STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

De Gereduceerde χ^2-Methode

J. Hemelrijk

MC
1950

J. HEMELRIJK

DE GEREDUCEERDE χ^2-METHODE

OVERDRUK UIT STATISTICA, JAARGANG 4, No. 6 1950

Overdruk uit Statistica, Jaargang 4, no. 6, 1950.

De gereduceerde χ^2-methode *)

door J. Hemelrijk
Mathematisch Centrum, Amsterdam

Summary

The method of reduced chi-square.

In these notes the author endeavours to describe in elementary form the theory of reduced chi-square, which formed the subject of two lectures given by Professor N e y m a n *in Amsterdam on the 8th and 9th of May, 1950.*

If the theoretical frequencies occurring as denominators in the usual expression of χ^2 are replaced by the observed frequencies, we obtain an expression χ'^2 which asymptotically possesses a χ^2-distribution. If the theoretical frequencies depend on a number of parameters these may be estimated by reducing either χ^2 or χ'^2 to a minimum; mathematically the latter method is often much more tractable.

The method of reduced χ^2 is illustrated by a simple example in § 2. The importance of this theory is explained in § 3. In §§ 4 and 5 the general formulation of N e y m a n's *theory and its application to the estimation of parameters and the testing of hypotheses are briefly summarized.*

§ 1. *Inleiding.*

De door Prof. N e y m a n in zijn voordrachten behandelde stof kan volledig teruggevonden worden in zijn artikel ,,Contribution to the theory of χ^2-tests'' [1]. Dit artikel is echter slechts leesbaar voor gevorderde mathematici, terwijl de door N e y m a n ontwikkelde theorie, die een generalisatie is van de minimum-χ^2-methode, zeer goed toegepast kan worden zonder bestudering van de vrij moeilijke bewijzen. Het is daarom wellicht gerechtvaardigd de resultaten van N e y m a n's onderzoek te releveren en toe te lichten aan de hand van een door hem in bovengenoemde twee voordrachten gebruikt voorbeeld. De bewijzen kunnen daarbij, evenals bij de voordrachten het geval was, gevoegelijk achterwege gelaten worden, daar deze alle in bovengenoemd artikel te vinden zijn.

§ 2. *Voorbeeld.*

2.1. Ter verduidelijking van de aard der problemen, die met de gereduceerde χ^2-methode kunnen worden opgelost, bespreken wij eerst een eenvoudig voorbeeld.

---

*) Deze aantekeningen zijn geschreven naar aanleiding van de beide voordrachten die Professor J. N e y m a n (Berkeley, California) op 8 en 9 Mei 1950 te Amsterdam heeft gehouden voor de Mathematisch-Statistische Sectie van de Vereniging voor Statistiek en voor het Mathematisch Centrum.

In een ziekenhuis worden de verpleegsters op geregelde tijden (met gelijke tussenpozen) onderzocht op de aanwezigheid van antistoffen tegen TBC-bacillen (Pirquet-reactie). Indien deze reactie eenmaal positief is uitgevallen, d.w.z. als de aanwezigheid van de antistoffen éénmaal is aangetoond, is dit bij een later onderzoek ook steeds het geval. Daarom worden bij ieder onderzoek alleen die verpleegsters opnieuw onderzocht, die bij het vorige onderzoek nog negatief reageerden. Overgang van negatieve naar positieve Pirquet-reactie vindt plaats, indien in de tussentijd besmetting is opgetreden.

Het eerste onderzoek doet dienst als vooronderzoek en zal dus verder buiten beschouwing gelaten worden; de verpleegsters, die bij dit voor-onderzoek reeds positief reageren, worden niet in aanmerking genomen. Laten wij het aantal der verpleegsters, die bij het voor-onderzoek negatief reageren n noemen en onderstellen, dat er na de voorcontrôle in het geheel $m—1$ maal onderzocht wordt. Van iedere verpleegster wordt genoteerd, bij welk onderzoek zij voor het eerst positief reageert, of eventueel, dat zij bij het laatste onderzoek nog negatief is. De vragen, die men met behulp van deze gegevens wenst te beantwoorden, zijn:

1e. Zijn voor alle verpleegsters de kansen op besmetting gelijk, of verschillend?

2e. Indien deze kansen voor alle verpleegsters redelijkerwijs gelijk mogen worden geacht, kan men dan een schatting van de grootte daarvan geven?

2.2. Uiteraard zullen deze vragen nader moeten worden gepreciseerd. Wij zullen in de eerste plaats nader ingaan op de eerste vraag. Uit statistisch oogpunt bezien, komt dit probleem neer op de *toetsing van de hypothese H*, dat de kans op besmetting voor alle verpleegsters dezelfde is. Indien wij onderstellen, dat H juist is, en verder aannemen, dat de kans op besmetting gedurende een zeker tijdsbestek evenredig is met de grootte daarvan *), en deze kans voor het tijdsbestek tussen twee op elkaar volgende onderzoeken gelijk aan p stellen, kunnen wij de kans, dat een verpleegster bij het i^e onderzoek voor het eerst een positieve Pirquet-reactie vertoont, in de onbekende parameter p uitdrukken. Noemen wij deze kans p_i dan geldt voor alle n verpleegsters, die bij het voor-onderzoek negatief reageren,

$$p_i = (1 - p)^{i-1}\, p \qquad i = 1, \ldots, m-1 \qquad (1)$$

*) In de discussie werd deze onderstelling terecht nogal bezwaarlijk geacht. Men kan dit bezwaar ondervangen door de onderstelling in de hypothese H op te nemen. Hierdoor zou het probleem echter aanzienlijk gecompliceerder worden. Wij zullen daarom niet verder op deze kwestie ingaan en de onderstelling handhaven.

terwijl de kans, dat in het geheel geen besmetting optreedt, zodat ook het $(m-1)^e$ onderzoek negatief uitvalt, gelijk is aan

$$p_m = (1 - p)^{m-1}. \tag{2}$$

Op deze wijze is de hypothese H uitgedrukt in m waarschijnlijkheden $p_1, \ldots, p_m$, met

$$\sum_{j=1}^{m} p_j = 1. \tag{3}$$

Bij uitvoering van het onderzoek vindt men m frequentiequotiënten $q_1, \ldots, q_m$, waarbij nq_i het aantal der verpleegsters voorstelt, die bij het i^e onderzoek voor het eerst (resp. in het geheel niet) positief reageren. Deze waarden $q_1, \ldots, q_m$ kunnen wij beschouwen als waarnemingen van m stochastische variabelen $\underline{q_1}, \ldots, \underline{q_m}$ *), die in de onderstelling, dat de hypothese H (of ook een van H verschillende hypothese) juist is, zekere waarschijnlijkheidsverdelingen bezitten.

Beschouwen wij nu de grootheid

$$\underline{\chi}^2{}_1 = n \sum_{j=1}^{m} \frac{(\underline{q_j} - p_j)^2}{p_j}, \tag{4}$$

en is $\underline{\hat{p}}$ die waarde van p, waarvoor deze uitdrukking minimaal is, dan is

$$\underline{\chi}^2_{1,H} = n \sum_{j=1}^{m} \frac{(\underline{q_j} - \underline{\hat{p}_j})^2}{\underline{\hat{p}_j}}, \tag{5}$$

waarin $\underline{\hat{p}_j}$ de waarde is, die p_j aanneemt voor $p = \underline{\hat{p}}$, een stochastische variabele, die, indien de hypothese H juist is, voor $n \to \infty$ asymptotisch verdeeld is volgens een χ^2-verdeling met $m - 2$ vrijheidsgraden (zie b.v. H. Cramér [2], p. 424).

Deze eigenschap kan men klaarblijkelijk gebruiken om de hypothese H te toetsen. Indien de gevonden waarde $\chi^2_{1,H}$ van $\underline{\chi}^2_{1,H}$ (die uit (4) en (5) verkregen wordt door voor $\underline{q_1}, \ldots, \underline{q_m}$ de gevonden waarden $q_1, \ldots, q_m$ in te vullen), zó groot is, dat de kans op deze of een nog grotere waarde kleiner is dan een van tevoren gegeven getal α (de onbetrouwbaarheidsdrempel), wordt H verworpen.

Neyman heeft nu bewezen, dat ook de grootheid

$$\underline{\chi}^2_{2,H} = \operatorname*{Min}_{0 < p \leq 1} n \sum_{j=1}^{m} \frac{(\underline{q_j} - p_j)^2}{\underline{q_j}} \tag{6}$$

*) Een stochastische variabele, d.i. een variabele die een waarschijnlijkheidsverdeling bezit, wordt aangegeven door een *onderstreepte* letter. Een door een stochastische variabele aangenomen waarde wordt aangegeven door dezelfde letter *zonder onderstreping*.

asymptotisch voor $n \to \infty$ verdeeld is volgens dezelfde χ^2-verdeling als $\underline{\chi}^2_{1,B}$. Ook deze grootheid kan dus gebruikt worden om H te toetsen. Nu staat echter q_j in de noemer in plaats van p_j. Indien het experiment is uitgevoerd zijn de door de $\underline{q}_j$ aangenomen waarden bekend, zodat de minimalisering veel eenvoudiger is dan in het vorige geval, waarbij de (niet stochastische, maar onbekende) variabelen p_j in de noemer staan; immers deze variabelen bevatten de p, waarover geminimaliseerd moet worden. Voor de toepassing van (6) is het echter noodzakelijk, dat alle $q_j \neq 0$ zijn. Het asymptotisch karakter van de toetsingsmethode legt overigens ook reeds de eis op, dat de q_j niet te klein mogen zijn.

2.3. Neyman heeft deze theorie nog verder ontwikkeld tot een methode om H te toetsen tegen een *gegeven klasse van alternatieve hypothesen*, die we met Ω zullen aangeven. In het onderhavige geval bestaat Ω uit hypothesen, inhoudende dat p niet voor alle verpleegsters gelijk is, maar stochastisch is en op de populatie van alle verpleegsters een bepaalde waarschijnlijkheidsverdeling bezit. Legt men aan deze waarschijnlijkheidsverdeling geen enkele beperking op, dan verkrijgt men het in 2.2 behandelde speciale geval, waarbij geen klasse van alternatieve hypothesen gespecificeerd is. Laat men echter slechts een speciale klasse van waarschijnlijkheidsverdelingen voor $\underline{p}$ toe (die dan tezamen de klasse Ω van alternatieve hypothesen vormen), dan verkrijgt men een toetsingsmethode die in het algemeen met betrekking tot de klasse Ω een groter onderscheidingsvermogen („power-function'') bezit, dan de eerstgenoemde. Dat wil zeggen, dat bij deze nieuwe methode de hypothese H, indien deze onjuist is en indien de *juiste* hypothese tot Ω behoort, sneller zal worden verworpen dan bij de oude.

Voor Ω zal men dus een klasse van alternatieve hypothesen kiezen, die „niet te gek'' worden geacht. In het onderhavige geval was Neyman b.v. geneigd alternatieve hypothesen, inhoudende, dat $\underline{p}$ een waarschijnlijkheids-

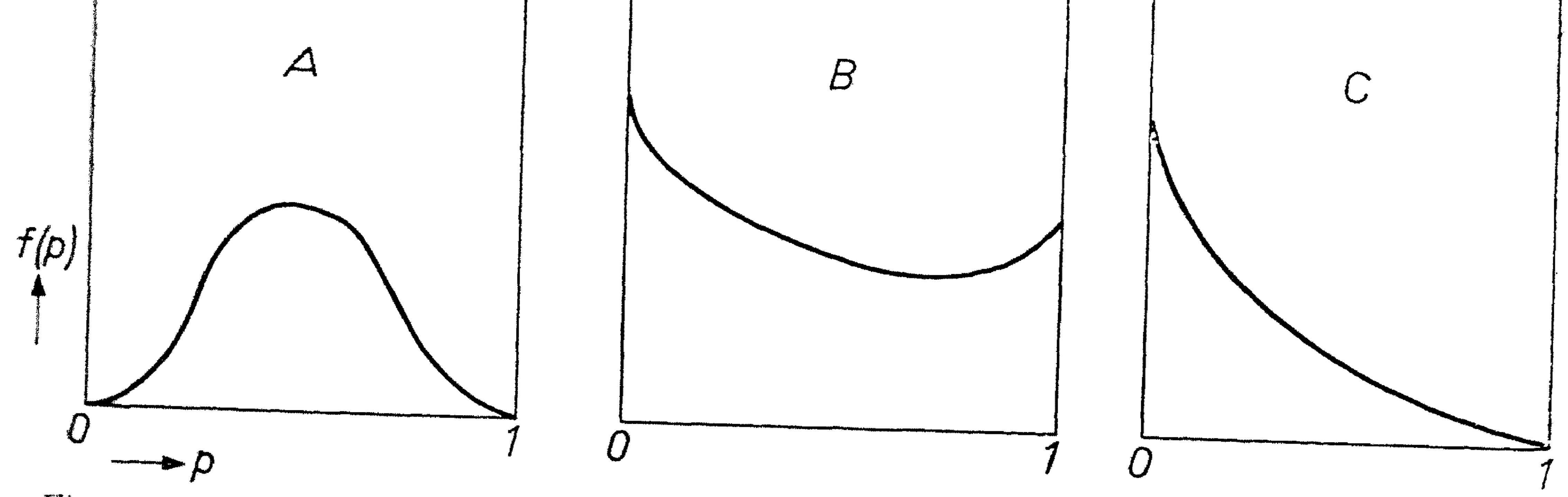

Fig. 1. Toegelaten alternatieve hypothesen; $f(p)$ is de waarschijnlijkheidsdichtheid van $\underline{p}$.

dichtheid van één van de in figuur 1 gegeven vormen bezit, aannemelijk te achten, maar een waarschijnlijkheidsdichtheid van een type als in figuur 2 geschetst, niet tot Ω toe te laten.

Beperkt men nu Ω tot waarschijnlijkheidsverdelingen als die van figuur 1, dan kan men deze gehele klasse vrij goed representeren door bêta verdelingen, d.w.z. door als waarschijnlijkheidsdichtheid te nemen de integrand van de onvolledige bêta-functie:

$$\frac{p^{\theta_1-1}(1-p)^{\theta_2-1}}{B(\theta_1, \theta_2)}, \qquad 0 \leqq p \leqq 1 \tag{7}$$

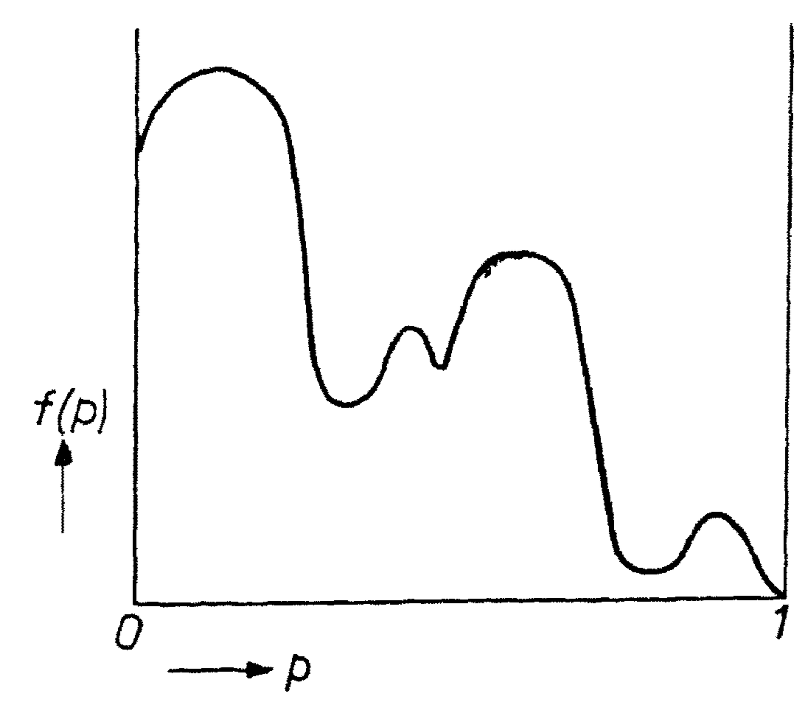

Fig. 2. Niet toegelaten alternatieve hypothesen; $f(p)$ is de waarschijnlijkheidsdichtheid van p.

waarin θ_1 en θ_2 twee reële positieve parameters zijn en $B(\theta_1, \theta_2)$ de bêta-functie voorstelt. Voor $\theta_1 > 2$ en $\theta_2 > 2$ b.v. verkrijgt men het meest linkse van de in figuur 1 aangegeven typen, voor $0 < \theta_1 < 1$ en $0 < \theta_2 < 1$ een U-vormige verdeling, die op het middelste type van figuur 1 lijkt en voor $0 < \theta_1 < 1$ en $\theta_2 > 2$ een J-vormig type, dat op de meest rechtse tekening van figuur 1 lijkt.

Neyman neemt (7) als klasse Ω en berekent daaruit de gemiddelde kans p_i', over alle verpleegsters genomen, op besmetting tussen het $(i-1)^e$ en het i^e onderzoek. Deze is voor $1 \leqq i \leqq m-1$:

$$p'_i = \int_0^1 \frac{p^{\theta_1-1}(1-p)^{\theta_2-1}}{B(\theta_1, \theta_2)} (1-p)^{i-1}\, pdp =$$

$$= \frac{B(\theta_1+1, \theta_2+i-1)}{B(\theta_1, \theta_2)} = \frac{\theta_1(\theta_2+i-2)!\,(\theta_1+\theta_2-1)!}{(\theta_2-1)!\,(\theta_1+\theta_2+i-1)!}, \tag{8}$$

terwijl

$$p'_m = 1 - \sum_{j=1}^{m-1} p'_j \tag{9}$$

is. Hierin zijn θ_1 en θ_2 dus nog onbekende parameters.

Naast de door (5) en (6) gedefinieerde vormen beschouwt Neyman dan nog de vormen

$$\underline{\chi}^2_{1,\Omega} = \underset{\substack{\theta_1>0\\ \theta_2>0}}{\mathrm{Min}}\; n \sum_{j=1}^{m} \frac{(q_j - p_j')^2}{p'_j} \tag{10}$$

en

$$\underline{\chi}^2{}_{2,\,\Omega} = \underset{\substack{\theta_1>0\\ \theta_2>0}}{\mathrm{Min}}\; n \sum_{j=1}^{m} \frac{(\underline{q}_j - p'_j)^2}{\underline{q}_j} \; . \tag{11}$$

De door N e y m a n bewezen stelling, waarop nu de toetsing van H tegen de klasse Ω van alternatieve hypothesen berust, houdt nu in dat

$$\underline{\chi}^2{}_{1,\,H} - \underline{\chi}^2{}_{1,\,\Omega} \tag{12}$$

en

$$\underline{\chi}^2{}_{2,\,H} - \underline{\chi}^2{}_{2,\,\Omega} \tag{13}$$

beide asymptotisch (voor $n \to \infty$) verdeeld zijn volgens een χ^2-verdeling met 1 vrijheidsgraad, indien de hypothese H juist is. Het *aantal vrijheidsgraden* is daarbij berekend als het aantal vrije parameters binnen Ω (in casu 2, nl. θ_1 en θ_2) verminderd met het aantal vrije parameters binnen H (i.c. 1, nl. p).

Men ziet tevens, dat het in 2.2 behandelde geval, waarbij geen alternatieve hypothesen werden gespecificeerd een speciaal geval van het hier besprokene is. Immers, indien aan de waarschijnlijkheidsverdeling van $\underline{p}$ geen beperking wordt opgelegd, kan men $p'_j = q_j$ nemen voor iedere j, zodat zowel $\chi^2{}_{1,\,\Omega}$ als $\chi^2{}_{2,\,\Omega}$ gelijk aan nul worden en (12) en (13) overgaan in (5) en (6). Het aantal vrije parameters in Ω is dan $m - 1$ (nl. $p'_1, \ldots, p'_{m-1}$, daar $\sum_1^m p'_j = 1$ is), zodat het aantal vrijheidsgraden $m - 2$ is, zoals reeds in 2.2 is opgemerkt.

2.4. Indien nu H niet voor verwerping in aanmerking blijkt te komen op grond van het uitgevoerde onderzoek, kan men de tweede in 2.1 vermelde vraag stellen: de grootte van p te schatten. Als *schattingen* van p neemt N e y m a n nu de grootheid $\underline{\hat{p}}$ (uit (5)) of de overeenkomstige grootheid $\underline{p}^*$, die de in (6) vermelde vorm minimaal maakt. Wordt H wel verworpen, dan kunnen θ_1 en θ_2 op analoge wijze geschat worden uit (10) of (11).

Zowel van de toetsingsmethoden als van deze schattingen heeft N e y m a n bepaalde eigenschappen aangetoond, die verderop kort besproken worden (zie § 5).

§ 3. *Betekenis van de theorie.*

De betekenis van deze theorie ligt in haar grote algemeenheid. Zeer vele problemen kunnen op deze wijze worden aangepakt. N e y m a n formuleert de theorie in een nog algemenere vorm, waarbij niet één, maar s series waarnemingen (in het voorbeeld s ziekenhuizen) beschouwd worden. Men kan dan b.v. onderzoeken, of de besmettingskans in de verschillende ziekenhuizen dezelfde is, of niet. Wij zullen hier niet op in gaan; de algemene mathematische formulering vindt men in § 4 besproken. Daar deze minder gemakkelijk te overzien is, hebben wij ons tot hier toe beperkt tot het bovenstaande speciale geval, waarin de gedachtengang van de methode volledig tot uitdrukking komt.

Een tweede punt van belang is, dat Neyman in deze theorie de schattings- en de toetsingsproblemen gezamenlijk oplost. Eenzelfde verband tussen schattings- en toetsingsmethoden vindt men in de theorie der aannemelijkste schattingen („theory of maximum likelihood") en de daarmee verbonden theorie der λ-toetsingsmethoden, die echter niet tegelijkertijd ontwikkeld zijn (de theorie der aannemelijkste schattingen is van R. A. Fisher en die der λ-toetsingsmethoden van J. Neyman en E. S. Pearson). Zoals uit de in § 5 opgesomde eigenschappen van de door Neyman afgeleide schattingen en toetsingsmethoden blijkt, worden de theorie der aannemelijkste schattingen, de λ-theorie, de minimum-χ^2-methode als schattings- en toetsingsmethode en de hier beschreven generalisatie van deze laatste theorie alle onder één hoofd samengevat; de schattingen zijn alle BAN-schattingen (vgl. § 5) en de toetsingsmethoden zijn alle asymptotisch aequivalent met de λ-toetsingsmethode.

Een nadeel van de theorie is, zoals Prof. Neyman opmerkte, dat de gehele theorie slechts asymptotisch geldt, d.w.z. dat de eigenschappen der schattingen slechts voor $n \to \infty$ gelden, evenals de toetsingsmethoden, terwijl men niet weet, hoe groot het aantal waarnemingen moet zijn, om deze asymptotische eigenschappen voldoende dicht te benaderen.

Dit neemt echter niet weg, dat deze nieuwe theorie als een zeer belangrijke bijdrage tot de schattings- en toetsingstheorie moet worden gezien.

§ 4. *Algemene formulering.*

4.1. De door Neyman gegeven algemene formulering van de theorie luidt als volgt:

Beschouw s onafhankelijke series van onafhankelijke experimenten (in het bovenstaande is $s = 1$), waarbij de i^e serie ($i = 1, \ldots, s$) bestaat uit n_i experimenten, die ieder één der m_i resultaten

$$R_{i,1}, \ldots, R_{i,m_i} \tag{14}$$

opleveren, met een kans $p_{i,j}$ op $R_{i,j}$, zodat

$$p_{i,1} + \ldots + p_{i,m_i} = 1 \qquad (i = 1, \ldots, s) \tag{15}$$

is. Laat het aantal experimenten, dat $R_{i,j}$ tot resultaat heeft, gelijk zijn aan $n_i \underline{q}_{i,j}$, zodat

$$\underline{q}_{i,1} + \ldots + \underline{q}_{i,m_i} = 1 \qquad (i = 1, \ldots, s) \tag{16}$$

is. De kansen $p_{i,j}$ zijn onbekend, de frequentiequotiënten $\underline{q}_{i,j}$ zijn stochastisch, maar de bij een experiment aangenomen waarden zijn bekend.

Zij verder

$$p_{i,j} \neq 0 \qquad \left\{ \begin{array}{l} i = 1, \ldots, s; \\ j = 1, \ldots, m_i \end{array} \right. \tag{17}$$

en

$$p_{i,j} = f_{i,j}(\theta_1, \ldots, \theta_\nu) \tag{18}$$

waarin de $f_{i,j}$ twee-maal continu-differentieerbare, gegeven functies zijn met ν onbekende parameters $\theta_1, \ldots, \theta_\nu$. Een stelsel waarden $\theta_1, \ldots, \theta_\nu$ kan aangegeven worden door een punt $(\theta_1, \ldots, \theta_\nu)$ in een ν-dimensionale ruimte R (de *parameterruimte*) en de verzameling van alle ,,toegelaten'' stelsels parameterwaarden is dan een deelverzameling Λ van deze ruimte. In het in de vorige paragrafen besproken geval zal Λ met H samenvallen, indien deze hypothese als juist wordt aanvaard, en anders b.v. met Ω.

4.2. *Schatting der parameters.*

Daar het in dit algemene verband eenvoudiger is, de bespreking van de schattingsmethoden aan die der toetsingsmethoden vooraf te laten gaan, zullen we eerst bespreken, hoe Neyman uit de gevonden frequentiequotiënten schattingen voor $\theta_1, \ldots, \theta_\nu$ (en dientengevolge ook voor de $p_{i,j}$) afleidt.

Neyman beschouwt de uitdrukking:

$$\underline{\chi}^2{}_1 = N \sum_{i=1}^{s} Q_i \sum_{j=1}^{m_i} \frac{(\underline{q}_{i,j} - p_{i,j})^2}{p_{i,j}} \tag{19}$$

waarin

$$p_{i,j} = f_{i,j}(\theta_1, \ldots, \theta_\nu), \quad N = \sum_{i=1}^{s} n_i \quad \text{en} \quad Q_i = n_i/N$$

is. Die waarden $\hat{\underline{\theta}}_1, \ldots, \hat{\underline{\theta}}_\nu$ met $(\hat{\underline{\theta}}_1, \ldots, \hat{\underline{\theta}}_\nu) \in \Lambda$, die $\underline{\chi}^2{}_1$ minimaal maken, zijn nu de schattingen van $\theta_1, \ldots, \theta_\nu$.

Geheel analoog zijn ook de waarden $\underline{\theta}_1^*, \ldots, \underline{\theta}_\nu^*$, met $(\underline{\theta}_1^*, \ldots, \underline{\theta}_\nu^*) \in \Lambda$, die de uitdrukking

$$\underline{\chi}^2{}_2 = N \sum_{i=1}^{s} Q_i \sum_{j=1}^{m_i} \frac{(\underline{q}_{i,j} - p_{i,j})^2}{q_{i,j}} \tag{20}$$

minimaal maken, als schattingen van $\theta_1, \ldots, \theta_\nu$ te beschouwen. Deze methode bezit boven de eerste het voordeel, dat $\underline{\theta}_1^*, \ldots, \underline{\theta}_\nu^*$ meestal gemakkelijker te berekenen zijn, dan $\hat{\underline{\theta}}_1, \ldots, \hat{\underline{\theta}}_\nu$. Zij kan echter slechts toegepast worden, indien alle bij het experiment gevonden waarden $q_{i,j} \neq 0$ zijn. Uit (17) volgt echter, dat dit, voor voldoende grote N (en bij constante Q_i) vrijwel zeker het geval zal zijn.

4.3. *Toetsingsmethoden.*

De meest algemene wijze, om een te toetsen hypothese aan te geven is, een deelverzameling H van R aan te wijzen. De te toetsen hypothese is dan: $(\theta_1, \ldots, \theta_\nu) \in H$, d.w.z. het punt $(\theta_1, \ldots, \theta_\nu)$ behoort tot de deelverzameling H. De verzameling van alternatieve hypothesen, waartegen H getoetst wordt, wordt op dezelfde wijze aangegeven door een deelverzameling Ω van R, die ook de gehele ruimte kan zijn *). In dit laatste geval kan men ook zeggen, dat geen alternatieve hypothese gespecificeerd is.

De toetsingsmethoden, die Neyman voor dit geval heeft ontwikkeld, berusten nu op de stelling, dat het verschil van de grootheden

$$\underline{\chi}^2_{1,H} = \operatorname*{Min}_{H} N \sum_{i=1}^{s} Q_i \sum_{j=1}^{m_i} \frac{(\underline{q}_{i,j} - p_{i,j})^2}{p_{i,j}} \tag{21}$$

en

$$\underline{\chi}^2_{1,\Omega} = \operatorname*{Min}_{\Omega} N \sum_{i=1}^{s} Q_i \sum_{j=1}^{m_i} \frac{(\underline{q}_{i,j} - p_{i,j})^2}{p_{i,j}} \tag{22}$$

waarin dus geminimaliseerd is met $(\theta_1, \ldots, \theta_\nu) \in H$ resp. $(\theta_1, \ldots, \theta_\nu) \in \Omega$, bij constante Q_i $(i = 1, \ldots, s)$ asymptotisch voor $N \to \infty$ verdeeld is volgens een χ^2-verdeling met als aantal vrijheidsgraden het verschil van het aantal onafhankelijk variabele parameters in Ω en in H. Grote waarden van $\underline{\chi}^2_{1,H} - \underline{\chi}^2_{1,\Omega}$ zijn kritiek, d.w.z. deze leiden tot verwerping van de hypothese H ten gunste van Ω.

Dezelfde stelling geldt voor de overeenkomstige grootheid $\underline{\chi}^2_{2,H} - \underline{\chi}^2_{2,\Omega}$, die uit $\underline{\chi}^2_{1,H} - \underline{\chi}^2_{1,\Omega}$ ontstaat, indien in (21) en (22) de $p_{i,j}$ in de noemers vervangen worden door $\underline{q}_{i,j}$. Hierbij dient weer het voorbehoud gemaakt te worden, dat de gevonden waarden $q_{i,j}$ alle $\neq 0$ moeten zijn; voor zover het de asymptotische verdeling van het verschil der twee grootheden $\underline{\chi}^2_{2,H}$ en $\underline{\chi}^2_{2,\Omega}$ betreft, houdt dit in, dat alle $p_{i,j} \neq 0$ moeten zijn. Immers dan geldt, voor $N \to \infty$ voor iedere i en j: $\lim_{N \to \infty} \mathrm{P}\,[\underline{q}_{i,j} = 0] = 0$.

§ 5. *Eigenschappen van de schattingen en toetsingsmethoden.*

5.1. De door Neyman afgeleide schattingen zijn, zoals door hem bewezen is, alle z.g. BAN-schattingen (BAN is afkorting van „best asymptotically normal"). Dit betekent, dat zij, voor $N = \sum_{i=1}^{s} n_i \to \infty$ bij constante $Q_i = n_i/N$ $(i = 1, \ldots, s)$ de volgende eigenschappen bezitten:

*) Eigenlijk zou men moeten zeggen: de gehele ruimte verminderd met H. Het is daarom misschien beter te spreken van de „toetsing van H *in* een verzameling Ω van *toegelaten* hypothesen". De uitdrukking in de tekst is echter de algemeen gebruikelijke.

a. $\hat{\theta}_k$ convergeert, voor iedere $k = 1, \ldots, \nu$ „in waarschijnlijkheid" tot de „ware" waarde θ°_k van θ_k. Dit betekent, dat er bij iedere $\varepsilon > 0$ en $\eta > 0$ een getal $N(\varepsilon, \eta)$ is te vinden met

$$P\,[\,|\,\hat{\theta}_k - \theta^\circ_k| > \varepsilon] < \eta \quad \text{voor} \quad N > N(\varepsilon, \eta)\,. \tag{23}$$

b. De waarschijnlijkheidsverdeling van $\hat{\theta}_k$ is, voor iedere k, asymptotisch normaal, d.w.z.

$$\lim_{N \to \infty} P \left[\frac{\hat{\theta}_k - \theta^\circ_k}{\sigma_k}\sqrt{N} < t\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{1}{2}x^2}\, dx \tag{24}$$

waarin de σ_k's niet van N afhangen.

c. Als $\theta'_k\,(k = 1, \ldots, \nu)$ een *willekeurige* andere schatting van θ°_k is, die aan a en b voldoet met σ in plaats van σ_k, geldt

$$\sigma \geqq \sigma_k. \tag{25}$$

5.2. Neyman bewijst verder, dat de besproken toetsingsmethoden voor $N \to \infty$ bij constante $Q_i\,(i = 1, \ldots, s)$ asymptotisch aequivalent zijn met de λ-toets voor de hypothese H tegen Ω. Deze λ-toets zullen wij hier niet bespreken. Voor het onderhavige geval vindt men deze beschreven in het onder [3] geciteerde artikel van J. Neyman and E. S. Pearson.

5.3. Een schattingsmethode, die eveneens BAN-schattingen levert, is de methode der aannemelijkste schattingen („maximum likelihood method"). Het voordeel van deze nieuwe methoden van Neyman is daarin gelegen, dat de bij deze methoden optredende minimum-problemen mathematisch van minder ingewikkelde aard zijn, zodat in gevallen, waarin de methode der aannemelijkste schattingen tot nog onopgeloste wiskundige vraagstukken aanleiding geeft, deze nieuwe methoden soms wel tot een oplossing leiden. Hetzelfde geldt voor de toetsingsmethoden in vergelijking met de λ-methode.

Afgezien van deze voordelen der nieuwe methoden is de door Neyman bereikte unificatie van de genoemde theorieën van groot belang.

LITTERATUUR:

[1] J. Neyman, *Contribution to the theory of the χ^2-test*, Proceedings of the Berkeley symposium on mathematical statistics and probability, University of California Press, Los Angeles 1949 pp 239-273.

[2] H. Cramér, *Mathematical methods of statistics*, Princeton University Press, Princeton 1946.

[3] J. Neyman and E. S. Pearson, *On the use of certain test criteria for purposes of statistical inference*, Biometrika **20A**, 175-240, 263-294, 1928.