

Karl Henrik Bentzon  
stud.mag.

## En kildekritisk analyse af den danske valgstatistik

### 1. Indledning

Den danske valgstatistik er en særlig gruppe inden for den officielle statistik, der udgøres af en række kvantitative data, der publiceres på det offentliges foranstaltning igennem Danmarks Statistik. Valgstatistikken har således en række specifikke egenskaber, som man bør holde sig for øje ved dens anvendelse som historisk kilde. Det er derfor min hensigt at opstille en række analysekategorier, der dels gør det muligt at vurdere valgstatistikken som kilde i sig selv, dels gør det muligt at placere denne i forhold til andre historiske kilder.

### 2. Kildeplacering

Den kildeanalyse man finder i fremstillinger om historisk metode beskæftiger sig næsten udelukkende med de såkaldte sproglige kilder. Det må efterhånden føles som et betydeligt savn, fordi anvendelsen af historisk statistik også inden for den mere traditionelle historieforskning spiller en stadig større rolle. Problemet forekommer mig heller ikke at være særligt stort, fordi de almindeligt anvendte kildedistinktioner uden større vanskelighed kan overføres på statistiske data.

#### 2.1. Ophav

Ottar Dahl <sup>1)</sup> skelner i spørgsmålet om kildens ophavsmand mellem personlige og institutionelle kilder. Valgstatistikken tilhører så udpræget den sidste gruppe.

---

1) Ottar Dahl: Grunntrekk i historieforskningens metode-  
lære p. 44.

Karl Henrik Bentzon  
stud.mag.

## En kildekritisk analyse af den danske valgstatistik

### 1. Indledning

Den danske valgstatistik er en særlig gruppe inden for den officielle statistik, der udgøres af en række kvantitative data, der publiceres på det offentliges foranstaltning igennem Danmarks Statistik. Valgstatistikken har således en række specifikke egenskaber, som man bør holde sig for øje ved dens anvendelse som historisk kilde. Det er derfor min hensigt at opstille en række analysekategorier, der dels gør det muligt at vurdere valgstatistikken som kilde i sig selv, dels gør det muligt at placere denne i forhold til andre historiske kilder.

### 2. Kildeplacering

Den kildeanalyse man finder i fremstillinger om historisk metode beskæftiger sig næsten udelukkende med de såkaldte sproglige kilder. Det må efterhånden føles som et betydeligt savn, fordi anvendelsen af historisk statistik også inden for den mere traditionelle historieforskning spiller en stadig større rolle. Problemet forekommer mig heller ikke at være særligt stort, fordi de almindeligt anvendte kildedistinktioner uden større vanskelighed kan overføres på statistiske data.

#### 2.1. Ophav

Ottar Dahl <sup>1)</sup> skelner i spørgsmålet om kildens ophavsmand mellem personlige og institutionelle kilder. Valgstatistikken tilhører så udpræget den sidste gruppe.

---

1) Ottar Dahl: Grunntrekk i historieforskningens metode-  
lære p. 44.

En kildes officielle karakter virker ikke altid befordrende på dens vidneværdi, men hvad angår f.eks. valgstatistikken er dette ikke noget problem - i hvert fald ikke i de vestlige industrilande. Tværtimod opnår man ved dens officielle karakter, at der stadig allokeres ressourcer til dens fremstilling, samt at den produceres af et kompetent organ, der dels sikrer en høj grad af nøjagtighed dels ensartethed over tid, der muliggør tidsrækkeanalyser, samt ikke mindst en betryggende procedure for indsamling af primærmaterialet. Det sidste forhold turde være af stor betydning. Inden for andre dele af den officielle statistik er disse fordele ikke lige iøjnefaldende: f.eks. har erhvervsstatistikken ændret sig meget gennem tidens løb, der vanskeliggør sammenligninger over tid. I det hele taget vil der næsten altid være problemer, når en procesgenereret kilde skal anvendes til et andet formål, i dette tilfælde forskning. For valgstatistikken er især dataaggregationen en hæmsko, et forhold, som vil blive belyst nedenfor.

## 2.2. Modtager

Når man spørger, hvilken person eller gruppe en kilde er tiltænkt er det betydningsfuldt at skelne mellem konfidentielle og offentlige kilder, idet vi herigennem kan sige noget væsentligt om ophavssituationen. Spørgsmålet om den tiltænkte modtager er med andre ord et normativt element ved kilden eller for at bruge Erslevs udtryk, en anvendelse af kilden som levning.

Det kan i denne forbindelse forekomme helt trivielt at fremhæve valgstatistikken som en offentlig kilde, fordi vi jo alene interesserer os for dens kognitive elementer, dvs. dens oplysninger om faktiske forhold, men det er værd at se offentlighedsprincippet som en del af den omhyggelige kontrol, der ifølge hele vor demokratiske ideologi bør udøves over valghandlingen. Det er i det hele taget interessant at se væksten i den officielle statistik siden folketællingen i 1769 som udtryk for to udviklingstendenser i samfundet. På den ene side

den økonomiske udvikling, der dels har gjort det nødvendigt at fremskaffe en række statistiske-økonomiske data, dels har muliggjort finansieringen af indsamling og bearbejdelse af det statistiske materiale, og på den anden side den politiske og organisatoriske udvikling i samfundet, som kraftigt har udvidet efterspørgslen efter statistiske data <sup>1)</sup>. I USSR finder man på trods af dets økonomiske udvikling stadig kun en pauver offentlig statistik. En medvirkende årsag hertil er utvivlsomt - ligesom forholdet var i det 18-århundredes Danmark -, at det ikke er i statens interesse, at befolkningen får et for intimt kendskab til samfundets indretninger.

### 2.3. Autencitet

Når man taler om en kildes autencitet mener man det mere eller mindre direkte forhold kilden har til det sociale fænomen, den søger at beskrive. Man skelner i denne sammenhæng bl.a. mellem primær- og sekundærkilder, hvor endvidere primærkilden enten kan være en førstehånds- eller en andenhåndskilde.

socialt objekt	primærkilde førstehåndskilde	sekundærkilde
valg handlingen	valgbøger	valgindberetninger valgstatistik

Der er forskel på den måde, materialeindsamlingen foregår ved henholdsvis kommunal- og folketingsvalg. Ved folketingsvalg forestås stemmeoptællingen af valgbestyrelser ved de enkelte opstillingskredse. Resultatet indføres i såkaldte valgbøger, der sendes over til Indenrigsministeriet, der igen lader Danmarks Statistik sammentælle stemmer, beregne mandatfordelingen, m.v. samt sørge for of-

1) Samfundet og Statistikken, St.M. 4 r. 139 bd. 1. hft. p. 5.

fentliggørelse af hele valgmaterialiet.

Kommunalvalget er derimod et lokalt anliggende, hvor stemmesammentælling og mandatfordeling foretages af de kommunale myndigheder og offentliggøres af disse. Kommunerne har heller ikke pligt til at opbevare valgbøgerne fra kommunalvalget, men vil i de fleste tilfælde gøre det, især i større kommuner. Når der alligevel foreligger en fyldestgørende kommunalvalgsstatistik, skyldes det, at Danmarks Statistik udsender en valgindberetning til kommunerne, hvorpå de vigtigste resultater fra valget skal påføres. På grundlag af disse valgindberetninger udarbejdes den kommunale valgstatistik.

Det ses af skemaet, at valgstatistikken er en sekundær kilde i forhold til primærkilden, valgbøgerne.

Inden for de verbale historiske kilder spiller denne distinktion en stor rolle, medens den er upågtet inden for statistikken. Det hænger sammen med, at man normalt interesserer sig for skellet mellem primær- og sekundærkilder i det omfang, der sker en informationsforvridning i form af enten tilfældige fejl i overleveringen eller bevidst fordrejning.

I almindelighed sker der ikke informationsforvridning inden for statistikken. Man har med visse undtagelser lov til at opfatte en publiceret statistik som en korrekt reproduktion af de tilsvarende data i primærmaterialet. Derimod sker der ofte en informationsreduktion, forstået på den måde, at den officielle statistik indeholder færre data end det tilgrundlæggende materiale. Reduktionen kan have to former:

- 1) Kategorireduktion. I dette tilfælde sker der det, at visse variabler, som optræder i primærmaterialet simpelthen udelades i statistikken. I kommunalvalgsstatistikken er f.eks. for visse år vælgere og afgivne stemmer ikke splittet op på mænd og kvinder, selvom oplysningerne kan findes i valgbøgerne.

- 2) Enhedsreduktion. Det sker meget ofte, at man for at gøre statistikken mindre voluminøs aggregerer grundenhederne til større enheder. Hermed går der naturligvis en betydelig information tabt, som i visse tilfælde kan gøre statistikken stort set uegnet til forskningsbrug. Det gælder f.eks. de ældre dele af den norske kommunalvalgsstatistik, hvor de kommunale valgoplysninger er blevet aggregeret på 5 landsdele, dvs. en reduktion fra ca. 1000 kommuner til 5 enheder.

#### 2.4. Finalitet

Valgstatistik som al anden statistik vil man være tilbøjelig til at karakterisere som rent kognitiv, fordi den i det omfang der ikke er tale om talfusk udelukkende beskriver et socialt fænomen. Spørgsmålet om kildens normative og kognitive elementer går med andre ord udelukkende på kildens faktiske indhold. Det er imidlertid karakteristisk for den kritiske videnskab, at den også spørger efter hvad der ikke står i kilden eller mere præcist: hvorfor der ikke står det og det <sup>1)</sup>.

Ser vi på den danske statistik i et historisk perspektiv er det iøjnefaldende, at den har produceret de data, som de til enhver tid magthavende grupper har haft brug for. Sammenhængen er især klar for den ældre tids statistik, men selv om forholdet er mere uklart i dag, er det ikke ensbetydende med, at sammenhængen er forsvundet, men det kræver en større forskningsindsats at dokumentere den.

Det fremgår af de overleverede kilder, at en væsentlig årsag til afholdelsen af de første danske folketællinger var et ønske hos den enevældige administration om at få et bedre grundlag for soldater- og skatteudskrivning. <sup>2)</sup> Da man i 1830-erne fik offentlige statsregnskaber

---

1) Øyvind Østerud: Randbemærkninger om statsvitenskap som ideologi, fagkritisk paper

2) st.m. op.cit. 6f.

og en handelsstatistik samt efter 1850 en erhvervs- og valgstatistik, må det ses på baggrund af, at andre samfundsgrupper havde fået politisk magt, nemlig det liberale borgerskab. Først ind i 1890-erne fik man en socialstatistik, dvs. en række oplysninger, som arbejderklassen kunne anvende i kampen om bedre økonomiske og sociale forhold.

Ved på den måde at interessere sig for, hvad der ikke står i en kilde, ledes man hen på det forhold, at selv tilsyneladende harmløse kilder set i et videre perspektiv produceres og anvendes af bestemte samfundsinteresser. Vi beskæftiger os med andre ord med kildens finalitet, dens formål i samfundet, og dette må karakteriseres som et normativt element ved kilden.

## 2.5. Variabeltype

Når man taler om datatyper skelnes der ofte mellem kvantitative og kvalitative data. Det forekommer mig imidlertid, at der bliver sagt meget vrøvl om dette skel, bl.a. fordi det indgår i den ideologiske strid mellem såkaldt borgerlig og kritisk videnskab. Jeg vil derfor indledningsvis give udtryk for nogle almindelige synspunkter på denne distinktion, inden valgstatistikken som sådan indrages.

Lad os som et eksempel se på Aristoteles berømte klassifikation af de politiske styreformer. Først udvalgte han sig en række enheder: de autonome politiske systemer i den kendte verden. Dernæst bestemte han, hvilke egenskaber ved disse enheder, han ville undersøge, det såkaldte variabelbegreb (meget af dette og det følgende er hentet fra Mats Fribergs fortrinlige lille bog: Teorikonstruktion för sociologer, 1. begrebslära. Særlig p. 13 ff), som i dette tilfælde blev politiske styreformer. Endelig - og det er det centrale - gav han variabelbegrebet en række forskellige værdier: monarki, aristokrati, demokrati, m.v., som tillod ham at anbringe de oprindelige enheder i en klassifikation. Tankegangen kan generaliseres på følgende måde:

Enheden  $x_j$  har værdien  $v_j$  på variabelen  $X$ .

Variabeltypen er givet med de værdier, den kan antage. I Aristoteles tilfælde danner værdierne en klassifikation, hvorfor man gerne taler om en nominalskalavariabel. Den karakteriseres ind imellem også som en kvalitativ variabel, fordi der er et såkaldt kvalitativt spring mellem de klasser, der tilsammen udgør klassifikationen.

Hvis man kan rangordne værdierne, eks. lav, middel, høj, betegnes variabelen som en ordinalskalavariabel, og hvis man kan angive den nøjagtige forskel på værdierne, dvs. karakterisere dem ved hjælp af vort talsystem, har man en intervalskalavariabel. Den sidste variabeltype kaldes også en kvantitativ variabel, medens ordinalskalavariablen nærmest er en mellemting.

Forholdet mellem variabeltyperne er sådan, at man altid kan reducere en intervalskala til en ordinalskala og videre en ordinalskala til en nominalskala, medens man ikke kan gå den anden vej.

Men hvilken skalatype skal man benytte? Her er vi ved stridspunktet, som kan udtrykkes i to problemkomplekser: validitet og relevans. Jeg vil alene beskæftige mig med begrebsdannelsen, men forholdene er analoge blot mere komplicerede hvad angår hypotesedannelse og teori-konstruktion.

Den skalatype, man kan anvende, er bestemt af de data, der er til rådighed. Aristoteles data var et almindeligt kendskab - efter vor målestok - til en række lande og bystater, der tillod ham at opstille en temmelig grov klassifikation. Men er vi under alle omstændigheder bedre stillet idag, hvor det er muligt på grundlag af statistiske data og matematiske hjælpemidler at udarbejde kvantificerede indeks over repræsentativitet, regerings-cirkulation, valgdeltagelse, politicering, osv.? Det afhænger af, om vore operationelle definitioner er valide i forhold til vore teoretiske overvejelser, eller sagt på en anden måde: om vi måler det, vi faktisk ønsker at måle. I den forbindelse er det værd at lægge mærke til



en forskel på kvalitative og kvantitative variablers anvendelse på sociale fænomener. Kvalitative variabler er i kraft af deres udifferentierethed relativt nemme at fortolke ind i den sociale virkelighed og giver således ikke så store validitetsproblemer. Men netop deres manglende stringens gør det vanskeligt at arbejde med dem på det operationelle niveau. Kvantitative data lider derimod af den modsatte skavank. Deres betydelige differentiationsniveau gør det vanskeligt at fortolke ind i større sociale sammenhænge. Validitetsproblemet er med andre ord stort. Til gengæld egner de sig fortrinligt til videnskabelig bearbejdelse. Denne egenskab har fået forskere til at oversætte kvalitative data til kvantitative, f.eks. kvantitative indholdsanalyser, hvad der specielt har bragt de kvantitative metoder i miskredit, fordi forskeren i realiteten skaber et nyt sæt data, hvis relation til det oprindelige materiale kan være meget prekær. Men man kan ikke lægge den kvantitative metode til last, at den anvendes illegitimt. I almindelige metodefremstillinger er man også blevet klar over dette misbrug, og det er f.eks. Blalocks fortjeneste, at han ikke mener, det kan siges ofte nok, at intervallskaleringer kun må anvendes, hvis materialet tillader det (Blalock: social statistics, passim).

Forkærlighed for kvantitative data har imidlertid også haft en anden konsekvens, som måske er alvorligere end validitetsproblemet. Man kan inden for den amerikanske inspirerede politologi se en tendens til at koncentrere opmærksomheden på samfundsfænomener, som lader sig kvantificere. Norge er et godt eksempel herpå. Indtil godt op i 1960-erne var politologien der ganske domineret af kvantitative analyser af vælgeradfærd og de politiske partier. Det er først i de seneste år, at den er blevet påvirket andetstedsfra. I Danmark er situationen noget anderledes. På grundlag af den udmærkede litteraturfortegnelse i Scandinavian Political Studies over, hvad der publiceres af skandinaviske forskere om samfunds-

videnskabelige emner, kan det beregnes, at for perioden 1960-69 var kun ca. 10-15 % af publikationerne af dansk oprindelse. Procenten er let stigende i perioden, hvad der måske i et vist omfang kan ses som udtryk for, at den danske forskning på dette område er kommet senere i gang, end i de øvrige nordiske lande. En nærmere analyse af værkernes titler viser endvidere, at den danske samfundsforskning specielt er svag hvad angår metateori, forskningsteknik, idehistorie, samt de mere sociologisk orienterede samfundsanalyser, herunder kvantitative undersøgelser. Kritikken skal ikke forstås på den måde, at det der faktisk forskes i Danmark er ringe i sig selv, men at den samlede indsats er temmelig lille i forhold til vore nordiske nabolande, samt at forskningen stort set centrerer om traditionelle emner, dvs. verbale fremstillinger af vore vigtigste politiske institutioner og deres virke, i denne sammenhæng har de politiske partier en dominerende stilling. Den endnu noget begrænsede anvendelse af kvantitative data i dansk samfundsvidenskabelig forskning er med andre ord ikke udtryk for nogen bevidst forskningsholdning, men blot udtryk for, at man endnu ikke besidder kapacitet til at indsamle og bearbejde store datamængder. - Men det kommer utvivlsomt.

En valgstatistik er kvantitativ af den simple grund, at enhederne er aggregerede. I det hele taget gælder det, at en nominalskala bliver omdannet til en intervalskala, hvis enhederne aggregeres på den ene eller anden måde. I valgstatistikken er aggregationen givet på grund af den hemmelige afstemning, der medfører, at de afgivne stemmer ikke kan henføres til de enkelte stemmende, men alene til det område inden for hvilket afstemningen foregår. Disse områder vil være bestemt af de administrative forhold der råder. For Danmarks vedkommende er den mindste enhed for såvel kommunal- som folketingsvalg kommunen. Anvendelsen af aggregerede data

stiller politologien over for en række metodologiske problemer, som gerne sammenfattes under begrebet den økologiske fejlslutning <sup>1)</sup>. Et andet spørgsmål er, om enhedernes størrelse er hensigtsmæssig ud fra et forskningsmæssigt synspunkt. Det er ønskeligt med så små enheder som muligt, fordi man hermed alt andet lige opnår en større ensartethed inden for den givne enhed. Denne ensartethed er af betydning for økologiske analyser, dels fordi den målelige variation mellem enhederne i almindelighed bliver større, dels fordi det bliver muligt at foretage visse forsigtige slutninger m.h.t. sammenhængen på individniveau. I den forbindelse har forskningen nydt godt af en forældet administrativ inddeling, d.v.s. et stort antal små kommuner. Kommunalinddelingsreformen vil imidlertid alt andet lige medføre en reduktion i de tilgængelige oplysninger.

#### 2.6. Objekt

Når man arbejder med et historisk materiale bliver den almindelige problematik omkring subjekt-objektrelationen yderligere kompliceret, fordi objektet ikke lader sig direkte iagttage, men kun indirekte via de kilder, som den historiske virkelighed har efterladt sig.

Det er indlysende, at sociale fænomener i meget varieret omfang afsætter sig spor i kilderne, ligesom overleveringen er meget uensartet. I dette forhold gør der sig en masse tilfældige faktorer gældende, men det er dog muligt at fremhæve nogle systematiske træk af interesse i denne sammenhæng.

---

1) Se Redder, Siune, Tonsgård: Introduktion til sociologisk metode, p. 207-09. En mere teknisk gennemgang på grundlag af variansanalyser findes hos H.R. Alker: A typology of ecological Fallacies i Quantitative ecological analysis in the Social Sciences, Dogan & Rokkan.

Man kan således skelne mellem på den ene side fænomener, hvis registrering er en almindelig norm inden for et samfund på et givet tidspunkt, og på den anden side fænomener, hvor en sådan norm ikke gør sig gældende. Det er karakteristisk for moderne industrisamfund, at denne norm har bredt sig til stadig flere områder, udviklingen i den danske offentlige statistik er et eksempel herpå. Med edb-teknikken er det ydermere blevet muligt at opbevare og behandle disse enorme datamængder. Dette forhold har også sat sig tydelige spor inden for samfundsvidenskaberne.

Ligesom indsamlingen af data på en lang række områder foregår efter ganske bestemte retningslinier, er det også sådan, at den adfærd der registreres, meget ofte sker på grundlag af en række detaljerede forskrifter. Det er på den baggrund, at man i sociologien skelner mellem institutionaliseret og ikke-institutionaliseret adfærd. Valghandlingen er institutionaliseret til det yderste. I det danske valgsystem har vælgeren fire handlingsmuligheder:

1. han kan undlade at stemme
2. han kan stemme på en liste
3. han kan stemme på en person og dermed også på en liste
4. han kan stemme blankt

Det forhold, at den sociale adfærd på forhånd er begrænset til et mindre antal veldefinerede handlingsalternativer, har den vigtige konsekvens, at registreringen af adfærden kan blive fuldstændig, nøjagtig samt utvetydig. Vælgeradfærdsforskningen har ingen validitets- eller reliabilitetsproblemer, hvad angår selve valghandlingen, men lige så snart der inddrages andet materiale såsom psykologiske-, erhvervs-, økonomiske, urbaniseringsvariabler, m.v. kommer man naturligvis ud i de sædvanlige målingsproblemer, ligesom hvis man vil fortolke vælgeradfærden i et videre perspektiv såsom udtryk for mobilisering, politiske holdninger, m.v.

### 2.7. Definition

Når man nu, som jeg har forsøgt det, har analyseret et fænomen ud fra forskellige synsvinkler, så er det det samme som at sige, at man har klassificeret fænomenet på grundlag af en række inddelingsgrunde. Der er med andre ord tale om en flerdimensioneret klassifikation. Herefter kan man foretage en definition på grundlag af de vigtigste klassebegreber.

En valgstatistik defineres flg. således som en institutionel, offentlig og sekundær kilde, som er opbygget på grundlag af aggregerede kvantitative data og som beskriver institutionel, lovbunden adfærd.