

# Hvad siger tallene?

Undersøgelser og tal fylder godt i de daglige nyheder. Men der er god grund til at kigge kritisk på de taltunge nyheder, for ofte er de skæmmet af fejlslutninger. Inge Henningsen guider her læseren gennem nogle af de mest almindelige fejl, man kan støde på.



## Forfatter



Inge Biehl Henningsen er lektor emeritus Matematisk Institut Københavns Universitet inge@math.ku.dk

“En ny undersøgelse viser, at...” Man skal ikke læse ret langt ind i dagens avis eller mange minutter ind i et nyhedsindslag på TV, før man bumper ind i en sætning i denne stil. Og så bliver man typisk bombarderet med tal, der skal dokumentere de påstande, der fremføres i historien. Tal giver nemlig historier en aura af præcision og vederhæftighed. Underkaster man disse taltunge historier et nærmere eftersyn, kniber det dog ofte med præcisionen. Der er nemlig mange faldgruber, når man beskæftiger sig med tal og statistik, og det kan være svært for menigmand – og journalister – at gennemskue.

I denne artikel vil jeg præsentere eksempler på mere eller mindre hensigtsmæssig talbrug i medierne, som jeg har samlet gennem tiden. At min primære kilde er *Politiken* er selvfølgelig ikke repræsentativt, men afspejler primært mine læsepræferencer. Jeg er dog overbevist om, at der kan findes mindst lige så mange eksempler på tvivlsom talbrug i de øvrige danske medier.

## Årsag og sammenhæng

I kvantitative analyser støder man ofte på den type fejl, hvor man forveksler sammenhæng med kausalitet og tror at finde en betydningsfuld årsagssammenhæng mellem to variable, mens sammenhængen i virkeligheden skyldes, at de begge påvirkes af den samme baggrundsvariabel. For at illustrere hvad der kan ske, når man “glemmer” at tage højde for sådanne komplicerende sammenhænge, vil man fx kunne se en interessant sammenhæng mellem indkomst og skostørrelse i den danske befolkning. Men det skyldes naturligvis, at mænd tjener ca. 20 % mere end kvinder, og mænd har også i gennemsnit større fødder. Derfor følges indtægt og sko-nummer til en vis grad ad.

Oplever man også den slags falske sammenhænge i nyhederne? Ja! Fx foreslog Ulla Tørnæs som undervisningsminister, at matematiske elever i gymnasiet ikke måtte vælge et nyt sprog som andet fremmedsprog, men skulle fortsætte med det, de havde haft i grundskolen.



Foto: Colourbox

Begrundelsen hentede hun i en rapport, der viste, at hvis man opgjorde studentereksamensgennemsnit i forhold til elevernes valg af andet fremmedsprog, så klarede elever med fortsættersprog sig bedre.

»Nye tal, der netop er offentliggjort af Undervisningsministeriet viser, at matematiske elever med begynderprog får de dårligste karakterer.«  
(Ulla Tørnæs, *Politiken* 11.5.2003)

Troede ministeren for alvor, at hvis elever fik fortsættersprog i stedet for begynderprog, så fik de automatisk en bedre studentereksamen? Det er svært at forestille sig, at Ulla Tørnæs (eller i hvert fald hendes embedsmænd) ikke har haft øje for, at valg af fremmedsprog er en kompliceret proces, hvor mange faktorer herunder elevens generelle dygtighed spiller ind, og hvor der måske kan være en tendens til, at svagere elever vælger begynderprog frem for fortsættersprog. Men hvis det er tilfældet, ændrer det jo ikke på noget at gennemtvinge et anderledes sproglig valg, for det er ikke

sprogvalget i sig selv, der er årsag til eventuelle dårlige karakterer. Man kan i øvrigt spekulere på, hvad meningen er med at lave opgørelsen, men forfatterne af rapporten var for så vidt enige med ministeren i fortolkningen, idet de skrev

»Men man må dog sige, at tallene er vand på møllen hos dem, som mener, at matematiske elever bør have fortsættersprog frem for begynderprog.«  
(Undervisningsministeriets pressefunktion 4. april 2003)

Hvis ikke denne påstand er direkte misvisende, så er den i hvert fald let at misforstå i retning af, at der er en direkte sammenhæng mellem valg af begynderprog og dårlige karakterer. Hvordan kan resultaterne ellers være »vand på møllen« for dem, der vil afskaffe begynderprog?

Et aktuelt eksempel, der er udtryk for samme tankegang, kan man finde omkring de facto afskaffelsen af førtidspension for unge under 40. Her er argumentet, at udelukkelsen fra arbejdsmarkedet fører til social eksklusion. Undersøgelser viser imidlertid, at det er handicap og psykiske vanskeligheder, der fører både til udelukkelse fra arbejdsmarkedet og social eksklusion, idet manglende arbejdsmarkedsdeltagelse ikke i sig selv påvirker den sociale eksklusion nævneværdigt.

### Årsag og virkning

Ovenfor så vi eksempler på sammenhænge, der kunne forklares ved tilstedeværelsen af en tredje faktor, der påvirkede de to man var i gang med at undersøge. Men nogle gange bliver årsag og virkning vendt helt om. De følgende citater er fra en undersøgelse af voksnes læsning fra 2001 af Torben Pilegaard Jensen og kolleger, hvor man bl.a. så på uddannelse og læsefærdigheder. Her står

- »... dårlige læsefærdigheder øger risikoen for at have oplevet ordblindhed.«
- »... det at være ufaglært er den mest afgørende faktor for at læse under minimumskravene.«

Er det virkelig den vej kausaliteten går? Prøv at sammenligne med følgende udsagn (som forfatteren selv står for)

- "en sygeplejeuddannelse øger chancen for at man oplever at være kvinde."
- "det at være fotomodel er den mest afgørende faktor for at se godt ud"

### Manglende repræsentativitet

I næsten alle statistiske undersøgelser prøver man at slutte fra egenskaber hos et mindre udvalg til en større population. Tænk fx på valgprognoser eller medicinske undersøgelser. I denne proces er *repræsentativitet*, det at det lille udvalg ligner den store population på væsentlige punkter, helt afgørende.

Manglende repræsentativitet kan ytre sig på mange måder. Nogle gange er der tale om manipulation, når man bevidst kasserer prøver og erstatter dem med andre for at få et ønsket resultat. Det kunne *Politiken* rapportere om i 2010, hvor den øverste chef på en boreplatform presser en ansat til at tage flere prøver og kun indrapportere dem med det laveste olieindhold til Miljøstyrelsen.

Et andet problem er, når udtagelse og/eller bortfald af prøver er knyttet sammen med det spørgsmål, man er ved at undersøge som i følgende uddrag fra *Politiken* den 11/4-1999:

#### Mistænkte æg havde salmonella

»Hvert tiende æg, der forlader de danske hønsebesætninger, er inficeret med salmonella. [...] Ud af de 50 besætninger, som Veterinær- og fødevarerdirektoratet gennem hele vinteren har haft mistanke til – uden at gribe ind – har myndighederne fået svar fra 18 af besætningerne. Heraf har ti besætninger fået konstateret salmonella, mens otte har klaret frisag. Holder tallene på landsplan, svarer det til, at mere end ti procent af de dansk-producerede æg er inficeret med salmonella, «

Mon ikke avisen kan være temmelig sikker på, at tallene ikke "holder på landsplan"? Er det et repræsentativt udvalg af besætninger, man har fået svar fra? Næppe. De oprindelige 50 besætninger må formodes at have en overgennemsnitlig chance for at levere inficerede æg, da de var udvalgt på grund af mistanke om salmonella. På den anden side må man formode, at de 18 besætninger, der faktisk har svaret, nok har troet, at det gik godt med salmonellaen. Hvordan man omsætter dette til en påstand om, at »Hvert tiende æg, der forlader de danske hønsebesætninger, er inficeret med salmonella« er næsten ubegribeligt. Det er ikke nogen undskyldning at gardere sig ved at skrive »Holder tallene på landsplan«. Man må have lov til at formode, at når en avis vælger at bringe en nyhed, så tror de selv på den.

#### Skævt fra starten

En anden side af repræsentativitetsproblemet er visse tværsnitsundersøgelser, hvor folk ikke har lige stor chance for komme med. Det kan fx illustreres af hospitalsindlæggelser. Hvis man vil undersøge, hvorfor folk kommer på hospitalet, nytter det ikke noget at tage alle, der er indlagt på en bestemt dag og se på deres diagnoser. Nogle sygdomme fører til længere indlæggelser end andre og disse patienter vil have en større sandsynlighed for at være indlagt på tællingsdagen. Så billedet bliver på den måde skævt, idet sygdomme med lang indlæggelsestid bliver overrepræsenteret.

Et andet eksempel kan findes i *Ugebrevet A4* fra den 10. august 2009, hvor man i en artikel: *Socialt udsatte fylder fængslerne*, konstaterede, at:

»Hver anden af de indsatte er blevet alvorligt svigtet som barn af sine forældre«.

Undersøgelsen var baseret på 3524 personer, der var indsat i april 2009. Men denne undersøgelse lider af den samme svaghed som hospitalsundersøgelsen ovenfor. Personer med lange domme har større chance for at være med i undersøgelsen. Samtidig ved man, at der er større chance for, at socialt udsatte begår lovovertrædelser, der fører til lange fængselsstraffe. Materialet er altså ikke repræsentativt i forhold til dem, der bliver dømt.

Man kan også møde forfattere, der har fået de statistiske regler galt i halsen og tror, at store datamængder opvejer manglende repræsentativitet. Det ser man fx i følgende citat, der stammer fra en undersøgelse af forsømmelser i gymnasiet fra år 2000.

»Den manglende tilfældige udvælgelse af klasser opvejes rigeligt af den store mængde modtagne spørgeskemaer.«

Som en studerende sagde, da jeg brugte ovenstående citatet i undervisningen:

»Det er rart for en fysiker at vide, at det ikke gør noget, at man måler forkert, bare man gør det gange nok.«

#### Undersøg et og udtal dig om noget andet

I sin fremragende bog *Innumeracy* har matematikeren John Allen Paulos følgende eksempel:

»Professor Dumholz tager tilfældigt en loppe fra en beholder på bordet, river bagbenene af den og kommanderer "spring". Der sker ingenting. Efter at have gentaget forsøget en lang række gange og anvendt de relevante statistiske metoder konkluderer han med overbevisning, at lopper hører med bagbenene.«

Et virkeligt eksempel på, hvordan man undersøge et og udtaler sig om noget andet, kan man finde i en meget omtalt undersøgelse fra Statens Institut for Folkehelse om prostitueredes sundhed fra 2013. I rapporten konkluderes bl.a., at prostituerede har et bedre psykisk arbejdsmiljø end danske kvinder generelt.

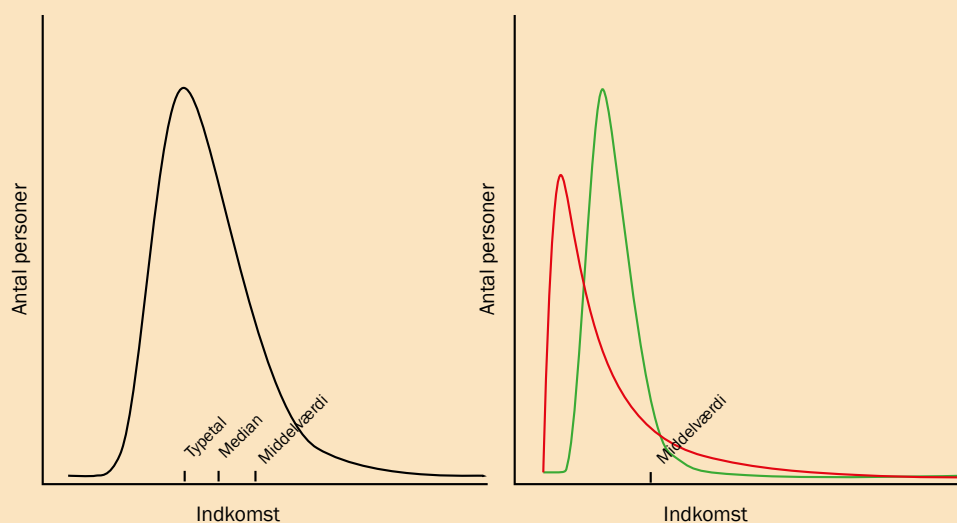
I artiklen har man brugt spørgsmålene fra en tilbagevendende landsdækkende undersøgelse af befolkningens sundhed og sygelighed (SUSY-undersøgelsen). Denne indeholder bl.a. en række spørgsmål om arbejdsmiljø, hvor det psykiske arbejdsmiljø er undersøgt ved to spørgsmål: »Om personen altid eller ofte ikke når sine arbejdsopgaver« og »Om man sjældent eller aldrig har indflydelse på, hvad man laver på sit arbejde.«

Disse to generelle spørgsmål adresserer helt oplagt ikke de særlige forhold, prostituerede arbejder under, og giver derfor heller ikke et retvisende

## Typetal, medianer og gennemsnit

Figuren til venstre viser en typisk indkomstfordeling og nogle størrelser, man kunne bruge til at karakterisere indkomstniveauet. På figuren er den typiske indkomst mindre end medianindkomsten, som igen er mindre end gennemsnitsindkomsten. Det gælder erfaringsmæssigt for de allerfleste indkomstfordelinger, og jo mere ulige indkomsterne er fordelt, des større forskel er der.

Figuren til højre viser to indkomstfordelinger med samme middelværdi.



billede af prostitueredes psykiske arbejdsmiljø. Det burde være indlysende, at man i en epidemiologisk undersøgelse af specielle grupper sikrer sig, at de indikatorer/markører, man bruger, har relevans for de pågældende gruppers situation. Det er ikke nogen undskyldning, at man i forvejen har målinger på noget andet (data fra SUSY-undersøgelsen). Det bliver let som Storm P's berømte tegning, hvor en mand leder efter sine nøgler under gadelygten »fordi der er så mørkt, der hvor han har tabt dem.«

### Tal taler ikke altid for sig selv

Den daglige avis er fuld af tal, der er med til at give nyhedsformidlingen et præg af præcision og vederhæftighed. Men tal er i virkeligheden kun informative, hvis man ved hvad de handler om. Som et eksempel kan man se på nedenstående uddrag af artiklen *Provinspiger vælger børn frem for abort* bragt i *Politiken* i maj 1998, der umiddelbart chokerede den på Frederiksberg bosiddende forfatter og flere af hendes veninder, da de læste om de næsten 4.000 aborter blandt unge kvinder i kommunen:

#### »Provinspiger vælger børn frem for abort

Unge piger i provinsen vælger oftere end deres medsøstre i hovedstadsområdet at gennemføre svangerskabet, hvis de bliver gravide. [...]

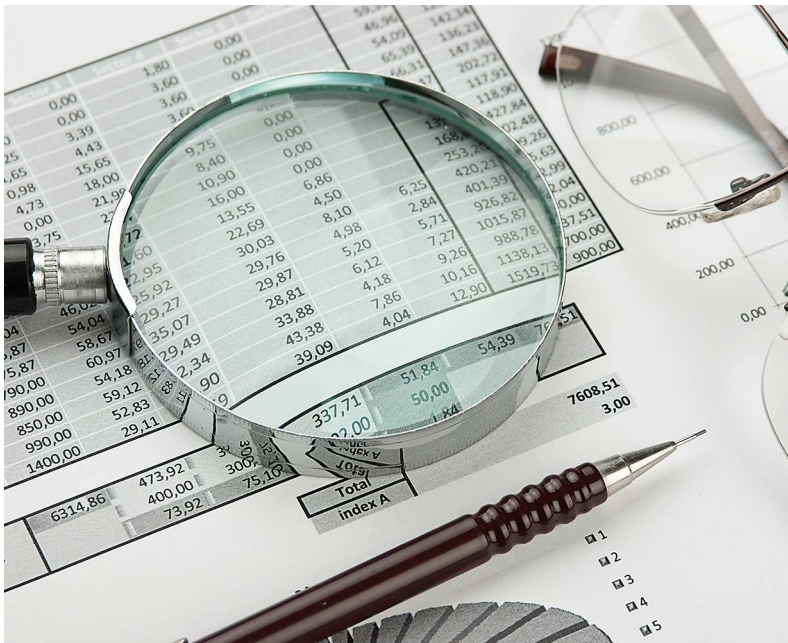
»Frederiksberg Kommune tegner sig for den største forskel. Her fik de 15-19-årige i 1996 3.769 aborter for hver 1.000 levendefødte børn. Københavns Kommune, Frederiksberg og Roskilde amter har alle over dobbelt så mange aborter som fødsler blandt teenagere. Vejle amt ligger i bund med 1.191 aborter for hver 1.000 børn, født af 15-19-årige mødre. [...]

Nu siger artiklen faktisk ikke noget om aborttallene blandt unge på Frederiksberg. Kun om det relative antal aborter i forhold til fødsler. Angivelsen "3.769 pr 1.000" ser imponerende nøjagtig ud – og aborttallet forekommer chokerende højt. Man kan imidlertid stille spørgsmålstegn ved om sammenligningen – aborter i forhold til fødsler – er særlig relevant for den almindelige avislæser og om den valgte basisværdi, "pr. 1000 levendefødte", er hensigtsmæssig. Et opslag i Statistisk Årbog viser, at der i 1996 var 13 fødsler og 49 aborter blandt de 15-19 årige på Frederiksberg. Selvfølgelig svarer "49 aborter og 13 fødsler" til "3.769 aborter og 1.000 fødsler". Men det budskab, der sætter sig i hukommelsen hos en person, som bare sidder og læser sin morgenavis, er meget forskellige afhængigt af hvilke forholdstal, man bruger. Artiklen er formentlig en direkte afskrift af en rapport fra Sundhedsstyrelsen eller en tilsvarende organisation. Om tallene er rigtige. Sikkert ja. Men er de informative. Helt sikkert nej.

### Typetal, medianer og gennemsnit

Når man skal beskrive en fordeling, fx indkomsten i et land, vil man gerne kunne nøjes med et enkelt tal, som man kunne kalde *middelindkomsten*. Men det viser sig i praksis, at sådan en middelindkomst ikke er noget entydigt begreb. Man kan vælge at angive den *typiske* indkomst, altså den indkomst som flest mennesker har (i matematikundervisningen kaldet *typetallet*). Man kan vælge at angive den *midterste* indkomst (*medianen*). Og endelig kan man vælge at angive *gennemsnitsindkomsten*. Men da de allerfleste indkomstfordelinger er karakteriseret ved en lang hale bestående af meget rige mennesker, så giver disse tre valg nok så forskellige resultater (se figur).





#### Videre læsning og kilder

Anders Arnfred, Pia Vivian Pedersen, Maria Holst Algren og Knud Juel (2013): Sundhed, sygelighed og trivsel blandt klinikprostituerede. Statens Institut for Folkesundhed (SIF), Syddansk Universitet

Jensen, Torben Pilegaard, Andersen, Angelo og Tue Halgreen (2001): Læsefærdigheder og deltagelse i samfundslivet. AKF Forlaget og SID, København.

Paulos, John Allen (1988): Innumeracy: Mathematical Illiteracy and its consequences, Hill and Wrang, New York.

Sinding et al (2000): "Rapport vedrørende elevforsømmelser i gymnasiet"

Derfor bruger folk typisk det tal, der passer dem bedst. Vil man vise velstand vælges gennemsnitsindkomsten. Vil man vise fattigdom vælges den typiske indkomst.

Da de tre begreber – *typetal*, *median* og *gennemsnit* – bruges i flæng i den offentlige debat, er der rige muligheder for misforståelser og manipulation. At journalister heller ikke altid har styr på begreberne, viser følgende citat fra *Politiken* 26/6-2000:

»Københavnere står i kø til ældreboliger I øjeblikket er der godt en måneds ventetid til en lille stue på et plejehjem, og det er den såkaldte medianventetid, som beskriver det almindelige forløb.

Bruges almindelig matematik, er den gennemsnitlige ventetid til plejehjem længere – omkring tre-fire måneder, fordi der er mennesker, der må vente endog meget længe, hvis de ønsker et helt bestemt plejehjem. Det trækker gennemsnittet op.«

I første afsnit forveksles den typiske ventetid og medianventetiden. I næste afsnit bruger journalisten ifølge artiklens eget udsagn "almindelig matematik" og beregner gennemsnitsventetiden. Den er i dette tilfælde større end både den typiske ventetid og medianventetiden, fordi der er nogen, der venter meget længe. Hvad der er den mest relevante oplysning kommer helt an på, fra hvis synsvinkel man ser sagen. Begge tal siger noget om ventetiden, men de måler ikke det samme, og forskellen på typetal og gennemsnit er i hvert fald bestemt af fordelingen og ikke af, om man bruger "almindelig matematik" eller ej.

#### Træn din talforståelse...

Med den omfattende brug af undersøgelser som baggrund for nyheder i medierne, er der god grund til at udvikle sin talforståelse og kritiske sans. Det er jo ikke kun journalister og menigmand, der kan gå galt i byen i brugen af tal – også fagfolkene bag undersøgelserne kan have fået dem galt i halsen. Og udover misforståelserne er der altid en fare for, at tallene bevidst fremstilles på en måde, så de er svære at gennemskue.

At se opmærksomt på tal er en træningssag og ikke noget, der bare kommer af sig selv.

Skal vi så ikke alle sammen til at lære mere statistik? Nej, det synes jeg egentlig ikke. At producere og formidle retvisende kvantitative opgørelser må først og fremmest påhvile de professionelle. Det er dem, der skal sørge for, at de ikke måler et og konkluderer om noget andet, at de bruger tal, der taler og ikke vrøvler, og at de ikke forveksler årsag og virkning eller finder sammenhænge, hvor der ingen er. Det kan ikke være den almindelige avislæser, der skal sidde og analysere selv banale talangivelser.

Men indtil vi oplever et sådant kvalitetsløft, kan man måske bruge artiklens eksempler som inspiration til at kaste et kritisk blik på de daglige nyheder. ■

## Kig nyhederne i kortene

Som forbruger af nyhedsmedier er det nyttigt at kende de relativt simple "greb", som man kan bruge til at screene for de mest almindelige fejlslutninger, man ser i kvantitative undersøgelser. Sådanne fejlslutninger kan fx skyldes, at man:

- sammenligner usammenlignelige grupper
- bytter om på årsag og virkning
- glemmer at tage hensyn til inhomogeniteter
- vælger en atypisk undersøgelsesgruppe, men generaliserer resultaterne.
- måler en ting og konkluderer om noget andet.

Når en undersøgelse er refereret i medierne, kan man som udgangspunkt starte med at spørge sig selv:

- Hvor mange personer har de spurgt?
- Hvordan er de udvalgt?
- Er de repræsentative?
- Hvad har man egentlig målt?
- Hvor stor usikkerhed er der på resultatet?

Det er ikke altid, at man kan læse sig til svaret – men så er der jo blot ekstra grund til at være kritisk!