

TEKST NR 16

1979

MICHAEL OLSEN

JØRN JENSEN

STATISTIK I

KRÆFTFORSKNINGEN

Projektrapport.

Vejleder:

JØRGEN LARSEN

TEKSTER fra

IMFUFA

ROSKILDE UNIVERSITETSCENTER
INSTITUT FOR STUDIET AF MATEMATIK OG FYSIK SAMT DERES
FUNKTIONER I UNDERVISNING, FORSKNING OG ANVENDELSER

-
- 1/78 "TANKER OM EN PRAKSIS" - et matematikprojekt
Anne Jensen, Marianne Kesselhahn, Lena Lindenskov og Nicolai Lomholt.
Vejleder: Anders Madsen.
- 2/78 "OPTIMERING" - Menneskets forøgede beherskelsesmuligheder af natur og samfund.
Projektrapport af Tom J. Andersen, Tommy R. Andersen, Gert Kreinø og
Peter H. Lassen. Vejleder: Bernhelm Booss
- 3/78 "Opgavesamling", breddekursus i fysik.
Lasse Rasmussen, Aage Bonde Kræmmer, Jens Højgaard Jensen.
- 4/78 "Tre essays" - om matematikundervisning, matematiklæreruddannelsen og
videnskabsrindalismen.
Mogens Niss.
- 5/78 "BIBLIOGRAFISK VEJLEDNING til studiet af DEN MODERNE FYSIKS HISTORIE"
Helge Kragh.
- 6/78 "Nogle artikler og debatindlæg om - læreruddannelse og undervisning i fysik,
og - de naturvidenskabelige fags situation efter studenterooprøret"
Karin Beyer, Jens Højgaard Jensen, Bent C. Jørgensen.
- 7/78 "Matematikens forhold til samfundsøkonomien"
B.V. Gnedenko.
- 8/78 "DYNAMIK OG DIAGRAMMER". Introduktion til energy-bond-graph formalismen.
Peder Voetmann Christiansen.
- 9/78 "OM PRAKSIS' INDFLYDELSE PÅ MATEMATIKKENS UDVIRKLING"
Motiver til Kepler's: "Nova Stereometria Doliorum Vinariorum"
Projektrapport af Lasse Rasmussen.
Vejleder: Anders Madsen.
-
- 10/79 "TERMODYNAMIK I GYMNASIET"
Projektrapport af Jan Christensen og Jeanne Mortensen
Vejledere: Karin Beyer og Peder Voetmann Christiansen.
- 11/79 "STATISTISKE MATERIALER"
red. Jørgen Larsen.
- 12/79 "Lineære differentiallyigninger og differentiallyigningssystemer"
Mogens Brun Heefelt.
- 13/79 "CAVENDISH'S FORSØG I GYMNASIET". Projektrapport af Gert Kreinø.
Vejleder: Albert Chr. Paulsen.
- 14/79 "Books about Mathematics: History, Philosophy, Education, Models, System
Theory, and Works of Reference etc. A Bibliography".
Else Høyrup.
- 15/79 "STRUKTUREL STABILITET OG KATASTROFER i systemer i og udenfor
termodynamisk ligevægt" Specialeopgave af Leif S. Striegler.
Vejleder: Peder Voetmann Christiansen.



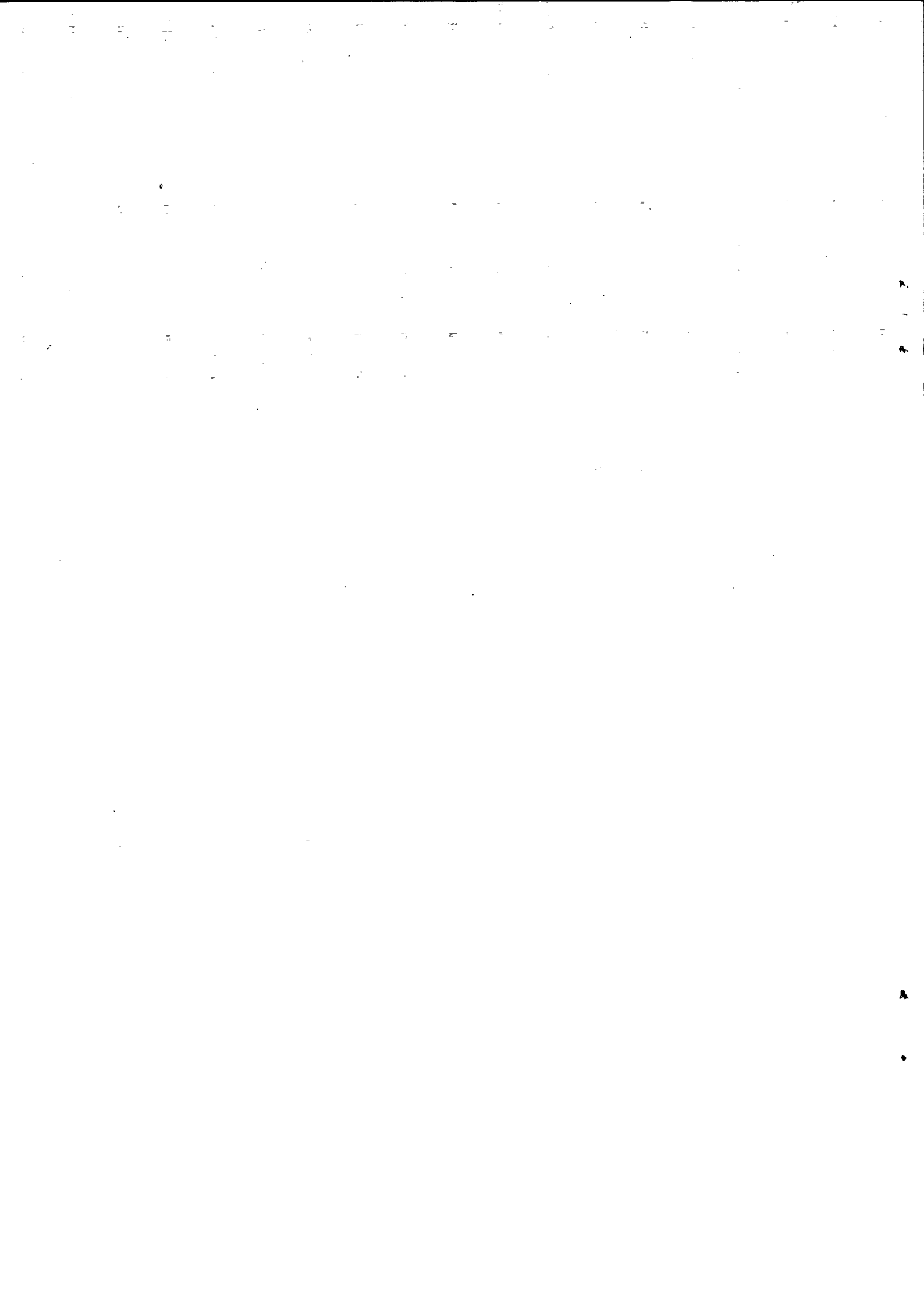
STATISTIK I KRÆFTFORSKNING

MICHAEL OLSEN

JØRN JENSEN

MAT / BIO JUNI 1978

Vejl. Jørgen Larsen



RETTELSESBLAD.

Ved en løs gennemgang af rapporten er fundet følgende væsentlige fejl:

- side 7, linie 8 fra oven: (o7) rettes til (lo)
- side 12, linie 8 fra neden: bilag 11 rettes til
bilag 10
- side 12, linie 4 fra neden: bilag 12 rettes til
bilag 11
- side 17, linie 4 fra neden: type I fejl stor rettes til type II fejl lille
- side 18, linie 1 fra oven: type II fejl rettes til type I fejl
- side 18, linie 10 fra oven: materiale rettes til problemstilling
- side 21, linie 3 fra oven: 100 cm^3 rettes til 100 m^3
- side 49, linie 8 fra neden: "variansen lig middelværdien....." rettes til "variansen lig middelværdien. Endvidere er variansen og middelværdien ikke uafhængige,....."
- side 73, linie 7 fra neden: note l) og note h) rettes til hhv. note i) og note e)
- side 75, linie 11 fra neden: lille rettes til stor

0 FORORD.

Dette projekt skal sammen med supplerende kurser opfylde de faglig-pædagogiske krav til det erkendelsesteoretiske/videnskabsteoretiske modul (modul 3) på den matematiske overbygning ved gymnasielæreruddannelsen på RUC. De specifikke knudepunkter, der skal behandles under dette modul er følgende:

Matematikkens erkendelsesteoretiske status, herunder

- i hvilken forstand og i hvilket omfang matematiske ræsonnementer i en model kan betragtes som meningsfulde og sande udsagn om den virkelighed, modellen foregiver at beskrive,
- hvorledes sammenhængen er mellem matematiske tankeformer og deduktive processer i andre forbindelser,
- hvorledes matematisk erkendelse frembringes bl.a. ved en vekselvirkning mellem matematikkens indre dynamik og dens modelfunktion, samt hvorledes der rent faktisk har fundet gensidig påvirkning sted mellem matematikkens udvikling og processer i samfundet.

Aksiomatisk behandling af områder i faget.

Kognitive problemer ved tilegnelse af matematik på forskellige udviklingsstrin og i forskellige socialgrupper.

Forskellige arbejdsformer til brug for skolens matematikundervisning, betragtet som et middel til bl.a. at give eleverne en fond af forståede færdigheder.

Som supplement til projektet har vi deltaget i kursus i logik og mængdelære samt i differentiaalligninger.

Vi er to studerende, der begge læser matematik og biologi på gymnasielæreruddannelsen. Valget af gymnasielæreruddannelsen skyldtes ikke så meget interesse i undervisning eller intentioner som undervisere som mangel på alternativer. Det er derfor hensigten at vi på

et tidspunkt efter bestået kandidateksamen vil søge videreuddannelse.

Valget af biologi skyldes dels en decideret interesse i faget, og dels at vi kan se nogle perspektiver i biologis rolle i samfundsudviklingen, og valget af matematik skyldes dels et ønske om en for fordelagtig fagkombination med hensyn til senere erhvervsmuligheder og dels et ønske om at kunne tage aktivt del i videreudviklingen af biologi, der netop i øjeblikket er udsat for en kraftig matematisering. Denne matematisering giver sig udslag i øget modelopstilling af biologiske fænomener og i øget brug af statistik og er på sin vis problematisk. For det første er det vores erfaring, at mange biologer kun har et ringe kendskab til matematik, og dette nødvendiggør et samarbejde med matematikere og statistikere, der ofte kun besidder sporadisk kendskab til områder inderfor biologien. Den kendte kommunikationskløft mellem eksperter fra forskellige fagområder giver sig i dette tilfælde af og til udslag i besynderlige modelopstillinger (både matematiske og statistiske), og en konsekvens heraf kan være en forringelse af biologis status som videnskabsfag. For det andet vil biologers manglende kendskab til matematik gøre undervisningen mangelfuld i områder, hvor netop matematik spiller en stor rolle.

Vi har som rød tråd gennem matematikstudiet valgt statistik, ud fra den begrundelse at statistikken floerer kraftigt i store dele af biologien.

Statistikken var altså udgangspunkt, da dette projekt påbegyndtes.

Som optakt til arbejdet foretog vi en litteratursøgning for at få et overblik over, hvor og i hvilket omfang statistik benyttedes indenfor biologisk forskning. Vi fandt ret hurtigt, at et velegnet område at beskæftige sig med var kræftforskning, og i denne optog epidemiologiske undersøgelser betydelig plads.

Som arbejdstitel valgtes "statistikks anvendelse i kræftforskning" og for yderligere at indsnævre området, valgte vi at koncentrere os om lungekræft og luftforurening. Dette valg begrundedes bl.a. ud fra et ønske om ikke at sprede os på for mange områder, men koncentrere os om en enkelt eller få problemstillinger. Vi fandt derfor Fredericiaundersøgelsen ideel. Navnlig da det gik op for os, at Fredericiaundersøgelsen er mere end den foretaget af Clemmesen et al. (14). Vi fik her lejlighed til at følge et helt forløb: først den meget debatterede Clemmesen et al.-undersøgelse (14), dernæst en Superfosbetalt modundersøgelse (01) og endelig undersøgelser iværksat af Sundhedsstyrelsen og Miljøstyrelsen (28), (36), (37). Men væsentligst af alt er det nok at erfare, hvordan gassen siver ud af mangel på velunderbyggede konklusioner. Ud fra Fredericiaundersøgelsen problemformulerede vi forløbet. Det vi ønskede at arbejde med, kunne formuleres i følgende tre punkter:

- 1) Hvordan tilrettelægges undersøgelser til påvisning af sammenhæng mellem lungekræftforekomst og luftforureningsgrad, og i hvor høj grad tillader undersøgelserne de dragne konklusioner.
- 2) Hvilken argumentation ligger til grund for anvendelsen af bestemte statistiske metoder indenfor lungekræftforskningen, og er denne argumentation holdbar.
- 3) Hvordan kan vores opnåede viden om 1) og 2) formidles til elever i HF/gymnasiet.

De mål, der bl.a. lå til grund for problemformuleringen var:

- at se på føudsætningerne for anvendelse af statistik,
- at trænge ned i den matematik, der ligger til grund for statistiske metoders anvendelighed,
- at få indsigt i hvordan undersøgelser tilrettelægges og hvilke komplikationer, der ligger heri, og

-at arbejdet skal åbne mulighed for et fælles matematik/biologi-speciale om en relevant biologisk problemstilling, hvori indgår et statistisk materiale.

Punkt 3) i problemformuleringen medtoges, som det fremgår tidligere i forordet, mere af pligt end af lyst. Vi har fundet det forsvarligt at nedprioritere dette område, idet vi i det tidligere matematik-modul udelukkende har arbejdet med amtematik i skolen, og idet vi gennem vore biologimoduler og gennem den obligatoriske pædagogiske dimension har beskæftiget os med generel pædagogik samt med få men relevante områder fra psykologien og sociologien.

I vores målformulering har vi nævnt, at vi ønskede at trænge ned i den matematik, der ligger til grund for statistiske metoders anvendelighed. Dette mål måtte vi bortrationalisere. Vi havde nemlig ikke realistiske muligheder for at nå at kunne opfylde dette mål. Målet er dog kun midlertidigt forkastet, da vi i vores næste projekt, som bliver et integreret specialeprojekt i biologi og matematik, vil lade det genopstå.

Man vil endvidere se, at vi ikke i problemformuleringen har nævnt statistikkens historiske tilblivelsesprocesser. Igen er argumentationen for kort tid og for få projektdeltagere. Vi finder emnet af stor vigtighed, men har måttet erkende, at vi kun har kunnet koncentrere os om de nævnte problemer og mål, da vi som højst prioriteret princip har, at vores arbejde skal indeholde konstruktiv kritik.

Som vi nævnte før, var vores udgangspunkt undersøgelser af sammenhængen mellem lungekræft og luftforurening med særligt henblik på Fredericia.

Blandt den udenlandske litteratur, vi kom i besiddelse af, var der enkelte, der også beskæftigede sig med problemet lungekræft og luftforurening. Undersøgelserne var tilrettelagt anderledes end de danske, og derfor har vi også arbejdet med disse. Dels for at udvide over-

blikket over, hvordan sådanne undersøgelser tilrettelægges og dels for om muligt at finde idéer til, hvordan ideelle undersøgelser tilrettelægges.

Til sidst skal blot nævnes, at store dele af dette semesters arbejde skal danne grundlag for bl.a. vores specialeprojekt. Vi vil derfor være taknemmelige for al form for kritik.

Kritik, kommentarer og anmodninger om rapporter kan sendes til

Jørn Jensen og Michael Olsen
MAT/BIO Hus 121
ROSKILDE UNIVERSITETSCENTER
4000 Roskilde.

INDHOLDSFORTEGNELSE.

0	FORORD	1
	INDHOLDSFORTEGNELSE	6
I	INDLEDNING	7
II	ANVENDELSE AF STATISTIK I EPIDEMIO- LOGISKE UNDERSØGELSER	10
III	UNDERSØGELSER AF SAMMENHÆNGEN MEL- LEM LUNGEKRÆFT OG LUFTFORURENING	19
IV	FREDERICIAUNDERSØGELSEN	36
V	KONKLUSION	63
VI	UNDERVISNING I STATISTIK OG SAND- SYNLIGHEDSREGNING I GYMNASIUM/HF	64
	NOTER	71
	LITTERATURFORTEGNELSE	77
	BILAG	82

I INDLEDNING:

I løbet af de sidste årtier er forskningen i kræft til stadighed blevet intensiveret, først og fremmest indenfor den medicinske behandling samt grundforskning i hvilke biologiske mekanismer, der har med kræft at gøre.

Først indenfor de seneste år er en ny gren af kræftforskningen kommet til: Forskning i miljøfaktorets betingning af kræft (07).

Denne forskning beskæftiger sig med, hvilke stoffer, eller påvirkninger i det hele taget, der fremkalder øget risiko for kræft -uden nødvendigvis at kunne angive en sikker årsagssammenhæng.

I lang tid har man kendt til særlige kræftfremkaldende (carcinogene) stoffer, som overført eller indgivet i en population (mus, mennesker etc.) medfører at en del af populationen dør af kræft.

De fleste af disse kendte carcinogener er endvidere kendte mutagener, dvs. de påvirker cellernes genetiske materiale DNA, og dette giver et solidt grundlag for at antage, at kræft skyldes uregelmæssigheder i cellernes genetiske materiale.

I meget korte træk kan den almindeligt accepterede teori fremstilles således: Der opstår i en eller flere af kroppens celler ændringer i det genetiske apparat, som medfører at cellen begynder at "gro vildt", hvorefter den ødelægger det omkringliggende væv; er dette et vigtigt organ ødelægges dets funktion, og individet dør. Kræftsvulster kan endvidere spredes via blodet og danne nye svulster i vigtige organer.

Det falder udenfor denne rapportes intentioner at give en mere tilbundsående behandling af grundforskningen samt den medicinske forskning, men vi skal alligevel lige nævne nogle forhold af betydning for den videre behandling.

For mange af de kendte mutagener gælder, at man ved hvordan de biokemisk indvirker på DNA'et (fejlparring af baser, inaktivering m.m.), men der er dels tale om et meget lille antal stoffer i forhold til det antal stoffer, man har mistanke til, og dels er der tale om ret kraftige virkninger, som let lader sig registrere. En stor del af den ovenfor anførte viden stammer direkte fra bakterieforsøg; nu er der imidlertid ret stor lighed på selve det genetiske apparat hos alle levende organismer, mens der med hensyn til optagelse, nedbrydning og udskillelse er stor forskel mellem dyregrupperne.

Fra bakterieforsøg vides, at selv om disse dyrkes på helt rene substrater, forekommer der en (omend lille) mutationshyppighed, som det formodentlig vil være umuligt at eliminere helt (vi vil ikke her komme ind på de nærmere forhold, men blot nævne, at f.eks. ultraviolet lys og kosmisk stråling vides at indvirke, ligesom det må formodes, at der altid vil forekomme en vis fejlfrekvens), disse forhold må også antages at gælde for mennesker.

De fleste forskere er dog efterhånden nået til det resultat, at i hvert fald 80% af al kræft hos mennesker er forårsaget af miljømæssige påvirkninger af forskellig art (10).

Ud fra de teorier, vi meget summarisk har gjort rede for, skulle det stå klart, at kræft først og fremmest er en miljøsygdom: Fremmede stoffer kommer via mad, luft, arbejde etc. ind i organismen, indvirker på nogle af kroppens celler, og skaber mulighed for at disse celler udvikler sig til kræft.

Når vi sammenholder dette med det faktum, at der årligt udvikles ca. 250.000 nye stoffer (40), hvoraf en stor del indføres uprøvede på markedet, hvor vi intet kender til dissestoffers mulige kræftfremkaldende virkning, må enhver erkende, at vi har behov for at forbedre vort

miljø ved at rense det for alle mistænkelige stoffer, samt at sørge for at nye stoffer testes før de kommer på markedet.

De muligheder vi i dag råder over til test for kræft-fremkaldende virkning, kan opdeles i tre hovedgrupper: test med bakterier, test med mus og befolkningstests udfra forskellige statistikker (epidemiologiske undersøgelser).

Test med bakterier viser sig meget lovende fordi der efterhånden er fundet stammer, som viser en stor (90% (09)) korrelation mellem mutagen effekt hos bakterierne, og carcinogen effekt på mennesker; endvidere er testene nemme at udføre, og vil kunne udføres i stort antal.

Test med mus har den fordel fremfor bakterietests at mus ligger mennesket betydeligt nærmere i taxonomi, hvilket gør at forholdene omkring optagelse, nedbrydning og udskillelse i kroppen er ret ens; de har dog den ulempe, at de er tidskrævende og dyre, hvis forsøgsmaterialet skal være tilstrækkeligt stort.

Epidemiologiske undersøgelser har den fordel at resultaterne kan anvendes direkte (selvfølgelig), men til gengæld en række svagheder, som vi i denne rapport vil belyse.

I afsnit II vil vi behandle det generelt, og i de følgende afsnit vil vi gå i detaljer udvalgte epidemiologiske undersøgelser, herunder bl.a. den såkaldte "Fredericiaundersøgelse".

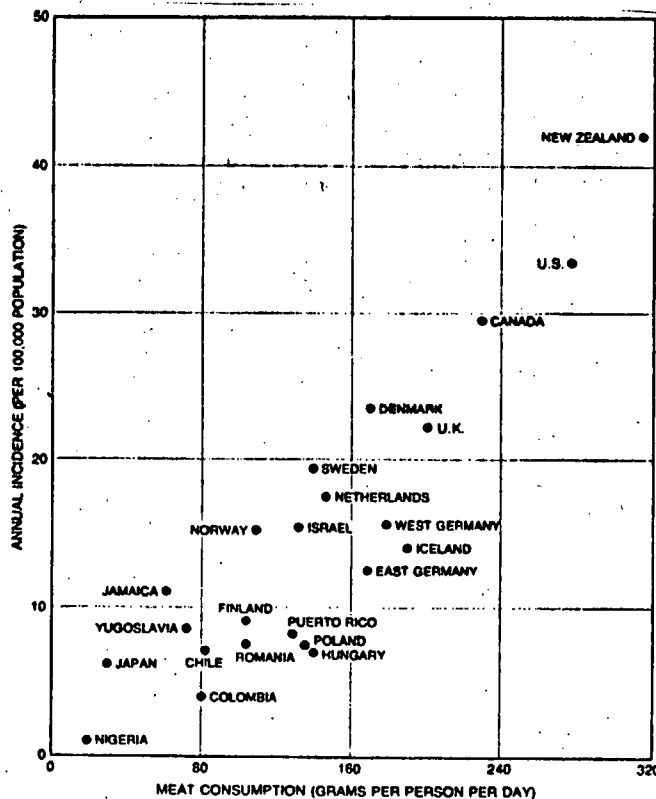
I rapportens sidste afsnit vil vi se lidt på, hvordea undervisningen i HF/gymnasiet i statistik kan tilrettelægges, så den tilfredsstillende det stigende behov for indsigt i anvendelsen af statistik.

II ANVENDELSEN AF STATISTIK I EPIDEMIOLOGISKE UNDERSØGELSER.

Vi skal i dette afsnit se generelt på hvordan statistikken bruges (og misbruges) indenfor kræftforskningen.

Ved hjælp af statistiske analyser og tests baseret på forskellige folkeregistre (såsom syge-og dødsjournaler, m.m., f.eks. (09)) undersøger man :

Om rygere har større risiko for at få lungekræft end ikke rygere; om folk i byer har større risiko for kræft end landboere; om arbejdere i visse industrier (svejser, malere etc.) udsættes for specielle kræftfremkaldende stoffer; om forskelle i socialgrupper giver forskelle i kræfthyppighed; eller som fig II.1., om folks spisevaner kan have indflydelse på hyppigheden af tarmkræft - vi skal senere kommentere fig II.1.



GEOGRAPHY OF A CANCER suggests a probable cause of the disease. The incidence of cancer of the large intestine among women in 23 countries is closely related to per capita meat consumption in those countries. The data are adjusted to eliminate differences in age distribution in the populations. An alternative explanation attributes cancer of the large intestine to a low consumption of cereals. The two hypotheses are hard to distinguish from each other because high meat consumption and low cereal consumption tend to go together.

Den sædvanlige metode for disse undersøgelser består i at udskilde en bestemt gruppe (f.eks. rygere), og sammenholde hyppigheden af kræft i denne gruppe med en kontrolgruppe (af ikke rygere); talmaterialet underkastes så en statistisk test, og som regel anvendes 5 % signifikansniveauet til at afgøre om der er forskel på grupperne.

De nævnte hypoteser er mere eller mindre begrundet ud fra den ufuldstændige viden vi har om de biologiske systemer i legemet, og for en stor del af hypoteserne gælder at de påståede sammenhænge hverken kan be- eller afkræftes på grundlag af eksisterende medicinsk viden, omend man i nogle tilfælde kan have en meget stærk mistanke.

Et godt eksempel på dette er rygning og lungekræft. Efterhånden er der lavet så mange undersøgelser om dette, at sammenhængen må betragtes som bevist, selvom det kan være svært at kvantificere forholdet på en rimelig måde, ((11) nævner noget med 20 %).

Man ved, at røgen består af gasser (CO, NO₂ m.fl.) som kan optages i blodet gennem alveolerne i lungerne; en andel del af røgen består af partikler, som aflejres i munden, luftrøret, bronchierne, og alveolerne afhængigt af partikelstørrelsen (kun de mindste når alveolerne); disse røgpartikler består af en række organiske forbrændingsprodukter "tjærestoffer", og om flere af dem ved man bla. fra forsøg med mus, at de er kræftfremkaldende.

De største partikler fanges i ciliesystemet i luftrøret og de øvre bronchier, og føres tilbage til munden, hvor de enten spyttes op, eller synkes og ender i maven. Ligeledes det finere af lungerne har et "clearingssystem", som er i stand til at absorbere fremmedstoffer og mindre partikler, og føre disse ud af organismen.

Nu er effektiviteten af disse forsvarsmekanismer imid-

lertid ret forskellig blandt mennesker, ligesom forskellige typer røg påvirker clearingssystemet forskelligt, og det gør, at det er umuligt at fastlægge en kausal sammenhæng mellem rygning og lungekræft, (hvem kender ikke historien om den 90-årige, som har røget og drukket hele livet, uden hverken at få lungekræft eller leverskader !).

Den påviste sammenhæng hviler altså dels på en generel medicinsk viden, og dels på en påvist stokastisk sammenhæng.

Den stokastiske sammenhæng påvises på følgende måde :

Vi vælger en population af rygere, og en af ikke rygere, og tester om hyppigheden af lungekræft er større i den første end i den sidste, hvilket er illustreret i nedenstående skema :

	lungekræft	ej lungekræft
rygere	p_{11}	p_{12}
ikke rygere	p_{21}	p_{22}

hvor p er sandsynligheden for at falde i de givne celler; p_{11} er da lig $(1 - p_{12})$.

Vi ser, at antal lungekræfttilfælde er binomialfordelt i de to populationer, og vi kan derfor sammenligne parametrene udfra hypotesen : $p_{11} > p_{21}$.

Som vi senere vil komme ind på, er vi nødt til at formulere nul-hypotesen $p_{11} = p_{21}$; må denne forkastes (f.eks. på 5 % niveau), har vi sandsynliggjort vor egentlige (alternative) hypotese.

Vi kunne også have valgt $p_{11} \neq p_{21}$ som vor alternative hypotese (bilag 11)

Vi vil i det følgende beskæftige os med, hvad det egentlig er for nogle hypoteser vi i et givet problem kan teste.

Indenfor den klassiske statistik (bilag 12) som vi har begrænset os til her, må vi først lave en model for de stokastiske variable; at konstruere en model vil sige at gøre sig nogle antagelser om fordelingen af de

stokastiske variable, begrundet i viden om det undersøgte fænomen.

I eksemplet rygere/ikke rygere vil man starte med at antage at kræfttilfældene er binomialfordelt (idet hver person enten får kræft eller ikke); nu er en poissonfordeling ofte nemmere at arbejde med, og da vi ved, at en binomialfordeling kan aproximeres med en poissonfordeling, hvis populationen er stor og sandsynligheden lille (bilag 12) vil vi nok ende med at antage, at antal kræfttilfælde er poissonfordelt.

Udfra stikprøven kan vi nu estimere parametrene i de to fordelinger (rygere og ikke rygere).

Ser vi nu på vor hypotese fra før : $p_{11} = p_{21}$, er det vi tester altså, om parameteren i de to poissonfordelinger er ens; eller sagt på en anden måde : om de to grupper kan antages at stamme fra den samme population. Dette forhold mellem parametre og tests vil blive behandlet eksemplarisk i forbindelse med bla. fredericiaundersøgelsen, men vi skal her hæfte os ved nogle generelle forhold vedrørende valget af parametre.

I ovenstående eksempel ville vi, hvis vi fandt at parametrene ikke var ens, kunne hævde at så er populationerne heller ikke ens, hvilket vil sige at der er forskel på sandsynligheden for at få lungekræft, afhængigt af om man ryger eller ikke.

I forbindelse med luftforureningsundersøgelser er parametrene dog langt fra i alle undersøgelser valgt hensigtsmæssigt, og det kan få store konsekvenser for fortolkningen af vor hypotese.

Den ufuldstændige viden om årsagssammenhængen mellem miljøpåvirkning og kræfthyppighed rejser efter vor mening nogle problemer, bla. med hensyn til formidlingen af resultaterne, men henleder endvidere opmærksomheden mod nogle af de grundlæggende forudsætninger, som nødvendigvis må være opfyldt, hvis statistikken skal an-

vendes på en "hæderlig" måde.

Et spørgsmål, som vi løb ind i adskillige gange underdiskussioner om projektet, var :

Hvis man udfra en stikprøve (se note a) kan finde en signifikant sammenhæng mellem to forhold, kan man da påstå, at der må være denne sammenhæng - uanset om man kan give en rimelig hypotese om sammenhængen begrundet i faglig viden om fænomenet ?

At problemet ikke er trivielt findes der talrige eksempler på; på fig II.1. så vi hvordan statistikken bliver anvendt til at forsøge at finde ud af, hvilke påvirkninger af legemet, som kan medføre øget risiko for kræft; man ser ved første øjekast, at en hypotese om at kød-spisning kan fremkalde tarmkræft synes at kunne bekræftes, uanset der os bekendt (samt iøvrigt beskrevet i artiklen) ikke findes nogen som helst medicinsk begrundelse for en sådan antagelse.

Man ser samtidigt, at hypotesen om at kornspisning kan nedsætte hyppigheden af tarmkræft ligesåvel synes at kunne blive bekræftet, og man spørger uvægerligt sig selv, om man ikke med denne form for statistik ville være i stand til at påvise de mest utrolige sammenhænge mellem fænomener som ikke har det ringeste med hinanden at gøre, (se note b).

Hvis hypotesen ikke er dannet på baggrund af faglig viden om fænomenet, er undersøgelsen nærmest af akademisk interesse, (hvem ville på baggrund af fig II.1. nægte at spise kød/fylde sig med korn), idet alle mulige faktorer, kendte som ukendte, ville kunne forklare sammenhængen lige så godt; dette forhold medfører automatisk, at en sådan test ville kunde angribes fra mange synsvinkler, således at det ville være umuligt at træffe nogen beslutning på baggrund af testen.

Ovenstående kunne også formuleres : Hvis vi vælger nogle parametre (her kødspisning), hvor vi ikke kender en

fagligt begrundet årsagssammenhæng, kan vi heller ikke gøre os håb om at testen belyser nogle relevante sider af problemet, samtidig med at fortolkningen af resultatet bliver suspekt.

Man må dog her ikke se bort fra den grundforskningsmæssige interesse det kan have, udfra et givet materiale at få ideer til relevante sammenhænge, som så kan danne grundlag for senere præcisering, såfremt de bliver sammenholdt med faglig viden, (se note c).

Et andet fundamentalt princip, som er vigtigt at holde sig for øje i kræftforskningen er :

Man må ikke danne sig en hypotese udfra en stikprøve, og bagefter teste denne i samme stikprøve, (se note d).

At dette princip skal overholdes findes ved en smule overvejelse ret logisk, men vi mener det kan volde en del problemer i praksis : Hvordan kan vi f.eks. kontrollere, om forskeren rent faktisk har dannet sig hypotesen inden stikprøven blev foretaget ?

Fig II.1. kunne tænkes (uden at vi dog på nogen måde ved noget om dette) at være opstået på følgende måde : Der har foreligget en hel masse tal over først og fremmest kræftdødeligheden i forskellige lande; en forsker har sat sig ned og sammenlignet tallene med tal over en masse andre forskelle i materialet, og nogle af disse (i hvert fald i første omgang) ubegrundede sammenhænge har vist sig at indicere en signifikant sammenhæng; herefter har han testet for denne sammenhæng, og fundet den signifikant.

En sådan fremgangsmåde overtræder begge de to principper vi har nævnt : for det første er hypotesen ikke baseret på faglig indsigt - for det andet testes hypotesen i det materiale, hvori den blev dannet.

Vi har hørt rygter (som vi dog ikke har fundet det umagen værd at checke efter) om, at det på denne måde skulle være muligt at påvise en signifikant sammenhæng mellem storkenes faldende antal i Danmark, og den faldende fødselsprocent.

At de ting vi har beskæftiget os med i dette afsnit måske for en statistiker er ret elementære ting, gør dem imidlertid ikke mindre interessante, idet statistikken langt fra altid kun bruges af fagfolk, samtidigt med at vi alle hver dag bombarderes med statistiske opgørelser over de mest utrolige ting.

Et andet princip som har betydning i kræftforskningen er :

Jo større testmateriale vi har, jo større bliver sikkerheden af testen, (se note e).

Dette princip kan imidlertid udtrykke flere forhold; det er således det samme princip, som ligger bag udsagnet :

Jo flere og større fejlkilder, jo større testmateriale er nødvendigt, samt udsagnet :

Jo mindre effekt der testes for, jo større testmateriale er nødvendigt, (se note f).

Vi har i indledningen gjort rede for, at kræft må antages at opstå fra en række miljøpåvirkninger, som hver især må tænkes at yde deres lille bidrag; det vil derfor kun være muligt at påvise de ret store effekter, (note g), og note h).

Testmaterialet, dvs. statistikker over dødsårsager, sygejournaler m.m. er i almindelighed ret stort, der indgår for det meste adskillige tusinde mennesker, men til gengæld er det for det meste behæftet med stor usikkerhed, f.eks. usikkerhed i diagnostiseringen af dødsårsagen, fejl eller mangler i spørgeskemaer m.m.

Endvidere må man ofte på grund af manglende statistik "sjusse" sig til nogle tal - f.eks. i stedet for at bruge luftmålinger, må man som mål for forureningen i en by anvende statistik over det samlede brændstofforbrug, (36).

Vi skal til slut se generelt på hvad det vil sige, at teste en given hypotese.

et er et generelt princip indenfor statistiske tests, (måske indenfor al naturvidenskabelig forskning), at en hypotese aldrig kan bevises - man kan kun forkaste en (nul) hypotese, (se note 1)

Men hvad vil det sige at forkaste en nulhypotese ?

Her fortæller den klassiske statistik os, at hvis vi kender den hypotetiske fordeling, kan vi opstille en teststørrelse, hvis fordeling vi kender og som er tabelleret; udfra tabellen kan vi så finde sandsynligheden for at vort sample kan stamme fra den hypotetiske fordeling - dvs. det vi i virkeligheden aflæser i tabellen, er sandsynligheden for at få en værre teststørrelse (se note a).

Vi kan nu f.eks. vælge, at hvis denne sandsynlighed (for at få en værre teststørrelse - eller for at få et mere ekstremt sample under hypotesen) er under 5 %, vil vi regne vor hypotese for bevist (sandsynliggjort).

Vi vover i virkeligheden den påstand, at ekstreme hændelser ikke forekommer i praksis - hvilket de selvfølgelig gør, og det er netop denne risiko for at begå fejl, som er det eneste argument vi kan bruge i fastlæggelsen af vort signifikansniveau.

Vi kan se, at der i virkeligheden er to typer fejl vi kan begå : type I fejl er, hvis vi forkaster vor nulhypotese, selvom den er sand; type II fejl er, hvis vi ikke forkaster den, selvom den er falsk.

Det vi skal gøre ved fastlæggelse af vort signifikansniveau, er altså at vurdere konsekvenserne af de to typer fejl i forhold til vore hypoteser.

Hvilket signifikansniveau skal vi da vælge i epidemiologiske undersøgelser ?

Man kunne argumentere, at hvis vi sørgede for at gøre sandsynligheden for type I fejl stor, f.eks. ved at vælge 10 % signifikansniveau, ville vi opnå at gøre det lettere at påvise en given hypotetisk sammenhæng (f.eks. forurening/kræft); vi ser imidlertid, at vi derved sam-

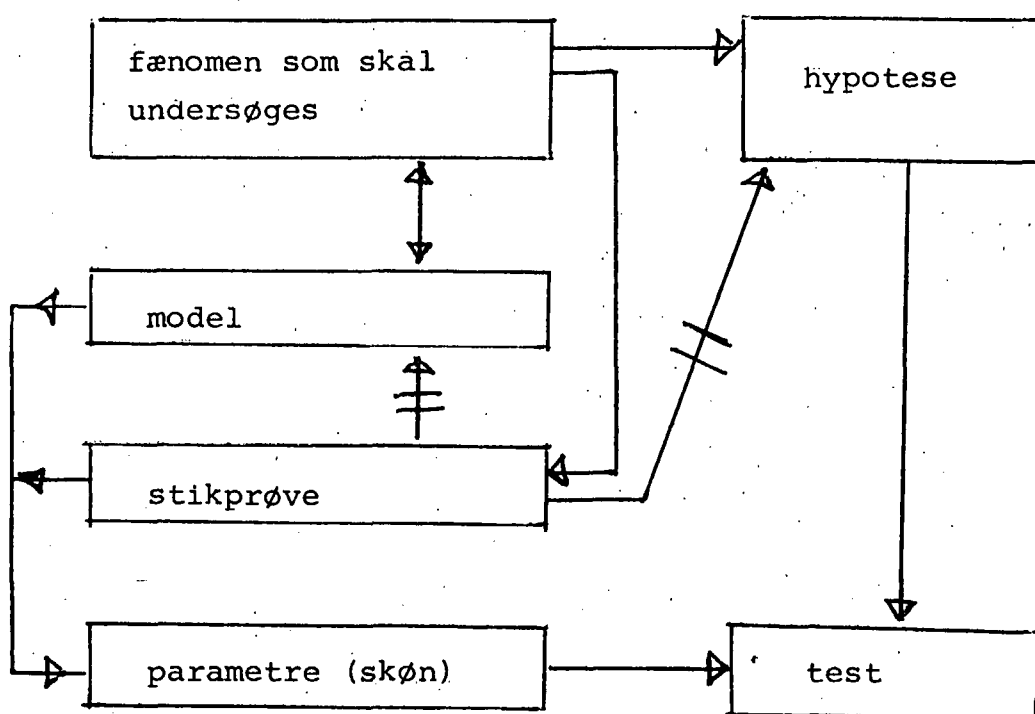
tidig øger muligheden for at begå type II fejl, hvilket vil medføre, at vi øger muligheden for at påvise alle mulige falske sammenhænge, (f.eks. kødspisning og kræft).

Hvis vi derimod i vor iver efter ikke at acceptere falske hypoteser gør vor type I fejl meget små (f.eks. ved at vælge 2 1/2 % niveau), opnår vi samtidigt, at gøre det svært at påvise sande sammenhænge, (se note j). Den eneste måde at løse dette dilemma på er altså i hver given sammenhæng at vurdere udfra materialet, hvilket signifikansniveau som bør anvendes.

Ve mener derfor det er vigtigt at angive den faktisk fundne sandsynlighed fra testen, således at det overlades til dem med faglig indsigt i problemet at fortolke resultatet, (se note k).

I nedenstående skema ses nogle af de ting vi har forsøgt at gøre rede for i dette afsnit:

Pilene viser slutningsrækkefølgen i almindeligt anerkendt statistisk inferens; de overstregede pile viser, hvor det ikke er tilladt at slutte noget fra og til - vi ser, at både model og hypotese ideelt set skal konstrueres udfra viden om fænomenet.



III UNDERSØGELSER AF SAMMENHÆNGEN MELLEM LUNGEKRÆFT OG LUFTFORURENING.

Begrundelse for afsnittet.

Begrundelsen for at beskæftige sig med et afsnit af denne type er dobbelt. For det første et ønske om at spore en argumentation for hypotesen: "Der findes en sammenhæng mellem grader af luftforurening og antallet af observerede lungekræfttilfælde" i en given formulering og for det andet at se, hvordan forskellige forskere har designet undersøgelser af denne type. Hvad den første begrundelse angår, er der ikke, i de artikler vi har behandlet, argumenteret medicinsk for hypotesens relevans. Dette kan hænge sammen med flere ting. Bl.a. at det kun har været muligt at påvise en forøget kræfthyppighed hos dyr, der er eksponeret for enkelte af de stoffer, der forefindes i forurenede luft, men ikke at forklare medicinsk, hvad det er for egenskaber ved netop disse stoffer, der fremkalder kræften, endog hvad kræft i det hele taget er. Denne diskussion vil det nok være muligt at gå i dybden med, men relevansen heraf i et matematikprojekt er tvivlsom. Derfor har vi, som de undersøgelser, vi behandler, accepteret det rimelige i at antage en sammenhæng mellem lungekræft og luftforurening.

Den anden begrundelse er mere relevant for projektet, idet et mål med arbejdet er at undersøge, hvorledes undersøgelser indenfor kræftforskningen designes. Hermed tænkes hovedsageligt på hvilke metoder og tests, der lægges til grund og på konklusionernes afhængighed af den valgte metode.

Selve afsnittet er bygget op omkring to udvalgte artikler, der ganske vist har et par år på bagen, men som på udmærket vis illustrerer problemerne med valg af metode.

Valget af netop de behandlede artikler skyldes ikke så

meget faglige hensyn som forhåndenværende materiale. Desværre er det ikke muligt i et projekt af et semesters varighed at spore samtlige relevante kilder, således at sammenhæng i artikelvalg kan opnås. Derfor har vi ud fra mængden af hjemskaffede artikler valgt to, der indeholder relevante problemstillinger. Endvidere har vi for at belyse egnetheden af indikatorstoffer for luftforurening behandlet nogle få artikler herom. Afsnittet er struktureret på følgende måde: først resumeres de to undersøgelser om sammenhængen mellem udvalgte parametre og lungekræftdødeligheden, og statistikken i de enkelte artikler behandles særskilt efter hvert resume. Dernæst diskuteres de statistiske metoders anvendelighed på det opstillede problem, og herunder omtales artikler om forskellige stoffers egnethed som luftforureningsindikatorer. Afsnittet afsluttes med en delkonklusion.

De to undersøgelser.

Den første undersøgelse, vi vil omtale, er af Donald O. Anderson (04), og den vil kort blive resumeret: Forfatteren tager udgangspunkt i en bestemt klassifikation af luftpollutanter, nemlig: partikulært stof, irriteranter, oxidanter og systemgifte. Artiklen gennemgår emner indenfor disse grupper, og vi har kun koncentreret os om polycykliske hydrocarboner (PH). (I rapporten er disse polycykliske hydrocarboner i flæng benævnt PH og PAH (A for aromatiske)).

Forfatteren indleder afsnittet med, at PH, mange med carcinogen aktivitet, kan findes i byluft. Graden af forekomst afhænger dog ikke blot af populationstæthed, men også af årstid, idet deres kemiske aktivitet er højt afhængigt af vejret. Dette, mener forfatteren, gør, at konklusioner på sammenhænge mellem kræft og koncentrationer af PH i luften kun vanskeligt kan drages.

Herefter omtales en strid mellem bl.a. forfatteren og Stocks. Stocks tillagde i 1955 en byluftkoncentration af 3:4-benzpyrene på 7.7 microgram pr 100 cm³ luft 50% af lungekræfttilfældene, og i en senere undersøgelse fra 1958 brugte han koncentrationer af 3:4-benzpyrene, 1:12-benzperylene, pyrene, flourethan, SO₂ og røg til at vise en korrelation mellem aldersstandardiserede dødsrater (note 1, side 75) blandt mandlige lungekræfttilfælde og alle disse substanser. (3:4-benzpyrene og benzo(a)pyrene bruges i flæng om det samme stof).

Stocks data er af bl.a. forfatteren blevet re-eksamineret med trinvis lineær regressionsanalyse for at udskille de faktorer, som leverer mest til variansen om regressionen. Denne analyse viser, at selv om SO₂ og PH er statistisk signifikante leverandører til multipel korrelationskoefficienten, er denne leverance infinitesimal i sammenligning med effekten af populationstæthed.

Stocks udvidede sin 1958-undersøgelse ved at inddrage et "socialt indeks" og viser ved partiel regressionsanalyse, at røg og 3:4-benzpyrene er relateret til lungekræft-rater, mens røg og 1:12-benzperylene kun er relateret til bronchitis og pneumonia-rater. Igen er Stocks data blevet re-analyseret med trinvis regressionsanalyse med henblik på at undersøge følgende uafhængigt variable: 3:4-benzpyrene, 1:12-benzperylene, pyrene, flourethan, populationstæthed og social klasse. Konklusionen er, at ingen PH fremstår som signifikant korrelerede med lungekræft, hvis røg og socialt indeks ikke medtages, selv om 3:4-benzpyrene er korreleret med andre kræfttyper. Derfor mener forfatteren ikke, der er belæg for at tro, at koncentrationer af PH i byluft er relateret til kræftdødsfald. Forfatteren nævner tobaksnydelse som en væsentlig årsag til lungekræft og slutter afsnittet med konklusio-

nen:

"...selv om man kan observere carcinogene PH i byluft, er der endnu ingen klar indikation for, at deres tilstedeværelse er signifikant, selv om deres samspil med biologiske og fysiske substanser må være mekanismen ved carcinogenesis. Disse andre substanser må, gennem deres irritant-egenskaber og svækkende effekt på det respiratoriske epitelforsvar i værten, skabe miljøet for den kumulative aktion af tilbageholdte carcinogener, som også findes i den omgivende luft. Sådan en hypotese er, trods alt, stadig meget spekulativ."

Det i denne sammenhæng interessante er uoverensstemmelsen mellem Stocks og Anderson. På det samme materialegrundlag opnås modstridende konklusioner afhængigt af analysemetode. Af denne grund har vi ønsket at præcisere, hvilke metoder parterne har benyttet. Præciseringen er vanskeliggjort af manglende originalmateriale, men vi har alligevel søgt at rekonstruere analyserne. Rekonstruktionerne skal naturligvis tages med et vist forbehold.

Stocks har ønsket at undersøge, om der er en sammenhæng mellem aldersstandardiserede dødsrater (forårsaget af lungekræft) og 3:4-benzpyrene, 1:12-benzperylene, pyrene, flourethan og røg samlet på 17 lokaliteter.

Metoden er korrelationsanalyse, men det fremgår ikke af artiklen hvilken type. Derfor skal her ridses op, hvilken mulighed vi kan forestille os.

Stocks har fundet korrelationskoefficienten mellem de opstillede parametre hver for sig og den aldersstandardiserede dødsrate stammende fra lungekræft.

Korrelationskoefficienten kan udtrykkes på følgende måde:

$$\rho = \frac{\text{cov}(x_1, x_2)}{\sqrt{\text{var}(x_1) \text{var}(x_2)}}$$

og skønnet over parameteren ρ er

$$r = \frac{s_{12}}{s_1 s_2}$$

hvor s_{12} betegner skønnet over kovariansen og kan til beregningsformål skrives på formen

$$r = \frac{\text{SAP}_{x_1 x_2}}{\sqrt{\text{SAK}_{x_1} \text{SAK}_{x_2}}}$$

Af korrelationskoefficienternes fordeling kan udledes et simpelt test for hypotesen $\rho=0$. Såfremt $\rho \neq 0$, er størrelsen

$$t = \frac{r}{\sqrt{1-r^2}} \sqrt{f}$$

t-fordelt med $f=n-2$ frihedsgrader.

Endiskussion af korrelationsanalysens anvendelighed vil blive fremført i et senere afsnit, hvor diskussionen af alle de behandlede metoder foreligger samlet.

Anderson har foretaget en re-eksamination af Stocks' data ved hjælp af trinvis regressionsanalyse. I vores kildemateriale (19) skelnes mellem forskellige analyser, der med god vilje kan kaldes trinvis. Trinvis regressionsanalyse er beskrevet i bilag 8, og der er grund til at antage, at det ikke er denne metode, Anderson har benyttet, idet multipel korrelationskoefficienten R^2 (bilag 7) ikke indgår i denne. Derimod indgår R^2 i "the forward selection procedure", og vi antager, det er denne Anderson har benyttet.

Fremgangsmåden er meget lig den, der er beskrevet for trinvis regressionsanalyse, idet der dog ikke i "the forward selection procedure" re-eksamineres på hvert trin af regressionen for de variable, der allerede er indsat i modellen på det foreliggende trin. Det vil sige, at en variabel, der først er indsat i ligningen, ikke kan risikere at blive forkastet på et senere tidspunkt.

De variable indsættes i ligningen efter samme kriterier som under trinvis regressionsanalyse. Når den første variabel er indsat, findes de mindste kvadraters ligning, og et F-test (bilag 4) viser, om regressionsligningen er signifikant. Den anden variable findes og indsættes, og igen findes de mindste kvadraters ligning. For denne ligning findes R^2 i procent, og det testes om ligningen er signifikant med et F-test.

For den ny-indsatte variabel undersøges, om den forårsager en signifikant formindskelse af den residuelle kvadratsum ved dens partielle F-værdi (bilag 5). Således fortsættes, indtil den partielle F-værdi for en variabel ikke er signifikant, og den variable må afvises i regressionsmodellen.

Stocks har udvidet antallet af variable ved at indføre et socialt indeks og har denne gang benyttet, hvad han kalder partiel regressionsanalyse. Denne metode antager vi er en almindelig multipel regressionsanalyse (bilag 3) med et efterfølgende test for, hvorvidt de enkelte koefficienter kan antages at være forskellige fra nul. Hvor dette er tilfældet tales om korrelation mellem den pågældende variabel og responsen.

Andersons reanalyse benævner han selv trinvis regressionsanalyse, men som tidligere nævnt antager vi, der er tale om "the forward selection procedure" (se omtale af denne).

Den anden undersøgelse, der skal omtales her, er af B.W. Carnow og P. Meier (11), og også denne vil blive resumeret:

Forfatterne har dels på baggrund af eget arbejde, dels på grund af andres, undersøgt sammenhængen mellem luftforurening, indiceret ved benzo(a)pyrene, og lungekræft ved at sammenligne byområde- med landområde-, migrant- med ikke-migrant- og ryger med ikke-rygerpopulationer.

Effekten af fast brændstof og cigaretter i 19 lande og af benzo(a)pyrenkoncentration og cigaretter i 48 tilstødende stater i USA undersøgtes ved benyttelse af multipel regressionsanalyse.

Forfatterne fokuserer på 4 typer af studier:

- 1) Sammenligning mellem bypopulationer og landpopulationer og undersøgelse af forskelle i dødsrater forårsaget af lungekræft, i de fleste tilfælde uden at berøre specifikke ætiologiske faktorer.
- 2) Sammenligning mellem lungekræft-dødsrater i migrantpopulationer af rater i hjemlandet med rater i modtagerlandet. Dette udtaler noget om forandring i dødsrater ved ændringer i miljøet.
- 3) Sammenligning mellem demografiske enheder og undersøgelser af sammenhængen mellem lungekræft-dødsrater og forskellige indicier for forurening ved brug af multipel regressionsanalyse for at adskille effekterne af miljømæssige- og andre faktorer.
- 4) Sampling studies, hvor karakteristikker af lungekræft-døde bestemmes ved interview med familie og sammenlignes med korresponderende karakteristikker af den resterende population.

Efter præsentation af disse fire typer undersøgelser, som alle behandles i artiklen, begrundes valget af benzo(a)pyrene som forureningsindikator. Benzo(a)pyrene forekommer i fast form i luften, er sædvanligvis adsorberet på partikler og derfor filtrerbar, er relativt let at måle og er godt korreleret med andre polycykliske organiske stoffer. Endvidere er det fundet carcinogent i dyr og antages at være det i mennesker. Herefter omtales undersøgelser af benzo(a)pyrene, der viser en kraftig årstidsvariation både i by- og landområder med et gennemsnitsindhold i byer på 6.6 microgram/1000 m³ luft og 0.4 microgram/1000 m³ luft i landområder.

1) By/land undersøgelser.

Forfatterne gennemgår undersøgelser af denne type. I mange er der ikke medtaget udtryk for forurening, og i de fleste er der ikke taget hensyn til cigaretter. De er altså rene sammenligninger mellem by- og land-dødsrater. Undersøgelsesernes konklusioner resumeres, og forfatterne laver til sidst en delkonklusion, der udtaler, at en 100% forøgelse i lungekræft-dødsrate hænger sammen med en forøgelse på 6.2 enheder benzo(a)pyrene, eller omregnet en omtrentlig forøgelse på 15% i dødsfald pr. enhed forøgelse i benzo(a)pyrene. Imidlertid er den før omtalte forøgelse på 100% fundet mellem de stærkest forurenede byer og de reneste landområder. Derfor er forureningseffekten sandsynligvis mindre end 15%.

2) Migrantundersøgelser.

Forfatterne peger selv på en svaghed ved disse undersøgelser, nemlig om migranter kan formodes at være ækvivalente til random eller repræsentative prøver af hjemlandspopulationen.

En gennemgang af flere undersøgelser af denne type munder ud i følgende delkonklusion: lungekræft-dødsrater hænger sammen med det generelle forureningsniveau i de undersøgte områder. Generelt viser undersøgelserne at migranternes dødsrate ligger mellem afgiver- og modtagerlandets. Dataene peger også på, at udsættelse for forurening tidligt i livet forårsager vedvarende effekter, dog mangler dataene stadig effekten af tobaksrygning. En enkelt undersøgelse, der inkluderer rygevaner, giver imidlertid basis for at konkludere, at forandringer i dødsrater er relateret til mere end tobaksrygning, dvs. til en byfaktor.

3) Regressionsanalyse.

Dette afsnit er tyngdepunktet i artiklen, idet forfatterne her har lavet en selvstændig undersøgelse. De har lavet multipel regressionsanalyse, fordi de har

ønsket at separere effekter af faktorer, der adskiller by- og landområder, med det sigte at identificere by-faktorer, som må gøres ansvarlige for forskelle i dødsrater fra kræft i lungerne.

Metoden har imidlertid nogle begrænsninger, hvoraf forfatterne nævner følgende som de vigtigste: 1) selv om relationen har den forventede form, kan fluktuationer i dataene, der skyldes tilfældige effekter af andre variable, gøre det vanskeligt at bestemme sammenhængskoefficienterne med rimelig præcision. Hvis det ikke lykkes at finde et signifikant resultat, er det ikke nødvendig et bevis for, at der ikke er sammenhæng. 2) Den camouflerende effekt som skyldes en variabel under et andet prædikat. Specielt når variable er højt korrelerede (som ved bymæssig crowding og forurening), og når der er væsentlige målefejl (som ved måling af pollutanter i luften), vil relationskoefficienten være vanskelig at tolke, og medtagning af en variabel i ligningen vil have effekt på koefficienterne for de andre.

Forfatterne har, i stedet for at bruge regressionsmetoder til at udvælge, hvilke sæt af forureningsvariable, der er mest direkte relateret til lungekræft, valgt et enkelt forureningsindeks, benzo(a)pyrene, til at repræsentere effekterne af alle korrelerede forureningsvariable kombineret.

Forfatterne har suppleret en tidligere undersøgelse i 19 lande af sammenhængen mellem et antal variable og lungekræftdødelighed. De har valgt cigaretforbrug og fast brændstofforbrug. Tidligere undersøgelser har nemlig vist, at bl.a. flydende brændstof ikke er relateret til lungekræft eller andre forureningsindikatorer.

Regressionen er af formen

$$Y = a + b_1 X_1 + b_2 X_2$$

hvor Y = alders-køns-specifikke lungekræft-dødsrate,

X_1 = cigaretforbrug i 1000 cigaretter/år/person over 15

år og X_2 =fast brændstofforbrug i 1000 kg/person/år.

b_1 udtrykker forøgelsen i dødsrate pr. enhed forøgelse i cigaretforbrug, og b_2 udtrykker forøgelse pr. enhed forøgelse i fast brændstofforbrug.

b_1 findes til ca. 15% og b_2 til ca. 20%. Forfatterne har ikke mulighed for at udtrykke koefficienterne i ækvivalente benzo(a)pyrene enheder.

Konklusionen på undersøgelsen er, at produkterne fra fast brændstofforbrænding, eller af variable højt korrelerede med fast brændstof, er en vigtig ætiologisk faktor i forbindelse med lungekræft.

Forfatterne sammenligner derefter deres undersøgelse med andre og slutter med at konkludere, at rimelige forøgelser i lungekræfthyppighed er 2-3% pr. benzo(a)pyrene enhed for hvide og 7-8% for ikke-hvide.

4) Sampling studies.

I denne type kan værdier af variable som rygevaner og bopæl bestemmes for hvert enkelt individ, og tilliden til tolkning af resultaterne vil være tilsvarende større end i regressionsanalysen.

Forfatterne gennemgår en del undersøgelser, der stort set blot bekræfter de tidligere konklusioner. Som nyt element kan dog nævnes en inddeling efter cigaretforbrug.

I de afsluttende kommentarer peger forfatterne på svagheder ved de omtalte undersøgelser generelt og peger på områder, hvor videre forskning er nødvendig. Endelig består deres konklusion af seks delkonklusioner, som skal nævnes:

- 1) Lungekræft er blevet den største enkeltårsag blandt kræftdødsfald hos mænd og en signifikant dødsårsag blandt kvinder i USA, og den udbredelse er øget i de sidste 30 år.
- 2) En væsentlig ætiologisk faktor udgør cigaretrykning, selv om rygning ikke kan gøres ansvarlig for hele

den øgede udbredelse af sygdommen.

- 3) Polycyklisk organisk stof, der findes i cigaretrøg i høj koncentration, forårsager lungekræft (og anden kræft) hos forsøgsdyr og findes i luften i store mængder i industrier, hvis arbejdere har høj lungekræft-dødsrate, ligesom det findes i byluft.
- 4) Byområder har omtrent dobbelt så høj udbredelse af lungekræft som landområder. Indenfor byområderne er udbredelsen større, hvor der er mere generel industriel forurening.
- 5) Generelt har migrantpopulationer en udbredelse af lungekræft mellem den, der er bestemt i deres oprindelige hjemland og den i det land, de er tilflyttet.
- 6) Et estimat på den effekt, forurening øver på lungekræft-dødsrater, er 5% forøgelse pr. enhed forøgelse i byforurening, indiceret ved benzo(a)pyrene.

Statistikken i denne undersøgelse fremgår af resuméet, og i bilag 1 er den generelle metode omtalt, ligesom de aktuelle beregninger er udført i bilag 2.

Statistikken vil blive behandlet i den efterfølgende diskussion, og her skal blot nævnes, at hvis man tester de fundne koefficienter i forhold til hypotesen $\beta_1, \beta_2 = 0$, findes t-værdier på henholdsvis 0.932 og 1.893, hvor det for den første værdi vedkommende gælder, at der ikke er grund til at antage koefficienten b_1 forskellig fra nul på 5% niveau.

Testes under hensyntagen til korrelationen mellem b_1 og b_2 , findes en F-værdi på 2.958, hvilket ikke giver statistisk signifikans på 5% niveau.

Diskussion.

I Andersons artikel finder Stocks ved korrelationsanalyse korrelation mellem 3:4-benzpyrene, 1:12-benzperylene, pyrene, flourethan, SO_2 , røg og standardiserede dødsrater forårsaget af lungekræft blandt mænd. Endvi-

dere findes en højere korrelation mellem lungekræft og røgkoncentrationer.

Ved re-eksamination med "the forward selection procedure" findes, at selv om SO_2 og polycykliske hydrocarboner giver et statistisk signifikant bidrag til multipel korrelationskoefficienten, er dette bidrag infinitesimalt i sammenligning med effekten af populationstæthed.

Kort sagt har vi her et sæt data, der giver anledning til forskellige konklusioner afhængigt af analysemetode.

Beregning af korrelationskoefficienten forudsætter, at iagttagelsesresultaternes fordeling er normal eller kan transformeres, således at den bliver normal. Denne betingelse er ikke altid opfyldt, men ved indsamling af data er det dog ofte muligt at fremskaffe et iagttagelsesmateriale, hvor begge variable har varieret tilfældigt under så vidt muligt ens forhold. Ved beregning af korrelationskoefficienten og ved efterfølgende test for hypotesen $\rho=0$, kan det lykkes at påvise en stokastisk sammenhæng mellem de to variable.

Det må understreges, at en stokastisk sammenhæng ikke er ensbetydende med en kausal sammenhæng. En signifikant korrelationskoefficient opfordrer altså til videre analyse af det pågældende problem for at finde ud af årsagssammenhængen. I denne videregående analyse bygges først og fremmest på faglig indsigt i problemet, og regressionsanalyse spiller ofte en væsentlig rolle i verifikationen af de opstillede hypoteser. Altså vil en korrelationsanalyse i sig selv sjældent være tilstrækkelig.

Endvidere vanskeliggøres tolkningen af korrelationskoefficienten ofte af, at andre faktorer end de iagttagne har varieret og dermed frembragt en falsk korrelation. Hvor dette er tilfældet, må materialet inddeles i undergrupper, indenfor hvilke de forstyrrende fakto-

rer ikke varierer, og korrelationskoefficienten må beregnes for hver undergruppe.

Sluttelig skal nævnes, at såfremt iagttagelsesresultaternes fordeling afviger betydeligt fra den todimensionale normale fordeling, er fortolkningen af korrelationskoefficienten højst usikker. Eksempelvis kan man komme ud for, at $\rho=0$ ikke er identisk med stokastisk uafhængighed.

Derimod er "the forward selection procedure" grundlæggende en god idé, idet den tillader at arbejde med flere variable end nødvendigt, da de variable testes i ligningen på hvert trin. Det havde dog nok været en forbedring af analysen at benytte trinvis regressionsanalyse, idet der her undersøges, hvilken effekt introduktionen af en ny variabel har på den rolle, en tidligere indsat variabel spiller.

Alt i alt må Andersons fremgangsmåde anbefales fremfor 'Stocks'. Det må dog fremhæves, at det er begrænset, hvilken nytte man kan have af de dragne konklusioner. I kræftforskningen må et af formålene være at kunne påvise en sammenhæng mellem indikatorer for luftforurening og lungekræftdødelighed (mht indikatorers egnethed se senere). Benyttes en variabel som befolkningstæthed, der ganske vist hænger sammen med den generelle luftforurening, vil det ofte være sådan, at denne variabel overskygger alle andre i ligningen. Dette er der i sig selv intet unaturligt ved, idet befolkningstæthed er godt korreleret med stort set alle kræftfremkaldende stoffer i forurenede luft, hvorimod en enkelt indikator som f.eks. benzpyrene, vil være dårligt korreleret med eksempelvis PVC, som også har effekt på kræftdødeligheden.

For at nå frem til pointen mener vi ikke, det er særligt hensigtsmæssigt at inddrage befolkningstæthed i sådanne undersøgelser. Hvordan skal de ansvarlige myndigheder forholde sig til en undersøgelse, der konkluderer

derer, at lungekræftdødeligheden stort set kun er godt korreleret med befolkningstæthed? Spredte befolkningen eller honorere et lavt fødselstal?

Undersøgelser af denne type er blot med til at bevare status quo og lægge en dæmper på den folkelige opinion til fordel for forureningskilderne.

En lignende kritik kan gives af Stocks' multipel regressionsanalyse og Andersons reanalyse med "the forward selection procedure".

"The forward selection procedure" er en videreudvikling af multipel regressionsanalyse, og er efter alt at dømme en mere pålidelig beskrivelse af tingenes tilstand, og derfor må Andersons konklusioner tillægges større vægt. Desværre, kunne man sige, fordi Andersons ærinde fremgår med tydelighed af hans samlede artikel, nemlig at undervurdere eller rettere underkende forurenede lufts indflydelse på vores helbred. Det problem, der efter vores skøn fremstår, er at få sat grænseværdier på tilstedeværelsen af kræftfremkaldende stoffer i luften. Det vil bl.a. sige, at forskningen i at finde et eller flere stoffer, der er godt korrelerede med størstedelen af tilstedeværende kræftfremkaldende stoffer, må intensiveres. Lykkes det at finde et godt indikatorstof, hvis koncentration i luften helt tydeligt er korreleret med lungekræftdødeligheden, vil der blot være en "cost-benefit"-analyse tilbage for at få sat en grænseværdi på forurenede luft.

Et skridt i denne retning mener vi, Carnow og Meier er gået. De har benyttet benzo(a)pyrene som indikator. Deres begrundelse var bl.a., at benzo(a)pyrene er godt korreleret med andre polycykliske organiske stoffer. Desværre indeholder deres artikel ikke kildehenvisninger til denne påstand, men vi har selv søgt litteratur, der kunne påvise benzo(a)pyrenes egenskaber som indikator.

Vi har fundet to artikler, der berører emnet perifert.

Den ene er af W.J. Blot et al. (08) og belyser sammenhængen mellem forskellige typer kræft og forefindelse af olieindustri i udvalgte områder i USA.

Der er ikke skelnet mellem forskellige stoffers tilstedeværelse, blot ved man, at polycykliske aromatiske hydrocarboner (PAH) findes i højere koncentrationer, hvor olieindustri forefindes. Resultatet af undersøgelsen er bl.a., at lungekræft var signifikant hyppigere i områder med olieindustri på 0.01 niveau.

En anden undersøgelse af Lee et al. (32) beskæftiger sig med analyse og bestemmelse af sammensætningen af de PAH-fraktioner, der kendes fra forbrænding af tre almindelige brændstoftyper, nemlig kul, træ og olie/petroleum. Endvidere sammenligner de fordelingen af PAH fundet ved de tre forskellige brændstoftyper med fordelingerne af partikulært stof fundet i luft fra et højt kulforbrugende område (Indiana) og et lavt kulforbrugende område (Massachusetts) for at se, om der kan laves korrelationer mellem brændstofudnyttelse og partikulært stof i luft. PAH's kendte bestanddel i partikulært stof i luft er mistænkt for at være hovedårsag til den observerede carcinogene aktivitet i områder med tung generel industri, og mange specifikke PAH vides at være carcinogene i dyr.

Disse effekter sammenholdt med en formodning om, at luftbåret PAH deltager i induktionen af lungekræft i mennesket, er begrundelsen for PAH som forskningsobjekt.

Af andre ting kan nævnes, at PAH stammer fra forbrændingsprocesser, der udnytter fossilt brændstof, og dannelsen er afhængig af forskellige forbrændingsparametre som f.eks. brændstoftype, temperatur og brændstof-til-luft ratio.

Endvidere skyldes ca. 90% af PAH-udsendelsen kulforbrændingsprocesser, og hovedparten af disse er koncentreret i midt-vestlige stater i USA.

Som resultat af undersøgelsen findes en bemærkelsesværdig overensstemmelse mellem kurver for kul-sod PAH og

Indianapolis luftpartikler så vel som mellem olie/petroleum-sod PAH og Boston luftpartikler. Denne sidste iagttagelse er konsistent med den kendsgerning, at olie er Bostons hovedvarme- og -energikilde, mens kulafbrænding praktisk talt ikke finder sted. Svovlholdige PAH i kulforbrændingsprodukterne er identiske med de, der er fundet i luftpartikler fra Indianapolis, og dette leder frem til den antagelse, at hovedleverandør til Indianapolis' luftbårne PAH er kulforbrændingsprocesser.

Denne sidste undersøgelse prøver altså ikke at finde et indikatorstof, men peger indirekte på, at tal for brændstofforbrug kan benyttes som indikator for luftens indhold af PAH.

Man fristes til at indskyde, at dette er en fjer i hatten på A-kraft tilhængere, dog må man have in mente, at selv om brændstofforbruget er godt korreleret med mængden af carcinogene PAH i luften, er det næppe godt korreleret med ioniserende stråling fra diverse kilder.

For at vende tilbage til Carnow og Meiers artikel finder vi selve idéen i undersøgelsen god, dog mangler der stadig dokumentation for benzo(a)pyrenes egenskaber som forureningsindikator.

Delkonklusion.

I undersøgelser, der har som mål at finde en sammenhæng mellem luftforurening og lungekræftdødelighed, vil det ofte være rimeligt at benytte regressionsanalyse i en eller anden form. I regressionsanalysen indgår et udvalgt antal parametre, og udvælgelsen af disse parametre må underkastes faglige kriterier. Kun på denne måde undgås parametre, der ikke indgår i kausal sammenhæng med lungekræftdødeligheden. Stokastiske sammenhænge kan give et fingerpeg om en mulig kausal sammenhæng, men visse parametres stokastiske sammenhæng er dog så indlysende tåbelige, at disse bør forkastes.

Vi mener, at en hensigtsmæssig måde at drive epidemiologisk kræftforskning på, er at finde frem til et eller flere indikatorstoffer for luftforurening, der er godt korreleret med andre kræftfremkaldende stoffer i luften, samtidigt med at der er god korrelation med lungekræftdødeligheden.

Det endelige mål må være at nå frem til acceptable grænseværdier for indikatorstofferne, således at det bliver en politisk beslutning, hvormange det hvert år skal dø af lungekræft og ikke som nu privat-økonomiske interesser.

IV FREDERICIAUNDERSØGELSEN.

Indledning.

Som vi nævnte i forordet, var Fredericiaundersøgelsen en af de motiverende faktorer, der forårsagede dette projekt, og den giver en mulighed for at tage udgangspunkt i en konkret og velbeskrevet statistisk undersøgelse, som netop på grund af resultaterne har været (og er) stærkt omdiskuteret.

Anvendelsen af statistik er langt fra entydig i den forstand, at et givet problem automatisk vil føre til anvendelsen af en given model, hypotese og test, uanset hvem der arbejder med det.

En stor del af arbejdet består af "ekstra statistiske" antagelser og vurderinger såsom valg af kontrolgruppe, præcisering af hypotese m.m. Heri er Fredericiaundersøgelsen særlig interessant, idet talmaterialet er af en sådan beskaffenhed, at tilsyneladende ens hypoteser henholdsvis accepteres og forkastes, afhængigt af hvilken statistisk fremgangsmåde, man gør brug af.

Alle disse ting vil vi se på i dette afsnit, og vi har valgt at strukturere på følgende måde:

Først gives en oversigt over indholdet i den samlede Fredericiaundersøgelse inklusiv debat (kronologisk oversigt findes i bilag 13). Dernæst behandles fem udvalgte artikler/undersøgelser. De to af undersøgelserne behandles dog ikke konsekvent adskilt, da problemstillingen stort set er ens (Clemmesen et al. og E.B. Andersen hhv. (14) og (15)). De øvrige undersøgelser behandles adskilt. Efter hver omtale diskuteres kritikpunkter, der er specifikke for den enkelte artikel, og til sidst kritiseres hele Fredericiaundersøgelsen. Afsnittes afsluttes med en delkonklusion.

Oversigt over Fredericiaundersøgelsen.

Den 30.09.1974 offentliggjortes i Ugeskrift for læger

en undersøgelse af Johannes Clemmesen et al. (14) om lungekræft og luftforurening i Fredericia, der bl.a. konkluderer, at det næppe kan skyldes tilfældigheder, at hyppigheden af lungekræft for mænd i Fredericia er højere end i Horsens, Kolding og Vejle. Denne undersøgelse er behandlet senere i rapporten.

En måned efter (04.11.) behandles problemet i korrespondancer. Peter Ernst (23) spørger, om hyppigheden af lungekræft blandt beboerne indenfor det gamle voldområde er korrigerede for alder og køn, idet det må formodes, at ældre aldersgrupper er overrepræsenteret i bykernen, hvilket kan forklare fundet af øget hyppighed.

Jens L. Hansen (27) kommer ind på rygevaner og nævner, at Fredericias vækst i forhold til kontrolbyerne på grund af industri og trafik måtte ske med tilflyttere fra andre egne. Tilflytterne kom til et fremmed miljø og blev derfor lettere ofre for cigaretrykning. Clemmesen et al. (15) svarer ikke direkte på spørgsmål og kritik, men nævner, at det kun er foreløbige resultater, der er forelagt i Ugeskrift for læger, og at de er uden ansvar for, at diskussionen er påbegyndt offentligt. Man har valgt at fremskynde forelæggelsen af de væsentligste træk i billedet sammen med de mulige forklaringer i en foreløbig rapport, så de ansvarlige myndigheder tidligst muligt kunne tage stilling til, om videregående undersøgelser var nødvendige. I samme nummer af ugeskriftet kritiserer miljøkonsulenterne fra LO og SID (22), at undersøgelsen ser bort fra muligheden af, at der kan være en væsentlig sammenhæng mellem arbejdspladspåvirkninger og ekstern forurening. Man efterlyser analyser af, hvad den ramte gruppe mennesker tidligere har været beskæftiget med, og om der på disse menneskers arbejdsplads har eksisteret og/eller stadig eksisterer forureningskilder, der kan medvirke til dannelse og udvikling af lunge-

kræft.

Endelig har læge Knud Wilson (43) hæftet sig ved, at Clemmesen et al. anbefalede en dyberegående analyse angående lungekræft i forskellige lokaliteter i landet og har fremsendt en lille statistik fra Strandby ved Frederikshavn, som kan have videregående interesse.

25.11.1974 offentliggør E.B.Andersen (01) en reanalyse af lungekræfttilfældene i Fredericia. Denne er behandlet senere i rapporten, og konklusionen skal blot gentages her. E.B.Andersen finder, at det fremlagte materiale ikke udgør nogen solid basis for at slutte, at hyppigheden af lungekræfttilfælde er større i Fredericia end i de tre kontrolbyer. Materialets mest markante træk er et uventet lavt antal lungekræfttilfælde i Horsens.

Arne Nielsen (34), der er en af deltagerne i den oprindelige undersøgelse, tager i samme nummer til genmæle overfor E.B.Andersens kritik. Nielsen argumenterer for Clemmesen et al.'s undersøgelsesmetode fremfor Andersens bl.a. ved, at Andersens test ikke tager hensyn til artiklens formål, nemlig at belyse om lungekræfthyppigheden i Fredericia kan antages at være højere end i andre byer med nogenlunde samme indbyggertal og alderssammensætning, fordi testet behandler alle fire byer på ensartet måde uden speciel hensyntagen til Fredericias særstilling i formålet, og det nævnes, at den i Andersens konklusion omtalte svage, generelle tendens giver en solid basis til at være på vagt, så man som foreslået sætter en større undersøgelse i gang.

Endelig forsøger Aage Lørring (33) at forklare Fredericias højere hyppighed af lungekræfttilfælde. Dette gør han ved, at et stort antal Huguenotter bosatte sig i Fredericia, og at byen og dens omegn endnu er præget af disse gamle slægter, der sikkert har medført avancerede rygevaner til Danmark. Endvidere nævner Lørring, at Fredericia gennem 100 år har været cen-

trum for Danmarks jernbanetrafik, og at et meget stort antal statsbanefunktionærer har haft deres bopæl i netop denne by og heriblandt særlig mange togførere, lokomotivførere og togfyrbødere. Kan man udelukke, at deres lungekræftrisiko er forøget?

02.12.1974 gør Per Eichner (21) fra Superfos opmærksom på, at vurderes luftkvaliteten i Fredericia og Vejle på basis af svovldioxidmålinger, tyder tallene på, at luftkvaliteten i Fredericia er bedre end i Vejle. Endvidere nævnes, at de fleste forskere har fundet det naturligt at sikre sig eller i det mindste sandsynliggøre, at der virkelig var forskel på luftkvaliteten i de områder, de sammenlignede, inden de accepterede luftforurening som en mulig forklaring på forskelle i sygeligheden.

10.02.1975 har E.B.Andersen (02) et indlæg i Ugeskrift for læger, der er et svar til Arne Nielsen. Andersen gentager Fredericiarapportens mest udtalte statistiske svagheder og kommer med en generel kommentar til de konklusioner, der drages. Efter Andersens skøn er det test Fredericiarapporten benytter ikke heldigt valgt, idet den anførte teststørrelse ikke, som anført i artiklen er normeret normalfordelt, hvilket gør det vanskeligt at vurdere dens relative størrelse. Den statistiske vurdering svækkes efter Andersens overbevisning altid af en kraftig inhomogenitet i den valgte kontrolgruppe. Konklusionen på den oprindelige undersøgelse må være, at materialet ikke kan påvise en forskel med nogen stor sikkerhed.

Første del af en større undersøgelse udført af Horwitz og Bach Pedersen (28) offentliggjortes i ugeskriftet 08.09.1975 under titlen "Sammenhængen mellem luftforurening og dødelighed med særligt henblik på lungesygdomme". Artiklen er ligesom anden del omtalt i denne rapport, og her skal blot nævnes, at dødeligheden i Fredericia er sammenlignet med dødeligheden i 6 andre

kommuner med samme indbyggertal.

Angående dødeligheden i de enkelte aldersgrupper indtog Fredericia en mellemstilling mellem de andre kommuner, både med hensyn til dødeligheden af alle årsager og med hensyn til dødeligheden af lungekræft og kronisk bronchitis. Medtages alle dødsfald, hvor kronisk bronchitis var anført som en dødsårsag, primær eller sekundær, var de aldersstandardiserede rater højere i Fredericia for mænds vedkommende på 0.1% niveau.

Anden del offentliggjortes samme sted 16.08.1976 under titlen "Sammenhængen mellem den almene dødelighed og industrialiseringen i en række danske provinsbyer" (36). Denne undersøgelse finder bl.a., at der ikke er sammenhæng mellem SO_2 -emissionen og dødeligheden. Udelukkes en enkelt kommune med særdeles høj SO_2 -emission af analysen, fandtes blandt mænd en positiv sammenhæng med alle lungesygdomme under et samt med lungekræft og kronisk bronchitis. Endvidere var der blandt mænd en positiv sammenhæng mellem tætheden af industrivirksomheder og dødeligheden af lungesygdomme samt en positiv tendens for lungekræft og bronchitis, der dog ikke var statistisk signifikant, og endelig var der en sammenhæng mellem andel af kommunen, der var dækket med by eller bymæssig bebyggelse, og dødeligheden af lungesygdomme, der dog ikke var statistisk signifikant for lungekræft og bronchitis alene.

Januar 1977 offentliggjorde E.B.Andersen (03) en artikel med titlen "Multiplicative Poisson Models with Unequal Cell Rates". Denne er omtalt senere i rapporten. Kort fortalt benyttedes den multiplikative Poisson-model med efterfølgende test til at konkludere, at hypotesen om ens byeffekter for Fredericia, Horsens, Kolding og Vejle ikke kan forkastes.

Horwitz og Bach Pedersen (37) havde 21.11.1977 en undersøgelse i Ugeskrift for læger om lungesygdomme og lungefunktion blandt skolebørn i Fredericia. Denne under-

søgelse viser, at den socio-økonomiske baggrund, vurderet ud fra boligens størrelse, forældrenes skolegang og rygevaner, var lavest i city, højere i forstæderne og højest i landdistrikterne. Langvarig forkølelse var hyppigst i city, astmatiske lidelser og lungebetændelse viste ikke noget entydigt mønster med hensyn til sygdomsmønster og svarhedsgrad, og børnenes højde, vægt og lungefunktion viste ingen systematisk afhængighed af boligområdet.

Johannes Clemmesen (16) mener i en korrespondance 13. 02.1978 ikke, at Horwitz og Bach Pedersens undersøgelse af skolebørnene i Fredericia bør stå ukommenteret. Clemmesens kritik angår både argumenter mod hans egen oprindelige Fredericiaundersøgelse og selve undersøgelsen af skolebørnene.

For det første nævnes, at det ikke var Clemmesen et al., der ansvarliggjorde en enkelt virksomhed for luftforureningen, og det fastholdes, at selvom man ved en sammenligning enkeltvis imellem Fredericia og henholdsvis Horsens, Kolding og Vejle ikke finder en signifikant forskel i lungekræftens hyppighed, kan dette ikke eliminere den påviste signifikans i forskellen mellem Fredericia og de tre nabobyer under et.

Endvidere har Clemmesen selv detaljeret bekræftet den i 1974 påviste ophobning af lungekræfttilfælde i Fredericias centrum i sammenligning med nabobyernes jævne indre fordeling. (Denne påvisning findes i Acta. path. microbiol. Scand. 1977. suppl. 261 og er i skrivende stund ikke hjemkommet).

Det nævnes bl.a., at man i København har påvist, at lungekræften viser en social gradering til ugunst for de socio-økonomisk dårligt stillede, og at man i Fredericia ikke fandt en så udtalt forskel i skattepligtig indtægt blandt lungekræftofre, at denne skønnedes at kunne forklare den fundne forskel.

Med hensyn til undersøgelsen af skolebørn skal her blot nævnes et par af Clemmesens kritikpunkter. Clem-

mesen mener, at de såkaldte landdistrikter næppe har interesse, og den fundne forskel i tobaksvaner hos forældre lader sig næppe anvende som forklaring på forskellen i lungekræft, og hvis man i diskussionen inddrager landdistrikter, er det fra alle undersøgelser venteligt, at f.eks. tobaksvanerne dér frembyder en mindre risiko for lungekræft, men dette er irrelevant i diskussionen om en eventuel virkning af luftforureningen i Fredericias centrum, modsat perifere byområder.

I samme korrespondance-rubrik fremfører Arne Nielsen (35), at de benyttede skoler, der repræsenterede oplandet, udviser de mest udprægede forskelle med hensyn til gennemsnitsvægt og -højder og med hensyn til lungefunktion. Disse forskelle, mener Nielsen, burde være taget i betragtning i analysen.

Som svar herpå påpeger Horwitz og Bach Pedersen (38), at de netop ikke har angivet data for lungefunktion m.m. område for område, men skole for skole, og derved er der klar overensstemmelse mellem analysen, og de krav Nielsen stiller.

20.02.1978 afsluttes (foreløbig) debatten om lungekræft-risikoen i Fredericia ved, at Karl M. Christensen og Martin Silberschmid (12) efterlyser, hvilken sammenhæng med kræftsygelighed det er, Horwitz og Bach Pedersen forventer, kan vises ved at undersøge børn fra forurenede/ikke-forurenede område. Desuden mener Christensen og Silberschmid der er to problemer, der gør artiklens værdi problematisk, nemlig 1) fremsættelse af postuletet, at børn forventes at være en fintmækkende indikator for luftforurening, og 2) anvendelsen af peak-flow som lungefunktionsvurdering.

Til dette sidste svarer Horwitz og Bach Pedersen (39), at det er standardmetoden ved de internationale studier, der udføres i WHO's og EF's regie.

Clemmesen et al./E.B.Andersen (14)/(01).

I januar/februar 1972 diagnosticerede man på Fredericia sygehus et større antal lungekræfttilfælde end tidligere, og dette sammeholdt med en formodning om at luftforureningen i Fredericia er særlig høj, førte til en hypotese om, at det høje antal kræfttilfælde kunne hæn-
ge sammen med forureningen.

Formodningen om særlig stor forurening i Fredericia er baseret på målinger af svævestøv og SO₂ på tre målestationer siden 1960.

Målingerne i Fredericia viser et årsgennemsnit af SO₂ på ca. 50 mikrogram/m³, og spidsværdier på ca. 100-200 mikrogram/m³; man har kun "sporadiske" målinger af tilsvarende art for resten af landet at sammenligne disse tal med, men fra andre undersøgelser ved man, at værdier på over 100 mikrogram/m³ i årsgennemsnit har medført hyppigere luftvejssygdomme og nedsat lungfunktion hos børn.

Man ved, at SO₂ i sig selv næppe har stor indflydelse på lungekræft, men man formoder, at SO₂ er godt korreleret med andre stoffer, som er kræftfremkaldende.

For at undersøge ovenstående hypotese fandt man antallet af lungekræftdødsfald i Fredericia i perioden 1968-71.

Man valgte nu at foretage en foreløbig undersøgelse, og for at finde en egnet kontrolgruppe tog man tre byer, som i visse henseender ligner Fredericia (bla. med hensyn til størrelse), nemlig Horsens, Kolding og Vejle, og fandt her de tilsvarende antal lungekræftdødsfald.

Det man ønskede at teste var altså :

Er lungekræfthyppigheden større i Fredericia end i andre lignende byer.

Clemmesen et al. foretog et statistisk test på materialet, og påviste (på 5 % niveau) at det samlede antal tilfælde i perioden for mænd er signifikant større end i en kontrolby, dvs. en "by" konstrueret på grundlag af et gennemsnit af Horsens, Kolding og Vejle (tallene for denne

"by" (i det fgl. kaldet 3-k byen) er normeret med hensyn til antal indbyggere og alderssammensætning, så den kan sammenlignes direkte med Fredericia).

Et par måneder senere offentliggjorde statistikeren E.B.Andersen en re-analyse af det samme materiale, og finder i denne analyse ikke signifikant forskel på antal lungekræfttildødsfald i de fire byer.

Vi vil i det følgende vurdere de to tests Clemmesen og E.B.Andersen har anvendt, samt føje endnu et par muligheder til for at vise, hvilken betydning den statistiske fremgangsmåde har på resultatet.

Materialet ser ud som følger :

Table II. Population by sex and age in the previous municipal areas. 1970.

	Fredericia		Horsens		Kolding		Vejle	
	Men	Women	Men	Women	Men	Women	Men	Women
0-24	6799	6959	7014	6493	7154	7256	5838	6111
25-29	1428	1319	1385	1128	1400	1328	1214	1205
30-34	990	949	841	804	1015	1023	844	902
35-39	915	948	758	860	957	989	775	807
40-44	976	1061	864	960	963	1132	762	843
45-49	1110	1115	997	1149	1084	1206	893	1006
50-54	973	980	1014	1118	1095	1213	865	1045
55-59	800	897	1083	1172	1050	1144	878	1018
60-64	710	887	923	1205	895	1025	839	1022
65-69	581	782	834	1126	702	901	631	873
70-74	509	667	634	914	535	797	539	770
75 and over	605	850	782	1173	659	1066	619	966
Total	16396	17414	17129	18102	17509	19080	14697	16568

Table III. Lung cancer. Number of cases and annual morbidity rates per 100,000 males 1968-1971 for various towns.

Age	Fredericia		Horsens		Kolding		Vejle	
	Number of cases	Annual morbidity rates per 100,000	Number of cases	Annual morbidity rates per 100,000	Number of cases	Annual morbidity rates per 100,000	Number of cases	Annual morbidity rates per 100,000
0-24	0	0	0	0	0	0	0	0
25-29	0	0	0	0	0	0	0	0
30-34	0	0	0	0	0	0	0	0
35-39	0	0	1	33	0	0	0	0
40-44	1	26	2	58	1	26	1	33
45-49	3	68	4	100	1	23	2	56
50-54	7	180	7	173	2	46	2	58
55-59	11	344	6	139	8	191	7	199
60-64	11	387	15	406	7	196	10	298
65-69	10	430	10	300	11	392	14	555
70-74	11	540	12	473	9	421	8	371
75 and over	10	413	2	64	12	455	7	283
Total	64	-	59	-	51	-	51	-

Clemmesen :

Clemmesen starter med at antage, at antal kræftdødsfald er poissonfordelt, idet hyppigheden er lille i forhold til antal indbyggere; endvidere må de være uafhængige, idet man ikke regner med at lungekræft smitter, (at kræftdødsfaldende er poissonfordelt ses også hvis man først antager, at de er binomialfordelt, enten kræft eller ikke, denne binomialfordeling kan så approximeres med en poissonfordeling, se bilag 12).

Disse poissonfordelinger (en svarende til hver aldersgruppe) aproximeres med normalfordelinger, og da vi ved at summen af en række normalfordelte observationer er en normalfordeling, får vi at summen af alle kræfttilfælde i hver by er normalfordelt, og endvidere ved vi, at variansen er lig middelværdien, idet de stammer fra poissonfordelinger.

Clemmesen slår som nævnt de tre kontrolbyer sammen til en "gennemsnitsby" (3-k by).

Ved at dividere med det samlede antal indbyggere (hver aldersgruppe for sig), og gange med antal indbyggere i Fredericia, får vi noget vi kunne kalde de forventede, og vi kan nu teste om disse stemmer overens med det faktisk observerede :

	Fredericia	3-k byen
0-24 år	0	0
25-29 -	0	0
30-34 -	0	0
35-39 -	0	0,37
40-44 -	1	1,51
45-49 -	3	2,61
50-54 -	7	3,60
55-59 -	11	5,58
60-64 -	11	8,55
65-69 -	10	9,38
70-74 -	11	8,64
75- -	10	6,17
	<hr/>	<hr/>
	64	46,41

Clemmesen foretager nu et u-test (t-test med kendt varians), idet han skønner variansen lig middelværdien, lig henholdsvis 64 og 46,4.

Vi trækker nu disse normalfordelte observationer fra hinanden, og får en ny normalfordeling - hvorefter vi kan teste om middelværdien i denne er 0.

Dette svarer til hypotesen :

Middelværdien i Fredericias population er lig middelværdien i 3-k by populationen, eller sagt på en anden måde : sandsynligheden for at få lungekræft er ens i Fredericia og 3-k byen.

Bemærk, at der er tale om et udtryk for den generelle sandsynlighed, dvs. sandsynligheden for på et eller andet tidspunkt i livet af få lungekræft, (forudsat at aldersfordeling og andre forhold i byen ikke ændres i en menneskealder).

Den normerede normalfordelte teststørrelse bliver :

$$z = \frac{64 - 46.4}{\sqrt{64 + 13.5}} = 2.00$$

13,5 er skønnet over variansen på det forventede antal, beregnet udfra

$$\text{var}(aX+bY) = a^2 \text{var}(X) + b^2 \text{var}(Y)$$

Ved opslag i tabellen fås : $P(z > 2,00) = 2,28 \%$

For kvinder bliver teststørrelsen $z = 1,09$, men da tallene her er temmeligt små, ses der bort fra disse.

Vi ser altså, at der for mænd er tydelig signifikans på 5 % niveau.

E.B.Andersen :

Andersen tager kort tid efter til genmåle; han foretager på de samme data en Q-test (chi-kvadrat-test), og finder ikke signifikant forskel.

Andersen er enig med Clemmesen i de første modelbetragtninger (poissonfordelingen), men han mener ikke det er rimeligt at slå Horsens, Kolding og Vejle sammen.

Han foretager i stedet en normering med hensyn til antal indbyggere i de forskellige byer, således at tallene for

for alle fire byer bliver direkte sammenlignelige.
 Endvidere slår han nogle af aldersgrupperne sammen for
 at få dem store nok, se nedenstående skema :

Tabel I. Sammenligning af forventede og observerede antal lungekræfttilfælde fordelt på aldersgrupper.

Aldersgruppe	Fredericia		Horsens		Kolding		Vejle		Q	P
	Observed	Forventet	Observed	Forventet	Observed	Forventet	Observed	Forventet		
40-54	11	8,7	13	8,2	4	8,9	5	7,2	6,8	10%
55-59	11	6,7	6	9,1	8	8,8	7	7,4	3,9	30%
60-64	11	9,1	15	11,8	7	11,4	10	10,7	3,0	40%
65-69	10	9,5	10	13,7	11	11,5	14	10,3	2,4	50%
70-74	11	9,2	12	11,4	9	9,7	8	9,7	0,7	80%
75-	10	7,0	2	9,1	12	7,7	7	7,2	9,2	2½%
I alt	64	50,2	58	63,3	51	58,0	51	52,5	26,0	10%

Kalder vi nu sandsynligheden for at få lungekræft i celle i_j for p_{ij} - hvor i er inddeling efter alder, og j er inddeling efter by, kan vi opstille følgende nulhypotese : $p_{i1} = p_{i2} = p_{i3} = p_{i4}$

Skal vi formulere dette i ord, bliver det :

Sandsynligheden for i en given alder at dø af lungekræft er ens i de fire byer.

Andersen skriver videre : "Det naturlige test for en sådan hypotese er et sæt uafhængige Q-test", hvilket han gør, med det resultat, at han kun finder en signifikant stor teststørrelse i aldersgruppen over 75 år.

Det ses imidlertid udfra dataene, at dette snarere stammer fra en markant lav kræfthyppighed i Horsens, end en markant høj i Fredericia.

Summeres alle teststørrelserne, fås igen en ikke signifikant forskel.

Diskussion :

Vi ser først, at Clemmesens og Andersens hypoteser ikke er ens, selvom det tilsyneladende er det samme overordnede problem, der undersøges.

Clemmesen har valgt at se på den generelle (dvs. alle aldersgrupper under et) sandsynlighed for kræft, mens Andersen mener, man bør teste hver aldersgruppe for sig.

Fra indledningen husker vi, at det var et spørgsmål, om der generelt var større risiko for kræft i Fredericia end andre steder, hvorfor Clemmesens test på dette område må synes at ligge tættest op af udgangspunktet. Det er jo sådan, at forureningens virkning er kumulativ, og det bliver derfor spørgsmålet, om risikoen i det lange løb er større i Fredericia end andre steder, vi ønsker svar på.

I Andersens test kan vi dog også hurtigt få et udtryk for det samme, idet vi kan summere Q-størrelserne, og øge antallet af frihedsgrader) - det bliver da hypotesen : $P_{i1} = P_{i2} = P_{i3} = P_{i4}$ for alle i
Dette giver en $Q = 26$; $P(Q > 26) = 10 \%$, altså heller ikke signifikant forskel.

Andersen argumenterer for at hans test er den bedste, bla. udfra at den tager hensyn til variationen i kontrolgruppen, dvs. kvaliteten af kontrolmaterialet tages med i betragtning, hvilket Clemmesens test ikke gør, idet disse simpelt hen slås sammen.

Om det ene eller det andet er det rimeligste, kan der ikke gives noget entydigt svar på - det må, som vi har gjort rede for i afsnit II afgøres af de hypoteser, vi formulerer.

Vi er dog på dette punkt mest tilbøjelige til at give Arne Nielsen (en af Clemmesens medarbejdere) ret i hans modkritik af Andersen, idet vi jo var interesseret i at sammenligne Fredericia med "andre steder", hvilket vel med en vis rimelig kan forstås som "gennemsnittet af tilsvarende byer".

Foretager vi endvidere et Q-test på de tre kontrolbyer efter Andersens metode, ser vi, at disse ikke er forskellige på 5 % niveau, omend det er lige ved :
vi får $Q = 21$, $P(Q > 21) = 5 \%$.

Slår vi kontrolbyerne sammen, og foretager et Q-test mellem Fredericia og 3-k byen, får vi :

$Q = 10,9$, $P(Q > 10,9) = 10 \%$.

Altså heller ikke signifikans.

Noget kunne altså tyde på, at det er forskelle i selve testprocedurene, som er årsag til divergensen.

En væsentlig forskel er, at Clemmesens test er ensidet, mens Andersens er tosidet (se bilag 10); her er vi igen mest enige med Clemmesen (og Arne Nielsen) i, at det på baggrund af vor alternative hypotese er mest rimeligt med et ensidet test.

I Andersens artikel (03) tester han i en multiplikativ poissonmodel, og får heller ikke signifikans, men det er også her et tosidet test.

Man kunne endvidere tænke sig, at man valgte ikke at slå kontrolbyerne sammen, men valgte at acceptere Clemmesens modelbetragtninger (med hensyn til normalfordelingsapproximationen) - man kunne så foretage en variansanalyse, hvilket vi har prøvet.

Vi får her en F-værdi mindre end 1; fortolkningen af en sådan værdi er imidlertid et udtryk for, at variansen indenfor grupperne (dvs. den udfra middelværdien skønnede varians) er større end variansen mellem grupperne, og dette kunne tyde på, at Clemmesens variansskøn ikke er så godt.

Dette er netop et af Andersens væsentligste kritikpunkter, idet det er det der ligger bag hans påstand om, at Clemmesens teststørrelse ikke er normeret normalfordelt; I den u-test, som Clemmesen foretager, er det nemlig en forudsætning, at variansen er kendt; Clemmesen skønner variansen = middelværdien, og variansen og middelværdien er derfor ikke uafhængige, hvilket de skal være i et u-test.

Man kan derfor sætte spørgsmål ved Clemmesens testprocedure, idet forudsætningerne ikke er helt i orden, men da Andersens test efter vort skøn ikke tester vor egentlige alternative hypotese, kan denne heller ikke siges at være uden svagheder.

E.B.Andersen. (03)

E.B.Andersen har i forbindelse med en artikel om den multiplikative Poissonmodel analyseret data fra Fredericiaundersøgelsen til illustration af modellen. Af denne grund og for at gøre billedet komplet er artiklen taget med i denne sammenhæng.

Artiklen omhandler den multiplikative Poissonmodel. Første del omfatter en beskrivelse af modellen, Maximumlikelihood estimation af parametrene i modellen og hypotesetestning.

Vi vil ikke gå i dybden med den multiplikative Poissonmodel, men koncentrere os om anden halvdel af artiklen, der analyserer data fra Clemmesen et. al.'s Fredericiaundersøgelse for at illustrere modellen.

Som udgangspunkt antager Andersen, at antallet af lungekræfttilfælde i hver celle i en tabel, der viser antal kræfttilfælde i aldersgrupper for fire forskellige byer, er Poisson-fordelt. Endvidere antages, at den individuelle risiko for at få lungekræft opstår som et produkt af en aldersfaktor og en byfaktor.

Disse antagelser bevirker, at hvis x_{ij} =antallet af lungekræfttilfælde i aldersgruppe i for by j , vil

$$E(x_{ij}) = N_{ij} \epsilon_i \delta_j$$

hvor N_{ij} er antal indbyggere i aldersgruppe i for by j , og ϵ_i og δ_j er henholdsvis aldersfaktor og byfaktor. (For alle δ gælder, at der er normeret, så $\sum \delta_j = 1$). Denne model er netop identisk med den multiplikative Poissonmodel.

Andersen gør opmærksom på, at han også antager uafhængighed mellem x_{ij} 'erne. Dette betyder, at lungekræft ikke skulle være smitsomt, hvilket må være en rimelig antagelse.

Løsningerne på likelihood-ligningerne benævnes $\hat{\epsilon}$ og $\hat{\delta}$, og disse findes til

$$\hat{\epsilon} = (0.01138, 0.03414, 0.05184, 0.06650, 0.07469, 0.04695)$$

og

$$\delta = (0.315, 0.227, 0.218, 0.240).$$

Først tester Andersen hypotesen H_0 : en Poisson-model med multiplikative middelværdier beskriver data.

Dette test er på grundlag af den oprindelige hypotese: alle celler er Poisson-fordelte. Det vil sige, at han ser om parametrene kan reduceres ved hjælp af den multiplikative Poissonmodel.

Testet udføres med både kvotient-test og med et χ^2 test. Begge test viser, at der på 5%-niveau er grund til at antage, at modellen fitter data.

Dernæst testes hypotesen om identiske kolonneeffekter (byeffekter). Testet er kvotient-test, og konklusionen er, at hypotesen om identiske kolonneeffekter ikke kan forkastes.

Til sidst i artiklen nævnes, at Clemmesen et al. antog at Fredericia havde en højere intensitet af lungekræfttilfælde, hvis der overhovedet var forskel.

Dette fordrer et en-sidet test for hypotesen

$$H_1^*: \delta_j = 1/4 \quad j=1,2,3,4$$

mod alternativet

$$\delta_1 > \delta_2, \delta_3, \delta_4.$$

Et sådant test kan ikke udføres med χ^2 -test, idet z_1 , som er kvotient-test variabel for hypotesen H_1^* ikke kan bestemme hvilke af de signifikante værdier, der er i overensstemmelse med

$$\delta_1 > \delta_2, \delta_3, \delta_4.$$

Alternativt kunne hypotesen omformuleres til

$$H_1^{**}: \delta_1 = (\delta_2 + \delta_3 + \delta_4) / 3$$

Dette kræver imidlertid at likelihood-ligningerne til bestemmelse af ε_i og δ_i ($i=1,2,3,4$) skal løses under indskrænkningen

$$\delta_1 = (\delta_2 + \delta_3 + \delta_4) / 3$$

En simplere metode ville være at gentage den komplette analyse med Kolding, Horsens og Vejle som en by.

Da der imidlertid efter Andersens opfattelse ikke er grund til at tro a priori, at Fredericia er en farligere by, har han ikke udført analysen.

Andersens metode virker umiddelbart tiltalende. Ved at opsplitte i række- og kolonneeffekter kan man koncentrere sig om i dette kolonne(by-)effekterne ved at eliminere række(alders-)effekten.

Som vi vil komme ind på i den generelle kritik af hele Fredericiaproblemet, mener vi, der er tale om et skinproblem, og et sådant kan en god metode ikke gøre relevant.

Vi er ikke kompetente til at gå i dybden med den multiplikative Poissonmodels anvendelighed, blot bemærker vi, at metoden fremtræder som en, der er værd at erindre.

I forgående underafsnit findes en diskussion af Andersens alternative hypotese.

Horwitz og Bach Pedersen. (28)

Den direkte anledning til, at dette studium blev taget op, var rapporterne om lungekræft i Fredericia. På initiativ af Sundhedsstyrelsen og Miljøstyrelsen besluttedes det at foretage en analyse af dødelighedsforholdene i Fredericia. Det foreliggende studium er det første i en række, hvori det skal belyses, om der eksisterer en sammenhæng mellem den almene sundhedstilstand og det industrielle miljø, og den her resumerede artikel belyser dødelighedsforholdene i Danmark 1970-72. Der redegøres for dels den almene dødelighed dels dødeligheden af lungesygdomme i bysamfund af forskellig størrelse.

Grundmaterialet er samtlige personer, der døde i årene 1970-72, ialt 147.430. Herudfra er yderligere udsorteret alle patienter, hvor lungelidelse var anført som primær dødsårsag, ialt 15.031 patienter, og

endelig er der foretaget en specialopgørelse for primær lungekræft og de kronisk obstruktive lungevejslidelser (fællesbetegnelse: kronisk bronchitis). Der var ialt 5.855 lungekræfttilfælde og ialt 3.986 tilfælde af kronisk bronchitis.

For at belyse dødelighedens sammenhæng med kommunestørrelse, blev kommunerne opdelt efter indbyggertal, og der fandtes ialt 7 kommuner i kategorien 40.000-49.000 personer, nemlig Tårnby, Hvidovre, Rødovre, Fredericia, Vejle, Silkeborg og Næstved.

Det har været af særlig betydning at sammenholde dødeligheden (årlige antal dødsfald pr. 100.000 personer) i Fredericia med dødeligheden i de seks andre byer med samme størrelse. Dette er gjort ved såkaldte "dødelighedsbånd", som er baseret på raterne i de seks byer. Båndene angiver, aldersklasse for aldersklasse, den højeste og den laveste dødelighed, der er iagttaget indenfor de seks kommuner. For at lette sammenligningen mellem dødeligheden i kommuner, hvis alderssammensætning er forskellig, er der beregnet standardiserede rater.

Det standardiserede antal dødsfald i den enkelte kommune fandtes ved at gange den aldersspecifikke dødelighed med antallet i den pågældende aldersklasse i standardbefolkningen. Ved at summere over de enkelte celler fås det standardiserede antal dødsfald.

Af resultater kan nævnes, at dødeligheden for hele landet ikke viser en systematisk variation med kommunestørrelse før 25-års alderen. Derefter er der et tydeligt mønster, idet raten er størst i de store kommuner og mindst i de små. Med hensyn til lungesygdomme var aldersmønstret hos mænd det samme som for alle sygdomme under ét, dog var stigningen med alder noget stejlere. Dødeligheden viste en tydelig afhængighed af kommunestørrelse: jo større kommune, des højere dødelighed. Lungekræft forekom ikke hos unge, men fra 40-års alder

ren sker en kraftig stigning, idet den højeste rate forelå hos de ældste. Raten var højest i hovedstaden og lavest i de små kommuner. Hos mænd var den standardiserede rate 2.8 gange højere i hovedstaden og hos kvinder 2.4 gange højere. Alderskurverne viser ikke systematisk forskel mellem de tre grupper af kommuner (over 500.000, 40.000-49.000, mindre end 10.000).

Det undersøgte om kommunerne med 40.000-49.000 indbyggere indtog en særstilling, og der var ingen grund til at antage dette.

Med hensyn til specielt Fredericia indtog dødeligheden blandt de nul-årige, der efter forfatterens skøn er fintmærende indikatorer for sundhedstilstanden, en mellemstilling.

Fredericias dødelighed af alle lungevejslidelser lå, bortset fra enkelte usystematiske afvigelser, på "dødelighedsbåndet". Dette gjaldt både for mænd og kvinder. Det samme gjaldt for dødeligheden af lungekræft og kronisk bronchitis. De enkelte afvigelser fra helheden forklares som tilfældige variationer betinget af små tal.

Der udførtes en særlig analyse for kronisk bronchitis, idet denne lidelse ofte er anført som sekundær dødsårsag. Fredericias standardiserede dødsrater viste her en signifikant forskel for mænd på 0.1% niveau, men ikke for kvinder på 5% niveau.

I diskussionen nævner forfatterne, at det ikke har været muligt at sammenholde data for luftforurening med mortalitetsdata, da der kun i få kommuner er foretaget luftforureningsmålinger. Derfor har man valgt at sammenligne mortalitetsdata inden for kommuner med nogenlunde samme indbyggertal, ud fra den betragtning, at et bestemt antal personer må kræve en bestemt mængde industri som eksistensgrundlag.

Sluttelig skal nævnes, at Horwitz og Pedersen finder det "berettiget at konkludere, at den forhåndenværen-

de viden om årsagsforhold og om dynamikken i befolkningsbevægelser ikke er tilstrækkelig til at forklare de iagttagne forskelle og resultater. De må foreløbig fremstå som kendsgerninger, hvis årsagsforhold må udredes af morgendagens forskning."

Diskussionen af denne undersøgelse gemmes til efter omtalen af den følgende anden del.

Bach Pedersen og Horwitz. (36)

Denne undersøgelse ligger i forlængelse af den lige omtalte. Man har også denne gang ønsket at undersøge den rolle industrialiseringen kan have på befolkningens sundhedstilstand. Det foreliggende studium omfatter tretten danske provinsbyer, og dødeligheden af alle årsager og af lungesygdomme i årene 1970-72 er sat i relation til den lokale industris art og størrelse.

For at sikre at grundbefolkningen og dens levevilkår var nogenlunde ensartet, skulle de studerede kommuner opfylde tre kriterier: 1) indbyggertallet skulle pr. 1.1.1971 ligge mellem 30.000 og 80.000, 2) kommunerne skulle have mindst 3.000 industriansatte og 3) kommunerne skulle ligge udenfor hovedstadsområdet.

Fjorten kommuner opfyldte kriterierne, men den ene, Køge, måtte udelukkes, da industrimønstret var helt afvigende. De resterende tretten kommuner var: Esbjerg, Horsens, Vejle, Fredericia, Silkeborg, Kolding, Herning, Roskilde, Næstved, Randers, Svendborg, Helsingør og Frederikshavn.

Dødsfaldene i de enkelte kommuner er opdelt efter køn og alder, og der blev udført en standardisering. Standardbefolkningen for mænd var det samlede antal mænd i de tretten kommuner, for kvinder var den det samlede antal kvinder.

I undersøgelsen har man ønsket at relatere dødeligheden til forskellige parametre, der er udtryk for industrialiseringsgrad. Desværre fremgår det ikke tyde-

ligt, hvilke parametre der er tale om, men her skal nævnes nogle stykker.

SO₂-emission.

Der foreligger ikke målinger, der angiver SO₂-koncentrationen i luften i de tretten kommuner. SO₂-emissionen er derfor beregnet ud fra det årlige forbrug af brændsel m.m. i de enkelte kommuner.

Industrialisering.

Industrivirksomheder er af Arbejdstilsynet opdelt i tre kategorier efter forureningsgrad, nemlig mindre forurenende virksomheder, middel forurenende virksomheder og stærkt forurenende virksomheder. Der foreligger specifikation af hver kategori.

Virksomhedernes fordeling efter forureningsgrad var ensartet for de tretten kommuner.

Antallet af beskæftigede giver et udtryk for virksomhedernes størrelse. Imidlertid er dette antal ikke registreret for virksomheder med mindre end 6 ansatte. Da man kender antallet af små virksomheder, har man beregnet en maksimums- og en minimumsværdi for antallet af ansatte i hver kommune og benytter et estimeret gennemsnit.

Funktionærernes og arbejderne aldersfordeling for de enkelte kommuner foreligger ikke belyst. Dette medfører, at man ikke kan beregne et standardiseret antal døde for de enkelte grupper. Analyserne må derfor betragtes som indikatorer og ikke som den udtømmende sandhed.

Sammenhængen mellem det standardiserede antal dødsfald og de forskellige parametre er belyst ved regressionslinier. Som angivelse af punkternes spredning omkring linierne benyttes korrelationskoefficienten. Et t-test er benyttet for at finde hvilken værdi af r (korrelationskoefficienten), der er grænse, og denne er 0.6. Det vil sige, at der er sammenhæng når r (numerisk) er 0.6 eller derover.

Undersøgelsen viser, at der ikke forelå sammenhæng mellem SO_2 -emissionen og dødeligheden, hverken den totale eller den, der skyldtes lungelidelser. Udelukkes Esbjerg fra undersøgelsen på grund af et ekstremt højt brændstofforbrug, får tydelig korrelation (positiv) blandt mænd for alle lungevejslidelser, for kronisk bronchitis og for lungekræft (denne oplysning findes i en fodnote). Blandt mænd var der tydelig sammenhæng mellem antallet af industrivirksomheder pr. km^2 og dødeligheden, både når man ser på dødeligheden af alle årsager og dødeligheden af lungesygdomme.

Det kan ikke afgøres om den iagttagne stigning i dødelighed blandt mænd alene er betinget af en forøgelse af de stærkt forurenende virksomheder, eller om den skyldes, at den totale mængde industri er forøget.

Antallet af beskæftigede i industrien forelå kun fordelt på køn for større virksomheder. Derfor måtte den samlede procentdel af befolkningen, der er beskæftiget indenfor industri, sammenholdes med det totale antal dødsfald, der er indtruffet, idet kønsopdeling og opdeling arbejder/funktionær måtte udelades. Der fandtes ingen sammenhæng hvad angik den totale dødelighed eller dødelighed af lungesygdomme.

Endelig fandtes for mænd en tydelig sammenhæng mellem urbaniseringsgraden og dødeligheden. Den fandtes også blandt mænd for alle lungevejslidelser.

I diskussionen kommer forfatterne ind på, hvor usikkert man står, når det ydre miljøes betydning skal vurderes. De nævner manglende eller mangelfulde data, og spørger om hvor mange og hvilke stoffer og partikler, der bør indsamles data for, når luftforureningen skal bedømmes. Efter at have peget på lignende konkrete problemer siges, at spørgsmålene må besvares, inden man kan drage konklusioner om, hvormeget forureningen påvirker befolkningens sundhedstilstand.

Hvorefter de konkluderer, at de steder, der kunne påvi-

ses en effekt af industrien på dødeligheden, var denne begrænset til mændene, og de mener, dette kunne tyde på, at industriens skadelige virkninger først og fremmest skal søges på arbejdspladsen, idet der, hvis dette ikke var tilfældet, også ville foreligge en positiv korrelation mellem industrialiseringsgrad og kvinders dødelighed.

Der er flere kommentarer at knytte til de to undersøgelser.

1) Hvorfor har man ønsket at belyse dødelighedens sammenhæng med kommunestørrelse? Der findes et utal af undersøgelser, der har påvist en sammenhæng. Der er muligvis ikke analyseret for danske kommuner, men industrimønstret (produktionsprocesserne) er stort set ens i hele den vestlige verden, og derfor kan man overføre resultater fra sådanne undersøgelser til danske forhold uden særlige modifikationer.

2) Det er en fejl, at man ikke i planlægningen har undersøgt hvilke oplysninger, det var muligt at tilvejebringe. Havde man gjort det, kunne man tilrettelægge undersøgelsen herefter. På den anden måde kan man ikke undgå overslag og spekulationer af enhver art, og det er inkrediensert i en usaglig undersøgelse, der ikke er modstandsdygtig over for kritik. Rent spekulative indvendinger må nødvendigvis tages for gode varer, så længe undersøgelsen selv benytter sig af sligt. I en form for efterrationalisering nævner Horwitz og Bach Pedersen nogle eksempler på, hvor usikkert man står, når det ydre miljø skal bedømmes. Havde denne viden været til stede allerede fra begyndelsen, var foreliggende undersøgelse næppe planlagt i den nuværende udformning.

3) Det fremgår ikke, om man har testet kontrolbyernes ensartethed.

4) Man skriver, at det har været af særlig betydning at sammenholde dødeligheden i Fredericia med dødelig-

ligheden i seks andre byer af samme størrelse. Som et udtryk for bystørrelse er benyttet indbyggertal i kommunen, hvilket er stærkt misvisende, idet forfatterne senere selv nævner, at den del af befolkningen, som boede i byen, kommunen er opkaldt efter, varierede fra 63% til 91%.

5) Der savnes en argumentation for, at de nul-årige er særligt fintmærende indikatorer for sundhedstilstanden.

6) Forfatterne påpeger, at der ikke er særlig grund til at tro, luftforureningen har indflydelse på dødeligheden, idet man da måtte forvente ens dødsrater for både mænd og kvinder.

Sammenligner man studier over lungekræftdødeligheden over en årrække, ser man en stigning i kvindernes dødelighed mod en udjævning af forskellen. Dette kan der gives flere spekulative forklaringer på, og en er, at hvor det før typisk var mænd, der var udsat for en "stresset" tilværelse, er det nu mindre og mindre et "privilegie" kun for mænd. Da man samtidigt ved, at alle former for stress nedsætter organismens modstanddygtighed, er det muligvis her, men skal søge en forklaring.

7) Der er ikke argumenteret for valg af SO_2 -emission som variabel. Hvis ikke man har formodninger om, at SO_2 er kræftfremkaldende, eller har indicier for at det er godt korreleret med f.eks. polycykliske hydrocarboner, søger man bevidst efter en stokastisk sammenhæng.

8) Det virker tåbeligt, at man udelukker Køge på grund af et afvigende industrimønster, når Esbjerg bibeholdes på trods af et afvigende brændstofforbrug.

Esbjergs høje forbrug hænger måske sammen med fiskeriet, og forbrændingen vil da hovedsageligt ske på havet. Altså sker der kun ringe påvirkning af byluften, og derfor vil Esbjerg have en lav dødelighed på trods af et højt brændstofforbrug.

Den beregnede SO_2 -emission vil også af denne grund være misvisende.

9) Endelig virker det besynderligt, at man taler om regressionslinier, når man benytter korrelationskoefficienten som udtryk for sammenhæng. Der savnes også en begrundelse for at benytte r som angivelse af punkternes spredning omkring regressionslinierne, når der for denne findes et standard-udtryk.

Samlet kritik af hele fredericiaundersøgelsen.

I dette afsnit har vi rettet en del kritik imod fredericiaundersøgelsen, og vi vil tilslut resumere de væsentligste punkter, samt med baggrund i afsnit II, og III trække nogle linier op, for hvordan vi mener fremtidige undersøgelser bør tilrettelægges.

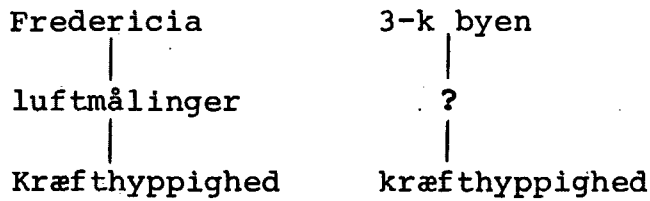
Man undersøger om Fredericia har større dødelighed end 3 kontrolbyer (14), og senere foranstalttes en endnu større undersøgelse (36), hvori man tester forskellige generelle hypoteser om kræftdødeligheden, bla. i forhold til befolkningstæthed og industrialiseringsgrad. Ingen af disse undersøgelser indeholder imidlertid pålidelige tal for de primære kilder til den øgede kræfthyppeghed, nemlig kemiske stoffer i luften (PAH, PVC m. fl.).

Fortolkningen af sådanne tests er som beskrevet i afsnit II suspekt, og dette fremgår da også med al ønskelighed af korrespondancen i ugeskrift for læger.

Som det også fremgår af afsnit II er det valget af parametre det er galt med.

I fredericiaundersøgelsen (14) har man kun en formodning om forureningen i Fredericia - denne formodning synes i øvrigt ikke at underbygges af visse senere offentliggjorte tal for Vejle (21), som viser et større indhold af SO_2 i Vejle end Fredericia; man skal dog tage disse tal med et vist forbehold, bla. af den grund, at SO_2 ikke er en primær kilde til kræft.

Vi kan opstille følgende skema for fredericiaundersøgelsen :

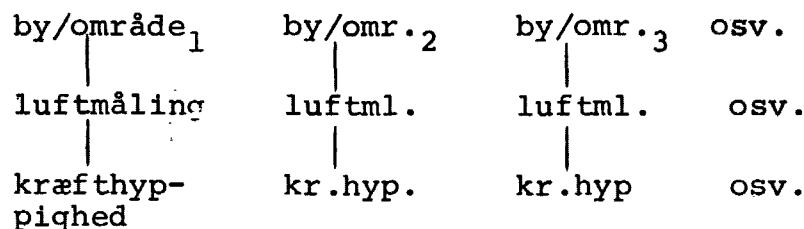


Vælger vi nu at godtage Clemmesens testresultat, er vi egentlig ikke kommet et skridt videre, idet vi ikke kan være sikre på at det er luftforureningen som har skylden, eller om det er andre faktorer (se gennemgang af korrespondancen).

Det er endvidere et faktum, at testen kun er signifikant for mænd, hvilket kunne tyde på, at arbejdsmiljøet også har indflydelse.

I Horwich (36) inddrager man et større materiale (befolkningstal), og forsøger på baggrund af statistik over brændstofforbrug at skønne forureningen. Et sådant skøn er behæftet med så stor usikkerhed, at det intet er værd, og man har altså heller ikke hernogle primære parametre at sammenligne med.

Den efter vor mening eneste relevante måde at tilrettelægge kommende undersøgelser på, er skitseret i følgende skema :



Som det fremgår af skemaet er luftmålinger de eneste relevante parametre i en forureningsundersøgelse; med luftmålinger forstår vi målinger af stoffer, hvis medicinske betingning af kræft er velbegrundet, eventuelt stoffer, som er godt korrelerede med kræftfremkaldende stoffer, men man må i så fald tage særlig hensyn til

eventuelle specielle industrier (f.eks. PVC).
Finder vi nu en signifikant stokastisk sammenhæng mellem koncentrationen af et forurenende stof og hyppigheden af kræft, kan vi med stor sikkerhed fastslå, at der må eksistere en kausal sammenhæng, der ikke lader sig bortforklare.

Delkonklusion.

For at undgå parametre, der ikke indgår i kausal sammenhæng med responsen og ikke er højt korrelerede med kræftfremkaldende stoffer i almindelighed, må de ansvarlige myndigheder hurtigst muligt igangsætte relevante luftkvalitetsmålinger i en stor række danske byer. Resultaterne herfra skal bl.a. indgå i epidemiologiske kræftundersøgelser.

Der foreligger en grundig registrering af dødsårsager i Danmark. Dette materiale giver en solid basis for netop epidemiologiske undersøgelser, og det er derfor beklageligt, at de parametre, dødsårsagerne skal sammenholdes med, er så ufuldstændige, som tilfældet er. For at gøre registreringen af dødsårsagerne fuldstændig kunne man ønske sig, at der på dødsattester blev anført, om den døde var ryger/ikke-ryger og eventuelt beskæftigelse.

I lang tid fremover vil forskningen i årsagerne til lungekræft bære præg af grundforskning. Myndighederne må derfor allerede nu påbegynde en omfattende litteraturresearch for at danne sig et billede af sammenhængen mellem luftforurening og lungekræft. Dette arbejde bør munde ud i fastsættelse af midlertidige grænseværdier for kræftfremkaldende stoffer i forurenede byluft.

V KONKLUSION.

Såvel udenlandske som danske undersøgelser af sammenhængen mellem luftforurening og lungekræfthyppighed bærer i unødvendig høj grad præg af grundforskning. Selv om en del af de udenlandske undersøgelser rent faktisk analyserer for parametre, der indgår i kausal sammenhæng med lungekræfthyppigheden, er analyserne stort set koncentreret om parametre som befolknings-tæthed, industrialiseringsgrad, SO₂-emission etc. Valget af sådanne parametre gør analysernes brugbarhed tvivlsom. De vil ikke være modstandsdygtige over for kritik af spekulativ art, og en sådan kritik er netop, hvad man må forvente, når analysernes emne er kontroversielt for stærke økonomiske interesser. For at undgå forfladigelse af en debat om sikkerhedsgrænser for færden i byluft må parametrene, der undersøges for, indgå i kausal sammenhæng med lungekræfthyppigheden og være højt korreleret med andre carcinogene stoffer.

Herhjemme må luftkvalitetsmålinger iværksættes i en stor række danske byer, og målingerne skal bl.a. benyttes i epidemiologiske undersøgelser.

Formålet med epidemiologiske kræftundersøgelser må være at få sat grænseværdier på carcinogene stoffer i forurenede byluft, og indtil pålidelige og relevante resultater foreligger, må der etableres midlertidige grænseværdier, således at der står en (politisk (og dermed ansvarlig) myndighed bag afgørelsen om, hvor mange der hvert år skal dø af lungekræft.

VI UNDERVISNING I STATISTIK OG SANDSYNLIGHEDS-
REGNING I GYMNASIUM/HF.

Vi havde egentlig tænkt os, at dette afsnit skulle indeholde et færdigudarbejdet undervisningsforløb i form af tema-arbejde; på grund af de forhold vi har nævnt i forordet, samt den evigt tilbagevendende tidsnød, har vi måttet nøjes med nogle generelle overvejelser over statistik og sandsynlighedsregningen i gymnasium/HF.

Disse overvejelser er dels gjort på grundlag af (17), dels på grundlag af vor sidste projektrapport (41), hvori vi argumenterer for en "virkelighedsrelatering" af matematikundervisningen - istedet for den aksiomatisk/deduktive, dvs. internt matematiske fremstilling som er repræsenteret i Kristensen og Rindungs bøger, og som hovedsagelig sigter mod at gøre alle eleverne til matematikere.

Endvidere har vi haft glæde af en anden RUC-rapport, (42), hvor man bla. vurderer tendenser i retning af større vægt på undervisningen i statistik i gymnasium/HF udfra bladet "Meddelelser fra matematiklærerforeningen".

I denne rapport påviser man, hvordan Kristensen og Rindung tager udgangspunkt i en aksiomatisk/deduktiv opbygning af sandsynlighedsregning og statistik, mens dette er ændret i Crone-Vibjergs bøger, hvor man tager udgangspunkt i den deskriptive statistik.

Denne udvikling falder i tråd med udviklingstendenserne med at gøre matematikken mere anvendelsesorienteret; denne anvendelsesorientering har hidtil hovedsageligt været rettet mod elevernes hverdag, samt anvendelse i andre fag, (jvfr. bekendtgørelserne, samt Crone-Vibjergs formulering : At matematikken skal være et redskab i hverdagen.

I vor tidligere rapport (41) argumenterer vi for, at man er nødt til at tage udgangspunkt i virkeligheden, hvilket vil sige den måde statistikken anvendes på, samt den måde, hvorpå eleverne præsenteres for anvendelse af statistik i aviser, tidsskrifter m.m.

Det er med andre ord ikke nok at sørge for, at ens undervisning tilfredsstillende kræver om, at eleverne skal kunne anvende statistikken i andre fag - de skal også lære noget om hvordan den anvendes i samfundet, hvordan hypotesetestning i forskellige sammenhænge danner baggrund for beslutninger i samfundet.

Det er efter vor mening deprimerende at se, hvordan statistikken misbruges af forskellige interesser (reklamefolk, politikere m.fl.), og vi mener den væsentligste grund til at dette kan lade sig gøre, er at den almindelige indsigt i statistik er et lukket land for alle andre end statistikere.

I (42) fremgår det, at statistik "er svært"; det er vi enige i, og vi mener derfor ikke det er en holdbar konklusion på ovenstående problem at tro, at man kan lære folk (gymnasieelever) tilstrækkelig statistik (i traditionel matematisk forstand) til at sætte dem i stand til at finde frem til de svage punkter i de statistiske tests - hvordan skal man opnå dette, når ikke engang statistikere kan blive helt enige om rækkevidden af deres argumenter.

I stedet må man udfra nogle "mere tilgængelige" betragtninger forsøge at lære dem noget om hvordan statistikken kan, og bliver misbrugt, således at de får et kritisk syn på anvendelse af statistik.

(17) giver en hel del gode råd om, hvordan dette kan opnås; forfatterens hovedholdning er, at statistik er meget nært forbundet med "common sense", og at det er muligt at lære noget om rækkevidden af statistiske argumenter, uden nødvendigvis at kunne gennemskue alle

de matematiske ræsonnementer (se note m), og efter læsning af bogen må vi give ham ret.

Vi har i afsnit II set på, hvordan overtrædelse af helt fundamentale statistiske forudsætninger kan føre til fejlslutninger, og vi har set hvordan det er muligt at påpege dette, uden at fordybe sig i matematiske argumenter.

Det er vor opfattelse, at en sådan viden om generelle statistiske principper rækker langt hen ad vejen til en kritisk holdning til anvendelsen af statistik, især hvis den er tilegnet i forbindelse med konkrete eksempler om forhold, hvor eleverne i forvejen har en del faglig viden om det undersøgte fænomen.

En måde hvorpå dette kan opnås er ved temaarbejde.

I temaarbejdet får eleverne mulighed for at beskæftige sig med et samfundsrelevant problem, hvor man må formode, at konklusionen af en given statistisk undersøgelse kan vurderes udfra den faglige indsigt eleverne får fra de andre fag i temaet.

Vi vil derfor skitsere, hvordan vi tænker os et tema vil være i stand til at opfylde de krav vi her har stillet til undervisningen (samt selvfølgelig de krav som stilles i bekendtgørelsen for undervisningen).

Vi har i dette afsnit især fokuseret på, hvordan vi mener eleverne skal få en forståelse af den samfundsmæssige anvendelse af statistik; vi anerkender selvfølgelig kravene om "anvendelse i andre fag", samt de behov som af hensyn til videregående studier samtisigt må tilgodeses; det er imidlertid vor opfattelse, dels at disse krav også tilgodeses ved ovenstående principper (eleverne vil senere styre uden om mange af de elementære bommerter, man kan komme ud for), dels at de hidtil har været altdominerende, hvorfor vi ikke nærer særlig stor bekymring for at de lige pludselig skulle forsvinde.

Skitsering af et temaarbejde.

Det følgende er et forsøg på at skitsere et temaarbejde for matematik, biologi og geografi. Som arbejdstitel har vi valgt "lungekræft i Fredericia", men selve temaet kunne være "Miljø og sundhed i samfundet", "Industrialisering" (med deltagelse af historie og samfundsfag), "Folkesundhed" eller lignende.

Ud fra den valgte arbejdstitel er det oplagt at integrere matematik, biologi og geografi. Geografi kunne behandle Fredericia, biologi kunne behandle lungesygdom og matematik kunne beskæftige sig med sammenhængen mellem lungesygdomshyppigheden og luftforureningen i Fredericia.

Indholdet i biologi- og geografiundervisningen lader vi stå åbent. Her vil vi blot give en gennemgang af indholdet i den del af temaet, der skal henlægges under matematik. Først skal vi præcisere niveau og principper for undervisningen og overordnede målformuleringer vil blive præsenteret.

Niveauet mener vi skal være HF-tilvalgsfag, naturfaglig gren, samfundsfaglig gren eller matematisk/fysisk gren, altså A eller B niveau. Imidlertid løber vi allerede her ind i vanskeligheder. For det første undervises der ikke i geografi på mat/fys grenen i 3. g. Hvis geografi skal medvirke, skal temaarbejdet foregå i 2. g, men her undervises ikke i biologi. Et temaarbejde på denne gren må altså ske på bekostning af enten biologi eller geografi. Et eller for den sags skyld flere fags udelukkelse vil dog ikke have indflydelse på indholdet i matematikundervisningen, idet temaet ikke er problemformuleret. Fagintegrationen har mere karakter af at skulle give sammenhæng i undervisningen.

Det andet problem ligger i tilvalgsordningen på HF. Man kan ikke forvente at eleverne vælger både matematik, biologi og geografi, og dette vil udelukke, at man benytter viden fra andre deltagende fag i sin undervisning. Dette problem er imidlertid også overskueligt, idet undervisningen som før nævnt ikke er problemformuleret.

Den del af statistikken, vi mener, eleverne skal stifte bekendtskab med, forudsætter direkte eller indirekte kendskab til infinitesimalregningen. Derfor er HF-fællesfag af sproglig linie udelukket. Endvidere indeholder den sproglige linie ikke statistik, men kun sandsynlighedsregning.

Konklusionen må altså være, at temaet ideelt kan tilpasses mat/nat grenen og mat/samf grenen, men med tillempler også mat/fys grenen og HF-tilvalgsfag.

Principperne for undervisningen er følgende:

- eleverne skal bibringes en opfattelse af drivkræfterne bag den historiske udvikling af statistikken
- udgangspunktet i statistik og sandsynlighedsregning skal være den deskriptive statistik
- undervisningen skal tage udgangspunkt i så vidt muligt den nære virkelighed

En udførlig argumentation for de to første principper findes i (42) og for det sidste i (41).

Målformuleringen for forløbet kan af gode grunde ikke blive særlig præcis, da en sådan skal forholde sig til bekendtgørelsen, som igen er forskellig fra niveau til niveau. Derfor er målene ikke formuleret i adfærdstermer, men blot som overordnede mål.

Målene kunne f.eks. være, at eleverne skal

- bibringes en opfattelse af drivkræfterne bag den historiske udvikling af statistikken
- kunne analysere og løse problemer af statistisk art, fra andre fag og fra den øvrige hverdag

-kritisk kunne analysere statistikkens anvendelse i samfundet

-bibringes kendskab til fundamentale statistiske begreber, tankegange og metoder.

Sådanne mål er vanskelige at evaluere, og det er begrundelsen for, at vi mener, målene for de enkelte sekvenser og de overordnede mål efter valg af niveau bør formuleres præcisere.

Den røde tråd gennem hele undervisningsforløbet skal være Fredericiaundersøgelsen (14), dog skal de enkelte begreber og metoder illustreres med velvalgte virkelighedsnære eksempler.

Det vil altid være hensigtsmæssigt at producere noter til eleverne. Disse noter skal indeholde de begreber, man som underviser forventer, eleverne skal være i besiddelse af efter afsluttet undervisningsforløb. Dette vil lette en eventuel eksamensrepetition og modgå den kritik, der ofte rettes mod undervisning, der bygger på eksempler og virkelige data, nemlig at eleverne har svært ved at skille væsentligt fra uvæsentligt i eksamenssammenhæng.

Tendenser mod skærper adgangsbegrænsning og øget konkurrence på det frie arbejdsmarked gør, at man som underviser ikke kan negligere elevernes krav i denne henseende.

Udgangspunktet for forløbet skal være de rå data fra Fredericiaundersøgelsen, og det første stykke tid bør bruges til at udvikle/præsentere forskellige fremstillingsformer. Følgende begreber, der hører til den deskriptive statistik, behandles:

hyppighedsfordeling, frekvensfordeling, stolpediagram, middelværdi, median, typetal, variationsbredde, fraktil, grupperede observationer.

Herefter behandles relevante dele af sandsynlighedsregningen for at opbygge et begrebsapparat til forståelse af Fredericiaundersøgelsen. Følgende begreber behandles mere eller mindre udførligt afhængigt af omstændighederne:

sandsynlighedsfelt, hændelse, sandsynlighedsfunktion, uafhængige hændelser, stokastisk variabel, sandsynlighedsfordeling for stokastisk variabel, middelværdi og spredning for stokastisk variabel, binomialfordeling, middelværdi og spredning for binomialfordelt stokastisk variabel, kontinuerte fordelinger, frekvensfunktion, fordelingsfunktion, normalfordeling, tilnærmelse med normalfordeling til binomialfordeling, tilnærmelse med binomialfordeling til Poissonfordeling, opstilling af hypotese, hypotesetestning.

Som man vil se, stemmer indholdet godt overens med bekendtgørelsens krav. Ligeledes er de overordnede mål også i overensstemmelse hermed.

I (41) findes nogle overvejelser over et samlet teamarbejdes organisering, og disse er ikke medtaget her.

NOTER.

- a) Ordet "stikprøve" skal her forstås bredt, idet der jo ikke her er tale om et eksperiment, hvorfra man får nogle målinger; man går ind i diverse statistiker, men rent statistisk er en sådan fremgangsmåde at betragte som en tilfældig stikprøve, såfremt udvælgelsesproceduren er uafhængig af hypotesen - dette gælder også i de tilfælde, hvor stikprøven omfatter hele den valgte population; i så fald må det betragtes som en stikprøve i forhold til "alle mennesker".
- b) Kilde (29) skriver om dette :

Til slut nogle ord om, hvorledes man vælger sine hypoteser. Den ideelle situation er, når parametrene har relation til sagligt relevante faktorer, der påvirker eksperimentet, i så fald vil det ofte være klart hvilke hypoteser, man er interesseret i at undersøge. Hvis man har to medikamenter, vil man ofte være interesseret i at undersøge, om de er ens, hvis man har garanteret en vis kvalitet af en vare, vil man ved en stikprøveundersøgelse teste, om der faktisk er en sådan kvalitet. Hvis man har en teori for de fænomener, der styrer eksperimenterne, kan man måske udlede, at visse målinger skal vokse lineært, og dette kan da testes på data.

I praksis har man ikke altid denne baggrundsviden, og man kan få gode ideer til hypoteser ved at betragte datas variation.

Hvis man ved at tegne et histogram for nogle kontinuert varierende observationer finder, at fordelingen ligner en normalfordeling, ja så kan man teste dette. Hvis man ved at afbilde et sæt målinger som funktion af tiden finder, at de vokser lineært, ja så kan dette testes.

Dette kan misbruges. Hvis man i 10 møntkast finder 4 plat, da kunne man jo få den ide, at teste at p var 0,4. Dette må være galt, foruden at være uinteressant. Man kan forsøge at sammenfatte ovenstående i et princip, der som alle andre principper skal benyttes med varsomhed:

Man må kun acceptere hypoteser, man kan forklare.

Kilde (30) skriver om samme :

The formulation and testing of hypotheses are the important features of the scientific method, and some general ideas on this process may be given. In the first place, familiarity with the subject matter of the problem under investigation is absolutely essential, and, on the basis of previous knowledge, which may deal with situations that seem to be analogous, certain elements in the subject matter are regarded as significant and relevant. These elements are chosen because the investigator is aware of theories dealing with these elements. Without some theories or ideas, the investigator will be unable to choose which elements in the situation to observe.

- c) Der tales her om et princip, og altså ikke om en regel som gælder enhver anvendelse af statistik - dette fremgår også af ovenstående citat fra Søren Johansen (29); Fisher skriver om dette (24):

"Inductive inference is the only proces known to us by which essential new knowledge comes into world"...

"Eksperimental observations ara only carefully planned in advance, and designed to form a secure basis of new knowledge".

Det er altså ikke "forbudt" at søge i diverse statistikker efter nye mulige sammenhænge; en sådan sammenhæng kan dog ikke kaldes kausal, medmindre der kan gives en faglig begrundelse.

- d) Om dette skriver (30) bla. :

It is of course essential, though frequently not realized, that the hypotheses and their outcomes should be formulated before verification is attempted. Hypotheses that are formulated from or modified by the observations are always suspect, and it is one of the elementary notions of statistical tests that probability statements cannot be made about statistical tests suggested by the data to which they are applied. It is possible that two or more hypotheses can give predictions that are experimentally indistinguishable, and in such cases appeal has to be made to notions of simplicity, that a simple hypothesis is in general to be preferred to a more complex one, and to esthetic appeal, though the simplicity or complexity of a hypothesis may not be obvious and may depend on the frame of reference in which it is formulated.

- e) På grund af en redaktionel fejl, er man her nødt til at læse note 1 først.

Fisher bruger sit eksempel (fra note 1) til at illustrere dette forhold :

I den beskrevne test ser vi, at hvis testen er baseret på 8 kopper (4 med hver slags the), vil hændelsen 4 rigtige forekomme i 1/70 af tilfældene, mens hændelsen 3 rigtige, 1 forkert vil forekomme i 16/70 af tilfældene; vi ser, at der ved "rene tilfældigheder" i $17/70 = 24,4 \%$ af tilfældene forekommer hændelser der er værre (i forhold til nulhypotesen) end 3 rigtige, 1 forkert, og vi kan derfor ikke på baggrund af en sådan hændelse forkaste nulhypotesen.

Havde vi derimod i stedet brugt 12 kopper (6 af hver slags), bliver de tilsvarende sandsynligheder :

6 rigtige 0,11 %

5 rigtige, 1 forkert 3,9 %

og vi ser altså, at et forsøg med 1 eller mindre fejl ville forkaste vor nulhypotese på 5 % niveau.

Det ses heraf : jo større testmateriale, jo større er sikkerheden af testen.

- f) I kilde (25) finder man på grundlag af 9.000 tvillinger en overdødelighed blandt rygere, men dataene indicerer, at denne overdødelighed ikke først og fremmest skyldes rygning, men det faktum, at rygere gennemgående fører en hårdere tilværelse end ikke rygere.

- g) Vi kan her bruge eksemplet fra note 1) og h) til at anskueliggøre at : jo større fejlkilder, eller jo mindre effekt vi tester for, jo større materiale kræves for en sikker test :

I eksemplet kunne vi tænke os, at der i nogle af kopperne var kommet mælkog/eller sukker i nogle af kopperne, hvilket vel sagtens ville påvirke smagerens re-

sultater negativt; eller vi kunne tænke os, at damen hævdede at hun ofte tog fejl, men at hun alligevel i det lange løb ville have flere rigtige end forkerte - begge disse forhold ville kræve en større test.

h) I fredericiaundersøgelsen er besvarelsen af spørgeskemaerne vedrørende rygning meget dårlig; når vi samtidig må antage, at rygning er en mindst lige så farlig påvirkning som luftforurening, og vi samtidig ved at der er andre støjkilder end rygning, ser vi, at materialet langt fra er et perfekt testmateriale, men det er en situation, man ofte befinder sig i.

i) Fisher giver i (24) et eksempel, som illustrerer logikken i dette :

En dame påstår, at hun kan smage forskel på to slags the; vi sætter et eksperiment op, hvor hun i vilkårlig rækkefølge skal smage på et antal kopper af hver slags.

Vi vil altså teste hypotesen : Damen kan smage forskel, men vi vælger i stedet at teste (nul) hypotesen : Damen kan ikke smage forskel.

Viser det sig nu, at damen svarer rigtigt i betydeligt flere end 50 % af tilfældene, forkaster vi nulhypotesen og sandsynliggør derved vor alternative hypotese.

Hvis vi tænker os, at vi skulle have bevist den alternative hypotese, er det let at se, at dette ville kræve, at hun svarede rigtigt hvergang, men på denne måde kan vi ikke med et endeligt antal forsøg få visshed, ligesom vi ser, at hvis det blot een gang hændte at hun svarede forkert, ville hypotesen være modbevist.

I (30) behandler Kempthorne spørgsmålet :

"Verifikation of a theory cannot be absolute; we can only show that the observations are compatible with the theory within the limits of error to which the observations are subject. This somewhat negative approach, i.e., being able only to prove a hypothesis false, is the major reason for the use of the null hypothesis in statistics....".

j) I en polemik i Information (45) mellem læge Mogens Fog og nogle yngre læger tages dette problem op :
De yngre læger argumenterer, at eftersom det er borgerne det går ud over, såfremt et stof ved en fejl (type II) ikke påvises at være kræftfremkaldende, selvom det senere viser sig at være det, må man gøre muligheden for sådanne fejl små - udfra synspunktet om, at "tvivlen bør komme borgerne tilgode", hvilket er et synspunkt vi i den foreliggende sammenhæng må være enige i.

k) Vi vil belyse dette lidt mere ved hjælp af et par eksempler :

Om produktionen af tændstikker i pakker ved vi, at der f.eks. er ca. 50 tændstikker i hver æske; hvis vi antager at antallet af tændstikker er normalfordelt med en kendt varians, kan vi altså teste i en stikprøve om middeltallet er 50.

Hvis dette er tilfældet i mere end 95 % af æskerne, accepterer vi produktionen, idet vi ved, at konsekvensen for brugeren er meget lille, selvom vi ved, at der i nogle af æskerne er færre end 50 tændstikker.

Hvis vi danner en parallel til f.eks. sterilisation af kirurgisk værktøj, ville vi absolut ikke acceptere, at der f.eks. kunne forekomme levedygtige smittekim i bare 1 % af pakningerne, og en test for fejl i en sådan produktion skulle altså anvende et meget lille signifikansniveau.

Eller hvad med de uhyre små sandsynligheder for en atomreaktorkatastrofe; hvad hjælper det at de er små, når konsekvenserne af et sådant uheld er uoverskuelige - her er det eneste rimelige at basere sin beslutning på ekstra-statistiske (kvalitative) overvejelser.

1) Den aldersstandardiserede dødsrate er forholdet mellem det observerede antal dødsfald i populationen og det forventede, hvis de forslellige aldersgrupper i populationen havde den aldersspecifikke dødsrate fundet i standardpopulationen.

m)

This is the age of numbers. They appear everywhere – in newspaper reports, broadcasts, advertising material – and are often essential to our work. Whether as laymen or professionals, we have to extract some meaning from all these figures.

Numbers tend to numb our critical faculties; yet, to understand what numbers can tell us, it is not necessary to be a mathematician. Statisticians collect, present, analyse and interpret numerical information. They use mathematical techniques which tend to scare away the non-mathematical mind. But in the statistician's world common sense reigns supreme.

This book describes the common sense of arguing with numbers, concentrating on why the statistician does what he does, rather than on what he does. This should be of interest to the general reader, of direct importance to all who have professional dealings with statisticians, and of value to students in providing a real framework within which to hang their technical studies.

LITTERATURFORTEGNELSE.

- 1) Andersen E.B.
En reanalyse af lungekræfttilfældene i Fredericia.
Ugeskr. Læg. 1974:136:2704-2705
- 2) Andersen E.B.
Ugeskr. Læg. 1975:137:409
- 3) Andersen E.B.
Multiplicative Poisson Models with Unequal Cell Rates.
Scand J Statist 4:153-158,1977
- 4) Anderson D.O.
The Effect of Air Pollution on Health: A Review.
Part II.
Canad. Med. Ass. J.:Sept. 9, 1967, vol 97:585-593
- 5) Barnett V.
Comparative statistical inference.
John Wiley & Sons/William Clowes & Sons
London 1973
- 6) Bickel P.J.
Another look at robustness....
Scand J Statist 3:145-168,1976
- 7) Blom Gunnar
Sannolikhetsteori och statistikteori med tillämpningar.
Lund 1973
- 8) Blot W.J., Brinton L.A., Fraumeni J.F., Stone B.J.
Cancer Mortality in US Counties with Petroleum Industries.
Science, 198,1977,51-53
- 9) Bridges B.A.
Short term screening tests for carcinogens.
Nature, vol 261, 20/5 1976

- 10) Cairns J.
The cancer problem.
Sci. Amer., 233, 64-78, 1975
- 11) Carnow B.W., Meier P.
Air Pollution and Pulmonary Cancer.
Arch Environ Health, 27, 1973, 207-218
- 12) Christensen K.M., Silberschmid M.
Ugeskr. Læg. 1978:140:441
- 13) Clemmesen J.
Statistical studies in malignant neoplasms, IV.
Munksgaard
København 1974
- 14) Clemmesen J., Hansen G., Nielsen A., Røjel J.,
Steensberg J., Sørensen S., Toustrup J.
Lungecancer og luftforurening i Fredericia. En
epidemiologisk undersøgelse.
Ugeskr. Læg. 1974:136:2260-2268
- 15) Clemmesen J., Nielsen A., Røjel J., Steensberg J.,
Sørensen S.
Ugeskr. Læg. 1974:136:2550
- 16) Clemmesen J.
Ugeskr. Læg. 1978:140:380-381
- 17) Cormack R.M.
The statistical argument.
Oliver and Boyd
Edinburgh 1971
- 18) Cox D.R., Hinkley D.V.
Theoretical statistics.
Chapman and Hall
London
- 19) Draper N.R., Smith H.
Applied Regression Analysis.
John Wiley & Sons, Inc.
1966

- 20) Dunn O.J., Clarck V.A.
Applied Statistics: Analysis of Variance and
Regression.
John Wiley & Sons, Inc.
1974
- 21) Eichner P.
Ugeskr. Læg. 1974:136:277o
- 22) Elikofer J., Demsitz M.
Ugeskr. Læg. 1974:136:255o-2551
- 23) Ernst P.
Ugeskr. Læg. 1974:136:255o
- 24) Fisher R.A.
The design of eksperiments.
Oliver and Boyd
Edinburgh 196o
- 25) Friberg L. et al.
Mortality in twins in relation to smoking habits
and alcohol problems.
Arch Environ Health, 27, 1973
- 26) Hald A.
Statistiske Metoder.
Det private ingeniørfond
København 1948
- 27) Hansen J.L.
Ugeskr. Læg. 1974:136:255o
- 28) Horwitz O., Pedersen E.B.
Sammenhængen mellem luftforurening og dødelighed
med særligt henblik på lungesygdomme.
Ugeskr. Læg. 1975:137:2122-2131
- 29) Johansen S.
Forelæsningsnoter i matematisk statistik.
Institut for matematisk statistik, KU
1974

- 30) Kempthorne O.
The design and analysis of eksperiments.
R.E.Krieger
New York 1975
- 31) Kendall M.G., Buckland W.R.
A Dictionary Of Statistical Terms.
Oliver and Boyd
Edinburgh 1971
- 32) Lee, M.L., Prado G.P., Howard J.B., Hites R.A.
Source Identification af Urban Airborne Polycyclic
Aromatic Hydrocarbons by Gas Chromatographic Mass
Spectrometry and High Resolution Mass Spectrometry.
Biomedical Mass Spectrometry vol 4, no. 3, 1977, 182-186
- 33) Lørring Aa.
Ugeskr. Læg. 1974:136:2706
- 34) Nielsen A.
Ugeskr. Læg. 1974:136:2705-2706
- 35) Nielsen A.
Ugeskr. Læg. 1978:140:381
- 36) Pedersen E.B., Horwitz O.
Sammenhængen mellem den almene dødelighed og in-
dustrialiseringen i en række danske provinsbyer.
Ugeskr. Læg. 1976:138:2071-2080
- 37) Pedersen E.B., Horwitz O.
Lungesygdomme og lungefunktion blandt skolebørn
i Fredericia.
Ugeskr. Læg. 1977:139:2833-2840
- 38) Pedersen E.B., Horwitz O.
Ugeskr. Læg. 1978:140:381
- 39) Pedersen E.B., Horwitz O.
Ugeskr. Læg. 1978:140:441

- 40) RUC-avisen
forår 1978
- 41) RUC-rapport, Jørn Jensen og Michael Olsen
Matematiks rolle i gymnasium og HF 1977.
- 42) RUC-rapport, Crilles Bacher et al.
Statistiske standardmetoder.
- 43) Wilson K.
Ugeskr. Læg. 1974:136:2551-2552
- 44) Yamane T.
Statistics, an introductory analysis.
Harper & Row
New York 1973
- 45) Gaarboe O. et al.
Sidste vers om lægevidenskab og politik.
(samt tidligere artikler i polemikken, angivet
i artiklen).
Information 15.05.1978

BILAG 1.

Lineær regressionsanalyse med to uafhængigt variable.

Forudsætningerne for regressionsanalysen er følgende:

- 1) For enhver fast værdi af (x_1, x_2) er y normalt fordelt.
- 2) Middelværdien af y , $\eta = E(y|x_1, x_2)$, er en lineær funktion af x_1 og x_2
$$\eta = \alpha + \beta_1(x_1 - \bar{x}_1) + \beta_2(x_2 - \bar{x}_2)$$
- 3) Variansen af y , $\text{var}(y|x_1, x_2)$, er uafhængig af (x_1, x_2)
- 4) Iagttagelsesresultaterne er stokastisk uafhængige.

Af iagttagelsesresultaterne skal beregnes skøn over parametrene α , β_1 , β_2 og σ^2 .

Kvadratsummen af iagttagelsesresultaternes afvigelser fra de teoretiske værdier, $\sum (y_i - \eta_i)^2$, er fordelt som $\sigma^2 \chi^2$ med n frihedsgrader, og skønnene over α , β_1 og β_2 bestemmes efter de mindste kvadraters metode, således at kvadratsummen af afvigelserne mellem de foreliggende iagttagelsesresultater og den empiriske regressionsplan bliver mindst mulig. Den empiriske regressionsligning betegnes

$$Y = a + b_1(x_1 - \bar{x}_1) + b_2(x_2 - \bar{x}_2)$$

Ud fra følgende tre ligninger beregnes skønnene a , b_1 og b_2 :

$$a = \bar{y}$$

$$b_1 \text{SAK}_{x_1} + b_2 \text{SAP}_{x_1 x_2} = \text{SAP}_{x_1 y}$$

$$b_1 \text{SAP}_{x_1 x_2} + b_2 \text{SAK}_{x_2} = \text{SAP}_{x_2 y}$$

$$(\text{SAK}_{x_1} = \sum (x_{1i} - \bar{x}_1)^2 \text{ og } \text{SAP}_{x_1 y} = \sum (x_{1i} - \bar{x}_1)(y_i - \bar{y}))$$

Variansen, der bygger på iagttagelsesresultaternes afvigelser fra de empiriske regressionsværdier, er

$$s^2 = \frac{1}{n-3} \sum (y_i - Y_i)^2$$

og er s^2 -fordelt ($1/\sigma^2 \chi^2$ -fordelt) med parametrene $(\sigma^2, n-3)$

Skønnet a er normalt fordelt med middelværdien a og

variansen σ^2/n .

Skønnene (b_1, b_2) er korrelerede med middelværdierne (β_1, β_2) , varianserne

$$\text{var}(b_1) = \sigma^2 \frac{\text{SAK}_{x_2}}{D_{x_1 x_2}} \quad \text{var}(b_2) = \frac{\text{SAK}_{x_1}}{D_{x_1 x_2}}$$

$$\text{og kovariansen } \text{cov}(b_1, b_2) = -\sigma^2 \frac{\text{SAP}_{x_1 x_2}}{D_{x_1 x_2}}$$

$$\text{hvor } D_{x_1 x_2} = \text{SAK}_{x_1} \text{SAK}_{x_2} - (\text{SAP}_{x_1 x_2})^2$$

Et test for signifikans kan opstilles, således at en hypotetisk værdi for f.eks. β_1 testes ved beregning af

$$t = \frac{b_1 - \beta_1}{s_{b_1}} \quad \text{hvor } s_{b_1} = s \frac{\sqrt{\text{SAK}_{x_1}}}{\sqrt{D_{x_1 x_2}}}, \quad f = n - 3$$

der er t-fordelt med $n-3$ frihedsgrader.

Ønskes et test for de bestemte værdier (β_1, β_2) , der tager hensyn til korrelationen mellem b_1 og b_2 , benyttes størrelsen

$$F = \frac{1}{2s^2} ((b_1 - \beta_1)^2 \text{SAK}_{x_1} + 2(b_1 - \beta_1)(b_2 - \beta_2) \text{SAP}_{x_1 x_2} + (b_2 - \beta_2)^2 \text{SAK}_{x_2})$$

Ønskes specielt at teste hypotesen $(\beta_1, \beta_2) = (0, 0)$ beregnes

$$F = \frac{1}{2s^2} (b_1 \text{SAP}_{x_1 y} + b_2 \text{SAP}_{x_2 y})$$

Dette sidste test beror på en spaltning af den totale variation af y 'erne i den del, der kan beregnes ud fra (x_1, x_2) -værdierne og restvariationen, der skyldes andre faktorer, idet de tilsvarende varianser sammenlignes ved hjælp af F-testet.

Ved anvendelse af regressionsanalyse til bestemmelse af sammenhængen mellem flere variable, i tilfælde hvor variationer af de uafhængigt variable i følge den faglige viden om sammenhængen må betragtes som årsag til til variationer af den afhængigt variable, har det ofte interesse ud fra regressionskoefficienternes indbyrdes størrelse at vurdere de forskellige faktors betydning for forklaringen af den afhængigt variables

variationer. Sammenligningen af regressionskoefficienternes størrelse kan suppleres med en undersøgelse af ændringerne i variansernes størrelse, idet tilføjelsen af en variabel, hvis variation har stor indflydelse på den afhængigt variables variation, bevirker en formindskelse i iagttagelsesresultaternes afvigelser fra regressionsplanen, dvs. i variansen af y .

BILAG 2.

Regressionsanalyse udført på data fra (11).

Alle tal i beregningerne er angivet med tre decimaler uden hensyntagen til den fjerde.

Country	Age, yr						Consumption	
	35-74	25-34	35-44	45-54	55-64	65-75	Cigarettes	Solid Fuel
Australia	622.8	8.67	60.0	305.4	1,100	1,771	2,040	2,396
Austria	1,186.7	17.34	58.8	571.8	2,217	3,343	1,520	1,231
Belgium	857.3	12.91	79.2	567.8	1,624	2,015	1,410	2,748
Canada	603.9	5.11	49.7	344.1	1,071	1,640	2,770	1,708
Denmark	674.0	1.98	74.0	399.2	1,264	1,664	1,220	1,290
Finland	1,465.6	10.58	110.5	767.2	2,647	4,061	1,850	625
France	516.3	8.16	53.9	317.5	969	1,258	1,290	1,713
Germany (West)	839.5	10.13	63.1	448.1	1,623	2,159	1,250	3,047
Ireland	609.8	8.55	111.4	402.3	1,130	1,351	2,430	621
Italy	497.9	12.20	64.4	377.4	1,026	924	1,480	232
Japan	219.1	4.48	23.1	97.5	349	695	1,630	565
Netherlands	973.3	14.10	96.2	593.8	1,847	2,364	1,750	1,589
New Zealand	712.4	12.58	49.8	385.7	1,175	2,117	1,860	638
Norway	286.9		21.3	157.9	659	578	550	411
South Africa	757.9	12.09	77.1	390.6	1,373	2,062	1,220	2,226
Sweden	337.8	7.06	36.8	153.4	636	918	1,050	742
Switzerland	745.7	8.26	77.2	475.0	1,304	1,923	1,920	580
United Kingdom	1,556.3	22.39	183.3	897.4	2,822	3,998	2,590	4,129
United States	772.8	13.31	100.5	472.6	1,337	2,005	3,440	2,260

Mortality values expressed as age-adjusted rates per 1 million population. Cigarette consumption expressed as pieces per capita over 15 years of age in 1957; solid fuel consumption, kilograms per capita in 1955 to 1958. Data from stocks.

	x_1	x_2	y
n=19			
S	33.220	28.751	14.236
SK	66.391	63.856	12.938
S^2/n	58.082	43.506	10.666
SAK	8.309	20.350	2.272
	$x_1 x_2$	$x_1 y$	$x_2 y$
SP	54.071	26.337	24.847
$S_1 S_j / n$	50.268	24.890	21.542
SAP	3.803	1.447	3.305
	x_1	x_2	y
S/n	1.748	1.513	0.749

$$b_1 \text{SAK}_{x_1} + b_2 \text{SAP}_{x_1 x_2} = \text{SAP}_{x_1 Y}$$

$$b_1 \text{SAP}_{x_1 x_2} + b_2 \text{SAK}_{x_2} = \text{SAP}_{x_2 Y}$$

$$8.309b_1 + 3.803b_2 = 1.447$$

$$3.803b_1 + 20.350b_2 = 3.305$$

$$\underline{b_1 = 0.109}$$

$$\underline{b_2 = 0.142}$$

$$Y = a + b_1(x_1 - \bar{x}_1) + b_2(x_2 - \bar{x}_2) \quad a = \bar{y}$$

$$= 0.749 + b_1(x_1 - 1.748) + b_2(x_2 - 1.513)$$

$$= \underline{0.345 + 0.109x_1 + 0.142x_2}$$

$$\text{SAK}_{Y|x_1, x_2} = \text{SAK}_Y - b_1 \text{SAP}_{x_1 Y} - b_2 \text{SAP}_{x_2 Y}$$
$$= 1.706$$

$$f = 19 - 3 = 16$$

$$\underline{s^2 = 0.106} \quad \underline{s = 0.326}$$

$$\underline{s_a^2 = 0.005} \quad \underline{s_a = 0.074}$$

$$D_{x_1 x_2} = 154.625$$

$$\underline{s_{b_1}^2 = 0.013} \quad \underline{s_{b_1} = 0.117}$$

$$\underline{s_{b_2}^2 = 0.005} \quad \underline{s_{b_2} = 0.075}$$

Tager vi nu som udgangspunkt $x_1 = x_2 = 1.000$, vil en forøgelse på en enhed x_1 give en forøgelse i y på 18.2%, og en forøgelse på en enhed x_2 vil give en forøgelse i y på 23.8%. Hvis udgangspunktet for x_1 og x_2 ikke er 1.000 vil forøgelsen blive mindre, dog ikke mere end nogle få procent.

Y -værdierne i atbellen er udtrykt i antal lungekræfttilfælde pr. 1000.

En kontrol af, om forudsætningerne for regressionsanalysen er opfyldt for det givne iagttagelsesmateriale, kan fås på følgende måde: størrelsen $(y - Y)/s$ udregnes for hver observation. I følge forudsætningerne skal de normerede afvigelser med tilnærmelse være normalt for-

delt med parametrene $(0,1)$. Ved iagttagelse af $(y-Y)/s$ fås et overblik over størrelsesvariationen og fortegnsvexlingen for $(y-Y)/s$ som funktion af (x_1, x_2) , således at eventuelle systematiske afvigelser fra forudsætningerne kan konstateres.

Y	y	$(y-Y)/s$
0.907	0.6228	-0.871
0.685	1.1867	1.538
0.888	0.8573	-0.094
0.884	0.6039	-0.859
0.661	0.674	0.039
0.635	1.4656	2.547
0.728	0.5163	-0.649
0.913	0.8395	-0.225
0.698	0.6098	-0.270
0.539	0.4979	-0.126
0.602	0.2191	-1.174
0.761	0.9733	0.651
0.638	0.7124	0.228
0.463	0.2869	-0.540
0.796	0.7579	-0.116
0.564	0.3378	-0.693
0.636	0.7457	0.336
1.213	1.5563	1.053
1.040	0.7728	-0.819

BILAG 3.

Lineær regressionsanalyse med m uafhængigt variable.

Denne regressionsanalyse er i princippet identisk med regressionsanalyse med to uafhængigt variable. (bilag 1).

Den uafhængigt variable forudsættes normalt fordelt med middelværdien

$$\eta = \alpha + \beta_1(x_1 - \bar{x}_1) + \beta_2(x_2 - \bar{x}_2) + \dots + \beta_m(x_m - \bar{x}_m)$$

og variansen σ^2

Skønnet over regressionsligningen betegnes

$$Y = a + b_1(x_1 - \bar{x}_1) + b_2(x_2 - \bar{x}_2) + \dots + b_m(x_m - \bar{x}_m)$$

og skønnet over parametrene bestemmes ved de mindste kvadraters metode til

$$a = \bar{y}$$

$$b_1 \text{SAK}_{x_1} + b_2 \text{SAP}_{x_1 x_2} + \dots + b_m \text{SAP}_{x_1 x_m} = \text{SAP}_{x_1 Y}$$

$$b_1 \text{SAP}_{x_2 x_1} + b_2 \text{SAK}_{x_2} + \dots + b_m \text{SAP}_{x_2 x_m} = \text{SAP}_{x_2 Y}$$

.

.

$$b_1 \text{SAP}_{x_m x_1} + b_2 \text{SAP}_{x_m x_2} + \dots + b_m \text{SAK}_{x_m} = \text{SAP}_{x_m Y}$$

Variansen

$$s^2 = \frac{1}{n-m-1} \sum (y_i - \hat{y}_i)^2$$

er s -fordelt ($1/f \cdot \sigma^2 \chi^2$ -fordelt) med parametrene $(\sigma^2, n-m-1)$.

Skønnet a er normalt fordelt med middelværdien a og variansen σ^2/n .

Skønnene (b_1, b_2, \dots, b_m) er normalt korrelerede med middelværdierne $(\beta_1, \beta_2, \dots, \beta_m)$ og varianserne

$$\text{var}(b_i) = c_{ii} \quad i=1, 2, \dots, m$$

c'erne defineres på følgende måde:

I førnævnte ligningssystem multipliceres den første af ligningerne med c_{11} , den anden med c_{12} osv., og efter addition af de m ligninger bestemmes $c_{11}, c_{12}, \dots, c_{1m}$ således, at koefficienterne til b_1 i sumligningen er

lig 1 og koefficienterne til b_2, \dots, b_m er lig nul. c 'erne bestemmes altså af det ligningssystem, der fremkommer af førnævnte ligningssystem når b 'erne erstattes af c 'er og tallene på ligningens højre side erstattes af tallene $0, 0, \dots, 0, 1, 0, \dots, 0$ hvor tallet står på den i 'te plads i rækken.

Endelig er skønnene s^2 , a og (b_1, b_2, \dots, b_m) stokastisk uafhængige, hvilket bl.a. medfører at et margialt test for hypotesen β_i er størrelsen

$$t = \frac{b_i - \beta_i}{s_{b_i}}, \quad s_{b_i} = s\sqrt{c_{ii}}$$

der er t -fordelt med $n-m-1$ frihedsgrader.

Et test for et specificeret værdisæt $(\beta_1, \beta_2, \dots, \beta_m)$ er størrelsen

$$F = \frac{1}{ms^2} \left(\sum (b_i - \beta_i)^2 \text{SAK}_{x_i} + 2 \sum \sum (b_i - \beta_i) (b_j - \beta_j) \text{SAP}_{x_i x_j} \right)$$

som er F -fordelt med $(m, n-m-1)$ frihedsgrader.

BILAG 4.

F-test.

I det foreliggende tilfælde interesserer vi os kun for test af hypotesen $\beta_i = 0, i=1,2,3,\dots$

Testet bygger, som nævnt under omtalen af lineær regression med to uafhængigt variable (bilag 1), på spaltning af den totale variation af y'erne i den del, der kan beregnes ud fra (x_1, x_2) -værdierne og restvariationen.

Sædvanligvis opstilles følgende skema:

VARIATION	KVADRATSUM	FRIHEDSGRADER	VARIANS	TEST
mellem regressionsværdierne	$\Sigma (Y_i - \bar{Y})^2$	p	s_2^2	$F = \frac{s_2^2}{s_1^2}$
restvariationen	$\Sigma (y_i - Y_i)^2$	n-1-p	s_1^2	
total	$\Sigma (y_i - \bar{Y})^2$	n-1		

hvor p er det antal variable, der er inkluderet i regressionsligningen.

Varianserne s_2^2 og s_1^2 fremkommer ved at dividere kvadratsummen med antal frihedsgrader, og F er tabellret.

BILAG 5.

Partielt F-test og sekventielt F-test.

Hvis der er flere udtryk i en regressionsligning, kan man forestille sig disse indsat i ligningen i en bestemt rækkefølge. Finder vi kvadratsummen af b_i under forudsætning af de andre værdier i regressionsligningen, altså

$SAK(b_i | b_0, b_1, \dots, b_{i-1}, b_{i+1}, \dots, b_k), i=1, 2, \dots, k$
vil vi have en kvadratsum med én frihedsgrad, som udtrykker bidraget til regressionskvadratsummen af hver koefficient b_i , under forudsætning af at alle de udtryk, som ikke involverer β_i , allerede er i modellen. Med andre ord skal vi have et udtryk for værdien af at tilføje et udtryk β_i til modellen.

Den korresponderende gennemsnitskvadratsum (der er lig kvadratsummen, idet der kun er en frihedsgrad), kan sammenlignes med s^2 ved et F-test.

Denne type F-test kaldes et partielt F-test af β_i . Når variable tilføjes en efter en i trin til regressionsmodellen tales om et sekventielt F-test. Dette er altså blot navnet for et partielt F-test på den variable, som netop er tilføjet regressionsligningen.

BILAG 6.

Simpel korrelationskoefficient.

Den simple korrelationskoefficient mellem X_i og X_j benævnes ofte ρ_{ij} , og skønnet herpå er

$$r_{ij} = \frac{\sum X_i X_j}{\sqrt{\sum X_i^2 \sum X_j^2}}$$

BILAG 7.

Multipel korrelationskoefficient.

For at få en idé om lineariteten mellem y og x_1, x_2, \dots, x_k beregnes de simple korrelationskoefficienter (bilag 6). Som sæt betragtet er dette imidlertid vanskeligt at tolke.

For de mindste kvadraters linie $Y = a + b(x - \bar{x})$ beregnes korrelationskoefficienten mellem y (det aktuelle y) og Y (y forudsagt for x fra den empiriske regressionslinie), hvilket er det samme som korrelationskoefficienten mellem y og x bortset fra en mulig fortegnforskelle, dvs. at

$$r_{yY} = \pm r_{xy}$$

Den fra prøven stammende multipel korrelationskoefficient mellem y og x_1, x_2, \dots, x_k defineres som den simple korrelationskoefficient mellem y og dens forudsagte værdi Y .

Den benævnes $R_{y:1\dots k}$ eller blot R .

Altså

$$R_{y:1\dots k} = r_{yY}$$

$$\text{hvor } Y = \bar{y} + b_1 x_1 + b_2 x_2 + \dots + b_k x_k$$

$R_{y:1\dots k}$ er den højeste korrelation mellem y og enhver lineær kombination af x_1, x_2, \dots, x_k .

Dette betyder, at man ved at udelukke et delsæt af X -variablerne fra regressionen ikke kan øge multipel korrelationskoefficienten.

Koefficienten behandles hovedsageligt som ikke-negativ, hvorfor R^2 oftest optræder i praksis.

BILAG 8.

Trinvis lineær regressionsanalyse.

Trinvis lineær regression er en specifik statistisk procedure til at udvælge variable i regressionen. Antages det, at man ønsker at etablere en lineær regressionsligning for en bestemt respons Y som funktion af uafhængigt variable X_1, X_2, \dots, X_k , vil to modstridende kriterier for at vælge en ligning sædvanligvis være involveret.

- 1) For at ligningen skal være god til forudsigelser, må modellen indeholde så mange X 'er som muligt, således at pålidelige værdier kan bestemmes.
- 2) På grund af omkostningerne ved at hente information fra et stort antal X 'er og ved behandlingen af disse, må man ønske at ligningen omfatter så få X 'er som muligt.

Kompromiset mellem disse to ekstremer kaldes sædvanligvis "udvælgelse af den bedste regressionsligning".

Der findes ingen entydig statistisk procedure til dette, og personlig bedømmelse er en nødvendig del af de statistiske metoder, der kan benyttes.

En metode, der kan anvendes, og som er pålidelig, er trinvis regression.

Proceduren i metoden er at indsætte variable en efter en, indtil regressionsligningen er tilfredsstillende. Rækkefølgen bestemmes ved benyttelse af den partielle korrelationskoefficient (bilag 9), som et udtryk for vigtigheden af de variable, der endnu ikke optræder i ligningen. En variabel, der er indsat i ligningen på et tidligt tidspunkt, kan dog senere forkastes, idet hver variabels betydning F -testes på hvert trin af analysen. Enhver variabel, som ikke giver et signifikant bidrag, fjernes fra modellen. Processen fortsættes, til der ikke er flere variable at tillade i ligningen, og ingen der kan fjernes.

Proceduren kan konkretiseres en smule mere:

1. trin En korrelationsmatrix opstilles, og den variabel der er højest korreleret med responsen, indsættes i ligningen. Den indsatte variabel testes altid for signifikans.

2. trin Partielle korrelationskoefficienter beregnes for alle de variable, der ikke optræder i regressionsligningen.

3. trin Her undersøges det bidrag det først indsatte X (f.eks. X_1) ville have givet, hvis den andet indsatte (f.eks. X_2) var kommet i regressionsligningen først. Det vil sige, at der laves partiel F-test på alle indsatte variable. Hvis der ikke er signifikans, forkastes pågældende variabel.

Proceduren fortsættes til der ikke er flere variable, der tillades indsættelse og ikke flere, der forkastes.

BILAG 9.

Partiel korrelationskoefficient.

Den partielle korrelationskoefficient er korrelationskoefficienten mellem to variable under forudsætning af andre variable.

Udtrykket for den partielle korrelationskoefficient for (x_i, x_j) i den betingede fordeling, når man betinger med x_k er følgende

$$\rho_{ij \cdot k} = \frac{\rho_{ij} - \rho_{ik} \rho_{jk}}{\sqrt{(1 - \rho_{ik}^2)(1 - \rho_{jk}^2)}}$$

Ønskes f.eks. korrelationen mellem X_i og X_j under forudsætning af X_k og X_m fås udtrykket

$$\rho_{ij \cdot km} = \frac{\rho_{ij \cdot k} - \rho_{im \cdot k} \rho_{jm \cdot k}}{\sqrt{(1 - \rho_{im \cdot k}^2)(1 - \rho_{jm \cdot k}^2)}}$$

Skønnene over parametrene $\rho_{ij \cdot k}$ og $\rho_{ij \cdot km}$ er henholdsvis

$$r_{ij \cdot k} = \frac{r_{ij} - r_{ik} r_{jk}}{\sqrt{(1 - r_{ik}^2)(1 - r_{jk}^2)}}$$

og

$$r_{ij \cdot km} = \frac{r_{ij \cdot k} - r_{im \cdot k} r_{jm \cdot k}}{\sqrt{(1 - r_{im \cdot k}^2)(1 - r_{jm \cdot k}^2)}}$$

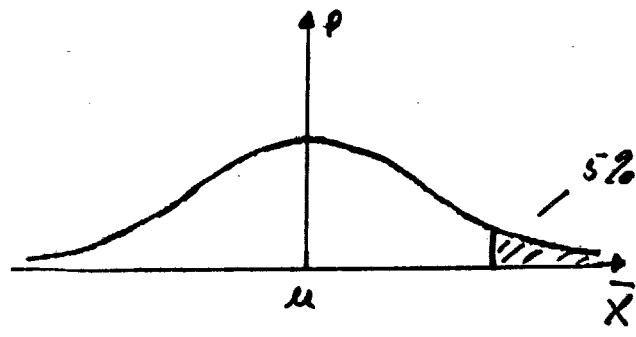
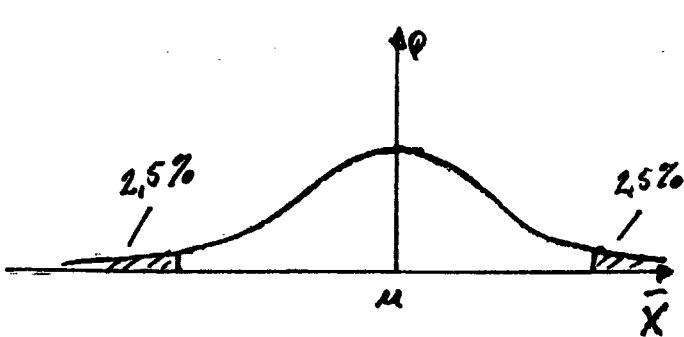
BILAG 10

For at forklare forskellen på en- og tosidede tests, har vi valgt at bruge et u-test som eksempel, bla. fordi det er et u-test som er brugt af Clemmesen i Fredericiaundersøgelsen.

vi tænker os at vi har nogle observationer fra en normalfordelt stokastisk variabel X; vi tænker os endvidere at variansen er kendt, og vi kan nu ønske at teste, om middelværdien i vore observationer er lig en given middelværdi under en hypotese, dvs. $H_0 : \bar{X} = u$, se fig.

tosidet

ensidet



$$H_0 : \bar{X} = u$$
$$H_1 : \bar{X} \neq u$$

$$H_0 : \bar{X} = u$$
$$H_1 : \bar{X} > u$$

Nu er man ofte mere interesseret i den såkaldte alternative hypotese, som sandsynliggøres (accepteres) såfremt vor nulhypotese forkastes; en alternativ hypotese til H_0 kunne være $H_1 : \bar{X} > u$.

Ved at udregne \bar{X} kan vi opstille teststørrelsen

$$z = \frac{\bar{X} - u}{\sigma} \quad \text{som så er normalfordelt } n(0,1).$$

Vi kan nu ud fra tabellen over samørende værdier af z og p (*) finde sandsynligheden for at vor hypotese H_0 holder: bliver z stor (det gør den, hvis $\bar{X} - u$, dvs. forskellen mellem den observerede og den hypotetiske middelværdi, er stor), aflæser vi en stor p, hvilket er et mål for sandsynligheden for at H_0 er

(*) næste side

falsk.

Vi kan imidlertid have en formodning om at \bar{X} er større end u , og vi opstiller derfor vor alternative hypotese $H_1 : \bar{X} > u$; skal vi teste denne, er vi altså kun interesseret i den ene hale af sandsynlighedsfordelingen (se fig).

En ensidet test forudsætter (selvfølgelig) at \bar{X} rent faktisk er større end u - dette er også tilfældet i fredericiaundersøgelsen (se side45).

Valget af ensidet/tosidet test må ske på baggrund af de alternative hypoteser man ønsker at teste; disse hypoteser skal dannes udfra fagligviden om fænomenet, (se note b).

I en x^2 -test er man kun i stand til at teste tosidet (idet forskellen mellem den observerede og den hypotetiske værdi kvadreres), hvilket forklarer, at E.B. Andersen i sin reanalyse af Clemmesens data i fredericiaundersøgelsen tester ensidet.

fodnote (*) : p er at forstå som sandsynligheden for at få en værre stikprøve, under den givne hypotese; i nogle tabeller aflæses $(1-p)$ - her skal ovenstående fortolkning bare "vendes om".

BILAG 11.

Foruden "den klassiske" statistik skabt af Pearson, Fisher, Yates m.fl. findes to mindre skoler, som bygger på nogle lidt andre principper for statistisk inferens, og vi vil her lige omtale dem meget summarisk af hensyn til ikke statistikere, som måtte læse rapporten.

Non-parametrisk statistik er som navnet antyder ikke baseret på teoretiske fordelinger, som kan beskrives ved hjælp af parametre (middelværdi, varians m.m.); i stedet forsøger man udfra forskellige antagelser om fordelingen af den (de) stokastiske variable (f.eks. symetri, linearitet m.m.) at danne sig et billede af den teoretiske fordeling; på grundlag af disse antagelser opstiller man en teststørrelse (som regel sørger man for, at denne teststørrelse har en kendt fordeling) hvorefter det bliver muligt at udføre visse tests, f.eks. fortegnstest.

Bayesiansk statistik bygger på en antagelse om, at tidligere statistisk materiale om et givet fænomen (f.eks. tidligere fundne middelværdier eller varians) bør indgå i den inferens om de teoretiske fordelinger, som er baseret direkte på stikprøven.

Denne antagelse kan synes meget rimelig, men nogle fortalere for den bayesianske skole går videre, og mener, at endvidere "intuition" - dvs. antagelser, som ikke er direkte begrundet i talmaterialet, men snarere begrundet i en forskers "dybgående indsigt" i en given problemstilling - kan indgå i inferensen om de teoretiske fordelinger.

Denne skole er genstand for en hæftig debat (bla. 06), idet de klassiske statistikere ofte forholder sig ret kritisk til denne form for statistisk inferens.

BILAG 12.

Vi har binomialfordelingen givet ved:

$$\binom{n}{x} \cdot p^x \cdot q^{n-x} = \frac{1}{x!} \cdot \frac{n!}{(n-x)!} \cdot p^x q^{n-x}$$

I de følgende betragtninger forudsætter vi, at p er lille, x er fast og n er meget stor:

$p \rightarrow 0$, $x = k \ll n$, $n \rightarrow \infty$, $n \cdot p \rightarrow m$ (m er en konstant).

Af ovenstående få ved division med n^x i første del af udtrykket og multiplikation i anden del:

$$\frac{n(n-1)(n-2)\dots(n-x+1)}{n^x} \cdot n^x \cdot \frac{1}{x!} \cdot p^x \cdot q^{n-x} =$$

$$1 \cdot \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{x-1}{n}\right) \cdot n^x \cdot p^x \cdot q^{n-x} \cdot \frac{1}{x!}$$

Vi ser nu, at når $n \rightarrow \infty$ går alle leddene $1 \cdot \left(1 - \frac{1}{n}\right) \dots$ mod 1.

Vi sætter endvidere $np = m$, og får:

$$\frac{1}{x!} \cdot m^x \cdot q^{n-x}$$

idet $q = (1-p)$ fås

$$q^{n-x} = \frac{(1-p)^n}{(1-p)^x}$$

men i følge vore forudsætninger er dette tilnærmelsesvis lig $(1-p)^n$. Endvidere er $(1-p)^n = e^{n \ln(1-p)}$.

Vi rækkeudvikler nu $n \ln(1-p)$ efter Taylor:

$$n \cdot \ln(1-p) \approx -np + n \cdot \frac{p^2}{2} + \dots \approx -np = -m$$

og vi får

$$e^{n \cdot \ln(1-p)} \rightarrow e^{-m}$$

Vi har nu i alt

$$\binom{n}{p} \cdot p^x q^{n-x} \rightarrow \frac{1}{x!} \cdot m^x \cdot e^{-m}$$

hvilket som bekendt er en Poissonfordeling.

Poissonfordeling/normalfordeling :

Et resultat af den centrale grænseværdisætning er at en Poissonfordeling kan aproximeres med en normalfordeling, og ræsonnementet er følgende :

En Poissonfordeling med parameter m kan opfattes som summen af en række Poissonfordelinger med parametre $\frac{m}{n}$; fra den centrale grænseværdisætning har vi, at summen af en række ens fordelte uafhængige observationer er tilnærmelsesvis normalfordelt.

BILAG 13.

Kronologisk oversigt over Fredericiaundersøgelsen og
hvad deraf fulgte.

- 30.09.74 Lungecancer og luftforurening i Fredericia.
En epidemiologisk undersøgelse.
- 04.11.74 Lungecancer i Fredericia.
a) korrespondance
b) korrespondance
c) svar
d) korrespondance
- 04.11.74 Lungecancer i Strandby (Frederikshavn).
- 25.11.74 En reanalyse af lungekræfttilfældene i
Fredericia.
- 25.11.74 Lungekræfttilfældene i Fredericia.
a) korrespondance
- 25.11.74 Lungecancer i Fredericia.
a) korrespondance
- 02.12.74 Lungecancer i Fredericia.
a) korrespondance
- 10.02.75 Lungekræfttilfældene i Fredericia.
a) korrespondance
- 08.09.75 Sammenhængen mellem luftforurening og døde-
lighed med særligt henblik på lungesygdomme.
- 16.08.76 Sammenhængen mellem den almene dødelighed og
industrialiseringen i en række danske pro-
vinsbyer.
- januar77 Multiplicative Poisson Models with Unequal
Cell Rates.
- 21.11.77 Lungesygdomme og lungefunktion blandt skole-
børn i Fredericia.

13.o2.78 Luftvejssygdomme blandt skolebørn i Fredericia.

a) korrespondance

13.o2.78 Lungesygdomme og lungefunktion blandt skolebørn i Fredericia.

a) korrespondance

b) svar

20.o2.78 Lungesygdomme og lungefunktion blandt skolebørn i Fredericia.

a) korrespondance

b) svar

BILAG 14.

Materiale til undervisning.

I forbindelse med vore overvejelser over relevant undervisningsmateriale foretog vi en "research" i aviskronikindexet på biblioteket, og fandt her en række artikler, som vil kunne bruges i undervisningen i statistik, særlig i forbindelse med temaarbejde.

Vi havde egentligt tænkt os at lave resumeer af artiklerne, men på grund af tidsnød har vi valgt i stedet blot at angive, hvor de kan findes :

Berlingske Tidende 27/9 - 74 : "Kræften i Fredericia" af Jørgen Røjel.

Information 17/5 - 78 : "Sidste vers om lægevidenskab og politik" af Ove Gaarde - samt tidligere artikler i polemikken, som kan findes i artiklen.

Politiken 26/4 - 74 : "At leve sundt - og længe, om industrikultur/folkesundhed, af G. Magnussen.

Aktuelt 6/10 - 74 : "Hvorfor bliver en normal celle en kræftcelle" af Karl Stefan.

Akruelet 10/9 - 74 : "Chokerende læsning - lyver statistikken om arbejdsulykker", af F. Bengtson.

Ålborg Stiftetidende 3/11 - 74 : "Den forkætrede positivliste", af S.S. Wittrup.

Børsen 16/11 - 74 : "Fagbevægelsen starter nu sin egen forskning i sygdomme på arbejdspladsen", af Hans Solholm.

Ovenstående vil hurtigt kunne suppleres med flere.