OVERSIGT OVER DET KGL. DANSKE VIDENSKABERNES SELSKABS FORHANDLINGER. 1908. N° 3

„DEN APOSTERIORISKE SANDSYNLIGHED"

AF

K. KROMAN

I.

Der har i den senere Tid hos flere danske Matematikere vist sig en ikke ringe Skepsis overfor den saakaldte aposterioriske Sandsynlighed og den dermed nøje sammenhængende Bayes' Regel. Saaledes har Professor Thiele baade i sin „Almindelig Iagttagelseslære" (S. 55 og 77) og i sin „Elementær Iagttagelseslære" (S. 118) benyttet forskellige stærkt afvisende Ytringer om hele denne Afdeling af Sandsynligheds-læren, og da General Zachariæ 1887 udsendte anden Udgave af sin „De mindste Kvadraters Metode", var Bayes' Regel, der fandtes i første Udgave, rent udeladt, og selve Metoden havde faaet en ny Begrundelse støttet til det nye Begreb „den relative Sandsynlighed". Den første Aarsag til denne Skepsis synes at ligge i en Afhandling af Direktør F. Bing i „Tidsskrift for Matematik" for 1879. Ogsaa her falder der meget haarde Ord om Bayes' Regel og den aposterioriske Sandsynlighed, og Prof. Thiele, der i sine nysnævnte Skrifter flere Gange henviser til denne Afhandling, udtaler endog (Alm. Iagttagelsesl. S. 77), at Forfatteren her har „modbevist denne Læres Rigtighed."

Der er efter min Opfattelse et Moment af Sandhed i alt dette; men det er ingenlunde ubetinget rigtigt. En Mængde af de benyttede Udtryk bliver i al deres Ubestemthed ganske uretfærdige. Der findes jo nemlig, som jeg skal søge at paavise, baade en fuldstændig rigtig og en ganske falsk „Bayes'

Regel", baade en værdifuld og en ganske meningsløs „aposteriorisk Sandsynlighed". Direktør Bing har klart set dette i Begyndelsen af sin Afhandling; men han har ikke klart fastholdt Forskellen. Ved at begaa forskellige Regnefejl støder han paa saa afskrækkende Modsigelser, at han tilsyneladende ender i udstrakt Tvivl overfor Reglen eller i ethvert Tilfælde udtaler sig, saa at hans Læsere nærmest maa faa denne Opfattelse.

Sagen har imidlertid saa stor baade matematisk og erkendelsesteoretisk Interesse, at det dog var værd at prøve, om ikke alle de fremhævede Modsigelser og øvrige Vanskeligheder strax forsvinder, saa snart den antydede Adskillelse virkelig gennemføres. I de fleste Lærebøger findes endnu den sande og den falske Regel i broget Blanding.

II.

Ligesom i de fleste reale Videnskaber, saaledes benævner man ogsaa ofte i Matematikken en Sætning, ikke efter dens Indhold, men efter den Person, der først har fremsat den. Men ligesom man paa matematisk Omraade aldrig godkender en Sætning alene af mulig stærk Tillid til dens Ophavsmand, saaledes finder man det her heller ikke fornødent vedblivende at fastholde Sætningen i netop dens oprindelige Fremtrædelsesform. Finder man en Dag et kortere og klarere Bevis for dens Gyldighed, tager man uden videre dette i Brug, og finder man et tydeligere og skarpere Udtryk for selve Paastanden, et Udtryk, der maaske afskærer en eller anden tidligere Urigtighed eller ombytter en mindre naturlig Afgrænsning for dens Omraade med en rimeligere, saa indfører man uden Betænkning saadanne og lignende Forbedringer uden derfor at lade Sætningen skifte Navn, saafremt dens væsenlige Indhold endnu er det samme. Det er Kærnen og ikke Skallerne, der her spiller den afgørende Rolle.

Ofte kan en saadan redaktionel Virksomhed endog være særlig fornøden. Vi kender saaledes ikke Bayes' Regel i Ophavsmandens egen Formulering. Den blev først offenliggjort i „Philosophical Transactions“ (Vol. 53) for 1763. Men Bayes var da nylig død. Det var hans Ven Richard Price, der havde fundet Manuskriptet blandt hans Efterladenskaber og indsendt det til „Royal Society“ med forskellige Rettelser og Tilføjelser. I den medfølgende Skrivelse fortæller Price, at Bayes havde lagt Afhandlingen til Side, fordi han endnu ikke var fornøjet med et vist Punkt i den og nødig vilde udgive noget, som kunde vente Indvendinger. Price's Rettelser angaar imidlertid ikke dette Punkt, men forskellige Udregninger. Reglen foreligger altsaa her i en for selve dens Ophavsmand mindre tilfredsstillende Form, og den er senere fremsat i flere stærkt forskellige Former. En vis Frihed i Formuleringen er derfor her dobbelt tilladelig, og jeg gaar da over til at fremstille den med den Skærpelse, jeg antager, Bayes selv vilde have funden fornøden, om det var blevet forundt ham yderligere at beskæftige sig med den.

Lad A være en mulig Tilstand, som kan optræde under den ene eller den anden af de hinanden udelukkende Former $A_1, A_2, \ldots A_k$, og ellers ikke. Lad A_1 have Existenssandsynligheden s_1 og, saafremt den existerer, med Sandsynligheden v_1 paa given Foranledning, som f. Ex. ved et eller flere Træk (med Tilbagelægning) fra en Urne med Kugler eller deslige, fremkalde Begivenheden B. Lad paa lignende Maade A_2 have „Værebrøken“ s_2 og „Virkebrøken for B“ v_2, og saaledes videre. Kan B nu kun komme fra A, saa er

Sandsynligheden for i et Træk at faa B fra A_1

$$S_1 = s_1 v_1,$$

Sandsynligheden for i et Træk overhovedet at faa B

$$S_2 = \Sigma sv,$$

Sandsynligheden for i n Træk at faa netop m Gange B fra A_1

$$S_3 = c \cdot s_1 v_1^{m} (1 - v_1)^{n-m},$$

idet, som bekendt,

$$c = \frac{|\underline{n}}{|\underline{n-m}\,|\underline{m}},$$

og Sandsynligheden for overhovedet at faa B netop m Gange i n Træk

$$S_4 = c\Sigma(sv^m(1-v)^{n-m}).$$

Er nu Begivenheden indtraadt f. Ex. netop m Gange i n Træk, saa er Sandsynligheden for, at den kom fra A_1

$$S_5 = \frac{s_1 v_1{}^m(1-v_1)^{n-m}}{\Sigma(sv^m(1-v)^{n-m})}, \tag{1}$$

og denne Sandsynlighed er da selvfølgelig tillige Sandsynligheden for, at A_1 samtidig existerede, altsaa A_1's nye Værebrøk, dens „aposterioriske Sandsynlighed“.

Undertiden kan Situationen være mere speciel. Tilstanden A kan f. Ex. omfatte en Mangfoldighed af i hinanden overglidende Nuancer, saa at den kan tænkes saaledes udstykket i en Uendelighed af Enkelttilfælde, at Virkebrøkerne v voxer fra Tilfælde til Tilfælde med den konstante Tilgift dv fra en vis Undergrænse a til en vis Overgrænse b. Sættes de tilsvarende smaa Værebrøker $s = udv$, bliver da f. Ex. Sandsynligheden for, at Begivenheden B vil komme fra et eller andet af Tilfældene med Virkebrøk mellem α og β, idet disse Størrelser tænkes liggende mellem *a og b*,

$$S_6 = \int_\alpha^\beta uv\,dv, \tag{2}$$

og er B kommet, er Sandsynligheden for, at det var fra et eller andet af disse Tilfælde og at et eller andet af disse altsaa existerede,

$$S_7 = \frac{\int_\alpha^\beta uv\,dv}{\int_a^b uv\,dv}. \tag{3}$$

En endnu speciellere Situation har man, hvis det tillige er givet, at samtlige $\frac{b-a}{dv}$ Enkelttilfælde skal dele Existenssandsynligheden 1 ligelig, saa at hvert faar $s = \frac{dv}{b-a}$. De to tilsvarende Formler bliver da

$$S_8 = \frac{1}{b-a}\int_{a}^{\beta} v dv, \quad S_9 = \frac{\int_{a}^{\beta} v dv}{\int_{a}^{b} v dv}. \tag{4, 5}$$

Og endelig kan man endnu speciellere have $a = 0$, $b = 1$ og altsaa

$$S_{10} = \int_{a}^{\beta} v dv, \quad S_{11} = \frac{\int_{a}^{\beta} v dv}{\int_{0}^{1} v dv}. \tag{6, 7}$$

I Ligning 1 har man imidlertid det almene Udtryk for Reglen, og denne kan derfor formuleres saaledes:

Kan et vist Resultat kun hidrøre fra den ene eller den anden af en vis Række hinanden udelukkende Tilstande, og er det indtruffet, saa er Sandsynligheden for, at det kom fra en bestemt af Tilstandene, samtidig denne Tilstands nye Existenssandsynlighed eller dens saakaldte aposterioriske Sandsynlighed.

Hvad Bayes efter min Formodning har haft en Fornemmelse af, men ikke klart set, er den Omstændighed, at samtlige Enkelttilstande ikke altid har en og samme Sandsynlighed og at Værebrøkerne s derfor maa tages med i Betragtning. At han ikke klart har set dette, skriver sig sagtens fra, at i det bestemte Exempel med Kugler, ved hvis Hjælp han udleder Reglen, har samtlige Enkelttilstande netop en og samme Existenssandsynlighed. Det er derfor nærmest Formel 7, han kommer til.

Men fjerner man nu denne Ufuldkommenhed, og passer man paa, at enhver af de speciellere Formler stiller sine bestemte Krav til den Situation, i hvilken den skal have Gyldighed, føres vi jo slet ikke ud over den rene Matematik eller Logik. Alt er fuldstændig exakt og vundet saa direkte fra velbekendte Sætninger, at al særlig Bevisførelse bliver overflødig. Den sædvanlige Definition af Begrebet Sandsynlighed: „Forholdet mellem gunstige og sideordnet mulige Tilfælde“ er ingenlunde overskredet i Udtrykket for den nye Sandsynlighed; man har de første i Tællerne og de sidste i Nævnerne. Hvad der endnu kunde trænge til at forandres, er blot selve de gængse Udtryk. Det er ikke heldigt paa den sædvanlige Maade at tale om apriorisk og aposteriorisk eller empirisk Sandsynlighed; thi al reel Sandsynlighed er i Virkeligheden empirisk. At der er Sandsynligheden $\frac{1}{6}$ for at slaa f. Ex. 3 med en Tærning, ved jeg først, efter at jeg har erfaret, hvad en Tærning er. Det eneste, der med Rette kan kaldes apriorisk Sandsynlighed, er vistnok den fritvalgte Sandsynlighed i Regneopgaver o. dl. Der er for saa vidt mere Grund til at tvivle om apriorisk end om aposteriorisk Sandsynlighed, og naturligst og rigtigst vilde det være ved Anvendelsen af Bayes' Regel bestandig blot at tale om den oprindelige og den nye eller den vundne Sandsynlighed.

III.

Det vil nu i det følgende blive Opgaven at undersøge, om nogen af de mange mod Bayes' Regel rettede Indvendinger rammer den i dens her givne Formulering, og særlig bliver det det afgørende Indlæg fra Direktør Bing, der maa prøves. Hans Afhandlings mange mærkelige Resultater maa føres tilbage til deres Kilder, og det maa paavises, at disse enten bestaa i Misbrug af Reglen, ɔ: Benyttelse af den falske Regel, Formel 7 taget som almengyldig, eller i Regnefejl, der ikke har nogen som helst nødvendig Forbindelse med den rette Regel.

Først vil det maaske være heldigt at imødegaa en almindelig nærliggende Indvending.

Man kunde mene, at Bayes' Regel ganske vist nu er gjort exakt, men at den samtidig er gjort omtrent ubrugelig. Thi alle disse Størrelser s og u kender man jo kun rent undtagelsesvis. Er ikke al Sandsynlighedslære netop en Lære om, hvorledes man ud fra utilstrækkelige Oplysninger vinder de bedst mulige Resultater? Og nu er her jo allersnarest krævet fuldt tilstrækkelige Oplysninger.

Svaret er dette: Her er intet som helst krævet. Her er blot vist, hvad der er nødvendig Forudsætning for et exakt Resultat. Og et saadant Ideal vil det altid være rigtigst at opstille alle Vegne, hvor det blot er muligt, for at man kan have noget at maale paa, noget, der kan antyde, hvor nær ved selve Sandheden eller hvor fjernt fra den man befinder sig. Hvor jeg ikke kan regne exakt, kan jeg godt endnu have fuld Frihed til at regne med Tilnærmelse eller under frit valgte Forudsætninger. Kun vil jeg gennem de exakte Formler da bestandig vide, at jeg er saa og saa langt borte fra det exakte. Jeg kan da udtale dette med i mit Facit eller maaske endog helt opgive at regne, fordi de givne Udgangspunkter er alt for utilstrækkelige. End ikke Sandsynlighedslæren er jo forpligtet til altid at skaffe Resultat. Betydningen af disse Almenbemærkninger vil bedst fremgaa af det følgende. Et ganske elementært Exempel vil passende kunne tjene som Indledning.

En Urne indeholder 2 Kugler, hvide, sorte eller begge Slags. Hvilken Sandsynlighed er der for at faa hvidt i første Træk? Og er dette sket og Kuglen lagt tilbage, hvilken Sandsynlighed er der da for at faa hvidt i næste Træk?

Denne eller en lignende Opgave findes i omtrent enhver Sandsynlighedslære og løses gennemgaaende saaledes:

Tre Tilstande er mulige i Urnen. Der kan være 0, 1 eller 2 hvide Kugler i den. Hver Tilstand har Værebrøken $\frac{1}{3}$, og Virkebrøkerne for hvidt er henholdsvis 0, $\frac{1}{2}$, 1. Den oprindelige Sandsynlighed for at faa hvidt er derfor

$$S_1 = \tfrac{1}{3} \cdot 0 + \tfrac{1}{3} \cdot \tfrac{1}{2} + \tfrac{1}{3} \cdot 1 = \tfrac{1}{2}.$$

Har man faaet hvidt, er Sandsynligheden for, at det kom fra A_1, A_2, A_3 henholdsvis

$$\frac{0}{\frac{1}{2}}, \quad \frac{\frac{1}{6}}{\frac{1}{2}}, \quad \frac{\frac{1}{3}}{\frac{1}{2}},$$

og dette er tillige de tre Tilstandes nye Værebrøker. Sandsynligheden for hvidt i andet Træk er derfor

$$S_2 = \frac{0}{\frac{1}{2}} \cdot 0 + \frac{\frac{1}{6}}{\frac{1}{2}} \cdot \frac{1}{2} + \frac{\frac{1}{3}}{\frac{1}{2}} \cdot 1 = \frac{5}{6}.$$

Denne Besvarelse af Opgaven er imidlertid urigtig; thi den er vilkaarlig. Det er vilkaarligt uden videre at give de tre Tilstande en og samme Værebrøk, $\frac{1}{3}$. Ligesaa rimeligt er det nemlig at antage, at enhver af de to Kugler med samme Sandsynlighed kan være sort som hvid, saa at vi faar de fire sideordnede Muligheder *ss*, *sh*, *hs*, *hh*, hver med Værebrøken $\frac{1}{4}$ og med Virkebrøkerne for hvidt henholdsvis 0, $\frac{1}{2}$, $\frac{1}{2}$, 1. Naturligvis kan de to mellemste af disse Tilstande godt slaas sammen til den ene: én Kugle af hver Farve. Men vi faar da tre Tilstande med Værebrøkerne $\frac{1}{4}$, $\frac{1}{2}$, $\frac{1}{4}$ og Virkebrøkerne for hvidt 0, $\frac{1}{2}$, 1.

Sandsynligheden for hvidt i første Træk bliver dermed som før

$$S_3 = \tfrac{1}{4} \cdot 0 + \tfrac{1}{2} \cdot \tfrac{1}{2} + \tfrac{1}{4} \cdot 1 = \tfrac{1}{2};$$

men de nye Værebrøker bliver nu

$$\frac{0}{\frac{1}{2}}, \quad \frac{\frac{1}{4}}{\frac{1}{2}}, \quad \frac{\frac{1}{4}}{\frac{1}{2}}$$

og Sandsynligheden for hvidt anden Gang altsaa kun

$$S_4 = \frac{0}{\frac{1}{2}} \cdot 0 + \frac{\frac{1}{4}}{\frac{1}{2}} \cdot \frac{1}{2} + \frac{\frac{1}{4}}{\frac{1}{2}} \cdot 1 = \frac{3}{4}.$$

Naturligvis vilde det ogsaa være galt at opstille denne sidste Løsning som enerigtig. Den rette Besvarelse af Opgaven er jo simpelthen denne, at den ikke er tilstrækkelig skarpt stillet, til at der kan svares entydigt paa dens sidste Spørgsmaal. Man maa svare betinget, at regner man Muligheden: en Kugle af hver Farve sideordnet med enhver af de to andre, bliver sidste Facit $\frac{5}{6}$; men regner man den sideordnet med de to andre tilsammen, bliver det $\frac{3}{4}$. Saaledes maa man jo imidlertid saa ofte svare betinget paa et matematisk Spørgsmaal, og at dadle Bayes' Regel af den Grund vilde være det samme som at blive fornærmet paa den pytagoreiske Læresætning, fordi den intet reelt kan sige os om Hypotenusen, naar vi intet reelt véd om Kateterne.

IV.

Den nys behandlede Opgave stillede os allerede overfor en tilsyneladende Modsigelse, idet vi fik to forskellige Svar paa dens sidste Spørgsmaal. Men det var en, om man saa maa sige, uskyldig og ingen „rigtig" Modsigelse. At man faar forskellige Resultater, naar man gaar ud fra forskellige Forudsætninger, er der jo god Mening i. Det er den Slags Modsigelser, selve Logikken gør til intet, idet den siger: Diversi respectus tollunt omnem contradictionem. Det følgende skal nu vise, at ganske af samme uskyldige Art er ogsaa alle de øvrige Modsigelser, man har ment at kunne paabyrde Bayes' Regel. Kun fordi Tanketraadene bliver lidt mere indviklede, faar disse Modsigelser, efterhaanden som vi skrider frem, et mere og mere afskrækkende Udseende.

En Grad alvorligere synes saaledes allerede det af Direktør Bing Side 9—13 behandlede Tilfælde at være. Med et Par ganske betydningsløse Ændringer kan det fremstilles saaledes:

Af en Urne med et overordenlig stort Antal, n, af i øvrigt ens Kugler har man i 10 Træk (med Tilbagelægning) faaet 7 med Paaskriften 1, 2 med Paaskriften 2 og 1 med Paaskriften 3. Hvilken Sandsynlighed er der for i næste Træk at faa en ny Slags?

Her maa altsaa skelnes mellem 4 Slags Kugler: Enere, Toere, Treere og „andre", og en almindelig Formel for samtlige i og for sig mulige Tilstande er den, at der er nx Enere, ny Toere, nz Treere og $nt = n\,(1 - x - y - z)$ andre Kugler. Lader man enhver af Brøkerne x, y og z efterhaanden repræsentere samtlige Værdier fra 0 til 1 med Springet $\frac{1}{n}$, dog saaledes at Summen af de tre samtidige Værdier ikke overstiger 1, faar vi dermed samtlige oprindelig mulige Enkelttilstande, der hver har sin oprindelige Værebrøk s og sine Virkebrøker x, y, z og $1 - x - y - z$ for henholdsvis Enere, Toere, Treere og andet. Den oprindelige Sandsynlighed for 7 Enere, 2 Toere og 1 Treer i 10 Træk er da

$$S_1 = c\Sigma\,(sx^7y^2z),$$

idet $c = \frac{10!}{7!\;2!\;1!} = 360$, og med større og større Tilnærmelse, jo større n er, kan dette Udtryk, idet vi sætter $s = u\,dx\,dy\,dz$, skrives

$$S_1 = c\int_0^1\!\int_0^p\!\int_0^q u\,x^7\,y^2\,z\,dx\,dy\,dz,$$

hvor $p = 1 - x$, $q = 1 - x - y$. Er nu hint Resultat kommet, bliver Sandsynligheden for en ny Slags Kugle i næste Træk

$$S_2 = \frac{\int_0^1\!\int_0^p\!\int_0^q u\,x^7\,y^2\,z\,(1 - x - y - z)\;dx\,dy\,dz}{\int_0^1\!\int_0^p\!\int_0^q u\,x^7\,y^2\,z\,dx\,dy\,dz}.$$

Men uheldigvis véd vi om Funktionen u aldeles intet. Maaske er der kun Kugler med de tre Paaskrifter 1, 2, 3; maaske er der i nogenlunde ligeligt Antal Paaskrifter op til 10 000. Kun under den ganske vilkaarlige Forudsætning, at u er ens for samtlige opstillede Tilstande, faar vi det bekendte Resultat $S_2 = \frac{1}{14}$, et Resultat, der altsaa kun tør opstilles, idet dets yderst betingede Gyldighed samtidig udtrykkelig fremhæves. Udgives Løsningen $\frac{1}{14}$ derimod for ubetinget, har man simpelthen gjort sig skyldig i en grov Fejlregning.

Og en saadan vil naturligvis let kunne skaffe os overraskende Konsekvenser paa Halsen. Forandrer man lidt paa

Opgaven, saa at den kommer til at lyde: Man har 10 Gange i Træk faaet en Kugle med Tal paa. Hvilken Sandsynlighed er der for i næste Træk at faa en uden Tal?, vil Svaret som bekendt ud fra Forudsætningen om lige Sandsynlighed for Kugler med og Kugler uden Tal paa blive $\frac{1}{12}$, og giver man nu de to Svar $\frac{1}{14}$ og $\frac{1}{12}$ uden med et Ord at nævne de ganske vilkaarlige og ganske forskellige, maaske yderst urigtige Forudsætninger, saa fremkommer der ved Forsømmelsen den Urimelighed, at der er større Sandsynlighed for i næste Træk at undgaa ethvert som helst Tal end for at undgaa Tallene 1, 2, 3, skønt Urnens Indhold er det samme, de gjorte Erfaringer ligesaa og Muligheden af, at der er Tal paa samtlige Kugler, slet ikke udelukket.

Men man ser uden Vanskelighed, at dette Exempel i alt væsenligt er identisk med det foregaaende. Hovedforskellen er den, at vi her har et Tilfælde, i hvilket de givne Oplysninger i Virkeligheden er saa bundutilstrækkelige, at det nærmest er meningsløst at indlade sig paa at regne.

Heller ikke Direktør Bing's næste Exempel fører os væsenlig videre. En vis Nuance, der giver Anledning til et Par Bemærkninger, rummer det dog. Jeg tager det derfor med.

En Købmand faar en Ladning Frugt hjem. Der er 100 000 Stykker. Er de alle sunde, kan deres Værdi sættes til 10 000 Kr.; men de mere eller mindre fordærvede maa alle regnes værdiløse. Manden har ikke tidligere modtaget den Art Frugt og véd derfor ikke, hvor godt den har taalt Rejsen. Til Prøve har han imidlertid rundt i Ladningen udtaget 30 Stykker. Disse viste sig alle friske, og han spørger nu en Matematiker, hvilken Værdi han derefter skal beregne Ladningen til.

Direktør Bing lader saa Matematikeren regne med den falske Regel, d. v. s. misbruge den speciellere Formel, hvorved Resultatet bliver 9 687,5 Kr., og da Købmanden imidlertid er kommet i Tanker om, at der egenlig er 30 forskellige Sorter Frugt og netop en Prøve svarende til hver Sort, gøres Regningen om paa lignende Maade, hvorved Værdien af Ladningen stiger til c. 9 836 Kr.

Vi har altsaa her nærmest blot en Gentagelse af, hvad der i forrige Exempel førte til Resultaterne $\frac{1}{12}$ og $\frac{1}{14}$.

Benyttes derimod den rette Regel, maa Ladningens Sundhedskvotient, S, i Henhold til Købmandens første Oplysninger sættes til

$$S = \frac{\int_0^1 uv^{31}\,dv}{\int_0^1 uv^{30}\,dv},$$

idet v er det almindelige Udtryk for en oprindelig mulig Sundhedskvotient mellem 0 og 1 og udv den tilsvarende Værebrøk. Har Matematikeren nu ikke den allerringeste Viden om dette u, vil det imidlertid være letsindigt af ham at indlade sig paa at regne; thi uden videre at tildele samtlige mulige Værdier af v en og samme Existenssandsynlighed, maa jo føre til et ganske vilkaarligt og maaske højst urigtigt Resultat. Men den Bemærkning fortjener aabenbart her en Plads, at vi jo i Virkeligheden temmelig sjelden er saa „absolut uvidende" om de paagældende Forhold, som Regningen med den falske Regel egenlig forudsætter. Har Købmanden ikke tidligere faaet netop den vedkommende Slags Frugt hjem, saa har han dog maaske mange Gange modtaget lignende Sorter, ja han har maaske endog hørt en hel Del om andres Ladninger af netop den omtalte Sort. Er det ham af Vigtighed at faa et rimeligt Svar paa sit Spørgsmaal, er det dog ganske naturligt, at Matematikeren søger at faa saa mange af disse reale Oplysninger at bygge paa som muligt, og kunde Købmanden nu f. Ex. meddele ham, at i en Mængde tidligere mere eller mindre nærbeslægtede Tilfælde havde Sundhedskvotienten bestandig ligget mellem Grænserne 0,9 og 1 med omtrent ligelig Fordeling, medens den aldrig havde været under 0,9, vilde der intet være i Vejen for, at han opstillede Regnestykket

$$S = \frac{\int_{0,9}^{1} v^{31}\,dv}{\int_{0,9}^{1} v^{30}\,dv} = 0,9726$$

og svarede Købmanden, at saafremt de meddelte Erfaringer ogsaa passede paa hans Tilfælde, vilde han passende kunne beregne sin Ladning til 9 726 Kr.

Kom Købmanden derpaa i Tanker om, at der egenlig var 30 Slags Frugt, maatte Matematikeren sikkert helst svare ham, at ud fra de givne ubestemte Forudsætninger vilde dette ingen Indflydelse kunne faa paa Beregningen. Der er jo ogsaa en Slags Fejlregning, der bestaar i, at man afleverer Resultater med 10 Decimaler, medens Udgangspunkterne kun er givne i groft afrundede hele.

V.

Mellem hele den første Afdeling af Direktør Bing's Afhandling og det her fremsatte er der i alt væsenligt Overensstemmelse. Forfatteren opstiller selv Side 4 den rigtige Regel og føjer endog til, at saaledes vil han i det følgende benytte Udtrykket „Bayes' Regel". Side 5 kalder han denne Regel exakt og fremhæver udtrykkelig, at den særlig ofte misbruges saaledes, at man uden videre regner alle Existenssandsynlighederne *s* lige store uden med et Ord at omtale denne Vilkaarlighed ved Fremsættelsen af de paagældende Resultater. At dette fører til Modsigelser, skal netop hans Exempler vise.

Alt dette er fortræffeligt. Men Forfatteren synes ikke tilstrækkelig skarpt og klart at fastholde dette Standpunkt, ifølge hvilket det altsaa blot er den meningsløse M i s b r u g af Bayes' Regel, der er uberettiget. Forestillingen om de nysnævnte saa betydningsfulde Værebrøker, *s*, træder efterhaanden tilbage i hans Bevidsthed. Allerede den Kritik, hvormed han ledsager Dobbeltresultaterne af de første Regneexempler, viser dette; thi han finder bestandig Fejlen særlig deri, at man har udstykket Hovedtilstanden A paa denne eller hin Maade. Saaledes har man f. Ex. i Opgaven med Kuglerne snart opstillet 4 Slags Kugler, nemlig Enere, Toere, Treere og andre, og snart to Slags, nemlig Kugler med og Kugler uden Tal paa. Men nøjagtig talt er denne Kritik egenlig uberettiget. En saadan Udstykning er simpelthen krævet af Opgaven og indeholder endnu intet som helst urigtigt. Fejlen indtræder først, idet man derpaa uvilkaarlig tillægger enhver af disse Enkelttilstande én og samme Existenssandsynlighed s, skønt den ene maaske er 1000 Gange saa sandsynlig som den anden. Som vi skal se, fører denne Overseen af Størrelserne s senere Forfatteren selv paa Vildspor.

Men ogsaa i en anden Henseende skifter Forfatteren i Løbet af Afhandlingen tilsyneladende Standpunkt, i alt Fald paa en temmelig forvirrende Maade i sin Udtryksform. Medens det Side 4 og 5 hed, at han ved „Bayes' Regel" vil forstaa den rette Regel, den Regel, der er „exakt" og altsaa uden Lyde, bruger han senere Gang paa Gang Udtrykket, hvor det i Virkeligheden er en falsk Regel, der er regnet med, og da denne Vending navnlig er fremtrædende overfor de to sidste, fra Dødelighedsstatistikken hentede Opgaver, hvor de opdukkende Modsigelser faar det allermest afskrækkende Udseende, bliver Læseren i Virkeligheden ret uheldig stillet. Thi det er nu vanskeligt for ham at afgøre, om Forfatteren trods de indtrædende stærkere og stærkere Modsigelser vedblivende holder paa, at al Ulempe blot skriver sig fra Misbrug af Reglen, eller han nu af de forøgede Vanskeligheder har set sig nødt til ogsaa at forkaste selve Reglen. At hele Forfatterens Afhandling blot skulde være en Advarsel mod Bortkastelsen af Størrelserne s og intet videre, det Indtryk vil næppe mange Læsere faa. Ogsaa Professor Thiele har jo omtalt den som langt videre gaaende. Men paa den anden

Side findes der adskillige Ytringer i dens sidste Afdeling, der kunde tyde paa, at det egenlig ikke er andet, Forfatteren har villet.

Dog, hvorledes det end forholder sig hermed, har vi her blot at forfølge vort Maal videre: at søge Kilderne til de af Direktør Bing fremstillede Modsigelser og at sammenholde dem med den her fremsatte Formulering af Reglen.

VI.

Vi tænker os en Kreds af n 40aarige Mennesker. Lever der gennemsnitlig efter henholdsvis et halvt og et helt Aars Forløb ny_1 og $ny_1y_2 = ny$ Individer af en saadan Kreds, kalder vi y_1 og y_2 første og andet Halvaars Levekvotient og y Helaarets. Lad nu y_1, y_2 og y staa som almene Betegnelser for de forskellige oprindelig mulige Værdier mellem 0 og 1 af saadanne Kvotienter, og lad $s_1 = u_1dy_1$, $s_2 = u_2dy_2$ og $s = udy$ være deres oprindelige Existenssandsynligheder. Lad der endvidere i første Halvaar være sket m_1, i andet Halvaar m_2, i Helaaret $m = m_1 + m_2$ Dødsfald i Kredsen. Man kan da spørge: Hvilken Sandsynlighed er der dermed for, at Helaarets Levekvotient for en saadan Kreds ligger mellem Grænserne α og β?

Søger man Svaret gennem Helaarserfaringerne, faar man ved Bayes' Regel

$$S_1 = \frac{\int_\alpha^\beta y^{n-m}(1-y)^m udy}{\int_0^1 y^{n-m}(1-y)^m udy}, \tag{1}$$

mens man, om man benyttede samtlige Halvaarserfaringer, vilde faa

$$S_2 = \frac{\iint y_1^{n-m_1}(1-y_1)^{m_1} y_2^{n-m}(1-y_2)^{m_2} u_1 u_2 dy_1 dy_2}{\int_0^1 y_1^{n-m_1}(1-y_1)^{m_1} u_1 dy_1 \int_0^1 y_2^{n-m}(1-y_2)^{m_2} u_2 dy_2}, \tag{2}$$

hvor Grænserne i Tælleren er bestemte ved, at $y_1 y_2$ skal falde mellem α og β.

Det er Direktør Bing, der stiller Opgaven, og han søger saa Svaret ad begge Veje. Idet han ligesom tidligere udelader Størrelserne u, kommer han til de to Ligninger:

$$S_1 = \frac{\int_\alpha^\beta y^{n-m}(1-y)^m\,dy}{\int_0^1 y^{n-m}(1-y)^m\,dy} \tag{3}$$

og

$$S_2 = \frac{\iint y_1^{\,n-m_1}(1-y_1)^{m_1} y_2^{\,n-m}(1-y_2)^{m_2} dy_1\,dy_2}{\int_0^1 y_1^{\,n-m_1}(1-y_1)^{m_1} dy_1 \int_0^1 y_2^{\,n-m}(1-y_2)^{m_2}\,dy_2}, \tag{4}$$

hvor Grænserne i sidste Tæller er som før. Sættes her $y_2 = \frac{y}{y_1}$ og derefter $1 - y_1 = z(1-y)$, gaar Udtrykket over til

$$S_2 = \frac{\int_\alpha^\beta y^{n-m}(1-y)^{m+1} dy \int_0^1 \frac{z^{m_1}(1-z)^{m_2} dz}{1-z(1-y)}}{\int_0^1 y_1^{\,n-m_1}(1-y_1)^{m_1} dy_1 \int_0^1 y_2^{\,n-m}(1-y_2)^{m_2}\,dy_2}, \tag{5}$$

og Direktør Bing gør nu den Bemærkning, at (5) altsaa vil give os et andet Resultat end (3). Derpaa stiller han det Spørgsmaal: Hvilken Sandsynlighed er der for, at en ny 40aarig vil leve et Aar ud, „idet vi for Simpelheds Skyld antage $m_1 = m_2 = 0$?“ Og herpaa faar han gennem (3) Svaret

$$S_3 = \frac{\int_0^1 y^{n+1} dy}{\int_0^1 y^n dy} = \frac{n+1}{n+2}, \tag{6}$$

medens (5) giver

$$S_4 = \frac{\int_0^1 y^{n+1}(1-y)\,dy \int_0^1 \frac{dz}{1-z(1-y)}}{\int_0^1 y_1^{\,n}\,dy_1 \int_0^1 y_2^{\,n} dy_2} = \left(\frac{n+1}{n+2}\right)^2. \tag{7}$$

Dette ser ved første Øjekast unægtelig noget overraskende ud, og Forfatteren fremhæver end yderligere, at delte man Aaret i uendelig mange Smaadele, vilde man ad samme Vej kunne godtgøre, at ikke én ny 40aarig vilde leve Aaret ud, og det, idet man støttede sig til den Erfaring, at af de n oprindelige var ikke én død i hele det betragtede Tidsrum.

„Grunden til de urimelige Resultater“, siger Direktør Bing (S. 17), „ligger i, at man ved at dele Aaret i Underafdelinger faar Tilfældet „død i Løbet af et helt Aar“ opløst i „død i første Periode“, „død i anden“, osv., medens noget lignende ikke sker for deres Vedkommende, der „leve et helt Aar““. Denne Forklaring er imidlertid, som nys fremhævet, ikke strengt rigtig. Har alle andre Videnskaber Lov til at dele Aaret i Underafdelinger, maa Statistik og Matematik ogsaa kunne gøre det, uden at der allerede dermed sker Ulykker. Nej, her er begaaet forskellige Fejl, og de urimelige Resultater har forskellige Grunde. Matematikken er, som Huxley siger, en fortræffelig Mølle, der leverer det allerfineste Mel; men én Ting formaar den ikke: Lægger man Rug ind, kan den ikke levere Hvedemel ud. Og netop det har man her forlangt af den.

Det er blandt andet urigtigt at vente ét og samme Svar ad de to benyttede Veje. Da y kan være bygget af mange andre Par Faktorer end netop y_1 og y_2, vilde Tilfældet y, om man regnede med bestemte Levekvotienter, være langt mere omfattende end Tilfældet $y_1 \cdot y_2$. Da man har regnet med samtlige mulige Kvotienter, har man imidlertid undgaaet de heraf følgende Konsekvenser. Men ganske paa samme Maade er Tilfældet m mere omfattende end Tilfældet $m_1 + m_2$, og denne Omstændighed bliver ikke betydningsløs før i sidste Øjeblik, da baade m_1 og m_2 sættes lig 0. Der er derfor intet mærkeligt i, at Ligningerne 3 og 5 giver forskellige Resultater; thi de svarer i Virkeligheden hver paa sit Spørgsmaal.

Men urigtige er de dernæst begge to, naturligvis fordi

Størrelserne u er udeladte. Thi dels indfører man dermed den ganske urimelige Forudsætning, at alle Værdier af Levekvotienten for et vist Tidsrum oprindelig er lige sandsynlige, skønt man meget godt véd, at det under nogenlunde almindelige Forhold dog er langt sandsynligere, at f. Ex. de 99 Procent af et større Antal 40aarige vil leve Halvaaret ud, end at slet ingen af dem eller hver og én skulde gøre det.

Men hertil kommer endnu det særlig for denne Opgave skæbnesvangre, at man ved at udelade Størrelserne u samtidig udsletter al Forskel mellem Halvaars- og Helaarskvotienter. Havde man i Overensstemmelse med de bedst mulige forhaandenværende Erfaringer sat Halvaarskvotienten til — lad os for Simpelheds Skyld sige — 0,99 med noget Spillerum til begge Sider. under aftagende eller endog blot konstant Sandsynlighed og Aarskvotienten paa lignende Maade til omkring $0{,}99^2$, saa havde man dog dermed faaet indført i Regningen, at det i det ene Tilfælde var, om ikke netop Halvaar, saa dog mindre Tidsrum, og i det andet omtrent dobbelt saa store Perioder, det drejede sig om, medens Kvotienterne nu er fuldstændig stumme angaaende denne Forskel.

En vis Antydning af en saadan Forskel har man imidlertid tilbage, saa længe Størrelserne m_1 og m_2 endnu spiller med. Men idet ogsaa denne sidste Rest af Oplysning fjernes, har man i Virkeligheden fuldstændig sat Tiden paa Døren og spurgt om Hvedemel uden at give Møllen det ringeste at male det af. I Ligning 6

$$S_3 = \frac{\int_0^1 y^{n+1}\,dy}{\int_0^1 y^n\,dy} = \frac{n+1}{n+2},$$

har man nu Formlen for et Helaar. Formlen for et Halvaar faas heraf, idet man forsyner y med Mærket 1 eller 2. Men dette Mærke kommer slet ikke til at spille med i Regningen, og Formlen er derfor i Virkeligheden lige saa vel Formel for

1 Dag som for 1000 Aar. Hvad der til sidst er spurgt om ved Opstillingen af Ligningerne, er derfor i Virkeligheden ikke andet end dette: Hvor stor Sandsynlighed er der for at overleve én ganske ubestemt Periode, og hvor stor Sandsynlighed er der for at overleve to saadanne, for øvrigt højst mærkværdige Perioder, hvori Dødskvotienten allernærmest er 50 Procent og hvori dog faktisk ikke en Sjæl dør? Herpaa er der svaret, at sidste Sandsynlighed er anden Potens af første, og et mere rimeligt Svar kan dog ikke ret vel forlanges. At selve Størrelsen $\frac{n+1}{n+2}$ er urigtig, er jo den simple Konsekvens af Værebrøkernes Urigtighed.

VII.

Direktør Bing fortsætter: „Naar man nu anerkender, at Bayes' Regel i alle saadanne Tilfælde, hvor man intet véd a priori om de søgte Aarsager, er ganske upaalidelig, saa føres man til at forsøge at opstille en anden, hvorved der kan undgaas Strid imellem de forskellige Løsninger."

Men dette er jo ganske uberettiget Tale. Der skulde f. Ex. have staaet: Naar det nu har vist sig, at man kan faa højst mærkværdige Resultater ved i Stedet for Bayes' Regel at benytte en ganske urigtig, der uden videre tillader os at regne med konstante Værebrøker, hver Gang man intet véd om deres Størrelse, ja endog hvor man godt véd, at de slet ikke er eller kan være konstante, saa føres man til at anbefale Folk at holde sig til den rette Regel og at opgive at regne, hvor det i Virkeligheden er ganske meningsløst at regne.

Direktør Bing har aabenbart her rent glemt, at han Side 4 og 5 fortalte os, hvad han vilde forstaa ved Bayes' Regel, og at han fandt denne Regel exakt, det vil dog vel sige: fuldstændig paalidelig og fejlfri. Han vil nu finde en ny og bedre. Men det er aabenbart lige saa svært som at finde en ny og bedre lille Tabel. Gaar man rigtigt frem, vil man upaatvivlelig komme til den gamle, og dette maa ogsaa gælde overfor Bayes' Regel. Regner Direktør Bing rigtigt, vil han afgjort komme tilbage til den af ham selv paa Side 4 anførte Regel, og kommer han ikke til den, maa han have regnet galt. Sandheden er ikke dobbelt. Da det øjensynlig er Afhandlingens Hovedslag, Forfatteren mener at levere paa de følgende Sider, vil det imidlertid have sin store Interesse at følge ham under hans Tankegang.

Han begynder med at omforme den almene Opgave til en mere speciel. Vi har atter for os en Kreds af n Mennesker. I Løbet af første Aar er der sket m_1 Dødsfald iblandt dem, i Løbet af det næste m_2. Det er antaget, at Dødskvotienten for første Aar er x_1 og at x paa lignende Maade er Forholdet mellem det gennemsnitlige Antal Dødsfald i andet Aar og det oprindelige Antal Individer i Kredsen. Hvilken Sandsynlighed er der nu for, at baade x_1 og x er rigtige?

Ved Behandlingen af denne Opgave vil Forfatteren altsaa søge at finde en Regel, som kan afløse Bayes', eller, hvad han dog vist nøjagtigere talt mener: en Regel, som kan supplere Bayes', og særlig en, som kan give rigtige Resultater paa Omraader, hvor Bayes' (vel at mærke ved først at lemlæstes) kun giver modsigelsesfulde.

Nu lykkes det virkelig Forfatteren u d f r a v i s s e v a l g t e U d g a n g s p u n k t e r paa en for øvrigt baade elegant og skarpsindig Maade at faa ført sin Regning til Ende. Men underligt nok kommer han til et Resultat, som han selv finder baade urimeligt og meningsløst og ganske ubrugeligt. Denne Omstændighed fortolker han saa, som om han dermed har „bevist, at der aldeles ikke existerer nogen aposteriorisk Sandsynlighed, naar der er Tale om Problemer, hvor man forud er absolut uvidende om de virkende Aarsager."

Men netop dette kunde jo Bayes' Regel — den rigtige, fra Side 4 — i Forvejen have sagt ham. Den forlanger jo netop, at man skal kende samtlige Enkelttilstandes baade Værebrøker og Virkebrøker, om man skal kunne naa til et berettiget Resultat. Selve det matematiske Udtryk for Reglen beviser jo Forfatterens Sætning saa kort og klart og afgørende, at man ikke kan ønske sig eller faa mere.

Det er derfor saa langt fra, at Forfatteren ved sit „ubrugelige Resultat“ i mindste Maade rammer Bayes' Regel, at man snarere maa sige, at han paa for øvrigt ganske overflødig Maade vilde have bekræftet dens Rigtighed, om han selv havde regnet rigtigt. Men dette har han uheldigvis ikke gjort, og ogsaa dette vil, som vi skal se, Bayes' Regel kunne lære ham.

Direktør Bing's Fremgangsmaade er nemlig denne:

Den søgte Sandsynlighed for, at x_1 og x begge er rigtige, siger han S. 18, maa have Formen $\varphi\,(x_1, x, m_1, m_2, n)\,dx_1\,dx$, „hvor φ er en ubekendt Funktion, om hvilken vi ikke vide andet end, at den maa være saaledes symmetrisk med Hensyn til de indgaaende Bogstaver, at den bliver uforandret ved, at man paa en Gang ombytter x_1 med x og m_1 med m_2, hvilket er umiddelbart indlysende.“

Dette er Regningens Grundlag. Det indeholder tre Paastande, som ved nærmere Undersøgelse alle tre vil vise sig urigtige.

Den søgte Sandsynlighed kan ikke have den angivne Form. Bayes' Regel (den „exakte“) saavel som almindelig sund Fornuft lærer os jo nemlig, at den nye Sandsynlighed bestandig ogsaa maa være Funktion af de oprindelige Existenssandsynligheder for de vedkommende Enkelttilstande. Vi maa derfor nødvendigvis have de oprindelige Sandsynligheder for Rigtigheden af x_1 og x med i Parentesen. Skal Differentialerne bag Parentesen betegne disse Sandsynligheder, saa er dermed jo allerede indført den gængse skæbnesvangre Fejltagelse, at de oprindelige Værebrøker uden videre, hver Gang det kniber, kan vælges frit og f. Ex. regnes konstante. Men derved vilde Forfatteren, om han ingen ny Fejl begik, blot naa til en Gentagelse af den gængse falske Regel.

Vilde Direktør Bing hertil indvende, at dette egenlig heller ikke vilde stride stort mod hans Ønske, da hans egenlige Øjemed netop er at komme til et umuligt Resultat for dermed at kunne slaa de vedkommende Sorter aposteriorisk Sandsynlighed grundigt ned, og at han desuden ikke kan indføre de rette Existenssandsynligheder for Værdierne af x_1 og x, fordi det netop er hans Forudsætning, at disse er ubekendte, saa vilde jeg dertil svare: At bygge nye og indviklede Beviser for, at man kommer til vilde Resultater ved vilkaarlig at regne med konstante Værebrøker, er det næppe Umagen værd at spilde Tid paa; man beviser dette aller kortest og klarest blot ved at opstille Reglen i dens rette Form. Og hvad den anden Indvending angaar, saa udelukker Ubekendtskabet med de faktiske Værdier af Værebrøkerne jo paa ingen Maade den Udvej, man ved lignende Lejligheder ellers altid benytter i Matematikken, nemlig den at indføre disse ubekendte i Regnskabet netop som ubekendte, der i Regningens Løb om muligt skal bestemmes. Havde Forfatteren gjort dette og i øvrigt regnet rigtigt, var han maaske kommet til et Resultat, der endnu indeholdt ubekendte, som ikke lod sig bestemme uden yderligere Oplysninger, men et Resultat, som netop derved vilde give os et langt frugtbarere Indblik i Problemets Natur end et slet og ret meningsløst.

Ogsaa mod Grundlagets anden og tredie Paastand maa der imidlertid rejses Indvending. Paastanden at den hævdede Symmetri er umiddelbart indlysende, maa vel nærmest selv kaldes umiddelbart indlysende urigtig. Er her virkelig Symmetri, er den i alt Fald alt andet end umiddelbart indlysende. Mig forekommer det snarere umiddelbart indlysende, at der ikke kan være Tale om Symmetri. Direktør Bing's egne Bogstavbetegnelser er uheldige, idet han har valgt symmetriske Tegn for usymmetriske Størrelser. Baade i denne og i den forrige Opgave er det sikkert heldigst at følge nedenstaaende Skema:

$$n = \begin{cases} nx_1 \\ ny_1 = \begin{cases} ny_1x_2 = nx. \\ ny_1y_2 = ny. \end{cases} \end{cases}$$

Af de n oprindelige Individer er efter første Tidsrums Forløb nx_1 døde, ny_1 endnu levende. Af disse sidste er efter andet Tidsrum endvidere ny_1x_2 døde og ny_1y_2 endnu levende. x_1 og y_1 er altsaa henholdsvis Døds- og Levekvotient for første Tidsrum og x_2 og y_2 det samme for andet, medens x er en Slags Dødskvotient for andet Tidsrum taget i Forhold til det oprindelige Antal levende. Det er med x_1 og x, Direktør Bing arbejder; men selv kalder han disse to Størrelser x_1 og x_2 og bestyrkes derved uvilkaarlig i den Antagelse, at den sidste af disse Størrelser staar i et lignende Forhold til m_2 og og andet Tidsrum, som x_1 staar i til m_1 og første. For ikke at føres i denne Fristelse vil vi her vedblivende benytte Skemaets Betegnelser.

Forfatterens Afhandling fremkaldte strax Indvendinger fra Professor L. Lorenz, og i et Par Artikler fra hver Side førtes der i den nævnte Aargang af „Tidsskrift for Matematik“ en varm Strid om den aposterioriske Sandsynlighed. Det lykkedes imidlertid ingen af Parterne at føre saa klare og afgørende Grunde i Ilden, at Modparten følte sig overbevist, og Kampen endte for saa vidt uden Resultat. Prof. Lorenz havde imidlertid en levende Fornemmelse af, at Direktør Bing misbrugte Størrelsen x i sin Regning, og søgte paa forskellige Maader at illustrere dette. Direktør Bing svarede igen med andre Billeder, der imidlertid ikke naaede ind til Problemets Kærne. Kun et enkelt Sted kommer han til at berøre det afgørende i Sagen paa mere matematisk Maade, og da han her tilmed synes at have opdaget nye Uheldigheder ved, hvad han nu kalder Bayes' Regel, hører en Undersøgelse af hans Ytringer her ganske naturligt hjemme paa det Punkt, hvortil vi nu er naaede.

Direktør Bing siger S. 128—29:

„Lad os atter tage Exemplet fra Dødelighedsstatistikken og anvende Bayes' Regel. Der er ved et Forsøg med n levende fra Begyndelsen død m_1 i første Aar og m_2 i andet; x_1 og x er for den hele Befolknings Vedkommende henholdsvis Forholdet mellem døde i første og andet Aar og det oprindelige Antal; vi vide forud intet om x_1 og x, hver Værdi har derfor forud Sandsynlighederne dx_1 og dx; Sandsynligheden efter Forsøget for x_1's og x's samtidige Rigtighed er da

$$x_1^{m_1} x^{m_2} (1 - x_1 - x)^{n-m_1-m_2} dx_1\, dx. \qquad (\alpha)$$

Dette støtter sig paa det almindelige Ræsonnement: da man intet kan sige, hvorfor en Værdi af x_1 eller x skulde være mere antagelig end en anden, saa maa vi sætte de aprioriske Sandsynligheder lige store, lig dx_1 og dx. Udtrykket (α) er symmetrisk med Hensyn til x_1, x, m_1 og m_2.

Nu kunne vi imidlertid ogsaa tage Sagen paa en anden Maade. Lad y_1 være Forholdstallet mellem levende ved Slutningen og Begyndelsen af første Aar og y_2 mellem levende ved Slutningen og Begyndelsen af andet Aar. Vi kunne da intet sige apriori om Værdierne af y_1 og y_2, hvorefter én Værdi skulde udmærke sig fra en anden, og vi faa derfor Udtrykket

$$y_1^{n-m_1} (1 - y_1)^{m_1} y_2^{n-m_1-m_2} (1 - y_2)^{m_2} dy_1\, dy_2,$$

der kan erstattes af

$$\frac{1}{1 - x_1} x_1^{m_1} x^{m_2} (1 - x_1 - x)^{n-m_1-m_2} dx_1\, dx. \qquad (\beta)$$

(α) og (β) stride imod hinanden, og navnlig har Udtrykket (β) den Egenskab ikke at være symmetrisk med Hensyn til x_1, x, m_1 og m_2. Med andre Ord, ved Brugen af Bayes' Regel er der stedse et Valg at træffe mellem uafhængige variable, for hvilke de aprioriske Sandsynligheder sættes lig Differentialerne; Valget er ganske frit, naar det blot af Problemets Natur fremgaar, at der intet taler for at foretrække visse Værdier af de variable for andre. Men dette frie Valg medfører netop, at der kommer Strid imellem Resultaterne, naar man snart vælger et, snart et andet Sæt variable. Hvad der er skrevet Side 19—21, gaar netop ud paa at finde, hvilke Udtryk der maatte sættes i Stedet for dem, der dannes efter Bayes' Regel, for at faa Uoverensstemmelsen fjernet. — At alle Uoverensstemmelser og al Strid, der fremkommer ved Anvendelsen af Bayes' Regel, umiddelbart eller middelbart ere Følger af det frie Valg af variable, synes Prof. L. ikke at kunne faa Øje paa."

Som man ser, staar vi hermed atter midt i vort Dødelighedsproblem, og det er vel tilmed ikke overdrevent dristigt at formode, at det netop er det symmetriske Udtryk α, der har ført Forfatteren til hans faste Tro paa, at den før omtalte Funktion φ paa den hævdede Maade maa være symmetrisk.

Men de anførte Udtalelser har en endnu langt videre rækkende Betydning. Thi pludselig har Forfatteren jo her rejst en splinterny Anke mod Bayes' Regel, nemlig den, at det under dens Brug skulde have en ganske ejendommelig Indflydelse, hvilke variable man valgte som uafhængige, selv om man ikke ved Valget forsyndede sig mod Matematikken i Almindelighed; ja alle Ulykker flydende fra Reglens Anvendelse erklæres nu udspringende fra dette Valg.

Denne Indvending er langt alvorligere end alle de tidligere tilsammen. Thi alle de tidligere var nøjagtigere talt jo blot Indvendinger mod den ganske enfoldige Regel, der tillader os ubetinget, ja mod bedre Vidende, at regne alle Værebrøker konstante; og samtlige tidligere Modsigelser lod sig jo forholdsvis let forklare blot ved Paavisningen af, at der var regnet med saadanne konstante Værebrøker paa Omraader, hvor alt talte for, at de næppe eller slet ikke kunde være konstante.

Saa let lader Forfatterens nye Anke sig ikke afvise. Thi ganske vist mener han ogsaa nu med Bayes' Regel den enfoldige Regel. Men Anken forsvinder nu ikke uden videre sammen med Enfoldigheden. Thi sæt — for at blive i det konkrete —, at Landet med de $m_1 + m_2 = m$ Dødsfald blandt de n Forsøgspersoner just havde været hærget f. Ex. af en lunefuld Kolera, der snart havde slaaet ned for Fode og snart vist sig næsten uskadelig, kort sagt opført sig saaledes, at Landets statistiske Institut ved omhyggelige Undersøgelser netop var kommet til det Resultat, at alle Værdier af Kvotienter som x_1 og x foreløbig maatte sættes lige sandsynlige, saa kunde der jo endnu spørges om den ved de $m_1 + m_2$ Dødsfald medbestemte nye Sandsynlighed for Rigtigheden af x_1 og x, og selv med den rigtige Regel maatte Regningen da føres med konstante Værebrøker. Spillede nu Valget af uafhængig variable den af Direktør Bing paastaaede Rolle, ja saa var det jo ogsaa ude med den rigtige Regel, saa maatte man jo ogsaa ved berettiget Anvendelse af den snart kunne komme til et symmetrisk Udtryk som α og snart til et dermed stridende usymmetrisk som β. Det er lidt svært at forstaa, at Direktør Bing ikke har set denne vidtrækkende Konsekvens af Paastanden om de variable; men antager vi, at han har set den, bliver det dobbelt uforstaaeligt, at han ikke er bleven forskrækket og skyndsomst har sluttet: Jeg maa dog vist have regnet galt.

Ja, vi maa endnu et Skridt videre. Nærmere beset har de to Udtryk α og β jo egenlig endnu intet med Bayes' Regel at gøre. De er nemlig ikke, som Forfatteren siger, Udtryk for Sandsynligheden „efter Forsøget“, men simpelthen (rigtige eller urigtige) Udtryk for Sandsynligheden af m_1 Dødsfald i første og m_2 i andet Aar under Forudsætning af Kvotienterne x_1 og x med konstante Værebrøker. Og da nu saadanne eller lignende Kvotienter — rent bortset fra Bayes' Regel — af og til faktisk kan have konstante Værebrøker, saa rækker Forfatterens Anke endog ud over baade den rigtige og den gale Regel. Allerede ved at søge en „sammensat Sandsynlighed“ er han jo kommet til de to med hinanden stridende

Udtryk. Det bliver altsaa intet mindre end Størstedelen af hele Sandsynlighedslæren, der maa falde, hvis han har Ret. Og naturligvis er Bayes' Regel, baade den rigtige og den gale vedblivende med i Skibbrudet. Thi det er jo ved Hjælp af det ene eller det andet af de to Udtryk, Brøken for den nye Sandsynlighed skal dannes, og det paa en saadan Maade, at Modsigelsen ogsaa her vil komme til at gøre sig gældende.

Allerede af dette vil det da med ret overvældende Sandsynlighed fremgaa, at Direktør Bing maa have regnet galt. Men endnu bedre vil det naturligvis være direkte at paavise Fejlen. Lettest vilde dette kunne gøres, idet vi direkte gik til Løsningen af den nysnævnte Koleraopgave. Men det vil være naturligere først at behandle den mere almene angaaende den nye Sandsynlighed for Rigtigheden af x_1 og x. Fra de almene Udtryk kan vi da aflede de særlige, svarende til konstante Værebrøker for x_1 og x, og saaledes komme tilbage til Betragtningen af de to Udtryk α og β og derigennem baade til Spørgsmaalet om de variable og om Symmetrien eller Ikke-Symmetrien hos Funktionen φ. Som alt bemærket er det Bayes' Regel, der vil kunne give os Svarene paa alle de nævnte Spørgsmaal.

Vi vælger da vore Betegnelser i Overensstemmelse med det før omtalte Skema

$$n = \begin{cases} nx_1 \\ ny_1 = \begin{cases} ny_1x_2 = nx \\ ny_1y_2 = ny \end{cases} \end{cases},$$

der definerer de forskellige Størrelser, vi faar Brug for under Regningen. Ved Betragtning af den hele Befolkning har man altsaa vundet Kvotienterne x_1 og x og samtidig sikkert ogsaa de øvrige. Vi maa da sagtens forudsætte, at Landet har et statistisk Institut, som har bearbejdet Forholdene og ikke blot antaget de bestemte Kvotientværdier x_1, y_1 osv., hver med en vis sandsynlig Rigtighed, men desuden dannet sig mere eller mindre nøjagtige Forestillinger om, hvorledes Sandsynligheden vil vexle, idet man lader Kvotientværdierne vexle. Om alt dette véd imidlertid kun Instituttet selv Besked. For os, som har faaet Opgaven til Behandling, er kun selve Bogstavbetegnelserne i Skemaet givne.

Dette forhindrer os imidlertid ikke i at antage, at enhver af disse Kvotienter maa have sin ejendommelige med Kvotientværdien aabenbart vexlende Sandsynlighed, og vi kan ogsaa give disse Sandsynligheder foreløbig tomme Navne. Lad Sandsynligheden for, at x_1 er rigtig, være $s_1 = u_1 dx_1$, og lad $s_2 = u_2 dx_2$ og $s = udx$ være de tilsvarende Størrelser for x_2 og x. Vi søger saa Sandsynligheden for Indtrædelsen af de to Sæt Dødsfald for derefter at anvende Bayes' Regel.

Sandsynligheden for, at x_1 er rigtig, er

$$s_1.$$

Sandsynligheden for, at der i Henhold hertil vil indtræffe netop m_1 Dødsfald i første Aar er

$$s_1 c_1 x_1^{m_1} (1 - x_1)^{n-m_1} = S_1, \quad \text{hvor } c_1 = \frac{\underline{|n}}{\underline{|n - m_1} \; \underline{|m_1}}.$$

Sandsynligheden for dette og for, at x er rigtig og andet Aars Dødskvotient altsaa $= \frac{x}{1 - x_1}$, er

$$S_1 s.$$

Sandsynligheden for dette og for, at der dermed vil dø netop m_2 i andet Aar, er

$$S_1 s c_2 \left(\frac{x}{1 - x_1}\right)^{m_2} \left(1 - \frac{x}{1 - x_1}\right)^{n-m} = S, \quad \text{hvor } c_2 = \frac{\underline{|n - m_1}}{\underline{|n - m} \; \underline{|m_2}}.$$

Den søgte Sandsynlighed er altsaa

$$\left.\begin{aligned} S &= c_1 c_2 x_1^{m_1} (1 - x_1)^{n-m_1} \left(\frac{x}{1 - x_1}\right)^{m_2} \left(\frac{1 - x_1 - x}{1 - x_1}\right)^{n-m} s_1 s \\ &= c_1 c_2 x_1^{m_1} x^{m_2} (1 - x_1 - x)^{n-m} u_1 dx_1 u dx. \end{aligned}\right\} \quad (1)$$

Havde vi som andet Aars Dødskvotient direkte benyttet Skemaets x_2, havde vi faaet

$$S = c_1 c_2 x_1^{m_1} (1 - x_1)^{n-m_1} x_2^{m_2} (1 - x_2)^{n-m} u_1 dx_1 u_2 dx_2 \quad (2)$$

eller, idet

$$x_2 = \frac{x}{1 - x_1}, \qquad dx_2 = \frac{1}{1 - x_1} dx,$$

$$S = c_1 c_2 x_1^{m_1} x^{m_2} (1 - x_1 - x)^{n-m} u_1 dx_1 u_2 \frac{1}{1 - x_1} dx. \quad (3)$$

Da Skemaet giver os $x_1 = 1 - y_1$, $x_2 = 1 - y_2$, $x = y_1 - y$, vil det være let at danne de tilsvarende Ligninger med Kvotienterne y_1, y_2 og y.

Sammenholdes (1) og (3), ses det, at skal man ad begge Veje komme til samme Værdi for S, maa man have

$$s = udx = u_2 \frac{1}{1 - x_1} dx = u_2 dx_2 = s_2, \quad (4)$$

altsaa begge de benyttede Kvotienter af samme Sandsynlighed eller bestemte med samme Nøjagtighed, hvilket da ogsaa er temmelig selvfølgeligt.

Alt dette har endnu ikke haft med Bayes' Regel eller aposteriorisk Sandsynlighed at gøre. Vi anvender nu Reglen og faar derved den søgte Funktion φ. Af de 6 Udtryk, vore Ligninger kunde give os, vil det være tilstrækkeligt at opstille de to, der nærmest svarer til Direktør Bing's α og β:

$$\left.\begin{aligned} \varphi &= \frac{x_1^{m_1} x^{m_2} (1 - x_1 - x)^{n-m} u_1 dx_1 udx}{\int_0^1 \int_0^{1-x_1} x_1^{m_1} x^{m_2} (1 - x_1 - x)^{n-m} u_1 dx_1 udx} \\ &= \frac{x_1^{m_1} x^{m_2} (1 - x_1 - x)^{n-m} u_1 dx_1 u_2 \frac{1}{1 - x_1} dx}{\int_0^1 \int_0^{1-x_1} x_1^{m_1} x^{m_2} (1 - x_1 - x)^{n-m} u_1 dx_1 u_2 \frac{1}{1 - x_1} dx}. \end{aligned}\right\} \quad (5, 6)$$

Af disse Udtryk følger nu ret umiddelbart de tilsvarende specielle for Koleraopgaven. For at gøre alt saa anskueligt som muligt skal jeg imidlertid Skridt for Skridt udlede dem.

Vi antager os altsaa i Besiddelse af saadanne Oplysninger, at vi med Rette kan sætte samtlige Værebrøker konstante. Nu kan x_1 jo voxe fra 0 til 1. Lader vi den voxe med Skridtlængden dx_1, faar den altsaa $\frac{1}{dx}$ Værdier, og da én af disse maa være rigtig, skal de altsaa dele Sandsynligheden 1 ligelig imellem sig, hvorved hver faar Sandsynligheden dx_1 og u_1 bliver $= 1$. Ganske det tilsvarende gælder om x_2. Hver af dens Værdier faar derfor Sandsynligheden dx_2, og ogsaa u_2 bliver $= 1$. Men anderledes forholder det sig med x. Ogsaa en af dens Værdier maa være rigtig; ogsaa alle dens Værdier skal altsaa dele Sandsynligheden 1 ligelig. Men lader vi den i Overensstemmelse med de andre voxe med Skridtlængden dx, faar den ikke $\frac{1}{dx}$ Værdier, hver med Sandsynligheden dx. Da nemlig x er Forholdet mellem det gennemsnitlige Antal Dødsfald i andet Aar og et større Antal Mennesker, end der gennemsnitlig findes i Kredsen i andet Aar, og da der ikke godt gennemsnitlig kan dø flere, end der gennemsnitlig findes, kan x følgelig heller ikke blive større end $\frac{ny_1}{n} = y_1 = 1 - x_1$. Den faar derfor kun $\frac{1-x_1}{dx}$ Værdier, og hver af disse maa følgelig have Sandsynligheden $\frac{1}{1-x_1} dx = s = udx$. Vi faar altsaa her ganske som i Ligning 4 $\frac{u}{u_2} = \frac{1}{1-x_1}$, og de to særlige Udtryk bliver

$$\left.\begin{aligned}
\varphi_1 &= \frac{x_1^{m_1} x^{m_2} (1 - x_1 - x)^{n-m} dx_1 \frac{1}{1 - x_1} dx}{\int_0^1\int_0^{1-x_1} x_1^{m_1} x^{m_2} (1 - x_1 - x)^{n-m} dx_1 \frac{1}{1 - x_1} dx} \\
&= \frac{x_1^{m_1} x^{m_2} (1 - x_1 - x)^{n-m} dx_1 \frac{1}{1 - x_1} dx}{\int_0^1\int_0^{1-x_1} x_1^{m_1} x^{m_2} (1 - x_1 - x)^{n-m} dx_1 \frac{1}{1 - x_1} dx},
\end{aligned}\right\} \quad (7)$$

altsaa fuldstændig identiske, akkurat ligesom (5) og (6) er fuldstændig identiske. Og havde vi benyttet de variable y_1, y_2 osv., vilde det, som let ses, være gaaet os ganske paa samme Maade.

Men det vil nu være let at fjerne alle de opstillede Modsigelser og at svare paa alle de foreliggende Spørgsmaal.

Naar Direktør Bing Side 128—29 er kommet til det i forhøjet Grad mærkelige Resultat, at det spiller en skæbnesvanger Rolle, hvilke variable vi knytter vor Regning til, ikke blot under Brugen og Misbrugen af Bayes' Regel, men allerede ved Bestemmelsen af sammensat Sandsynlighed, saa skriver denne Fejltagelse sig fra den Omstændighed, at Forfatteren nu har regnet, ikke blot som tidligere med vilkaarlige konstante Værebrøker, men slet og ret med en absolut umulig Værebrøk. Det er den ganske usymmetriske Størrelse x, der har vildført ham. Han har dannet sig den Opfattelse, at den stod i et lignende Forhold til m_2 og andet Aar, som x_1 staar i til m_1 og første Aar. Han har overset, at den i Modsætning til x_1 og x_2, de to Tidsrums Dødskvotienter, er Udtryk for Forholdet mellem de gennemsnitlige Dødsfald i en vis Periode og et Antal Mennesker fra en anden og gennemsnitlig menneskerigere Periode, at den derfor slet ikke kan naa Størrelsen 1 og at dens Værebrøk derfor ikke med nogen som helst Mening kan sættes $= dx$, men, hvis den skal være konstant, nødvendigvis maa sættes $= \frac{1}{1-x_1} dx$. Retter han denne Fejl, som han har arbejdet med ved Udledelsen af Udtrykket α, bliver dette identisk med Udtrykket β, hvor Fejlen ikke har faaet Indpas, fordi der her er regnet med Størrelserne y_1 og y_2 og Faktoren $\frac{1}{1-x_1}$ derfor er kommet med. Men hermed er altsaa den vidtrækkende Paastand om de variable tilintetgjort. Den havde, som det vil ses, sit hele Udspring fra det dobbelt uheldige Dogme, at det er tilladt vilkaarligt at regne ikke blot med konstante, men med absolut umulige Værebrøker.

Og samtidig er det brydsomme Spørgsmaal angaaende Symmetrien i Funktionen φ løst. Saa snart man har faaet Øje paa, hvilken ejendommelig Rolle Størrelsen x spiller, bliver det jo virkelig, som tidligere fremhævet, umiddelbart indlysende, at den paastaaede Symmetri ikke kan finde Sted, og det samme fremgaar med fuldstændig Tydelighed saa vel af Ligningerne (5) og (6) som af (7). Hvad enten man regner med konstante Værebrøker eller ej, faar man Usymmetriskhed. Dette viser sig umiddelbart i alle de tre Nævnere som

Forskellen i Integrationsgrænserne for x_1 og x, og at allerede Tællerne er usymmetriske, ses for de to sidstes Vedkommende ligeledes rent umiddelbart, ligesom det for den førstes Vedkommende fremgaar af Ligning (4), der jo gør Tælleren i (5) identisk med den i (6).

Ligningerne (5—7) viser endvidere, hvad der jo egenlig ogsaa er yderst selvfølgeligt, at φ faktisk er medbestemt ved de oprindelige Værebrøker, hvad enten disse er variable eller konstante. Det er derfor en ganske naiv Udvej uden videre at give disse Brøker vilkaarlige Størrelser alene for at faa muliggjort en maaske ellers umulig Regning, og naar denne naive Udvej faktisk ofte er bleven benyttet, maa dette Fænomen aabenbart have sine ganske ejendommelige Aarsager, hvorom et Par Bemærkninger siden. Idet vi her har fulgt den sædvanlige matematiske Fremgangsmaade og indført de vedkommende Værebrøker som ubekendte og nu aabenbart i den almenere Opgave mangler Oplysninger, som kunde muliggøre deres Bestemmelse, lærer vi, at denne Opgave i Virkeligheden er uløselig ud fra de givne Udgangspunkter alene, hvorimod Ligning (7) viser os, at den speciellere meget godt kan løses, uden at vi støder paa nogen som helst Modsigelse eller anden Ulempe.

VIII.

I det foregaaende er vistnok alle de mod Bayes' Regel rejste Indvendinger gennemgaaede, og det er paavist, at ingen af dem rammer den virkelige Regel. De angaar enten udelukkende det falske Dogme om de vilkaarlige Værebrøker, eller de beror paa tilfældige Regnefejl. Der lader sig da for saa vidt opstille en sand og berettiget Regel og en til denne svarende berettiget aposteriorisk Sandsynlighed. Den, der endnu med Rette skulde kunne benægte dette, maatte kunne paavise Urigtigheder i de i andet Afsnit opstillede Formler. Den rette Regel deler den Egenskab med enhver anden matematisk Sætning, at den ikke kan skaffe os reale Resultater ud fra hvilke som helst Udgangspunkter. De givne Oplysninger kan, som vi har set, være saa utilstrækkelige, at det bliver meningsløst at regne. Men de kan ogsaa være saa udtømmende, at vi kan faa ganske ubetingede Resultater, og mellem disse Ydergrænser findes der en hel Række Tilfælde, i hvilke Reglen med vexlende Grad af Nøjagtighed og Frugtbarhed lader sig anvende, blot man opgiver Fantasien om den „absolutte Uvidenhed“ og støtter sig til den reale Viden, der i alt Fald i mange Tilfælde virkelig staar til vor Raadighed.

Endnu staar imidlertid den store Gaade tilbage, hvorledes det er gaaet til, at en Fejltagelse som Dogmet om de konstante Værebrøker i det hele har kunnet skaffe sig Fodfæste indenfor en Disciplin af saa matematisk Beskaffenhed som Sandsynlighedslæren. At erklære Bayes selv for Ophavsmand til dette Dogme, vilde være uretfærdigt. Alle hans Udtalelser om Reglen knytter sig nøje til det konkrete Exempel, ved hvis Hjælp han udleder den, og her er de konstante Værebrøker som sagt berettigede. I det højeste vilde man derfor om Bayes kunne sige, at han egenlig kun har opstillet den mere specielle Regel og endnu ikke den almene. Men hvad der har Interesse, er jo overhovedet slet ikke dette, hvem der først har fundet paa hint Dogme, men meget mere den Omstændighed, at et saadant Dogme i det hele har kunnet vinde Anerkendelse, som det synes endog blandt de mest fremragende i Videnskaben.

Og dette lader sig sikkert kun forklare ud fra den Antagelse, at en stærk historisk Faktor her har gjort sig gældende. Medens Discipliner som Filosofi, Historieforskning o. dl. i Almindelighed gør alle Slægtens Svingninger under dens Udvikling med og i Reglen fremtræder stærkt farvede af den netop herskende Tidsaand, synes Matematikken at staa ganske uberørt af al saadan Ensidighed og at udvikle sig ganske retlinjet og rationelt, højt hævet over al menneskelig Ufuldkommenhed. En nærmere Undersøgelse viser alligevel, at dette ikke gælder i fuld Udstrækning og særlig ikke overfor de matematiske Fag, der, som f. Ex. Sandsynlighedsregningen, mest nærmer sig de reale Videnskaber. Ogsaa her gør Tidsaanden til en vis Grad sin farvende Indflydelse gældende, og et interessant Exempel herpaa giver upaatvivlelig det ovennævnte Dogme os.

Det var i den Kulturperiode, man har kaldt Rationalismen, Bayes' Regel fremkom. Det store hos denne Periode bestod i, at den paany opdagede og hævdede Tankens Magt og Tankens Betydning, der nærmest havde været glemt siden Grækernes Blomstringstid. Det ufuldkomne hos den laa i, at den ikke formaaede at give den nye Ide den rette Afgrænsning. Der er i den hele Periode en stærk og voxende Tendens til at ville udlede hele Tilværelsens Beskaffenhed gennem „den rene Tænkning“ alene og en ligeledes voxende Ulyst til at bekymre sig videre om de positive empiriske Udgangspunkter. Først senere bliver man ganske klar over, at saa vist som enhver logisk Slutning maa have sine Forudsætninger, saa vist maa ogsaa det hele Tankearbejde, der skal bestemme Virkeligheden, have sine tilstrækkelige Udgangspunkter i den samme positive Virkelighed. Men netop en tilsvarende Tendens og en tilsvarende Uklarhed er det jo, der karakteriserer det føromtalte Dogme, og det bliver dermed ret forklarligt, at det overhovedet har kunnet vinde Fodfæste indenfor et Omraade, hvor der ellers i Almindelighed holdes skarp kritisk Grænsevagt. At der ikke desto mindre af og til kan trænge sig uberettigede Anskuelser ind, derpaa frembyder jo for øvrigt i vore Dage hele det Fænomen, der er kaldt „den ikke-euklidiske Geometri“, et ganske tilsvarende, endnu mere fremtrædende Exempel.