Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

MC SYLLABUS 1.3

J. HEMELRIJK
J. KRIENS

LEERGANG BESLISKUNDE

DEEL 3

STATISTIEK

4e DRUK

MATHEMATISCH CENTRUM AMSTERDAM 1976

AMS(MOS) subject classification scheme (1970): 62-01

ISBN 90 6196 016 9

1e druk 1966
2e verbeterde druk 1969
3e ongew. druk 1972
4e ongew. druk 1976

Inhoud

	blz.
Voorwoord	1
1. Grafisch onderzoek naar de aard van een verdeling	3
2. Inleiding tot de schattingstheorie	16
3. Nauwkeurigheid van schatters; betrouwbaarheidsintervallen	23
4. Meest aannemelijke schatters	29
5. Kleinste kwadraten; regressie	32
6. De momentenmethode	38
7. Inleiding tot de toetsingstheorie	40
8. Grondbegrippen van de toetsingstheorie	43
9. Binomiale toets met normale benadering	55
10. Chikwadraat-toets voor aanpassing	58
11. Verband tussen toets en betrouwbaarheidsinterval	64
12. Verdelingsvrije toetsen	65
13. Hoogwaterstanden, exponentiële verdeling	73
14. Enige statistische opgaven	80
Literatuur over statistiek	94

oorwoord

Voortbouwend op de wiskundige basiskennis (deel 1) en vooral op de ansrekening (deel 2) bespreken wij in dit derde deel van de Leergang Beliskunde de belangrijkste statistische begrippen en technieken. Stonden n deel 2 theoretische begrippen zoals kans, stochastische grootheid, moent en limietstelling centraal, in dit deel gaan wij bij voortduring uit an waarnemingen of metingen.

Deze kunnen dienen voor het grafisch onderzoek naar de aard van een erdeling, voor het schatten van onbekende parameters of voor het toetsen an hypothesen. Onze behandeling van schattings- en toetsingstheorie is n de eerste plaats gericht op het bijbrengen van inzicht in de begrippen n in de gedachtengang. De talrijke voorbeelden zijn vooral met dit doel ogenomen; dit boekje is geen uitvoerige opsomming van recepten voor allerei voorkomende schattings- en toetsingsproblemen. Naar enige boeken die el dit karakter hebben, wordt verwezen in de literatuurlijst aan het slot an dit deel.

Het boekje is in hoofdzaak een bewerking door Prof. J. KRIENS van e hoofdstukken 12 en 13 van rapport S 200 (C 8) van de Statistische Afdeing van het Mathematisch Centrum, getiteld: Syllabus van een oriënterende ursus Mathematische Statistiek, door Prof.Dr. J. HEMELRIJK, (1956). Paraaaf 1 is een bewerking van een tekst van Mevrouw G. KLERK-GROBBEN. De agrafen 7, 10, 11 en 12 werden herschreven door Drs. W. MOLENAAR, die k de eindredactie verzorgde. Enige voorbeelden werden ontleend aan rapport S 258 (C 12) van de Statistische Afdeling van het Mathematisch Centrum, titeld: Syllabus van de cursus "Enkele elementaire statistische methoden" or Dr. J. FABIUS, (1959).

sterdam, 1965.

.. Grafisch onderzoek naar de aard van een verdeling.

Het komt voor dat men bepaalde kenmerken wenst te bestuderen bij ·en groot aantal elementen (dit kunnen personen, produkten, proef-·emingen of nog andere zaken zijn). In vele gevallen is het te duur, ·f zelfs onmogelijk, alle elementen te onderzoeken. Men beperkt het ·nderzoek dan tot de elementen van een steekproef, d.w.z. een aselect ·it de populatie van alle elementen gekozen groep (zie deel 2, o.a. ·lz. 34-35). Het is de taak van de statisticus op grond van de waar-·emingen aan deze steekproef uitspraken te doen over de populatie. Deze ·itspraken zullen niet met zekerheid algemeen geldig kunnen zijn, omdat e steekproef nu eenmaal minder informatie verschaft dan de populatie. n het algemeen zullen wij zien dat steeds stelliger uitspraken mogelijk ·orden naarmate de steekproefomvang toeneemt.

oorbeeld 1.1

Voor het beantwoorden van de vraag, welke fractie der rijks-mbtenaren academicus is, zullen wij een steekproef uit de populatie an alle rijksambtenaren afzonderen en via de personeelsdiensten of oor ondervraging nagaan hoe hoog het percentage afgestudeerden in de teekproef is. Wij moeten dan eerst een paar lastige vragen beantwoorden, oals: geeft aselecte trekking hier een representatieve steekproef?, elke studies zijn precies universitair?, hoeveel ambtenaren moeten ndervraagd worden? Op dit moment laten wij deze technische kwesties erzijde.

Als men aan elk element van de steekproef een meting of waarneming erricht, kan men in het wiskundig model het te meten of waar te nemen enmerk bij een aselect uit de populatie getrokken element beschouwen ls een stochastische grootheid $\underline{x}$, die voor elk element van de populatie en vaste getalwaarde x aanneemt. Zo kunnen we in voorbeeld 1.1 af-preken: $\underline{x}$ = 1 resp. 0 als de ambtenaar wel, resp. niet academicus is. aren we niet geïnteresseerd in zijn titel maar in de oppervlakte van

zijn bureau, dan zou $\underline{x}$ allerlei niet-negatieve waarden (het aantal cm^2) kunnen hebben. Trekt men aselect zonder teruglegging n elementen en is n klein t.o.v. de populatie-omvang, dan is het redelijk de waarnemingen aan deze n elementen te zien als n stochastische grootheden die allen dezelfde verdelingsfunctie hebben en stochastisch onafhankelijk zijn. De n waarnemingen die we zullen doen als we aselect gaan trekken, worden dus als stochastische grootheden $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ beschouwd; de getalwaarden die ontstaan als we inderdaad trekken en waarnemen geven we met $x_1, x_2, \ldots, x_n$ aan. Het is van belang dat elk getal x_i een realisering is van de stochastische grootheid $\underline{x}_i$; m.a.w. dat de getalwaarde anders had kunnen uitvallen als we een ander element ter waarneming hadden getrokken.

Bij het opstellen van de uitspraak die we tenslotte over de populatie zullen doen speelt de gemeenschappelijke verdelingsfunctie van $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ een grote rol. In deze paragraaf zullen wij enige grafische methoden bespreken die ons een globale indruk geven of de waargenomen waarden $x_1, x_2, \ldots, x_n$ bij een bepaald type verdeling (bijv. normaal of exponentieel) passen. Als dit het geval is kan men vervolgens de nog onbekende parameter(s) van de verdeling schatten, of bepaalde hypothesen omtrent die parameter(s) toetsen. Deze begrippen worden in de volgende paragrafen uitvoerig besproken.

Voorbeeld 1.2. Levensduur van een serieprodukt.

Eén der belangrijkste eigenschappen van veel produkten, zoals gloeilampen, onderdelen van machines en apparaten, is de levensduur. Het is dan ook van groot belang te beschikken over een methode, waarmee een goed inzicht in deze eigenschap verkregen kan worden. In dit voorbeeld zullen wij de voorgaande beschouwingen toelichten aan de hand van gloeilampen.

De levensduur van een gloeilamp kan men bepalen door hem te laten branden tot hij doorbrandt. Uiteraard is deze methode ongeschikt wanneer men een nog in de handel te brengen partij wil onderzoeken en daarom is men in dit geval gedwongen de eigenschappen van een gegeven soort gloeilampen met behulp van steekproeven te onderzoeken.

Wij beschouwen de levensduur als een stochastische variabele $\underline{x}$, die voor iedere lamp een bepaalde waarde aanneemt. De bij een steekproef van lampen gevonden waarden vormen dan een reeks van stochastisch onafhankelijke waarnemingen van $\underline{x}$. Veelal wordt bij dergelijke problemen verondersteld dat $\underline{x}$ een (bij benadering) normale verdeling bezit. Soms geldt dit niet voor $\underline{x}$ zelf maar wel voor log $\underline{x}$. Speciaal voor levensduren veronderstelt men ook wel een exponentiële verdeling.

De vraag is nu hoe men in zo'n geval kan nagaan, welke verdeling het meest in aanmerking komt. In sommige gevallen kan men op theoretische gronden tot een bepaalde vorm besluiten. Vaker echter zal men de aard van de kansverdeling uit de waarnemingen moeten afleiden. Gewoonlijk is dit een kwestie van proberen, waarbij de volgende richtlijnen van nut kunnen zijn.

Suggereren de waarnemingen een symmetrische verdeling, dan kan men een normale verdeling veronderstellen. Het gebruik van waarschijnlijkheidspapier (hieronder behandeld) is hierbij een vlugge grafische methode om de normaliteit van de waarnemingen te beoordelen.

Bij scheve verdelingen kan men (bijv. met logarithmisch waarschijnlijkheidspapier) nagaan of misschien log $\underline{x}$ normaal verdeeld is. Ook de gamma-verdeling komt hier in aanmerking. Als de frequentie der waarnemingen monotoon dalend is, denkt men aan een exponentiële verdeling en als de waargenomen grootheid duidelijk aan twee kanten begrensd is, zou men een bêta-verdeling kunnen proberen. Deze verdelingen zijn besproken in deel 2, §7 (vgl. blz. 46 e.v.).

Waarschijnlijkheidspapier

We beschouwen eerst een stochastische grootheid $\underline{u}$ met de gestandaardiseerde normale verdeling en geven de verdelingsfunctie aan met $\Phi(u)$; dus

$$\Phi(u) = P\left[\underline{u} \leq u\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{1}{2}v^2} \, dv \,. \tag{1.1}$$

De functie $\Phi(u)$ op gewoon millimeterpapier tegen u uitgezet heeft de gedaante die in figuur 1.1 geschetst is.

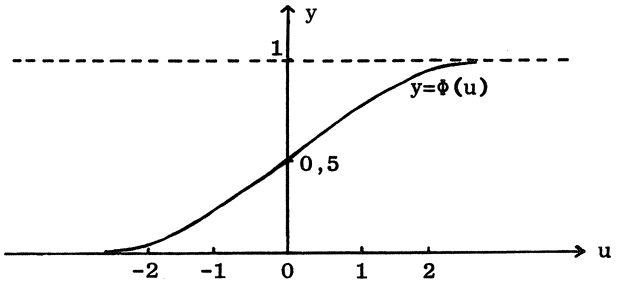


fig. 1.1

Gestandaardiseerde normale verdelingsfunctie op mm-papier

Door transformatie van de y-schaal in een z-schaal met $y = \Phi(z)$ gaat deze kromme over in de rechte lijn $z = u$ (zie figuur 1.2).

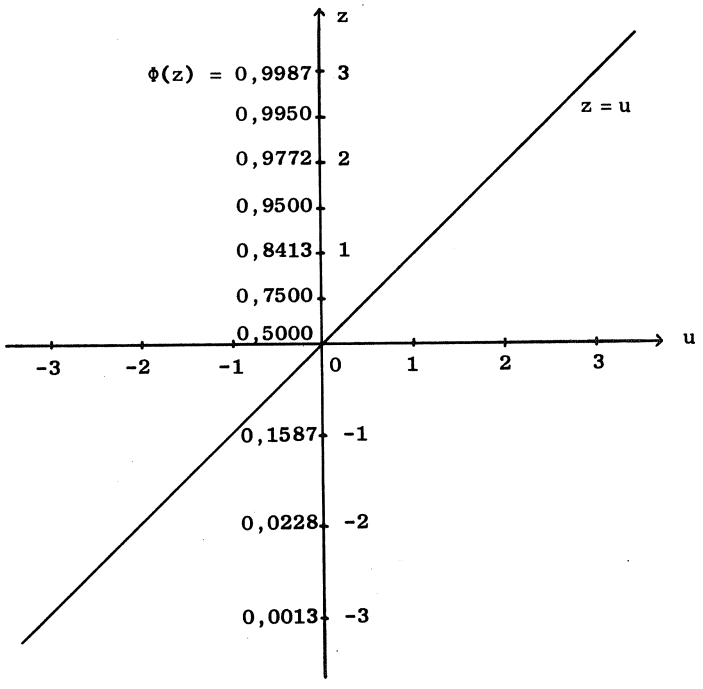


fig. 1.2

Gestandaardiseerde normale verdelingsfunctie
op waarschijnlijkheidspapier

Bij $y = 0{,}5$ behoort $z = 0$; $y = 1$ gaat over in $z = \infty$ en $y = 0$ in $z = -\infty$. De staarten van de verdelingsfunctie komen nu dus niet meer op het papier. Op het waarschijnlijkheidspapier dat in de handel verkrijgbaar is, is deze transformatie reeds uitgevoerd. Langs de z-schaal zijn hierop de bijbehorende waarden $y = \Phi(z)$ bijgeschreven. De y-waarden lopen gewoonlijk van 0,0002 tot 0,9998.

Als $\underline{x}$ $N(\mu,\sigma)$-verdeeld is, voldoet zijn verdelingsfunctie $F(x)$, zoals we in stelling 7.1 van deel 2 hebben gezien, aan $\Phi(u) = F(\mu+u\sigma)$; dus

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) , \tag{1.2}$$

waarin $\Phi(u)$ weer de verdelingsfunctie van de $N(0,1)$-verdeelde $\underline{u}$ voorstelt. $y = F(x)$ tegen x uitgezet heeft een vorm als in figuur 1.1, met dit verschil dat nu $y = 0{,}5$ bij $x = \mu$ optreedt en de horizontale schaal met σ vermenigvuldigd is t.o.v. dit punt. Toepassing van de transformatie $y = \Phi(z)$ maakt deze kromme weer tot een rechte lijn: $z = \frac{x-\mu}{\sigma}$, waarvan de plaats en de helling bepaald zijn door μ en σ (figuur 1.3); voor $x = \mu$ wordt $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ gelijk aan $\Phi(0) = 0{,}5$ en voor $x = \mu+\sigma$ wordt dit $\Phi(1) = 0{,}8413$.

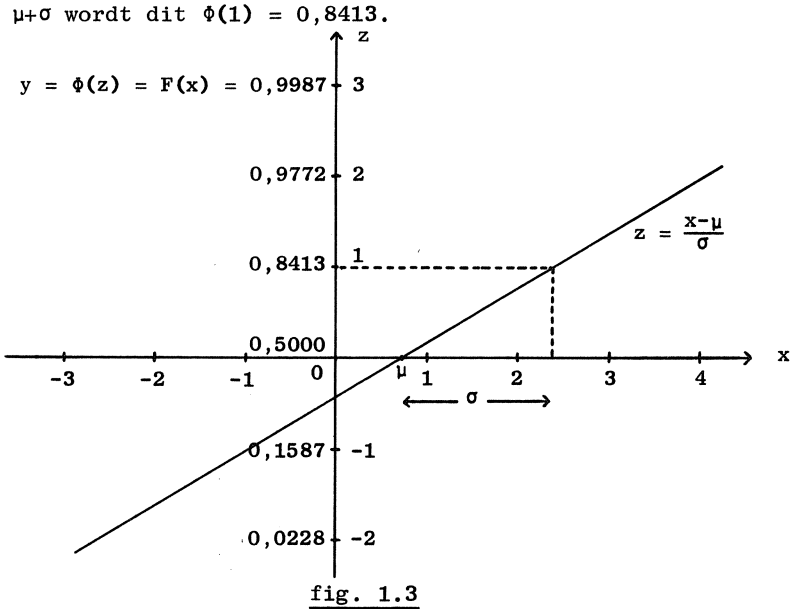


fig. 1.3

De verdelingsfunctie $F(x)$ van de $N(\mu,\sigma)$-verdeling op waarschijnlijkheidspapier, voor $\mu = 0{,}75$ en $\sigma = 1{,}6$

Meestal zijn μ en σ en dus ook $\Phi(\frac{x-\mu}{\sigma})$ onbekend; om F(x) tegen x uit te zetten zal men $F(x) = \Phi(\frac{x-\mu}{\sigma})$ moeten schatten uit de waarnemingen $x_1, x_2, \ldots, x_n$ van $\underline{x}$. Hiertoe worden de waarnemingen naar opklimmende grootte gerangschikt; wij stellen de geordende waarnemingen voor door $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$. Men kan dan de onbekende waarde $F(x_{(i)})$ schatten door $\frac{i}{n}$. Dit heeft het bezwaar dat $F(x_{(n)})$ door $\frac{n}{n} = 1$ geschat wordt, dus door $z = \infty$ (wegens $\Phi(\infty) = 1$) en buiten het papier valt. Dit is een slechte schatting, want bij de rechte lijn van de echte F(x) treedt $z = \infty$ pas bij $x = \infty$ op.

Voor een steekproef van omvang n van een stochastische grootheid met willekeurige verdelingsfunctie F kan men bewijzen, dat de i^{de} waarneming naar grootte-orde een stochastische grootheid $\underline{x}_{(i)}$ is, waarvoor geldt

$$\mathcal{E}\, F(\underline{x}_{(i)}) = \frac{i}{n+1} \qquad (i=1,2,\ldots,n) \; . \tag{1.3}$$

De schatting $\frac{i}{n+1}$ levert a.h.w. gemiddeld over alle denkbare steekproeven de juiste waarde op (is zuiver, met een in paragraaf 2 in te voeren term). Men kan echter bewijzen dat de lijn door de punten $(\underline{x}_{(i)}, \frac{i}{n+1})$ vaak (maar vrij weinig) steiler loopt dan de onbekende ware lijn en zelden (maar vrij veel) te vlak. Omdat een te steile lijn bij veel onderzoekingen bezwaren oplevert, geeft men dikwijls de voorkeur aan een lijn die de helft van de keren te steil uitvalt en de helft te vlak. BENARD en BOS-LEVENBACH *) vonden dat de lijn door $(\underline{x}_{(i)}, \frac{i-0,3}{n+0,4})$ vrijwel aan deze eis voldoet. Voor grote steekproeven zijn de drie genoemde quotiënten voor bijna alle i vrijwel gelijk, zodat de keuze dan niet veel ter zake doet.

Zijn $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ de naar opklimmende grootte gerangschikte waarnemingen en liggen de punten $(x_{(i)}, \frac{i-0,3}{n+0,4})$ op het waarschijnlijkheidspapier ongeveer op een rechte lijn ("ongeveer", want $\frac{i-0,3}{n+0,4}$ is slechts een schatting voor $F(x_{(i)})$), dan mag verondersteld worden dat de waarnemingen uit een normale verdeling afkomstig zijn.

---

*) A. BENARD en E.C. BOS-LEVENBACH, Het uitzetten van waarnemingen op waarschijnlijkheidspapier, Statistica Neerlandica 7,(1953), blz. 163-173.

Met de op het oog zo goed mogelijk aangepaste rechte lijn kan men
ngeveer de waarden van μ en σ vinden. Immers μ is de x-waarde die
olgens de ideale rechte lijn bij Φ(z) = 0,5 hoort en σ is het verschil
ussen de x-waarden bij Φ(z) = 0,84 en Φ(z) = 0,5. In paragraaf 2
ullen wij nauwkeuriger middelen bespreken om μ en σ te bepalen uit een
tel waarnemingen van een N(μ,σ)-verdeling.

Systematische afwijkingen van een rechte lijn zullen er op wijzen
at de verdeling van x niet normaal is. Zo ontstaat bij een verdeling
aarvan de staarten dunner zijn dan bij de aangepaste normale verdeling,
o het waarschijnlijkheidspapier een kromme zoals in figuur 1.4 geschetst

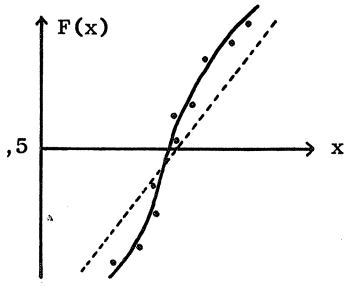


fig. 1.4
Staarten te dun

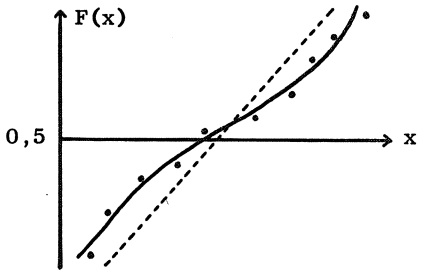
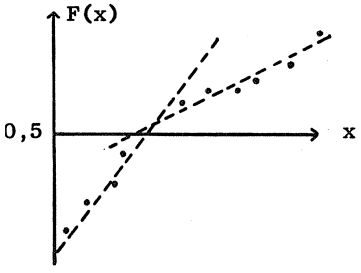


fig. 1.5
Staarten te dik



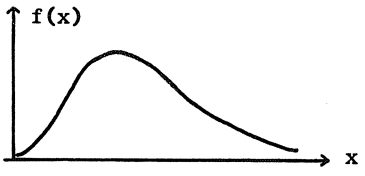
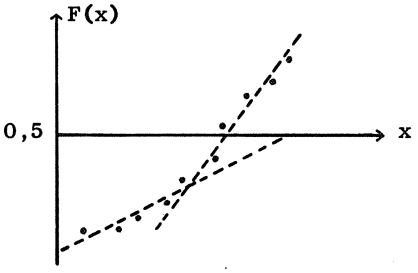


fig. 1.6
Verdeling scheef naar rechts



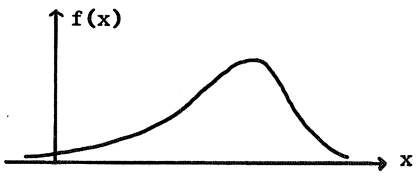


fig. 1.7
Verdeling scheef naar links

is. Zijn de staarten juist te dik, dan ontstaat figuur 1.5. Figuur 1.6 en 1.7 schetsen het resultaat bij scheve verdelingen, resp. voor een verdeling scheef naar rechts en voor een verdeling scheef naar links; zie de verdelingsdichtheden die erbij getekend zijn.

Gewoonlijk is de spreiding van de punten om de lijn zo groot, dat alleen sterke afwijkingen van een normale verdeling zichtbaar worden, vooral bij kleine steekproeven. De grafische methode is daarom alleen geschikt om snel een eerste indruk van de verdeling te verkrijgen.

Is de verdeling van $\underline{x}$ scheef naar rechts, dan kan geprobeerd worden of wellicht log $\underline{x}$ normaal verdeeld is, of algemener log $(\underline{x}-a)$ voor een geschikte constante a. Het verschuiven over een afstand a is nodig als niet alle x_i positief zijn, en wenselijk als de waarnemingen een duidelijk positieve ondergrens hebben. Met $x_i' = \log (x_i-a)$ kan bovenstaand onderzoek dan op waarschijnlijkheidspapier worden uitgevoerd. Het is daarbij niet nodig $\log (x_i-a)$ werkelijk op te zoeken, daar er naast het gewone waarschijnlijkheidspapier met lineaire x-schaal ook logarithmisch waarschijnlijkheidspapier bestaat, waarbij de logarithmische transformatie van de x-schaal reeds is uitgevoerd. Hierop kan men dus weer rechtstreeks $x_{(i)}-a$ en de bijbehorende schattingen $\frac{i-0,3}{n+0,4}$ voor $F(\log (x_{(i)}-a))$ aangeven. Voor een meer uitgebreide behandeling van deze methode verwijzen wij naar het boek van AITCHISON en BROWN *).

Tenslotte bekijken we nog het aanpassen van een exponentiële verdeling voor $\underline{x}$, dus een verdeling met verdelingsfunctie (zie deel 2, (7.21))

$$F(x) = \int_0^x \lambda e^{-\lambda v}\, dv = 1 - e^{-\lambda x} . \tag{1.4}$$

Hierbij kan gewoon logarithmisch papier voor een grafisch onderzoek gebruikt worden, daar

---

*) J. AITCHISON and J.A.C. BROWN, The lognormal distribution, Cambridge University Press (1957).

1.5) $z = \log(1-F(x)) = -\lambda x$

s, dus z een lineaire functie van x. Op dit papier kan dan langs de
ogarithmische schaal een schatting voor $1-F(x_{(i)})$, dus bijv. $1 - \frac{i-0,3}{n+0,4} = \frac{n-i+0,7}{n+0,4}$ en langs de lineaire schaal $x_{(i)}$ uitgezet worden, waarbij
e $x_{(i)}$ weer naar opklimmende grootte gerangschikt zijn. Men kan ook
e $x_{(i)}$ naar dalende grootte rangschikken en dan schatten met $\frac{i-0,3}{n+0,4}$.

Er bestaat ook dubbel logarithmisch papier (beide schalen loga-
ithmisch), waarmee men zou kunnen nagaan of log x exponentieel
erdeeld is.

Andere transformaties van de x-schaal en ook transformaties van de
-schaal die op een andere dan de normale of exponentiële verdeling
erusten, zijn mogelijk om de verdelingsfunctie tot een rechte lijn
n te vormen. Hiervoor bestaat in het algemeen helaas geen speciaal
afiekenpapier; de transformaties moet men dan zelf berekenen, hetgeen
e methode zeer bewerkelijk maakt.

orbeeld 1.3. De verdeling van het gezinsinkomen.

In januari 1957 is in opdracht van de Contactgroep Opvoering
oductiviteit (C.O.P.) door de Stichting voor Statistiek een onderzoek
daan naar de grootte van het gezinsinkomen door middel van een
nquête, waarbij 4400 gezinnen en alleenstaanden werden ondervraagd.
n het volgende wordt een alleenstaande ook als "gezin" aangemerkt.
e resultaten, samengevat in de eerste en tweede kolom van tabel 1.1,
ebben wij ontleend aan het Handboek voor Marktanalytische gegevens,
eel I.

Bij de verdere behandeling van deze gegevens hebben we de 13%
.et opgegeven inkomens buiten beschouwing gelaten (het zou niet juist
.jn hen gelijkelijk over de andere inkomensklassen te verdelen!).
et gevolg hiervan is dat de percentages in de tweede kolom van
abel 1.1 vermenigvuldigd moeten worden met $\frac{100}{87} = 1,149$ om weer een
nsverdeling te krijgen (zie de derde kolom van tabel 1.1; de gecumu-
erde percentages vindt men in de vierde kolom).

Tabel 1.1

Gezinsinkomens (januari 1957)

Inkomensklasse (in guldens)	Percentage gezinnen	Gewijzigde percentages	Idem cumulatief
niet opgegeven	13	-----	-----
≤ 2.400	13	14,94	14,94
2.400 - 3.600	18	20,69	35,63
3.600 - 4.800	22	25,29	60,92
4.800 - 6.600	17	19,54	80,46
6.600 - 9.600	11	12,64	93,10
> 9.600	6	6,90	100

De vraag is nu welke kansverdeling bij deze waarnemingen past. In vele publikaties over inkomens gebruikt men de logarithmisch normale verdeling. Voor de inkomens boven een bepaalde ondergrens x_0 vindt men vaak de verdeling van PARETO, met verdelingsfunctie $1 - (x_0/x)^{\alpha}$ voor $x \geq x_0$, waar α en x_0 positieve parameters zijn. Wij willen nu nagaan of een log-normale verdeling passend is en wel een normale verdeling van de logarithme van het inkomen (dus zonder de behandelde verschuiving).

Omdat de exacte waarnemingen $x_1, x_2, \ldots, x_{4400}$ ons niet bekend zijn, werken we met de gegroepeerde waarnemingen. Het heeft dan geen zin $\frac{i-0,3}{n+0,4}$ uit te zetten; wij weten alleen dat 14,94 procent van de steekproefwaarnemingen links van x = 2,4 (in duizenden guldens) valt, 35,63 procent links van x = 3,6, enz. Op logarithmisch waarschijnlijkheidspapier zetten wij deze gegevens uit (figuur 1.8).

De gevonden punten liggen in goede benadering op een rechte lijn. Hieruit kunnen wij concluderen dat de gezinsinkomens in eerste benadering log-normaal verdeeld zijn.

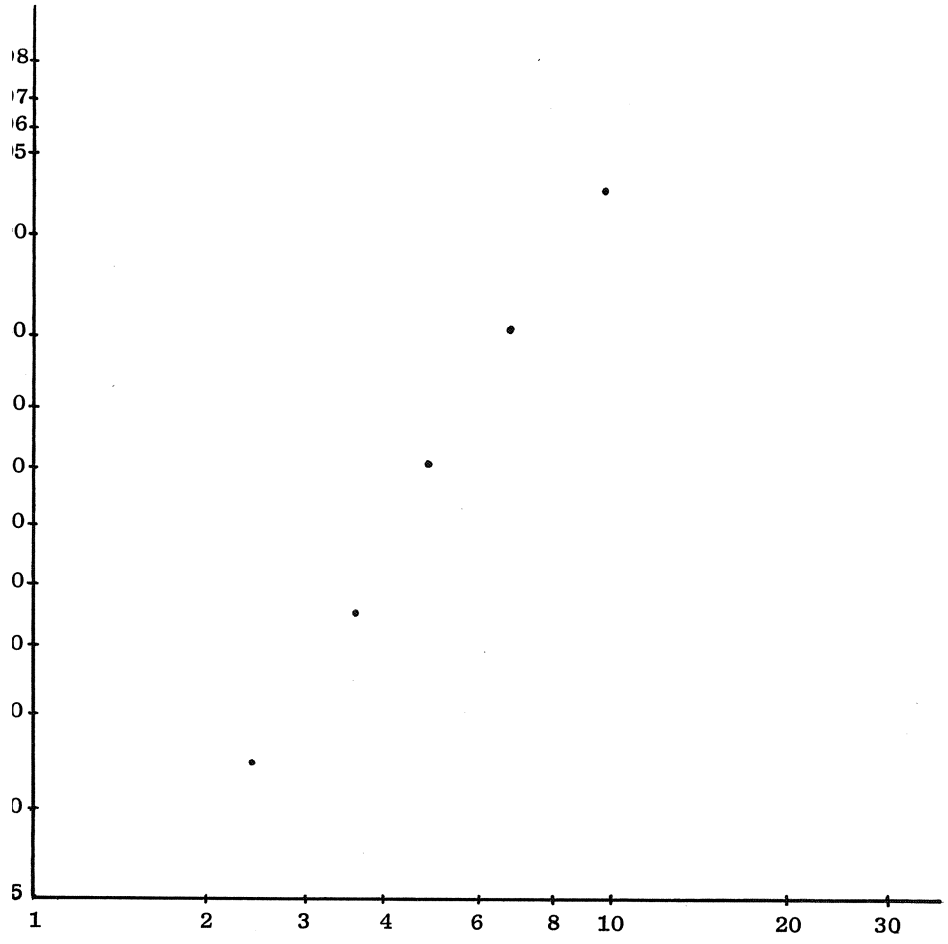


fig. 1.8

Verdeling der gezinsinkomens in januari 1957

voorbeeld 1.4. De verdeling van het inkomen per belastingplichtige.

De gegevens die wij nu bekijken zijn ontleend aan een publikatie van het Centraal Bureau voor Statistiek: Inkomensverdeling 1955. Tabel 1.2 bevat de gegevens betreffende de inkomensverdeling per belastingplichtige over 1955, voor geheel Nederland.

Tabel 1.2 *)

Inkomensverdeling naar het aantal belastingplichtigen (1955)

Cumulatieve inkomensklasse (in guldens)	Aantal belastingplichtigen (cumulatief)	Percentage belastingplichtigen (cumulatief)
≤ 1.000	319.195	7,457
" 2.000	1.036.336	24,212
" 3.000	1.703.132	39,790
" 4.000	2.569.616	60,034
" 5.000	3.237.187	75,630
" 6.000	3.596.688	84,029
" 7.000	3.797.103	88,711
" 8.000	3.918.769	91,554
" 9.000	3.999.459	93,439
" 10.000	4.054.144	94,717
" 15.000	4.179.841	97,653
" 20.000	4.222.372	98,647
" 50.000	4.270.770	99,778
" 100.000	4.278.169	99,951
" + ∞	4.280.283	100,000

In figuur 1.9 zijn deze gegevens uitgezet op logarithmisch waarschijnlijkheidspapier. De rechterklassegrenzen zijn horizontaal uitgezet langs de logarithmische schaal en de bijbehorende cumulatieve percentages belastingplichtigen vertikaal langs de normale waarschijnlijkheid schaal.

Het valt meteen op dat we niet kunnen zeggen dat de gevonden punte in goede benadering op een rechte lijn liggen. Door vergelijken van figuur 1.4 en figuur 1.9 zien we dat de staarten van de verdeling, vooral de rechter staart, te dun zijn voor een log-normale aanpassing.

*) Deze gegevens zijn gebaseerd op de definitieve aanslagen betreffende lo en inkomstenbelasting over 1955. Bij de indeling der inkomens in klasse is in sommige gevallen een correctie in het werkelijke jaarinkomen aang bracht (bijv. bij sterfgevallen). Het hier opgegeven jaarinkomen van de belastingplichtigen is door extrapolatie berekend uit het werkelijke jaarinkomen.

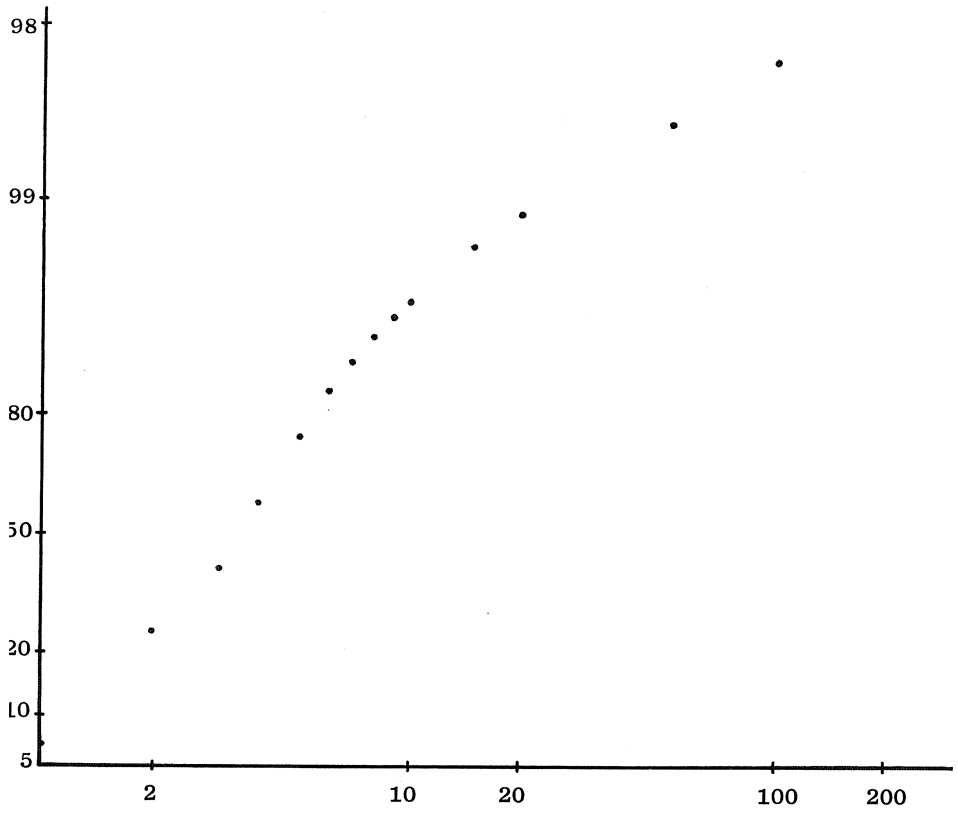


fig. 1.9

Verdeling van het inkomen per belastingplichtige (1955)

Wij komen dus tot de conclusie dat de inkomens per belasting-
.ichtige in eerste benadering niet log-normaal verdeeld zijn. We
)uden nu kunnen trachten een gamma-verdeling aan te passen, aangezien
:ze meestal dunnere staarten heeft dan de log-normale. Omdat hier geen
:nvoudige grafische methode beschikbaar is, zullen wij dit onderzoek
.et uitvoeren.

De in paragraaf 10 te bespreken χ^2-toets kan eveneens dienen om
: onderzoeken of een bepaalde verdeling bij de waarnemingen past.

2. Inleiding tot de schattingstheorie.

Wanneer de aard van de kansverdeling is vastgesteld, kunnen wij proberen de onbekende parameter(s) uit de waarnemingen te schatten. Als θ een parameter voorstelt en $x_1,x_2,\ldots,x_n$ de waarnemingen, zoeken wij een functie $t(x_1,x_2,\ldots,x_n)$ die θ zo goed mogelijk benadert.

Stellen wij ons weer op het standpunt dat het aselect trekken en waarnemen nog niet geschied is, dan zijn $\underline{x}_1,\underline{x}_2,\ldots,\underline{x}_n$ onafhankelijke stochastische grootheden met dezelfde verdeling en $\underline{t} = t(\underline{x}_1,\underline{x}_2,\ldots,\underline{x}_n)$ is eveneens stochastisch. Wij noemen $\underline{t}$ een schatter; de getalwaarde $t = t(x_1,x_2,\ldots,x_n)$ die ontstaat als we voor alle $\underline{x}_i$ de waargenomen getalwaarden x_i invullen, heet een schatting. Men zou kunnen zeggen dat de schatter een methode aangeeft om bij elk stel waarnemingen een schatting te produceren. In het spraakgebruik wordt het onderscheid tussen schatter en schatting niet altijd in acht genomen.

In principe kan nu iedere functie t van n variabelen gekozen worden om de onbekende parameter θ te schatten. Van deze vrijheid maakt men gebruik door te eisen dat de schatter één of meer gunstige eigenschappen bezit; men zoekt schatters die bijv. zuiver zijn en asymptotisch raak. Deze begrippen worden nu gedefinieerd, waarbij we in plaats van $\underline{t}$ het symbool $\underline{t}_n$ zullen gebruiken om duidelijk aan te geven dat de schatter gebaseerd is op n waarnemingen.

Definitie 2.1

Een schatter $\underline{t}_n$ van een onbekende parameter θ wordt zuiver genoemd, als zijn verwachting gelijk is aan deze onbekende parameter; in formule

(2.1) $$\mathcal{E}\,\underline{t}_n = \theta \ .$$

Definitie 2.2

Een schatter $\underline{t}_n$ van een onbekende parameter θ wordt asymptotisch raak genoemd, als de rij $\{\underline{t}_n\}$ (n=1,2,...) voor een onbegrensd stijgend aantal waarnemingen n stochastisch convergeert naar die onbekende parameter; in formule

(2.2) $$\lim_{n\to\infty} P\left[|\underline{t}_n-\theta| > \varepsilon\right] = 0 \quad \text{voor iedere } \varepsilon > 0.$$

Het begrip zuiver hangt niet samen met de steekproefomvang; het etekent ruw gezegd dat de schatting gemiddeld over alle denkbare selecte steekproeven de juiste waarde oplevert. Asymptotische raakheid s een eigenschap die alleen van belang is voor een grote steekproef-mvang.

Bij vele verdelingen zijn de parameters juist de verwachting en e variantie, of eenvoudige functies van deze momenten. Met name de ormale verdeling is door μ en σ^2 volkomen vastgelegd. Wij zullen aarom eerst schatters behandelen voor de verwachting, de variantie en e standaardafwijking van een willekeurige verdeling.

Het ligt voor de hand als schatter voor het gemiddelde μ te nemen et steekproefgemiddelde *); dus

2.3) $$\bar{\underline{x}} = \frac{1}{n} \sum_{i=1}^{n} \underline{x}_i \ .$$

olgens stelling 12.2 uit deel 2 is deze schatter asymptotisch raak ı in het bewijs van deze stelling bleek $\mathcal{E}\,\bar{\underline{x}}$ gelijk aan μ te zijn, odat $\bar{\underline{x}}$ ook zuiver is.

Als schatter voor de variantie σ^2 onderzoeken wij de steekproef-ıriantie *)

2.4) $$\underline{s}'^2 = \frac{1}{n} \sum_{i=1}^{n} (\underline{x}_i - \bar{\underline{x}})^2$$

) zijn merites.

: verwachting van $\underline{s}'^2$ berekenen wij als volgt:

---

j de waarnemingen $x_1, x_2, \ldots, x_n$ kunnen wij een nieuwe stochastische ootheid $\underline{x}'$ invoeren, die elk van deze n waarden aanneemt met kans $\frac{1}{n}$. n vindt dan $\mathcal{E}\,\underline{x}' = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}$ en $\sigma^2(\underline{x}') = \mathcal{E}(\underline{x}' - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 =$ $'^2$; vandaar de namen steekproefgemiddelde en steekproefvariantie. n noemt de verdeling van $\underline{x}'$ de steekproefverdeling; deze is dus altijd screet ook als de gemeenschappelijke verdeling van $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ ntinu is.

$$\sum_{i=1}^{n} (\underline{x}_i - \bar{\underline{x}})^2 = \sum_{i=1}^{n} (\underline{x}_i - \mu + \mu - \bar{\underline{x}})^2 =$$

(2.5)

$$= \sum_{i=1}^{n} (\underline{x}_i - \mu)^2 - 2(\bar{\underline{x}} - \mu) \sum_{i=1}^{n} (\underline{x}_i - \mu) + n(\bar{\underline{x}} - \mu)^2 .$$

Hierin is

$$\sum_{i=1}^{n} (\underline{x}_i - \mu) = n(\bar{\underline{x}} - \mu) , \tag{2.6}$$

zodat

$$\sum_{i=1}^{n} (\underline{x}_i - \bar{\underline{x}})^2 = \sum_{i=1}^{n} (\underline{x}_i - \mu)^2 - n(\bar{\underline{x}} - \mu)^2 . \tag{2.7}$$

Daar

$$\mathcal{E}(\underline{x}_i - \mu)^2 = \sigma^2(\underline{x}) = \sigma^2 \quad \text{en} \quad \mathcal{E}(\bar{\underline{x}} - \mu)^2 = \sigma^2(\bar{\underline{x}}) = \frac{\sigma^2}{n} , \tag{2.8}$$

geldt

$$\mathcal{E}\underline{s}'^2 = \frac{n-1}{n} \sigma^2 . \tag{2.9}$$

De schatter $\underline{s}'^2$ is dus niet zuiver, maar

$$\underline{s}^2 \overset{\text{def}}{=} \frac{1}{n-1} \sum_{i=1}^{n} (\underline{x}_i - \bar{\underline{x}})^2 \tag{2.10}$$

is dat wel. Voor praktische toepassingen wordt daarom veelal $\underline{s}^2$ als schatter van σ^2 gebruikt.

Opmerking 2.1

Het argument dat $\underline{s}^2$ een zuivere schatter van σ^2 is en $\underline{s}'^2$ niet, verliest zijn waarde indien men niet σ^2 maar σ wenst te schatten, hetgeen in de praktijk gewoonlijk het geval is. Beschouwt men immers $\underline{s}$ als schatter van σ, dan is $\underline{s}$ weer onzuiver. Men kan dit inzien door in deel 2 (10.26) voor $\underline{x}$ te substitueren $\underline{s}$, waardoor men vindt

$$\sigma^2 = \mathcal{E}\underline{s}^2 > (\mathcal{E}\underline{s})^2; \quad \text{dus} \quad \sigma > \mathcal{E}\underline{s} , \tag{2.11}$$

zodat de verwachting van $\underline{s}$ kleiner is dan de te schatten parameter σ.

let gebruik van $\underline{s}^2$ in plaats van $\underline{s}'^2$ berust dan ook vooral op traditie.

In het speciale geval van een normaal verdeelde $\underline{x}$ kan men de faktor $c_n \overset{\text{def}}{=} \frac{\sigma}{\mathcal{E}\,\underline{s}}$ berekenen; dan is $c_n\underline{s}$ een zuivere schatter van σ. Een tabel van c_n vindt men bij E.S. PEARSON and H.O. HARTLEY, Biometrika Tables for Statisticians, vol.I, Cambridge, (2e druk,1958). Zij geven c_n onder de naam "expectation" als functie van $\nu = n-1$ (bij 10 waarnemingen moet men dus $\nu = 9$ opzoeken), voor $\nu = 1\ (1)\ 20\ (5)\ 50\ (10)\ 100$ *). Omdat $\lim_{n\to\infty} c_n = 1$ is, kan men voor $n \geq 100$ wel $c_n = 1$ stellen (de tabel geeft $c_{99} = 0{,}9975$).

Wij zullen nu bewijzen dat $\underline{s}$, $\underline{s}'$, $\underline{s}^2$ en $\underline{s}'^2$ allen asymptotisch raak zijn. Omdat uit $|\underline{s}-\sigma| > \varepsilon$ volgt $|\underline{s}^2-\sigma^2| > \varepsilon|\underline{s}+\sigma| > \varepsilon(2\sigma-\varepsilon)$, blijkt dat $\underline{s}$ asymptotisch raak is zodra dit voor $\underline{s}^2$ geldt; evenzo $\underline{s}'$ en $\underline{s}'^2$. Verder zijn $\underline{s}^2$ en $\underline{s}'^2$ voor $n \to \infty$ equivalent, omdat $\frac{n-1}{n} \to 1$. Zo blijkt het voldoende om

2.12) $$\underline{s}'^2 = \frac{1}{n}\sum_{i=1}^{n}(\underline{x}_i-\mu)^2 - (\underline{\bar{x}}-\mu)^2$$

te beschouwen. Volgens deel 2, stelling 12.2 convergeert $\underline{\bar{x}}$ stochastisch naar μ, dus $|\underline{\bar{x}}-\mu|$ naar 0, dus ook $(\underline{\bar{x}}-\mu)^2$ naar 0. Deze term kan dus verwaarloosd worden: behoudens een willekeurig kleine kans wordt hij zelf willekeurig klein. Noemen wij verder

2.13) $$\underline{y}_i \overset{\text{def}}{=} (\underline{x}_i-\mu)^2 \qquad (i=1,2,\ldots,n)\ ,$$

dan convergeert, wederom volgens de reeds aangehaalde stelling,

2.14) $$\frac{1}{n}\sum_{i=1}^{n}\underline{y}_i = \frac{1}{n}\sum_{i=1}^{n}(\underline{x}_i-\mu)^2$$

stochastisch naar de verwachting $\mathcal{E}\underline{y}_i = \sigma^2$ **). Daarmede is (zij het niet volledig exact) bewezen dat $\underline{s}'^2$ asymptotisch raak is.

---

*) Zoals in deel 2 reeds gesteld is betekent de notatie a (c) b : vanaf a met stapgrootte c tot en met b. De grootheid ν doorloopt hier dus alle waarden 1 t/m 20, dan 25,30,35,40,45,50,60,70,80,90,100.

**) Volgens de genoemde stelling moet de voorwaarde vervuld zijn dat $\sigma^2(\underline{y}_i)$ eindig is; deze voorwaarde kan echter vermeden worden.

Wij zijn er dus in geslaagd schatters te vinden voor μ en σ^2, die zowel zuiver zijn als asymptotisch raak. Terloops merken wij op dat het in sommige gevallen niet mogelijk is schatters voor onbekende parameters van kansverdelingen te vinden, die èn zuiver èn asymptotisch raak zijn.

Voorbeeld 2.1

Van 10 lampen van een bepaald type, aselect gekozen uit de produktie, is de levensduur (in uren) gemeten. De uitkomsten zijn vermeld in de eerste kolom van tabel 2.1 *).

Tabel 2.1

Berekening van gemiddelde en variantie van een reeks waarnemingen

x_i	x_i-1000	$(x_i-1000)^2$
982	- 18	324
996	- 4	16
1047	47	2209
980	- 20	400
1057	57	3249
1025	25	625
990	- 10	100
1058	58	3364
994	- 6	36
1036	36	1296
totaal	165	11619

Om de berekening van $\bar{x}$ en s^2 te vereenvoudigen kiezen wij eerst een zg. "voorlopig gemiddelde"; d.i. een getal x_0 (het doet er niet veel toe welk precies), dat vermoedelijk in de buurt van het werkelijke gemiddelde $\bar{x}$ ligt. In dit geval ligt het voor de hand $x_0 = 1000$ te nemen. In de tweede kolom staat x_i-1000 berekend, in de derde (x_i-1000)

*) Deze gegevens zijn niet aan de praktijk ontleend.

lu is

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} (x_i - x_0) + nx_0 ,$$

us

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} (x_i - x_0) + x_0 .$$

n ons geval is

$$\bar{x} = \frac{1}{10}.165 + 1000 = 1016,5 .$$

erder geldt (zie (2.5))

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} (x_i - x_0)^2 - n(x_0 - \bar{x})^2 .$$

n ons geval is dus

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = 11619 - 10(16,5)^2 = 8896,5 .$$

erhalve is

$$s^2 = \frac{1}{9}.8896,5 = 988,5 \rightarrow s = 31,4 .$$

Willen wij nu, op grond van de gevonden schattingen voor μ en σ, ijv. schatten hoe groot het percentage lampen is dat een levensduur van inder dan 950 uur bezit, terwijl al is vastgesteld dat de levensduur ngeveer normaal verdeeld is, dan berekenen wij $\frac{1016,5-950}{31,4} = 2,12$ n zoeken dit getal op in de tabel van de N(0,1)-verdeling. De uitkomst s 0,017, dus 1,7%. Van de steekproef is geen enkele uitkomst zo laag.

pmerking 2.2

De hier aangegeven methode om het gemiddelde te bepalen vormt lleen een vereenvoudiging wanneer men de berekeningen met potlood en apier uitvoert. Wanneer men een tafelrekenmachine gebruikt, kan men et beste $x_0 = 0$ kiezen en dus uitgaan van de formule

2.15)
$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n} .$$

Men berekent dan de som van kwadraten $\sum_{i=1}^{n} x_i^2$ en trekt er profijt van dat de rekenmachine de som $\sum_{i=1}^{n} x_i$ tegelijkertijd noteert.

Voorbeeld 2.2. Enquête.

Wij beschouwen vervolgens een enquête die erop gericht is de kennis van een bepaalde groep mensen te onderzoeken. Voor de eenvoud houden wij één vraag in het oog, die juist of onjuist beantwoord kan worden. Deze vraag wordt voorgelegd aan een steekproef van n personen, aselect genomen uit de te onderzoeken groep (de populatie). Als model nemen wij aan dat een fractie p van de personen uit deze groep de vraag goed zou beantwoorden en de overige (fractie q = 1-p) fout. Het gaat er nu om deze fracties te schatten op grond van de antwoorden van de n personen van de steekproef.

Kiezen wij aselect één persoon uit de populatie, dan is de kans dat wij een juist antwoord verkrijgen gelijk aan p. Bestaat de populatie uit N personen, dan bevinden zich daaronder dus Np, die een juist antwoord zouden geven en Nq, die de vraag onjuist zouden beantwoorden. Hieruit zijn er nu n aselect en zonder teruglegging genomen en wij bevinden ons dus in de situatie van voorbeeld 5.7 van deel 2. Het aantal juiste antwoorden in de steekproef van omvang n, dat wij met $\underline{k}$ aangeven, heeft dus een hypergeometrische verdeling, waarvoor geldt

$$P[\underline{k} = k] = \frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}} . \tag{2.16}$$

Volgens tabel A aan het slot van deel 2 is de verwachting van $\underline{k}$, wanneer wij de hier gebruikte symbolen substitueren

$$\mathcal{E}\,\underline{k} = np . \tag{2.17}$$

Dus

$$\mathcal{E}\,\frac{\underline{k}}{n} = p . \tag{2.18}$$

De fractie juiste antwoorden in de steekproef is dus een zuivere schatter voor de onbekende fractie p. De vraag naar het asymptotisch

;edrag doet zich hier niet op natuurlijke wijze voor, daar n > N niet :an optreden. Voor n = N vindt men p reeds exact.

Het kan wel gebeuren dat N niet bekend is. Volgens (2.18) kan $\frac{k}{n}$ lan toch als schatter gebruikt worden. Is N >> n, dan is volgens (5.14) ıit deel 2 de kansverdeling van k bij benadering binomiaal met para- ıeters n en p, waarbij (2.17) geldig blijft. Aan de schatter verandert lus niets. Volgens stelling 12.3 uit deel 2 (de theoretische wet van de rote aantallen) is de schatter in het geval van een binomiale verdeling symptotisch raak. (In het hier beschouwde geval impliceert $n \to \infty$ dan ok $N \to \infty$; is p de onbekende kans op succes bij een experiment, dat illekeurig vaak herhaald kan worden, dan doet deze complicatie zich iet voor.)

Hiermee besluiten wij de voorbeelden, waarin goede schatters van e onbekende parameters alle min of meer voor de hand lagen. Naast euristische redeneringen om schatters te vinden, bestaan er ook erschillende algemene theorieën om schatters af te leiden. De belang- ijkste is de theorie van de meest aannemelijke schatters, waarvan de rondslag gelegd is door R.A. FISHER; enkele elementen van deze theorie ehandelen wij in paragraaf 4. Op een andere methode wordt in paragraaf in het kort ingegaan.

. Nauwkeurigheid van schatters; betrouwbaarheidsintervallen.

Wij hebben in paragraaf 2 al opgemerkt dat een schatter een tochastische grootheid is. De kansverdeling van de schatter geeft e volledige beschrijving van al zijn eigenschappen. De speciale eigen- chap zuiverheid heeft betrekking op de verwachting, asymptotische aakheid op het gedrag van de kansverdeling voor $n \to \infty$. Een belangrijke ɔg niet beschouwde eigenschap is de nauwkeurigheid van een schatter, le men bijv. door zijn variantie kan aangeven.

Voorbeeld 3.1. Variantie van het steekproefgemiddelde.

Als $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ onafhankelijke trekkingen uit een verdeling met verwachting μ en variantie σ^2 voorstellen, dan hebben we μ geschat door het steekproefgemiddelde

$$\bar{\underline{x}} = \frac{1}{n} \sum_{i=1}^{n} \underline{x}_i \quad . \tag{3.1}$$

Zoals in het bewijs van stelling 12.2 uit deel 2 is afgeleid, geldt $\mathcal{E}\,\bar{\underline{x}} = \mu$ en $\sigma^2(\bar{\underline{x}}) = \frac{\sigma^2}{n}$. De schatter $\bar{\underline{x}}$ heeft dus een standaardafwijking $\frac{\sigma}{\sqrt{n}}$.

Wanneer de standaardafwijking σ niet bekend is, geeft deze uitkomst weinig informatie. Als het er alleen om te doen is een globale indruk van de waarde van $\sigma(\bar{\underline{x}})$ te krijgen, kan men voor σ de schatting s of s' invullen (zie paragraaf 2). Op de bezwaren van dit invullen komen wij nog terug.

Bij het onderling vergelijken van verschillende schatters voor μ is de uitkomst $\sigma(\bar{\underline{x}}) = \frac{\sigma}{\sqrt{n}}$ nuttig. Zo kan men de verwachting μ van een verdeling ook schatten door de steekproefmediaan $\underline{m}$ (de middelste waarneming, of het gemiddelde van de beide middelste als n even is). Voor een symmetrische verdeling kan men aantonen dat $\underline{m}$ zuiver is; voor een normale verdeling $N(\mu,\sigma)$ bovendien dat $\underline{m}$ asymptotisch raak is en dat voor grote n $\sigma(\underline{m}) \approx \sigma\sqrt{\frac{2\pi}{n}}$ is. Bij gelijkblijvend aantal waarnemingen is het steekproefgemiddelde dus een nauwkeuriger schatter voor de verwachting van een normale verdeling dan de steekproefmediaan. Bij sommige onderzoekingen verkiest men toch de mediaan, omdat deze eenvoudiger te bepalen is.

Men kan bewijzen dat er voor een steekproef uit een normale verdeling geen enkele zuivere schatter voor μ bestaat die een kleinere variantie heeft dan $\bar{\underline{x}}$; men noemt $\bar{\underline{x}}$ daarom de meest doeltreffende zuivere schatter.

Voorbeeld 3.2. Binomiale verdeling.

Bij een binomiale verdeling met parameters n en p vonden wij in voorbeeld 2.2 als schatter voor p de grootheid $\frac{\underline{k}}{n}$. De variantie hiervan is volgens stelling 10.5 uit deel 2

3.2) $$\sigma^2\left(\frac{\underline{k}}{n}\right) = \frac{1}{n^2} \cdot \sigma^2(\underline{k}) = \frac{1}{n^2} \cdot npq = \frac{pq}{n} .$$

eze variantie, die weer de kleinst mogelijke voor een zuivere schatter an p is, is weer in de regel onbekend, doch kan uit de waarnemingen eschat worden. Een zuivere schatter ervan is

3.3) $$\frac{\underline{k}(n-\underline{k})}{n^2(n-1)} ,$$

etgeen volgt uit

3.4) $$\mathcal{E}\,\underline{k}(n-\underline{k}) = n\,\mathcal{E}\underline{k} - \mathcal{E}\underline{k}^2 = n\,\mathcal{E}\underline{k} - \sigma^2(\underline{k}) - (\mathcal{E}\underline{k})^2 = \\ = n^2p - npq - (np)^2 = n(n-1)pq .$$

Een schatter $\underline{t}$ geeft bij elke serie getalwaarden voor de waarnemingen een getalwaarde $t(x_1,x_2,\ldots,x_n)$, de schatting of <u>puntschatting</u> voor de onbekende parameter θ. Als we de gehele kansverdeling van de schatter $\underline{t}$ kennen, is het vaak mogelijk via $\underline{t}$ een <u>interval</u> te construeren waarbinnen de onbekende waarde θ vermoedelijk ligt. Hiertoe hebben wij de volgende definitie nodig, waarin α_ℓ, α_r en α vaste getallen tussen 0 en 1 zijn. (Gangbare waarden zijn 0,05 of 0,025 of 0,01.)

<u>efinitie 3.1</u>

Een functie $\underline{g}_\ell = g_\ell(\underline{x}_1,\ldots,\underline{x}_n)$ van de waarnemingen is een <u>betrouwbaarheidsondergrens</u> voor een onbekende parameter θ, met onbetrouwbaarheidsdrempel α_ℓ, indien geldt

3.5) $$P\left[\underline{g}_\ell < \theta\right] \geq 1 - \alpha_\ell ,$$

waarin θ de (onbekende) werkelijke waarde van de parameter voorstelt.

Analoog geldt voor een <u>betrouwbaarheidsbovengrens</u> $\underline{g}_r$, met onbetrouwbaarheidsdrempel α_r,

3.6) $$P\left[\underline{g}_r > \theta\right] \geq 1 - \alpha_r .$$

Tezamen vormen $\underline{g}_\ell$ en $\underline{g}_r$ een <u>betrouwbaarheidsinterval</u> voor θ, met onbetrouwbaarheid α, indien geldt

$$P\left[\underline{g}_\ell < \theta < \underline{g}_r\right] \geq 1-\alpha; \tag{3.7}$$

hierin is $\alpha = \alpha_r + \alpha_\ell$, indien $P\left[\underline{g}_\ell < \underline{g}_r\right] = 1$.

Indien de verdeling van $\underline{g}_\ell$ en $\underline{g}_r$ continu is, kan men zorgen dat de kansen in (3.5) t/m (3.7) exact gelijk zijn aan de rechterleden. Dan spreekt men van onbetrouwbaarheid i.p.v. onbetrouwbaarheidsdrempel.

Wij zullen nu aan een voorbeeld toelichten, hoe men van een schatter tot een betrouwbaarheidsinterval kan komen. Na de behandeling van de toetsingstheorie komen wij nog op de constructie van betrouwbaarheidsintervallen terug.

<u>Voorbeeld 3.3.</u> <u>Betrouwbaarheidsinterval voor de verwachting van een normale verdeling.</u>

Wanneer $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ een steekproef uit een $N(\mu,\sigma)$-verdeling zijn, heeft

$$\underline{\bar{x}}^* = \frac{\underline{\bar{x}} - \mu}{\sigma} \cdot \sqrt{n} \tag{3.8}$$

de N(0,1)-verdeling (zie deel 2, stelling 11.4 en bewijs van stelling 12.2). Volgens tabel B van het slot van deel 2 geldt dus bijv.

$$P\left[\left|\underline{\bar{x}}^*\right| < 1{,}96\right] = 0{,}95\ . \tag{3.9}$$

Wij kunnen dit omrekenen tot

$$P\left[\mu - 1{,}96\,\frac{\sigma}{\sqrt{n}} < \underline{\bar{x}} < \mu + 1{,}96\,\frac{\sigma}{\sqrt{n}}\right] = 0{,}95\ . \tag{3.10}$$

Het interval $\left(\mu - 1{,}96\,\frac{\sigma}{\sqrt{n}}\ ,\ \mu + 1{,}96\,\frac{\sigma}{\sqrt{n}}\right)$ wordt nu een <u>voorspellingsinterval</u> voor $\underline{\bar{x}}$ met onbetrouwbaarheid 0,05 genoemd; de voorspelling dat $\underline{\bar{x}}$ in dit interval ligt is met kans 1-0,05 juist. Maar wij kunnen (3.9) ook omrekenen tot

3.11) $$P\left[\underline{\bar{x}} - 1{,}96\,\frac{\sigma}{\sqrt{n}} < \mu < \underline{\bar{x}} + 1{,}96\,\frac{\sigma}{\sqrt{n}}\right] = 0{,}95\ .$$

Dus bepalen $\underline{g}_\ell = \underline{\bar{x}} - 1{,}96\,\frac{\sigma}{\sqrt{n}}$ en $\underline{g}_r = \underline{\bar{x}} + 1{,}96\,\frac{\sigma}{\sqrt{n}}$ een <u>betrouwbaarheidsinterval</u> voor μ met onbetrouwbaarheid 0,05. De interpretatie van o'n interval willen wij toelichten aan de hand van voorbeeld 2.1, aarbij $n = 10$ en $\bar{x} = 1016{,}5$ was.

De ongelijkheid tussen de vierkante haken van (3.11) gaat bij ubstitutie van deze waarden over in

3.12) $$1016{,}5 - 1{,}96\,\frac{\sigma}{\sqrt{10}} < \mu < 1016{,}5 + 1{,}96\,\frac{\sigma}{\sqrt{10}}\ ,$$

aarmee wij nog niet veel opschieten als σ onbekend is. Nemen wij even an dat σ op grond van vroegere waarnemingen bekend is en bijv. 33 edraagt, dan vinden wij na invullen van deze waarde

3.13) $$996 < \mu < 1037\ .$$

Deze uitspraak, waarbij de onbekende parameter μ tussen twee renzen ingesloten wordt, kan juist of onjuist zijn. Op grond van (3.11) ou men de neiging kunnen hebben te zeggen dat er een kans 0,95 is dat ij juist is. Strikt genomen is dit niet waar: (3.13) is òfwel juist en dan is de kans dat hij juist is gelijk aan 1), òfwel onjuist (en dan s die kans gelijk aan 0). Maar wèl volgt uit (3.11) dat de gevolgde <u>ethode</u> een kans 0,95 bezit op een juiste uitkomst. Beperkt men dus zijn eschouwingen niet tot één geval (zoals de uitkomst (3.13)), maar eschouwt men de methode als zodanig, of een lange reeks toepassingen aarvan, dan kan men zeggen dat de kans op een juiste uitspraak 0,95 edraagt. Een betrouwbaarheidsinterval is een interval met stochastische indpunten, dat met kans minstens $1-\alpha$ de (vaste maar onbekende) paraeterwaarde bevat.

Meestal is σ niet bekend. In eerste instantie substitueert men dan elal een schatting van σ, bijv. $s = \sqrt{s^2}$ (zie (2.10)) voor σ. Zoals ij hebben gezien is de schatter <u>s</u> niet zuiver; gemiddeld wordt σ door onderschat. Dit doet vermoeden dat de grenzen waartussen μ op deze

wijze ingesloten wordt, wel eens te dicht bij elkaar kunnen komen, of ook dat de kans op een verkeerde uitspraak bij toepassing van deze methode groter dan α zal zijn. Dit is inderdaad juist en men kan de substitutie van s voor σ dan ook slechts als een eerste benadering beschouwen, die voor grote steekproeven ($n \to \infty$) goed bruikbaar is. Dan convergeert nl. $\underline{s}$ stochastisch naar σ. Voor kleine steekproeven bestaat - in het normale geval - een exacte methode, die op de zg. verdeling van STUDENT berust. Wij gaan daar niet op in.

Uit (3.11) is (nogmaals) te zien dat μ in principe met willekeurige nauwkeurigheid bepaald kan worden door n voldoende groot te nemen. Immers voor $n \to \infty$ naderen de beide grenzen tot elkaar. Dit is equivalent met het vroeger bewezene, dat $\underline{\bar{x}}$ voor $n \to \infty$ stochastisch naar μ convergeert.

In de praktijk dient deze stelling echter met voorzichtigheid gehanteerd te worden. Indien men bijv. de lengte van een voorwerp een aantal malen opmeet met een meetinstrument dat niet nauwkeuriger dan in cm. afleesbaar is en de werkelijke lengte van het voorwerp is gelijk aan 10,9 cm., dan zullen alle waarnemingen (of vrijwel alle) de waarde 11 cm. aangeven en men behoeft er niet op te hopen dat het gemiddelde tot 10,9 cm. zal convergeren. De kansverdeling der waarnemingen is dan ook niet normaal, zelfs niet continu, daar alleen gehele waarden aangenomen kunnen worden. In de praktijk zijn echter waarnemingen altijd van discreet karakter. De moraal daarvan is, dat het in de regel zinloos is het gemiddelde van een aantal metingen in meer decimalen op te geven dan op het meetinstrument afgelezen kunnen worden. Hiertegen wordt vaak gezondigd.

Wij merken nog op, dat het construeren van een betrouwbaarheidsinterval op grond van een schatter niet steeds uitvoerbaar is. Het kriterium voor de toepasbaarheid van de methode ligt in de mogelijkheid om een voorspellingsinterval zoals gegeven in (3.10) om te zetten in een betrouwbaarheidsinterval zoals gegeven in (3.11).

. Meest aannemelijke schatters.

Het principe van de theorie zetten wij uiteen voor continue erdelingen. Voor discrete is het analoog (met kansen in plaats van erdelingsdichtheden).

Laten

4.1) $$\underline{x}_1,\underline{x}_2,\ldots,\underline{x}_n$$

(al of niet onafhankelijke) waarnemingen voorstellen, die een imultane verdeling bezitten met verdelingsdichtheid

4.2) $$f(x_1,x_2,\ldots,x_n \mid \theta_1,\theta_2,\ldots,\theta_h) ,$$

aarin $\theta_1,\theta_2,\ldots,\theta_h$ onbekende parameters voorstellen, die uit de waarneingen geschat moeten worden. Stellen nu $x_1,x_2,\ldots,x_n$ de waarden voor ie in werkelijkheid gevonden zijn, zodat het dus getallen zijn (geen ariabelen meer), dan is (4.2) een functie van de onbekende parameters $_j$ (j=1,2,...,h). Deze functie wordt de aannemelijkheidsfunctie of ɔrtweg aannemelijkheid genoemd en aangegeven als

4.3) $$L(\theta_1,\theta_2,\ldots,\theta_h) \overset{\text{def}}{=} f(x_1,x_2,\ldots,x_n \mid \theta_1,\theta_2,\ldots,\theta_h) .$$

ɜwoonlijk (doch niet voor alle verdelingsdichtheden f) bezit L een ɑximum. Het punt waarin dit wordt aangenomen, wordt aangegeven met $_1,\hat{\theta}_2,\ldots,\hat{\theta}_h$ en deze waarden - die functies van de x_i (i=1,2,...,n) ɬjn - worden de meest aannemelijke schattingen van $\theta_1,\theta_2,\ldots,\theta_h$ ɜnoemd. Zolang de x_i nog stochastisch zijn spreken we van meest ɑnnemelijke schatters.

ɔorbeeld 4.1. Normale verdeling.

Laat $\underline{x}$ een normale verdeling bezitten met onbekend gemiddelde μ ɪ onbekende variantie σ^2 en laten $x_1,x_2,\ldots,x_n$ onderling onafhanke-.jke waarnemingen van deze grootheid zijn. Wegens de onafhankelijkheid ɪnnen wij stelling 9.3 uit deel 2 op (4.3) toepassen, zodat wij .nden met $\theta_1 = \mu$ en $\theta_2 = \sigma^2$:

$$(4.4) \qquad L(\mu,\sigma^2) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \prod_{i=1}^{n} e^{-\frac{1}{2}\frac{(x_i-\mu)^2}{\sigma^2}} .$$

In deze vorm beschouwen wij σ^2 (en niet σ) als onbekende parameter, daar dit de berekeningen vereenvoudigt. Om het maximum van L te vinden bij gegeven waarden der x_i, kunnen wij de partiële afgeleiden van L naar μ en σ^2 nul stellen (vgl. deel 1, pag. 105). Ter vereenvoudiging differentiëren we liever log L; de lezer kan direkt inzien dat wegens $L > 0$ bijv. de beweringen $\frac{\partial L}{\partial \mu} = 0$ en $\frac{\partial \log L}{\partial \mu} = 0$ equivalent zijn.

$$(4.5) \qquad \log L = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log \sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 ,$$

$$(4.6) \qquad \frac{\partial \log L}{\partial \mu} = \frac{\sum_{i=1}^{n}(x_i-\mu)}{\sigma^2} ,$$

$$(4.7) \qquad \frac{\partial \log L}{\partial (\sigma^2)} = -\frac{n}{2\sigma^2} + \frac{1}{2}\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^4} .$$

Stellen wij deze afgeleiden gelijk aan 0, dan verkrijgen wij twee vergelijkingen voor $\hat{\mu}$ en $\hat{\sigma}^2$, en wel

$$(4.8) \qquad \sum_{i=1}^{n}(x_i-\hat{\mu}) = 0$$

en

$$(4.9) \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2 .$$

Uit de eerste volgt

$$(4.10) \qquad \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x} \qquad \text{(vgl. (2.3))}$$

en vervolgens uit de tweede

4.11) $$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 = s'^2 \qquad \text{(vgl. (2.4))} .$$

en kan nu nagaan dat de functie $L(\mu,\sigma^2)$ voor de waarden $\mu = \bar{x}$ en $^2 = s'^2$ inderdaad een maximum heeft (vgl. deel 1, stelling 12.7). e meest aannemelijke schattingen $\hat{\mu}$ en $\hat{\sigma}^2$ zijn voor de normale verdeling lijkbaar identiek met de in paragraaf 2 gevonden asymptotisch rake chattingen.

pmerking 4.1

Indien wij niet σ^2 maar σ zelf als onbekende parameter gekozen adden, zou dezelfde uitkomst verkregen zijn. Dit is een speciaal geval an een algemene stelling, inhoudende dat de meest aannemelijke schatter an een continue en omkeerbare (één-éénduidige) functie van een nbekende parameter verkregen wordt door dezelfde functie te nemen van e meest aannemelijke schatter van die parameter. Dus: is $\hat{\underline{\theta}}$ de meest annemelijke schatter van θ, dan is $Q(\hat{\underline{\theta}})$ de meest aannemelijke schatter an $Q(\theta)$, mits Q een continue en ondubbelzinnig omkeerbare functie is.

pmerking 4.2

Wij zien tevens dat meest aannemelijke schatters niet zuiver ehoeven te zijn. Dit volgt trouwens reeds uit opmerking 4.1, daar uiverheid, zoals wij in opmerking 2.1 gezien hebben, bij transformatie an de parameter verloren kan gaan.

oorbeeld 4.2. Binomiale verdeling.

Is $\underline{x}$ binomiaal verdeeld met onbekende p, maar bekende n, dan geldt

4.12) $$L(p) = P\left[\underline{x}=k \mid p\right] = \binom{n}{k} p^k q^{n-k} ,$$

4.13) $$\log L(p) = \log \binom{n}{k} + k \log p + (n-k) \log (1-p) .$$

ifferentiëren naar p geeft

4.14) $$\frac{d \log L}{dp} = \frac{k}{p} - \frac{n-k}{1-p} ,$$

zodat wij om $\hat{p}$ te vinden de vergelijking

$$\frac{k}{\hat{p}} = \frac{n-k}{1-\hat{p}} \tag{4.15}$$

moeten oplossen. Dit geeft

$$\hat{p} = \frac{k}{n} \tag{4.16}$$

en deze waarde geeft inderdaad een maximum voor log L.

Opmerking 4.3

Ook hier vinden wij dus de vroeger reeds vermelde schatter. Voor de hypergeometrische verdeling is dit niet het geval; daarbij treden complicaties op, o.a. doordat Np in dat geval geheel moet zijn, zodat p niet alle waarden tussen 0 en 1 kan bezitten.

De gevonden meest aannemelijke schatters voor μ in voorbeeld 4.1 en voor p in voorbeeld 4.2 zijn, zoals in paragraaf 3 werd opgemerkt, meest doeltreffend, d.w.z. zij hebben van alle zuivere schatters de kleinste variantie. Dit hangt samen met een hier niet behandelde stelling, dat meest aannemelijke schatters onder vrij algemene voorwaarden asymptotisch voor $n \to \infty$ minimale varianties bezitten en normaal verdeeld zijn. In vele gevallen gelden deze eigenschappen ook voor elke eindige n (de normaliteit bijv. voor $\hat{\underline{\mu}}$ en de minimale variantie zowel voor $\hat{\underline{\mu}}$ als voor $\hat{\underline{p}}$). Het zijn o.a. deze eigenschappen die de meest aannemelijke schatters aantrekkelijk maken.

5. Kleinste kwadraten; regressie.

Wij beschouwen in deze paragraaf een stochastische grootheid $\underline{y}$, waarvan de verdeling afhangt van een andere, niet stochastische, grootheid x. Zo kan $\underline{y}$ bijv. de in een week door een koe geleverde hoeveelheid melk voorstellen (die van week tot week en van koe tot koe verschilt) en x de leeftijd (in jaren) van de koe. Of $\underline{y}$ kan de lengte van de remweg van een auto voorstellen en x de snelheid bij het begin

ın het remmen. In deze beide gevallen kan men x zelf kiezen - in het
ɜrste geval door een koe van de gewenste leeftijd uit te zoeken, in
ɜt tweede door bij de gewenste snelheid te gaan remmen -, terwijl $\underline{y}$
ıargenomen wordt.

In een dergelijk geval willen wıj onderzoeken, hoe de verdeling
ın $\underline{y}$ van x afhangt. Wij beschouwen hier alleen een eenvoudig geval;
.j onderstellen nl. dat de verwachting van $\underline{y}$ lineair van x afhangt

.1) $$\mathcal{E}(\underline{y} \mid x) = \alpha + \beta x.$$

.s wij ons voorstellen dat we een aantal x-waarden $x_1, x_2, \ldots, x_n$ kiezen
ı bij elke x_i de bijbehorende $\underline{y}_i$ waarnemen (i=1,2,...,n), dan zullen
.j in ons model de $\underline{y}_i$ weer beschouwen als onafhankelijke stochastische
'ootheden. Deze hebben nu <u>niet</u> dezelfde verdeling (hun verwachtingen
$\underline{y}_i = \mathcal{E}(\underline{y} \mid x_i) = \alpha + \beta x_i$ zijn bijv. al verschillend).

Als een onderzoeker nu een rij van n waarnemingsuitkomsten

.2) $$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

ɜeft gevonden, kan hij op de volgende wijze een schatting voor de
ıbekende parameters α en β van het lineaire verband (5.1) vinden. Men
.n de waarnemingen grafisch uitzetten, zoals in figuur 5.1 is gedaan.

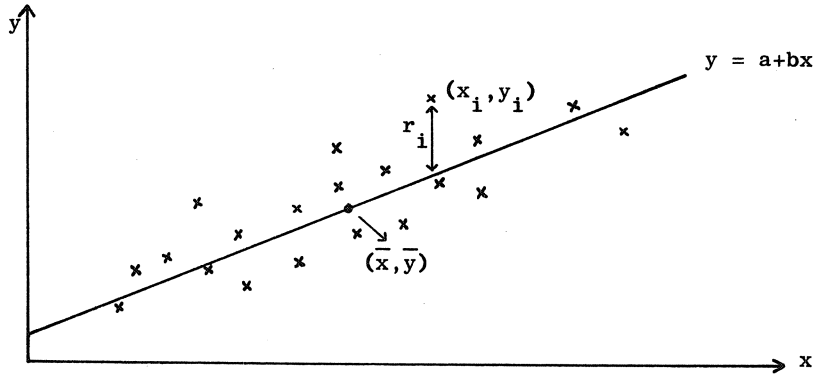


fig. 5.1

Lineaire regressie van y op x

De figuur geeft een aanwijzing dat $\underline{y}$, hoewel aan toevalsfluctuaties onderhevig, in het algemeen met x toeneemt; een lineair verband tussen x en $(\underline{y} \mid x)$ lijkt geen dwaze veronderstelling. Tekent men nu op het oog een lijn y = a+bx die goed bij de puntenwolk aansluit, dan behoort bij ieder punt (x_i, y_i) een vertikale afstand

$$r_i = |y_i - a - bx_i| \tag{5.3}$$

tot de lijn; deze afstand wordt residu genoemd. Men vindt nu schattingen voor α en β door a en b zo te kiezen, dat de som van de kwadraten van de residuen zo klein mogelijk wordt. Deze zg. methode van de kleinste kwadraten, al zeer lang bekend, wordt bijv. door natuur- en sterrenkundigen als een onderdeel van de foutentheorie toegepast op vele problemen.

Wij zoeken dus de waarden van a en b waarvoor

$$Q = \sum_{i=1}^{n} (y_i - a - bx_i)^2 \tag{5.4}$$

minimaal wordt en stellen daartoe $\frac{\partial Q}{\partial a}$ en $\frac{\partial Q}{\partial b}$ gelijk aan nul (deel 1, blz. 105):

$$-\frac{1}{2}\frac{\partial Q}{\partial a} = \sum_{i=1}^{n} (y_i - a - bx_i) = 0 \ , \tag{5.5}$$

$$-\frac{1}{2}\frac{\partial Q}{\partial b} = \sum_{i=1}^{n} x_i(y_i - a - bx_i) = 0 \ . \tag{5.6}$$

Als we schrijven $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ en $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, volgt uit (5.5)

$$a = \bar{y} - b\bar{x} \ . \tag{5.7}$$

Substitueren wij dit in (5.6), dan krijgen wij

$$\sum_{i=1}^{n} x_i \{ y_i - \bar{y} + b\bar{x} - bx_i \} = 0 \ , \tag{5.8}$$

is

.9) $$b = \frac{\sum_{i=1}^{n} x_i(y_i-\bar{y})}{\sum_{i=1}^{n} x_i(x_i-\bar{x})} .$$

ar $\sum_{i=1}^{n} \bar{x}(y_i-\bar{y})$, $\sum_{i=1}^{n} \bar{x}(x_i-\bar{x})$ en $\sum_{i=1}^{n} (x_i-\bar{x})\bar{y}$ nul zijn, kan men ook hrijven

.10) $$b = \frac{\sum_{i=1}^{n} (x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n} (x_i-\bar{x})^2} = \frac{\sum_{i=1}^{n} (x_i-\bar{x})y_i}{\sum_{i=1}^{n} (x_i-\bar{x})^2} .$$

or de tweede partiële afgeleiden van Q naar a en b te bepalen, kan n aantonen dat Q inderdaad minimaal wordt bij substitutie van de arden (5.7) en (5.10) voor a en b. Deze minimale waarde van Q zullen j Q_0 noemen. De optimale lijn die wij gevonden hebben wordt de gressielijn van $\underline{y}$ op x genoemd en heeft tot vergelijking

.11) $$y = a+ bx, \quad \text{d.w.z.} \quad y-\bar{y} = b(x-\bar{x}) ;$$

lijn gaat dus door het punt $(\bar{x},\bar{y})$. Bij een willekeurige x hebben wij voor de onbekende getalwaarde $\mathcal{E}(\underline{y} \mid x)$ de schatting a+bx.

Als we de $\underline{y}_i$ weer als stochastisch beschouwen (de x_i blijven tallen!), hebben we voor β en α de schatters

.12) $$\underline{b} = \frac{1}{\sum_{i=1}^{n} (x_i-\bar{x})^2} \sum_{i=1}^{n} (x_i-\bar{x})\underline{y}_i ,$$

sp.

.13) $$\underline{a} = \bar{y} - \underline{b}\bar{x} = \bar{y} - \frac{\bar{x}}{\sum_{i=1}^{n} (x_i-\bar{x})^2} \sum_{i=1}^{n} (x_i-\bar{x})\underline{y}_i .$$

ze zijn zuiver, want wegens $\mathcal{E}\underline{y}_i = \alpha+\beta x_i$ is

$$(5.14)\qquad \mathcal{E}\,\underline{b} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(\alpha+\beta x_i)}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \frac{\alpha\sum_{i=1}^{n}(x_i-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} + \frac{\beta\sum_{i=1}^{n}(x_i-\bar{x})x_i}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \beta$$

$$(5.15)\qquad \mathcal{E}\,\underline{a} = \mathcal{E}\,\underline{\bar{y}} - (\mathcal{E}\,\underline{b})\bar{x} = \frac{1}{n}\sum_{i=1}^{n}(\alpha+\beta x_i) - \beta\bar{x} = \alpha\ .$$

Vaak veronderstelt men niet alleen dat de verwachting van $\underline{y}$ bij gegeven x een lineaire functie van x is, maar tevens dat de variantie van $\underline{y}$ bij gegeven x een niet van x afhankelijke waarde σ^2 heeft. Men heeft nu drie onbekende parameters α, β en σ^2, waarvan α en β weer zuiver geschat worden door $\underline{a}$ en $\underline{b}$. Wij vermelden zonder bewijs, dat

$$(5.16)\qquad \underline{s}^2 \overset{\text{def}}{=} \frac{\underline{Q}_0}{n-2} = \frac{1}{n-2}\sum_{i=1}^{n}(\underline{y}_i-\underline{a}-\underline{b}x_i)^2$$

een zuivere schatter voor σ^2 is. Men kan verder aantonen dat $\underline{b}$ en $\underline{\bar{y}} = \underline{a}+\underline{b}\bar{x}$ ongecorreleerd zijn (covariantie nul hebben, zie deel 2 blz. 92-94) en dat $\underline{a}$ en $\underline{b}$ van alle schatters voor α en β die lineair zijn in de $\underline{y}_i$ de kleinste variantie hebben (zg. stelling van GAUSS-MARKOV). Deze varianties van $\underline{a}$ en $\underline{b}$ kunnen wij nu berekenen; volgens (5.12) is

$$(5.17)\qquad \underline{b} = \sum_{i=1}^{n} c_i\underline{y}_i \qquad \text{met} \quad c_i \overset{\text{def}}{=} \frac{(x_i-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\ .$$

Dus $\underline{b}$ is een lineaire combinatie van de onafhankelijke $\underline{y}_i$, zodat (deel 2, stelling 10.4 en 10.5)

$$(5.18)\qquad \sigma^2(\underline{b}) = \sum_{i=1}^{n} c_i^2\,\sigma^2(\underline{y}_i) = \sigma^2\sum_{i=1}^{n} c_i^2 = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\ .$$

Omdat $\sigma^2(\underline{\bar{y}}) = \frac{\sigma^2}{n}$ is en $\underline{\bar{y}}$ en $\underline{b}$ covariantie nul hebben, volgt uit (5.13) met de opmerking onder stelling 10.4 van deel 2:

5.19) $$\sigma^2(\underline{a}) = \sigma^2(\underline{\bar{y}}) + \bar{x}^2\sigma^2(\underline{b}) = \sigma^2\left\{\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right\}.$$

Tenslotte willen wij de nog sterkere onderstelling maken, dat $\underline{y}$ ij gegeven x de $N(\alpha+\beta x,\sigma)$-verdeling heeft. Onder deze eis van normali- eit kunnen wij de meest aannemelijke schatters voor α, β en σ^2 bepalen ij onafhankelijke waarnemingen $(x_i,\underline{y}_i)$ (i=1,2,...,n). De verdelings- ichtheid van $\underline{y}_i$ is

5.20) $$f(y_i \mid x_i) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left\{\frac{y_i-(\alpha+\beta x_i)}{\sigma}\right\}^2}.$$

egens de onafhankelijkheid is

5.21) $$L(\alpha,\beta,\sigma^2) = \prod_{i=1}^{n} f(y_i \mid x_i)\ ;$$

5.22) $$\log L = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\alpha-\beta x_i)^2 .$$

n de schattingen $\hat{\alpha}$ en $\hat{\beta}$ te vinden die log L bij vaste σ^2 maximaliseren, oeten wij blijkbaar $Q = \sum_{i=1}^{n}(y_i-\alpha-\beta x_i)^2$ minimaliseren als functie van en β. Maar dit houdt in dat de meest aannemelijke schatters juist e kleinste-kwadratenschatters zijn: $\hat{\underline{\alpha}} = \underline{a}$ en $\hat{\underline{\beta}} = \underline{b}$, gegeven door 5.13) en (5.12). De meest aannemelijke schatter voor de parameter $E(\underline{y} \mid x)$ bij een gegeven waarde x is nu

5.23) $$\underline{Y}_x = \hat{\underline{\alpha}} + \hat{\underline{\beta}}x .$$

ij substitutie van $\hat{\alpha}$ en $\hat{\beta}$ in (5.22) gaat Q over in Q_0, dus

5.24) $$\log L(\sigma^2) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{Q_0}{2\sigma^2} .$$

e afgeleide naar σ^2 is $-\frac{n}{2\sigma^2} + \frac{Q_0}{2\sigma^4}$, zodat

(5.25) $$\hat{\underline{\sigma}}^2 = \frac{\underline{Q}_0}{n}$$

de meest aannemelijke schatter voor σ^2 is. Deze is niet zuiver maar heeft verwachting $\frac{n-2}{n}\sigma^2$; een zuivere schatter $\underline{s}^2$ voor σ^2 is al genoemd in formule (5.16).

Als lineaire combinaties van normaal verdeelde $\underline{y}_i$ zijn de schatters $\underline{a}$ en $\underline{b}$ normaal verdeeld; zij zijn ook asymptotisch raak. Omdat $\bar{\underline{y}}$ en $\underline{b}$ nu normaal verdeeld zijn en covariantie nul hebben zijn zij onafhankelijk (deel 2, stelling 10.10).

Wij merken tenslotte nog op dat er ook lineaire modellen bestaan waarbij beide variabelen $\underline{x}$ en $\underline{y}$ stochastisch zijn. Als de voorwaardelijke verdeling van $\underline{y}$ bij gegeven x voldoet aan de genoemde eisen kan men de theorie vrijwel onveranderd toepassen. Men zou nu ook de regressielijn van $\underline{x}$ op $\underline{y}$ kunnen bepalen, die voor $|\rho(\underline{x},\underline{y})| < 1$ niet met de regressielijn van $\underline{y}$ op $\underline{x}$ samenvalt. In het speciale geval van een twee-dimensionale normale verdeling voor $\underline{x}$ en $\underline{y}$ kan men laten zien dat $\sigma^2(\underline{y} \mid \underline{x}=x) = (1-\rho^2)\sigma^2(\underline{y})$; de onvoorwaardelijke verdeling van $\underline{y}$ heeft dus een grotere variantie dan de voorwaardelijke bij gegeven x. Het splitsen van de onvoorwaardelijke variantie in een variantie "om de lijn" en een verdere term wordt onder veel algemenere omstandigheden bestudeerd in de hier niet behandelde variantieanalyse.

6. De momentenmethode.

De oudste methode om onbekende parameters van een verdeling te schatten is de zg. momentenmethode, welke afkomstig is van K. PEARSON.

PEARSON stelde voor de onbekende parameters zodanig te kiezen, dat zoveel mogelijk momenten in de steekproef dezelfde waarde bezitten als in de verdeling. Men zorgt er dan eerst voor dat de eerste momenten (verwachtingen) gelijk zijn, vervolgens indien mogelijk de tweede gereduceerde momenten (varianties), dan indien mogelijk de derde momenten, enz. Bij een normale verdeling, die twee onbekende parameters

ieeft (μ en σ^2) worden dus twee momenten gelijk gemaakt: de verwachting
:n de variantie. Men eist

6.1) $$\frac{1}{n}\sum_{i=1}^{n} x_i = \mu$$

:n

6.2) $$\frac{1}{n}\sum_{i=1}^{n} (x_i-\bar{x})^2 = \sigma^2$$

:n geeft dus als schatter voor μ op $\bar{\underline{x}} = \frac{1}{n}\sum_{i=1}^{n} \underline{x}_i$ en als schatter voor
^{2}: $\frac{1}{n}\sum_{i=1}^{n}(\underline{x}_i-\bar{\underline{x}})^2$. Hier leidt deze methode dus tot de reeds eerder
evonden schatters.

In het algemeen behoeft dit niet het geval te zijn. Een belangrijk
oordeel van deze schattingsmethode is dat de schatters vaak veel
envoudiger berekend kunnen worden dan de meest aannemelijke schatters.
uist om deze reden wordt ze meestal toegepast bij gamma-verdelingen.
tellen wij de gamma-verdeling weer voor door

6.3) $$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} e^{-\lambda x} x^{\alpha-1}$$ (vgl. (7.25) van deel 2),

an is het eerste moment

6.4) $$\mathcal{E}\underline{x} = \frac{\alpha}{\lambda}$$ (vgl. (10.9) van deel 2)

n de variantie

6.5) $$\sigma^2(\underline{x}) = \frac{\alpha}{\lambda^2}$$ (vgl.(10.25) van deel 2).

Wanneer de steekproefuitkomst $x_1,\ldots,x_n$ gegeven is en dus het
ekenkundig gemiddelde $\frac{1}{n}\sum_{i=1}^{n} x_i$ en de variantie $\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$ in de
teekproef bekend zijn, dan worden op eenvoudige wijze schattingen
oor α en λ gevonden door deze waarden gelijk te stellen aan (6.4)
esp. (6.5). De schattingen luiden dan

(6.6) $$\frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} (x_i-\bar{x})^2} \quad \text{voor } \lambda$$

en

(6.7) $$\frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n \sum_{i=1}^{n} (x_i-\bar{x})^2} \quad \text{voor } \alpha .$$

De meest aannemelijke schattingen voor λ en α kunnen in dit geval niet expliciet opgeschreven worden. Wij zullen daarom in zo'n geval gebruik maken van schattingen (6.6) en (6.7).

Tegenover het genoemde voordeel van de hier behandelde schattingsmethode staat het nadeel dat de schatters in het algemeen bij gelijke steekproefomvang een grotere variantie bezitten dan de meest aannemelijke schatters en dus onnauwkeuriger zijn; dit blijft vaak gelden wanneer de steekproefomvang willekeurig toeneemt. Wanneer de verdeling in kwestie drie of meer parameters heeft, is het bovendien bezwaarlijk dat derde en hogere momenten van de steekproef pas bij zeer grote steekproefomvang weinig afwijken van de momenten van de verdeling.

7. Inleiding tot de toetsingstheorie.

Onder de elementaire toepassingen van de statistiek neemt het toetsen van hypothesen de belangrijkste plaats in, hoewel ook het schatten van parameters zeer vaak voorkomt. Nadat men een mathematisch-statistisch model voor een experimentele situatie heeft opgesteld, kan men binnen dit model een veronderstelling of vermoeden formuleren en nagaan of de waarnemingen deze veronderstelling, die we de nulhypothese zullen noemen en met H_0 zullen aangeven, bevestigen of tegenspreken. Zodra de verschijnselen niet deterministisch zijn, is van een volmaakt

)evestigen of weerleggen geen sprake; de toetsingstheorie onderzoekt 'ooral in welke mate de waarnemingen en de nulhypothese met elkaar 'erenigbaar zijn.

Wij geven enige voorbeelden:

Model	Nulhypothese H_0
A) Er is een grootheid gemeten aan aselect gekozen objecten.	De verdeling van de gemeten grootheid is exponentieel.
B) (x_i, y_i) $(i=1,\ldots,n)$ zijn waarnemingen van twee kenmerken aan aselect gekozen elementen.	$\underline{x}$ en $\underline{y}$ zijn stochastisch onafhankelijk.
C) De gemeten grootheid is in elk van twee gegeven populaties normaal verdeeld.	De verwachting is in beide populaties even groot.
D) De experimenten zijn onafhankelijk en hebben dezelfde kans p op succes.	$p = \frac{1}{4}$.

Bij het confronteren van H_0 en de waarnemingen zijn er nu twee conclusies mogelijk:

Conclusie I: Als H_0 juist is, is het zeer onwaarschijnlijk dat de gevonden waarnemingen zouden optreden. Bij deze conclusie verwerpt men H_0; wat dan wel een juiste hypothese is, valt nader te bezien. Soms geeft de toets hierover duidelijke aanwijzingen, soms niet.

Conclusie II: Als H_0 juist is, zijn de gevonden waarnemingen niet zo onwaarschijnlijk. Bij deze conclusie is H_0 verenigbaar met de waarnemingen en wordt H_0 niet verworpen. Dit wil nog niet zeggen dat H_0 nu bewezen is en zonder meer aanvaard moet worden; wij weten alleen dat onze waarnemingen H_0 niet tegenspreken.

Bij de nu volgende bespreking van de toetsingstheorie zullen wij ons meestal beperken tot nulhypothesen van het soort genoemd in voorbeeld D. Wij gaan uit van een steekproef uit een verdeling met parameter θ, en beschouwen de nulhypothese $\theta = \theta_0$, waarin θ_0 een gegeven getal is. Er blijkt dan een nauw verband met de schattingstheorie te bestaan.

Immers als een schatter $\underline{t}$ voor θ een kansverdeling heeft die dicht om de ware waarde geconcentreerd ligt en een gegeven waarnemingsreeks geeft een waarde $t(x_1,x_2,\ldots,x_n)$ die sterk van θ_0 afwijkt, dan is het intuïtief duidelijk dat men tot verwerping van θ_0 zal moeten overgaan. Vindt men bijv. voor een steekproef van 100 lampen een gemiddelde levensduur van 700 uur met een geschatte standaardafwijking van 100, dan is de waarde $\theta_0 = 2000$ voor de verwachting van de levensduur niet acceptabel.

Bij de toepassingen is er echter een duidelijk verschil tussen toetsen en schatten; een goede formulering van het probleem en het doel van het onderzoek leidt meestal tot een duidelijke keuze tussen beide methoden. Vergelijkt men bijv. twee geneesmiddelen (een klassiek probleem) wat betreft hun effekt bij de bestrijding van een bepaalde ziekte, dan zal men het beste van de twee middelen gaan gebruiken zodra men de onderstelling van gelijkwaardigheid van deze middelen verworpen heeft. Toepassing van de toetsingstheorie is hier dan de aangewezen weg. Wenst men daarentegen de premie voor een bepaalde verzekering te berekenen, dan heeft men een kwantitatieve schatting van het risico nodig en is het niet genoeg (of zelfs niet van belang) om te weten dat dit risico groter is dan een bepaald bedrag of groter dan het risico dat aan een andere situatie verbonden is.

Bij de toetsingstheorie wordt een aantal nieuwe begrippen gebruikt, die wij in algemene bewoordingen zullen definiëren en daarnaast zullen toelichten aan de hand van voorbeeld D (de zg. binomiale toets). Terwille van de overzichtelijkheid zullen wij niet de meest algemene formulering van de toetsingstheorie geven; wij zullen in de loop van ons betoog een aantal onderstellingen invoeren, die in een volkomen algemene opzet niet thuis horen, maar waar men in de praktijk vaak van uitgaat.

. Grondbegrippen van de toetsingstheorie.

et waarnemingsmateriaal

Het waarnemingsmateriaal, dat uit één of meer waarnemingsreeksen an bestaan, zullen we aangeven met

8.1) $$\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_n \ .$$

et is niet noodzakelijk dat alle $\underline{w}_i$ onafhankelijk zijn en evenmin dat ij dezelfde verdeling bezitten; vaak zullen wij dit echter wel veron- erstellen. De grootheden $\underline{w}_i$ kunnen meer-dimensionaal zijn; elke $\underline{w}_i$ kan ijv. een paar waarnemingen $(\underline{x}_i, \underline{y}_i)$ voorstellen.

oegelaten hypothesen en nulhypothese

Als de kansverdelingen der $\underline{w}_i$ volledig bekend waren, zou er geen oetsingsprobleem bestaan. In de regel is er over deze kansverdelingen el ièts bekend, d.w.z. er worden bepaalde onderstellingen over gemaakt, ie als gegeven worden beschouwd. Hieronder valt meestal de onderstelling in onafhankelijkheid van de waarnemingen die we boven reeds vermeldden verder ook een onderstelling omtrent de vorm van de verdelingen. Zo rusten veel toetsen op de veronderstelling van normaliteit van de rdelingen. Wij zullen het geval beschouwen dat de verdelingen der $\underline{w}_i$ g van één onbekende parameter afhangen, maar overigens volledig bekend jn *). Geven wij de onbekende parameter aan met het symbool θ, dan nnen we de simultane verdelingsfunctie van $\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_n$ voorstellen or

8.2) $$F(w_1, w_2, \ldots, w_n \mid \theta) \ .$$

Als wij aan θ een bepaalde waarde geven is F geheel vastgelegd en kend. Zo'n hypothese omtrent θ die F volledig bepaalt wordt een enkel- udige hypothese genoemd; een hypothese waarbij men alleen onderstelt t θ in een zeker interval ligt heet samengestelde hypothese. Welke

---

paragraaf 12 zullen wij enige voorbeelden bespreken waarbij niet veron- rsteld wordt dat de vorm van F bekend is (zg. verdelingsvrije toetsen).

waarden we voor θ mogelijk achten, hangt af van de aard van de parameter θ. Is θ een kans dan kan θ alleen waarden tussen 0 en 1 aannemen. Stelt θ de parameter λ van de gamma-verdeling voor (pag. 55 van deel 2), dan moet $\theta > 0$ zijn. Is θ de verwachting van een normale verdeling, dan zijn alle waarden $-\infty < \theta < +\infty$ mogelijk. Bij iedere mogelijke waarde van θ hoort dus een enkelvoudige hypothese en de verzameling van al deze hypothesen wordt de verzameling van toegelaten hypothesen genoemd.

Wij beschouwen meestal als nulhypothese H_0: $\theta = \theta_0$. De toetsingstheorie is ook toepasbaar op samengestelde nulhypothesen, zoals $\theta \leq \theta_0$ of $\theta_0' \leq \theta \leq \theta_0''$.

Zoals in paragraaf 7 al werd opgemerkt, zal men bij sommige waarden van de waarnemingen de nulhypothese verwerpen. Bij verwerpen van $\theta = \theta_0$ luidt de conclusie uiteraard $\theta \neq \theta_0$; bij verwerpen van $\theta \leq \theta_0$ is de conclusie $\theta > \theta_0$. De hypothesen die aanvaard worden als H_0 verworpen wordt, worden de alternatieve hypothesen genoemd.

Voorbeeld 8.1

Fraaie voorbeelden van het optreden van kansen met bekende waarden vindt men in de erfelijkheidsleer. Aan de proeven van MENDEL ligt een erfelijkheidstheoretisch model ten grondslag, dat bijv. leidt tot de uitkomst, dat bij het kruisen van bepaalde planten één van de twee verschijnselen A en B zich voordoet en dat de kans op A bij iedere proef een bekende waarde p_0 (bijv. $\frac{1}{4}$) bezit. De proeven van MENDEL dienden nu om te onderzoeken of de waarde van zo'n kans, en daarmee de ten grondslag liggende erfelijkheidstheorie, juist kon zijn. De verschijnselen A en B kunnen bijv. twee kleuren (van erwtenbloesems) zijn of iets dergelijks. De n waarnemingen, verkregen bij n onafhankelijke proeven van deze aard, nemen in dit geval de "waarden" A of B aan; men kan dit in cijfers omzetten door het aantal malen A bij het i^{de} experiment w_i te noemen. Elke $\underline{w}_i$ is dus 1 met kans $p = P[A]$ en 0 met kans $q = 1-p$; de som

$$\underline{x} = \sum_{i=1}^{n} \underline{w}_i \tag{8.3}$$

heeft een binomiale verdeling met parameters n en p (vgl. deel 2, voorbeeld 8.7):

3.4) $$P\left[\underline{x}=x \mid p\right] = \binom{n}{x} p^x q^{n-x} \qquad (x=0,1,\ldots,n) \ .$$

In dit voorbeeld speelt p de rol van de onbekende parameter θ
iet aantal experimenten n is uiteraard bekend). De verzameling van
oegelaten hypothesen is bepaald door de eis dat p een kans voorstelt:
$\leq p \leq 1$. De nulhypothese is hier enkelvoudig, nl. $p = p_0$ (waarbij p_0
en getal tussen 0 en 1 is); wij zullen speciaal de nulhypothese H_0:
$= \frac{1}{4}$ toetsen. Verwerping van H_0 leidt tot de conclusie $p \neq \frac{1}{4}$. In
raktische bewoordingen: de erfelijkheidstheorie, die tot de waarde
$= \frac{1}{4}$ leidde, is niet juist en moet door een andere vervangen worden.
e nieuwe hypothese zal dan aan hernieuwde proefnemingen getoetst moeten
orden.

e toetsingsgrootheid

De beoordeling van de te toetsen hypothese geschiedt met behulp
an een toetsingsgrootheid, die uit de waarnemingen berekend wordt:

$$\underline{t} = t(\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_n) \ .$$

eze grootheid kan in principe meer-dimensionaal zijn, maar is gewoon-
ijk één-dimensionaal. Meestal is $\underline{t}$ een schatter van de onbekende
rameter θ, of - wat hiermee equivalent is - een monotone functie van
en schatter van θ. Geven de waarnemingen nu een getalwaarde
$= t(w_1, w_2, \ldots, w_n)$ die veel van de te toetsen waarde θ_0 afwijkt, dan
al men de nulhypothese verwerpen. Bij het beoordelen van de mate van
fwijking baseert men zich op de verdelingsfunctie $F(t \mid \theta_0)$ van de
oetsingsgrootheid $\underline{t}$ als θ_0 de parameterwaarde is. Uit de verdelings-
nctie $F(w_1, w_2, \ldots, w_n \mid \theta_0)$ valt deze verdelingsfunctie van $\underline{t}$ af te
eiden.

oorbeeld 8.1 (vervolg)

Bij de proeven van MENDEL is

3.5) $$\hat{\underline{p}} = \frac{\underline{x}}{n} = \frac{1}{n} \sum_{i=1}^{n} \underline{w}_i$$

de meest aannemelijke schatter voor de onbekende kans p. Bij het toetsen van H_0: $p = \frac{1}{4}$ zullen wij

(8.6) $$\underline{t} = \underline{x} = n\underline{\hat{p}}$$

als toetsingsgrootheid gebruiken; de verdeling van $\underline{t}$ onder H_0 is

(8.7) $$P\left[\underline{x}=x \mid p=\frac{1}{4}\right] = \binom{n}{x}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{n-x} \qquad (x=0,1,\ldots,n) .$$

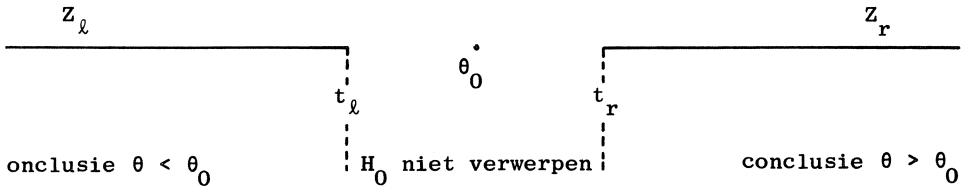
De kritieke zone

Ter uitvoering van een toets gaan wij nu preciseren, welke waarden van de toetsingsgrootheid $\underline{t}$ onder H_0 zó onwaarschijnlijk zijn, dat ze tot verwerping van H_0 zullen leiden. Hiertoe stelt men eerst een onbetrouwbaarheidsdrempel α vast (een traditionele keuze is $\alpha = 0{,}05$ of $0{,}01$). In de verzameling van mogelijke waarden voor $\underline{t}$ kiest men nu een zg. kritieke zone Z van t-waarden waarbij men H_0 zal verwerpen, en wel zodanig dat $P\left[\underline{t} \in Z \mid H_0\right] \leq \alpha$ is.

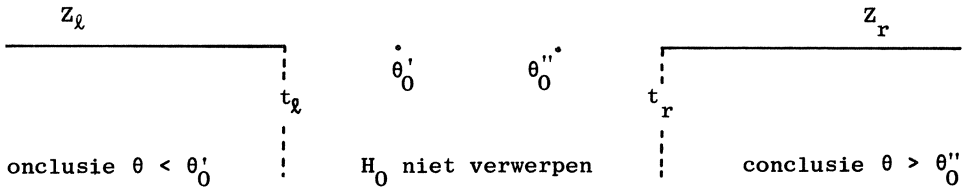
Gemakshalve nemen wij even aan dat $\underline{t}$ een schatter voor θ is. Voor nulhypothesen van de vorm $\theta = \theta_0$ kiest men de kritieke zone Z tweezijdig d.w.z. Z valt uiteen in een rechter kritieke zone $Z_r = [t_r,\infty)$ en een linker kritieke zone $Z_\ell = (-\infty,t_\ell]$. Op de keuze van de zg. kritieke waarden t_ℓ en t_r komen wij nog terug. Men toetst in dit geval tweezijdig zowel zeer grote als zeer kleine waarden van $\underline{t}$ spreken de nulhypothese tegen. Als $\underline{t}$ een waarde t aanneemt met $t \geq t_r$ zal men H_0: $\theta=\theta_0$ verwerpen ten gunste van $\theta > \theta_0$, terwijl voor $t \leq t_\ell$ de conclusie zal zijn $\theta < \theta_0$. Ook H_0: $\theta_0' \leq \theta \leq \theta_0''$ wordt tweezijdig getoetst.

Toetst men H_0: $\theta \geq \theta_0$, dan ligt het voor de hand eenzijdig te toetsen en uitsluitend H_0 te verwerpen als $t \leq t_\ell$ is (linkseenzijdige toets). Evenzo toetst men $\theta \leq \theta_0$ rechtseenzijdig met een kritieke zone $t \geq t_r$. Wij verwijzen verder naar figuur 8.1, waar voor de vier genoemde nulhypothesen de kritieke zones zijn geschetst, met de conclusies die men trekt als $\underline{t}$ in de diverse gebieden valt.

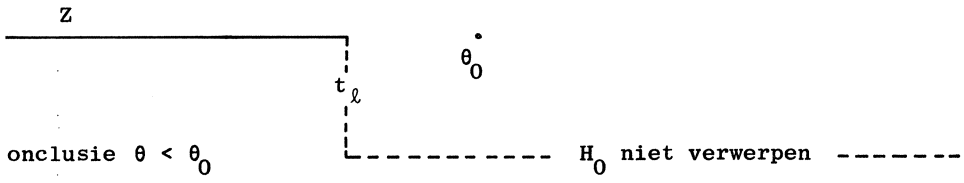
H_0: $\theta = \theta_0$ (tweezijdig)



H_0: $\theta'_0 \leq \theta \leq \theta''_0$ (tweezijdig)



H_0: $\theta \geq \theta_0$ (linkseenzijdig)



H_0: $\theta \leq \theta_0$ (rechtseenzijdig)

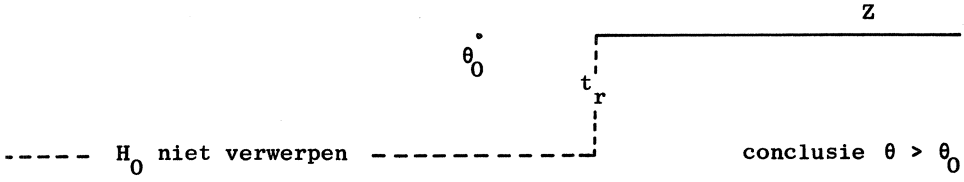


fig. 8.1

Kritieke zones bij diverse nulhypothesen

Bij de meeste problemen is de keuze tussen tweezijdige en eenzijdige toetsing niet moeilijk. Zo toetst men bij de proeven van MENDEL $p = \frac{1}{4}$. Worden nu zeer weinig proeven met A als uitkomst verkregen (dus heeft $\underline{x}$ een kleine waarde), dan wordt $p = \frac{1}{4}$ verworpen ten gunste van $p < \frac{1}{4}$; worden er echter zeer veel gevonden dan concludeert men $p > \frac{1}{4}$. Wij gebruiken hier dus een tweezijdige toets. In het volgende voorbeeld blijkt eenzijdige toetsing wenselijk te zijn.

Voorbeeld 8.2

Een bedrijf overweegt de aanschaf van een nieuwe machine, maar men wil hiertoe alleen overgaan indien de nieuwe machine beter is dan de oude. Van de oude machine is bekend dat deze een fractie p_0 aan uitval geeft. Om tot een beslissing te komen toetst men de hypothese H_0: $p \geq p_0$ (waarin p de fractie uitval van de nieuwe machine voorstelt). Alleen bij verwerping van H_0 gaat men tot aanschaf van de nieuwe machine over. Een tweezijdige toets zou hier niet op zijn plaats zijn, daar verwerping van H_0: $p = p_0$ ten gunste van $p > p_0$ evenmin tot aanschaf van de nieuwe machine leidt als niet-verwerpen van H_0.

De onbetrouwbaarheid

Indien men op de boven beschreven wijze te werk gaat is het blijkbaar niet uitgesloten, dat H_0 verworpen wordt hoewel deze hypothese juist was. Deze foute uitkomst (H_0 ten onrechte verwerpen) wordt een fout van de eerste soort genoemd. Er is meestal slechts één methode om deze fout met zekerheid uit te sluiten en dat is: H_0 nooit verwerpen. Immers zodra men een niet-lege kritieke zone Z gebruikt is er in de meeste problemen, ook als H_0 juist is, een positieve kans dat $\underline{t}$ in Z valt en H_0 verworpen wordt.

Wij voeren in bij een enkelvoudige nulhypothese H_0:

de linkeronbetrouwbaarheid $\alpha_\ell \overset{\text{def}}{=} P\left[\underline{t} \in Z_\ell \mid H_0\right] = P\left[\underline{t} \leq t_\ell \mid H_0\right]$,

de rechteronbetrouwbaarheid $\alpha_r \overset{\text{def}}{=} P\left[\underline{t} \in Z_r \mid H_0\right] = P\left[\underline{t} \geq t_r \mid H_0\right]$,

de onbetrouwbaarheid $\alpha_0 \overset{\text{def}}{=} P\left[\underline{t} \in Z \mid H_0\right] = \alpha_\ell + \alpha_r$.

Wij herinneren er aan dat het kritieke gebied Z zo gekozen is, dat e onbetrouwbaarheid α_0 hoogstens gelijk is aan de onbetrouwbaarheids-rempel α. Wij kunnen de kans op een fout van de eerste soort door het aststellen van een kleine waarde voor α zeer gering maken; als zo'n out ernstige gevolgen heeft kiest men wel $\alpha = 0{,}001$. Bij de bespreking an het onderscheidingsvermogen zal blijken dat zo'n kleine waarde van het bezwaar heeft dat H_0 vrijwel nooit meer verworpen wordt, ook niet ls H_0 wel degelijk onjuist is.

Bij een tweezijdige toets mag men α_ℓ en α_r willekeurig kiezen, its hun som hoogstens α is. Men noemt een toets <u>symmetrisch tweezijdig</u>, ls voor t_ℓ en t_r de grootste resp. kleinste waarde van t genomen wordt, aarvoor α_ℓ en α_r beide $\leq \frac{1}{2}\alpha$ zijn.

Het onderscheid tussen α_0 en α lijkt overbodig; waarom niet $\alpha_0 = \alpha$ emaakt? Inderdaad doet men dit steeds wanneer het mogelijk is. Maar ls de toetsingsgrootheid onder H_0 een discrete verdeling heeft, moet en ermee tevreden zijn α_0 zo dicht bij α (maar kleiner dan α) te iezen als deze verdeling toestaat.

Bij een meervoudige nulhypothese, bijv. $\theta \geq \theta_0$, definieert men de inkeronbetrouwbaarheid als $\alpha_\ell = \max_{\theta \geq \theta_0} P\left[\underline{t} \in Z_\ell \mid \theta\right]$; analoog voor r en α_0.

oorbeeld 8.1 (vervolg)

Bij de proeven van MENDEL hadden wij als toetsingsgrootheid voor $_0$: $p = \frac{1}{4}$ het aantal malen $\underline{x}$ dat verschijnsel A (succes) is opgetreden. oor 50 proeven is de verdeling van $\underline{x}$ onder H_0 in tabel 8.1 in 4 deci-alen gegeven, voorzover de kansen niet verwaarloosbaar klein zijn.

Bij de keuze $\alpha = 0{,}05$ kunnen wij nu voor de linker kritieke zone $_\ell$ nemen $x \leq 6$ en voor de rechter kritieke zone Z_r $x \geq 19$; dan is olgens de tabel

$$\alpha_0 = \alpha_\ell + \alpha_r = P\left[\underline{x} \leq 6 \mid H_0\right] + P\left[\underline{x} \geq 19 \mid H_0\right] =$$

$$= 0{,}0194 + 0{,}0287 = 0{,}0481 < 0{,}05 \ .$$

Tabel 8.1

Binomiale verdeling met $n = 50$ en $p = \frac{1}{4}$

x	$P[\underline{x} = x]$	$k_\ell = F(x)$	$k_r = 1 - F(x-1)$
0	≈ 0	≈ 0	1
1	≈ 0	≈ 0	≈ 1
2	0,0001	0,0001	≈ 1
3	0,0004	0,0005	0,9999
4	0,0016	0,0021	0,9995
5	0,0049	0,0070	0,9979
6	0,0123	0,0194	0,9930
7	0,0259	0,0452	0,9806
8	0,0463	0,0916	0,9548
9	0,0721	0,1637	0,9084
10	0,0985	0,2622	0,8363
11	0,1194	0,3816	0,7378
12	0,1294	0,5110	0,6184
13	0,1261	0,6370	0,4890
14	0,1110	0,7481	0,3630
15	0,0888	0,8369	0,2519
16	0,0648	0,9017	0,1631
17	0,0432	0,9449	0,0983
18	0,0264	0,9713	0,0551
19	0,0148	0,9861	0,0287
20	0,0077	0,9937	0,0139
21	0,0036	0,9974	0,0063
22	0,0016	0,9990	0,0026
23	0,0006	0,9996	0,0010
24	0,0002	0,9999	0,0004
25	0,0001	≈ 1	0,0001
26,...,50	≈ 0	≈ 1	≈ 0

(Deze tabel is ontleend aan H.G. ROMIG, 50-100 Binomial Tables, Wiley, N.Y.& Chapman and Hall, London, 1953).

eze kritieke zone is voor $\alpha = 0{,}05$ de meest geschikte, als men beide elen ongeveer $\frac{1}{2}\alpha = 0{,}025$ tot de onbetrouwbaarheid wil laten bijdragen. oor $\alpha = 0{,}01$ vindt men bij $t_\ell = 5$ en $t_r = 22$ een onbetrouwbaarheid $_0 = 0{,}0070 + 0{,}0026 = 0{,}0096$. Men kan ook nemen $t_\ell = 4$ en $t_r = 21$, an is $\alpha_0 = 0{,}0021 + 0{,}0063 = 0{,}0084$. Er is geen algemeen aanvaarde egel om in dergelijke gevallen tussen de verschillende mogelijkheden e kiezen. Bij een continu verdeelde toetsingsgrootheid doet deze oeilijkheid zich niet voor, omdat men daar α_ℓ en α_r beide exact elijk aan $\frac{1}{2}\alpha$ kan maken.

e overschrijdingskans

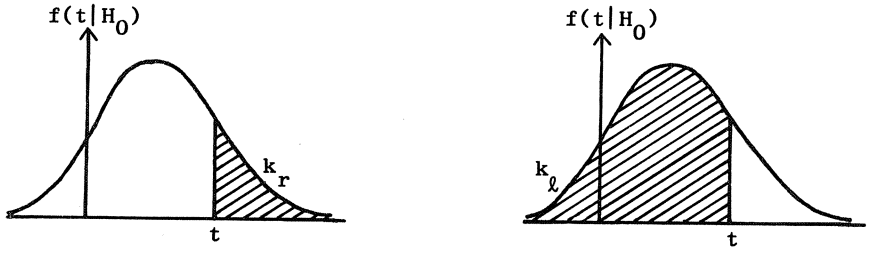
Bij een waarde t van de toetsingsgrootheid definiëren wij:

e linker overschrijdingskans $\quad k_\ell(t) \stackrel{def}{=} P\left[\underline{t} \leq t \mid H_0\right]$,

e rechter overschrijdingskans $\quad k_r(t) \stackrel{def}{=} P\left[\underline{t} \geq t \mid H_0\right]$,

e tweezijdige overschrijdingskans $\quad k(t) \stackrel{def}{=} 2 \min\{k_\ell(t), k_r(t)\}$.

In de meetkundige voorstelling van de verdelingsdichtheid van $\underline{t}$ nder H_0 (zie figuur 8.2) komen deze kansen overeen met de gearceerde ppervlakken.



rechtseenzijdig linkseenzijdig

fig. 8.2

Overschrijdingskansen bij gevonden waarde t van $\underline{t}$

Als $k_\ell(t) \leq \alpha_\ell$ dan ligt t in Z_ℓ; als $k_r(t) \leq \alpha_r$ dan ligt t in Z_r. Als $k(t) \leq \alpha_0$ en de toets is symmetrisch tweezijdig, dan ligt t in Z.

De overschrijdingskans is van belang omdat men er niet alleen aan kan zien of de gevonden waarde van t in de kritieke zone van de toets ligt, maar ook of dit maar juist of ruimschoots het geval is. Hoe kleiner de gevonden overschrijdingskans is, met des te meer overtuiging zal men H_0 verwerpen.

Voor vele bekende toetsen bestaan tabellen voor t_ℓ en t_r bij de gangbare onbetrouwbaarheidsdrempels $\alpha = 0{,}05$, $\alpha = 0{,}025$ en $\alpha = 0{,}01$. Als men de toetsingsgrootheid berekend heeft, kan men in zo'n tabel zien of H_0 verworpen moet worden. Wanneer de gehele verdelingsfunctie van $\underline{t}$ onder H_0 getabelleerd is, kan men de overschrijdingskans opzoeken. Zo vindt men bij voorbeeld 8.1 uit tabel 8.1 bijv. $k_\ell(3) = 0{,}0005$; $k_r(21) = 0{,}0063$; $k(4) = 0{,}0042$. Voor toetsen die herhaaldelijk op grote aantallen waarnemingen worden losgelaten gebruikt men vaak een elektronische rekenmachine; men kan een machineprogramma schrijven waarbij zowel de toetsingsgrootheid als de overschrijdingskans worden berekend.

Het onderscheidingsvermogen

Wij beschouwen de enkelvoudige nulhypothese H_0: $\theta = \theta_0$. De functie

$$\pi(\theta) \stackrel{\text{def}}{=} P\left[\underline{t} \in Z \mid \theta\right] \tag{8.8}$$

wordt het onderscheidingsvermogen (Engels: power) van de toets genoemd. Krachtens zijn definitie is $\alpha_0 = \pi(\theta_0)$. Voor $\theta_1 \neq \theta_0$ is $\pi(\theta_1)$ de kans dat H_0 verworpen wordt als θ_1 de ware waarde van de parameter is (dus H_0 inderdaad verworpen diende te worden). Dit is om zo te zeggen de kans, dat de toets θ_1 van θ_0 zal weten te onderscheiden.

Het ten onrechte verwerpen van H_0 als H_0 juist is - de fout van de eerste soort - vindt nu zijn tegenhanger in het ten onrechte niet verwerpen van H_0 als H_0 onjuist is; dit heet een fout van de tweede soort. Een fout van de eerste soort kan men dus alleen maken als H_0 juist is, en een fout van de tweede soort uitsluitend als H_0 onjuist is, d.w.z. als de parameter θ een waarde $\theta_1 \neq \theta_0$ heeft. De kans op een fout van de tweede soort is dan $1-\pi(\theta_1)$, omdat H_0 met een kans $\pi(\theta_1)$ zal worden verworpen.

Wij verkeren nu in een onaangename situatie; het verkleinen van
e kans op een fout van de eerste soort komt neer op het kiezen van
en kleinere kritieke zone Z, waardoor $\pi(\theta)$ eveneens kleiner wordt en
e kans op een fout van de tweede soort toeneemt. Men zoekt dus naar
en compromis waarbij beide fouten een niet te grote kans hebben;
eestal fixeert men $\alpha = 0{,}05$ en neemt er genoegen mee dat het onder-
cheidingsvermogen dan redelijk goed is voor θ-waarden die niet te
icht bij de getoetste waarde θ_0 liggen. Beide kansen op fouten gelijk-
ijdig verminderen kan men alleen door een betere toetsingsgrootheid
e gebruiken (als die bestaat) of door de toets op meer waarnemingen
e baseren.

Bij een tweezijdige toets voeren wij nog in:

8.9) $$\pi_\ell(\theta) \stackrel{\text{def}}{=} P\left[\underline{t} \in Z_\ell \mid \theta\right] \quad \text{en} \quad \pi_r(\theta) \stackrel{\text{def}}{=} P\left[\underline{t} \in Z_r \mid \theta\right],$$

et linker- en rechteronderscheidingsvermogen. Uiteraard is
$(\theta) = \pi_\ell(\theta) + \pi_r(\theta)$.

oorbeeld 8.1 (vervolg)

Het onderscheidingsvermogen van een binomiale toets - een functie
an de onbekende kans p - kan men nu opzoeken in een tabel van de
inomiale verdeling. Voor het toetsen van $p = \frac{1}{4}$ uit 50 experimenten,
et onbetrouwbaarheidsdrempel 0,05 vonden wij Z_ℓ: $x \leq 6$ met $\alpha_\ell = 0{,}019$
n Z_r: $x \geq 19$ met $\alpha_r = 0{,}029$. Wij moeten dan in de tabel opzoeken

$$_\ell(p) = P\left[\underline{x} \leq 6 \mid p\right] \quad \text{en} \quad \pi_r(p) = P\left[\underline{x} \geq 19 \mid p\right] = 1 - P\left[\underline{x} \leq 18 \mid p\right].$$

eze waarden en hun som $\pi(p)$ zijn weergegeven in figuur 8.3. In
iguur 8.4 vindt men geschetst hoe het linkeronderscheidingsvermogen π_ℓ
teeds beter wordt naarmate de steekproefomvang toeneemt. Dat de limiet
an $\pi_\ell(p)$ gelijk is aan 1 resp. 0 voor $p < \frac{1}{4}$ resp. $p > \frac{1}{4}$ wordt hier
iet bewezen. Voor π_r geldt uiteraard een analoog resultaat en ook de
euze $p_0 = \frac{1}{4}$ doet niet terzake.

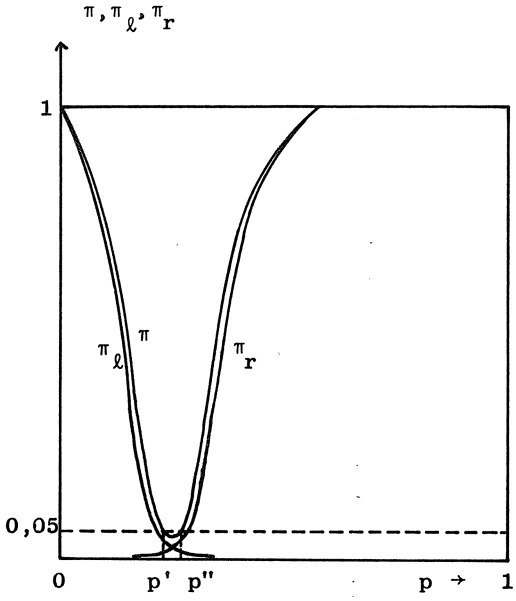


fig. 8.3

Onderscheidingsvermogen van de binomiale toets voor $p = \frac{1}{4}$ bij n=50, $\alpha = 0,05$, $t_\ell = 6$ en $t_r = 19$.

dat $\pi_\ell(p)$ voor $p < \frac{1}{4}$ snel afneemt. Voor $p = 0,3$ is $\pi_\ell(0,3) = 0,002$ de kans om de falikant verkeerde conclusie $p < \frac{1}{4}$ te trekken; $\pi_r(0,3) = 0,141$ is daarentegen de

Wij merken nog op, dat de hier beschreven toets ook beschouwd kan worden als een toets voor $p = 0,24$ met $\alpha_0 = \pi(0,24) = 0,047$, of als een toets voor $p' \leq p \leq p''$ met $\alpha = 0,05$ waarbij p' en p" de p-coördinaten zijn van de snijpunten van de kromme $\pi(p)$ met de horizontale lijn $\pi = 0,0$ Deze fijne onderscheidingen worden gewoonlijk niet gemaakt.

Omdat $\pi_\ell(p)$ de kans voorstelt dat wij tot $p < \frac{1}{4}$ besluiten, terwijl p de ware waarde is, is het verheugend

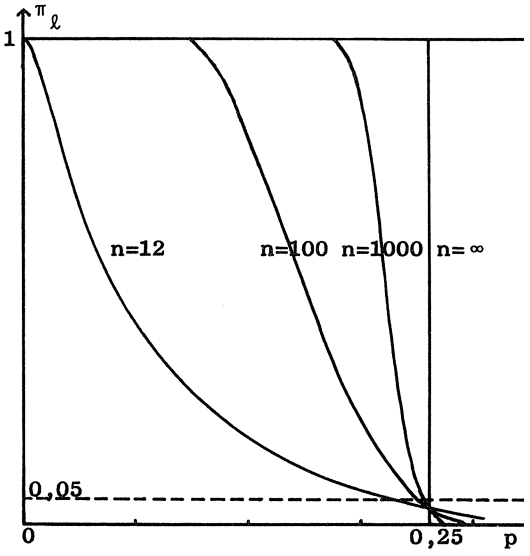


fig. 8.4

Onderscheidingsvermogen bij linkseenzijdige toetsing van $p \geq \frac{1}{4}$, voor $\alpha = 0,05$ en diverse waarden van n.

ans dat wij de (juiste) conclusie $p > \frac{1}{4}$ zullen trekken. Dus is $\pi(0,3)$
e som van de kansen op een falikant onjuiste en op een juiste conclu-
ie. Het is daarom vaak beter π_r en π_ℓ afzonderlijk te beschouwen.

. Binomiale toets met normale benadering.

Aan het slot van deel 2 hebben wij al gezien, dat de binomiale
erdeling voor grote n benaderd kan worden door de normale. Deze
igenschap zullen wij nu gebruiken om voor de binomiale toets van
$= p_0$ bij benadering kritieke zones en overschrijdingskansen te
epalen. Zo is bijv.

).1) $$\alpha_\ell = P\left[\underline{x} \leq x_\ell \mid p_0\right] \approx P\left[\underline{u} \leq \frac{x_\ell + \frac{1}{2} - np_0}{\sqrt{np_0q_0}}\right] ;$$

renals in deel 2 op blz. 115 hebben we bij de vervanging van de
iscreet verdeelde $\underline{x}$ door de continu verdeelde $\underline{u}$ de continuïteits-
rrectie toegepast. Is nu u_β het (in tabel B van deel 2 te vinden)
tal waarvoor geldt

).2) $$P\left[\underline{u} \geq u_\beta\right] = \beta ,$$

n moet bij benadering voor de kritieke waarden x_ℓ en x_r gelden

).3) $$-u_{\alpha_\ell} \approx \frac{x_\ell - np_0 + \frac{1}{2}}{\sqrt{np_0q_0}} \quad ; \quad \text{dus} \quad x_\ell \approx np_0 - \frac{1}{2} - u_{\alpha_\ell}\sqrt{np_0q_0} ,$$

sp.

).4) $$u_{\alpha_r} \approx \frac{x_r - np_0 - \frac{1}{2}}{\sqrt{np_0q_0}} \quad ; \quad \text{dus} \quad x_r \approx np_0 + \frac{1}{2} + u_{\alpha_r}\sqrt{np_0q_0} .$$

t geeft ons de kritieke zones voor links- en rechtseenzijdige toet-
ng met onbetrouwbaarheid ongeveer α_ℓ resp. α_r. Voor een symmetrische

tweezijdige toets met onbetrouwbaarheid ongeveer α, vervangt men α_ℓ en α_r door $\frac{1}{2}\alpha$ in (9.3) en (9.4). Om aan de goede kant van de onbetrouwbaarheidsdrempel te blijven, rondt men de hier verkregen x_ℓ naar beneden en x_r naar boven af.

Volkomen analoog vindt men voor de overschrijdingskansen

$$(9.5) \qquad k_\ell(x) = P\left[\underline{x} \leq x \mid p_0\right] \approx P\left[\underline{u} \leq \frac{x - np_0 + \frac{1}{2}}{\sqrt{np_0q_0}}\right] ;$$

$$(9.6) \qquad k_r(x) = P\left[\underline{x} \geq x \mid p_0\right] \approx P\left[\underline{u} \geq \frac{x - np_0 - \frac{1}{2}}{\sqrt{np_0q_0}}\right] .$$

In voorbeeld 8.1 ($n = 50$, $p_0 = \frac{1}{4}$) vindt men uit (9.6) bijv. $k_r(18) \approx P\left[\underline{u} \geq 1{,}63\right] = 0{,}0516$, terwijl volgens tabel 8.1 de exacte waarde 0,0551 is.

Voorbeeld 9.1

Bij een cursus statistiek wenste men ter demonstratie enige steekproefexperimenten uit te voeren. Daartoe werd een vaas gevuld met 1201 erwten, waarvan er 94 rood geverfd waren. Men nam nu steekproeven van 25 stuks door de vaas in de hand te nemen, goed te schudden en vervolgens met een plankje, voorzien van 25 uithollingen ter grootte van één erwt, 25 erwten uit de vaas te scheppen. Op deze wijze werden 100 steekproeven van 25 verkregen. In tabel 9.1 zijn de resultaten weergegeven.

Tabel 9.1

Aantallen rode erwten in 100 steekproeven van 25

aantal rode erwten	frequentie	aantal rode erwten	frequentie
0	2	4	18
1	21	5	3
2	28	6	4
3	24	≥ 7	0

In deze tabel staat in de eerste kolom het aantal rode erwten en
ı de tweede kolom het aantal steekproeven waarin dit aantal rode
ɾwten gevonden is. In totaal werden dus 2.0 + 21.1 + 28.2 + 24.3 +
18.4 + 3.5 + 4.6 = 260 rode erwten gevonden op de 2500 erwten; d.w.z.
ɜn fractie van 0,104. Aangezien in de vaas slechts een fractie
$\frac{94}{201}$ = 0,0783 rode erwten voorkomt, lijkt dit aantal rode erwten vrij
ɔog. Dit steekproef-experiment doet derhalve het vermoeden rijzen dat
ɜ trekkingswijze niet aselect is. Om dit na te gaan mogen we niet een
ɔets toepassen op de gegevens van tabel 9.1, aangezien ons vermoeden
ıist hierop gebaseerd is. Wij hebben daarom dezelfde persoon nogmaals
)0 steekproeven van 25 laten trekken op dezelfde wijze als tevoren.
ıder de 2500 zo verkregen (en weer teruggelegde) erwten bevonden zich
39 rode erwten. Formule (9.6) geeft nu

$$k_r(229) \approx P\left[\underline{u} \geq \frac{228{,}5 - 0{,}0783.2500}{\sqrt{2500.0{,}0783.0{,}9217}}\right] = P\left[\underline{u} \geq 2{,}44\right] = 0{,}0073 \ .$$

.j rechtseenzijdige toetsing met onbetrouwbaarheidsdrempel 0,01 zouden
ɜ de hypothese $p \leq 0{,}0783$ dus moeten verwerpen ten gunste van
$> 0{,}0783$. Dit wijst er dus op, dat de trekkingswijze niet aselect is,
.. dat om de een of andere reden systematisch teveel rode erwten
ɜtrokken worden. De volgende stap is nu het opsporen van de oorzaak
ɜn dit verschijnsel. Deze werd in dit geval inderdaad gevonden. De 94
ɔod geverfde erwten waren gemiddeld groter dan de andere erwten. Bij
ɜt schudden zakken de kleine erwten gemakkelijker naar de bodem dan
ɜ grote. Het gevolg hiervan is dat men, vooral als men de pot in de
ɜnd en dus onwillekeurig wat scheef houdt, systematisch teveel grote,
ı dus ook systematisch teveel rode erwten vindt.

10. Chikwadraat-toets voor aanpassing.

a) Bekende verdeling

Laat een verdelingsfunctie F(x) gegeven zijn (zowel de vorm als de parameters zijn dus bekend) en een steekproef van n onafhankelijke waarnemingen $x_1, x_2, \ldots, x_n$. Wij willen toetsen of deze waarnemingen uit de gegeven verdeling afkomstig kunnen zijn.

Hiertoe verdeelt men de verzameling van waarden die de stochastische grootheid $\underline{x}$ kan aannemen, in r klassen K_i (meestal intervallen). Laat f_i het aantal waarnemingen in K_i voorstellen en laat $p_i = P\left[\underline{x} \in K_i \mid F\right]$ zijn. Dan is $\sum_{i=1}^{r} f_i = n$ en $\sum_{i=1}^{r} p_i = 1$. Onder de nulhypothese dat de waarnemingen een kans p_i hadden om in de klasse K_i te vallen, zal het gevonden aantal waarnemingen f_i weinig afwijken van het verwachte aantal waarnemingen np_i. Men kan nu bewijzen dat de limiet voor $n \to \infty$ van de verdeling van

$$\underline{t} \stackrel{\text{def}}{=} \sum_{i=1}^{r} \frac{(\underline{f}_i - np_i)^2}{np_i} \tag{10.1}$$

onder H_0 de χ^2-verdeling met r-1 vrijheidsgraden is (vgl. deel 2,(7.22)) Voor eindige n wijkt de verdeling van $\underline{t}$ weinig van deze limietverdeling af, mits alle np_i niet te klein zijn. Als H_0 niet juist is, wijken één of meer f_i vermoedelijk drastisch van np_i af. Ongeacht het teken van $f_i - np_i$ zal de toetsingsgrootheid (10.1) dan groter zijn. Wij toetsen de nulhypothese dus rechtseenzijdig: als de waarde t van $\underline{t}$ zo groot is dat voor een grootheid $\underline{\chi}^2_{r-1}$ met de χ^2-verdeling met parameter r-1 geldt $P\left[\underline{\chi}^2_{r-1} \geq t\right] \leq \alpha$, zullen wij de nulhypothese verwerpen en daarmee ook verwerpen dat de waarnemingen afkomstig waren uit de gegeven verdeling.

De keuze van de klassen K_i is bij deze χ^2-toets willekeurig. Omdat de toets echter geen onderscheid maakt tussen de gegeven verdeling en elke andere verdeling die dezelfde kansen p_i op de klassen K_i geeft, wil men het aantal klassen zo groot mogelijk maken. Aan de andere kant heeft $\underline{t}$ onder de nulhypothese alleen asymptotisch de χ^2-verdeling en wijkt de verdeling voor eindige n hiervan sterker af naarmate de

'erwachte aantallen waarnemingen np_i kleiner worden. Het is in ieder ;eval wenselijk alle np_i groter dan 1 te maken; sommige leerboeken ':isen groter dan 5. Als ruwe richtlijn dient de volgende tabel van ıanbevolen waarden voor np_i bij verschillende waarden van n.

Tabel 10.1

n	aanbevolen np_i
100	5
200	10
400	20
1000	30

'oms maakt men np_i in de staarten van de verdeling wat kleiner dan in ıet midden.

In het bovenstaande is nergens vermeld of de verdelingsfunctie F liscreet of continu is; dit doet voor de toets niet terzake. Ook mogen le waarnemingen in gegroepeerde vorm gegeven zijn, dus bijv. n_j /aarnemingen in het interval I_j (j=1,2,...,J). In zo'n geval zal men 'oor de klassen K_i de intervallen I_j gebruiken en als dit tot te kleine ıp$_i$ zou leiden, er eventueel enige samenvoegen.

Voorbeeld 10.1. Eindcijfers.

In een biologische publikatie worden de gemeten diameters van ;eer kleine schelpdiertjes vermeld: de 874 gemeten waarden variëren van ',8 tot 12,0 micron. Het valt een kritische lezer op, dat nogal veel /aarnemingen een nul achter de komma vertonen. Hij vraagt zich af of le onderzoeker het aantal tienden van microns wel nauwkeurig kon ıflezen. Als het laatste cijfer ten dele gegist is was er misschien ;en onbewuste neiging om tot een geheel aantal microns af te ronden.

Het gebruik van toetsingstheorie om deze vraag te beantwoorden ıeidt tot een moeilijkheid. Hier zijn nl. eerst metingen gedaan, waarna)ns aan de meetresultaten een verschijnsel is opgevallen dat wij /ervolgens willen toetsen. Strikt genomen zouden wij dan een andere ;erie metingen van dezelfde bioloog voor de toets moeten gebruiken, ın analogie met hetgeen bij voorbeeld 9.1 is opgemerkt.

In dit geval was er geen tweede serie beschikbaar, zodat wij ons met de oorspronkelijke metingen moesten behelpen.

Met het oog op het grote verschil tussen beide uiterste waarnemingen mogen wij veronderstellen dat bij exacte meting elk eindcijfer even vaak zal voorkomen. Als nulhypothese kiezen wij daarom, dat het eindcijfer $\underline{x}$ elk der waarden $0,1,\ldots,9$ met kans $\frac{1}{10}$ aanneemt (discrete homogene verdeling, deel 2,(7.10)). In tabel 10.2 vindt men de waargenomen frequenties van de eindcijfers.

Tabel 10.2

Eindcijfer i	Frequentie f_i	f_i^2	f_i in procenten
0	176	30976	20,1
1	63	3969	7,2
2	109	11881	12,5
3	90	8100	10,3
4	84	7056	9,6
5	82	6724	9,4
6	73	5329	8,4
7	59	3481	6,7
8	89	7921	10,2
9	49	2401	5,6
Totaal	874	87838	100,0

Wij toetsen de nulhypothese met de χ^2-toets. De indeling in klassen is hier al gegeven; wij beschouwen de waarden $0,1,\ldots,9$ elk als een aparte klasse. Voor de bepaling van de toetsingsgrootheid bedenken wij dat $n = 874$ en $p_i = 0,1$ voor elke i. Dus geldt

$$t = \sum_{i=0}^{9} \frac{(f_i - 87,4)^2}{87,4} = \frac{\sum_{i=0}^{9} f_i^2}{87,4} - \frac{2.87,4 \sum_{i=0}^{9} f_i}{87,4} + 10.87,4 =$$

(10.2)

$$= \frac{1}{87,4} \sum_{i=0}^{9} f_i^2 - 874.$$

.j vinden dus $t = \frac{87838}{87,4} - 874 = 131,011$. Volgens een tabel van de
-verdeling met 9 vrijheidsgraden (er zijn immers 10 klassen) heeft
:ze waarde een rechteroverschrijdingskans van minder dan 0,0001. Wij
ınnen de nulhypothese dus met grote stelligheid verwerpen. Het is
.termate onwaarschijnlijk dat zo grote afwijkingen tussen f_i en np_i
)or toeval tot stand zijn gekomen. In de tabel hebben wij f_i ook nog
ı procenten uitgedrukt, waarbij men duidelijk ziet dat het eindcijfer
veel te vaak voorkomt en de eindcijfers 1 en 9 betrekkelijk weinig.

Verdeling met aangepaste parameters

De χ^2-toets voor het vergelijken van waarnemingen met een volledig
:kende verdeling wordt niet zo vaak gebruikt. Meestal wil men nl.
.derzoeken of de waarnemingen afkomstig kunnen zijn uit een verdeling
ın bekende vorm maar met onbekende parameters. Men wil bijv. weten
' de waarnemingen uit een normale verdeling komen en welke μ en σ die
:rdeling heeft doet niet terzake. Wij kunnen dan die parameters zo
ınstig mogelijk kiezen (d.w.z. schatten uit de steekproef) om een
:ede aanpassing tussen waarnemingen en normale verdeling te krijgen.

Laat $F(x \mid \theta_1,\theta_2,\ldots,\theta_s)$ een kansverdeling van bekende vorm zijn,
ıar met s onbekende onafhankelijke parameters; $\underline{x}_1,\underline{x}_2,\ldots,\underline{x}_n$ zijn
: waarnemingen. Men kiest weer een indeling in klassen K_i, waarbij
het aantal waarnemingen in K_i voorstelt. De kansen

.0.3) $$\pi_i \overset{\text{def}}{=} P\left[\underline{x} \in K_i \mid F(x \mid \theta_1,\theta_2,\ldots,\theta_s)\right]$$

.jn niet bekend, omdat wij de parameters niet kennen. Schatten wij
.ke θ_j door $\underline{t}_j = t_j(\underline{x}_1,\underline{x}_2,\ldots,\underline{x}_n)$, dan is

.0.4) $$\underline{p}_i \overset{\text{def}}{=} P\left[\underline{x} \in K_i \mid F(x \mid \underline{t}_1,\underline{t}_2,\ldots,\underline{t}_s)\right]$$

:n schatter voor π_i. Als nu de schatters $\underline{t}_j$ aan zekere eisen voldoen,
.n heeft

$$\underline{t} = \sum_{i=1}^{r} \frac{(\underline{f}_i - n\underline{p}_i)^2}{n\underline{p}_i}$$

onder de nulhypothese ongeveer een χ^2-verdeling met r-1-s vrijheidsgraden. Wij zullen dit niet bewijzen, maar we merken op dat het aantal vrijheidsgraden nu niet r-1 is omdat de p_i hier stochastisch zijn en uit dezelfde steekproef worden geschat die ons ook de f_i geeft. Men kiest de klassen weer zo dat het verwachte aantal waarnemingen per klasse niet te klein is (zie tabel 10.1). De eisen waaraan de schatters voor θ_j (j=1,2,...,s) moeten voldoen zijn te gecompliceerd om hier weer te geven; vaak geven dergelijke schatters aanleiding tot moeizaam rekenwerk. Men neemt daarom wel zijn toevlucht tot eenvoudiger schattingsmethoden (bijv. de in paragraaf 6 behandelde momentenmethode) en toetst dan toch met de χ^2_{r-1-s}-verdeling, hoewel het bovenstaande dan niet meer geheel juist is. Wanneer de zo gevonden overschrijdingskans duidelijk boven of duidelijk onder de onbetrouwbaarheidsdrempel ligt, kan men het resultaat van de toets toch wel met vertrouwen bezien.

Voorbeeld 10.2. Verkoop van flesjes was.

In een warenhuis wil men in verband met het bepalen van een optimale voorraadpolitiek nagaan of de frequentieverdeling van de verkopen per dag van flesjes vloerwas aansluit bij een Poisson-verdeling. Gedurende een reeks weken wordt daartoe iedere dag opgenomen hoeveel flesjes er verkocht worden. In de tweede kolom van tabel 10.3 vindt men het aantal dagen f_i waarop i flesjes verkocht zijn.

Tabel 10.3

Aanpassing van een Poissonverdeling

aantal flesjes i	frequentie f_i	$P[\underline{x}=i \mid \lambda=0,805]$	np_i	$\frac{(f_i-np_i)^2}{np_i}$
0	22	0,4471	18,33	0,735
1	10	0,3599	14,76	1,533
2	7	0,1449	5,94	0,189
3	1	0,0389		
4	0	0,0078		
5	0	0,0013	1,97	0,000
6	1	0,0002		
> 6	0	0,0000		
Totaal	41		41	2,457

De parameter λ van de Poissonverdeling is onbekend en wordt eschat met de momentenmethode. Omdat de verwachting van de Poisson-erdeling juist λ is (deel 2, tabel A), gebruiken wij als schatting oor λ het steekproefgemiddelde

10.5) $$\bar{x} = \frac{1}{41}(22.0 + 10.1 + 7.2 + 1.3 + 1.6) = 0{,}805 \ .$$

n een tabel van de Poissonverdeling zijn de individuele kansen voor = 0,805 opgezocht (derde kolom). Als klassen voor de χ^2-toets ebruiken wij de afzonderlijke waarden 0,1,2; om een redelijk grote np_i e krijgen voegen wij de waarden $\geq$ 3 samen tot één klasse.

Met 4 klassen en 1 aangepaste parameter vindt men een toetsings-rootheid t = 2,46. Onder de nulhypothese heeft t een χ^2-verdeling met -1-s = 4-1-1 = 2 vrijheidsgraden. De overschrijdingskans van 2,46 edraagt dan 0,29. Wij zien dat er geen reden is de hypothese van een oissonverdeling te verwerpen; de aansluiting is heel behoorlijk.

euze tussen verschillende verdelingen

De χ^2-toets wordt onderscheidender naarmate men van grotere teekproeven uitgaat. Dit betekent dat ook kleine afwijkingen van H_0 leine overschrijdingskansen geven, indien n maar groot genoeg is. et is daarom moeilijk een onbetrouwbaarheidsdrempel α vast te leggen; eze grens zou afhankelijk van n gekozen moeten worden. Van deze oeilijkheid heeft men minder last indien men een keuze moet doen ussen verschillende mogelijke verdelingen. De χ^2-toets kan dan als riterium gebruikt worden door de toets op ieder van de verdelingen oe te passen en die verdeling te kiezen, die de grootste overschrij-ingskans geeft. Men heeft dan nl. die verdeling uitgezocht waarbij en, althans op grond van de χ^2-toets, het minst geneigd is de ypothese, dat de waarnemingen $x_1,\ldots,x_n$ uit deze verdeling afkomstig ijn, te verwerpen.

11. Verband tussen toets en betrouwbaarheidsinterval.

Wij beschouwen een verdelingsfunctie $F(x \mid \theta)$ met één parameter θ; laat $\underline{t}$ een schatter voor θ zijn. Gemakshalve wordt ondersteld dat $\underline{t}$ en θ alle reële waarden kunnen aannemen. Als de verdeling van $\underline{t}$ voor $\theta = \theta_0$ bekend is, kunnen wij met $\underline{t}$ als toetsingsgrootheid een symmetrische tweezijdige toets voor H_0: $\theta = \theta_0$ construeren met onbetrouwbaarheid α. Als $Z_\ell = (-\infty, t_\ell]$ en $Z_r = [t_r, \infty)$ de kritieke zones van deze toets zijn, hangen de getalwaarden van t_ℓ en t_r uiteraard af van de getoetste waarde θ_0: wij schrijven $t_\ell(\theta_0)$ en $t_r(\theta_0)$. Als wij zo'n toets opstellen voor iedere waarde θ_0, kunnen wij de kritieke waarden verbinden tot twee curven $t = t_\ell(\theta_0)$ en $t = t_r(\theta_0)$. Deze curven zullen ongeveer verlopen zoals in figuur 11.1 is aangegeven. Het punt met gelijke coördinaten $t = \theta$ ligt steeds in het interval waarbinnen H_0 niet verworpen wordt.

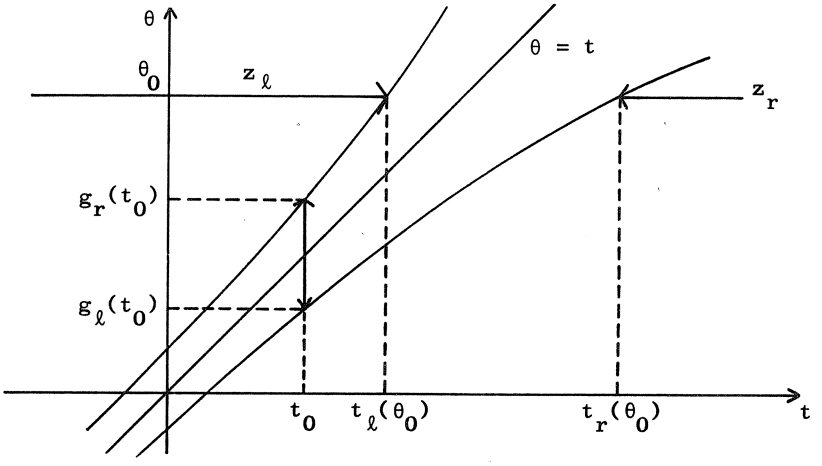


fig. 11.1

Toets en betrouwbaarheidsinterval

Wij kunnen ook een vaste waarde t_0 beschouwen en het interval $(g_\ell(t_0), g_r(t_0))$ definiëren als de verzameling van die waarden θ_0 van θ waarvoor H_0: $\theta = \theta_0$ niet verworpen wordt als de toetsingsgrootheid de waarde t_0 heeft. Het is duidelijk dat dit interval begrensd wordt

or dezelfde twee curven; d.w.z. de relaties

1.1) $$t_\ell(\theta) < t < t_r(\theta) \quad \text{en} \quad g_\ell(t) < \theta < g_r(t)$$

jn logisch equivalent. Wij beweren nu dat $(g_\ell(\underline{t}), g_r(\underline{t}))$ een trouwbaarheidsinterval voor θ is met onbetrouwbaarheid α (zie finitie 3.1). Immers wegens de genoemde logische equivalentie is

1.2) $$P\left[g_\ell(\underline{t}) < \theta < g_r(\underline{t})\right] = P\left[t_\ell(\theta) < \underline{t} < t_r(\theta)\right] =$$
$$= 1 - P\left[\underline{t} \in Z_\ell\right] - P\left[\underline{t} \in Z_r\right] = 1 - \alpha \ .$$

Is omgekeerd een willekeurig betrouwbaarheidsinterval $(a(\underline{t}), b(\underline{t}))$ geven, dan kunnen wij een toets voor de hypothese H_0: $\theta = \theta_0$ stellen. Wij verwerpen H_0 nl. als de toetsingsgrootheid een waarde t eft waarbij θ_0 niet in het interval $(a(t), b(t))$ ligt.

Men kan gemakkelijk nagaan dat er een soortgelijk verband bestaat ssen een rechtseenzijdige toets en een betrouwbaarheidsondergrens tussen een linkseenzijdige toets en een betrouwbaarheidsbovengrens. die gevallen ontbreekt in figuur 11.1 de curve linksboven resp. de rve rechtsonder.

De onderstelling dat de toetsingsgrootheid $\underline{t}$ een schatter voor de rameter θ is, werd alleen voor de eenvoud van formulering gemaakt kan gemist worden. Het gebied van θ_0-waarden waarbij $\theta = \theta_0$ niet rworpen wordt als de toetsingsgrootheid de waarde t heeft, bepaalt eeds een betrouwbaarheidsinterval voor θ, en omgekeerd.

. Verdelingsvrije toetsen.

Bij de schattings- en toetsingstheorie hebben wij steeds verondereld, dat de verdelingsfunctie van de waarnemingen op één of meer rameters na bepaald was. Wanneer een experiment een aantal waarnengen oplevert, is aan deze eis natuurlijk niet voldaan. Wel kan men

met grafisch onderzoek (paragraaf 1) of een χ^2-toets (paragraaf 10) nagaan of de waarnemingen bijv. passen bij een normale verdeling, maar de conclusie is dan toch hoogstens dat een normale verdeling niet onaanvaardbaar is; normaliteit bewijzen is niet mogelijk. Bovendien kan het gebeuren dat geen enkele bekende verdeling redelijk bij de waarnemingen past, of dat men geen tijd heeft allerlei verdelingen te proberen.

Voor dit soort omstandigheden bestaan er toetsen die geen veronderstellingen maken over de verdelingsfunctie van de waarnemingen. Wij zullen enige voorbeelden van zulke verdelingsvrije toetsen noemen, zonder in details te treden. In het algemeen kan men zeggen, dat een verdelingsvrije toets voor een bepaalde hypothese een wat slechter onderscheidingsvermogen zal hebben dan een toets voor dezelfde hypothes onder veronderstelling van bijv. normaliteit. Als men minder onderstellingen maakt kan men minder conclusies trekken. Aan de andere kant verdient een verdelingsvrije toets de voorkeur wanneer een onderstelling, zoals bijv. die van normaliteit, in feite ongefundeerd is.

Tekentoets

Wanneer bekend is dat de waarnemingen afkomstig zijn uit een symmetrische verdeling met verwachting μ, kan men H_0: $\mu = \mu_0$ toetsen door gebruik te maken van het feit, dat voor symmetrische verdelingen geldt

$$P\left[\underline{x} < \mu_0 \mid H_0\right] = P\left[\underline{x} > \mu_0 \mid H_0\right] . \tag{12.1}$$

Is de verdeling continu, dan zijn beide leden gelijk aan $\frac{1}{2}$ omdat $P\left[\underline{x} = \mu_0 \mid H_0\right] = 0$ is. De relatie (12.1) geldt echter ongeacht continuïteit ook onder de extra voorwaarde $\underline{x} \neq \mu_0$. Dus als we eventuele waarnemingen $x = \mu_0$ weglaten, geldt voor de overige waarnemingen dat zij onder H_0 met kans $\frac{1}{2}$ kleiner dan μ_0 zijn.

Wij kunnen dus voor

$$p \stackrel{\text{def}}{=} P\left[\underline{x} < \mu_0 \mid \underline{x} \neq \mu_0\right] \tag{12.2}$$

e nulhypothese formuleren als H_0^*: $p = \frac{1}{2}$, en de binomiale toets ebruiken (paragraaf 8 en 9). Omdat $\mu > \mu_0$ correspondeert met $p < \frac{1}{2}$ en $< \mu_0$ met $p > \frac{1}{2}$, weten wij bij verwerpen van H_0^* ook welke alternaieve hypothese van H_0 wij moeten aanvaarden.

De binomiale toets voor $p = \frac{1}{2}$ is eveneens bruikbaar om bij paren aarnemingen $(\underline{x}_i, \underline{y}_i)$ na te gaan of er een systematisch verschil in rootte bestaat tussen $\underline{x}_i$ en de bijbehorende $\underline{y}_i$. Onder de nulhypothese at $\underline{x}_i > \underline{y}_i$ even vaak voorkomt als $\underline{x}_i < \underline{y}_i$ geldt ook hier $p = \frac{1}{2}$, als

12.3) $$p \overset{\text{def}}{=} P\left[\underline{x}_i > \underline{y}_i \mid \underline{x}_i \neq \underline{y}_i\right] .$$

ij moeten dus de paren waarnemingen met $x_i = y_i$ schrappen en bij de verige paren nagaan of $x_i > y_i$ is (succes) dan wel $x_i < y_i$ (mislukking). et is bij deze toets niet nodig dat alle $\underline{x}_i$ dezelfde kansverdeling ebben en evenmin de $\underline{y}_i$.

In beide gevallen spreekt men van de tekentoets, omdat alleen het eken van $\underline{x}_i - \mu_0$ resp. $\underline{x}_i - \underline{y}_i$ gebruikt wordt bij het bepalen van de oetsingsgrootheid. Of een dergelijk verschil groot of klein is heeft us op het resultaat van de toets geen invloed.

oorbeeld 12.1. Lengten van echtparen.

Een onderzoeker heeft de lengten gemeten van de mannen en vrouwen an 21 aselect gekozen echtparen, met de bedoeling om na te gaan of ehuwde mannen in het algemeen groter zijn dan hun vrouwen. De (ficieve) resultaten van dit (fictieve) onderzoek zijn samengevat in abel 12.1, waarvan de laatste kolom later gebruikt zal worden.

Wij willen $p = \frac{1}{2}$ tweezijdig toetsen met $\alpha = 0{,}05$. Het 18^{de} chtpaar wordt geschrapt omdat $x_i = y_i$ is. Er blijven 20 echtparen ver. Uit een tabel van de binomiale verdeling voor $n = 20$ en $p = \frac{1}{2}$ inden wij de kritieke waarden 5 en 15. Wij hebben 14 echtparen met $_i > y_i$, zodat de nulhypothese niet wordt verworpen. Volgens de inomiale tabel is de overschrijdingskans van 14 (vgl. paragraaf 8)

(12.4) $k(14) = 2.\min\{k_\ell(14), k_r(14)\} = 2.\min\{0,979; 0,058\} = 0,116.$

Tabel 12.1

Lengten van echtgenoten (in meters)

Nummer van het echtpaar	$x_i \equiv$ lengte van de man	$y_i \equiv$ lengte van de vrouw	teken van $x_i - y_i$	T_i (zie later)
1	1,59	1,63	-	15
2	1,62	1,66	-	12
3	1,63	1,53	+	18
4	1,65	1,61	+	15
5	1,66	1,56	+	16
6	1,67	1,72	-	7
7	1,69	1,65	+	11
8	1,70	1,74	-	6
9	1,71	1,64	+	10
10	1,72	1,62	+	10
11	1,73	1,68	+	8
12	1,74	1,59	+	9
13	1,75	1,70	+	7
14	1,76	1,71	+	6
15	1,77	1,67	+	6
16	1,79	1,77	+	3
17	1,81	1,75	+	4
18	1,82	1,82	=	2
19	1,83	1,84	-	1
20	1,85	1,76	+	1
21	1,88	1,89	-	0
				167

De hypothese dat er geen systematisch verschil bestaat mag dus niet verworpen worden. Wij hebben tweezijdig getoetst, omdat $p < \frac{1}{2}$ en $p > \frac{1}{2}$ beide denkbaar zijn. Men mag nimmer op grond van de waarnemingen concluderen dat een eventuele afwijking van $p = \frac{1}{2}$ wel $p > \frac{1}{2}$ zal betekenen, en om die reden rechtseenzijdig toetsen. Slechts als $p < \frac{1}{2}$ tevoren uitgesloten wordt geacht of oninteressant, kan men besluiten een eenzijdige toets te gebruiken.

angcorrelatietoetsen

De onderzoeker uit het laatste voorbeeld vraagt zich af of het aar is, dat lange mannen vaak met lange vrouwen trouwen. Er zijn ook ensen die beweren dat kleine mensen juist een lange huwelijkspartner oeken. Als de lengten van de man en de vrouw van het i^{de} aselect ekozen echtpaar weer door $\underline{x}_i$ resp. $\underline{y}_i$ worden voorgesteld, kan men de ulhypothese toetsen dat $\underline{x}$ en $\underline{y}$ ongecorreleerd zijn.

Wanneer men weet dat de grootheden $\underline{x}$ en $\underline{y}$ een twee-dimensionale ormale verdeling hebben, kan men uit de steekproef-correlatiecoëffiiënt (vgl. deel 2,(10.41))

12.5)
$$r = \frac{\sum_{i=1}^{n} (x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i-\bar{x})^2 \sum_{i=1}^{n} (y_i-\bar{y})^2}}$$

en betrouwbaarheidsinterval voor de correlatiecoëfficiënt ρ van de opulatie bepalen. Ligt de waarde $\rho = 0$ in dit interval, dan zullen wij) (ongecorreleerdheid) niet verwerpen; bevat het interval uitsluitend ositieve waarden, dan verwerpen wij H_0 ten gunste van de hypothese an positieve correlatie en analoog als het interval alleen negatieve aarden bevat. Wij gaan hier niet in op de constructie van zo'n etrouwbaarheidsinterval; er bestaan grafieken waaruit men bij gegeven teekproefomvang en gegeven r het interval direkt kan aflezen (o.a. in EARSON & HARTLEY, Biometrika Tables, tabel 15, voor $\alpha = 0,05$ en = 0,01).

Ook voor deze nulhypothese bestaat er een verdelingsvrije toets, aarbij men geen normale verdelingen voor $\underline{x}$ en $\underline{y}$ behoeft aan te nemen. Lj merken op, dat de x_i in tabel 12.1 naar opklimmende grootte zijn erangschikt. Als het nu waar is dat lange mannen met lange vrouwen rouwen (dat $\underline{x}$ en $\underline{y}$ positief gecorreleerd zijn), dan zullen er aan het gin van de tabel veel kleine waarden voor y_i staan en aan het eind ornamelijk grote. Hoe sterker de positieve correlatie, hoe minder

zullen de getallen in de y_i-kolom afwijkingen van de natuurlijke volgorde vertonen. Is de correlatie-coëfficiënt bijna -1, dan zullen de y_i ongeveer in dalende volgorde staan als de x_i opklimmen.

M.G. KENDALL heeft nu voorgesteld bij elke x_i het aantal T_i te bepalen van de y_j-waarden die onder y_i staan en groter zijn dan y_i, en vervolgens te bepalen

$$\tau = \frac{4 \sum_{i=1}^{n} T_i}{n(n-1)} - 1 \; . \tag{12.6}$$

Als de y_i in strikt dalende volgorde staan is $T_i = 0$ voor elke i, dus $\tau = -1$. Vertonen de y_i een strikt stijgende volgorde, dan is $T_i = n-i$, dus $\sum_{i=1}^{n} T_i = n-1+n-2+\ldots+1+0 = \frac{1}{2}n(n-1)$ (som van een rekenkundige reeks, deel 1,(7.12)). In dat geval is $\tau = +1$. Men kan nagaan dat τ in het algemeen kleiner is naarmate de y_i meer van de natuurlijke (stijgende) volgorde afwijken.

Onder de nulhypothese van ongecorreleerdheid is elke volgorde der y_i even waarschijnlijk; men kan de verdeling van $\underline{\tau}$ onder H_0 beschouwen en laten zien dat de waarden dicht bij $\tau = 0$ de grootste kansen hebben. Omdat elke $\underline{T}_i$ alleen gehele waarden aanneemt heeft $\underline{\tau}$ een discrete verdeling op de getallen $\frac{4j}{n(n-1)} - 1$ $(j=0,1,\ldots,\frac{1}{2}n(n-1))$.

Voorbeeld 12.1 (vervolg)

In de laatste kolom van tabel 12.1 staan de grootheden T_i (aantal $y_j > y_i$ waarvoor $j > i$). Wij vinden uit (12.6) met $n = 21$

$$\tau = \frac{4.167}{21.20} - 1 = 0{,}59 \; . \tag{12.7}$$

Om na te gaan of H_0 verworpen moet worden, d.w.z. of deze waarde significant van 0 afwijkt, werkt men meestal met

$$\underline{S} \stackrel{\text{def}}{=} 2 \sum_{i=1}^{n} \underline{T}_i - \tfrac{1}{2}n(n-1) \; ; \tag{12.8}$$

eze grootheid is voor grote n ongeveer normaal verdeeld met verwach-
ing 0 en variantie $\frac{n(n-1)(2n+5)}{18}$. Omdat $\underline{S}$ voor even waarden van
(n-1) uitsluitend even waarden aanneemt, passen wij nog een conti-
uïteitscorrectie toe (vgl. deel 2,blz. 115)

12.9) $$k_r(S) = P\left[\underline{S} \geq S\right] \approx P\left[\underline{u} \geq \frac{S-1}{\sigma(\underline{S})}\right] ;$$

12.10) $$k_\ell(S) = P\left[\underline{S} \leq S\right] \approx P\left[\underline{u} \leq \frac{S+1}{\sigma(\underline{S})}\right] .$$

ij onze gegevens was $\sum T_i = 167$ en n = 21, dus S = 334 - 210 = 124,
n $\sigma(\underline{S}) = \sqrt{21.20.\frac{47}{18}}$. Substitutie geeft $k_r(124) \approx P\left[\underline{u} \geq 3{,}71\right] = 0{,}0001$.
et is direkt duidelijk dat $k_\ell(124)$ veel groter zal zijn. De tweezijdige
verschrijdingskans is dus 0,02 procent, zodat men H_0 met grote
telligheid kan verwerpen. Omdat de uitslag van de toets zo duidelijk
s levert het geen bezwaren op dat de verdeling van $\underline{S}$ voor n = 20
og niet zo heel dicht bij de normale verdeling ligt. Uit de waarne-
ingen van tabel 12.1 (die overigens fictief waren!) zou dus volgen
at kleine mensen zich vaak een kleine huwelijkspartner zoeken. In
iguur 12.1 is dit verschijnsel duidelijk zichtbaar.

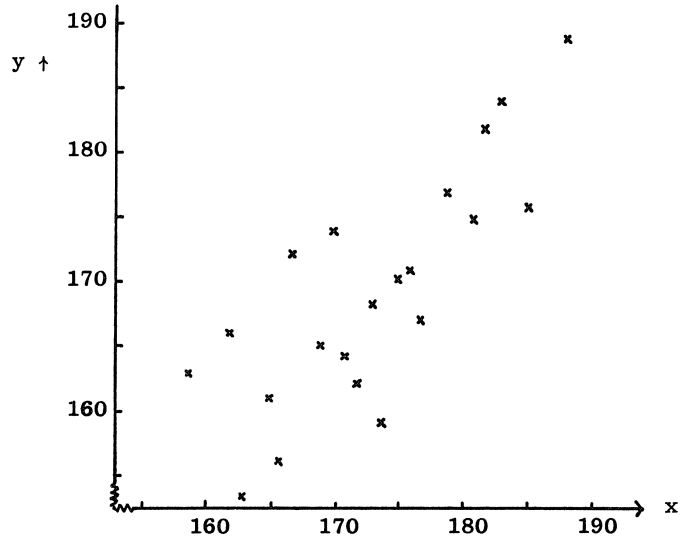


fig. 12.1

De lengten van mannen (x) en hun echtgenotes (y),
zoals vermeld in tabel 12.1

Wanneer onder de x_i of onder de y_i gelijke waarnemingen voorkomen, moet men enige modificaties aanbrengen. Wij verwijzen verder naar M.G. KENDALL, Rank correlation methods, Ch. Griffin & Co., London, (1955; 2nd edition).

De toets van WILCOXON voor twee steekproeven

De hypothese C genoemd in paragraaf 7 kan eveneens verdelingsvrij getoetst worden (de aldaar gemaakte veronderstelling van normaliteit kan dus gemist worden). Laat $x_1,x_2,\ldots,x_m$ resp. $y_1,y_2,\ldots,y_n$ een steekproef uit de eerste resp. de tweede populatie voorstellen. Wij rangschikken dan alle waarnemingen (x_i en y_j door elkaar) naar opklimmende grootte en tellen bij elke x_i het aantal y_j dat ervóór staat. Men kiest twee maal de som van deze aantallen als toetsingsgrootheid W. Onder de nulhypothese (gelijke verwachtingen) is de verdeling van W bekend voor kleine m en n, terwijl W voor grote steekproefomvangen vrijwel normaal verdeeld is. Ook hier geven gelijke waarnemingen aanleiding tot complicaties. Voor een uitvoerige behandeling van deze toets verwijzen wij naar de literatuur *); wij volstaan met een voorbeeld.

Voorbeeld 12.2

Een textielfabrikant wil door toepassing van een prepareermethode zijn garens sterker maken. Trekproeven, verricht op monsters onbehandelde en behandelde garens, geven de volgende resultaten (uitgedrukt in één of andere sterkte-eenheid):

Behandelde garens (x)	Onbehandelde garens (y)
11,79	10,34
11,21	11,40
13,20	10,19
12,66	12,10
13,37	11,46

*) Bijv. Ir. D. WABEKE en Dr. C. VAN EEDEN: Handleiding voor de toets van Wilcoxon, rapporten S 176 (M 65 en M 65A) van de Statistische Afdeling van het Mathematisch Centrum, 1e druk 1955; 2e druk 1970; en verder: A. BENARD en Dr. C. VAN EEDEN, Handleiding voor de symmetrietoets van Wilcoxon, rapport S 208 (M 76), idem, 1956.

ij toetsen de nulhypothese dat behandelde en onbehandelde garens
ven sterk zijn. Aangezien de fabrikant de prepareermethode echter
lleen wil invoeren indien deze methode de garens sterker maakt, is
ij er alleen in geïnteresseerd H_0 te verwerpen ten gunste van de
lternatieve hypothèse dat de behandelde garens systematisch sterker
ijn dan de niet behandelde. De twee mogelijkheden: 1) de prepareer-
ethode helpt niet, en 2) de prepareermethode schaadt, leiden beide tot
ezelfde beslissing nl. de methode niet in te voeren, zodat zij bij
it probleem, praktisch gezien, equivalent zijn. Wij behoeven deze
wee mogelijkheden dus niet van elkaar te onderscheiden. Het linkerdeel
an de kritieke zone van een tweezijdige toets zou dus doelloos zijn.
ij toetsen daarom rechtseenzijdig, en wel met $\alpha = 0{,}05$.

Rangschikking van de waarnemingen naar opklimmende grootte geeft
e volgende rij, waarin wij de y_i een halve regel lager hebben gezet
an de x_i:

		11,21			11,79		12,66	13,20	13,37
10,19	10,34		11,40	11,46		12,10			
jdrage:		2			4		5	5	5

der elke x_i staat het aantal voorafgaande y_j. Wij vinden
= 2(2+4+5+5+5) = 42. Volgens een tabel van de verdeling van $\underline{W}$ voor
= n = 5 is de rechter kritieke waarde voor $\alpha = 0{,}05$ eenzijdig juist
lijk aan 42, zodat we H_0 nog net verwerpen en de prepareermethode
s een verbetering beschouwen. De lezer had misschien een kleinere
erschrijdingskans verwacht, maar bij zulke kleine aantallen waarne-
ngen hebben bijna alle toetsen een gering onderscheidingsvermogen.

. Hoogwaterstanden, exponentiële verdeling.

Tot besluit behandelen wij een voorbeeld waarbij grafische
npassing, schatting en toetsing gebruikt zullen worden.

In opdracht van de na de overstroming van 1953 ingestelde Delta-
mmissie heeft het Mathematisch Centrum de kansverdeling van de

waargenomen hoogwaterstanden te Hoek van Holland onderzocht. Deze blijkt vrij goed aan te sluiten bij een exponentiële verdeling, wanneer men de lage peilen buiten beschouwing laat *). Dit betekent dat voor de hoogwaterstand $\underline{x}$ bij een geschikte niet te lage drempelwaarde x_0 geldt

$$(13.1) \qquad P\left[\underline{x} > x \mid \underline{x} \geq x_0\right] = \frac{P\left[\underline{x} > x \wedge \underline{x} \geq x_0\right]}{P\left[\underline{x} \geq x_0\right]} = \frac{P\left[\underline{x} > x\right]}{P\left[\underline{x} \geq x_0\right]} = e^{-\lambda(x-x_0)}, \quad (x > x_0) .$$

Om te laten zien dat de keuze van x_0 hier niet van belang is, hebben wij de volgende merkwaardige stelling nodig.

Stelling 13.1

Als $\underline{y}$ exponentieel verdeeld is, dan is voor elke $a > 0$ de voorwaardelijke verdeling van $\underline{y} - a$, gegeven $\underline{y} \geq a$, dezelfde als die van $\underline{y}$; d.w.z. als

$$(13.2) \qquad P\left[\underline{y} \leq y\right] = 1 - e^{-\lambda y} \qquad (y \geq 0) ,$$

dan is voor elke $a > 0$ en $y > 0$

$$(13.3) \qquad P\left[\underline{y} > y+a \mid \underline{y} \geq a\right] = P\left[\underline{y} > y\right] .$$

Bewijs:

$$(13.4) \qquad P\left[\underline{y} > y+a \mid \underline{y} \geq a\right] = \frac{P\left[\underline{y} > y+a \wedge \underline{y} \geq a\right]}{P\left[\underline{y} \geq a\right]} = \frac{P\left[\underline{y} > y+a\right]}{P\left[\underline{y} > a\right]} = \frac{e^{-\lambda(y+a)}}{e^{-\lambda a}} = e^{-\lambda y} = P\left[\underline{y} > y\right] ,$$

wegens de definitie van voorwaardelijke kans en de eigenschap $P\left[\underline{y} = a\right] = 0$.

---

*) Vergelijk: P.J. WEMELSFELDER, Wetmatigheden in het optreden van stormvloeden, "De Ingenieur", 9 (1939), Bouw- en Waterkunde 3.

Met deze stelling tonen wij aan dat de juistheid van (13.1) mpliceert dat deze relatie (13.1) correct blijft als men x_0 vervangt oor een willekeurige $x_1 > x_0$. Immers volgens (13.1) is $\underline{y} = \underline{x} - x_0$ xponentieel verdeeld. Men past nu stelling 13.1 toe op $\underline{y}$ en $a = x_1 - x_0$: e verdeling van $\underline{y} - a$ (d.w.z. $\underline{x} - x_1$) gegeven $\underline{y} \geq a$ (d.i. $\underline{x} \geq x_1$) s weer exponentieel. Maar dit betekent dat (13.1) blijft gelden bij ervanging van x_0 door x_1. Door de keuze van het beginpunt x_0 wordt e vorm van de verdeling en de waarde van λ dus niet beïnvloed. Dit aakt het mogelijk de lage waarden van x, waar de verdeling niet ekend is, uit te schakelen zonder dat de betrekkelijk willekeurige euze van het beginpunt veel invloed op de uitkomst uitoefent.

De genoemde eigenschap van de exponentiële verdeling is daarom oor het statistische onderzoek van de hoogwaterstanden aan de ederlandse kust - en daarmede voor de bepaling van dijkhoogten, die en voldoende bescherming tegen overstromingen bieden - van veel belang. en verklaring voor het experimenteel vastgestelde feit, dat de verdeing der hoogwaterstanden exponentieel is, is (nog?) niet bekend.

In verband met de onderzoekingen van de Deltacommissie was men oornamelijk geïnteresseerd in de hoogwaterstanden gedurende de maanden ovember, december en januari, die optraden tijdens depressies van een otentieel gevaarlijk karakter. In overleg met het K.N.M.I. werd aarom een groep waarnemingen geselecteerd die onder deze omstandigeden waren gedaan. In tabel 13.1 vindt men voor deze geselecteerde roep de frequenties waarmee peilen boven 1,69 meter bereikt werden, n het aantal overschrijdingen van elk peil. De standen zijn opgegeven ot in centimeters nauwkeurig. Daar gedurende de Tweede Wereldoorlog een weerkaarten getekend konden worden, was het onmogelijk de selectie oor de winters 1939/40 t/m 1944/45 toe te passen.

Tabel 13.1

Frequenties van de hoogwaterstanden boven 1,69 meter, opgetreden tijdens depressies met een potentieel gevaarlijk karakter, in de winters 1888/89 t/m 1939/40 en 1945/46 t/m 1956/57

Hoogwater-standen in meters	Frequentie	Aantal over-schrijdingen	Hoogwater-standen in meters	Frequentie	Aantal over-schrijdingen
3,85	1	1	2,07	2	61
3,28	1	2	2,05	1	62
3,00	2	4	2,04	2	64
2,96	2	6	2,02	2	66
2,74	1	7	2,01	4	70
2,68	2	9	2,00	1	71
2,66	1	10	1,99	2	73
2,63	1	11	1,98	1	74
2,62	2	13	1,97	2	76
2,54	1	14	1,96	2	78
2,53	1	15	1,95	7	85
2,52	1	16	1,94	1	86
2,50	1	17	1,93	2	88
2,44	1	18	1,92	1	89
2,39	1	19	1,91	2	91
2,38	1	20	1,90	3	94
2,37	1	21	1,89	2	96
2,36	1	22	1,88	4	100
2,33	2	24	1,87	1	101
2,26	1	25	1,86	3	104
2,25	2	27	1,85	4	108
2,24	1	28	1,84	3	111
2,22	3	31	1,83	2	113
2,21	1	32	1,82	4	117
2,20	1	33	1,81	4	121
2,18	3	36	1,80	8	129
2,16	2	38	1,79	1	130
2,15	2	40	1,78	6	136
2,14	1	41	1,77	1	137
2,13	1	42	1,76	7	144
2,12	8	50	1,75	1	145
2,11	1	51	1,74	3	148
2,10	2	53	1,73	3	151
2,09	3	56	1,72	5	156
2,08	3	59	1,71	1	157
			1,70	9	166

Wij zullen nu door grafisch onderzoek nagaan, of ook de geselec-
eerde hoogwaterstanden van tabel 13.1 bij een exponentiële verdeling
assen. Zoals in paragraaf 1 is opgemerkt moeten wij daartoe de
aarnemingen uitzetten op enkelvoudig logarithmisch papier.

Alle waargenomen punten worden in een rij geplaatst naar afdalende
ootte, zodanig dat een peil dat k maal is waargenomen ook k maal in
e rij voorkomt. Wij nummeren de zo verkregen rij en zetten bij elk
eil langs de as met de logarithmische schaalverdeling uit: $\frac{i-0,3}{166+0,4}$,
aarbij i het rangnummer van het betreffende peil is (d.w.z. het aantal
verschrijdingen). Bij sommige peilen worden dus verschillende punten
evonden. Deze procedure houdt verband met het feit dat de waarnemingen
iet exact zijn. Bij het peil 3,00 meter wordt dus tweemaal een punt
itgezet en bij het peil 2,12 meter zelfs 8 maal.

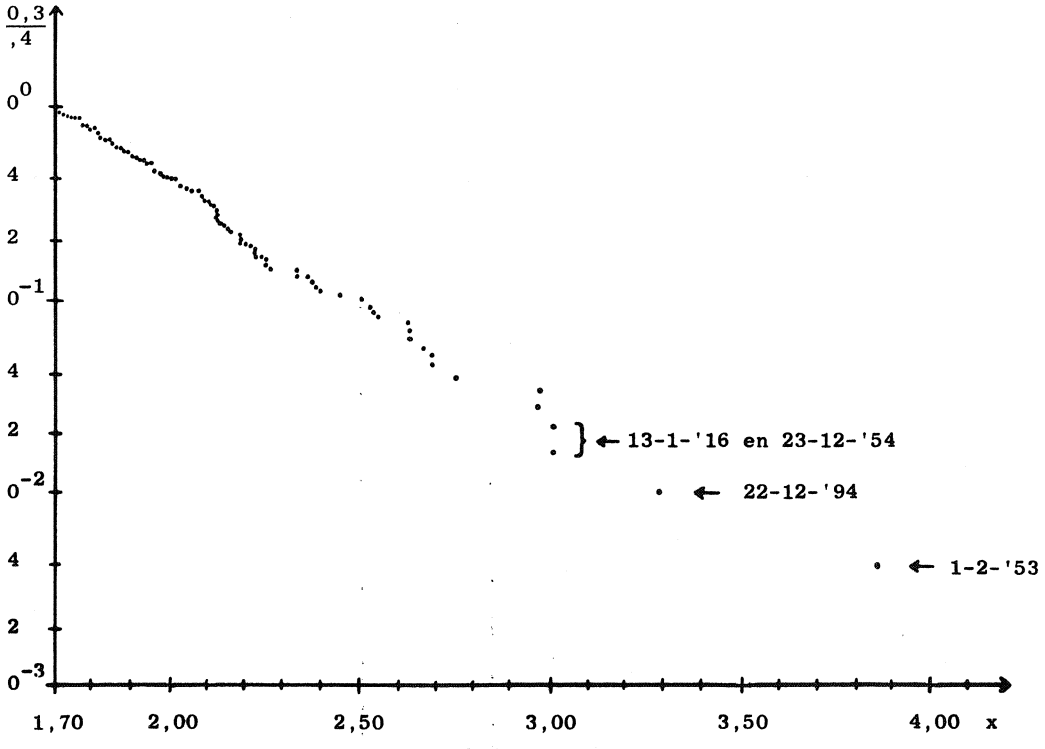


fig. 13.1

Verdeling der hoogwaterstanden boven 1,69 meter

De exponentiële aanpassing blijkt nu heel behoorlijk te zijn; immers de gevonden punten liggen bij benadering op een rechte lijn $y = ax+b$. Daaruit is de formule voor de verdelingsfunctie $F(x)$ van de hoogwaterstanden $\underline{x}$ af te leiden. Zij $n(x)$ het aantal overschrijdingen van het peil x; dan geldt volgens de gemaakte grafiek

$$\log \frac{n(x)-0,3}{166,4} \approx ax + b \quad , \tag{13.5}$$

of

$$1 - \frac{n(x)-0,3}{166,4} \approx 1 - e^{ax+b} \quad . \tag{13.6}$$

Voor $x = 1,70$ moet het rechterlid van deze vergelijking nul zijn en voor $x = +\infty$ gelijk aan één. Derhalve is $b = -1,70\,a$, terwijl a negatie is. We schrijven daarom $a = -\lambda$ met $\lambda > 0$.

De geselecteerde hoogwaterstanden boven 1,70 meter in de beschouwd periode passen dus bij een exponentiële verdelingsfunctie

$$F(x) = P\left[\underline{x} \leq x \mid \underline{x} \geq 1,70\right] = 1 - e^{-\lambda(x-1,70)} \quad . \tag{13.7}$$

Voor de onbekende parameter λ bepalen wij nu de <u>meest aannemelijke schatter</u> (paragraaf 4), waarbij wij de drempelwaarde 1,70 terwille van de algemeenheid weer als x_0 schrijven. De verdelingsdichtheid, steeds onder de voorwaarde $\underline{x} \geq x_0$, is

$$f(x) = F'(x) = \lambda e^{-\lambda(x-x_0)} \quad , \tag{13.8}$$

dus als $x_1, x_2, \ldots, x_n$ de waarnemingen zijn, is

$$L(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda(x_i-x_0)} = \lambda^n e^{-\lambda \sum_{i=1}^{n} (x_i-x_0)} \quad . \tag{13.9}$$

Dus

$$\log L = n \log \lambda - \lambda \sum_{i=1}^{n} (x_i-x_0) \tag{13.10}$$

en

$$\frac{d \log L}{d \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} (x_i-x_0) \quad . \tag{13.11}$$

erhalve is

13.12) $$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n}(x_i - x_0)} = \frac{1}{\bar{x} - x_0},$$

is $\frac{1}{\hat{\lambda}}$ is het gemiddelde der waarnemingen boven x_0 verminderd met x_0.
e waarde $\lambda = \hat{\lambda}$ geeft inderdaad een maximum voor log $L(\lambda)$.

Men kan laten zien dat voor de gevonden schatter geldt

13.13) $$\mathcal{E}\,\underline{\hat{\lambda}} = \mathcal{E}\left(\frac{1}{\underline{\bar{x}} - x_0} \mid \underline{x} \geq x_0\right) = \frac{n}{n-1}\lambda .$$

e meest aannemelijke schatter is dus niet zuiver, terwijl

13.14) $$\underline{\ell} = \frac{n-1}{n} \cdot \frac{1}{\underline{\bar{x}} - x_0} = \frac{n-1}{\sum_{i=1}^{n}(\underline{x}_i - x_0)}$$

el een zuivere schatter van λ is.

De schatters zijn beide asymptotisch raak. Uit deel 2 stelling 12.2 olgt namelijk, dat $\underline{\bar{x}} - x_0$ stochastisch convergeert naar $\mathcal{E}(\underline{x} - x_0 \mid \underline{x} \geq x_0) = \frac{1}{\lambda}$ en dus convergeert $\underline{\hat{\lambda}} = \frac{1}{\underline{\bar{x}} - x_0}$ stochastisch naar λ; enzo voor $\underline{\ell}$.

In tabel 13.1 is de gemiddelde hoogwaterstand $\bar{x} = 2{,}0318$ en de empelwaarde $x_0 = 1{,}70$, zodat de meest aannemelijke schatting is

13.15) $$\hat{\lambda} = \frac{1}{\bar{x} - x_0} = \frac{1}{2{,}0318 - 1{,}70} = 3{,}014 .$$

Ter bevestiging van het grafisch onderzoek zullen wij nu nog met en χ^2-toets nagaan of de waarnemingen uit tabel 13.1 passen bij een xponentiële verdeling met $\lambda = 3{,}014$. Het interval $x \geq 1{,}70$ verdelen ij hiertoe zó in klassen, dat het verwachte aantal hoogwaterstanden er klasse ongeveer gelijk aan 10 wordt (vgl. paragraaf 10). In tabel 3.2 worden deze klassen gegeven met de bijbehorende kansen en erwachte aantallen, terwijl in de kolommen 4 en 8 de waargenomen antallen hoogwaterstanden vermeld zijn.

Tabel 13.2

Aanpassing van een exponentiële verdeling

klasse	p_i	$166.p_i$	f_i	klasse	p_i	$166.p_i$	f_i
1,70 -1,725	0,0725	12,03	15	1,995-2,045	0,0576	9,56	9
1,725-1,745	0,0543	9,02	6	2,045-2,105	0,0585	9,72	11
1,745-1,765	0,0512	8,50	8	2,105-2,185	0,0632	10,49	18
1,765-1,785	0,0480	7,97	7	2,185-2,280	0,0577	9,58	9
1,785-1,815	0,0670	11,12	13	2,280-2,41	0,0562	9,33	6
1,815-1,845	0,0611	10,14	9	2,41 -2,64	0,0590	9,79	8
1,845-1,875	0,0559	9,28	8	2,64 - ∞	0,0588	9,76	10
1,875-1,915	0,0670	11,12	11				
1,915-1,955	0,0594	9,86	11				
1,955-1,995	0,0526	8,73	7				

Bij berekening blijkt

$$t = \sum_{i=1}^{17} \frac{(f_i - np_i)^2}{np_i} = 10{,}125 \ .$$

Het aantal vrijheidsgraden van de χ^2- verdeling is 15 (want er zijn 17 klassen en we hebben één parameter geschat; vgl. paragraaf 10). De overschrijdingskans van 10,125 is dan ongeveer 0,81, zodat de exponentiële aanpassing zeer goed genoemd mag worden.

14. Enige statistische opgaven.

Door het oplossen van vraagstukken dient de lezer zich vertrouwd te maken met het gebruik van statistische methoden. In het boekje "Elementaire statistische opgaven met uitgewerkte oplossingen", door Prof.Dr. J. HEMELRIJK en Ir. D. WABEKE, J. Noorduijn & Zoon, Gorinchem 1957, vindt men honderd opgaven met oplossingen, enige tabellen en een overzicht van notaties en formules. Met een enkele uitzondering (bijv. vraagstukken over de toets van STUDENT) sluiten deze opgaven aan bij de

n deel 2 en 3 van deze leergang behandelde stof. De meeste boeken uit
e literatuurlijst aan het slot van dit deel bevatten ook een collectie
efeningen. In deze paragraaf geven wij nog enige opgaven (met oplos-
ingen) over het aanpassen van verdelingen, omdat dit onderwerp in de
eeste vraagstukken-collecties niet erg sterk vertegenwoordigd is.

pgave 1. Bij 120 worpen met een dobbelsteen zijn de volgende resul-
aten verkregen:

worpresultaat	frequentie
1	15
2	20
3	13
4	15
5	25
6	32

oets met onbetrouwbaarheid 0,05 of de dobbelsteen zuiver is.

pgave 2. Bij 54 onafhankelijke waarnemingen van een grootheid x heeft
en de volgende uitkomsten verkregen:

21,57	12,97	16,66	4,23	8,18	9,41
14,37	16,61	10,95	22,94	15,12	12,88
11,74	12,75	13,85	13,10	16,31	12,00
15,30	10,20	12,92	18,50	10,42	8,40
6,56	16,26	23,02	9,94	11,59	23,37
15,16	10,95	17,05	19,38	12,09	18,68
13,28	16,70	13,89	13,36	11,43	13,85
15,73	13,28	8,89	14,29	7,48	12,88
18,68	7,48	13,63	14,59	12,31	14,73

oets met onbetrouwbaarheid 0,05 of deze waarnemingen uit een normale
erdeling afkomstig kunnen zijn.

pgave 3. Van een stochastische grootheid x heeft men bij onafhankelijke
etingen de volgende waarden gevonden:

18,92	5,70
27,11	36,60
0,49	37,34
0,98	192,48
39,65	78,26

Ga na welke verdeling $\underline{x}$ zou kunnen hebben. Schat de verwachting en de variantie van $\underline{x}$ en van log $\underline{x}$. Opmerking: er zijn verschillende methode

Opgave 4. De exploitant van een autoverhuurbedrijf met vijf auto's heeft op honderd aselect gekozen dagen het aantal verhuurde auto's geboekt. Het bedrijf verhuurt uitsluitend auto's voor één dag (niet korter of langer). De geboekte cijfers zijn:

aantal verhuurde auto's	aantal dagen (frequentie)
0	1
1	3
2	10
3	12
4	20
5	54

a) Toets met onbetrouwbaarheid 0,05 of het aantal klanten dat per dag om een auto komt vragen een Poissonverdeling heeft. Opmerking: merk op dat het bedrijf slechts vijf auto's bezit!

b) Als de eigenaar een zesde auto aanschaft, schat dan de kans dat op een willekeurige dag alle zes auto's verhuurd worden.

Aanwijzing: Bij het bepalen van de meest aannemelijke schatting voor de parameter van een eventuele Poissonverdeling kan men gebruik maken van de volgende stelling. Als er klassen K_i $(i=1,2,\ldots,r)$ gegeven zijn zodanig dat de grootheid $\underline{x}$ met kans p_i in de klasse K_i valt, en er zijn m onafhankelijke waarnemingen van $\underline{x}$ gegeven, dan is de kans dat hiervan voor $i=1,2,\ldots,r$ precies m_i waarnemingen in de klasse K_i vallen gelijk aan

$$\frac{m!}{m_1!m_2!\ \ldots\ m_r!}\ p_1^{m_1}\ p_2^{m_2}\ \ldots\ p_r^{m_r}\ .$$

Dit is de zg. multinomiale verdeling, waarbij is verondersteld dat de m_i gehele getallen zijn die voldoen aan

$$m_i \geq 0, \qquad \sum_{i=1}^{r} m_i = m\ ,$$

terwijl voor de kansen p_i moet gelden

$$0 < p_i < 1, \qquad \sum_{i=1}^{r} p_i = 1\ .$$

oor r= 2 ontstaat de in formule (7.6) van deel 2 ingevoerde binomiale erdeling.

pgave 5. Uit de verzameling van de eerste 20.000 natuurlijke getallen eeft men aselect 490 getallen getrokken. Men heeft deze naar opklimende grootte gerangschikt en van elk paar opeenvolgende getallen het erschil x bepaald. Men vond de volgende tabel van verschillen met hun requenties:

erschil	frequentie	verschil	frequentie	verschil	frequentie
0- 4	47	75- 79	6	150-154	4
5- 9	50	80- 84	12	155-159	0
10-14	41	85- 89	6	160-164	1
15-19	40	90- 94	3	165-169	2
20-24	45	95- 99	8	170-174	0
25-29	30	100-104	5	175-179	1
30-34	30	105-109	4	180-184	1
35-39	22	110-114	4	185-189	0
40-44	20	115-119	2	190-194	0
45-49	20	120-124	2	195-199	2
50-54	19	125-129	4	200-204	0
55-59	13	130-134	1	205-209	3
60-64	13	135-139	2	210-214	0
65-69	12	140-144	1	215-219	1
70-74	9	145-149	3		

nderzoek grafisch en met de χ^2-toets of x exponentieel verdeeld kan ijn.

plossing 1. Onder de nulhypothese dat de dobbelsteen zuiver is zijn e kansen op 1,2,3,4,5 en 6 ogen allen gelijk aan $\frac{1}{6}$. Voor een χ^2-toets et zes klassen is het verwachte aantal waarnemingen dus voor elke lasse gelijk aan 20. Men vindt

$$t = \sum_{i=1}^{6} \frac{(f_i-20)^2}{20} = \frac{5^2+0^2+7^2+5^2+5^2+12^2}{20} = 13{,}4 \ .$$

nder H_0 heeft t ongeveer een χ^2-verdeling met 5 vrijheidsgraden. olgens een tabel van deze verdeling is de rechteroverschrijdingskans an 13,4 gelijk aan 0,019. Wij zullen de nulhypothese dus verwerpen en esluiten dat de dobbelsteen niet zuiver is. Men kan ook een tabel

gebruiken die bij gegeven α en gegeven aantal vrijheidsgraden de rechter kritieke waarde vermeldt. Voor $\alpha = 0,05$ en 5 vrijheidsgraden is dit 11,07, zodat de gevonden waarde 13,4 in de kritieke zone $t \geq 11,07$ valt en H_0 verworpen wordt.

Oplossing 2. Men vindt $\sum_{i=1}^{54} x_i = 741,91$ en $\sum_{i=1}^{54} x_i^2 = 11.085,5325$.
Dus geldt (zie ook (2.15))

$$\begin{cases} \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{741,91}{54} = 13,74 \ ; \\ s^2 = \frac{1}{n-1} \left\{ \sum_{i=1}^{n} x_i^2 - \frac{1}{n} \left(\sum_{i=1}^{n} x_i \right)^2 \right\} = 16,837 \ ; \\ s = 4,10 \ . \end{cases}$$

Om na te gaan of de waarnemingen uit een N(13,74;4,10)-verdeling afkomstig kunnen zijn, kan men de reële as in tien klassen verdelen die elk een verwacht aantal waarnemingen $\frac{54}{10} = 5,4$ zullen bevatten, en vervolgens de χ^2-toets gebruiken. Voor het construeren van de klassen gebruiken wij de tabel van de N(0,1)-verdeling, die aangeeft dat elk van de volgende intervallen met kans 0,1 de gestandaardiseerd normaal verdeelde grootheid u bevat:

$-\infty$ tot -1,28	0 tot 0,25
-1,28 tot -0,84	0,25 tot 0,52
-0,84 tot -0,52	0,52 tot 0,84
-0,52 tot -0,25	0,84 tot 1,28
-0,25 tot 0	1,28 tot ∞

Deze intervallen transformeren wij door met $s = 4,10$ te vermenigvuldigen en bij de uitkomsten $\bar{x} = 13,74$ op te tellen. Men vindt dan de volgende klassen en de volgende aantallen waarnemingen in die klassen:

klasse	freq.	klasse	freq.
$-\infty$ tot 8,49	6	13,74 tot 14,77	7
8,49 tot 10,30	4	14,77 tot 15,87	4
10,30 tot 11,61	5	15,87 tot 17,18	6
11,61 tot 12,71	4	17,18 tot 18,99	3
12,71 tot 13,74	10	18,99 tot ∞	5

mdat het verwachte aantal waarnemingen in elke klasse 5,4 is, vindt
en

$$t = \sum_{i=1}^{10} \frac{(f_i - 5,4)^2}{5,4} = \frac{36,40}{5,4} = 6,74 \ .$$

mdat er 10 klassen en 2 aangepaste parameters zijn heeft t onder H_0
symptotisch een χ^2-verdeling met 7 vrijheidsgraden. De rechtseenzij-
ige overschrijdingskans van 6,74 is 0,46, zodat de hypothese van
ormaliteit niet wordt verworpen. Men kan desgewenst de kritieke zone
epalen; dit blijkt het interval $t \geq 16,01$ te zijn.

plossing 3. Van de tien waarnemingen liggen er drie tussen 0 en 10,
én tussen 10 en 20 en tussen 20 en 30, drie tussen 30 en 40, terwijl
e waarnemingen 78,26 en 192,48 aanmerkelijk groter zijn. De verdeling
ijkt te scheef voor een normale en te weinig dalend voor een exponen-
iële. Wij vragen ons af of wij misschien te maken hebben met een
og-normale verdeling, hoewel de "toppen" tussen 0 en 10 en tussen 30
n 40 ook deze veronderstelling niet erg aantrekkelijk maken. Wij
ebruiken de in paragraaf 1 beschreven grafische methode; d.w.z. op
ogarithmisch waarschijnlijkheidspapier zetten wij $x_{(i)}$ uit tegen
$\frac{-0,3}{+0,4}$ (figuur 14.1).

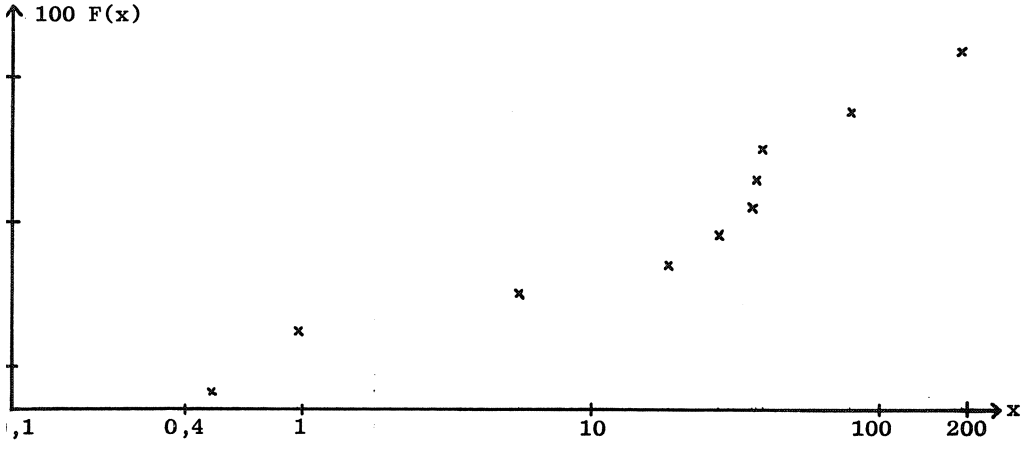


fig. 14.1

e waarnemingen van opgave 3, uitgezet op log. waarschijnlijkheidspapier

Van een fraai lineair verband is geen sprake. Toch willen wij de parameters van een eventuele log-normale verdeling trachten te schatten, en wel op drie manieren.

a) Grafische schatting. Een op het oog getrokken rechte lijn die zo goed mogelijk bij de tien punten aansluit snijdt het niveau $F(x) = 0,5$ in $x = 15$, $F(x) = 0,84$ in $x = 123$ en $F(x) = 0,16$ in $x = 1,9$. Als $\log \underline{x}$ normaal verdeeld is wordt μ geschat door $\log 15 = 2,71$ en σ door $\log 123 - \log 15 = 2,10$ of $\log 15 - \log 1,9 = 2,07$ *). Gezien de onnauwkeurigheid van zo'n grafische schatting zullen we als uitkomst opgeven $m_0 = 2,7$ en $s_0 = 2,1$.

b) Momentenschatting. Wij berekenen

$$\begin{cases} \bar{x} = \frac{1}{10} \sum_{i=1}^{10} x_i = \frac{437,53}{10} = 43,753 \ ; \\ s^2 = \frac{1}{9}\left(\sum_{i=1}^{10} x_i^2 - 10\bar{x}^2 \right) = 3273,61 \ . \end{cases}$$

Volgens paragraaf 6 vindt men de momentenschattingen voor μ en σ door $\bar{x}$ en s^2 gelijk te stellen aan de verwachting resp. de variantie van een log-normale verdeling met parameters μ en σ. Uit tabel A aan het slot van deel 2 blijkt

$$\mathcal{E}\underline{x} = e^{\mu + \frac{1}{2}\sigma^2} \ , \tag{14.1}$$

$$\sigma^2(\underline{x}) = e^{2\mu+\sigma^2}(e^{\sigma^2}-1) \ . \tag{14.2}$$

Deel (14.2) door het kwadraat van (14.1):

$$e^{\sigma^2} - 1 = \frac{\sigma^2(\underline{x})}{(\mathcal{E}\underline{x})^2} \ . \tag{14.3}$$

---

*) Zoals steeds in deze leergang hebben de logarithmen het grondtal e.

Als schattingen m_1 voor μ en s_1 voor σ vindt men door substitutie van de steekproefmomenten in (14.3) en (14.1)

$$\begin{cases} s_1^2 = \log\left(1 + \dfrac{s^2}{\bar{x}^2}\right) = \log\left(1 + \dfrac{3273{,}61}{1914{,}325}\right) = 0{,}9969 , \\ m_1 = \log \bar{x} - \dfrac{1}{2}s_1^2 = 3{,}7786 - 0{,}4985 = 3{,}2801 , \end{cases}$$

dus na afronding $m_1 = 3{,}28$ en $s_1 = 1{,}00$.

c) Meest aannemelijke schatting. Men kan direkt nagaan dat de meest aannemelijke schattingen voor μ en σ bij een log-normale verdeling ontstaan door volgens (4.10) en (4.11) de meest aannemelijke schatters voor de waarnemingen $\log x_i$ uit een normale verdeling te bepalen:

$$\begin{cases} \hat{\mu} = \dfrac{1}{10}\sum_{i=1}^{10} \log x_i , \\ \hat{\sigma}^2 = \dfrac{1}{10}\left\{\sum_{i=1}^{10} (\log x_i)^2 - 10\hat{\mu}^2\right\} . \end{cases}$$

Na opzoeken van $\log x_i$ in een tabel van natuurlijke logarithmen vindt men $\hat{\mu} = 2{,}78$ en $\hat{\sigma} = 1{,}80$.

De drie schattingen voor μ en σ vertonen zoals men ziet nogal grote verschillen. Een log-normale verdeling past volgens figuur 14.1 nogal slecht bij de waarnemingen en men zou kunnen zeggen dat elke schattingsmethode op deze afwijking anders reageert. Het merkwaardige is dat wij bij het opstellen van dit vraagstuk wel degelijk van een log-normale verdeling zijn uitgegaan. In een tabel van aselecte trekkingen uit een N(0,1)-verdeling *) hebben wij een willekeurige plaats gekozen en vanaf het aangewezen getal tien opvolgende getallen u_i overgenomen. Deze getallen hebben wij getransformeerd met de formule

$$x_i = e^{2u_i+3} .$$

---

*) H. WOLD, Tracts for Computers, no. XXV: Random Normal Deviates, Dept. of Statistics, University of London, Cambridge University Press, Cambridge, 1948.

Omdat de getallen u_i trekkingen uit een N(0,1)-verdeling zijn, stellen de getallen x_i een aselecte steekproef uit een log-normale verdeling met parameters $\mu = 3$ en $\sigma = 2$ voor, immers log $\underline{x}$ is N(3,2)-verdeeld. De getallen x_i zijn echter volgens figuur 14.1 niet erg representatief voor een steekproef uit een log-normale verdeling. De parameters en hun schattingen waren:

ware waarde	grafisch	momenten	meest aannemelijk
$\mu = 3$	$m_0 = 2{,}7$	$m_1 = 3{,}28$	$\hat{\mu} = 2{,}78$
$\sigma = 2$	$s_0 = 2{,}1$	$s_1 = 1{,}00$	$\hat{\sigma} = 1{,}80$

Uit een en ander blijkt dat de resultaten uit kleine steekproeven niet altijd betrouwbaar zijn.

Het aantal waarnemingen is zo klein dat een χ^2-toets voor aanpassing weinig zinvol is. Met minstens vijf waarnemingen per klasse zouden wij hoogstens twee klassen kunnen vormen, en na het aanpassen van twee parameters zou van dit aantal drie moeten worden afgetrokken!

Oplossing 4. Op 54 van de 100 dagen waren alle vijf auto's verhuurd, maar het is niet bekend hoeveel klanten zich op deze dagen hebben aangemeld. Het kunnen er precies vijf geweest zijn, maar ook meer dan vijf. Dit geeft enige moeilijkheden bij het schatten van de parameter van de Poissonverdeling.

Als alle 54 waarnemingen $\underline{x} \geq 5$ precies $\underline{x} = 5$ zouden betekenen, was het steekproefgemiddelde

$$\frac{1}{100}(1.3 + 2.10 + 3.12 + 4.20 + 5.54) = 4{,}09 \ .$$

De werkelijke waarde van het steekproefgemiddelde is dus minstens 4,09. Proberen wij een aantal waarden voor de parameter λ die hier niet te ver boven liggen, dan geeft een tabel van de Poissonverdeling voor de grootheden

$$(14.4) \qquad \begin{cases} p_k \overset{\text{def}}{=} P\left[\underline{x} = k \mid \lambda\right] = \dfrac{\lambda^k e^{-\lambda}}{k!} & (k=0,1,\ldots,4) \\ p_5 \overset{\text{def}}{=} P\left[\underline{x} \geq 5 \mid \lambda\right] = \displaystyle\sum_{k=5}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \ , \end{cases}$$

e volgende waarden

	$\lambda = 4{,}5$	$\lambda = 5$	$\lambda = 5{,}5$	$\lambda = 6$	$\lambda = 7$
$100p_0$	1,1	0,7	0,4	0,2	0,1
$100p_1$	5,0	3,4	2,2	1,5	0,6
$100p_2$	11,2	8,4	6,2	4,5	2,2
$100p_3$	16,9	14,0	11,3	8,9	5,2
$100p_4$	19,0	17,5	15,6	13,4	9,1
$100p_5$	46,8	56,0	64,2	71,5	82,7

Omdat de kolom $\lambda = 5$ het meest lijkt op de waargenomen aantallen ,3,10,12,20,54 is $\lambda = 5$ een niet onredelijke schatting. Wij zullen nu e meest aannemelijke schatting voor λ bepalen om na te gaan of die nigszins met deze door proberen gevonden waarde overeenstemt. Omdat e rubriek $\underline{x} \geq 5$ niet nader onderverdeeld is, gebruiken wij de in de anwijzing genoemde stelling.

Als $\underline{x}$ een Poissonverdeling met parameter λ heeft, zijn de groot-eden p_k gegeven door (14.4). De kans dat er van n waarnemingen n_k de aarde k hebben (k=0,1,...,4) en n_5 een waarde ≥ 5, is dan

$$L(\lambda) = P\left[\underline{n}_0=n_0, \underline{n}_1=n_1, \underline{n}_2=n_2, \underline{n}_3=n_3, \underline{n}_4=n_4, \underline{n}_5=n_5 \mid \lambda\right] =$$

$$= \frac{n!}{n_0!n_1!n_2!n_3!n_4!n_5!} \prod_{k=0}^{4} \left(\frac{\lambda^k e^{-\lambda}}{k!}\right)^{n_k} \left(\sum_{k=5}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!}\right)^{n_5} .$$

De meest aannemelijke schatting voor λ vindt men door $\frac{d \log L}{d\lambda} = 0$ e stellen en de waarden $n = 100$, $n_0 = 1$, $n_1 = 3$, $n_2 = 10$, $n_3 = 12$, $_4 = 20$, $n_5 = 54$ in te vullen.

$$\text{og } L = \text{constante} + \sum_{k=0}^{4} n_k(k \log \lambda - \lambda - \log k!) + n_5\left(-\lambda + \log \sum_{k=5}^{\infty} \frac{\lambda^k}{k!}\right).$$

$$\frac{\log L}{d\lambda} = \frac{1}{\lambda} \sum_{k=0}^{4} kn_k - \sum_{k=0}^{4} n_k + n_5 \left(-1 + \frac{\sum_{k=5}^{\infty} \frac{k\lambda^{k-1}}{k!}}{\sum_{k=5}^{\infty} \frac{\lambda^k}{k!}} \right) .$$

In de teller van de laatste breuk kiest men j = k-1 als sommatie-index; de faktor tussen haken wordt dan

$$\frac{-\sum_{k=5}^{\infty}\frac{\lambda^k}{k!}+\sum_{j=4}^{\infty}\frac{\lambda^j}{j!}}{\sum_{k=5}^{\infty}\frac{\lambda^k}{k!}}=\frac{\frac{\lambda^4}{24}}{e^{\lambda}-(1+\lambda+\frac{\lambda^2}{2}+\frac{\lambda^3}{6}+\frac{\lambda^4}{24})}\ .$$

Na substitutie van de gegeven waarden voor n_k vindt men

$$0=\frac{d\log L}{d\lambda}=\frac{139}{\lambda}-46+\frac{54}{24}\,\frac{\lambda^4}{e^{\lambda}-(1+\lambda+\frac{\lambda^2}{2}+\frac{\lambda^3}{6}+\frac{\lambda^4}{24})}\ .$$

Vermenigvuldigen met $\frac{\lambda}{46}$ geeft: $\lambda = f(\lambda)$, waarbij

$$f(\lambda)\overset{\text{def}}{=}\frac{139}{46}+\frac{54}{24.46}\,\frac{\lambda^5}{e^{\lambda}-(1+\lambda+\frac{\lambda^2}{2}+\frac{\lambda^3}{6}+\frac{\lambda^4}{24})}\ .$$

De meest aannemelijke schatting $\hat{\lambda}$ is dus de oplossing van de transcendente vergelijking $\lambda = f(\lambda)$. Direkte oplossing van deze vergelijking is niet mogelijk, maar als men over een beginschatting λ_0 beschikt, kan men achtereenvolgens bepalen

$$\mu_0 = f(\lambda_0);\quad \lambda_1 = \tfrac{1}{2}(\lambda_0+\mu_0);$$
$$\mu_1 = f(\lambda_1);\quad \lambda_2 = \tfrac{1}{2}(\lambda_1+\mu_1);$$
$$\text{enz.}$$

Het is geenszins zeker dat deze iteratieve procedure convergeert, d.w.z. dat de opeenvolgende getallen λ_j steeds dichter bij elkaar komen te liggen. De meeste procedures van deze aard convergeren òf zeer snel òf helemaal niet, en het hangt vooral van de beginschatting af of er convergentie optreedt.

Het lijkt niet onverstandig als beginschatting de door proberen gevonden waarde $\lambda_0 = 5$ te gebruiken. Men berekent dan achtereenvolgens

$$\mu_0 = f(5) = 4{,}8624;\quad \lambda_1 = 4{,}9312;$$
$$\mu_1 = f(4{,}9294) = 4{,}9024;\quad \lambda_2 = 4{,}9168.$$

mdat λ al vrijwel constant blijft staken wij hier de berekening. Als fgeronde waarde voor $\hat{\lambda}$ gebruiken wij 4,9.

Om na te gaan of het aantal klanten per dag aansluit bij een oissonverdeling passen wij de χ^2-toets toe.

klasse	$100p_j$	f_j
0 en 1	4,39	4
2	8,94	10
3	14,60	12
4	17,89	20
$\geq$ 5	54,18	54

e kolom $100p_j$ is gevonden uit een tabel van de Poissonverdeling voor = 4,9; de klassen 0 en 1 zijn samengevoegd om het verwachte aantal aarnemingen niet te klein te maken. Voor de toetsingsgrootheid vindt en t = 0,87. Asymptotisch heeft $\underline{t}$ een χ^2-verdeling met drie vrijheidsraden (nl. vijf klassen en één aangepaste parameter). De overschrijingskans van 0,87 is dan 0,83, zodat een Poissonverdeling heel goed ast bij de waarnemingen.

Het antwoord bij vraag b) is

$$P\left[\underline{x} \geq 6 \mid \lambda = 4,9\right] = 0,3665 .$$

plossing 5. Om na te gaan of $\underline{x}$ bij benadering exponentieel verdeeld is, ebruiken wij evenals in paragraaf 13 enkelvoudig logarithmisch papier. angs de lineaire schaal (horizontaal) zetten wij de verschillen x uit, n langs de logarithmische schaal het aantal verschillen dat groter dan f gelijk aan x is; wij noemen deze grootheid n(x). Omdat de gegevens egroepeerd zijn kunnen wij voor x alleen gehele veelvouden van 5 kiezen.

De gevonden punten liggen in goede benadering op een rechte lijn. p het grafiekenpapier kunnen wij aflezen dat deze lijn het niveau og 1 (dus nul) ongeveer snijdt in x = 260 en de y-as in y = log 489, odat zijn richtingscoëfficiënt ongeveer gelijk is aan

$$-\frac{\log 489}{260} = -\frac{6,1936}{260} = -0,0238.$$

oals in paragraaf 13 is opgemerkt is de richtingscoëfficiënt van de 'ideale lijn" gelijk aan $-\lambda$, zodat wij de parameter van de exponentiële erdeling grafisch schatten als 0,024.

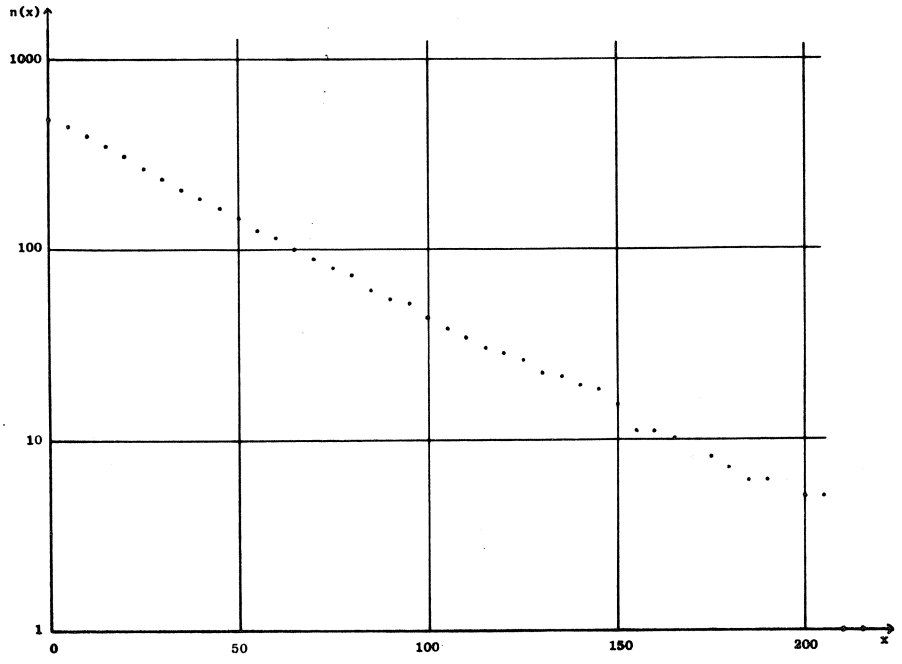


fig. 14.2

De waarnemingen van opgave 5 uitgezet op enkelvoudig logarithmisch papi

Vervolgens willen wij λ schatten uit het gemiddelde van de gegeven verschillen. Omdat deze gegroepeerd gegeven zijn met een klassebreedte van vijf, nemen wij voor elke klasse het midden (dus 2,7,12,...) en vermenigvuldigen deze getallen met de bijbehorende frequenties (47,50, 41,...). De som van deze produkten blijkt dan 20033 te zijn. Omdat het aantal verschillen 489 is vinden wij als schatting voor de verwachting van de exponentiële verdeling

$$\frac{20033}{489} = 40{,}967 \ .$$

Volgens tabel A aan het slot van deel 2 is deze verwachting juist $\frac{1}{\lambda}$. Volgens de momentenmethode (paragraaf 6) is dan

$$\frac{1}{40{,}967} = 0{,}02441$$

een schatting voor λ.

Tenslotte gebruiken wij de χ^2-toets om na te gaan of $\underline{x}$ exponen-
ieel verdeeld kan zijn. Wij gebruiken in principe de gegeven indeling
n klassen ter breedte 5, maar voegen vooral aan het slot een aantal
lassen samen om het verwachte aantal waarnemingen per klasse (bij de
xponentiële verdeling met parameter 0,02441) niet lager dan ongeveer
0 te laten worden, in overeenstemming met tabel 10.1. Zo vinden wij
e onderstaande tabel.

klasse	$489p_i$	f_i
0 tot 4,5	50,87	47
4,5 tot 9,5	50,34	50
9,5 tot 14,5	44,55	41
14,5 tot 19,5	39,44	40
19,5 tot 24,5	34,90	45
24,5 tot 29,5	30,90	30
29,5 tot 34,5	27,34	30
34,5 tot 39,5	24,21	22
39,5 tot 44,5	21,42	20
44,5 tot 49,5	18,96	20
49,5 tot 54,5	16,78	19
54,5 tot 64,5	28,00	26
64,5 tot 74,5	21,94	21
74,5 tot 89,5	24,33	24
89,5 tot 109,5	21,25	20
109,5 tot 149,5	21,05	19
149,5 tot ∞	12,72	15

De toetsingsgrootheid voor de χ^2-toets blijkt bij berekening de
aarde 5,31 te hebben. Omdat er zeventien klassen zijn en één aangepaste
arameter, is $\underline{t}$ onder H_0 ongeveer χ^2-verdeeld met vijftien vrijheids-
raden. De overschrijdingskans van 5,31 is dan ongeveer 0,99. De aan-
assing blijkt dus uitermate goed te zijn.

Literatuur over statistiek

Elementaire inleidingen

J.L. HODGES Jr. and E.L. LEHMANN, Basic Concepts of Probability and Stati tics, Holden-Day, San Francisco, 1964 [een voortreffelijke inleiding].

W.A. WALLIS and H.V. ROBERTS, Statistics: A new approach, The Free Press, Glencoe, Illinois, 1956 [een boek met talrijke uitgewerkte voorbeelden en vraagstukken].

H. DE JONGE, Inleiding tot de Medische Statistiek, Nederlands Instituut v Praeventieve Geneeskunde, Leiden; deel 1 (2e druk; 1963); deel 2 (1960) ; [een goed "receptenboek" zonder veel theorie].

W.J. DIXON and F.J. MASSEY, An Introduction to Statistical Analysis, McGraw-Hill, New York, 2nd edition, 1957 [idem].

M.L. WIJVEKATE, Verklarende Statistiek, Aula-pocket no. 39, Het Spectrum, Utrecht, 1961 en latere drukken [een populaire inleiding voor mensen met weinig wiskundekennis].

M.J. MORONEY, Facts from Figures, Pelican Book no. A 236, Harmondsworth, 3rd edition, 1962; Nederlandse vertaling: Feiten uit Cijfers; hoe te werk met formules, grafieken en tabellen, Marka-pocket no. 71, Het Spectrum, Utrecht, 1967 [een populaire inleiding voor mensen met weinig wiskundeken

Wiskundige leerboeken en handboeken

M.G. KENDALL and A. STUART, The advanced Theory of Statistics, Griffin & London; volume 1, 1958; volume 2, 1960; volume 3, 1966 [een belangrijk ha boek].

H. CRAMÉR, Mathematical Methods of Statistics, Princeton University Press Princeton, 1946 [een duidelijk, maar enigszins verouderd leerboek].

S.S. WILKS, Mathematical Statistics, Wiley, New York, 1962 [een modern handboek].

L. LEHMANN, Testing Statistical Hypotheses, Wiley, New York, 4th edition, 66 [in dit boek worden de algemene principes van de toetsingstheorie bendeld].

SVERDRUP, Laws and Chance Variations - Basic Concepts of Statistical ference, North Holland Publishing Cy., Amsterdam; volume I, 1967; lume II, 1967 [een goed leerboek].

M. MOOD and F.A. GRAYBILL, Introduction to the Theory of Statistics, Graw-Hill, New York, 2nd edition 1963 [een goed leerboek; ook verkrijgar in de uitvoering van een International Student Edition].