MATHEMATISCH CENTRUM
AMSTERDAM

Enige statistische annotaties betreffende de lengten van vogelsnavels.

1. Gegeven is een serie van "snavels van de volgende lengten (in halve mm):
   69, 71, 72, 73, 74, 75, 76, 77, 78, 79, 84,
   alle met de frequentie 1. Gevraagd is of de snavel van 84 halve mm tot dezelfde populatie behoort als de overige snavels.
2. Gewoonlijk past men aan een dergelijke empirische verdeling een theoretische verdeling aan, bijv. een normale of een binomiale of een Poisson-verdeling. De hier gegeven verdeling telt echter zo weinig waarnemingen, dat het moeilijk wordt, van een dergelijke handelwijze gebruik te maken. De normale verdeling, die de opdrachtgever gebruikt heeft, is niet zeer geschikt, want de serie vertoont niet in het minst de bekende klokvorm. Maar zelfs, al zou men uit andere hoofde weten, dat de serie beschouwd moest worden als een steekproef uit een normale verdeling (bijv. omdat regelmatig gevonden is, dat de snavellengten van vogels van die soort normaal verdeeld zijn), dan nog zou men uit deze steekproef weinig kunnen afleiden omtrent het gemiddelde en de spreiding van de gehele populatie. Immers, wanneer men de overschrijdingskansen wenst te berekenen, moet men in principe uitgaan van gemiddelde en spreiding van de gehele populatie en niet van deze steekproef; nu doet men dit laatste in feite toch, zodat de uitkomsten alleen maar approximatief geldig zijn. Maar deze benadering is slechts redelijk als de steekproef ongeveer dezelfde (i.c. klokvormige) frequentiecurve geeft als de gehele populatie. Wanneer men dus aanneemt, dat de serie een steekproef is uit een normaal verdeelde collectie - zoals de opdrachtgever gedaan heeft - dan is juist deze steekproef ongeschikt om gemiddelde en spreiding een daarmee overschrijdingskansen te bepalen.
3. Niet alleen de normale verdeling is ongeschikt om deze serie te beschrijven, maar eigenlijk elke bepaalde "standaardverdeling". De oorzaak hiervoor ligt in het zeer geringe aantal waarnemingen, hetgeen maakt, dat men van geen enkele standaardverdeling kan zeggen, dat zij beter bij de empirische verdeling aansluit dan andere verdelingen. Nu is er wel een methode, die gebruikt kan worden onafhankelijk van de aard van de verdeling; maar deze methode is zo onscherp, dat zij, toegepast op het hier beschouwde materiaal, geen succes opleverde. Wij geloven dat het op statistische gronden niet mogelijk is, uit te maken of de snavel van 84 halve mm. tot de populatie behoort. Men kan

dit - helaas negatieve - standpunt als volgt illustreren: beschouw 11 mensen in een tram en bepaal hun gewicht; één van hen is bijzonder zwaar. Kunnen we nu zeggen, dat hij tot een andere populatie behoort? Hèt is mogelijk, maar de gegevens zijn veel te gering.

4. Tenslotte de beantwoording van een aantal vragen, door de opdrachtgever gesteld:

<u>Ad (1)</u> Men moet uitgaan van de hypothese, dat de snavel 84 tot de populatie behoort. Wanneer men nu op grond van deze hypothese vindt, dat datgene wat men in werkelijkheid aantreft (nl. de gevonden serie vogelsnavels) een zeer ~~xxx~~ onwaarschijnlijke gebeurtenis is, verwerpt men de hypothese. Op grond van waarschijnlijkheidsoverwegingen kan de statistiek er toe leiden een hypothese te <u>verwerpen</u>; het <u>aanvaarden</u> ener hypothese kan alleen in de negatieve zin geschieden van: het statistisch onderzoek geeft geen aanleiding de hypothese te verwerpen.

<u>Ad (2)</u> Een vollediger tabel wordt U hierbij als bijlage toegezonden.

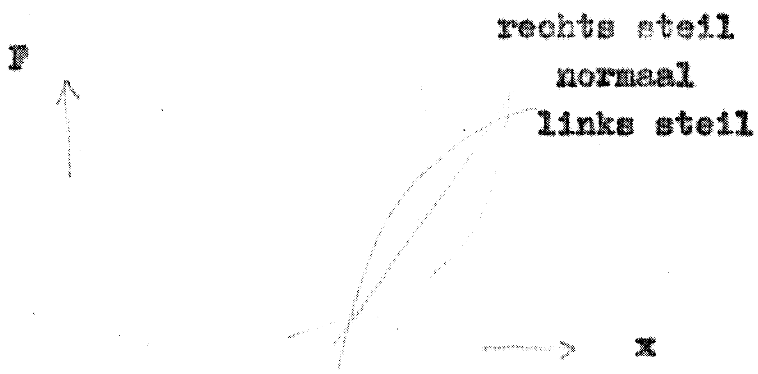
<u>Ad (3)</u> Het aangeven van grenzen, binnen welke men variaties aanvaardt en buiten welke men ze verwerpt, is tamelijk willekeurig. Soms werkt men met grenzen, die behoren bij een overschrijdingskans 0.01, soms 0.05; ook wel werkt men met de 3σ - of 2σ -grenzen (dan zijn de overschrijdingskansen bij een <u>normale</u> verdeling 0.00135 resp. 0.02275). Wat hier in feite gekozen wordt, hangt af van de eisen van zekerheid die de onderzoeker in kwestie stelt. Een keuze van b.v. 0.05 houdt in dat men bereid is 5% foutieve beslissingen te aanvaarden.

<u>Ad (4) en (5)</u>. Wanneer men aan een gegeven empirische verdeling een normale verdeling wil aanpassen, kan men als volgt te werk gaan. Men bepaalt van de empirische verdeling het gemiddelde (μ) en de standaarddeviatie (σ). De normale verdeling heeft dan algebraisch deze vorm

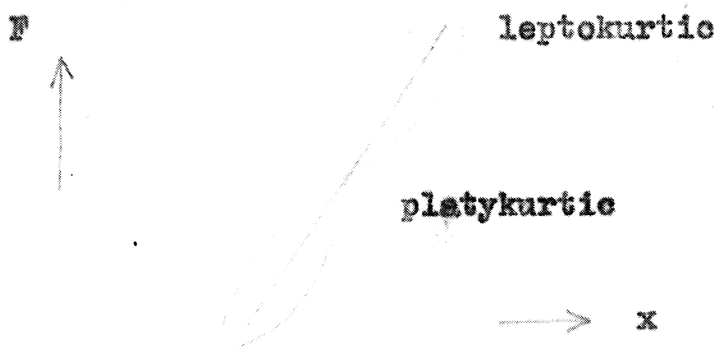
$$y = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

waarbij π = 3.14159... en e = 2.71828... Men vult in deze formule voor x verschillende waarden in (zoals in het geval van de snavellengten: 68, 69, 70,...., 82, 83, 84, 85...) en vindt aldus evenzovele waarden voor y. Dit al zet men af in een grafiek, waarin tevens de empirische verdeling wordt afgezet. Komen beide curven nu min of meer samen te vallen, dan is de veronderstelling, dat de empirische verdeling afkomstig is van een steekproef uit een normaal verdeelde collectie, houdbaar te noemen. Daarbij moet men rekening houden met het feit, dat bij een gering aantal waarnemingen grotere onregelmatigheden getolereerd kunnen worden dan bij een groter aantal. Een exacter wijze van toetsing is de χ^2-test, die o.a. beschreven wordt in Yule and Kendall, Introduction to the theory of statistics, chapter 22.

Een andere, vaak aan te bevelen methode, maakt gebruik van zgn. waarschijnlijkheidspapier. De schaalverdeling van dit papier is zodanig, dat de gecumuleerde normale verdelingskromme als een rechte lijn verschijnt. Een steekproef van een normaal verdeelde populatie zal in het algemeen geen volkomen rechte lijn geven, maar tengevolge van de steekproefonregelmatigheden, een serie van punten niet ver van deze rechte. Als de verdeling asymmetrisch is, vindt men in plaats van een rechte lijn een holle of bolle lijn. Is de waarschijnlijkheidsverdeling links steil en rechts minder steil, dan vindt men een bolle lijn, in het andere geval een holle lijn:



Is de verdeling zeer plat (platykurtic) óf heeft zij een zeer geprononceerde steile top (leptokurtic), dan vindt men een golvende lijn:



Als er onzekerheid bestaat omtrent de waarde van de hoogste en de laagste waarnemingen, kan men volstaan met aanpassing aan het middenstuk, zeker als dit behoorlijk recht is. Noodzakelijk is echter, dat men met redelijk grote aantallen te doen heeft.

Ad (6). De gegeven getallenreeks is inderdaad te klein om tot een normale verdeling te besluiten; bovendien wijkt zij zeer sterk van een normale verdeling af. De conclusie, dat de wiskunde dan misschien in het geheel geen hulp kan bieden, is niet volkomen juist; zij kan bijv. ook van andere pogingen tot beredeneren van een vermoeden de onvoldoende betrouwbaarheid aantonen. De mogelijkheid scheppen, uit onvoldoende waarnemingsmateriaal betrouwbare conclusies te trekken, kan de wiskunde uiteraard evenmin als enige andere wetenschappelijke methode.