EUR 2224.d

EUROPÄISCHE ATOMGEMEINSCHAFT - EURATOM

PRAKTISCHE STATISTIK UND SUBJEKTIVISMUS

von

P. IHM

1965

Gemeinsame Kernforschungsstelle
Forschungsanstalt Ispra - Italien

Zentralstelle für die Verarbeitung wissenschaftlicher Information - CETIS

HINWEIS

Dieser Bericht wird zum Preise von 40,— bfrs. verkauft. Bestellungen sind zu richten an : PRESSES ACADEMIQUES EUROPEENNES — 98, chaussée de Charleroi, Brüssel 6.

Die Zahlung ist zu leisten durch Überweisung an die :

— BANQUE DE LA SOCIETE GENERALE (Agence Ma Campagne) - Brüssel - Konto Nr. 964.558 ;

— BELGIAN AMERICAN BANK AND TRUST COMPANY - New York - Konto Nr. 22.186 ;

— LLOYDS BANK (Europe) Ltd. - 10, Moorgate, London E.C.2 ;

als Bezug ist anzugeben : « EUR 2224.d - PRAKTISCHE STATISTIK UND SUBJEKTIVISMUS ».

Das vorliegende Dokument wurde an Hand des besten Abdruckes vervielfältigt, der zur Verfügung stand.

Gedruckt von L. VANMELLE, S.A., Gent.
Brüssel, Januar 1965.

EUR 2224.d

PRAKTISCHE STATISTIK UND SUBJEKTIVISMUS von P. IHM

Europäische Atomgemeinschaft — EURATOM
Gemeinsame Kernforschungsstelle
Forschungsanstalt Ispra (Italien)
Zentralstelle für die Verarbeitung wissenschaftlicher Information (CETIS)
Brüssel, Januar 1965 — 17 Seiten

Es wird in der Arbeit gezeigt, wie Konfidenzbereiche, die an sich auf einer objektivistischen Wahrscheinlichkeitstheorie begründet sind, subjektivistisch interpretiert werden können. Es wird an einem Fall gezeigt, dass Konfidenzbereiche Quantile von Likelihoodfunktionen und, unter gewissen Voraussetzungen, von subjektiven Wahrscheinlichkeitsverteilungen sein können; auf Verallgemeinerungen wird hingewiesen.

Der P-Wert wird ebenfalls subjektivistisch interpretiert.

EUR 2224.d

PRACTICAL STATISTIC AND SUBJECTIVISM by P. IHM

European Atomic Energy Community — EURATOM
Joint Nuclear Research Center
Ispra Establishment (Italy)
Scientific Data Processing Center (CETIS)
Brussels, January 1965 — 17 pages

It is shown how confidence regions which generally are based on objectivistic probability can be interpreted in a subjectivistic sense.

It is shown in a particular case that confidence regions are quantiles of likelihood functions or, under certain assumptions, of subjective probability distributions ; possible generalizations are recorded.

The P-value is also interpreted in the subjectivistic way.

EUR 2224.d

PRACTICAL STATISTIC AND SUBJECTIVISM by P. IHM

European Atomic Energy Community — EURATOM
Joint Nuclear Research Center
Ispra Establishment (Italy)
Scientific Data Processing Center (CETIS)
Brussels, January 1965 — 17 pages

It is shown how confidence regions which generally are based on objectivistic probability can be interpreted in a subjectivistic sense.

It is shown in a particular case that confidence regions are quantiles of likelihood functions or, under certain assumptions, of subjective probability distributions ; possible generalizations are recorded.

The P-value is also interpreted in the subjectivistic way.

EUR 2224.d

PRACTICAL STATISTIC AND SUBJECTIVISM by P. IHM

European Atomic Energy Community — EURATOM
Joint Nuclear Research Center
Ispra Establishment (Italy)
Scientific Data Processing Center (CETIS)
Brussels, January 1965 — 17 pages

It is shown how confidence regions which generally are based on objectivistic probability can be interpreted in a subjectivistic sense.

It is shown in a particular case that confidence regions are quantiles of likelihood functions or, under certain assumptions, of subjective probability distributions ; possible generalizations are recorded.

The P-value is also interpreted in the subjectivistic way.

EUR 2224.d

EUROPÄISCHE ATOMGEMEINSCHAFT - EURATOM

PRAKTISCHE STATISTIK UND SUBJEKTIVISMUS

von

P. IHM

1965

Gemeinsame Kernforschungsstelle
Forschungsanstalt Ispra - Italien

Zentralstelle für die Verarbeitung wissenschaftlicher Information - CETIS

Manuskript erhalten am 24.11.1964.

INHALTSVERZEICHNIS

1. Einleitung 5

2. Die subjektivistische und dualistische Theorie 6

3. Der Konfidenzschluss 8

4. Die subjektivistische Interpretation 12

5. Der P-Wert 15

Bibliographie 16

PRAKTISCHE STATISTIK UND SUBJEKTIVISMUS

1. Einleitung

In letzter Zeit sind eine Reihe von Veröffentlichungen erschienen, in denen eine Theorie der Statistik auf der Basis einer nichtobjektivistischen und vornehmlich subjektivistischen Wahrscheinlichkeitstheorie versucht wird (RAIFFA u. SCHLAIFER 1961, RICHTER 1954 a,b, SAVAGE 1954, SCHLAIFER 1959, 1961). Sie werden gestützt durch Arbeiten von CARNAP (1950), sowie CARNAP und STEGMÜLLER (1959) über induktive Logik, in denen eine solche aufgrund einer objektivistischen Wahrscheinlichkeitstheorie für unmöglich gehalten wird. Der praktische Statistiker hat sich angesichts dieser Fakten die Frage vorzulegen, inwiefern seine Arbeit dadurch beeinflusst wird. Wie in dieser Arbeit gezeigt werden soll, lautet die Antwort dahingehend, dass keine Beeinträchtigung der bisherigen Methoden vorliegt, dass sogar gewisse Verhaltensweisen, die im Widerspruch zu den Forderungen der objektivistischen NEYMAN-PEARSONschen Theorie stehen, zwanglos erklärt werden können.

Es wird zunächst ein kurzer Abriss der Grundlagen der subjektivistischen und dualistischen Theorie gegeben, dann werden P-Wert und Konfidenzschluss im Sinne dieser Theorie gedeutet.

Einige verwendete Bezeichungsweisen seien erklärt. Vektoren $\mathfrak{n}$ sind Spaltenvektoren, $\mathfrak{n}'$ bezeichnet den Zeilenvektor mit den gleichen Komponenten. Mengen S mit Elementen s werden durch $S=\{s:Ps\}$ erklärt, wobei Ps eine Eigenschaft von s bezeichnet, d.h. S ist die Menge der Elemente s, die die Eigenschaft P haben. $s\in S$ bedeutet, dass s in S enthalten ist. $\bar{S}$ ist das Komplement von S, d.h. die Menge der Elemente, die nicht in S enthalten sind. $S=\emptyset$ ist die sog. leere Menge, d.h. die Menge, die überhaupt kein Element enthält. Für zwei Mengen S und T bezeichnet $S\cup T$ die Vereinigung, d.h. die Menge der Elemente, die in S oder T oder beiden enthalten sind. Für eine Klasse von Mengen S_τ, abhängig von einem Index τ, ist $\cup S_\tau$ die Vereinigung aller S_τ. Der Durchschnitt von S und T, $S\cap T$ (oder auch ST) ist definiert als die Menge der Elemente, die sowohl in S als auch in T liegen. Ist $S\cap T=\emptyset$, so heissen S und T zueinander fremd oder disjunkt. $\cap S_\tau$ wird analog zu $\cup S_\tau$ definiert.

Ist S' eine Teilmenge von S, so schreibt man $S' \subset S$ ("S' enthalten in S"), im Grenzfalle $S'=s$, d.h., wenn S' nur ein Element enthält, $s \in S$. Eine gute Einführung in die Mengenlehre, sowie in die Vektoren und Matrizenalgebra findet sich bei CRAMER (1946).

2. Die subjektivistische und dualistische Theorie

POISSON schreibt 1836 in seinen "Recherches sur la probabilité des jugements en matière civile et criminelle":

"La probabilité d'un événement est la raison que nous avons de croire qu'il aura ou qu'il a eu lieu...
La probabilité dépendant des connaissances que nous avons sur un événement, elle peut être inégale pour un même événement et pour diverses personnes...
Dans le langage ordinaire les mòts chance et probabilité sont à peu près synonymes. Le plus souvent nous emploierons indifféremment l'un et l'autre; mais lorsqu'il sera nécessaire de mettre une différence entre leurs acceptions, on rapportera, dans cet ouvrage, le mot chance aux événements en eux-mêmes et indépendamment de la connaissance que nous en avons, et l'on conservera au mot probabilité sa définition précédente. Ainsi, un événement aura, par sa nature, une chance plus ou moins grande, connue ou inconnue; et sa probabilité sera relative à nos connaissances, en ce que le concerne".

Hier tritt eine dualistische Auffassung zu Tage, die zwei Arten von Wahrscheinlichkeiten unterscheidet, nämlich die probabilité, die von Beobachter zu Beobachter variieren kann, also subjektiv ist, und die chance, die eine in der Natur feststehende physikalische Grösse ist. Diese beiden Wahrscheinlichkeiten seien hier mit 1-Wahrscheinlichkeit bzw. 2-Wahrscheinlichkeit bezeichnet. Eine mathematischen Ansprüchen genügende dualistische Theorie ist die RICHTERs (1954 a,b), auf die hier zunächst eingegangen werden soll.

Es sei eine 2-Wahrscheinlichkeitsdichte $f(x,\vartheta)$ einer aleatorischen Grösse x, abhängig von einem Parameter ϑ, gegeben, $x \in X$, $\vartheta \in \Theta$; x und ϑ können mehrdimensional sein. Auf Θ sei eine subjektive Funktion $g(\vartheta) \geqslant 0$ definiert, die Glaubwürdigkeitsdichte, mit der Mengenfunktion

$$G(\Theta) = \int_{\Theta} g(\vartheta)\, d\vartheta \leq \infty .$$

Die Möglichkeit $G(\Theta)=\infty$ schliesst die Normierbarkeit im allgemeinen aus.

$$l(\vartheta,x) = f(x,\vartheta)$$

heisst als Funktion von ϑ bei festem x Likelihood von ϑ. Wird ein x beobachtet, so geht $g(\vartheta)$ in ein $g^*(\vartheta)$ über vermöge

2.1 $$g^*(\vartheta) \propto l(\vartheta,x)\; g(\vartheta),$$

(RICHTER 1954 a,b). Man geht in der Theorie davon aus, dass das wahre ϑ nicht bekannt ist. Daher kann auch die 2-Wahrscheinlichkeit

$$p_{\vartheta}(X') = \int_{X'} f(x,\vartheta)\; dx$$

des Ereignisses $x \in X'$ nicht berechnet werden. Der Mensch nimmt nach RICHTER als Ersatz hierfür die 1-Wahrscheinlichkeit

2.2 $$\chi(X' \mid g) = \int_{\Theta} p_{\vartheta}(X')\; g(\vartheta)/G(\Theta).$$

Das ist der Erwartungswert von $p_{\vartheta}(X')$ bezüglich $g(\vartheta)$. Die 1-Wahrscheinlichkeit für die Hypothese, das wahre ϑ sei in Θ' enthalten, ist nach RICHTER unter gewissen Voraussetzungen

2.3 $$\chi(\Theta' \mid g) = \int_{\Theta'} g(\vartheta)\; d\vartheta/G(\Theta) = G(\Theta')/G(\Theta).$$

Gleiches gilt natürlich auch für g^*. (2.3) ist formal mit dem BAYESschen Theorem identisch, RICHTER besteht jedoch auf dem inhaltlichen Unterschied, der dadurch gegeben ist, dass subjektive und objektive Grössen darin nebeneinander vorkommen. In der Tat enthält

2.4 $$\chi(\Theta' \mid g^*) = \int_{\Theta'} l(\vartheta,x)\; g(\vartheta)\; d\vartheta \Big/ \int_{\Theta} l(\vartheta,x)\; g(\vartheta)\; d\vartheta$$

sowohl die subjektive Dichte $g(\vartheta)$ als auch die von einer 2-Wahrscheinlichkeitsdichte abgeleitete Likelihood $l(\vartheta,x)$.

2.5 $$h(\vartheta) = g(\vartheta) \;/\; G(\Theta)$$

heisst 1-Wahrscheinlichkeitsdichte. Analog lässt sich h^* definieren.

In einer reinen 1-Wahrscheinlichkeitstheorie kann man auch davon ausgehen, dass $g(\vartheta)$ und $g^*(\vartheta)$ dem Beobachter bekannt sind und die Likelihood bei $g(\vartheta) > 0$ bis auf einen konstanten Faktor aus (2.1) folgt. Sonst ändert sich im wesentlichen nichts.

3. Der Konfidenzschluss

Der Konfidenzschluss wurde 1928 von NEYMAN angegeben, um von einem Versuchsergebnis x auf den Parameter ϑ schliessen zu können. Dies geschieht wie folgt: Man definiert für jedes ϑ einen Annahmebereich A_ϑ mit den Eigenschaften

3.1 $$P_\vartheta(A_\vartheta) = \gamma,$$

3.2 $$\bigcup A_\vartheta = X.$$

(3.2) besagt, dass die Vereinigung aller Annahmebereiche die Definitionsmenge X von x sein soll; dies stellt sicher, dass jedes x wenigstens in einem A_ϑ enthalten ist. γ heisst Konfidenzkoeffizient, $\alpha = 1-\gamma$ Irrtums-2-Wahrscheinlichkeit. Wird in einem Experiment ein Ergebnis x erhalten, so ist der Konfidenzbereich definiert durch

$$K_x = \{\vartheta : x \in A_\vartheta\},$$

d.h. der Menge der ϑ, die die Eigenschaft haben, dass x in dem für die definierten Annahmebereich liegt. Es ist leicht einzusehen, dass, sofern das Konfidenzverfahren vorgegeben, d.h. die Klasse der A_ϑ vor Erhalt von x festgelegt ist, das wahre ϑ mit der 2-Wahrscheinlichkeit γ in K_x enthalten ist. Wenn man also die Konfidenzbehauptung "Das wahre ϑ liegt in K_x" aufstellt, so ist diese mit 2-Wahrscheinlichkeit γ_0 richtig, bzw. man irrt sich dabei mit 2-Wahrscheinlichkeit α. Wesentlich dabei ist, wie gesagt, dass das Konfidenzverfahren vorgegeben ist, d.h. nicht in Abhängigkeit vom Ergebnis x festgesetzt wird.

Für die weiteren Erörterungen ist es zweckmässig, sich des Beispiels der n-dimensionalen Normalverteilung zu bedienen. Für x sei dann $\mathfrak{x}$, für ϑ $\mathfrak{w}$ geschrieben. Ist $\mathcal{L}$ die (nichtsinguläre) Kovarianzmatrix, so ist die 2-Wahrscheinlichkeitsdichte

$$f(\mathfrak{x};\mathfrak{m},\mathfrak{C}) = \frac{1}{(2\pi)^{\frac{n}{2}}\,|\mathfrak{C}|^{\frac{1}{2}}}\; e^{-\frac{1}{2}(\mathfrak{x}-\mathfrak{m})'\mathfrak{C}^{-1}(\mathfrak{x}-\mathfrak{m})} \tag{3.3}$$

mit der Determinante $|\mathfrak{C}|$ von $\mathfrak{C}$. Daraus ergibt sich als Likelihood von $\mathfrak{m}$:

$$l(\mathfrak{m};\mathfrak{x},\mathfrak{C}) = f(\mathfrak{m};\mathfrak{x},\mathfrak{C}). \tag{3.4}$$

Annahmebereiche können vermöge

$$A_{\mathfrak{m}} = \left\{\mathfrak{x} : (\mathfrak{x}-\mathfrak{m})'\,\mathfrak{C}^{-1}\,(\mathfrak{x}-\mathfrak{m}) < \chi^2_{n,\gamma}\right\} \tag{3.5}$$

definiert werden mit dem Wert $\chi^2_{n,\gamma}$ der χ^2-Verteilung für n Freiheitsgrade, für den

$$p(\chi^2 < \chi^2_{n,\gamma}) = \gamma \tag{3.6}$$

ist. Dann ist nach Erhalt von $\mathfrak{x}$

$$K_{\mathfrak{x}} = \left\{\mathfrak{m} : (\mathfrak{m}-\mathfrak{x})'\,\mathfrak{C}^{-1}\,(\mathfrak{m}-\mathfrak{x}) < \chi^2_{n,\gamma}\right\}. \tag{3.7}$$

Dies ist nicht die einzige Möglichkeit, n-dimensionale Konfidenzbereiche zu definieren - bezüglich der sog. simultanen Konfidenzintervalle vgl. z.B. IHM (1955,1957), KEULS (1952,1962), SCHEFFÉ (1953), TUKEY (1949) - doch genügt die Klasse (3.7) für die Zwecke der Darstellung.

Marginale Konfidenzbereiche $K^*_{\mathfrak{y}}$ lassen sich wie folgt festlegen: Es sei für $r \leq n$

$$\mathfrak{M}'_r = \begin{pmatrix} \breve{\mathfrak{m}}'_1 \\ \cdots \\ \breve{\mathfrak{m}}'_2 \\ \cdots \\ \vdots \\ \cdots \\ \breve{\mathfrak{m}}'_r \end{pmatrix} \tag{3.8}$$

eine rxn-Matrix mit linear unabhängigen Zeilen. Dann sind die Abbildungen der $\mathfrak{m}$ auf die von den $\breve{\mathfrak{u}}_i$ aufgespannte r-dimensionale Hyperebene H^r durch

$$3.10 \qquad \mathfrak{u} = \mathfrak{U}_r' \, \mathfrak{m}$$

gegeben. Ist weiter

$$3.11 \qquad \mathfrak{y} = \mathfrak{U}_r' \, \mathfrak{m}$$

und

$$3.12 \qquad f = \mathfrak{U}_r' \, \mathfrak{L} \, \mathfrak{U}_r \, ,$$

so ist der marginale Konfidenzbereich die Abbildung von $K_{\mathfrak{x}}$ auf H^r, nämlich

$$3.13 \qquad K^*_{\mathfrak{y}} = \left\{ \mathfrak{u} : (\mathfrak{u}-\mathfrak{y})' \, f^{-1} \, (\mathfrak{u}-\mathfrak{y}) < \chi^2_{n,\gamma} \right\}.$$

Hätte man es von vornherein mit den r linearen Funktionen $\mathfrak{u}$ von $\mathfrak{m}$ zu tun gehabt, so wäre der Konfidenzbereich für $\mathfrak{u}$

$$3.14 \qquad K_{\mathfrak{y}} = \left\{ \mathfrak{u} : (\mathfrak{u}-\mathfrak{y})' \, f^{-1} \, (\mathfrak{u}-\mathfrak{y}) < \chi^2_{r,\gamma} \right\}$$

gewesen mit

$$3.15 \qquad \chi^2_{r,\gamma} < \chi^2_{n,\gamma}$$

bei $r < n$. (3.14) gibt für $\mathfrak{u}$ also eine günstigere Lösung als (3.13). Da (3.13) für alle Matrizen $\mathfrak{U}_r$ mit den geforderten Eigenschaften gilt, kann man aus (3.7) Aussagen (3.13) über beliebige $\mathfrak{u}_1, \mathfrak{u}_2, \ldots$ ableiten. Sie sind in ihrer Gesamtheit richtig mit 2-Wahrscheinlichkeit $\gamma^* \geqslant \gamma$. Ist speziell r=1, so erhält man die simultanen Konfidenzintervalle SCHEFFÉs (1953).

Hat man ein p-dimensionales $\mathfrak{x} = (x_1, x_2, \ldots x_p)$ beobachtet, so kann für die Komponenten von $\mathfrak{m}$, $m_1, m_2, \ldots m_p$, jeweils K_{x_i} mit n=1 nach (3.7) angegeben werden. Dann ist der Erwartungswert einer richtigen Aussage gleich γ.

Eine zweite Möglichkeit ist die Formulierung einer gemeinsamen Aussage über $\mathfrak{M}$, d.h. nach (3.7) über die Gesamtheit der m_i mit $n=p$; dann ist der Konfidenzkoeffizient für diese pauschale Aussage gleich γ. Betrachtet man die Irrtums-2-Wahrscheinlichkeit, so ist mit den Worten von TUKEY (1949) im ersten Falle der Fehler 'experimentwise' gleich α, im zweiten Falle der 'overall error'. Macht man in diesem zweiten Falle nur eine Aussage über $\mathfrak{m}$, so ist bei $r < n$ der 'overall error' $\alpha^* < \alpha$, man hat also Sicherheit gewonnen. Da man aber genau α verlangt hat, wäre es besser gewesen, von vornherein mit (3.13) zu arbeiten. Das geht aber dann nicht, wenn man vor dem Experiment $\mathfrak{N}'_r$ noch nicht gekannt hat, ~~und~~ also erst nach Betrachtung des Resultates $\mathfrak{y}$ diejenigen linearen Funktionen auswählt, für die man sich a posteriori interessiert. Die Entscheidungsfreiheit, die man bei Verwendung von (3.7) für $n=p$ bewahrt, bewirkt auf der anderen Seite einen ungünstigeren, d.h. grösseren Konfidenzbereich, nämlich (3.13). In der Praxis wird die Forderung nach a-priori-Festsetzung des Konfidenzverfahrens aber nicht eingehalten: man entscheidet sich nach dem Experiment für das interessierende $\mathfrak{m}$ und verwendet (3.14) mit $r=p$. Diese a-posteriori-Festsetzung des Konfidenzverfahrens durchlöchert natürlich das Konzept der Irrtums-2-Wahrscheinlichkeit.

RICHTER (1954 b) macht auf folgende Eigenschaften der Konfidenzverfahren aufmerksam:

1. sie berücksichtigen das Vorwissen, wie es sich in g (Abschn.2) aus drückt, nicht;

2. die Angabe eines Konfidenzbereiches stellt einen Abschluss dar.

Die zweite Eigenschaft wirkt sich besonders ungünstig aus, denn sie hat zur Folge, dass das durch einen Konfidenzschluss erhaltene Wissen nicht in einem zweiten verwendet werden kann, ohne dass weitere Komponenten hinzu genommen werden als das Konzept der Irrtums-2-Wahrscheinlichkeit. Zwar kann man Sequentialverfahren angeben, aber diese enden mit 2-Wahrscheinlichkeit Eins nach endlich vielen Schritten, und man ist anschliessend in der gleichen Situatuion wie beim gewöhnlichen Konfidenzschluss. Diese Schwierigkeit umgeht man nur, wenn man $\mathfrak{M}$ genügend hoch dimensioniert, z.B. derart, dass es die Parameter eines ganzen in sich geschlossenen Wissensgebietes umfasst, aber mit n wächst auch $\chi^2_{n,\gamma}$, so dass marginale Konfidenzaussagen schliesslich trivial werden.

4. Die subjektivistische Interpretation

Diese soll wieder am Beispiel der Normalverteilung erfolgen. Wegen (2.1) kann, im Gegensatz zum Konfidenzschluss, die Kenntnis über $\mathcal{M}$, die Menge der zugelassenen μ, beliebig verbessert werden. Hat man $\mathcal{Y}$ beobachtet, so ergibt sich nach (2.1)

$$g^*(\mu) \propto l(\mu; \mathcal{Y}, \mathcal{L})\ g(\mu),$$

bzw. nach (2.5)

4.1 $$h^*(\mu) = l(\mu; \mathcal{Y}, \mathcal{L})\ g(\mu)\ /\ G^*(\mu).$$

Interessiert man sich nur für $\mu \in \mathcal{N}$, so hat man es mit der marginalen Wahrscheinlichkeitsdichte

4.2 $$\bar{h}^*(\mu) = l(n; y, f)\ \bar{g}(\mu)\ /\ \bar{G}^*(\mu)$$

zu tun, ganz unabhängig davon, ob zuerst $\mathcal{Y}$ oder gleich y beobachtet wurde. Diese Unabhängigkeit der Entscheidung existiert beim Konfidenzschluss nicht, bei dem man im Voraus festsetzen muss, ob man K^*_y (3.13) oder K_y (3.14) nehmen will.

Die Konfidenzbereiche haben bezüglich h^* eine besondere Eigenschaft, zu deren Erklärung folgende Definition nötig ist:

Sei eine Wahrscheinlichkeitsdichte f(x) gegeben. Ein Bereich A mit der Eigenschaft

$$\int_A f(x)\ dx = \gamma$$

heisst γ-Quantil bezüglich f(x).

Es gilt:

Ist $g(\mu) = g_0(\mu) = 1$ für alle $\mu \in \mathcal{M}$ und $g^*(\mu) = g_0^*(\mu) = l(\mu; \mathcal{Y}, \mathcal{L}) = h_0(\mu)$, so ist $K_{\mathcal{Y}}$ ein γ-Quantil bezüglich $h_0^*(\mu)$, d.h. die 1-Wahrscheinlichkeit von $K_{\mathcal{Y}} \subset \mathcal{M}$ ist

4.3 $$\chi(K_{\mathcal{Y}} \mid g_0^*) = \gamma.$$

Beweis: Wegen der Definition der A_m gilt

$$\chi(K_y \mid \varepsilon_0) = \int_{K_y} l(m; y, \mathcal{L})\, dm = \int_{A_m} f(y; m, \mathcal{L})\, dy = \gamma,$$

woraus die Behauptung folgt. Daraus ergibt sich, dass K_y γ-Quantil der marginalen Likelihoodfunktion $l(m; y, \gamma)$ ist. Quantile sind Maße für Lokation und Dispersion. Es lässt sich also im vorliegenden Falle sagen, dass die Konfidenzbereiche jedenfalls die Likelihoodfunktion beschreiben. Nähme man an, der Mensch schlösse auf m nicht mittels des Konfidenzschlusses, sondern unter Verwendung von h^* (4.2), so liesse sich, falls $g(m) = 1$ für alle $m \in \mathcal{M}$ ist, das Verhalten der a-posteriori-Festsetzung des Konfidenzverfahrens so erklären, dass man für die marginale 1-Wahrscheinlichkeitsdichte $\bar{h}^*$ (4.3) ein γ-Quantil angeben möchte.

Wenn auf diese Eigenschaft der Konfidenzbereiche hingewiesen wird, so darf keineswegs der Eindruck entstehen, man wolle behaupten, die a-priori-Dichte müsse auf $\mathcal{M}$ konstant sein (BAYESsches Postulat). $g(m)$ ist, je nach den Umständen, verschieden. Es ist aber unter gewissen Voraussetzungen möglich, mit der Annahme seiner Konstanz praktisch zu arbeiten. Dazu sei angenommen, y habe die Dichte $f(y; m, \mathcal{L}/N)$, was z.B. dann der Fall ist, wenn y ein Mittelwert aus N Werten z_i ist mit Dichten $f(z_i; m, \mathcal{L})$. Es gelte:

4.4 $$0 \leq g(m) \leq C,$$

4.5 $$\frac{|g(m + \Delta m) - g(m)|}{|\Delta m| \; g(m)} \leq D$$

für alle $m \in \mathcal{M}$. Dann gilt für beliebig vorgegebenes $\delta > 0$ für $B = \{m : |m - y| \leq \delta\}$:

$$G^*(\bar{B}) = \int_{\bar{B}} g^*(m)\, dm \leq C \int_{\bar{B}} l(m; y, \mathcal{L}/N)\, dm = C \int_{|z| > \delta\sqrt{N}} l(z; v, \mathcal{L})\, dz$$

für

$$g^*(m) = l(m; y, \mathcal{L}/N)\, g(m)$$

und

$$z = (m - y)\sqrt{N}.$$

Ferner ist

$$G^{*}(B) = \int_{B} g^{*}(\mathfrak{m})\, d\mathfrak{m} = g(\mathcal{M})\left\{\int_{|z| \leq \delta\sqrt{N}} l(z\,; v, \mathcal{L})\, dz + \frac{\vartheta D}{\sqrt{N}}\int_{|z| \leq \delta\sqrt{N}} l(z\,; v, \mathcal{L})\, dz\right\}$$

mit geeignetem $|\vartheta| \leq 1$. Für festes δ wird also $G^{*}(\bar{B})$ gegenüber $G^{*}(B)$ bei entsprechender Wahl von N beliebig klein, so dass es nur auf $g(\mathfrak{m})$ in B ankommt. Dort gilt dann

$$4.6 \qquad g^{*}(\mathfrak{m}) = l((\mathfrak{m}-\mathcal{M})\sqrt{N}\ ;\ v, \mathcal{L})\ g(\mathcal{M})\ \left\{1 + O(\tfrac{\sqrt{N}}{D})\right\},$$

d.h. man kann bei genügend grossem N mit der Likelihoodfunktion allein arbeiten. Man muss im Einzelfall aber prüfen, ob man (4.4) und (4.5) annehmen kann. Das ist nicht mehr der Fall, wenn g unstetig ist, da dann (4.5) nicht mehr gilt. Ein wichtiger Fall liegt vor, wenn zudem $\mathfrak{M}$ eine Menge $\{\mathfrak{m}_i\}$ diskreter Massenpunkte besitzt, das heisst Punkte mit $G(\mathfrak{m}_i) > 0$, in denen $g(\mathfrak{m})$ unstetig ist. MAINLAND (196) schildert einen Fall sukzessiver Erhebung, die so lange fortgesetzt wird, bis der Konfidenzbereich $\mathfrak{m} = v$ nicht mehr enthält, wonach man $\mathfrak{m}$ als signifikant von v abweichend betrachtet. *) Dieses Verfahren ist statthaft, wenn die Voraussetzungen von (4.6) erfüllt sind, wird aber im obengenannten Fall falsch, wenn G(v) erheblich grösser als $g(\mathfrak{m})$ auf $\mathfrak{M} - v$ ist, was sehr häufig zutrifft. In jedem Fall bleibt die Quantileigenschaft der Konfidenzbereiche für die normierte Likelihoodfunktion. Oft wird empfohlen, man solle das Ergebnis eines Konfidenzschlusses vernunftgemäss beurteilen. Dies lässt sich als Berücksichtigung von $g(\mathfrak{m})$ interpretieren und reflektiert subjektivistisches Verhalten.

Die hier angegebenen Resultate lassen sich für Wahrscheinlichkeitsverteilungen auf lokal kompakten topologischen Gruppen verallgemeinern (IHM, unveröffentlicht).

---

*) Er zitiert: "We experimented on a few animals and then applied the t-test. If it did not show significance we added a few more animals and tested again, and so on, until the answer popped over the 5 p.c.line". Dann fährt er fort: "I wonder how many of the 'significant' differences in medical literature were created in this way".

5. Der P-Wert

Ist ein $\mathfrak{x}$ beobachtet worden, so pflegt man häufig nicht $K_{\mathfrak{x}}$, sondern für ein festes $\mathfrak{m} = \mathfrak{m}_0$ und

$$J = \{\mathfrak{z} : (\mathfrak{z} - \mathfrak{m}_0)' \mathfrak{L}^{-1} (\mathfrak{z} - \mathfrak{m}_0) \geq (\mathfrak{x} - \mathfrak{m}_0)' \mathfrak{L}^{-1} (\mathfrak{x} - \mathfrak{m}_0)\}$$

die 2-Wahrscheinlichkeit

$$P = \int_J f(\mathfrak{z}; \mathfrak{m}_0, \mathfrak{L})\, d\mathfrak{z}$$

anzugeben. Je kleiner P ist, desto weniger traut man der Hypothese, das wahre $\mathfrak{m}$ sei gleich $\mathfrak{m}_0$, oder aber man verwirft $\mathfrak{m}_0$ bei $P \leq \alpha$. Letzteres Verfahren ist mit einem Konfidenzschluss identisch, bei dem ja gefragt wird, ob $\mathfrak{m}_0$ im Konfidenzbereich liegt oder nicht. Definiert man

$$K_{\mathfrak{x},P} = \left\{\mathfrak{m} : (\mathfrak{m} - \mathfrak{x})' \mathfrak{L}^{-1} (\mathfrak{m} - \mathfrak{x}) < \chi^2_{n,1-P}\right\},$$

und ist $g(\mathfrak{m}) = 1$ für alle $\mathfrak{m} \in \mathfrak{M}$, so folgt für die 1-Wahrscheinlichkeit in der Bezeichnungsweise des Abschnittes 3

$$\chi(K_{\mathfrak{x},P} \mid g_0^*) = 1 - P.$$

Bei Vorliegen der Bewertung g_0^* ist also P die 1-Wahrscheinlichkeit dafür, dass $\mathfrak{m}_0$ gerade nicht mehr in $K_{\mathfrak{x},P}$ liegt. Für die Frage der Konstanz von $g(\mathfrak{m})$ gelten die gleichen Argumente wie in Abschnitt 3. Man kann unter den gleichen Einschränkungen den P-Wert also ebenfalls subjektiv interpretieren. In diesem Sinne ist das Arbeiten mit diesem Werte der Angabe eines Konfidenzbereichs überlegen.

Bibliographie

1. CARNAP, R.: 1950 - Logical Foundations of Probability. Chicago.

2. CARNAP, R. & W. STEGMÜLLER: 1959 - Induktive Logik und Wahrscheinlichkeit. Wien.

3. CRAMÉR, H.: 1946 - Mathematical Methods of Statistics. Princeton.

4. IHM, P.: a) 1955 - Eine exakte Methode als Ersatz für die Varianzanalyse in bestimmten Fällen. Der Züchter 25:365-368.

5. IHM, P.: b) 1957 - Varianzanalyse und Konfidenzbehauptungen. Der Züchter 27:172-177.

6. KEULS, M.: 1952 - The use of the "studentized range" in connection with an analysis of variance. Euphytica 1:112-122.

7. KEULS, M.: 1962 - Weten in aansluiting op een variantie-analyse. Statistica Neerlandia 16:373-387.

8. MAINLAND, D.: 1963 - Why, or why not, use a sequential design? Notes from a laboratory of medical statistics No.39.

9. POISSON, S.-D.: 1837 - Recherches sur la probabilité des jugements en matière civile et criminelle. Paris.

10. RAIFFA, H. u. R. SCHLAIFER: 1961 - Applied statistical decision theory. Division of Research, Graduate School of Business Administration, Boston.

11. RICHTER, H.: 1954 - Zur Begründung der Wahrscheinlichkeitsrechnung. (Zur Frage der Grundlagen der Wahrscheinlichkeitsrechnung II). Dialectica 8(29), 48-77.

12. RICHTER, H.: 1954 - Zur Grundlegung der Wahrscheinlichkeitstheorie. V. Indirekte Wahrscheinlichkeitstheorie. Math. Annalen 128, 305-339.

13. SCHEFFÉ, H.: 1953 - A method for judging all contrasts in the analysis of variance. Biometrika 40, 87-104.

14. SAVAGE, L.S.: 1954 - The foundations of statistics. New York.

15. SCHLAIFER, R.: 1959 - Probability and statistics for business decisions.
McGraw-Hill Book Company, Inc., New York-Toronto-London.

16. SCHLAIFER, R.: 1961 - Introduction to statistics for business decisions.
McGraw-Hill Book Company, Inc., New York-Toronto-London.

17. TUKEY, J.W.: 1949 - Comparing individual means in the analysis of variance.
Biometrics 5, 99-114.

Erkenntnisse verbreiten ist soviel wie Wohlstand verbreiten — ich meine den allgemeinen Wohlstand, nicht den individuellen Reichtum — denn mit dem Wohlstand verschwindet mehr und mehr das Böse, das uns aus dunkler Zeit vererbt ist.

Alfred Nobel

EURATOM — C.I.D.
51 - 53, rue Belliard
Bruxelles (Belgique)