

Iakttakelser og erfaringer vedrørende representative utvalg

Gjennom forskjellige epoker i vitenskapens utvikling, har det vært visse ideer som har dominert i en slik grad at de har overskygget andre ideer som også har vært svært viktige, men som ikke har blitt verdsatt, det være seg fordi de ikke har hatt nok gjennomslagskraft, eller fordi en har sluttet å gi dem oppmerksomhet.

En slik dominerende ide innenfor statistisk forskning er formulert i de store talls lov. Jeg vil på ingen måte fornekte denne loven, som man kan kalle grunnleggende innenfor statistikken, men det synes for meg som om den kanskje har vært for enerådende, og at en ikke har gitt den ønskede oppmerksomhet til den representative metodes viktighet for statistikken. I mange tilfeller er dette den eneste mulige framgangsmåten.

Disse generelle betraktningene kan anvendes innenfor mange områder i statistikken, men kanskje med aller størst rett innenfor de sosiale spørsmål. For å belyse disse kompliserte forholdene, er det nødvendig å samle inn svært detaljerte opplysninger, og anvende skjema med et så stort antall spørsmål at det umulige i å gjennomføre en slik undersøkelse for et helt land, viser seg fra første stund. For å få til en praktisk løsning knyttet til de sosiale spørsmål, blir det nødvendig å gi avkall på de strikte krav som de store talls lov synes å pålegge, og ta tilflukt til undersøkelser basert på representative utvalg.

Det er derfor klart at følgende spørsmål er av stor viktighet:

Hvordan skal det representative utvalg være for å kunne gi et så godt bilde som mulig av samfunnet som helhet?

Oversatt av
Anne Gro Hustoft

Den representative utvalgsmetoden ble for første gang presentert for et internasjonalt publikum på møtet til det Internasjonale statistiske institutt (ISI) i Bern i 1895. I det følgende presenteres Kiærs opprinnelige foredrag, som nå for første gang er blitt oversatt fra fransk til norsk.

Kiær viste selv til undersøkelser han hadde foretatt i 1890-årene ved hjelp av denne metoden. Men bruken av representative utvalg i norsk statistikk var eldre enn dette. Allerede sent på 1870-tallet og tidlig på 1880-tallet gjennomførte Jakob Neumann Mohn (1838-1882) flere slike undersøkelser. Tradisjonen ble fulgt opp av Mohns etterfølger, Boye Christian Riis Strøm, gjennom flere utvalgsundersøkelser i 1880-årene. Så vidt vi vet, er det ingen andre land som gjennomførte store utvalgsundersøkelser på et så tidlig tidspunkt. Grunnene til at Norge ble et pionerland i på dette området, er nok sammensatte. Men det er liten tvil om at den særegne kombinasjonen av et bredt ansvar for statistikkinnsamlingen og en velutviklet forskningsvirksomhet i tilknytning til denne – som har kjennetegnet Statistisk sentralbyrå i hele sin historie – er en viktig medvirkende årsak.

Historien om diskusjonene omkring den representative undersøkelsesmetoden har i ettertid tiltrukket seg stor interesse. Dels er dette fordi metoden i dag har en så viktig plass i statistikkproduksjon, forskning, meningsmålinger og markedsundersøkelser. Men historien er også av andre grunner en godbit for vitenskapshistorikere. Som det fremgår av diskusjonen som fulgte foredraget, ble metoden fra flere hold møtt med betydelig skepsis. Også ved senere møter i ISI - i 1897, 1899, 1901 og 1903 - ble metoden lansert av Kiær. Enkelte støttet tanken om at en slik metode burde inngå som et supplement til fullstendige tellinger. Men kritikken vedvarte. Og etter 1903 forsvant temaet helt fra agendaen. Først i 1925, etter Kiærs død, ble den representative undersøkelsesmetoden igjen drøftet i instituttets møter. På dette tidspunkt var motstanden liten. Vi har derfor å gjøre med en metode som ettertiden har oppfattet som selvfølgelig nyttig og utsagnskraftig, men som likevel led nederlag i møtet med en sentral vitenskapelig forsamling - for altså siden å gjenoppstå en generasjon senere.

Det er dette som gjør metodens forhistorie så spennende: Hvorfor ble metoden møtt med så stor skepsis i 1895, og hvorfor sluttet Kiær å fremme metoden etter 1903? Svaret på det første spørsmålet kan man i noen grad lese seg til i det følgende, og posisjonene i debatten er ytterligere drøftet i boken Faktisk talt, om norsk statistikkshistorie (av Einar Lie og Hege Roll-Hansen (2001), utgitt på Universitetsforlaget). Der går vi også grundig inn på spørsmålet om hvorfor Kiær sluttet å fremme metoden, og hvorfor den for en lang periode ikke ble brukt i store sosiale og økonomiske undersøkelser i Norge.

dr. philos. Einar Lie er forsker ved Senter for teknologi, innovasjon og kultur ved Universitetet i Oslo.

Anne Gro Hustoft er seniorrådgiver i Statistisk sentralbyrå, Seksjon for metoder og standarder (anne.gro.hustoft@ssb.no)

Arbeidskontorene som, spesielt i Amerika, fortjener ros for å ha belyst arbeiderklassens forhold gjennom mange rapporter, har kommet fram til sine resultater via spesialundersøkelser og en slags representative utvalg. Det er imidlertid mulig – selv om jeg ikke kan si noe sikkert om dette – at en i disse så interessante undersøkelsene, ikke har beskjeftiget seg nok med den representative metode som sådan.

Hvordan det nå enn er med dette, synes det meg at denne metoden fortjener å diskuteres allment, og framfor alt her ved instituttet (l'Institut international de statistique).

Det som framfor alt forbauser meg, er at de detaljerte undersøkelser som er gjort med hensyn til arbeiderklassen vedrørende inntekter, boliger og andre økonomiske og sosiale forhold, ikke er gjort på en tilsvarende måte for alle klasser i samfunnet. Det synes klart at på samme måte som en har behandlet spørsmål vedrørende arbeiderne, kan en også sammenlikne den økonomiske, sosiale, moralske etc. situasjonen til arbeiderne med tilsvarende for middelklassen og overklassen. I et land der overklassen er svært rik, og middelklassen meget velstående, vil arbeiderklassens krav mht. lønn, boliger etc. ligge på en annen skala enn i et land (eller et område) der mesteparten av personer som tilhører overklassen ikke er rike eller der middelklassen har dårlig råd.

Ut fra denne påstand, som synes helt innlysende, følger det at en for å kunne vurdere arbeiderklassens situasjon, også må kjenne til de tilsvarende forhold hos de andre klassene. Men en må ta enda et skritt og si at siden samfunnet ikke bare består av arbeiderklassen, kan en i sosiale undersøkelser ikke utelate noen av de andre sosiale klasser.

Et representativt utvalg for et land kan konstrueres på ulike måter, og den metode som bør anbefales i hvert tilfelle, vil i stor grad avhenge av de fenomen eller forhold en vil studere.

Jeg hadde ønsket at spørsmålet om representative utvalg skulle behandles på en mer generell og internasjonal måte, og hadde i det øyemed invitert vår lærde kollega, H. Westergaard fra København, hvis teoretiske kunnskaper innenfor dette området er allment kjent, til å introdusere dette emnet, men fordi det var umulig for ham å komme til Bern, har jeg måttet løpe risikoen med å presentere spørsmålet ut fra – jeg må innrømme det – et begrenset synspunkt.

Med disse unnskyldninger går jeg over til å redegjøre for de to representative undersøkelser som vi den siste tid har brukt i Norge; jeg tror at det på denne måten vil være lettere for meg å forklare forskjellige aspekter ved det spørsmål som er stilt.

Siden begynnelsen av dette året har det i vårt land vært arbeidet med et representativt utvalg med det mål å kunne belyse forskjellige spørsmål knyttet til etablering av en generell folkepensjon (både invalide- og alderspensjon). Dette arbeidet gjøres under tilsyn av en parlamentarisk komité nedsatt for å undersøke disse spørsmålene, og jeg er medlem av denne. Det individuelle spørreskjemaet som brukes ved denne anledningen, inneholder 34 spørsmål til hvert voksent individ. Noen av disse spørsmålene er imidlertid igjen delt på en slik måte at det reelle antall spørsmål beløper seg til flere enn 60.

Det spørres bl.a. om yrke i forskjellige epoker av livet, og om inntektene i disse, om størrelsen på de viktigste utgiftene i 1894, om antall arbeidsdager og dager en har mistet pga. sykdom eller andre årsaker; om helsetilstand og uførhet, når det sistnevnte inntraff og grunnen til dette. Individene i utvalget bes også om opplysninger vedrørende sivilstand, årstall for ekteskapsinngåelse og opphør av ekteskap, antall levende og døde barn, om deres fødsels- og dødsår etc. Disse skjemaene fylles ut av folk som er spesielt ansatt for å hjelpe til med innsamling av opplysninger på det enkelte sted. En foreslår å samle inn skjema fra et totalt antall på omtrent 120 000 voksne individer, av dem er 80 000 framkommet ved et egentlig representativt utvalg, og 40 000 er et tilsvarende spesialutvalg gjort i områder bebodd av arbeidere.

Når det gjelder det representative utvalget, har vi først delt det totale antall i utvalget proporsjonalt mellom byer og landsbygd, basert på forholdstall fra folketellingen i 1891 slik at omtrent 20 000 skjema skal fylles ut i byene og 60 000 på landsbygda.

Systemet for fordeling mellom byer er forskjellig fra systemet for fordeling mellom de ulike distriktene på landsbygda. For byene har en først valgt som representative de 5 byene som har 20 000 eller flere innbyggere, så 8 som er mellomstore eller små, i alt 13, eller nesten en femtedel av det totale antall byer som er 61.

Antall skjema er ikke fordelt mellom de representative byene eksakt i forhold til innbyggertall. Siden de små og mellomstore byene skal representere ikke bare sine egne innbyggere, men også dem i andre små og mellomstore byer, har en funnet det korrekt å tildele disse byene en antall skjema relativt sett større enn det de store byene får. Altså, mens en har tildelt Kristiania 6 350 skjema, noe som tilsvarer omtrent $1/16$ av den voksne befolkningen i 1891, har en valgt $1/9$ som basis for den voksne befolkningen i de mellomstore byene, og inntil $1/3$ som basis i de minste byene. Dette fordelingssystemet, som synes fornuftig i seg selv, har også visse fordeler ved optelling av opplysningene.

Og nå til hvordan fordelingen er gjort innenfor vår hovedstad Kristiania. Fordi det hvert år er en folketelling her, har vi hatt fordelingen av å kjenne folketallet i hver gate etter tellingen i 1892 og 1893. Det var i 1892 400 ulike gater som var delt i 4 ulike kategorier: 1) 100 gater, hver med en befolkning på 100 eller færre, 2) 187 gater med fra 101 til 500 innbyggere, 3) 80 gater med fra 501 til 1 000 innbyggere og 4) 33 gater med 1 001 eller flere innbyggere. Fordelingen av skjema mellom disse gatene er foretatt etter følgende system; en har talt opp 1) hele den voksne befolkningen i en tjuendedel av de minste gatene, dvs. 5 gater, 2) befolkningen i halvparten av husene i $1/10$ av gatene i den andre kategorien, 3) befolkningen i $1/4$ av gatene i tredje kategori ved å ta en femtedel av husene i disse gatene og 4) i halvparten av de tettest befolkede gatene (kategori 4) har en sendt tellere til $1/10$ av husene. I valget av gater som skal inngå i det partielle utvalget, har en gjort seg umak med å få til en jevn fordeling over hele byens utstrekning.

I de mellomstore byene har en fulgt en tilsvarende, men likevel mye enklere, prosedyre, og i de små byene har man simpelthen talt den voksne befolkningen i hvert tredje eller fjerde hus.

Når det gjelder landsbygda, har det vært nødvendig å anvende et annet system. Vi fordelte først skjemaene som skulle fylles ut, på kongedømmets 18 landlige amt i henhold til den voksne befolkningen i 1891. Så, etter å ha kommet fram til en så korrekt representasjon som mulig av hvert amt, gikk vi fram på følgende måte:

Først ble de ulike herredene i amtet klassifisert i henhold til viktigste næring; i landbrukshereder, skogbrukshereder, industrihereder, sjøfartshereder og fiskerierhereder. Så, ved å se på antall hereder tilhørende hver kategori, valgte man representative hereder, og man fastsatte antall skjema som skulle tildeles hvert av dem slik at det samlede antall skjema representerte andelen av landbruks-, skogbruks-, industri- etc. distrikter i hele amtet. Samtidig har en passet på at de ulike geografiske regioner er representert på en så lik måte som mulig.

Det totale antall representative hereder eller kirkesogn beløper seg til 109, det er i gjennomsnitt 6 pr. amt, mens det totale antall hereder i kongedømmet var 498 i 1891.

Antall skjema tildelt hvert representative herred er ikke proporsjonalt med dets befolkning; det hender i virkeligheten ofte at noen av de valgte hereder representerer en for liten eller for stor befolkning blant de valgte hereder sett i forhold til deres viktigste næringsvei. Det kan f.eks. hende at et eller annet herred framstår som representativt ut fra et geografisk synspunkt, men at en spesiell næringsvei er for dominerende der i forhold til yrkesfordelingen i de hereder det skal representere.

En har imidlertid eliminert de oppståtte vanskeligheter ved å tildele slike hereder et antall skjema relativt sett mindre enn de ellers skulle ha hatt, og vice versa i det motsatte tilfellet. Derfor er det i samme amt representative hereder der en har tatt med inntil halvparten av den voksne befolkningen, mens en i andre ikke har tatt med mer enn en tredje- eller en sjettedel, en tidel og enda færre.

Det gjenstår å forklare hvordan det partielle utvalget er foretatt innenfor de representative heredene selv. I dette henseende har man bestrebet seg på å anvende, i så høy grad som mulig, det samme prinsipp som er vist ovenfor ved først å fordele de individuelle skjema mellom ulike kirkesogn i herredet i forhold til befolkningsstørrelse, i det en skiller mellom dem som er spredtbygde, og dem som er tettbygde. Den videre ordning av utvalget i hver del av heredene ble overlatt til tellernes skjønn. Disse hadde imidlertid mottatt instruksjoner mht. valg av "reiserute" som i så høy grad som mulig skulle korrespondere med målet om å ha et mest mulig representativt utvalg. Slik skulle en være påpasselig med å besøke ikke bare det som ut fra et sosialt synspunkt er gjennomsnittsboliger, men generelt de hus som representerer de ulike sosiale og økonomiske forhold som finnes i herredet.

Ved hjelp av alle disse forholdsreglene, under hvilke det ovenfor beskrevne systemet er blitt utført, og fremdeles utføres, er det å håpe at utvalget gir en god representasjon av hele befolkningen som resultat, selv om utvalget ikke omfatter mer enn knapt 6 prosent. Vi har dessuten tatt en spesiell forholdsregel som jeg også tror det er nyttig å nevne.

I betraktning av at resultatet som er oppnådd ved hjelp av den nevnte framgangsmåten, er en ufullkommen, eller for å si det slik, grov, representasjon, har vi for hvert amt reservert et større eller mindre antall skjema for til en viss grad å kunne korrigere den avstand eller avvik som måtte vise seg vedrørende fordelingen av skjema etter yrke. Vi har ut fra en kortfattet klassifisering fra folketellingen i 1891, som ble utført i hvert amt, beregnet fordelingen av de individuelle skjema i forhold til yrke. Vi sammenlikner disse beregnede tallene med tallene som framkommer i det partielle utvalget, og vi supplerer de sistnevnte tallene ved å samle inn flere skjema pr. yrke dersom skjemaene, som er samlet inn etter det grove utvalget, ikke når opp til de tall som indikeres i den generelle folketellingen.

Hvis derimot det partielle utvalget, helt unntaksvis, skulle komme ut med for store tall innenfor noen yrker i forhold til det en skulle ha, kan vi fjerne et visst antall skjema; men hvis forskjellene ikke er store, kan man neglisjere dem i det man tar i betraktning at dette på den ene siden sannsynligvis jevner seg ut i totalsummen for amtene, og at det på den annen side aldri er verken nødvendig eller mulig at et representativt utvalg, på en helt eksakt måte, skal stemme overens med en folketelling i alle detaljer.

Til slutt, for å illustrere metoden bedre, viser jeg i det følgende de foreløpige resultater fra det første amtet, Akershus, som det representative utvalget har blitt prøvd ut i, og som fikk tildelt et totalt antall på 3 850 individuelle skjema. Slik framkommer fordelingen etter yrke for de 3 258 valgte voksne individene i amtet, bosatt på ulike steder i de 7 representative herredene.

Ved å sammenlikne første og andre kolonne i tabellen, finner man, tross en viss generell overensstemmelse, mange betydelige avvik mellom tallene fra utvalget og dem fra folketellingen. Det er sant at viktigheten av disse avvikene fra et metodesynspunkt i en viss forstand minsker hvis en tar med i betraktningen at klassifiseringen etter yrke (i de to kolonnene), kanskje ikke er helt den samme fordi skillet mellom ulike grupper alltid er utydelig pga. uklar angivelse av yrke. Men selv om en kan forklare noe av avviket ut fra dette, gjenstår det fortsatt visse avvik som en må ta med i beregningen.

Og her er noen av de sannsynlige årsakene.

Tabell 1. Fordeling, etter yrke			
Yrke	Beregnet fra utvalg (3 258 individer)	Ifølge folketellingen	
		Justert til utvalgsstørrelse 3 258 individer	Justert til utvalg + tillegg 3 850 individer
Menn			
Funksjonærer i det offentlige	29	30	35
Jordbrukere	332	287	339
Fiskere	3	9	11
Jordbruksarbeidere			
Forpaktere	120	101	119
Sønner som arbeider hjemme	213	150	177
Drenger	160	103	122
Andre	123	112	132
Håndverkere			
Mestere	62	93	110
Arbeidere	93	77	90
Fabrikkarbeidere	137	180	212
Handelsmenn o.l.	24	30	36
Ansatte innenfor handel	33	28	33
Andre yrker	224	388	461
Totalt	1 553	1 588	1 877
Kvinner			
Giftede kvinner uten eget yrke	743	743	877
Kvinner beskjeftiget med jordbruk ..	28	34	40
Kvinner ansatt i småindustri	20	52	61
Diverse sysselsetting	40	81	96
Døtre som hjelper til i husholdningen	254	199	236
Hushjelper	454	253	299
Fabrikkarbeidere	25	46	54
Andre	141	262	310
Totalt	1 705	1 670	1 973
Samlet antall for menn og kvinner ..	3 258	3 258	3 850

Først og fremst kan ikke de representative herredene gi mer enn tilnærmede resultater siden utvalget som er brukt er partielt, mens det var fullstendig i folketellingen.

Det partielle utvalget kunne kanskje ha fungert bedre dersom tellerne hadde hatt mer erfaring. Som resultatene fra utvalget viser, har de besøkt litt for mange hus bebodd av selveiende bønder og for få av arbeidernes hus. Som følge av denne erfaringen mottar intervjuerne nå instruksjoner i den hensikt å eliminere denne feilen i de amt eller utvalg der undersøkelser ennå ikke er igangsatt.

En vil se, ved å sammenlikne 1. og 3. kolonne, at i de fleste tilfeller vil det være lett å forbedre og komplettere det partielle utvalget vha. de skjema som er reservert til dette formålet.

Dersom en legger til de skjema som allerede er samlet inn for hver yrkesgruppe, et antall tilleggsskjema som svarer til differansen mellom første og tredje kolonne, og omvendt, hvis en fjerner det overflødig antall som utvalget ga i visse yrkesgrupper, vil en få en eksakt representasjon av befolkningen fordelt etter yrke. Det vil dessuten ikke bli nødvendig å rette opp resultatene fra det partielle utvalget for hvert amt spesielt nøyaktig fordi mange av avvikene er uten noen betydning, og fordi de sannsynligvis vil bli mer eller mindre eliminert ved utjevning mellom tall oppnådd for kongeriket som helhet.

Det andre eksempelet, som jeg hentydet til ved starten av mitt innlegg, er helt annerledes, og det er nyttig å nevne det til slutt for å vise at et representativt utvalg kan etableres vha. mange forskjellige metoder.

Vi ønsket via den representative metode å bestemme inntektsfordelingen for voksne menn etter yrke, alder og sivilstand. For å nå dette målet, gikk vi fram på følgende måte:

Etter å ha valgt ut 127 representative herreder og 23 representative byer, tok vi først de individuelle skjemaene samlet inn i folketellingen for menn med alder 17, 22, 27, 32, 37 osv. til 97 år og bosatt i de aktuelle herreder/byer. På denne måten fikk vi med 1/5 av den voksne mannlige befolkningen i de representative herreder/byer, men fordi vi ønsket å redusere antallet ytterligere, brukte vi ikke straks alle de utpekte skjemaene, men bare dem der mannens navn begynte på henholdsvis A, B, C, L, M eller N (i Kristiania og noen andre byer med stor befolkning brukte vi bokstavene L, M og N). Ved å kopiere de viktigste opplysningene gitt i folketellingen (alder, sivilstand, næringsvei/yrke) fra disse individuelle skjemaene, kunne vi vha. en spesiell undersøkelse legge til angivelser av formue og inntekt for de respektive individer. På denne måten oppnådde vi et spesielt representativt utvalg som omfatter i snitt 3,3 prosent av den mannlige befolkningen i byene, og 1,6 prosent av mennene på landsbygda.

For til slutt å vise i hvilken grad dette utvalget virkelig er representativt, har jeg gleden av å presentere tabell 2 der en har koblet sammen den prosentvise fordelingen av den mannlige befolkningen etter yrke; den første kolonnen i henhold til folketellingen, den andre etter det spesielle representative utvalget.

En ser her at bortsett fra for sjøfolk, er det ikke stor forskjell mellom resultatene fra folketellingen og resultatene fra det representative utvalget, og avviket for sjøfolkene kan også forklares ut fra et spesielt forhold, nemlig at tallene fra folketellingen relaterer seg til det faktiske befolkningstallet, mens tallene fra det representative utvalget refererer seg til befolkningen som er til stede, hvilket i Norge vil si at et stort antall sjømenn som oppholder seg i utlandet, ikke tas med.

Dette viser at utvalget, på tross av at det er utført i liten skala (det omfatter ikke mer enn 10 800 individer), og på tross av at den anvendte metode, spesielt den som er knyttet til valget av de representative herreder/byer, kunne vært gjort på en mer fornuftig måte, likevel har gitt tilfredsstillende resultater.

Tabell 2. Fordeling, etter yrke og bosettingsområde

Yrke	Befolkning i landlige områder (prosent)		Befolkning i byer (prosent)	
	Ifølge folketellingen	Ifølge utvalget	Ifølge folketellingen	Ifølge utvalget
Funksjonærer i det offentlige	1,5	1,8	4,8	4,8
Jordbrukere	23,9	23,7	0,5	0,2
Fiskere	8,3	7,4	1,9	1,3
Jordbruksarbeidere				
Forpaktere	5,8	5,5	0,0	0,0
Sønner som arbeider hjemme ...	9,6	8,4	0,1	0,0
Drenger	4,5	4,9	0,1	0,1
Andre	5,2	4,4	0,6	0,7
Håndverkere				
Mestere	4,7	4,5	7,4	6,8
Arbeidere	3,5	3,7	17,8	18,1
Fabrikkarbeidere	4,3	5,6	12,1	13,1
Kjøpmenn, redere etc.	1,3	1,3	6,5	6,6
Ansatte innenfor handel	0,7	0,8	6,2	5,4
Arbeidere innenfor handel	0,4	0,3	5,9	4,6
Kapteiner, styrmenn og maskinister	0,9	1,8	3,2	4,6
Matroser	2,0	4,3	5,5	7,6
Andre yrker	23,4	21,6	27,4	26,1
Totalt	100,0	100,0	100,0	100,0

Jeg vil oppsummere dette som følger:

1) de representative metoder kunne få stor betydning for utviklingen av statistikken, framfor alt dersom de spesielle utvalgsundersøkelsene kunne lages på en slik måte at de på viktige punkter kunne kontrolleres mot fulltelling, 2) alt etter omstendighetene, kan det være ulike metoder å velge mellom, og 3) følgelig anbefales statistikerne å studere og diskutere fordelene og ulempene ved de forskjellige metodene.

(Herr Kiær's spørsmål ble av generalforsamlingen på møtet 30. august sendt til den kommende sesjon i l' Institut international de statistique).

Diskusjon

Herr Kiær: Siden kontoret ved vårt institutt har vært så vennlig å forhåndstrykke innlegget om iakttakelsene og erfaringene vedrørende representative utvalg, som jeg har hatt æren av å holde, vil jeg ikke misbruke deres tid med en lang innledning, men begrense meg til å peke på noen essensielle punkter. Jeg tror ikke en kan beregne forholdene i et helt land ut fra funnene i et enkelt område. Det dreier seg om å velge som basis et stort nok antall av små geografiske enheter eller lokaliteter spredd over hele utstrekningen til et land eller en by. Jeg forsøker å bevise at samlingen av et slikt antall enheter til en viss grad kan tjene til å representere hele landet så fremt valget av enheter er gjort etter en god metode. Det sier seg selv at de representative utvalg ikke har som mål å erstatte fulltelling, tvert imot, de har som mål, slik herr Bodio så godt uttrykte det, å gi kjøtt og blod til det skjelett som fulltellingene forsyner oss med.

Det vil være nyttig, før en snakker om representative utvalg eller representative undersøkelser, å rette oppmerksomheten mot hvilke områder av statistikken metodene kan egne seg for.

Som eksempel vil jeg først tillate meg å nevne fordelingen av individuelle inntekter etter yrke, alder og sivilstand i ulike deler av et land. Deretter boligens egenart, rommenes antall og størrelse, sett i forhold til befolkningens yrke og inntekt. En kan ta med mange eksempler; en kan også nevne jordbruksstatistikk, lønninger og utgifter for arbeiderklassen osv. Det er innlysende at mange av de nødvendige detaljer i denne statistikken ikke kan samles inn for et helt land, ei heller for en stor by, men i disse og analoge tilfeller, tror jeg en kan oppnå svært gode resultater vha. representative utvalg eller representative partielle undersøkelser.

Det er en fundamental iakttakelse som gjelder for de undersøkelser som ikke omfatter et helt land. Det er at nøyaktigheten i resultatene ikke så mye avhenger av et større eller mindre antall observasjoner, men av en god metode for å etablere en korrekt representasjon. En må framfor alt unngå faren ved å velge de størrelser som er *lette* å samle inn, hvis en gjør det, vil en ofte få et uriktig utvalg.

Etter dette, ga taleren et kort resymé av rapporten som var blitt trykket og distribuert til forsamlingen.

Herr Guillaume: Jeg er herr Kiær takknemlig for at han har reist dette spørsmålet. Nylig lovpriste herr professor Westergaard interpolasjonsteorien i Revue Conrad, og gjorde oppmerksom på de omkostninger som gir støtet til statistisk arbeid. Dette er omkostninger som kommer til å øke i alle land. Selv om teorien knyttet til representative utvalg aldri kan være en erstatning for den vitenskapelige metode, og fullt ut erstatte fulltellingene, bør vi, fordi det nå er praktisk, være enige om å ikke vise oss som umedgjørlige dogmatikere, men uten fordommer undersøke de forslag som presenteres oss i den hensikt å redusere utgiftene til et minimum. Herr Kiær forlanger ikke at metodens prinsipp skal adopteres i denne sesjonen, men at den skal undersøkes, for til slutt å finne ut om instituttet kan anbefale den og om den kan være tjenlig innenfor det ene eller det andre området i statistikken. Jeg støtter derfor herr Kiær's forslag om å oppnevne en komité som skal legge fram en rapport på neste sesjon.

Herr v.Mayr: Jeg vil ikke la herr Kiær's forslag passere uten å erklære at jeg anser de synspunkter jeg finner i hans arbeid, for meget farlige. Jeg forstår svært godt at undersøkelsene som kalles representative utvalg, kan ha verdi i spesielle tilfeller, men denne verdien vil alltid være begrenset til området belyst av de statistiske observasjoner. En kan ikke erstatte reelle observasjoner av faktiske forhold med beregninger. Hvis en har gjort en partiell observasjon av et område i stedet for en generell, har en nok statistikk for den observerte delen, men en har ikke en korrekt statistikk for hele området. Det er framfor alt farlig å være for teorien om representative undersøkelser midt i en forsamling av statistikere. Man forstår at et slikt begrenset utvalg kan være nyttig i lovgivende eller administrativt øyemed, men en må ikke glemme at denne aldri kan erstatte fullstendige statistiske observasjoner. Det er desto mer nødvendig å understreke det forannevnte, fordi det i disse dager blant matematikere hos oss finnes ideer i omløp om at de i mange henseende heller vil gjøre beregninger enn å observere. Men det er nødvendig å stå fast og si: ingen beregninger der det kan gjøres observasjoner.

Herr Bodio: Jeg støtter herr v.Mayr's synspunkter. Etter mitt syn vil det være farlig å anbefale en indirekte metode, basert på et begrenset antall forsøk,

hvis det dreier seg om å bestemme antall innbyggere i et land eller i hver regional enhet som det består av. En kan benytte seg av sonderinger i større eller mindre avstand fra hverandre for å måle dybden i en sjø eller innsjø eller geologiske lag, fordi noe annet ikke er mulig. Det samme gjelder innenfor demografi. Der finnes det komplekse forhold som ikke kan studeres på annen måte enn ved dyptgående studier av et begrenset antall familier av dem som står oppført i visse folkeregistre, men en kan neppe oppnå resultatene fra en folketelling på noen annen måte enn gjennom en virkelig telling. Ingen induktiv prosedyre, ingen manipulering av tall vedrørende bevegelser i sivilstand eller av registre (som alltid er ufullstendige) vedrørende flytting kan erstatte det effektive og ekte arbeidet som gjøres i en folketelling. Jeg vet imidlertid godt at det finnes saker der forvaltningen kan nøye seg med omtrentlige vurderinger. Dette gjelder f.eks. telling av brev som sendes med posten. Det er innlysende at en ikke kan kreve at all brev- og kortkorrespondanse skal telles hver dag på hvert postkontor i landet. Omkostningene ville bli altfor store, og ute av proporsjon, i forhold til nytten en ville ha av en eksakt telling. Det er derfor en generelt velger noen timer i hvert kvartal, og så ut fra proporsjonalitetsregler beregner korrespondansen som ekspederes gjennom hele året. Når det ved bruk av metoden opplyses at dette ikke dreier seg om noe annet enn en viss grad av sannsynlighet, er det ikke rimelig å forlange mer. Men, for å gjenta, vi må være på vakt mot å kalle et slikt overslag for ekte statistikk.

Herr Rauchberg: Jeg tror ikke det er behov for en ytterligere vurdering av dette spørsmålet. Det dreier seg om å skille mellom statistikk som omfatter alle tilfeller, og den som ikke tar med andre enn dem som representerer et tilfeldig utvalg. Hvis vi begrenser oss til sistnevnte, kan ikke resultatene tjene til å fortelle om noe annet enn de delene av helheten som utvalget er tatt fra. De kan ikke representere helheten. Ingen trenger tvile på hvilken av disse to metodene som er å foretrekke, det dreier seg om prinsipper som er etablert a priori, en diskusjon om disse prinsippene er overflødig. Den første metoden er statistikkens, den andre er det ikke. Jeg er derfor imot å behandle dette spørsmålet videre.

Herr Millet: Jeg synes ikke det er riktig at kongressen med sin stemme skal gi den representative metode (som kort og godt ikke kan være noe annet enn en nødløsning) en viktighet som den seriøse statistikken aldri vil tilkjenne den. Uten tvil gir statistikk basert på denne metoden, som jeg vil kalle en *Pars pro toto* (en del representerer helheten)-statistikk, interessante opplysninger, men dens prinsipper er så i strid med de krav som den statistiske metode stiller, at vi som statistikere ikke kan gi de samme borgerrettigheter, for å si det slik, til denne ufullstendige metode, som vi gir til det ideal som vi på vitenskapelig vis har satt oss fore å nå opp til.

Herr Cheysson: Jeg ønsker ikke at denne diskusjonen skal resultere i en forhåndsdom over spesialundersøkelser, eller den monografiske metode. Da han forleden dag presenterte Dr. Engels arbeid, definerte herr Bodio mesterlig rollene til hhv. den monografiske og den statistiske metode. Den førstnevnte går i dybden på områder, den andre arbeider med helheten. De to metodene er en gjensidig støtte for hverandre, og en kan ikke ustraffet vrake den ene til fordel for den andre. Den generelle statistikken belyser helheter, kartlegger et område og rettleder i valg av spesielle variabler, monografien utdyper disse variablene og gjennomtrenger dem til bunns. Begge

disse metodene er nødvendige og fruktbare, og den ene bør ikke framfor den andre utsettes for angrep i den pågående debatten.

Herr Levasseur: Jeg tror det er nødvendig å skille mellom tre ting; den generelle statistikken som i egentlig forstand har med helheten til fenomener av samme art i en region, land, provins, by etc. å gjøre, monografien som er en detaljert beskrivelse av et objekt eller et unikt fenomen og statistiske undersøkelser.

Den generelle statistikk er den eneste som gir helhetlig (jeg sier ikke komplett) kunnskap om et emne. Monografien gir presis kunnskap om en av emnets enheter, sett i forhold til alle dens aspekter. De to metodene utfyller hverandre, og begge er uunnværlige verktøy i undersøkelser. Den generelle statistikk vil alltid forbli framgangsmåten framfor alle andre når det gjelder å få fram informasjon om svært komplekse forhold, sammensatt av et stort antall enheter, som f.eks. befolkningens tilstand. Jeg er overbevist om at herr Kiær deler herr Bodios synspunkt på dette.

Men utenom disse to metodene, er det ikke også plass til en tredje framgangsmåte; den statistiske undersøkelse, dvs. å undersøke forholdene rundt en type fenomener vha. statistikk som ikke anvendes på helheten av fenomener, men på en antall bestemt av og begrenset til noen fenomener eller grupper av fenomener. Herr Carrol D. Wright har anvendt dette i USA og oppnådd tilfredsstillende resultater. Jeg tror at l'Institut international de statistique gjør klokt i å oppnevne en komité for å studere spørsmålet.

Herr Schmoller legger merke til at statistikkens fanatikere, som han vil kalle dem, ikke vil la seg overbevise om at tellinger, i sak etter sak, viser seg å være unyttige og gi dårlige resultater når det ikke dreier seg om enkle fenomener. Jo mer komplekse sakene er, og jo større antall spørsmål som må stilles i hvert tilfelle, desto mer nødvendig blir det å ikke telle mer enn en enkelt del av tilfellene, men å telle den med desto større presisjon. Hesiod advarte oss mot de dumrianer som aldri vil forstå at ofte er halvdelen bedre enn helheten.

Herr v.Mayr: Det er nå svært nødvendig for meg å ta ordet, for dypest sett er jeg enig med foregående taler. Jeg har også det syn at i et stort antall saker som er undersøkt ved den monografiske metode, er det kommet fram interessante isolerte resultater av stor nytte. Det dreier seg om, med størst mulig nøyaktighet og ved bruk av alle undersøkelsesverktøy, selv uten tellinger og nøyaktige mål, å vinne kunnskap om de sosiale faktorer som en studerer; dette gjelder framfor alt i tilfellet nevnt av herr Schmoller, dvs. tilfeller der det dreier seg om objekter en ikke kan telle eller måle. Men dette er ikke tilfellet med objektene som inngår i herr Kiærs representative utvalg ("stikkprøver"), som man kaller det. Dette dreier seg om objekter som i sin helhet kan telles, men som med overlegg bare telles delvis. Som herr Levasseur så godt har observert, er disse studiene nyttige som "undersøkelser knyttet til en bestemt sak", i den forstand gir de "skinnresultater".

Men det er en helt annen sak å se disse resultatene som en tilstrekkelig erstatning for fulltallige studier. Disse studiene som alene er de sanne statistiske undersøkelser, kan aldri erstattes vha. disse partielle observasjoner som kalles representative. I prinsippet er jeg derfor mot forslaget. Dog ser jeg ingen reell fare i å akseptere forslaget slik det er framsatt, fordi det ikke

dreier seg om annet enn å forelegge spørsmålet for komiteen som vil fortsette med studien. Det er nødvendig at dette spørsmålet klarlegges, og det er også ønskelig at saken diskuteres i dybden i den kommende sesjon for å forebygge at "Hesiøds tid" fortsetter.

Herr Kiær delte til et visst punkt de betraktninger som var kommet fram mot den representative metode, og han mente ikke å anbefale den reservasjonsløst. Likevel opprettholdt han at metoden med fordel kunne brukes i mange tilfeller. Og en har egentlig anvendt denne metoden innenfor visse greiner av statistikken; f.eks. i jordbruksstatistikken hvor man ikke teller opp grøden fra hvert bruk, men nøyer seg med å fastslå kvantiteten for et større eller mindre antall typiske bruk.

(Til slutt vedtok forsamlingen, etter forslag fra herr Kiær, at spørsmålet skulle sendes over til den kommende sesjonen i l'Institut international de statistique.)