

Jan Erik Kristiansen



Jan Erik Kristiansen er sosiolog og seniorrådgiver i Statistisk sentralbyrå, Formidlingsavdelingen. Han har lang erfaring i å presentere statistikk på en brukervennlig måte og har holdt en rekke kurs i statistikkforståelse og bruk av statistikk.

Han utga i 2007 boken «Tall kan temmes!» (U-forlaget), som delvis danner grunnlaget for denne spalten.

(jan.erik.kristiansen@ssb.no)

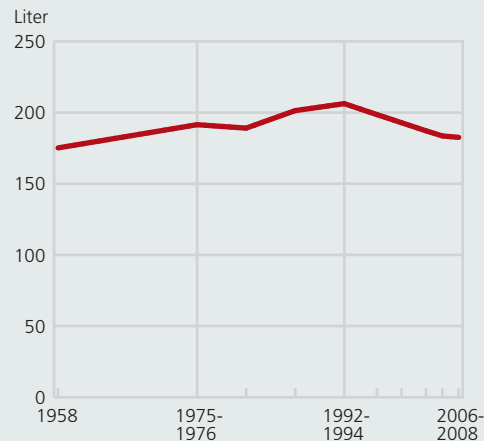
Når tallene har noe å skjule

Tallene lyver ikke, men de kan av og til skjule sannheten. Ikke sjelden finner vi bare små eller ingen endringer eller forskjeller når vi studerer utviklingen av et fenomen over tid eller sammenlikner ulike grupper. Men dette er ofte bare tilsynelatende. Om vi går «bak» totaltallene eller gjennomsnittet, finner vi ofte store endringer eller forskjeller. Helheten varierer mindre enn delene.

Endrede drikkevaner

Ser vi på forbruket av alkoholfrie drikkevarer som melk, brus, mineralvann og juice, har dette vært nokså konstant siden 1958. Det økte noe i begynnelsen av perioden, men har siden sunket til nesten samme nivå (se figur 1).

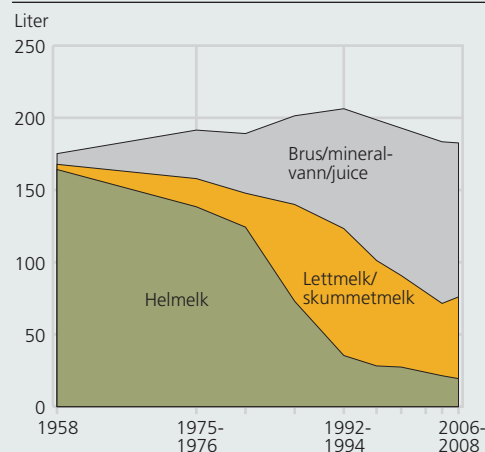
Figur 1. Samlet forbruk av melk og mineralvann. 1958-2008. Liter per person



Kilde: Forbruksundersøkelsene, Statistisk sentralbyrå.

Under denne tilsynelatende stabiliteten skjuler det seg imidlertid store endringer i våre drikkevaner. Dette blir tydelig når vi fordeler det samlede forbruket på ulike typer drikkevarer. Melkeforbruket er redusert – fra nesten 170 liter per person til knappe 80. Og mens vi tidligere for det meste drakk H-melk, er det nå lettmeik og skummet melk som dominerer (se figur 2).

Figur 2. Forbruk av melk og mineralvann. 1958-2008. Liter per person



Kilde: Forbruksundersøkelsene, Statistisk sentralbyrå.

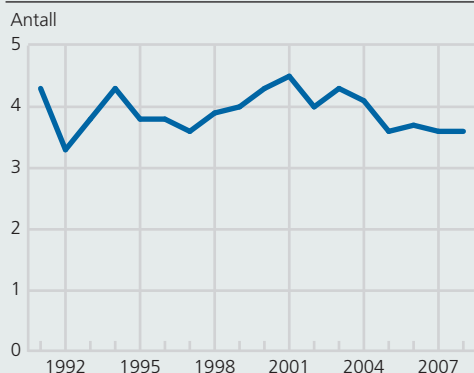
Til gjengjeld er konsumet av alkoholfrie drikkevarer (mineralvann, juice og brus) mangedoblet siden 1958, og vi drikker i gjennomsnitt 106 liter per år. Denne økningen tilsvarer omtrent nedgangen i melkeforbruket.



Stabilt kinobesøk, men ...

Siden begynnelsen av 1990-tallet har antallet kinobesøk per person vært relativt stabilt og har ligget på rundt fire kinobesøk per person per år (se figur 3).

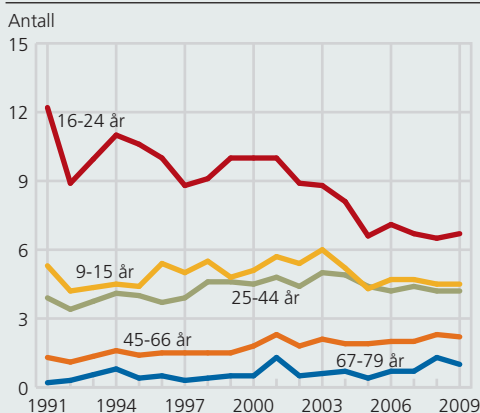
Figur 3. **Antall kinobesøk per person. 1991-2009**



Kilde: Kultur- og mediebruksundersøkelsene, Statistisk sentralbyrå.

Men om vi ser på ulike aldersgrupper, finner vi at utviklingen har vært svært forskjellig. I aldersgruppen 16-24 år, som er de ivrigste kinogjengerne, er kinobesøket nesten halvert i løpet av perioden. For de aller yngste er tallene stabile, og for de øvrige aldersgruppene er kinobesøket økende, særlig for de aller eldste (se figur 4).

Figur 4. **Antall kinobesøk per person i ulike aldersgrupper. 1991-2009**



Kilde: Kultur- og mediebruksundersøkelsene, Statistisk sentralbyrå.

Vi kan altså si at det som framstår som en tilsynelatende stabil utvikling, viser ofte en annen tendens når vi tar med en tredje variabel, slik som kjønn, alder eller bosted.

Og omvendt: Forskjeller blir ofte borte eller endret når vi tar andre variabler med i analysen. Vi har tidligere sett hvordan summariske (enkle) rater kan være

bedrageriske og skjule store forskjeller, for eksempel mellom land eller mellom menn og kvinner, på grunn av forskjeller i alderstrukturen. Her skal vi se på noen flere eksempler.

Simpsons paradoks

Berkeley universitet i California ble i 1973 beskyldt for å diskriminere kvinnelige søkere ved opptak, fordi en langt større andel menn enn kvinner ble opptatt som studenter.

La oss for enkelthets skyld si at dette universitetet har 700 mannlige søkere og 500 kvinnelige. Av mennene blir 70 prosent tatt opp, mens bare 56 prosent av kvinnene. Altså en langt større andel menn (se tabell 1).

Tabell 1. **Eksempel: søkere og opptak ved et universitet**

	Søkere	Opptak	Andel opptatt
I alt	1 200	770	64
Menn	700	490	70
Kvinner	500	280	56
Jus			
I alt	400	110	28
Menn	100	10	10
Kvinner	300	100	33
Sosiologi			
I alt	800	660	83
Menn	600	480	80
Kvinner	200	180	90

La oss også for enkelthets skyld anta at universitetet bare har to institutter: sosiologisk og juridisk institutt. Kvinnene søker seg i større grad enn menn til juridisk fakultet, som har færre studie plasser og en lavere andel opptatte studenter. Når vi derfor beregner opptaket for menn og kvinner separat for de to studieretningene, finner vi at kvinner oftere enn menn blir opptatt ved begge fakultetene.

Dette er et eksempel på et mer generelt fenomen, som går under navnet «Simpsons paradoks», og som innebærer at en



sammenheng eller forskjell ofte forsvinner eller reverseres når flere grupper slås sammen (aggregeres). Eller motsatt: at en forskjell først kommer til syne når vi studerer ulike undergrupper.

Menn fullfører oftere?

Statistikk over fullførte utdanninger i videregående opplæring og annen videregående utdanning viser for skoleåret 2008/09 at det var en litt større andel menn enn kvinner som fullførte slik utdanning (se tabell 2).

Tabell 2. Andel fullførte videregående utdanninger. 2008/09. Prosent

I alt	80,0
Menn	80,6
Kvinner	79,5

Kilde: Videregående opplæring, Statistisk sentralbyrå.

Tallene er overraskende: Annen statistikk viser at kvinner har langt mindre frafall underveis i videregående opplæring enn menn, og vi vet også at kvinner gjennomgående får bedre karakterer. Hvordan henger dette sammen?

Tallene i tabell 2 omfatter to typer utdanning: videregående opplæring og fag- og svenneprøver. Hvis vi beregner fullføringsandeler separat for disse to hovedtypene, blir bildet et helt annet. I videregående opplæring er det klart flere kvinner som fullfører, mens det for fag- og svenneprøver ikke er noen forskjell (se tabell 3). Bildet blir altså det motsatte av førsteinntrykket. Kvinner fullfører oftere enn menn.

Tabell 3. Andel fullførte videregående utdanninger, etter type utdanning. 2008/09. Prosent

	Videregående opplæring	Fag-/svenneprøver	I alt
I alt	74,9	92,6	80,0
Menn	72	92,7	80,6
Kvinner	76,8	92,4	79,5

Kilde: Videregående opplæring, Statistisk sentralbyrå.

Forklaringen på dette tilsynelatende paradokset ligger i to forhold: nemlig at fag- og svenneprøver som er dominert

av menn, har en langt høyere fullføringsandel enn videregående opplæring som er dominert av kvinner. Samtidig er det langt flere elever som avslutter videregående opplæring, enn som tar fag- og svenneprøver. Totaltallene for menn og kvinner i tabellen ovenfor kan betraktes som veide gjennomsnitt for de to typene utdanninger. For mennenes del veier da avlagte fag- og svenneprøver (med langt høyere fullføringsgrad) tyngst, mens det er omvendt for kvinnene. Langt de fleste kvinner tar videregående opplæring, og det er færre som fullfører.

Når vi på denne måten «kontrollerer for» eller tar hensyn til type utdanning, får vi altså et helt annet bilde av kjønnsforskjellene enn om vi bare ser på forskjellene mellom menn og kvinner totalt. Igjen et eksempel på at det å gå bak tallene kan være både nødvendig og fruktbart.

Alle bedre enn gjennomsnittet?

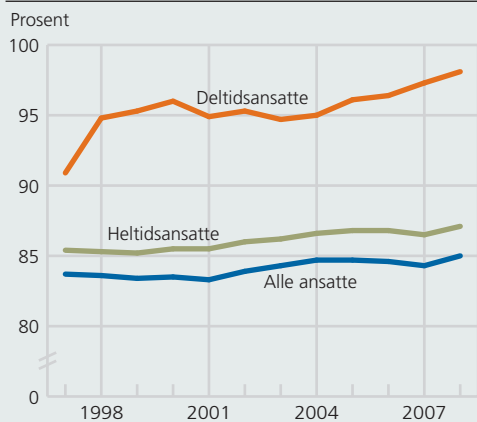
Det fortelles mange vitser om statistikkens gjennomsnittsberegninger, for eksempel at alle menn tror at de er en bedre sjåfør enn gjennomsnittet! Underforstått: Dette er selvfølgelig ikke mulig. Når noen er bedre enn gjennomsnittet, må også noen være dårligere.

Men noen ganger kan det tilsynelatende virke som om alle kan være bedre enn gjennomsnittet. Et eksempel kan være når det gjelder spørsmålet om likelønn. For å belyse dette beregner man gjerne kvinners månedslønn som en prosentandel av menns lønn. Om man regner om deltidsansattes lønn til det den ville vært hvis de arbeidet heltid (heltidsekvivalenter), så utgjør kvinners lønn i dag omtrent 85 prosent av menns lønn. 85 prosent er dermed på en måte et gjennomsnitt for alle kvinner (og menn) med ulik arbeidstid og i ulike næringer, yrker og stillinger (se figur 5).

Men om vi så ser på kvinners månedslønn som andel av menns for heltids- og deltidsansatte hver for seg, finner vi at *både* heltidsansatte og særlig deltidsansatte kommer bedre ut enn gjennom-



Figur 5. **Kvinners månedslønn som andel av menns lønn blant alle ansatte, heltids- og deltidsansatte. 1997-2008. Prosent**



Kilde: Lønnsstatistikk alle ansatte, Statistisk sentralbyrå.

snittet, idet deres lønn utgjør henholdsvis 87 og 98 prosent av heltids- og deltidsansatte menns lønn. Altså: Alle kvinner – både heltids- og deltidsansatte – kommer bedre ut enn gjennomsnittet! Hvordan er dette mulig?

Ulike grupper, ulik arbeidstid

Svaret er at prosenten for alle kvinner (85 prosent) *ikke* er et gjennomsnitt av prosenten for deltidsansatte og heltidsansatte, men et gjennomsnitt for alle ansatte. De to gruppene – heltidsansatte og deltidsansatte – har imidlertid svært ulik størrelse og sammensetning. Blant de heltidsansatte er flertallet menn, mens kvinner utgjør langt de fleste deltidsansatte. Siden deltidsansatte tjener mindre enn heltidsansatte, også når vi regner om til heltidsekvivalenter, vil deltidsansatte kvinners lønn veie mye tyngre enn deltidsansatte menns når vi beregner gjennomsnittet for alle.

Sagt på en annen måte: Når vi kontrollerer for (tar hensyn til) arbeidstid, blir noe av lønnsforskjellene borte.

Ceteris paribus?

Dette latinske uttrykket kan oversettes med «under ellers like forhold» og er på en måte uttrykk for statistikkens forsøk på å gjøre tallene så sammenliknbare som mulig, som når vi kontrollerer for al-

derstrukturen ved beregning av dødelighetsrater for menn og kvinner. På samme måte justerer vi for befolkningsstørrelse når vi sammenlikner biltettheten eller kriminaliteten over tid eller i to land.

Hvis vi igjen bruker lønnsforskjellene mellom kvinner og menn som eksempel, er kvinners månedslønn i gjennomsnitt om lag 85 prosent av menns. Som vi så ovenfor, ble forskjellene mindre når vi så på heltids- og deltidsansatte hver for seg (kontrollerer for arbeidstid).

På samme måte vil forskjellen bli mindre om vi tar hensyn til andre forhold, som utdanningslengde, at kvinner og menn er ansatt i ulike stillinger, yrker og bransjer/næringer og den ulike fordelingen på privat og offentlig sektor. Om vi tar hensyn til *alle* disse faktorene, vil sannsynligvis lønnsgapet reduseres noe.

I Sverige publiseres statistikk som viser kvinners månedslønn som prosent av menns når man standardiserer for alder, utdanning, arbeidstid, sektor og yrke, og prosentandelen øker da fra 84 til 92 (Statistiska centralbyråen: *På tal om kvinnor och män 2008*).

Et annet spørsmål er i hvilken grad det er naturlig å ta hensyn til alle disse ulikhetene mellom kvinners og menns situasjon. Ved å kontrollere for disse faktorene fjerner vi også effekten av dem, uten at disse faktorenes betydning for den enkelte forsvinner.

