

Fiduciële Statistiek

H.N. Linssen*

1. Inleiding

Het nu volgende verhaal is de tekst van een voordracht over fiduciële statistiek voor de sectie Mathematische Statistiek van de Vereniging voor Statistiek. De bedoeling van de voordracht was niet om in detail te treden over de wiskundige aspecten van de fiduciële statistiek, noch om nieuwe toepassingen ervan aan te geven. Wel was de bedoeling om aannemelijk te maken, dat de fiduciële statistiek waard is om serieus te worden genomen, door het geleerde gehoor de centrale concepten ervan voor ogen te voeren.

Ik heb de vrijheid genomen hier en daar van de gesproken tekst af te wijken.

Tengevolge van de onvruchtbare polemiek tussen Fisher en Neyman en anderen, die in de dertiger jaren en later plaatsvond, hebben vele statistici zich van de fiduciële statistiek afgekeerd. De polemiek was mede daarom zo onvruchtbaar omdat Fisher niet precies wist, wat fiduciële statistiek is (daarover later meer) en zo de kritiek van zijn opposenten niet bevredigend kon pareren. Dit en het grote aantal paradoxen, die uit de fiduciële theorie schijnen te volgen, heeft ertoe geleid dat vele statistici de fiduciële statistiek niet serieus nemen (zie bijvoorbeeld van Dantzig³).

2. Empirie

Uitgangspunt is dat statistiek een empirische theorie is van de statistische besluitvorming. De centrale probleemstelling is: Wat zeggen de waarnemingen over de parameter(s) van belang? Een methode voor een specifiek geval wordt gezien als een hypothese over hoe de statistische besluitvorming in dat geval moet plaatsvinden, en kan verworpen worden als zij leidt tot duidelijk absurde resultaten in speciale (of 'cruciale') gevallen. In de fysica geldt analoog dat een hypothese alleen kan worden gefalsificeerd (en niet geverifieerd) en wel naar aanleiding van de resultaten van cruciale experimenten.

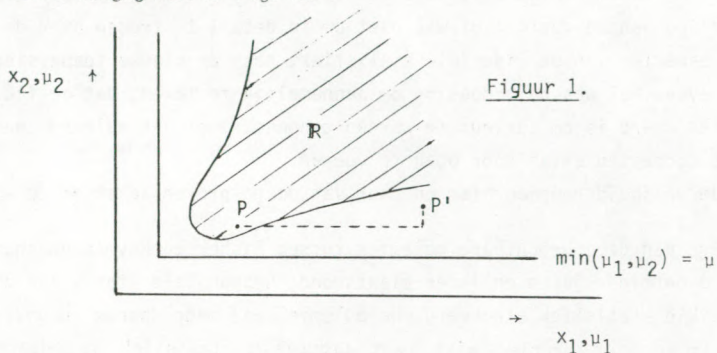
* Onderafdeling der Wiskunde en Informatica, Technische Hogeschool Eindhoven, Postbus 513, 5600 MB EINDHOVEN.

Een mooi voorbeeld van een cruciaal geval wordt gegeven door Chambers¹. Zij $\underline{x}_i \sim N(\mu_i, 1)$ ($i = 1, 2$) en zij tevens $\min(\mu_1, \mu_2) \leq \mu$ de interessante hypothesen. In de Neyman-Pearson-Wald ('frequentistische') aanpak wordt deze hypothesen getoetst door middel van een kritiek gebied $\mathbb{R}$, zodanig dat

$$P((\underline{x}_1, \underline{x}_2) \in \mathbb{R} \mid \mu_1, \mu_2) \leq \alpha$$

voor alle (μ_1, μ_2) met $\min(\mu_1, \mu_2) \leq \mu$.

Bij alle toelaatbare (admissible) toetsen hoort dan een kritiek gebied van de vorm zoals aangegeven in Fig. 1.



Het absurde karakter van dit gebied wordt duidelijk als we naar de punten P en P' kijken. $P \in \mathbb{R}$, dus wordt de hypothese verworpen, maar $P' \notin \mathbb{R}$ zodat de hypothese niet verworpen kan worden, hoewel in P' beide waarnemingen groter zijn dan in P.

Een ander mooi voorbeeld van statistische empirie vinden we bij Cox². Hij wilde daarmee laten zien dat je moet conditioneren op ancillaire grootheden.

De waarneming bestaat uit twee componenten, $\underline{x}$ en $\underline{t}$. De grootte $\underline{t}$ kan twee waarden aannemen, zeg 1 en 2, ieder met kans $\frac{1}{2}$. Er geldt:

$$\text{als } t = 1 \quad \text{dan} \quad \underline{x} \sim N(\mu, 1) ;$$

$$\text{als } t = 2 \quad \text{dan} \quad \underline{x} \sim N(\mu, \sigma^2),$$

waarbij $\sigma^2 \gg 1$ en bekend.

(Het is alsof er twee meetinstrumenten zijn, waarvan één zeer onnauwkeurig.

Aan de meting gaat dan een aselechte keuze van het meetinstrument vooraf.)

De nulhypothese is $H_0 : \mu = 0$ en de alternatieve is $H_1 : \mu = 1$. Het kritieke gebied behorende bij de frequentistische meest onderscheidende toets met een

betrouwbaarheid van 95% wordt gegeven door:

$$\text{als } t = 1 \quad \text{dan} \quad x > 1.28;$$

$$\text{als } t = 2 \quad \text{dan} \quad x > 5\sigma.$$

Inderdaad geldt $P(\text{verwerpen} \mid H_0) = \frac{1}{2} \cdot 10\% + \frac{1}{2} \cdot 0\% = 5\%$. Tevens

$P(\text{verwerpen} \mid H_1) = \frac{1}{2} \cdot 38\% + \frac{1}{2} \cdot 0\% = 19\%$. Het komt er op neer, dat de nul-hypothese vrijwel nooit verworpen wordt als $t = 2$.

De toets zou zin kunnen hebben in een kwaliteitscontrole-context, daar ze dan wellicht zeer vaak herhaald zal worden en 'vreemde' beslissingen in een deel van de gevallen toegestaan zijn, omdat daarmee het overall-onderscheidingsvermogen opgevoerd wordt. Voor één steekproef echter, leidt ze tot absurde resultaten. Bij de voorwaardelijke toets wordt er geconditioneerd op de ancillaire grootheid t . Het kritieke gebied wordt dan:

$$\text{als } t = 1 \quad \text{dan} \quad x > 1.645;$$

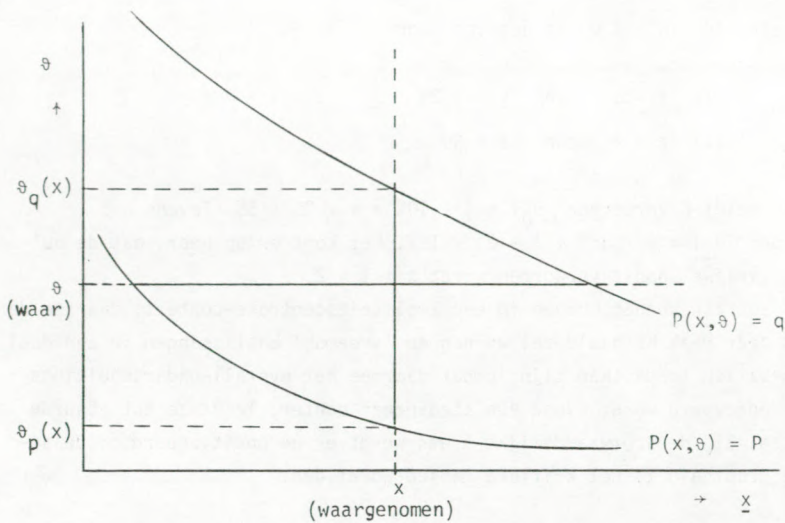
$$\text{als } t = 2 \quad \text{dan} \quad x > 1.645\sigma.$$

Er geldt: $P(\text{verwerpen} \mid H_1) = 15\%$.

Inderdaad is het onderscheidingsvermogen m.b.t. H_1 van de voorwaardelijke toets kleiner dan die van de onvoorwaardelijke. Dat is geen wonder, want de onvoorwaardelijke aanpak mikt nou net op het maximaliseren van onderscheidingsvermogen. Dit voorbeeld laat zien dat deze aanpak tot absurde resultaten leiden kan.

3. Fiduciële inversie

Bij de fiduciële inversie worden frequentieuitspraken over waarnemingen gegeven de parameters getransformeerd tot fiduciële uitspraken over de parameters gegeven de waarnemingen. Hoe de inversie tot stand komt, wordt getillustreerd aan de hand van Figuur 2 voor het 1 waarneming-1 parameter-geval. Laat x een continue stochastische variabele zijn met verdelingsfunctie $P(x; \vartheta)$. Voor de eenvoud nemen we aan, dat ϑ ook continu is. Een noodzakelijke voorwaarde voor de fiduciële inversie is, dat de vergelijking $P(x; \vartheta) = p$ voor alle $p \in (0,1)$ een monotoon dalend verband tussen x en ϑ definiëert. Bij een waargenomen x wordt het p -quantiel ϑ_p van de 'fiducieverdeling' van ϑ impliciet gegeven door $P(x, \vartheta_p) = p$, geheel analoog aan de constructie van betrouwbaarheidsintervallen voor een parameter. (Indien $P(x, \vartheta) = p$ een monotoon stijgend i.p.v. dalend verband tussen x en ϑ definiëert, dan geldt $P(x, \vartheta_{1-p}) = p$.)



Figuur 2.

De horizontaal getrokken lijn in Figuur 2 is een frequentie-interval voor x gegeven θ , terwijl de verticaal getrokken lijn een fiduciële interval voor θ gegeven x is, met precies de juiste betrouwbaarheidsfrequentie. We drukken deze eigenschap uit door te zeggen, dat de fiducieverdeling een 'directe frequentie-interpretatie' toelaat.

4. Onsamenhangendheid

Nu is het niet zo, dat de fiducieverdeling van een functie van θ met behulp van de wetten van de kansrekening kan worden afgeleid uit de fiducieverdeling van θ . Dit is alleen mogelijk voor eeneenduidige functies. Dit op het eerste gezicht vreemde feit volgt rechtstreeks uit het 'relevantie-beginsel', dat zegt dat alles wat relevant is moet worden meegenomen, en alles wat niet relevant is genegeerd moet worden bij de statistische besluitvorming. Een voorbeeld ter illustratie: Als $\underline{x}$ normaal verdeeld is met verwachting μ en variantie 1, dan is het resultaat van de fiduciële inversie dat de fiducieverdeling van θ normaal is met verwachting x en variantie 1. Notatie:

$$\underline{x} \sim N(\theta, 1) \xrightarrow{\text{f.i.}} \bar{\theta} \sim N(x, 1), \quad (1)$$

waarbij het streepje boven ϑ aangeeft dat we niet met een kans- maar met een fiducievariabele te maken hebben. Als echter niet ϑ maar ϑ^2 de parameter van belang is, dan is het teken van ϑ irrelevant en het teken van x een irrelevant deel van de waargenomen x . Het is duidelijk, dat de statistische besluitvorming m.b.t. ϑ^2 gebaseerd moet worden op de frequentieverdeling van $|x|$ of x^2 . Dit resulteert in:

$$\underline{x}^2 \sim \chi_1^2(\vartheta^2) \xrightarrow{\text{f.i.}} F(\vartheta^2) = P(\chi_1^2(\vartheta^2) > x^2) . \quad (2)$$

Rechts van de inversie-pijl staat dat de p-quantielen van de fiducieverdeling van ϑ^2 gegeven worden door die waarde van ϑ^2 waarvoor x^2 precies het $(1-p)$ -quantiel is van de kansverdeling van x^2 . Als fiducieverdelingen aan de wetten van de kansrekening zouden voldoen, dan zou uit (1) volgen:

$$\overline{\vartheta^2} \sim \chi_1^2(x^2) ,$$

hetgeen strijdig is met (2).

Dit fenomeen drukken we uit door te zeggen, dat fiducieverdelingen niet 'samenhangen'. Alleen voor onderling eeneenduidige parametertransformaties hangen de fiducieverdelingen samen.

Wilkinson⁷ heeft het niet samenhangen van fiducieverdelingen verheven tot het 'principe van onsamenhangendheid' (incoherence). Op het eerste gezicht misschien een vreemd principe, maar bij nader inzien toch voor de hand liggend. Immers, als een andere parameterfunctie van belang is, zullen andere componenten van de waarnemingen relevant zijn. In de beslissingstheorie treffen we een geheel analoge situatie aan. Niemand hoort het vreemd te vinden, dat een schatter toelaatbaar is met betrekking tot de ene verliesfunctie en ontoelaatbaar met betrekking tot de andere. Net zo min is het vreemd, dat een waarnemingscomponent relevant is met betrekking tot de ene parameterfunctie en irrelevant met betrekking tot de andere en dat dus de fiducieverdelingen niet hoeven samen te hangen. Het is aannemelijk, dat Fisher zich niet voldoende heeft gerealiseerd, dat onsamenhangendheid een essentiële kenmerk is van fiducieverdelingen.

Een interessant punt van (2) is, dat er condensatie van fiducie plaatsvindt in $\vartheta^2 = 0$. Deze zgn. 'niet-toegewezen fiducie' is gelijk aan:

$$F(0) = P(\chi_1^2(0) > x^2) .$$

De interpretatie van deze fiducie is afhankelijk van de context. De fiduciever-

deling heet dan 'incomplete'.

Dat fiducieverdelingen niet kunnen samenhangen wordt op dramatische wijze duidelijk bij de 'paradox' van Stein⁶. Hij redeneerde als volgt:

Stel dat voor de onderling onafhankelijke waarnemingen $x_1, \dots, x_n$ geldt:

$$x_i \sim N(\mu_i, 1)$$

en dat $\sum \mu_i^2$ de parameter van belang is. Er geldt:

$$\underline{x} \sim N_n(\underline{\mu}, 1) \xrightarrow{\text{f.i.}} \bar{\mu} \sim N_n(x, 1) .$$

Stein stelde:

$$\bar{\mu} \sim N(x, 1) \longrightarrow \bar{\mu}^T \bar{\mu} \sim x_n^2(x^T x) , \quad (3)$$

en merkte op, dat de actuele betrouwbaarheid van fiduciële intervallen, die op (3) gebaseerd zijn, naar 0 gaat voor n naar ∞ . Dat is de inhoud van de 'paradox'.

In (3) is onuitgesproken aangenomen, dat de fiducieverdelingen van μ en $\mu^T \mu$ samenhangen. Echter, relevant voor $\mu^T \mu$ is $x^T x$ en niets anders, zodat:

$$\underline{x}^T \underline{x} \sim x_n^2(\mu^T \mu) \xrightarrow{\text{f.i.}} F(\mu^T \mu) = P(x_n^2(\mu^T \mu) > x^T x) . \quad (4)$$

Fiduciële intervallen, gebaseerd op (4), hebben de correcte betrouwbaarheids-frequentie. Als de niet-toegewezen fiducie

$$F(0) = P(x_n^2(0) > x^T x)$$

te groot is, bestaat er reden het waarnemingsmodel in zijn geheel te verwerpen⁷.

5. Directe en verwachte frequentie-interpretatie

Fiduciële statistiek is een empirische theorie, waarbinnen naar de oplossing van een gegeven statistisch probleem wordt gezocht. Die oplossing wordt dan bepaald zonder gebruik te maken van criteria, zoals maximaal onderscheidingsvermogen of minimaal verwacht verlies, en zonder gebruik te maken van irrelevante priorverdelingen. De fiduciële statistiek gaat over de logica (eenduidigheid, gedetermineerdheid) van de statistische besluitvorming.

Voor n waarnemingen, 1 parameter ziet die logica er als volgt uit:

Zij $\underline{x} \in \mathbb{R}^n$ en zij tevens $\underline{x}$ equivalent met het paar $(\underline{y}, \underline{z})$, met $\underline{y} \in \mathbb{R}^1$ en $\underline{z} \in \mathbb{R}^{n-1}$, waarvoor geldt:

$$dP(\underline{y}, \underline{z}; \vartheta) = dP(\underline{y} | \underline{z}; \vartheta) dP(\underline{z}) , \quad (5)$$

waarbij dP staat voor dichtheid. De waarnemingscomponent z wordt dan 'ancillair' genoemd en y 'voorwaardelijk voldoende'. Als nu $P(y|z;\theta)$ een monotoon verband tussen $y|z$ en θ definiëert, is fiduciële inversie mogelijk en ligt dus de fiducieverdeling van θ vast.

Voor n waarnemingen en 2 parameters geldt analoog:

$\underline{x} \in \mathbb{R}^n$, $\underline{x}$ equivalent met (y, z) met $y \in \mathbb{R}^2$, $z \in \mathbb{R}^{n-2}$, waarvoor (3) geldt. Om de fiduciële inversie mogelijk te maken, moet $dP(y|z;\theta)$ verder kunnen worden gefactoriseerd:

$$dP(y;\theta) = dP(y_1;\theta_1) dP(y_2|y_1;\theta_1,\theta_2),$$

waarbij alle dichtheden voorwaardelijk z zijn.

Het is nodig, dat $P(y_1;\theta_1)$ een monotoon verband tussen y_1 en θ_1 definiëert, zodat de fiducieverdeling $F(\theta_1)$ kan worden bepaald, en dat $P(y_2|y_1;\theta_1,\theta_2)$ een monotoon verband tussen $y_2|y_1$ en θ_2 definiëert bij elke θ_1 , zodat de voorwaardelijke fiducieverdeling $F(\theta_2|\theta_1)$ kan worden bepaald.

$F(\theta_1)$ heeft een directe frequentie-interpretatie, $F(\theta_2|\theta_1)$ ook, maar dan gegeven θ_1 . De voorwaardelijke fiducieverdeling is niet relevant, eenvoudig omdat θ_1 niet bekend is. Het ligt voor de hand het resultaat van de besluitvorming te representeren met behulp van de 'verwachte fiducieverdeling':

$$\bar{F}(\theta_2) = \int F(\theta_2|\theta_1) dF(\theta_1),$$

die dan geen directe frequentie-interpretatie meer heeft, maar een verwachte (en een directe in een hypothetische referentieverzameling^{7,5}). Het is mogelijk dat $F(\theta_1)$ niet relevant is voor de fiducieverdeling van θ_2 . Dit is bijvoorbeeld het geval als $F(\theta_2|\theta_1)$ een functie is van alleen θ_1^2 . Dan is $F(\theta_1)$ niet relevant, omdat $F(\theta_1)$ en $F(\theta_1^2)$ niet samenhangen. Het is nodig, dat $F(\theta_2|\theta_1)$ en θ_1 samenhangen (de zgn. 'kruissamenhang'⁷).

Een en ander kan worden geïllustreerd aan de hand van het Behrens-Fisher probleem met $n = n_1 = n_2$. De Behrens-Fisher statistiek

$$d = \frac{(\mu_1 - \mu_2) - (\bar{x}_1 - \bar{x}_2)}{\sqrt{(s_1^2 + s_2^2)/n}}$$

is, als we aannemen dat $\sigma_1 = \sigma_2$, t_{2n-2} -verdeeld, zodat fiduciële uitspraken over $\mu_1 - \mu_2$ mogelijk zijn met een directe frequentie-interpretatie. Als we echter niet weten dat $\sigma_1 = \sigma_2$, hoewel dat feitelijk het geval is, dan moet de fiducieverdeling van $\mu_1 - \mu_2$ diffuser zijn dan in het geval dat we dat wel

weten. Deze fiducieverdeling heeft dan geen directe frequentie-interpretatie, omdat immers de betrouwbaarheidsfrequentie van een fiduciële interval afhangt van σ_1^2 en σ_2^2 .

De oplossing die Welch geeft voor het Behrens-Fisher probleem is in strijd met het principe, dat gebrek aan relevante informatie de statistische besluitvorming diffuser moet maken; immers de Welch test geeft een kleiner betrouwbaarheidsinterval dan de t-toets (waarbij aangenomen $\sigma_1 = \sigma_2$), indien tenminste $s_1^2/s_2^2 \approx 1$.

Fisher's oplossing voldoet niet aan de eis van kruissamenhang. Fisher realiseerde zich niet, dat het niet uitmaakt dat $\sigma_1^2/\sigma_2^2 = k$ of $1/k$ en dat het teken van $\log s_1^2/s_2^2$ een irrelevante component van de waarnemingen is*.

6. De lineaire functionele relatie

Het is mogelijk dat U niet overtuigd bent, dat bij simultane besluitvorming de directe frequentie-interpretatie verloren kan gaan. U zou kunnen zeggen: "Het kan mij niet schelen wat de fiduciële statistiek zegt. Ik voor mij sta erop, dat betrouwbaarheidsintervallen een directe frequentie-interpretatie hebben".

De Neyman-Pearson-Wald (frequentistische) aanpak van de statistische besluitvorming eist dit inderdaad van betrouwbaarheidsintervallen. Waar dat toe kan leiden, althans bij het toetsen van hypothesen, hebt U kunnen zien aan het voorbeeld waar ik mee begonnen ben. Ik eindig met een voorbeeld over betrouwbaarheidsintervallen op het gebied van de functionele relaties. Het model is:

$$\eta = \alpha + \beta \xi \quad \text{met } \eta, \xi \in \mathbb{R}^n \quad \text{en} \quad \alpha, \beta \in \mathbb{R}^1.$$

Waarnemingen worden gegeven door

$$\underline{x} = \xi + \underline{e} \quad ; \quad \underline{y} = \eta + \underline{e}',$$

waarbij $\underline{e}$ en $\underline{e}'$ onderling onafhankelijk standaard normaal verdeeld zijn (zie Figuur 3).

* Wilkinson stuurde mij onlangs de sheet-teksten van een lezing over fiduciële statistiek, die hij heeft gehouden op de Second International Meeting on Bayesian Statistics te Valencia, met daarin een nieuwe oplossing van het B-F probleem naast andere interessante voorbeelden van toepassingen van de theorie. Een kopie van de teksten stuur ik op verzoek toe.



Figuur 3.

Als we definiëren

$$\gamma^2 = \sum_{i=1}^n (\xi_i - \bar{\xi})^2 / n ,$$

dan wordt de parameterruimte gegeven door:

$$\{(\alpha, \beta, \gamma^2) \mid -\infty < \alpha, \beta < \infty, \gamma^2 > 0\} .$$

Gleser⁴ bewijst de volgende stelling:

"Ieder betrouwbaarheidsinterval $[L, W]$ voor β met betrouwbaarheid (Gleser: 'coverage probability') tenminste $1 - \alpha$ voor alle parameterwaarden voldoet aan

$$P\{(L, W) = (-\infty, \infty)\} > 0 ,$$

voor iedere parameterwaarde (en heeft dus oneindige verwachte lengte)...".

Je kunt erover van mening verschillen of dit nu een slechte eigenschap is van zo'n betrouwbaarheidsinterval of niet. Het fenomeen ontstaat doordat γ^2 willekeurig klein kan zijn en dat betekent dat je dan in wezen maar één punt van de lijn waarneemt, waardoor de helling onbepaald wordt.

Als echte frequentist probeert Gleser onder deze door hem als ongewenst ervaren eigenschap uit te komen door de parameterruimte te beperken en te eisen dat $\gamma^2 \geq \tau > 0$. Hij leidt dan betrouwbaarheidsintervallen met eindige lengte af, die functies zijn van τ en nog steeds voldoen aan de frequentistische eis dat de betrouwbaarheid minstens $1 - \alpha$ is voor alle waarden in de parameterruimte. Het probleem is, dat van de experimentator een waarde voor τ gevraagd wordt, iets wat hij waarschijnlijk niet dacht te hoeven weten en wat een subjectief element in de besluitvorming kan introduceren.

De fiducialist daarentegen zal opmerken, dat de waarnemingen aanleiding geven om meer fiducie te hebben in de ene waarde van γ^2 dan in de andere. Met name zal hij weinig fiducie hebben in de mogelijkheid dat $\gamma^2 = 0$, omdat het immers onaannemelijk is, dat de in Figuur 3 schetsmatig weergegeven steekproef afkomstig is uit een (2-dimensionale) populatie die, afgezien van het gemiddelde, standaard-normaal is. Hij zal proberen een fiducieverdeling voor γ^2 te bepalen en met behulp daarvan een verwachte fiducieverdeling voor β . Hoe dat precies moet in dit geval, weet ik niet.

7. Slotwoord

Een empirische theorie doet niet altijd wat je verwacht. Fisher schijnt eens tegen L.J. Savage gezegd te hebben⁸: "I do not understand yet what fiducial probability does. We shall have to live with it for a long time before we know what it is doing for us. But it should not be ignored just because we do not yet have a clear interpretation".

Dames en heren, ik hoop dat ik U duidelijk heb gemaakt, dat de fiduciële statistiek niet pure onzin is en dat zij waard is serieus genomen te worden. Het is nodig, dat de vanzelfsprekendheid, waarmee de frequentistische of beslissingstheoretische aanpak in de leerboeken als evangelie wordt aangeboden, wordt ondergraven en dat er meer aandacht komt in het onderzoek voor statistiek zonder verliesfunctie, zelfs zonder alternatieve hypothese en zonder (subjectieve) priorverdelingen.

Het resultaat van de fiduciële statistiek is een fiducieverdeling voor de parameter(s) van belang (en dat is Bayesiaans in de geest), die echter niet hoeft samen te hangen met de fiducieverdeling voor een niet-eenduidige functie van de parameter. Met de fiducieverdeling is een frequentie-interpretatie verbonden (en dat is frequentistisch in de geest), die echter niet direct hoeft te zijn.

Referenties

1. Chambers, H.L., A simple problem with strikingly different frequentist and bayesian solutions. J.R. Statist. Soc. B, 32, 278-82, 1970.
2. Cox, D.R., Some problems connected with statistical inference, Ann. Math. Statist., 29, 357-72, 1958.
3. Dantzig, D. van, Statistical priesthood (Sir Ronald on scientific inference). Statistica Neerlandica 11, 185-200, 1957.

4. Gleser, L.J., Confidence regions for the slope in a linear errors-in-variables regression model. Purdue University, Dept. of Statistics, Technical Report #82-23, 1982.
5. Pedersen, J.G., Fiducial inference. International Statistical Review, 46, 147-70, 1978.
6. Stein, C.H., An example of wide discrepancy between confidence intervals and fiducial intervals. Ann. Math. Statist., 30, 877-81, 1959.
7. Wilkinson, G.N., On resolving the controversy in statistical inference. J.R. Statist. Soc. B, 38, 119-71, 1977.
8. Idem, pagina 145.

Ter lezing aanbevolen

fiduciële theorie en veel referenties in:

- * Wilkinson, G.N., referentie 7.

bibliografie in:

- * Turkey, J.W., Some examples with fiducial relevance. Ann. Math. Statist., 28, 687-95, 1957.

en in:

- * Brillinger, D.R., Examples on the definition of fiducial probability, Ann. Math. Statist., 33, 1349-55, 1962.

additionele referenties, fiduciële referentieverzamelingen en relevante deelverzamelingen in:

- * Pedersen, J.G., referentie 5.

Statistische besluitvorming min of meer volgens fiduciëel recept in:

- * Barndorff-Nielsen, O., Information and exponential families in statistical theory. Wiley 1978.

Een fiduciële kijk op robuustheid in:

- * Wilkinson, G.N., Robust inference - the Fisherian approach. In: Robustness in statistics. Academic Press 1979.

Een vergelijkende studie van de 'scholen' in de statistiek in:

- * Barnett, V., Comparative statistical inference. Wiley 1973.