. I mange land arbeider en systematisk for å fremme en mer kunnskapsbasert

praksis i helse- og sosialtjenestene. Alt for ofte viser det seg å være en kløft mellom det vi egentlig vet (den kunnskapen som finnes i verden) og det vi gjør i praksis. Pålitelig forskningsbasert informasjon blir ikke tatt i bruk og kommer ikke til nytte. En av årsakene er svak kompetanse i kunnskapshåndtering; de fleste utøvere i tjenestene har ikke lært hvordan de kan bli effektive, kritiske og konstruktive brukere av forskning. Kunnskap i statistikk inngår som en naturlig del av faget kunnskapshåndtering[1].

Vi har hatt både forskere, studenter og praktikere i tankene når boken ble til, selv om boken har slagside mot forskernes behov – med hovedinnretning mot enkle forskningsarbeider basert på egen praksis, og masteravhandlinger og hovedoppgaver.

1.2 Bokens siktemål

Vårt første hovedanliggende er å gi leseren innsikt i tilfeldig variasjon som fenomen. Tilfeldig variasjon er uløselig knyttet til enhver måling. Behovet for å skille den rene tilfeldighet fra den viktige tendens er selve statistikkens begrunnelse. Vår neste ambisjon er å bistå den som vil utforske strukturen i sin datamengde og klarlegge eventuelle sammenhenger mellom observasjoner.

Disse to målene gjenspeiler to analytiske perspektiver, det medisinske og det samfunnsvitenskapelige. Selv om det ligger samme form for rasjonalitet i bunnen av de to fagene, og samme ivrige bruk av statistikk blant forskerne, bruker en ofte forskjellige design og ulike analytiske framgangsmåter. Vi vil vise fram begge disse kvantitative tradisjonene.

I jakten på disse målene vil vi veilede i valg av statistisk teknikk. En ting er å forstå isolerte omtaler av spesielle emner, en annen er å finne fram til den rette analyseformen av et datasett. Det handler om å kjenne metodenes forutsetninger og begrensninger. I datamaskinens tidsalder er regnefeil en sjeldenhet (i hvert fall blir de sjelden oppdaget), mens feilaktig applikasjon av bestemte metoder forekommer hyppig.

Et tredje mål er av pedagogisk art. Vi er mer opptatt av å beskrive ideer enn å framvise kompliserte utregninger. Det betyr at vi bruker ord mer enn formler. Nå kan det være grunn til å minne om at formler bare er korte og konsise uttrykk for et resonnement. Men for mange er matematiske uttrykk en snublestein og ikke en hjelp. Vår erfaring fra flere års undervisning i statistikk er at de fleste som jobber i og helse- og sosial-

1 Bjørndal, A., Flottorp, S. og Klovning, A. (2000): *Medisinsk kunnskapshåndtering*. Oslo: Gyldendal Akademisk.

tjenesten, har hatt lite opplæring i statistikk og er lite motivert for å gi seg i kast med statistiske formler. Boken er skrevet på grunnlag av denne erfaringen. Den er ikke fri for formler, men alle ledsages av et intuitivt resonnement i en vanlig språkdrakt. Av og til er det viktig å vise fram utregningsmåter for å illustrere prinsipper. Likevel har vi forsøkt å ta hensyn til at vi lever i pc-ens tidsalder, og at forskeren sjelden nøyer seg med blyant og kalkulator som hjelpemidler.

1.3 Statistikkens begrensninger

«Hur mäter en vackert» er tittelen på en bok om kvalitativ metode. Den illustrerer at det finnes sider ved livet og menneskene som ikke så lett lar seg fange i tall og tabeller. Etter vår vurdering går det ikke noe prinsipielt viktig skille mellom kvalitative og kvantitative metoder; begge er framgangsmåter for å skape erkjennelse ut over det fenomenet en akkurat nå har informasjon fra. De er velegnet for å besvare hvert sitt sett av spørsmål. Denne boken handler imidlertid om kvantitative metoder.

Med statistikken som utgangspunkt kan en bli mer opphengt i forskjeller enn i likheter. At kvinner bruker mye mer psykofarmaka enn menn, kan få en til å overse det forhold at de fleste mennesker overhodet ikke bruker slike legemidler. En kan også glemme verdien av negative funn. Studien som viser at et tiltak ikke virker, kan være like verdifull som undersøkelsen som viser at det virker.

Det statistiske faget beskjeftiger seg i hovedsak med data som allerede er samlet inn. En god analyse forutsetter imidlertid at grunnleggende krav til undersøkelsesopplegg og datainnsamling er oppfylt. Her grenser statistikken opp mot andre fag rundt et slags fellesområde der grunnleggende krav til forskningsprosessen, framgangsmåter for å samle inn data (design) og, mer generelt, forskningsmetode er temaet.

1.4 Om denne og andre lærebøker

Det foreligger mange lærebøker i statistikk. Når denne boken likevel er blitt til, er årsaken todelt. For det første foreligger det få bøker på norsk om statistikk anvendt på helse- og sosialfaglige problemstillinger. For det andre er det vår klare erfaring at det eksisterer et behov for en ikke-matematisk introduksjonstekst.

Bakerst i denne boken vil leseren finne en kort omtale av en rekke lærebøker i statistikk, hvorav flere har vært til inspirasjon for denne teksten (se appendiks 2). Hvis en virkelig ønsker å lære statistikk, men

synes det er vanskelig, er det en god idé å konsultere flere bøker. Det er ofte en hjelp å lese forskjellige formuleringer av samme problemstilling. Ordene varierer, eksemplene likeså, og forskjellige forfattere kan derfor kaste lys over et emne fra ulike vinkler. Gjenkjenning er en viktig form for læring.

1.5 Bokens oppbygging

I *kapittel 2* omtaler vi sentrale typer av design i anvendt forskning og viser hvordan forskningsspørsmålet avgjør valget av design (som i sin tur preger valget av statistisk analyseteknikk).

Som en introduksjon til de senere kapitlene går vi i *kapittel 3* gjennom en del basale begreper og ideer, som enhet, variabel, verdi; målenivåer; populasjon og utvalg; variasjon; tilfeldighet og sannsynlighet.

I *kapittel 4* presenteres noen viktige teknikker for oversiktlig og lettfattelig framstilling av fordelingers tyngdepunkt (gjennomsnitt, median, modus) og spredning (modalprosent, variasjonsvidde/interkvartilbredde, prosentiler og standardavvik).

I *kapittel 5* viser vi hvordan en utnytter kunnskap om den tilfeldige variasjonen til å beregne feilmarginer. Forskeren har for eksempel beregnet et gjennomsnitt eller en prosentandel i et utvalg, men ønsker å generalisere fra utvalget til hele befolkningen. Da må han/hun ta hensyn til at utvalgsresultatet er beheftet med usikkerhet, og ikke bare angi det gjennomsnittet eller den andelen han/hun fant i utvalget, men også tallets feilmargin.

I *kapittel 6* går vi videre med resonnementene om feilmarginer og appliserer dem på forskjeller mellom gjennomsnitt og mellom prosentandeler.

Kapitlene 7 og 8 dreier seg om signifikanstesting. Kapitlene om konfidensintervaller har svart på spørsmålet om feilmarginers bredde. Signifikanstesting gir svar på et litt annet spørsmål, nemlig spørsmålet om hvor sannsynlig det er at et observert resultat ikke er annet enn en utvalgstilfeldighet. *Kapittel 7* gjør greie for prinsippene for signifikanstesting, mens *kapittel 8* gjennomgår ulike signifikanstester.

I *kapittel 9* tar vi opp grunnleggende teknikker for analyse av sammenheng mellom krysstabulerte variabler. Vi viser noen mål for sammenheng i bivariate krysstabeller (gamma og prosentdifferanse). Kapitlet viser også hvordan mer kompliserte kryssfordelinger kan leses. Vi kontrollerer sammenhenger for tredje og fjerde variabler, og introduserer begrepet statistisk interaksjon.

Kapittel 10 viser hvordan en tester for om fordelingsulikheter kan tilbakeføres til ren utvalgstilfeldighet, eller om en kan ta sjansen på å tro at de viser at variablene virkelig samvarierer.

Kapittel 11 handler om korrelasjonskoeffisienter – og om hvordan en undersøker faren for at korrelasjon i et utvalg bare skyldes utvalgstilfeldighet.

Kapittel 12 gir en innføring i variansanalyse og introduserer begrepene totalvarians, restvarians og forklart varians.

I *kapittel 13* tar vi opp enkel og multippel regresjon. Vi diskuterer resonnementene bak denne analyseteknikken, viser hvordan en beregner konstantledd og regresjonskoeffisienter – og diskuterer hvordan en undersøker risikoen for at utvalgsfunnene våre kan komme av rene utvalgstilfeldigheter.

Kapittel 14 handler om logistisk regresjon: standardmetoden for analyse av dikotome avhengige variabler, dvs. av sannsynligheten for at en begivenhet skal inntreffe, og hva som påvirker den.

I *kapittel 15* gir vi en kort beskrivelse av overlevelsesanalyse og Cox-regresjon, som i ett perspektiv er en fortsettelse av logistisk regresjon. Den viktige forskjellen er at en ikke nøyer seg med å studere om en begivenhet inntreffer eller ikke, men også spør hvor lang tid det tar før den (eventuelt) inntreffer, og hva som påvirker hvor lang tid det tar.

Resonnementene i kapittel 6 omkring konfidensintervall for andeler og gjennomsnitt danner utgangspunkt for svaret på spørsmålet om hvor stort utvalget må være for at en skal kunne bestemme populasjonsandeler og -gjennomsnitt med tilfredsstillende sikkerhet og presisjon. Forskeren kan uttale seg tryggere og med mindre feilmargin om han/hun har data fra et større utvalg, og i *kapittel 16* viser vi hvordan utvalgsstørrelsen forhåndsfastlegges.

Kapittel 17 introduserer begrepene statistisk styrke og type 1- og type 2-feil. Til nå har vi bare drøftet hvordan en unngår å gjøre den feil at en fester lit til funn som skyldes utvalgstilfeldigheter (type 1-feil). Det er viktig. Men en må også vokte seg for den motsatte feil; å la være å tro på funn som er riktige (type 2-feil). Problemet oppstår fordi en aldri vet sikkert om et funn er en utvalgstilfeldighet eller ei. Balanseringen av faren for godtroenhet mot faren for overforsiktighet er vanskelig, for en øker den ene når en vil minske den andre – med mindre en øker størrelsen på utvalget. Kapittel 17 viser derfor også hvordan en forhåndsfastlegger utvalgsstørrelse ut fra hensynet til teststyrke; undersøkelsens evne til å unngå type 2-feil.

I *kapittel 18* forsøker vi å samle trådene og gi en oppsummering av hva statistikken kan hjelpe oss med.

Kapittel 2

Forskningsprosessen

2.1 Krav til forskning

Selv om dette ikke er en bok som pensler ut detaljene i hvordan forskeren planlegger og gjennomfører sitt prosjekt, ønsker vi i dette kapitlet å sette innhentingen og tolkningen av tallene inn i en større sammenheng. Forskning er en viktig kilde til erkjennelse i helse- og sosialfagene – mange vil si den viktigste. Det betyr ikke at vi mangler respekt for andre erkjennelsesformer, for eksempel tro eller praksiserfaring. Men vi er opptatt av å jobbe med innholdet i våre fag og ønsker å avvise for eksempel hjernebalansering og kraniosakral massasje som aktuelle behandlingsformer – inntil noen eventuelt legger dokumentasjon på bordet som viser at det virker. Med dokumentasjon mener vi noe annet enn enkeltstående beretninger om at noen har hatt nytte av tiltaket, og da er vi straks over i spørsmålet om hvordan vi organiserer en systematisk og uhildet innhenting av informasjon om effekter, det vil si kriterier for hva som er forskning.

Rute 2.1 Forskning og kritisk vurderingsevne som beskyttelse

Science is more than a body of knowledge; it is a way of thinking. I have a foreboding of an America in my children's or grandchildren's time – when the United States is a service and information economy; when nearly all the key manufacturing industries have slipped away to other countries; when awesome technological powers are in the hands of a very few, and no one representing the public interest can even grasp the issues; when the people have lost the ability to set their own agendas or knowledgeably question those in authority; when, clutching our crystals and nervously consulting our horoscopes, our critical faculties in decline, unable to distinguish between what feels good and what's true, we slide, almost without noticing, back into superstition and darkness.

Carl Sagan[2]

2 Sagan, C. (1997): *The demon-haunted world*. New York: Ballantine Books.

To sentrale krav som skiller forskning fra andre erkjennelsesformer er:

- universalisme
- systematisk skepsis

Med universalisme menes at hvorvidt ny kunnskap anerkjennes eller forkastes, ikke skal være avhengig av forskerens personlige og sosiale forhold, bl.a. nasjonalitet, religion og utdanningsbakgrunn. Det avgjørende er forskningens kvalitet, ikke forskerens individualitet. Det er imidlertid ikke grunn til å stikke under en stol at det i kunnskapssamfunnet er flere oppfatninger av hva som er god og pålitelig forskning. Ikke bare varierer synet mellom ulike fag – for eksempel mellom samfunnsøkonomer og medisinere – også innen fagene er det varierende meninger. Felles for alle er imidlertid et grunnsyn som handler om at det må foreligge kriterier som skiller mellom forskning og «synsing».

Systematisk skepsis beskriver et syn på kunnskapsutviklingen som forutsetter logisk, kritisk og konsistent tenkning – og systematisk praktisk arbeid (for eksempel datainnhenting), der ingenting tas for gitt. Å tilhøre et forskningsbasert fag er derfor noe ganske annet enn å tilhøre et trossamfunn. «In God we trust. All others: bring your data». Det betyr ikke at vi glemmer at vi også opererer innenfor et tankesett – et paradigme – og at det fra tid til annen kommer paradigmeskifter. Men det betyr at vi tilhører – per fagdefinisjon og av intellektuell legning – skeptikerne i verden.

2.2 Noen grunnleggende trekk ved forskningsprosessen

Den klassiske beskrivelsen av naturvitenskapelig forskning har vært at forskeren iakttar naturen, gjør sine observasjoner og deretter stiller opp sine hypoteser (antakelser) og teoretiske, sammenfattende betraktninger. Observasjonen er ikke preget av forskerens eget ståsted, forventninger eller verdinormer; alle som betraktet det samme fenomenet ville komme til det samme resultat. Enkeltobservasjoner leder til hypoteser gjennom en prosess som gjerne betegnes som *induksjon*, det vil si å slutte fra enkelttilfellet til det allmenne. Ut fra de framsatte hypotesene eller antakelsene kan en *dedusere* hva en forventer av undersøkelser i framtiden, det vil si slutte fra det allmenne til det spesielle, og – hvis forventningene holder – har en verifisert (bekreftet) hypotesen.

I kontrast til en slik forenklet beskrivelse står *kravet om falsifikasjon* slik det er beskrevet framfor alt av Karl Popper (1902–1994). For Popper innledes vitenskapsprosessen ved at en framsetter en hypotese. Fruktbare hypoteser kan ha rot i mer eller mindre systematiske observasjoner eller for så vidt være hentet ut av den kreative fantasien. Det avgjørende for at vi skal kalle det forskning, er kravet om at forskeren utleder sannsynlige utfall på grunnlag av sin antakelse og deretter gjør det han/hun kan for å avkrefte hypotesen. Dette kriteriet, hvorvidt en teori faktisk kan testes, bruker Popper for å avgrense hva som er vitenskap (og slik avviser han Freuds teorier og det marxistiske historiesynet som vitenskap fordi disse teoriene ikke lar seg teste). Vår viten på et bestemt tidspunkt utgjøres derfor i en viss forstand av de hypoteser en ikke har lykkes med å falsifisere.

Etter falsifikasjonsteorien består kunnskapsproduksjon ikke av å bekrefte, men å avkrefte ideer om sammenhenger mellom årsak og virkning. Den som ser at en pasient blir frisk etter en behandling, vil – hvis en følger en induktiv tankegang – føle seg mer overbevist om at behandlingen virker. Den som lar seg inspirere av Poppers tanker, vil i stedet sette denne teorien på prøve og etterspørre et vel utført kontrollert forsøk som har potensial til å avkrefte sammenhengen mellom behandling og vellykket utfall. Men i dette ligger det et problem hvis vi (som Popper) skal betrakte den deduktive logikken som tilstrekkelig for å beskrive all vitenskap. En trenger i tillegg en mulighet for bekreftelse, ikke bare avkreftelse. I så fall må vi ta i bruk det induktive resonnementet, uttrykt i form av sannsynligheter.

I praksis følger helse- og sosialfaglig forskning begge disse sporene. Det er rimelig å skille ut to hovedtyper: beskrivende (deskriptiv) og hypotetisk-deduktiv forskning. I den første tradisjonen er det ikke så viktig å ha oppstilt klare hypoteser; en gjør en rekke observasjoner av hvordan naturen er og forsøker å skape et bilde ut fra enkeltobservasjoner. Slutninger fra enkeltobservasjoner til mer omfattende utsagn (fra det spesielle til det allmenne) kalles som nevnt over for induksjon. Induktive slutninger er i sin natur ikke logiske nødvendigheter, men sannsynlighetsbetraktninger. Et induktivt argument går fra en avgrenset liste med observasjoner (premissene) og trekker en konklusjon ved hjelp av det vi i dagligtalen kaller generalisering. I motsetning til det deduktive argumentet har vi ingen garanti for at slutningen er sann.

Hvis du ser gjennom tidligere eksamensoppgaver og finner at studentene hver gang får i oppgave å reflektere over WHOs definisjon av helse, er det rimelig å forberede seg på dette (du slutter fra enkeltobservasjoner til det allmenne). Men du har ingen garanti for hva som hender når det er