MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM
STATISTISCHE AFDELING

Leiding: Prof. Dr D. van Dantzig
Chef van de Statistische Consultatie: Prof. Dr J. Hemelrijk

Rapport S182 (V12)

Wat is en waarvoor dient de statistiek

door

Prof.Dr J. Hemelrijk

October 1955

Wat is en waarvoor dient de statistiek

door Prof.Dr J. Hemelrijk

1. Inleiding.

In kranten en tijdschriften treft men soms lijsten met vragen aan, die van een aantal antwoorden voorzien zijn, waaruit men dan zijn keuze kan maken. De lezer kan daar dan zijn algemene ontwikkeling aan toetsen. Het zou interessant zijn de antwoorden van een groot aantal mensen te hebben op de volgende vragen.

Vraag	Mogelijke antwoorden *)
Wat is Statistiek?	1. Een bepaald soort straatverlichting 2. Een ander woord voor dictatuur 3. Een tabel met cijfers 4. Een bepaald soort journalistiek 5. Een onderdeel der toegepaste wiskunde 6. Een goocheltruc 7. Een bureau in Den Haag 8. Een onderdeel van het leger 9. Een administratie methode 10. ---------- **)
Waarvoor dient Statistiek?	1. Voor vermaak van het publiek 2. Voor het verkrijgen van overzicht over massale gegevens 3. Om na te gaan of een geneesmiddel werkt 4. Als hulpwetenschap op vrijwel ieder gebied 5. Nergens voor 6. Voor snelle verplaatsing van zwaar materiaal 7. Voor bevolkingsadministratie 8. Voor de verkeersveiligheid 9. Om reclame te maken 10. ---------- **)

De uitslag van een dergelijk onderzoek is niet gemakkelijk te voorspellen, maar van de juiste antwoorden zou toch vermoedelijk het merendeel betrekking hebben op de verzamelende en beschrijvende statistiek.

---

*) Verschillende antwoorden kunnen als juist worden aangegeven.

**) Desgewenst zelf invullen.

Men denkt bij het woord "statistiek" vaak in de eerste plaats aan boeken vol tabellen, massale administratieve gegevens en bevolkingsadministratie. Verder aan grafieken, soms beeldgrafieken met hele en halve mannetjes, indexcijfers en opinieonderzoek. Inderdaad vormen deze gebieden een belangrijk onderdeel van de statistiek en hun maatschappelijke betekenis is groot. De statistiek bestrijkt echter een veel breder terrein dan dat der massale gegevens en wordt tegenwoordig als hulpwetenschap op zeer vele gebieden van wetenschap, techniek en samenleving toegepast, ook op waarnemingsmateriaal van kleine omvang.

De bekendheid van de methoden en principes van de statistiek is in het algemeen nog zeer gering, evenals het besef, dat een kritische instelling bij de toepassing en interpretatie van het grootste belang is. Dit besef treft men overigens in de historie hier en daar wel aan, soms zelfs in extreme vorm. Getuige b.v. het uitentreuren aangehaalde gezegde, dat aan Disraely wordt toegeschreven: "There are lies, damned lies and statistics".

De noodzaak van een kritische instelling wordt uitvoerig naar voren gebracht in een kort geleden verschenen boekje van D. HUFF, "How to lie with statistics" (N.Y.1954), waarin vele voorbeelden van bewuste en onbewuste fouten worden gegeven. Wij ontlenen aan dit boekje de volgende anecdote.

De "Reader's Digest" liet eens van een aantal sigaretten van verschillende merken het gehalte aan nicotine en andere vergiften analyseren en publiceerde het resultaat in extenso. Dit was, dat ze allemaal ongeveer even veel gif bevatten en dat er slechts verwaarloosbare kleine verschillen waren gevonden, in tegenstelling tot de reclamecampagnes der fabrikanten. Typografie en kleine verschillen in de waarnemingsresultaten brachten echter tezamen toch één merk onderaan de lijst. Deze gelukkige omstandigheid werd door de betrokken fabrikanten aangegrepen om een grootse reclamecampagne op te zetten, waarbij - onder weglating van alle cijfers - betoogd werd, dat dit merk de minste ongezonde stoffen bevatte en dat dus iedere roker met gezond verstand zich tot dit merk moest beperken. Hetgeen aan de "New Yorker" de verzuchting ontlokte: "There 'll always be an ad man."

HUFF wijst in dit verband op het oude gezegde: "A difference is a difference only if it makes a difference" en wij kunnen hier nog aan toevoegen: Een verschil is alleen een verschil als iets het een verschil maakt. Aan de statistiek is de taak, om na te gaan of een verschil inderdaad een verschil is, d.w.z. om te onderscheiden tussen systematische en toevallige verschillen.

Een ander voorbeeld ter illustratie van de noodzaak van een kritische instelling, deze keer bij de interpretatie van statistisch goed verzamelde en bewerkte gegevens, is het volgende. Over een periode van enige tientallen jaren werd in Engeland een onmiskenbare gelijktijdige stijging geconstateerd van enerzijds het aantal radiobezitters en anderzijds het aantal verpleegden in inrichtingen voor psychisch gestoorden. Men zegt dan, in statistisch jargon, dat deze twee aantallen gedurende de beschouwde tijdsperiode sterk gecorreleerd waren. Ontstaat deze correlatie nu doordat speciaal psychopaten een radio aanschaffen of is de stijging van het aantal psychopaten een gevolg van het luisteren naar de radio? Betrof het de Nederlandse radio dan zou er wel reden zijn om in het bijzonder het tweede als een reële mogelijkheid te beschouwen. Zet men zelf de radio af, dan wordt men nog gek van die van de buren. En is men zelf bestand tegen de affreuze Nederlandse programma's, dan maakt men er de buren gek mee. Of dit in Engeland, het land der uiterste beleefdheid, - waar bovendien de programma's zoveel beter zijn - , ook zo is, valt te betwijfelen. En in ieder geval is er een veel meer voor de hand liggende indirecte oorzaak voor de correlatie aan te wijzen: gedurende de jaren, waar het onderzoek over ging, was de radio juist in opkomst en steeg dus het aantal luisteraars monotoon. Iedere andere grootheid, die in die tijd een stijging of daling vertoonde, om welke reden dan ook, vertoont dus een grote positieve of negatieve correlatie daarmee. En de stijging van het aantal verpleegden kan b.v. het gevolg zijn geweest van het toenemende levenstempo en de steeds groeiende gejaagdheid om ons heen.

Bij dit voorbeeld heeft men geen scherpe blik nodig om in te zien, dat een correlatie niet op een direct oorzakelijk verband tussen de beide onderzochte grootheden behoeft te wijzen. De gevallen zijn echter legio, waarbij dit minder voor de hand ligt en waarbij een verkeerde interpretatie het resultaat is.

Intussen zijn wij nog niet gevorderd met de beantwoording der in de titel gestelde vragen. Wij zullen daar nu aan beginnen.

2. Voorbeelden van statistische problemen.

De statistiek wordt wel eens kortweg beschreven als een methode om langs exacte weg tot een beslissing of een conclusie te geraken in onzekere situaties. De onzekerheid is het centrale kenmerk van de statistiek, zowel wat de gegevens als wat de conclusies betreft. Nu is het niet moeilijk om van onzekere gegevens tot onzekere conclusies te komen. Dit gebeurt dagelijks (men denke b.v. aan de beoordeling van leerlingen) en men kan zich afvragen, welke voordelen dan wel verbonden zijn aan een exacte methode hiervoor. Dit is een punt, dat later ter sprake zal komen.

De onzekere gegevens hebben betrekking op grootheden of verschijnselen, die bij herhaalde waarneming een onregelmatige variabiliteit vertonen en zich daardoor als moeilijk of slechts gedeeltelijk voorspelbaar voordoen. Enkele voorbeelden zijn:

Het weer,
het aantal auto-ongelukken in een weekend,
het aantal beschuiten in een rol,
de wachttijd bij de Hembrugpont,
het jaarlijkse nationale inkomen,
de opbrengst van een akker,
de straling van een vlam,
de lengte van een walvis,
de hoogte van een stormvloed,
de snelheid en mate van herstel van een patient,
enz.

Enkele typische problemen, ontleend aan verschillende gebieden van statistiek en toepassingsterrein, zijn, kort beschreven, de volgende.

1. De uitwerking te bepalen van een gif op een bepaald soort insect. Welke concentratie en hoeveel van het gif moet men gebruiken om een insectenplaag van bepaalde omvang doeltreffend te bestrijden. (Schattingstheorie)

2. Te onderzoeken of de werkzaamheid van twee geneesmiddelen voor een bepaalde ziekte verschilt, indien een eventueel verschil niet zo sterk is, dat men met een grof experiment reeds tot een conclusie kan geraken. (Toetsingstheorie)

3. De ontwikkeling te voorspellen van de bevolkingsaanwas van een land. (Tijdreeks-analyse, groeikrommen)

4. Een maatsysteem voor confectie te ontwerpen, dat met zo weinig mogelijk kosten voor een zeer groot percentage der bewoners van een land in passende confectiecostuums voorziet. (Regressie-en correlatierekening; "Operations-research", over een Nederlandse term is men het nog niet eens.)

5. Het ontwerpen van een eenvoudig routineschema voor de controle van een snelverlopende serieproductie, waardoor de kwaliteit der afgeleverde producten op een bepaald niveau gebracht of gehouden wordt. (Kwaliteitsbeheersing)

6. Het onderzoek naar de invloed, die verschillende factoren, zoals de soort en hoeveelheid van brandstof en snelheid en hoeveelheid van toegevoerde lucht, hebben op de straling van een vlam in een oven, en de

wisselwerking van deze factoren. (Variantieanalyse)

7. Voorspelling van de invloed, die een bepaalde loonsverhoging zal hebben op het prijspeil van bepaalde artikelen. (Econometrie)

8. Het reduceren van een groot aantal kenmerken, die waargenomen zijn aan ieder exemplaar van een verzameling objecten (b.v. proefpersonen), tot een kleiner aantal, met een zo gering mogelijk verlies aan informatie. (Factoranalyse)

Deze voorbeelden bestrijken slechts een kleine fractie zowel van de mathematische statistiek als van het toepassingsgebied. Voor een eerste indruk van de aard der problemen zijn zij echter wel voldoende.

3. De statistische gedachtegang uiteengezet aan de hand van een voorbeeld.

De eerste stap van een experimenteel onderzoek is het vaststellen van het doel, dat men wenst te bereiken. Het belang daarvan wordt wel eens onderschat. Een vage doelstelling houdt echter het gevaar in, dat men achteraf niet in staat blijkt te zijn bepaalde belangrijke vragen te onderzoeken, omdat met deze vragen bij de verzameling van het waarnemingsmateriaal geen rekening is gehouden.

Wij maken het ons, wat dit betreft, gemakkelijk door als voorbeeld te kiezen een onderzoek naar de al of niet werkzaamheid van een nieuw geneesmiddel (G) voor een ziekte, die voorheen met een ander middel (A) werd behandeld. Het doel is dan na te gaan, of één van beide middelen beter werkt dan het andere.

De tweede stap is het opstellen van een plan voor het verrichten der proeven en/of waarnemingen. Daarover bestaat een uitgebreide statistische litteratuur, die zijn bestaansrecht ontleent aan het feit, dat alleen indien aan bepaalde voorwaarden voor de experimenten voldaan is een verantwoorde statistische analyse mogelijk is. Bovendien zijn er proefschema's ontwikkeld, die een zo groot mogelijke doeltreffendheid bezitten, d.w.z. die erop gericht zijn het aantal te verrichten waarnemingen - met gelijke kans op een duidelijke en practisch bruikbare conclusie - zo klein mogelijk te maken. Het te kiezen schema hangt van allerlei omstandigheden af. Eén der belangrijkste is de aard der waarnemingen en de maat, waarin deze worden uitgedrukt. Ook daarin heeft men veelal de keuze tussen verschillende mogelijkheden. Voor de eenvoud nemen we aan, dat bij ons voorbeeld als maat genomen wordt: het aantal dagen dat verloopt tussen de eerste toediening van het geneesmiddel (G of A) en de eerste koortsvrije dag van de patient. De volgende proefopzet kan dan worden gebruikt. Deel de te behandelen patiënten door loting in in

twee groepen, A en G, waarbij de patiënten van groep A geneesmiddel A en die van groep G geneesmiddel G toegediend krijgen. Deze loting dient te voldoen aan de volgende eis. Indien er in totaal N patiënten zijn, die verdeeld worden in twee groepen van m resp. n (m+n=N), dan moeten alle mogelijke verdelingen (dat zijn er dus $\binom{N}{m}$) even waarschijnlijk zijn.[1] Bij voorkeur zorge men er verder voor dat m en n ongeveer (of precies) gelijk zijn, daar dit de kans op het ontdekken van een verschil tussen A en G, als dat aanwezig is, maximaliseert.

De derde stap is nu het verrichten van het experiment en het noteren der uitkomsten. Het resultaat hiervan zijn twee reeksen van waarnemingen:

$x_1,\ldots,x_m$ voor groep A

$y_1,\ldots,y_n$ voor groep G,

getallen, die dus het aantal dagen aangeven tussen het begin van de behandeling en de eerste koortsvrije dag van de patiënt en die we, tezamen genomen, aangeven met $z_1,\ldots,z_N$.

Vervolgens komt de statistische bewerking, die hier het karakter draagt van een toetsing. Men volgt daarbij ongeveer het schema van een bewijs uit het ongerijmde en gaat daarom (voorlopig) uit van de hypothese H_o, dat middel A niet beter is dan G en G niet beter dan A, kortom: dat het om het even is welk middel men gebruikt. Indien deze onderstelling juist was had men even goed aan alle patiënten A (of G) toe kunnen dienen en dan is de verdeling van de waarnemingen $z_1,\ldots,z_N$ in twee groepen een zuiver toevallige, n.l. tot stand gekomen door de loting, die de verdeling in de twee groepen heeft bepaald.

Nu is het intuitief duidelijk, dat H_o verworpen zal moeten worden als de waarden $x_1,\ldots,x_m$ alle of bijna alle kleiner zijn dan de waarden $y_1,\ldots,y_n$ (dat wijst er nl. op, dat A beter is dan G) en eveneens indien het omgekeerde het geval is. Wij zoeken daarom een maat voor de ligging der beide groepen waarnemingen t.o.v. elkaar. Hiervoor bestaan vele mogelijkheden. Wij kiezen, om de gedachten te bepalen, een eenvoudige, nl. het aantal der positieve verschillen onder

$$v_{ij}=x_i-y_j \qquad (i=1,\ldots,m;\ j=1,\ldots,n).$$

---

1) Het waarschijnlijkheidsbegrip bij een loterij veronderstellen wij voldoende bekend. Gelijke waarschijnlijkheden van twee of meer gebeurtenissen betekent, dat deze gebeurtenissen bij een lange reeks lotingen ongeveer even vaak optreden.

Dit aantal noemen wij U [2] en, de niet essentiële complicatie van gelijke waarnemingen buiten beschouwing latende, is het duidelijk dat U een waarde tussen 0 en mn aanneemt. Grote (resp. kleine) waarden van U vormen een aanwijzing dat G beter is dan A (resp. A beter dan G). Indien een waarde in de buurt van $\frac{1}{2}$mn gevonden wordt zal men noch tot het één noch tot het ander kunnen besluiten en zal men dus de hypothese H_o niet kunnen verwerpen. Het onderzoek eindigt dan onbeslist.

Een precisering van deze gedachtegang wordt als volgt verkregen. Bij ieder der $\binom{N}{m}$ splitsingen van $z_1,\ldots,z_N$ in twee groepen $x_1,\ldots,x_m$ en $y_1,\ldots,y_n$ behoort één waarde van U, die men kan berekenen. Ieder van deze $\binom{N}{m}$ splitsingen heeft echter, indien H_o juist is, dezelfde waarschijnlijkheid. Op grond hiervan kan men de waarschijnlijkheidsverdeling van U direct opstellen, door voor iedere waarde van U te tellen bij hoeveel der splitsingen deze waarde behoort. Daarbij blijkt, dat de waarden in de buurt van $\frac{1}{2}$mn het meest frequent optreden en dat naar beide zijden de frequentie monotoon afneemt. Voor ieder getal a kan men nu uitrekenen welke fractie der splitsingen een $U \leqq a$ geeft en men kiest nu een getal α (de onbetrouwbaarheidsdrempel) en bepaalt twee getallen U_l en U_r zodanig, dat de fractie der splitsingen met $U \leqq U_l$ en die der splitsingen met $U \geqq U_r$ beide $\leqq \frac{1}{2}\alpha$ zijn en deze grens zo dicht mogelijk naderen. Vindt men nu bij het experiment een waarde van U, die $\leqq U_l$ is, dan besluit men tot: A beter dan G, en bij $U \geqq U_r$ tot: G beter dan A. Waarden tussen U_l en U_r leiden niet tot een beslissing.

De consequentie van deze handelwijze is nu uit het voorafgaande duidelijk. Indien H_o juist is bezitten alle splitsingen gelijke waarschijnlijkheden en dus is de kans om een waarde van $U \leqq U_l$ of $\geqq U_r$ te vinden $\leqq \alpha$. Dit betekent, dat men een kans $\leqq \alpha$ bezit om H_o te verwerpen, indien hij juist is. Voor α kiest men een niet te groot getal, veelal 1/20 of 1/100. Bij toepassing van deze methode komt men dan slechts zelden (nl. gemiddeld hoogstens 1 op de 20 resp. 1 op de 100 keer) tot een onjuiste uitspraak.

Heeft nu de statisticus zijn berekeningen uitgevoerd dan is de interpretatie van de resultaten, strikt genomen, niet meer zijn taak, maar - in ons voorbeeld - die van de medicus. Deze interpretatie lijkt in dit geval zeer eenvoudig en is in het bovenstaande ook reeds verwerkt. Toch zijn er ook in een zo eenvoudig geval nog haken en ogen. Indien de beide behandelingen duidelijk verschillend van karakter zijn, b.v. als G een behandeling met injecties is en A niet, dan kunnen psychologische factoren in het geding komen. Deze factoren kunnen soms grote invloed uitoefenen en het is verre van ondenkbaar, dat men dientengevol-

---

2) Dit is "de U van WILCOXON", de ontwerper van deze toets.

ge een verschil tussen de groepen A en G zou vinden. Een ongeveer even groot verschil zou dan echter wellicht ook gevonden zijn indien men inplaats van G dummy-injecties gegeven had, b.v. van physiologisch zout. Men moet dus toch nog voorzichtig zijn met de conclusie en indien het gevaar van een dergelijk psychologisch (of een ander) effect aanwezig is, moet men de proefopzet nog uitbreiden, b.v. door een derde groep patiënten met dummy-injecties te behandelen en deze groep met G te vergelijken.

Deze beschrijving van een statistisch onderzoek is uiterst summier en onvolledig. De tijd laat niet anders toe. Zo zijn vele details nog onbesproken, waaraan wij slechts een kort woord kunnen wijden, doch die bij de concrete uitvoering van een onderzoek van belang zijn.

We hebben als maat voor de ziekteduur genomen: het aantal dagen tussen het begin van de behandeling en de eerste koortsvrije dag van een patiënt. Ook andere factoren, zoals leeftijd, geslacht, hoogte der koorts en het aantal koortsdagen voor het begin van de behandeling zijn van belang. Het is dan ook aan te raden deze gegevens ook op te nemen; bij de statistische verwerking kan men daar dan rekening mede houden volgens verschillende, hier niet te beschrijven methoden. De beschreven toets is echter ook steeds toepasbaar in de boven beschreven vorm, waarbij met deze factoren geen rekening is gehouden. Daarvoor dient nl. juist die loting!

Een tweede kwestie is, of de gekozen maat medisch gezien van veel betekenis is. Bij malaria b.v. zal dit wel niet het geval zijn, bij andere ziekten, waarbij het dalen van de koorts op volledige genezing wijst, echter wel.

Verder zijn wij voorbij gegaan aan de motieven, die de keuze tussen de toetsingsgrootheid U of een andere (b.v. het verschil tussen de gemiddelden van $x_1,\ldots,x_m$ en $y_1,\ldots,y_n$) bepalen. Wij hebben niet besproken hoe groot de beide reeksen waarnemingen moeten zijn om een bepaald verschil tussen A en G te ontdekken en wij hebben alleen in principe aangegeven hoe de waarschijnlijkheidsverdeling van U wordt bepaald, terwijl daarvoor veel snellere methoden ontwikkeld zijn.

Een opmerking van algemene aard, die voor de interpretaties van het resultaat van belang is, is verder de volgende: indien een verschil tussen A en G is gevonden, b.v. ten gunste van G, wil dit nog niet zeggen, dat G voor iedere patiënt beter is dan A. "G is beter dan A" betekent: indien men vele patiënten met G behandelt zal in de regel de ziekteduur korter zijn dan bij behandeling met A. Echter niet altijd. Voor één apart geval kan men dit niet voorspellen. De "onvoorspelbaarheid in het klein", die aan statistische verschijnselen eigen is, kan door de statistiek niet worden opgeheven.

En tenslotte - en daaraan willen wij nog enige aandacht wijden - wij hebben het waarschijnlijkheidsbegrip gebruikt alsof dit bekend ondersteld mocht worden. Dit kan voor een zo eenvoudig geval wel zonder veel gevaar voor verkeerd begrip geschieden, maar een goed inzicht in de statistiek, zoals die tegenwoordig wordt beoefend, is niet mogelijk zonder een redelijke kennis van de waarschijnlijkheidsrekening.

4. De plaats van het waarschijnlijkheidsbegrip in de statistiek.

Het begrip "kans" of "waarschijnlijkheid" (afkorting: wh, meervoud: whn) kan in principe bij de statistiek wellicht worden gemist. Het kan uit de vorige paragraaf gemakkelijk worden geweerd door aan de loterij de eis te stellen, dat bij een lange reeks lotingen alle mogelijke splitsingen ongeveer even vaak voorkomen, terwijl toch de resultaten van reeds verrichte lotingen geen invloed uitoefenen op de volgende lotingen. Een dergelijke loterij "zonder geheugen" is zeer wel te construeren, zoals wij straks zullen laten zien. Men kan dan het gehele betoog stellen op de experimentele basis van een dergelijke loterij en komt dan tot precies dezelfde conclusie als eerst: dat H_o slechts ongeveer in een fractie $\leqq \alpha$ van een lange reeks toepassingen van deze methode ten onrechte zal worden verworpen.

Er zou wellicht enige reden zijn te trachten het kansbegrip uit de statistiek te weren, daar er over de philosophische zijde van dit begrip zeer uiteenlopende meningen bestaan. Volgt men hiertoe de vorengeschetste weg, dan vermijdt men echter alleen de term en gebruikt het begrip toch in één van zijn interpretaties, nl. de frequentie-interpretatie. Deze houdt in, dat men een kans p op een verschijnsel V interpreteert als een bij een bepaalde situatie behorend getal, dat uitdrukt in welke fractie v ongeveer van de gevallen, waarin deze situatie optreedt het verschijnsel V zich zal voordoen. Deze fractie v ,die een frequentie-quotiënt (afkorting: fq, meervoud: fqn; notatie: fq(v)) wordt genoemd, is voor concrete reeksen van waarnemingen te berekenen en wordt dan als een meting van p beschouwd, die, zoals alle of vrijwel alle metingen, een zekere onnauwkeurigheid bezit. Dit laatste heeft tot gevolg dat men in alle voorafgaande formuleringen het woord "ongeveer" tegenkomt.

Uitgaande van dit standpunt is er echter geen dringende reden om de term "wh" te vermijden; men zou de resultaten der waarschijnlijkheidsrekening (afkorting: whr) toch gebruiken, zij het in vertaalde vorm.

De beantwoording van de vraag: "wat is statistiek" kan dus ook kort als volgt worden gegeven. Statistiek is toegepaste waarschijnlijkheidsrekening en whr is een (b.v. axiomatisch opgezet) onderdeel van de zuivere wiskunde, in het bijzonder van de maattheorie. Men ontleent aan de

whr een mathematisch model voor het beschouwde probleem en de whn hebben daarbij het karakter van modelwaarden van fqn. Het ligt daarom voor de hand de axiomas te ontlenen aan eigenschappen van fqn en dit geschiedt dan ook. Een summiere toelichting van dit procédé volgt hieronder.

Een fq is steeds het quotient van het aantal elementen van een verzameling, die een zeker kenmerk vertonen, en het totale aantal elementen van die verzameling. Dergelijke verzamelingen, die in de statistiek populaties worden genoemd, hebben wij in het voorafgaande telkens ontmoet zonder ze steeds expliciet te noemen. Voorbeelden: een reeks experimenten, een groep patiënten, de insecten van een bepaalde soort, de producten van een fabriek, etc. Ook toekomstige of denkbeeldige elementen (zoals experimenten die men zal of zou kunnen verrichten, nog te fabriceren exemplaren van een artikel) kunnen tot een populatie behoren en dit is zeker zo, indien men tot voorspellingen wil komen.

Voor de eenvoud beperken wij ons tot eindige populaties. Laat nu $A_1, A_2, \ldots$ kenmerken voorstellen, die de elementen van een bepaalde populatie al of niet bezitten. Geven wij de aantallen elementen, die deze kenmerken bezitten, aan met $n(A_1)$, $n(A_2)$, etc. en het totale aantal met N, dan is dus

$$fq(A_i) = n(A_i)/N \qquad i=1,2,\ldots$$

en de volgende eigenschappen volgen direct uit deze definitie:

1) $0 \leq fq(A_i) \leq 1$ voor alle i;
2) $fq(A_i)=0$ dan en slechts dan als A_i niet voorkomt;
3) $fq(A_i)=1$ dan en slechts dan als A_i steeds voorkomt;
4) $fq(A_i \text{ of } A_j)=fq(A_i)+fq(A_j)-fq(A_i \text{ en } A_j)$,

waarin "of" betekent: één van beide of beide.

Het wh-theoretische model wordt nu geheel analoog hieraan opgebouwd. Een modelverzameling met elementen γ vormt de basis. Deze elementen stellen nu echter niet, zoals bij de populatie, concrete objecten of uitkomsten voor, maar mogelijke uitkomsten, eventualiteiten.

Een eenvoudig voorbeeld, waaraan het verschil tussen populatie en collectie goed te zien is, is het werpen met een knoop. Stellen wij de beide mogelijke uitkomsten voor door H (holle kant boven) en B (bolle kant boven), dan zal de populatie b.v. bestaan uit een reeks worpen,

waaronder weliswaar toekomstige kunnen zijn, maar die toch ieder één der uitkomsten H of B hebben of zullen hebben. De modelverzameling bestaat echter, voor iedere worp, uit twee elementen H en B, de beide mogelijke uitkomsten en het model voor een reeks van n worpen is het product van n dergelijke collecties van ieder 2 elementen en bevat dus 2^n elementen, die de mogelijke uitkomsten van de gehele reeks voorstellen.

Op een dergelijke verzameling Γ zijn nu kenmerken A_i gedefinieerd, d.w.z. dat van ieder element gegeven is, of het A_i ($i=1,2,\ldots$) bezit of niet. De deelverzamelingen van Γ, bestaande uit alle elementen, die het kenmerk A_i bezitten, geven wij kortweg ook met A_i aan. De collectie wordt nu een wh-veld genoemd, als op Γ een verzamelingsfunctie P gedefinieerd is, die aan de volgende axiomas[3] voldoet:

1) $0 \leqq P[A_i] \leqq 1$ voor alle i;

2) $P[A_i]=0$ als A_i onmogelijk is;

3) $P[A_i]=1$ als A_i zeker is;

4) $P[A_i \text{ of } A_j] = P[A_i] + P[A_j] - P[A_i \text{ en } A_j]$.

$P[A_i]$ wordt "de kans op A_i" genoemd (P=probability).

De tijd laat niet toe, dat wij volledig bespreken, hoe uit dit model weer de weg terug kan worden gevonden naar de practijk, d.w.z. hoe stellingen, afgeleid uit deze axiomas, in de vorm van uitspraken met fqn in plaats van whn geïnterpreteerd kunnen worden. Het ligt voor de hand dat dit geschiedt door een wh weer te interpreteren als een benadering van een fq in een lange reeks waarnemingen of experimenten, maar de rechtvaardiging, die de theorie daarvoor biedt kunnen wij niet bespreken. Wel willen wij even aangeven, als zeer simpel voorbeeld van de rekenwijze met deze axiomas, hoe men op eenvoudige wijze kansen van voorgeschreven grootte kan construeren met primitieve materiële hulpmiddelen. Daarvoor hebben wij nog één extra begrip nodig, dat van stochastische onafhankelijkheid.

Twee eventualiteiten A_1 en A_2 heten stochastisch onafhankelijk indien $P[A_1 \text{ en } A_2]=P[A_1]\ P[A_2]$ is. De practische interpretatie hiervan is, dat het al of niet optreden van A_1 het al of niet optreden van A_2 niet beïnvloedt. Beschouwen wij nu twee opeenvolgende worpen met één onverslijtbare knoop, verricht op dezelfde wijze, dan is bij ieder van die worpen de kans op H dezelfde, b.v. p:

$$P[H] = p$$

---

3) Behalve de hier genoemde axiomas is er nog een limietaxioma, dat uitsluitend van belang is indien men oneindige collecties beschouwt, hetgeen bepaalde wiskundige voordelen meebrengt doch voor de principes der opzet niet van belang is.

en $$P\left[B\right] = q = 1 - p \ ,$$

zoals gemakkelijk uit de axiomas af te leiden is. Aannemende, dat de uitkomst van de eerste worp die van de tweede niet beïnvloedt - een onderstelling, die zeer aannemelijk lijkt, maar die desgewenst experimenteel geverifieerd kan worden -, verkrijgen wij als wh-veld voor een paar van worpen in korte notaties:

$$\left.\begin{aligned} P\left[HH\right] &= p^2, \\ P\left[HB\right] &= pq, \\ P\left[BH\right] &= qp, \\ P\left[BB\right] &= q^2. \end{aligned}\right\} \quad \Gamma_2.$$

Nu is $pq=qp$, dus $P\left[HB\right]=P\left[BH\right]$, wat ook p moge zijn. Verkleinen wij nu Γ_2 tot Γ_2' door de elementen HH en BB weg te laten - wat in de practijk betekent, dat wij paren worpen, die HH of BB geven, buiten beschouwing laten dus doorgaan met het verrichten van paren worpen tot wij ofwel HB ofwel BH verkrijgen - dan geldt voor Γ_2' evenals voor Γ_2:

$$P\left[HB\right] = P\left[BH\right] \ ,$$

maar daar HB en BH de enige elementen van Γ_2' zijn, is nu $P\left[HB\right]=1-P\left[BH\right]$, zodat beide $=\frac{1}{2}$ zijn.

Wij hebben nu dus, met behulp van een knoop, een kans $=\frac{1}{2}$ geconstrueerd; een "zuivere" munt is daarvoor dus niet nodig. Langs soortgelijke weg kunnen wij nu iedere rationale wh exact construeren en iedere reële willekeurig dicht benaderen. De betekenis hiervan voor het opzetten van experimenten, die de mogelijkheid van een verantwoorde statistische verwerking garanderen, is in par.3 toegelicht.