

Bylaag tot „Koers” Deel XXVI, September 1958.

Die Eie Karakter en die Terrein van Statistiek

(Inougurele Rede gehou by die aanvaarding van die professoraat in Statistiek aan die P.U. vir C.H.O., op 8 Augustus 1958, deur prof. dr. J. M. de Wet).

Hooggeagte Here Lede van die Raad van die P.U. vir C.H.O.,

Hooggeleerde Rektor, Waarnemende Rektor, Professore en

Dames en Here Dosente van die P.U. vir C.H.O.,

Hooggeagte Dames en Here Studente, Dames en Here.

Die onderwerp wat ek by hierdie geleentheid, die aanvaarding van 'n professoraat in Statistiek aan die P.U. vir C.H.O., met u wil bespreek is:

DIE EIE KARAKTER EN DIE TERREIN VAN STATISTIEK.

INLEIDING.

In 1907 skryf die gevierde Nederlandse wiskundige, dr. L. E. J. Brouwer¹), as volg: Die mense is 'n vermoë eie wat al hul wisselwerkinge met die natuur begelei, die vermoë nl. tot die wiskundige bekyk van hul lewe.

Die invloed van hierdie vermoë laat hom steeds sterker en sterker geld in die wetenskappe wat van kwalitatiewe en kwantitatiewe waarnemings gebruik maak.

Waar Wiskunde tot aan die begin van hierdie eeu hoofsaaklik in die fisiese en tegniese wetenskappe en op enkele ander terreine as belangrike hulpmiddel gebruik was, het dit op hierdie en op nuwe terreine in 'n nuwe gedaante van belangrikheid te voorskyn getree. Waar Wiskunde tot enkele dekades gelede nog hoofsaaklik gebruik was as hulpmiddel om min of meer eksakte verskynsels en natuurwette te ontleed, het dit nou 'n belangrike instrument geword vir die analise van stogasties variërende kenmerke. Die matematiese behandeling van hierdie verskynsels het die toepassingsgebied van Wiskunde baie uitgebrei, soos verder aan sal blyk, en Statistiek het 'n belangrike invloed op die toenemende gebruik van Wiskunde uitgeoefen, en doen dit tans nog. Meer algemeen word op al hoe meer vakgebiede 'n al hoe meer grondige kennis van Wiskunde vereis — 'n verskynsel wat noodwendig in die nabye toekoms sy invloed op die aard en omvang van Wiskunde in skool- en universiteitsleergange sal moet laat geld.

ENKELE GRONDBEGRIPPE EN DIE EIE KARAKTER VAN STATISTIEK.

Statistiek gaan uit van waarneminge en lei, deur induksie, tot uitsprake en voorspellinge — die metode wat meeste ervaringswetenskappe ten grondslag lê en waarop tewens baie van ons alledaagse handelinge gebaseer is.

1. Brouwer, L. E. J.: *Over de Grondslagen der Wiskunde*, p.81. P. Noordhoff-Groningen. 1907.

Die nie-statistiese geval.

Wanneer 'n eksperimenteerder sekere omstandighede skep kry hy 'n bepaalde resultaat of verskynsel. Herhaling van die eksperiment lewer by benadering dieselfde resultaat. Deur induksie word dit as 'n reël aanvaar dat altyd wanneer dieselfde omstandighede geskep word, die bepaalde verskynsel verwag kan word. Wanneer die verkreë verskynsels die omstandighede vir nuwe verskynsels skep, ontstaan 'n kousale reeks.

Brouwer sê die volgende: Die mens is 'n vermoë eie, tot die sien in die wêreld vanherhalinge van volgreekse, van kousale sisteme op grond waarvan lewensmomente uiteenval as volgreekse van kwalitatief verskillende dinge En die lewensgedrag van die mens soek om soveel moontlik van die wiskundige volkreekse te kan waarneem, of telkens waar in die werklikheid by 'n vroeëre element van so 'n reeks met meer sukses skyn ingegryp te kan word as by 'n latere,, die eerste te kies as rigting vir sy daad. Hy noem laasgenoemde stap die sprong van „doel op middel.”)

Ons onderskei die volgende drie stappe:

- (a) Herhaalde waarneming van 'n kousale reeks.
- (b) Induksie: dat die volgreekse hom, na verwagting, sal herhaal.
- (c) Die kies van 'n rigting vir toekomstige daad.

In die wetenskappe bestaan waarneming uit die skep van 'n kousale reeks: die kenmerk(e) wat die resultaat is van die omstandighede wat geskep of gekies is, word dikwels in getalvorm uitgedruk, nl. as die waarde van 'n parameter — en dit is die geval wat nou beskou word. Herhaling lewer 'n deelversameling, waarvan die elemente getalwaardes kry.

Nou bestaan daar 'n basiese onderskeid in die aard en/of benadering van die waarnemingsmateriaal tussen wat ons hier en verder die nie-statistiese geval noem en die statistiese geval.

Op talle wetenskaplike terreine is kenmerke met 'n skerp geïsoleerde aard, d.w.s. met 'n definitiewe kousale karakter, die onderwerp van belangstelling. Indien genoeg eksperimentele voorsorg getref word kan die kenmerk noukeurig genoeg vir praktiese bruikbare doeleindes waargeneem word, bv. die afstand van plek A na plek B is 10 myl; die versnelling van 'n voorwerp wat vry val is 32 voet per sekonde per sekonde; die volume van .08 pond lug onder 14 lb. per vierkante duim druk is 1 kubieke voet. Die beoefenaar is bewus van sy beperkte waarnemingsvermoë en van kleinere steurende effekte in sy eksperiment wat dit onmoontlik maak om 'n bepaalde noukeurigheid te oortref. Maar dit mag gebeur en dit is dikwels so, dat hierdie steurende effekte nie intree op die noukeurigheidsvlak wat vereis word nie, of dat nuttige resultate verkry kan word met die noukeurigheidspeil wat wel bereik kan word deur verwaarlosing van hierdie steurende effekte.

Dit is die nie-statistiese geval en gee aanleiding tot 'n nie-statistiese probleem.

Gevolgtik ontstaan deur herhaling 'n deelversameling waarvan die elemente konstante waardes het of waaraan hoogstens, deur die berekening van 'n gemiddelde, konstante waardes toegeken word.

Volgens die induksiemetode word die denkbeeldige resultate van alle moontlike verdere eksperimente as gelykwaardige elemente by dié van die waargenome deelversameling gevoeg om 'n omvattende versameling te vorm waaruit die resultaat van enige soortgelyke toekomstige eksperiment tevoorskyn tree en met vertroue voorspel kan word.

Die metode kan uitgebrei word deur (a) 'n aantal verskillende reekse wat op dieselfde parameters betrekking het, waar te neem en die kenmerke van elke reeks as 'n stel paramterwaardes te beskou en (b) die resultate saam te vat en deur induksie as 'n wet te beskou.

Brouwer stel dit as volg: „Sy groot mag egter kry die wiskundige natuurwetenskap nog nie deur die opmerk van vir die instink ongeveer gelykwaardige volgreekse nie, maar deur die samevatting van 'n baie groot aantal van sulke volgreekse onderèèn gesigspunt deur middel van 'n met behulp van matematies induksie opgeboude sisteem wat wet genoem word.”)

Voorbeeld uit die natuurwetenskappe:

1. Eenvoudige geval:

- (a) 'n Gas by temperatuur t_1 en druk p_1 het volume v_1 , herhaalde kere waargeneem.
- (b) Verwagting: Toekomstige eksperimente sal dieselfde resultaat lewer (induksie).

2. Uitbreiding:

- (a) 'n Gas by temperatuur t_2 en druk p_2 het volume v_2 (tweede volgreekse) en so meer.
Hieruit blyk $p_1 v_1 = p_2 v_2 = \dots = p_n v_n$.
- (b) Wet: $p v =$ konstant (induksie).

Die statistiese geval.

Soos reeds aangedui, is die „eksaktheid” van waarnemings van beperkte aard en word die geskiktheid van die noukeurighedsgraad dikwels met praktiese behoeftes as maatstaf bepaal, waar moontlik.

In werklikheid is hetrogeniteit altyd teenwoordig.

Die waarnemingsresultaat vir 'n fisiese parameter kan bv. in 'n reeks eksperimente, tot 2 desimale, die getalle 6.52, 6.52, , 6.52 lewer, en as geen

2. Brouwer, L. E. J.: Over de Grondslagen der Wiskunde, p.84. P. Noordhoff-Groningen. 1907.

verdere eise gestel word nie sal statistiese metodes nie nodig wees nie, en die resultaat van 'n verdere eksperiment kan met vertroue as 6.52 voorspel word. Wanneer die bepaling egter tot 4 desimale gedoen word, kan die derde en vierde desimale by verskillende eksperimente verskil en statistiese metodes kan moontlik meer lig op die aard van die parameter werp.

Verder, in teenstelling met die min of meer eksak bepaalbare kenmerke, is daar 'n groot aantal wetenskapsgebiede waar die kenmerke oorheersend stogasties van aard is en die variasie groot is met betrekking tot die meetvermoë van die eksperimenteerder. Wanneer telkens getrag word om dieselfde omstandighede te skep of te kies, is die resultaat nie elke keer dieselfde nie maar varieer dit op 'n wyse buite die beheer van die eksperimenteerder en sodanig dat die resultaat van 'n volgende eksperiment nie met sekerheid voorspel kan word nie. Twee voorbeelde word genoem:

1. Tien muntstukke word herhaalde kere gewerp: meestal toon daar 5 munt, dikwels 4 of 6 munt, minder dikwels 3 of 7 munt, selfde 2 of 8 munt, minder selde 1 of 9 munt en baie selfde 0 of 10 munt.

2. Die liggaamslengte van manspersone word ondersoek en 99 lengtes gemeet. Die 99 gemete lengtes sal waarskynlik so onreëlmatig varieer dat die lengte van die 100ste persoon nie met vertroue voorspel kan word nie.

Waar in die eenvoudigste nie-statistiese geval die reeks waarnemings 'n deelversameling konstantes lewer, kry ons in die statistiese geval 'n deelversameling statistiese waarnemings waarvan die elemente waardes van 'n stogastiese of kansveranderlike is. Dit bepaal die toepassingsterrein van Statistiek, en die analise van stogastiese materiaal vereis 'n denkwyse en 'n teorie met 'n eie karakter.

Histories spruit moderne Statistiek uit twee wortels nl. „politieke rekenkunde”) en waarskynlikheidsrekening, en laasgenoemde was van die sewentiende tot die negentiende eeu sterk deur kansspele beïnvloed. Die gedrag van 'n stogastiese veranderlike moet egter nie aan „noodlot” of „toeval” toegeskryf word nie. Die ware statistikus is nie 'n dobbelaar nie. Ook hierdie — vir ons komplekse — verskynsels vind plaas deur die skepping van die Almagtige en is onderhewig aan die wetsorde van God. Die Christenwetenskaplike sien die hand van God in alle verskynsels. Slegs die mens se onvolmaaktheid verplig hom om waar sy waarnemingsvermoë te kort skiet of waar kontrole gedeeltelik of onmoontlik is, gebruik te maak van metodes om saamgestelde verskynsels te bestudeer.

Die wetenskapsterreine waar waarnemings van 'n sterk stogastiese aard te voorskyn tree, laat hulle nie in erkende vakgebiede afbaken nie. Belangrike toepassingssfeer word gevind op gebiede van die demografiese, ekonomiese, biologiese, ingenieurs- en fisiese wetenskappe, in laasgenoemde geval bv. tot 'n

3. Die volgende publikasies het bv. gedurende die sewentiende eeu verskyn:

Sir W. Petty: *Essays in Political Arithmetick* 1623-1682.

Graunt, J.: *Natural and Political Observations upon the Bills of Mortality*. 1662.

sterk mate op die mikrosiesie gebied. Ten einde die grondprobleme van die geval onder beskouing te bespreek, is dit nodig om enkele grondbegrippe te verduidelik.

Soos die naam aandui word die waargenome deelversameling beskou as 'n deel van 'n omvattende versameling genoem die populasie'). Die populasie bestaand dus uit alle waargenome en nie-waargenome elemente, met kenmerke van stogastiese aard, en wat deur voorgeskrewe omstandighede aan die groep wat ondersoek word, behoort. Die waarnemings vorm 'n steekproef uit die populasie. In die eeuvoedigste geval, nl. wanneer èn kenmerk ondersoek word, kan 'n sekere waarde van die kenmerk besonder baie voorkom, 'n volgende waarde dikwels, 'n ander waarde minder dikwels, en so aan. Die aantal waarnemings met 'n sekere waarde, of waarvan die waarde in 'n sekere gebied lê, word die frekwensie van daardie waarde of gebied genoem en die verhouding hiervan tot die totale aantal waarnemings, die frekwensiekwosiënt. Die frekwensie-kwosiënte vir al die moontlike waardes van die kenmerk gee aanleiding tot waardes van 'n frekwensiefunksie, 'n digtheidsfunksie en van 'n verdelingsfunksie van die kenmerk wat 'n stogastiese veranderlik is.

Die begrip verdelingsfunksie, wat na twee en meer kenmerke uitgebrei kan word, is eie aan Statistiek en verleen 'n eie karakter aan Statistiek.

Die grondprobleme kan nou as volg gestel word):

In die definiëte geval word gevind dat sekere funksies van die parameters, bv. 'n funksie, $f(x,y)$, 'n konstante, numeriekbepaalde waarde besit, in ons geval besit dit 'n verdelingsfunksie.

Die eerste grondprobleem van die Statistiek lui:

Om uit die gegewe waarnemingsmateriaal hierdie verdelingsfunksie van die parameters te bepaal.

Die tweede grondprobleem bestaan in die definiëte geval uitsluitend uit die oplossing na die ontbrekende parameterwaardes van die vergelykings wat ontstaan deur die substitusie van gegewe waardes van die parameters in die konstante funksies.

Statisties is dit die moeilikste probleem, nl. om uit die waarnemingsmateriaal verdelingsfunksies af te lei vir sekere funksies van hierdie parameters. Hierdie probleem is in baie gevalle onopgelos.

Verder blyk dit, soos verwag kan word uit die stogastiese aard van die waarnemings, dat parameters van die omvattende versameling 'n verdelingsfunksie besit wat nie eksak genoeg hoef ooreen te stem met die verdelingsfunksie van die deelversameling nie. Metodes wat hierdie verskil in aanmerking

4) Ander name: Universum, Kollektief.

5. Van Dantzig, D.: Mathematisch Statistiek. Mathematisch Centrum. Amsterdam.

neem gaan egter met groot matematiese moeilikhede gepaard, sodat ons dikwels op die minder eksakte metode moet terugval waarby die frekwensieverdeling van die kenmerke, soos dit by die verrigte waarnemings opgetree het, ten grondslag aan die omvattende versameling of populasie gelê word.

Op grond van die induksiemetode word omtrent definiete kousale verskynsels definiete prognoses, en omtrent stogasties kousale verskynsels stogastiese prognoses opgestel. Dit gee aan Statistiek 'n eie karakter.

Dit is 'n deel van die taak van die statistikus om uitsprake in kwantitatiewe vorm te verskaf en daarnaas die sekerheidsgraad te konstateer. Die statistikus aanvaar die stogastiese aard van die waarnemingsmateriaal en bou 'n wetenskap op wat by die veralgemening van steekproef na populasie die betroubaarheid of ontroubaarheid van die proses in kwantitatiewe vorm uitdruk. In die voorbeeld oor liggaamslengte kan die volgende soort uitspraak gegee word: omdat net 99 persone se lengtes gemeet is kan die gemiddelde lengte van al die persone nie eksak voorspel word nie, maar 'n bepaalde lengte-interval wat op 'n sekere manier bereken word, sal in 99% van sodanige eksperimente die regte gemiddelde insluit.

STATISTIEK AS WISKUNDIGE WETENSKAP.

Die wetenskappe wat Wiskunde as hulpmiddel gebruik volg in hooftrekke, met verskillende beklemtoning, dieselfde metode en dit word deur Van Dantzig¹) as volg gestel:

1. Die waarneming van bepaalde groepe verskynsels. (Versameling van waarnemingsmateriaal).
2. Beskrywing van die waarnemingsmateriaal in geskikte vorm vir matematiese bewerking. (Inskakeling van die matematiese formalisme).
3. Groepering van die waarnemingsmateriaal en die opmerk van bepaalde reëlmatighede daarin, waarby werklike waarnemingsmateriaal deur 'n fiktief vervang moet word deur verwaarlosing van klein of seldsame onreëlmatighede en sonder uitsondering geldig-verklaring van bepaalde opgemerkte reëlmatighede. (Kies van 'n matematiese model deur regularisering van waarnemingsmateriaal.)
4. Die selekteer van bepaalde reëlmatighede in die matematiese model op deduktiewe wyse op grond van die veronderstelde geldigheid sonder-uitsondering van die opgemerkte reëls. (Toepassing van die matematiese formalisme).
5. Die omsit van die geselekteerde reëlmatighede in prognoses. (Uitskakeling van die matematiese formalisme).
6. (Bygevoeg). Toetsing of verifiëring van die matematiese model aan die hand van die getoetste geldigheid van uitsprake.

Om te onderskei tussen Statistiek en ander vakwetenskappe met betrekking tot die plek wat Wiskunde daarin bekleed is moeilik. Selfs matematiese statistici verskil ingrypend, terwyl sommige toepassers die Wiskunde sover

moontlik vermy. Wanneer die verskillende dele van 'n wetenskaplike ondersoek, soos hierbo genoem, op die horisontale as, en die mate van beklemtoning van die dele op die vertikale as van 'n grafiek voorgestel word, sal die grafiek wat vir verskillende vakwetenskappe gekry word varieer vanaf dié met 'n hoë piek in die middel (toepassing van die matematiese formalisme) tot dié waarvan die grafiek na die middel daal. Wiskunde, wat los staan van enige fisiese objek, sou op hierdie grafiek te voorskyn tree as 'n enkele kolom wat op die middel van die horisontale as staan. Daarna volg Toegepaste Wiskunde, wat by die getal en ruimte van Wiskunde ook massa en tyd as objekte het. Die grafiek vir Toegepaste Wiskunde sal 'n hoë piek in die middel hê.

Vervolgens kry ons Teoretiese Fisika, wat ook abstraheer, maar waar die Wiskunde reeds 'n middel tot 'n doel word; en dan volg die Fisiese en Tegniese Wetenskappe, en so aan.

Volgens my mening en wens neem Statistiek 'n plek in net na Toegepaste Wiskunde. Sommige skrywers meen dat Teorie van Statistiek 'n vertakking van Toegepaste Wiskunde is, maar dan is dit gewoonlik nie duidelik wat onder Toegepaste Wiskunde verstaan moet word nie. Vir Toegepaste Wiskunde word hier dan die beskouing van W. P. Robbertse as maatstaf geneem. Dit word deur hom as volg gestel: „Onder Toegepaste Wiskunde wil ons dus verstaan: Meganika, sowel die klassieke as die moderne, Die Meganika voer behalwe hierdie twee (objekte, nl. getal en ruimte) nog twee nuwe objekte in, nl. massa en tyd. Die getalle-kontinuum en die objek tyd leen hulle tot onderlinge vereenselwiging. Hierdie nuwe objekte van massa en tyd word voorasnog nie gedefinieer as, of vereenselwig met bepaalde fisiese entiteite nie. Hulle leen hulle wel daartoe”). Die Statistiek tree in laasgenoemde opsig 'n stap nader aan die Fisika. Behalwe matematiese abstrahering en wiskundige denke beskou Statistiek ook die in- en uitskakeling van formalisme.

Toegepaste Wiskunde leen hom tot toepassing via die fisiese en tegniese wetenskappe. Statistiek staan op eie bene in veel meer wetenskapsgebiede en wedersydse beïnvloeding tussen teorie en ervaring is noodsaaklik. Statistiek kan hom nie geheel en al van die werklikheid skei nie. Die aard van statistiese materiaal bied egter 'n uitstekende geleentheid vir die beoefening van moderne Wiskunde. As vakwetenskap wat aan 'n universiteit doseer word, vorm die matematiese formalisme die ruggraat van statistiek en dit is my vaste oortuiging dat in hierdie rigting volgehou moet word. Die matematiese statistikus behou hom dus die reg voor, en sekerlik nie ten nadele nie, om soos D. J. van Rooy dit uitdruk: „Soos die fantasie in onbelemmerde vlug bo die werklike uit te styg”).

6. Robbertse, W. P.: Die Terrein, Taak en Perspektiewe van die Toegepaste Wiskunde. Rede, Koers, Des. 1950, p.103.

1. Van Rooy, D. J.: Beginsels en Metodes van die Hoër Onderwys. Saamgestel onder redaksie van J. C. Coetzee en D. J. van Rooy. P.253.

Miskien moet ons 'n aksioma-sisteem in Statistiek dan 'n hipotese-sisteem noem?

Die grafiek vir Statistiek sal dus in die middel minstens so hoog soos dié van Toegepaste Wiskunde styg maar na beide kante meer prominent vertoon.

Die „of” en „hoe” van aksiomatiese benadering in Statistiek is die onderwerp van bitsige argumente. Dit blyk duidelik uit die volgende aanhaling, wat die aksiomatiese benadering, met 'n ongeregverdigde motief, veroordeel. Die aanhaling is uit „The Advanced Theory of Statistics” (Vol. I, p.167) van Kendall, 'n vooraanstaande kontemporêre statistikus. Hy sê, verwysende na besware teen 'n definisie van waarskynlikheid: „Die matematiese teoretikus wat met waarskynlikheid te doen kry, kom die moeilikheid te bowe, op die wiskundige se gewone gerieflike manier, deur of die sirkeldefinisie te aanvaar of deur waarskynlikheid bloot as 'n eienskap van puntversamelings te beskou . . . en bou sodoende 'n teorie oor waarskynlikheid op as 'n vertakking van die teorie van versamelingsfunksies. Enige verifikasie van die teorie, enige toets of dit 'n redelik suiwer beeld van die werklike gebeure in hierdie wêreld gee, word na eksperimentele fisika verwys. Die wiskundige is natuurlik gewoon aan hierdie afskuiwing van verantwoordelikhede, maar die statistikus is gemoeid met ooreenstemming tussen teorie en praktyk en kan nie altyd eksperimentele verifikasie aan andere oorlaat nie.”

Dit word gesê nieteenstaande 'n halwe bladsy terug staan: „Vir ons teenswoordige doel kan die fundamentele probleme ter syde gestel word aangesien alle partye dit eens is oor die reëls wat regstreekse kansrekening beheer. Ons noem dus hierdie reëls sonder om 'n poging aan te wend om hulle uit meer primitiewe proposisies af te lei.” Aldus Kendall. Laasgenoemde hou tot 'n mate 'n aksiomatiese benadering in, en boonop is een van die reëls wat dan genoem word juis dié een waarna in die eerste aanhaling verwys word. Die twee aanhalings weerspreek mekaar dus. Dit is in elk geval onverdiende kritiek op wiskundiges en nie ons opvatting van die plek wat Wiskunde in Statistiek inneem nie. Selfs Oliver Heaviside, wat deur selfstudie die werke van Maxwell bemeester het en toe belangrike wiskundige bydraes tot die elektrisiteitsleer gelewer het, maar gedurig met universiteitswiskundiges oor rigoriesheid oorhoops gelê het, het gesê: „Even Cambridge mathematicans deserve credit.”

Die grondliggende moeilikheid is dat niemand nog ooit daarin geslaag het om 'n definisie van waarskynlikheid, wat 'n basiese rol in Statistiek bekleed, te gee wat volkome by al die gebruike van die woord pas nie.

Sedert die sewentiende eeu het baie beroemde wiskundiges en filosowe hul aan die probleem bewaag. U aandag word op slegs enkele van hierdie gedagterigtings gevestig. Die eerste ernstige poging om waarskynlikheid te definieer

verskyn in 'n boek deur die beroemde Franse wiskundige Laplace as volg: Die waarskynlikheid van 'n gebeurtenis is die verhouding van die aantal gunstige gevalle tot die totale aantal ewe moontlike gevalle⁸).

So lank kansrekening sy vernaamste toepassingsgebied in die kansspele gehad het, was hierdie definisie gangbaar en was algemeen deur wiskundiges van die negentiende eeu aanvaar. Die eerste beswaar teen Laplace se definisie, behalwe dat die woorde „ewe moontlik” die indruk van 'n sirkelredenasie skep, is dat dit di sogenaamde indifferensieprinsiep, of beginsel van gelyke onbekendheid bevat, wat tot kontradiksies lei. Die besware het egter nog groter afmetings aangeneem tot die begrip waarskynlikheid in biologiese en sosiologiese en soortgelyke waarnemings begin gebruik is. In 'n bewering soos „die waarskynlikheid dat 'n sekere saadsoort sal ontkiem is .25”, is dit moeilik om die probleem met 'n aantal „ewe moontlike” gevalle te assosieer. Die saad ontkiem of dit ontkiem nie en die gebeurtenisse is nie „ewe waarskynlik” nie.

Richard von Mises⁹) het die nuwe gedagtes wat na aanleiding van kritiek op Laplace se definisie na vore gekom het, in die vorm van 'n frekwensielimiet-teorie uitgewerk. Hy beskou 'n oneindige ry gebeurtenisse, genoem 'n kollektief, en definieer die waarskynlikheid van 'n kenmerk A as die matematiese limiet van die verhouding van die aantal elemente met kenmerk A tot die totale aantal elemente, d.w.s. as die limiet van 'n frekwensieverhouding of -kwosient. Hierdie teorie het regstreeks en onregstreeks veel tot die studie van waarskynlikheid bygedra maar lei tot matematiese moeilikhede. Een beswaar is bv. dat die limiet van 'n oneindige ry nie verander as 'n eindige aantal elemente deur ander vervang word nie, terwyl in praktiese toepassinge altyd eindige rye waargeneem word.

Alhoewel geen teorie suiwer objektief of subjektief kan wees nie, staan teenoor bogenoemde meer objektiewe opvattinge die subjektiewe skool waarvan Bruno de Finetti van Triest 'n verteenwoordiger is¹⁰). Volgens hom is die waarskynlikheid van 'n gebeurtenis E met betrekking tot 'n bepaalde proefpersoon die verhouding $\frac{k}{k+1}$, wanneer die persoon bereid is om k teen een te wed dat E sal plaasvind. Die waarde van die waarskynlikheid sal dus baie daarvan afhang of die persoon 'n optimis of 'n pessimis is. Dit sou beter wees om 'n meer stabieledefinisie van waarskynlikheid te gebruik, tesame met Frinette se formule, om die swartgalligheid van die proefpersoon te meet. D. van Dantzig sê dat dit hom laat dink aan 'n arts wat aan sy pasiënt vra hoe warm

8. Laplace, Pierre S.: *Theorie Analitique des Probabilites*, 1814.

9. Von Mises, R.: *Wahrscheinlichkeitsrechnung*. 1931.

10. Aangehaal deur D. van Dantzig in referensie II.

hy voel en dan die antwoord gebruik, nie alleen om die pasiënt se koorstoestand vas te stel nie, maar ook om sy koorspen te yk¹¹). Teenoor die objektiewe rigting staan ook die logistiese skool wat die formele metodes van Logika op die waarskynlikheidsrekening toepas. Keynes¹²) en Reichenbach¹³) het interessante teorië ontwerp en waardevolle bydraes gelewer, maar wat uit 'n matematiese oogpunt nie as geslaagd beskou kan word nie. Volgens hierdie opvatting bestaan daar 'n bepaalde waarskynlikheid vir 'n gebeurtenis, en die onsekerheid word aan die bewering omtrent die waarskynlikheid toegeskryf. Waarnemings dui egter daarop dat dit nie so eenvoudig aanvaar kan word nie.

Uit matematiese oogpunt het die volgende belangrike bydrae, na Von Mises, van Kolmogoroff¹⁴) gekom. Terwyl wiskundiges en statistici dit almal eens was omtrent die grondbewerkinge van die waarskynlikheidsrekening maar nie kon saamstem omtrent die in- en uitskakeling van die model nie, het Kolmogoroff deur van versamelingsfunksies gebruik te maak die matematiese formalisme tot 'n pragtige teorie ontwikkel wat aan moderne wiskundige standaarde voldoen. Oor die in- en uitskakeling is nog geen eenstemmigheid bereik nie.

Die frekwensieteorie, wat nou volg, het die meeste aanhangers en is myns insiens tot nou toe die geskikste vir 'n statistiese teorie, wat met stogastiese waarnemings skakel. Die frekwensiekwosiënt f/n , d.i. die verhouding van die aantal kere wat 'n kenmerk te voorskyn tree tot die totale aantal waarnemings word statisties stabiel met groterwordende aantal waarnemings; dit is die „wet van groot getalle”. Op grond van hierdie regulariteit word in die matematiese model 'n getal w met die gebeurtenis geassosieer en die waarskynlikheid van die gebeurtenis genoem. As grondslag vir die bewerkinge met w word die eenskappe van frekwensiekwosiënte aksiomaties aanvaar. Die begrip waarskynlikheid tree dus na vore as 'n matematiese model en dit word dikwels uit die oog verloor — wat vir die model geld, geld nie noodwendig vir enige praktiese geval nie — hoe goed die teorie in 'n bepaalde geval toegepas kan word, word bepaal deur die geskiktheid van die model in die betrokke geval.

Verder moet onderskei word tussen die betekenis van waarskynlikheid in die aldus gedefinieerde vorm en enige ander betekenis wat die woord in die alledaagse gebruik mag besit.

Die wet van groot getalle is lank reeds waargeneem en verskyn sover vasgestel kan word die eerste keer gedurende 1741 in 'n boek van J. Süssmilch, wat dit as volg stel: „Hoe groter die somme aangroei, hoe nader kom hul aan die ware wette van die natuur en hoe meer word die ongereëldhede in klein getalle as't ware verslind.” Die „nadering” vind nie plaas in die gewone

11. Van Dantzig, D.: Waarschijnlijkheidsrekening. Mathematische Centrum. Amsterdam. 1946.
12. Keynes, J. M.: Treatise on Probability. 1931.
13. Reichenbach, H.: Wahrscheinlichkeitslehre. 1935.
14. Kolmogoroff, A.: Grundbegriffe der Wahrscheinlichkeitsrechnung. Berlyn. 1933.

matematiese sin, soos Von Mises dit wou sien nie. Daar kan uit statistiese oogpunt nie aanvaar word dat na 'n groot genoeg aantal waarnemings die verskil tussen die frekwensiekwosent f/n en w met sekerheid kleiner as 'n gegewe getal is nie, d.w.s. daar bestaan 'n persentasie gevalle, hoe klein ook al, wanneer weens die stogastiese aard van die waarnemingsmateriaal, oorskreiding van die vasgestelde grens verwag kan word.

Dit is 'n grondprobleem met eie statistiese karakter ten opsigte van die aksiomatiese benadering.

Daar kan weliswaar 'n teorie van statistiek ontwikkel word sonder die inskakeling van waarskynlikheid — dit sou dan grotendeels uit frekwensierekening bestaan het — maar of dit wesentlik van waarskynlikheidsrekening sou verskil, is te betwyfel.

Opponente van die frekwensie-teorie voer aan dat sommige gebeurtenisse glad nie of slegs op verwronge wyse as elemente van 'n herhaalbare reeks eksperimente gesien kan word, m.a.w. dat die frekwensiebasis van waarskynlikheid nie standhou nie. Voorbeeld: Wat is die betekenis van „waarskynlikheid” in die volgende vraag: „Wat is die waarskynlikheid dat die Libanese opstandelinge die stryd gaan wen?” Aan die beswaar moet toegegee word maar die probleem val myns insiens buite die terrein van Matematiese Statistiek. Wat ook al die betekenis van die woord waarskynlikheid in die voorbeeld is, dit is nie 'n statistiese waarskynlikheid wat geassosieer kan word met die regulariteit van frekwensiekwosiënte van stogastiese waarnemings nie.

TOEPASSINGS.

Vervolgens veroorloof ek my enkele gedagtes oor die ietwat ongewone aspek by 'n geleentheid soos hierdie, nl. toepassing van Statistiek.

Statistiek as die wetenskap wat numeriese waarnemings hanteer en in oorsigtelike vorm rangskik en opsom, is roetine in enige goed georganiseerde hedendaagse onderneming en vereis nie matematiese statistiek van 'n hoë kaliber nie. Dit is onteenseglik 'n belangrike aspek, maar die opvatting dat dit al is, is 40 jaar outyds. Met toepassing word dan eintlik bedoel die bydrae wat die statistikus vanweë sy teoretiese opleiding en bepaalde denkwyse kan lewer. Die metodes wat die afgelope 30 tot 40 jaar ontwikkel is, word toegepas om met behulp van steekproewe gevolgtrekkinge en uitsprake te maak omtrent karakteristieke van die populasies waaruit die steekproef kom. Uitsprake word op 'n bepaalde betroubaarheidsdrempel gegee. Wanneer bv. 'n hipotese getoets word en die aannames van die hipotese lei tot 'n gebeurtenis — in die vorm van die steekproef — wat slegs 1 uit 1000 keer vir so 'n stogastiese veranderlike sou plaasvind, sal die hipotese verwerp word op 'n bepaalde betroubaarheidsdrempel. Die uitspraak is dus dat die verkreë gebeurtenis nie toegeskryf kan word aan stogastiese variasie nie, maar aan 'n verkeerde hipotese. Ooreenkomstige karakteristieke van verskillende steekproewe kan vergelyk word en vasgestel word of die verskil duidend is, d.w.s. of die verskil aan stogastiese

variasie toegeskryf kan word en of dit dui op 'n werklike verskil. Die waarde van uitsprake van hierdie aard op aktuariële gebied, in verband met opnames en ondersoeke op sosiologiese, ekonomiese, opvoedkundige en natuurwetenskaplike terreine is voor die hand liggend.

In natuurwetenskaplike eksperimente kan dikwels ook nog kontrole uitgeoefen word en in mediese, landboukundige en biologiese eksperimente behandelinge vergelyk word.

Uitsprake kan gemaak word omtrent enkelvoudige en meervoudige korrelasie. Die teorie van skatting kan aangewend word om bloot die waardes van sekere karakteristieke te bepaal en so doeltreffend moontlik te bepaal.

In groot ondernemings is daar sommige persone wat iets van alles, andere wat alles van iets maar selfde iemand wat alles van alles weet. Om 'n oorsig van 'n bepaalde aspek te kry is moeilik. Die statistikus, met sy teoretiese opleiding en steekproeftegniek, kan hulp verleen deur verskynsels vas te stel, te ontleed en die aandag daarop te vestig. Baie mense reken dat 'n statistikus se geluk eweredig is met die aantal syfers wat hy tot sy beskikking het; dit is verkeerd, dit is net andersom, sy geluk is omgekeerd eweredig met die aantal syfers tot sy beskikking — hy word opgelei en voel gelukkig as hy meer inligting met minder waarnemings kan kry.

Daar was die geval van 'n groot onderneming wat 'n personeel van 100 gebruik het om by 'n vervaardigingsproses defektiewe artikels uit te soek. 'n Statistiese ondersoek het aan die lig gebring dat selfs die mees ervare van hierdie werkers vir hierdie doel geen onderskeidingsvermoë besit het nie. In 'n ander geval het 'n fabriek die eerste jaar na statistiese kwaliteitskontrole ingevoer was een miljoen dollar bespaar. Uit statistiese oogpunt is dit 'n elementêre prosedure wat, in beeld, die aandag vestig op 'n resultaat wat nie aan stogastiese variasie toegeskryf kan word nie.

In die meeste gevalle is die uitsprake nie so skouspelagtig as in die genoemde gevalle nie, en wat nog swaarder weeg, nie so populêr nie. 'n Statistikus is nie 'n towenaar wat subjektiewe verwagtinge uit 'n stel waarnemings kan presipiteer nie.

'n Minder bekende toepassingsrigting is funksioneringsnavorsing, wat gedurende die tweede wêreldoorlog ontwikkel het toe 'n aantal wetenskaplikes afgesonder was met die oog op sg. „Operational Research”.

Uit waarnemings van verkenningsvlugte is bv. 'n verandering in die vyand se taktiek met behulp van hierdie metodes doeltreffend vasgestel, en teenmaatreëls kon onmiddellik toegepas word; die effek van die teenmaatreëls kan dan waargeneem word, en die siklus herhaal hom weer. 'n Studie van die grootte en samestelling van konvoie het 'n beslissende invloed op die verloop van die see-oorlog in die Atlantiese Oseaan gehad. Die aard van die waarnemings vereis meesal statistiese metodes.

Funksioneringsnavorsing betree nou 'n baie breër terrein as die militêre. Dit word beskou as 'n wetenskaplike metode wat 'n kwantitatiewe basis skep

vir die gebruik van persone of liggame met die oog op beslissinge i.v.m. die kontrole van prosesse onder hulle beheer.

'n Vertakking wat die afgelope aantal jare heelwat aandag geniet is lineêre programmering. Basies bestaan dit uit die bepaling van 'n optimale waarde van 'n lineêre vorm onderhewig aan lineêre dwangvoorwaardes. Die metode kan hulp verleen wanneer die optimale waarde bv. die minimumkoste in 'n vervoerprobleem, of maksimum-wins in 'n produksieproses is. Die optimale waarde kan nie met die gewone wiskundige metodes bepaal word nie.

Stogastiese prosesse is die naam wat gebruik word vir prosesse wat of met behulp van suiwer kansrekening bestudeer kan word of wat 'n statistiese struktuur het. Tyd tree dikwels, dog nie altyd nie, as parameter op en na hierdie afdeling kan verwys word as die Dinamika van Statistiek. Probleme is meesal gerig op voorspellingsuitsprake.

Van die groot hoeveelheid onsin wat omtrent prediksie verkondig is, het statistici ongetwyfeld 'n groot deel bygedra. Wanneer 'n tydreeks reëlmatigheid vertoon haas die Statistikus en andere om matematiese funksies daarop te pas, wat nie onverstandig is nie, maar om deur ekstrapolasie op grond van hierdie funksies voorspelling te waag is dikwels rampspoedig.

Ter versagting kan aangevoer word dat die matematikus net sistematies doen wat andere intuïtief sou doen. Dit was die toedra gvan sake tot ongeveer 1930, toe dit uiteindelik besef is dat hierdie prosesse 'n sterk stogastiese element besit. In baie gevalle is die probleme nog onopgelos, in sommige gevalle werp die teorie lig op die aard van die probleem en in ander gevalle word die wisselvalligheid van 'n uitkoms aangetoon. Die volgende is enkele terreine waarop die teorie toegepas word: Versekeringsrisiko, groei en mutasie by bakterieë en waggprobleme soos verkeers- en ander ophopinge en toustaanverskynsels.

As die groot verskeidenheid van toepassingsterreine in oënskou geneem word, behoort besef te word dat die grense van resepsstatistiek beperk is; dit is dikwels minder doeltreffend en moet in elk geval gediagnoseer kan word. Hedendaagse en toekomstige ontwikkeling is en sal tot 'n groot mate spesialisering wees. Dit geld ook vir die toepassingsterreine. Dikwels is daar meer as een metode waaruit gekies moet word of soms moet 'n metode aangepas word.

Ons kom tot die gevolgtrekking dat alhoewel Statistiek 'n eie konseptuele sfeer van abstrakte objekte besit, dit op toepassingsterrein nie die studie van eie, bepaalde fisiese objekte ten doel het nie, behalwe in soverre dat waarnemings sekere kenmerke moet besit. Dit en die veelvuldigheid van sy toepassingsterreine en die diversiteit van sy toepassinge besorg hom 'n eie karakter in vergelyking met meeste ander wetenskappe.

Onderrig in Statistiek aan 'n universiteit moet eerstens daarop gerig wees om teorie van statistiek te leer en te bevorder en tweedens om instrumente beskikbaar te stel vir empiriese navorsing op ander vakgebiede.

Hooggeagte Here Lede van die Raad van die P.U. vir C.H.O., van hierdie geleentheid maak ek graag gebruik om u te bedank vir die eer wat u my aangedoen het om my te benoem tot professor in Statistiek. Dit is 'n voorreg en dit stem my tot groot dankbaarheid dat God in Sy genade beskik het dat hierdie verantwoordelike amp aan my opgedra is, en dat dit juis aan die P.U. vir Christelike Hoër Onderwys is, is voorwaar 'n voorreg wat sal aanspoor tot ywer en toewyding.

Hooggeleerde prof. D. J. van Rooy,

Ontvang hiermee my besondere en opregte waardering vir u prinsipiële leiding, u navolgingswaardige voorbeeld as leermeester en vakwetenskaplike. Dit is 'n voorreg om met u saam te werk.

Hooggeleerde Professore, Dames en Here Lektore,

Aan u almal betuig ek my hartlike dank vir wat u vir my beteken het en nog steeds beteken, sommige as leermeesters, sommige as leiers, sommige as mededosente in die Departement Wiskunde, Toegepaste Wiskunde en Statistiek — aan u almal vir u kollegiale samewerking. Ons beywer ons om die wetenskap op ons verskillende vakterreine te bevorder. Dit is my vaste oortuiging dat ons hiermee ook nasiediens te verrig het — alles tot eer van God. Die ontwikkeling van Suid-Afrika betree 'n belangrike fase — 'n fase wat 'n verantwoordelike taak op die skouers van wetenskaplikes aan ons Suid-Afrikaanse Universiteite plaas. Mag dit ons erns wees.

Hooggeagte Dames en Here Studente en Dames en Here, ek dank u vir u teenwoordigheid en u welwillende aandag.

Algemene referensies:

Cramèr, H.: *Mathematical Methods of Statistics*. 1954.

Kendall, M. G.: *Modern Statistics in Business and Commerce*. The Journal of the Institute of Actuaries. 1956.

Van Dantzig, D.: Referensies 5 en 11 hierbo.