

## STATISTIKENS FREMTID.

Afskedsforelæsning den 28. Maj 1924.

Af Harald Westergaard.

Naar man som jeg staar ved Afslutningen af mange Aars Tjenestetid ved Universitetet, er det naturligt at dvæle ved det Fag, man fortrinsvis har skullet vie sine Kræfter, og det maa da for mit Vedkommende blive Statistiken. Mange vilde ved en saadan Lejlighed føle sig tilskyndede til at kaste Blikket tilbage paa den stærke Udvikling, som har fundet Sted i dette Tidsrum. Jeg vil foretrække at se fremad og spejde efter Fremtidsmuligheder. Paa en vis Maade er det vel iøvrigt det samme. Man vil altid bedømme Fremtidsudsigterne paa Baggrund af den foregaaende Udvikling.

Ved første Øjekast vil man vistnok faa et lyst Syn paa Statistikens Fremtid, naar man ser den store Skare af skolede Kræfter, der Verden over staar til Raadighed for dette Fag. Men noget maa man straks trække fra for at naa ned til Virkeligheden. Det er noget lignende som, naar en Landmand en smuk Dag i Forsommeren ser paa sine Marker og tænker paa alt, hvad de kan bære, hvis alle tilstedeværende Muligheder kan komme til at udfolde sig. Denne Rigdom opnaas aldrig. Der vil altid være en Række Uheld, der formindsker det forventede rige Udbytte. Hvad der hindrer Udviklingen i Statistiken er den Planløshed, som ganske naturlig er opstaaet der, en vis Mangel paa Arbejdsdeling, som fører til, at en Mængde Mennesker paa en og samme Tid kaster sig over samme

Emne, medens man lader andre Emner ligge, der kunde være mere frugtbringende. Det gaar her som i den økonomiske Verden, hvor man saa ofte ser en hel Skare af Producenter blindt kaste sig ind i en Forretningsgren, der har givet en enkelt Foregangsmand et godt Udbytte, men ikke vil være lige lønnende, naar der er mange om det. For Tiden er saaledes Begrebet Korrelation paa Mode; mange Mennesker arbejder med de forskellige Kombinationer, som dette Problem fører til, saa at de statistiske Tidsskrifter er opfyldte med Afhandlinger derom, men til Gengæld er der andre Emner, man kunde ønske at se noget mere om.

I det hele taget er det interessant at se, hvor mange Matematikere der i den sidste Tid har kastet sig over Statistiken. Særlig efter Verdenskrigen er Sluserne lukket op, og Lærebog paa Lærebog i matematisk Statistik er udkommet. Der er ikke noget at sige til denne frodige Lærebogsproduktion. Det er ikke mere end rimeligt, at den, der har tilegnet sig et eller andet Stof og lært at se paa Sagen med sine egne Øjne, nu ogsaa trænger til at udtrykke dette paa den Maade, som falder ham naturligt. Jeg nærer ikke Tvivl om, at der, i alt Fald indirekte, derved vil opnaas et Resultat af Betydning. Lad mig drage en Parallel fra Nationaløkonomiens Historie, fra Epigonernes Tid i Midten af forrige Aarhundrede. Det gjaldt for disse Forfattere at fremstille Nationaløkonomiens Hovedsætninger i den klareste mulige Form, og det lykkedes, men Formen blev saa gennemskinnende, at man ikke kunde undgaa at se Skavankerne ved disse Sætninger, deres Mangel paa Virkelighedsindhold. Det er maaske denne Bebrejdelse, man først og fremmest kan rette mod Nutidens Dyrkere af matematisk Statistik, at der er en Mangel paa Virkelighedssans hos den, saa at de tænkte Tilfælde kan komme til at betyde mere end de virkelige. Hvis Nutidens matematiske Statistikere vil gøre sig dette klart, saa at de først og fremmest spørger om Kendsgerningerne

og hvad de har at fortælle, og derefter vil skride til Løsningen af de matematiske Problemer, som Behandlingen af disse Kendsgerninger medfører, vil meget være vundet for Fremtidens Statistik. Den samme Kritik kunde i og for sig have været anvendt overfor mange af de Matematikere, der fortsatte *Pascals* og *Fermats* Arbejde med Sandsynlighedsregningen. Ganske naturlig førte dette Arbejde til Løsningen af mange Problemer, som frembød stor teoretisk Interesse, men ofte førte bort fra Virkelighedens Verden, saa at Sandsynlighedsregningen fik mindre Betydning for Statistikkens Udvikling, end den kunde have faaet.

Ved den matematiske Statistiks Behandling af Iagttagelserne kan man gaa to Veje. Man kan enten opstille Formler for den faktiske Gruppering, f. Eks. af et Antal Arbejdere efter Arbejds løn, af et Lands Befolkning efter Indtægt. Eller man kan søge at bearbejde Iagttagelserne saaledes, at man nærmer sig det, der er den simpleste Forudsætning i Sandsynlighedsregningen, at der er en enkelt fremherskende Aarsag, en enkelt Type, omkring hvilken Iagttagelserne samler sig. Naar man har naaet en saadan Rendyrkning af Materialet, fremkommer den velkendte Binomialkurve, en lille Smule skæv, dog ikke mere skæv end, at man i Regelen i Praksis kan behandle den som symmetrisk. Er der flere Typer, fremkommer der en Kurve, som mere eller mindre afviger fra Binomialkurven.

Jeg ser nu ikke rettere end, at det først og fremmest er Iagttagelsernes Bearbejdelse med Binomialkurven for Øje, man maa ind paa. Studiet af de omtalte mere uregelmæssige Kurver frembyder vel megen Interesse, men det fører os ikke umiddelbart til det Spørgsmaal, som maa være det væsentligste i Statistiken: Aarsagssammenhængen. Naar man søger at underkaste saadanne Kurver en matematisk Behandling i Stedet for først at prøve paa at opløse dem i deres enkelte Elementer, ledes man for-

modentlig ofte uvilkaarlig af en Tro paa, at denne uregelmæssige Form i sig selv er faststaaende, typisk. Men selv om man skulde have Ret i dette, er det urigtigt at standse her, i Stedet for at gaa paa Jagt efter de forskellige Forhold, der er Skyld i Kurvernes ejendommelige Form. Deres Skævhed afgiver et ubedrageligt Vidnesbyrd om, at der er flere Hovedaarsager i Virksomhed, som man da bør sætte sin Kraft ind paa at finde. Dette er et særdeles frugtbringende Grundproblem, og det fører os lige ind i Virkelighedens Verden.

Det gør ikke dette Arbejde mindre interessant, at det kræver meget stor Forsigtighed. Det kan se ud, som om man er ved at naa sit Maal, at Erfaringerne virkelig grupperer sig typisk om et enkelt Punkt. Men der vil altid være smaa Uregelmæssigheder, som hindrer os i at faa Kurven frem med fuldstændig Nøjagtighed, og der er da en Mulighed for, at dette skyldes, at der i Virkeligheden endnu er en Flerhed af Typer tilbage, som man ikke naar at faa fat paa, før der foreligger nye lagttagelser.

Medens det forekommer mig, at et Studium af skæve Kurver af den nævnte Art ikke er særlig lønnende, er der andre Kurver, som det i høj Grad lønner sig at undersøge. Lad os f. Eks. tænke paa Dødelighedssandsynlighedernes Afhængighed af Alderen. Kurven er et Resultat af en Række Sandsynligheder, der hver for sig er iagttaget med den størst mulige Nøjagtighed, men som dog kan fremvise Bakker og Dale, som efter alt at dømme ikke hører med til Kurvens typiske Forløb og derfor bør fjernes ved en Udjævning. Paa Statistikkens nuværende Standpunkt er man for saa vidt Herre over Opgaver af denne Art, at man kan angive praktisk brugbare Metoder til den anførte Udjævning, men der vedbliver at være uløste Opgaver ved disse tilsyneladende simple Kurver, som det kan lønne sig for den matematiske Statistik at give sig i Kast med.

Hvad Indsamlingen og Bearbejdelsen af lagt-

tagelserne angaar, har man i de sidste Decennier gjort meget store Fremskridt. Man har lært at stille Spørgsmaalene koncist og klart, og man har udmærkede tekniske Hjælpemidler, elektriske Tællemaskiner o. s. v. Men man synes ofte for lidet at beskæftige sig med Spørgsmaalet om Nøjagtighedsgraden. Der trænges til mere systematisk Undersøgelse, end man ofte foretager, for at komme paa det rene med Fejlkilderne og Grænserne for de Unøjagtigheder, der ikke har kunnet undgaas. At faa et absolut rent lagttagelsesmateriale er umuligt, Grænserne er mere eller mindre flydende, ofte er der Lakuner. Det bliver da Statistikerens Opgave at vinde Klarhed over, om han trods saadanne mere eller mindre iøjnefaldende Mangler kan drage nogen Slutning af Materialet. Det er ikke mindst i den økonomiske Statistik, at Materialet bereder os saadanne Vanskeligheder, vi behøver blot at tænke paa Handelsstatistikken; men Vanskeligheder er jo til for at overvindes, og her er da et frugtbart Felt for Fremtidens Statistikere.

Et simpelt og nærliggende Middel ved Renselsesprocessen overfor det statistiske Materiale er Anvendelsen af Stikprøver. Vi kan ikke — for at tage et Eksempel — rekonstruere en hel Folketælling for at komme efter de fejlagtige Aldersangivelser, men vi kan undersøge smaa Udsnit af Befolkningen for at faa at vide, hvor stor en Tilbøjelighed der er til at angive Alderen galt.

Men bortset fra dette nødvendige Revisionsarbejde har Stikprøverne, Repræsentativtællingerne, utvivlsomt en Fremtid i Statistiken. Vi lever i en Spareperiode, det er nødvendigt at økonomisere med Kræfterne, men samtidig vokser de statistiske Opgaver fra Dag til Dag. Der er mangfoldige Undersøgelser, som det kunde have den største Interesse at foretage. Her gælder det da, om man ikke ved at gøre et Udsnit af Helheden, og gøre dette til Genstand for nærmere Undersøgelse, kan faa en tilstrækkelig nøjagtig Viden. Lad os som Eksempel nævne

Høsten. Hvis man med tilstrækkelig stor Nøjagtighed kan beregne den samlede Høst i Danmark ved blot at udtage en Brøkdæl af Sognene, i Stedet for at rette Spørgeskemaer til dem alle, saa kan der spares en hel Del Ulejlighed og Omkostning. Eller hvis det gælder om at belyse Indtægtsforholdene, og man kan gøre dette ved efter et vist System at udvælge en Del Gader eller Huse i Byerne til grundig Undersøgelse, saa vil ogsaa her meget være vundet. Det er den gamle politiske Aritmetik om igen, men under langt mere videnskabelige Former og med en langt større Tryghedsfølelse, fordi man har saa mange Helhedsundersøgelser at støtte sig til og har saa mange Midler til at kontrollere sine Metoders Brugbarhed. Der foreligger her en rig Mark til Opdyrkning, og det vil kunne beskæftige mange Statistikere, hvis man skal komme til fuld Klarhed over de mest brugbare Repræsentativmetoder og de Grænser, indenfor hvilke Fejlene vil komme til at ligge. Der vil sikkert vise sig mange Uligheder fra det ene Emne til det andet og fra Land til Land; det vil ikke være gjort med et enkelt vellykket Forsøg; der skal et langvarigt og taalmodigt Arbejde til, for at man tilfulde kan blive Herre over alt dette.

Ved en saadan Anvendelse af Repræsentativtællinger rykker Statistiken saa at sige op i en højere Plan, man faar mindre mekanisk Indsamling af Stof, men stiger op til en maalbevidst Udvalgelse og Bearbejdelse af enkelte Dele af Stoffet, idet man stadig har Blikket henvendt paa Helheden og søger af de enkelte udplukkede Billeder at danne et Totalbillede, som med tilstrækkelig Nøjagtighed kan gengive Virkeligheden. Følger man denne Vej, vil man naa til den mest økonomiske Udnyttelse af de Kræfter, som staar til Raadighed.

Et af de Omraader, hvor man vil kunne drage Nytte af en systematisk Udvikling af Repræsentativtællingernes Teori og Metode, er Studiet af Fortidens statistiske For-

hold. Der ligger rundt om i Arkiverne store Samlinger af Materiale vedrørende svundne Tiders økonomiske og sociale Forhold, som det har den største Interesse at bearbejde: Mandtal, Kvægtællinger, Lønningslister og Regnskaber o. s. v., og som vil kunne udnyttes med langt større Sikkerhed end tidligere, fordi vi har Nutidens Helhedsbilleder at sammenligne med. Her er Arbejde nok for mange Kræfter.

Og ved Siden af denne stille System med Fortiden, frembyder Nutiden en Opgave, som kalder paa mange Mennesker, det er den at samle og sigte de mange enkelte mere eller mindre brudstykkeagtige Oplysninger fra Jordklodens forskellige Lande, for saaledes at faa et Overblik over Verdensbevægelsen. Der var en Tid, da man sværmede for „sammenlignende Statistik“, der skulde give ensartede Billeder af Forholdene i de enkelte Lande. Det blev ofte en Skuffelse. I vore Dage vil man med langt større Sikkerhed kunne magte denne Opgave, og paa Grundlag af disse Oplysninger tilvejebringe de samlede Verdensoversigter, som har saa stor Betydning i vort febrilsk pulserende Nutidsliv: Verdenshøsten, Skibsfarten, Handelsomsætningen, for ikke at tale om Forskydninger mellem Folkeslag og Racer o. s. v. Lad mig i denne Sammenhæng nævne den triste Opgave, hvortil der endnu kun har kunnet gøres Forstudier, at udfinde Tabene af Menneskeliv ved Verdenskrigen og dens Indflydelse paa den almindelige Sundhedstilstand rundt om i Landene.

Spørger man om Opgaverne i Statistiken, ligger det i det hele taget nær at tænke paa Befolkningen og dens Bevægelser. Her er særdeles meget at tage fat paa. Længst er man kommen med Hensyn til Dødelighedsundersøgelser, selv om der uden Tvivl stadig vil være mange Spørgsmaal, der trænger til at belyses, for ikke at tale om, at man selvfølgelig stadig maa have Føling med Udviklingens Gang. Men ved Siden af Dødeligheden gælder det i høj Grad at udvikle Sygeligheds-

og Invaliditetsstatistikken og tillige at undersøge Erhvervsevnen i de forskellige Aldere. Vielses- og Fødselsstatistikken er ikke udtømt, lige saa lidt som Spørgsmaalet om Overgangen fra en Livsstilling til en anden. Og hertil kommer det store og vidtforregnede Spørgsmaal om Vandringer, som endnu kun er delvis bearbejdet. Teoretisk set er det let nok at komme Spørgsmaalet ind paa Livet, selv om man ikke naar til en udtømmende Løsning. Det er klart, at naar man har Kendskab til en Befolknings Forskydninger indenfor et vist Tidsrum, har man Midler i Hænde til at bestemme den samlede Afgangintensitet i hver Alder, ved Dødelighed og Vandringer, og fradrager man da Dødelighedsintensiteten, er Vandringsintensiteten med det samme bestemt.

I denne Sammenhæng kan dernæst nævnes den medicinske Statistik. Der var en Tid, da det, der gik under dette Navn, for en stor Del var et værdiløst Sammensurium, f. Eks. Optegnelser af Hospitalsdødsfald eller Gennemsnitsalderen for Døde i forskellige Samfundsklasser. Det er nu nærmest en Saga. Der er i den medicinske Faglitteratur en voksende Forstaaelse af Betydningen af rationelle statistiske Undersøgelser. Det gælder i vore Dage ikke blot Dødeligheden, f. Eks. de forskellige Behandlingsmaaders Virkning, men ogsaa en Mængde fysiologiske Forhold, saaledes Blodlegemers Antal under forskellige Betingelser. En skolet Statistiker vil her kunne være til stor Hjælp ved Opstillingen af de fornødne Formler og ved Bedømmelsen af Resultaterne, og man trænger til rationelle Lærebøger om dette Emne. Hertil kommer to Emner af stor Interesse: Antropometrien og Arveligheden. Den første har havt sin store Betydning i Statistikens Udviklingshistorie, idet man i den menneskelige Legemshøjde fandt et ypperligt Eksempel til Belysning af lagttagelsers Fordeling om deres Gennemsnit. Selve Spørgsmaalet om Legemshøjde er nu stærkt gennemarbejdet, men der er mange andre Omraader af Antropometrien,



hvor der er meget at udrette, bl. a. ogsaa saadanne Spørgsmaal, hvor man vanskelig kan finde en numerisk Maalestok, f. Eks. Haarets og Øjnenes Farve. Man kommer ogsaa ganske naturlig til at tænke paa Egenskabernes Arvelighed, Spørgsmaalet om, i hvilken Udstrækning vi genfinder Forældrenes Egenskaber hos Børnene o. s. v. Det er Problemer, hvor det er vanskeligt at isolere Aarsagerne. Spørger man f. Eks., om Mennesker med lang Livsvarighed faar Børn med stor Vitalitet, saa maa man huske paa de mange Forhold, der griber ind i et Menneskes Livsbane, at tilfældige Aarsager kan bortrive en Mand tidlig, som maaske efter hele sin Konstitution kunde have ventet et langt Liv. Man faar altsaa aldrig helt rene lagttagelser, men maa være tilfreds, hvis man kan konstatere, om der har været en Indflydelse eller ikke. Bedre vil man kunne være i Stand til at regulere Aarsagerne, hvor det drejer sig om Planter og Dyr.

Baade i Antropometri og ved Arvelighedsundersøgelser anvender man ofte Korrelationsformler, og dette Apparat har efterhaanden, som jeg berørte, været Genstand for en alsidig Behandling. Da Korrelationsmetoden principmæssig ikke adskiller sig fra de tidligere sædvanlig benyttede Metoder, kan det nærmest synes at være en Smagssag, i hvilken Udstrækning man vil bruge den. Det maa dog indvendes mod den, at den indeholder en Fristelse til at arbejde for mekanisk, idet man saa at sige kaster Tallene ind i en Maskine, der leverer det færdige Produkt, medens det i Virkeligheden gælder om at fjerne sig saa lidet som muligt fra sit lagttagelsesmateriale. De mere gammeldags Metoder giver os ofte langt tydeligere Begreb om Tallenes indbyrdes Sammenhæng. Jeg kan ikke frigøre mig fra den Tro, at Korrelationsteorien aldrig var kommen saaledes paa Mode, som den er, hvis de mange matematiske Problemer, den kan give Anledning til, ikke havde øvet deres Tiltrækningskraft paa Matematikere, og jeg tror, at disse Formler efterhaanden vil blive

lagt til Side som altfor uhaandterlige. Den største Fare ved dem ligger i, at de let kan friste til den Tro, at der er fundet en dyberegaaende Sammenhæng mellem de iagttagne Størrelser, noget i Retning af Naturlove. Det er den samme Tillid til Tallene og deres typiske Karakter, der i sin Tid gav sig Udslag i den ivrige Søgen efter en matematisk Lov for Dødelighedens Afhængighed af Alderen.

Det kan se ud, som om der ofte er en lignende Tankegang, bevidst eller mere eller mindre langt nede i Underbevidstheden, med Hensyn til de periodiske Bevægelser i de økonomiske og sociale Forhold. Problemet om periodiske Bevægelser møder vi allerede i Meteorologien. Det er klart, at denne Videnskab maa staa i nøje Forbindelse med Statistiken. Spørgsmaalet om der f. Eks. er Perioder i Regnmængden fra Aar til Aar er et statistisk Problem. Dette Spørgsmaal har givet Anledning til Opstilling af interessante periodiske Funktioner. Dog vil man ofte gøre den Erfaring, at man i mange Tilfælde faar nok saa meget at vide ved en ganske elementær Behandling af Tallene. Og man gør i hvert Fald vel i at begynde dermed, fordi de meteorologiske Forhold paavirkes af en Mængde forskellige Aarsager, som det vil være lettere at faa Øje paa, naar man ikke straks binder sig ved Opstilling af matematiske Love for Bevægelsen.

Fra Meteorologien sker Overgangen ganske naturlig til de periodiske Bevægelser i alle de økonomiske Forhold. Her er et Omraade, som mange Statistikere i de sidste Aaringer har kastet sig over med stor Iver. Ogsaa her har man ment, at der kan opnaas noget ved Hjælp af periodiske matematiske Funktioner, men Aarsagerne er ogsaa her saa mangfoldige og saa stærkt varierende — man tænke blot paa Verdenskrigen — at en Formel let bliver en Prokrustesseng, at man hindres i at faa Øje paa Tingenes øjeblikkelige Sammenhæng. Det lader sig uden Vanskelighed gøre at finde simple og fuldkommen elementære Metoder til Belysning af dette Spørgs-

maal. Men det er iøvrigt godt og naturligt at faa mange Arbejdskræfter sat i Gang her. Det er formodentlig nok saa vanskeligt at forudsige økonomiske Storme, som det er for Meteorologerne at profetere om Vejret, men noget vil man dog være i Stand til at faa Klarhed over, og en Sammenstilling af alle de mange lagttagelser, man raader over i den økonomiske Statistik angaaende periodiske Bevægelser, vil kunne gøre stor Nytte. Spørgsmaalet er, hvorvidt der er nogen Sammenhæng mellem to Rækker af Tal, der belyser forskellige økonomiske Forhold, hvorvidt man af en Stigen eller Dalen det ene Sted vil kunne forudsige en tilsvarende Bevægelse et andet Sted, indtrædende efter en kortere eller længere Tid. Det drejer sig ikke blot om de maanedlige eller ugentlige Bevægelser indenfor et Kalenderaar, der ofte har en ejendommelig Rytme, ganske som Fødsler og Dødsfald, men ogsaa om Bevægelserne indenfor længere Tidsrum. Her synes at være Udsigt til et godt Udbytte. Overhovedet kan man vistnok sige, at efter at den økonomiske Statistik nu i saa mange Aar har været Stedbarn, er dens Tid nu ved at komme, og hele det uhyre Stof vil give Anledning til en uoverskuelig Række af Opgaver.

Man kunde nævne adskillige andre Omraader, hvor statistiske Undersøgelser vil kunne give frugtbringende Resultater. Jeg har saaledes f. Eks. ikke omtalt Kriminalstatistiken, et af de mange Afsnit af den saakaldte Moralstatistik, der i tidligere Tid laa i Kaos. Skolevæsenet frembyder en Række yderst interessante statistiske Opgaver, som f. Eks. Resultatet af Standpunkts- og Intelligensprøver. I det hele taget kan man sige, at overalt, hvor der spørges om, hvilke Aarsager der paavirker en Række forskellige Elementer, bliver der Anledning til en Gruppering af lagttagelserne, d. v. s. til en statistisk Undersøgelse. Saadanne Undersøgelser vil altid faa deres Præg af Spørgsmalets specielle Natur, men i Hovedsagen vil de blive ført efter de samme Principer.

I hvert Fald turde det fremgaa af, hvad jeg her har udviklet, at en Statistiker i vore Dage ikke behøver at frygte for at blive arbejdsløs. Jeg har søgt at skitsere nogle af de Opgaver, der ligger for, og det er mit Haab, naar jeg nu staar ved Afslutningen af min Løbebane ved Universitetet, at denne Skitse maa indeholde en og anden frugtbringende Arbejdsanvisning. Skulde der være Impulser for Statistikere i, hvad jeg her har udviklet, vil jeg anse det for en lykkelig Udgang af min Gerning ved Universitetet.