Zeitschrift: Mitteilungen / Vereinigung Schweizerischer Versicherungsmathematiker = Bulletin / Association des Actuaires Suisses = Bulletin / Association of Swiss Actuaries

Herausgeber: Vereinigung Schweizerischer Versicherungsmathematiker

Band: 70 (1970)

Artikel: Statistische Rückschlüsse beim Behrens-Fisher-Problem

Autor: Streit, Franz

DOI: https://doi.org/10.5169/seals-967036

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 08.12.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Statistische Rückschlüsse beim Behrens-Fisher-Problem*

Von Franz Streit, Bern

Die Theorie der statistischen Rückschlüsse ist seit ihrer systematischen Begründung, um die sich in erster Linie R. A. Fisher und J. Neyman sehr verdient gemacht haben, gewaltig ausgebaut worden und hat als Forschungszweig zunehmend an Bedeutung gewonnen. Einerseits sind die klassischen Verfahren weiter erforscht und ihre Nutzanwendungen eingehend studiert worden; anderseits wurde besonders in den letzten Jahren auf verschiedene andere Methoden aufmerksam gemacht und ihre vermehrte Verwendung propagiert.

Angesichts dieser Sachlage scheint es zweckmässig zu sein, sich Rechenschaft darüber abzulegen, welche Techniken zur Lösung von statistischen Rückschlussproblemen grundsätzlich zur Verfügung stehen und welches die Voraussetzungen sind, unter welchen sie angewandt werden dürfen. Ferner sollte abgeklärt werden, welche Aspekte des gemeinsamen Grundproblems mit den einzelnen Verfahren bewältigt werden können. Zwar gibt es viele Statistiker, die eine Methode, mit der alle auftretenden Probleme zufriedenstellend behandelt werden können, suchen oder sogar glauben, eine solche gefunden zu haben. Angesichts der grossen Vielfalt verschiedenster statistischer Fragestellungen ist es aber eher unwahrscheinlich, dass ein derartiges universales Lösungsprinzip existiert [32, p. 921; 23, p. 373]. Anstatt in dogmatischer Weise die Unfehlbarkeit eines einzelnen Verfahrens zu verfechten, scheint es daher gewinnbringender zu sein, durch eine systematische Analyse die Leistungsfähigkeit und Besonderheiten der verschiedenen Varianten zu analysieren und auf diese Weise die notwendigen Kenntnisse, die zur Wahl der brauchbarsten Methode im Einzelfall benötigt werden, bereit zu stellen.

---

* Diese Arbeit entstand durch Überarbeitung eines am 26. Juni 1970 vor der Mathematischen Vereinigung in Bern gehaltenen Vortrages.

In dieser Arbeit, die ein Beitrag zu derartigen Studien sein soll, wird in einem ersten Abschnitt (§ 1) das Grundproblem der Theorie der statistischen Rückschlüsse kurz beschrieben. Als Illustrationsbeispiel dient uns das sogenannte Behrens-Fisher-Problem. Einleitende Ausführungen zu dieser Fragestellung, die sowohl in theoretischer Hinsicht als auch vom Standpunkt der praktischen Anwendungen bedeutsam ist, findet man in § 2. Der restliche Teil der Arbeit ist einer Übersicht über eine Auswahl der wichtigsten Rückschlussverfahren und einer Zusammenstellung ihrer Nutzanwendungen auf das Behrens-Fisher-Problem gewidmet. In jedem Abschnitt wird die allgemeine Problemlage dargelegt und Folgerungen, die sich daraus für das Behrens-Fisher-Problem ergeben, hergeleitet oder auf entsprechende Resultate in der Fachliteratur hingewiesen. Der Reihe nach gelangen die Methode der parametrischen Schätzungen (§ 3), die Methode der Konfidenzmengen (§ 4), die statistischen Testverfahren (§ 5), das Verfahren von Bayes (§ 6) und die Methode der strukturellen Rückschlüsse (§ 7) zur Darstellung.

§ 1 Zum Grundproblem der Theorie der statistischen Rückschlüsse

Zum besseren Verständnis der nachstehend angeführten Formalisierung des allgemeinen Rückschlussproblems ist es vorteilhaft, auf die Unterschiede in den Verhältnissen, die bei wahrscheinlichkeitstheoretischen und statistischen Methoden vorliegen, hinzuweisen.

Wahrscheinlichkeitstheoretische Betrachtungen nehmen Bezug auf ein wahrscheinlichkeitstheoretisches Modell, das in den einfachsten Fällen durch den sogenannten Wahrscheinlichkeitsraum gegeben ist. Bekanntlich wird ein solcher Wahrscheinlichkeitsraum $(\mathfrak{X}, \mathfrak{U}, P)$ durch die Wahl einer abstrakten Menge $\mathfrak{X}$, einer σ-Algebra $\mathfrak{U}$ von $\mathfrak{X}$ und ein auf dem Messraum $(\mathfrak{X}, \mathfrak{U})$ definiertes Wahrscheinlichkeitsmass P festgelegt. Aufgabe der Wahrscheinlichkeitstheorie ist es, Wahrscheinlichkeitsräume oder allgemeiner Klassen von Wahrscheinlichkeitsräumen zu erforschen.

Bei statistischen Untersuchungen – jedenfalls soweit sie nicht bloss deskriptiver Natur sind – spielen wahrscheinlichkeitstheoretische Überlegungen eine grosse Rolle. Indessen sind zwei grundlegende Unterschiede festzustellen. Einerseits werden Vergleiche der Modellbe-

rechnungen mit den wirklichen Gegebenheiten vorgenommen. Dies erfolgt durch Sammeln von experimentell gewonnenen Beobachtungsdaten, die zur Kontrolle der theoretischen Berechnungen dienen. Bei statistischen Rückschlussproblemen besteht ein weiterer Unterschied darin, dass wir uns nicht von vornherein auf ein bestimmtes wahrscheinlichkeitstheoretisches Modell festlegen. Tatsächlich ist dies vielfach nicht möglich oder zum mindesten nicht ratsam. Oft können eine Vielzahl von wahrscheinlichkeitstheoretischen Modellen als vernünftig und realistisch angesehen werden. Es gilt dann auf Grund von wissenschaftlichen Betrachtungen Aussagen über die Eignung der einzelnen Wahrscheinlichkeitsräume zu machen. Dies führt uns auf den Begriff des statistischen Modells.

Ein *statistisches Modell* besteht aus einer Klasse $\mathfrak{M}$ von Wahrscheinlichkeitsräumen bezüglich desselben Messraumes und einer endlichen Anzahl von experimentell gewonnenen Beobachtungswerten, die in der Regel erst nach der Festsetzung von $\mathfrak{M}$ gesammelt werden sollten. In symbolischer Schreibweise:

$$\mathfrak{M} = \{M_\theta = (\mathfrak{X}, \mathfrak{U}, F(\underset{\sim}{x} : \theta))\ [\theta \in \Omega]\}\ ; \quad \underset{\sim}{x} = (x_1, \ldots, x_n)$$

Es haben sich dabei in der Statistik die folgenden Bezeichnungsweisen eingebürgert, die z. T. von der Terminologie der Wahrscheinlichkeitstheorie abweichen: Man nennt $\mathfrak{X}$ den Stichprobenraum, $\mathfrak{U}$ die Familie der zufälligen Ereignisse, $F(\underset{\sim}{x} : \theta)$ die Klasse der in Betracht gezogenen Verteilungsfunktionen, θ den Parameter und Ω den Parameterraum. Die Aufgabe der Theorie der statistischen Rückschlüsse besteht darin, auf Grund der Beobachtungswerte $\underset{\sim}{x}$ Aussagen über die Eignung der einzelnen Modelle M_θ zur Beschreibung des untersuchten Zufallsvorganges zu machen. Setzen wir voraus, dass das den Beobachtungswerten effektiv zugrunde liegende Modell zu $\mathfrak{M}$ gehört, so soll dieses realisierte Modell M_{θ_0} und der wahre Wert θ_0 von θ auf Grund von $\underset{\sim}{x}$ soweit wie möglich bestimmt oder charakterisiert werden.

In den Abschnitten 3–7 werden einige der wichtigsten Techniken, die zur Bewältigung dieses Problems zur Verfügung stehen, kurz besprochen. Die Wahl des Verfahrens wird vernünftigerweise weitgehendst durch die spezifische Problemstellung, die der Untersuchung zugrundeliegt, und eventuell durch das Vorhandensein zusätzlicher Information über den Zufallsprozess beeinflusst.

§ 2 Das Behrens-Fisher-Problem

Es sind heute verschiedene Formulierungen des Behrens-Fisher-Problems bekannt. In ihrer ursprünglichen, einfachen Fassung lautet die Fragestellung wie folgt: Gegeben seien zwei unabhängige Stichproben

$$x_{11}, \ldots, x_{1n_1},$$

$$x_{21}, \ldots, x_{2n_2},$$

die von zwei normalverteilten Grundgesamtheiten mit Mittelwerten μ_1 und μ_2 und mit Streuungen σ_1^2 und σ_2^2 stammen. Wir nehmen also an, dass es sich bei den Beobachtungswerten $x_{11}, \ldots, x_{1n_1}$ um Realisierungen von Zufallsvariablen mit der Dichtefunktion $\frac{1}{\sqrt{2\pi}\sigma_1} \cdot \exp\left[-\frac{1}{2}\left(\frac{x_{1i}-\mu_1}{\sigma_1}\right)^2\right]$ und bei den Beobachtungswerten $x_{21}, \ldots, x_{2n_2}$ um Realisierungen von Zufallsvariablen mit der Dichtefunktion $\frac{1}{\sqrt{2\pi}\sigma_2} \cdot \exp\left[-\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]$ handelt. Ferner wird vorausgesetzt, dass alle Beobachtungswerte stochastisch voneinander unabhängig seien und dass wir die effektiven Parameterwerte $\mu_1 = \mu_{10}$, $\mu_2 = \mu_{20}$, $\sigma_1^2 = \sigma_{10}^2$ und $\sigma_2^2 = \sigma_{20}^2$ der untersuchten Grundgesamtheiten nicht kennen. Die Aufgabe besteht darin, auf Grund des beschriebenen Modells und der Beobachtungswerte Aussagen über den wahren Wert ζ_0 der Grösse $\zeta = \mu_1 - \mu_2$, also über den Unterschied zwischen den Mittelwerten der beiden Grundgesamtheiten zu machen. Es liegt in der Natur einer derartigen Problemlage, dass es natürlich nicht möglich ist, ζ_0 in unfehlbarer Weise anzugeben, etwa zu berechnen. Vielmehr muss unter den gegebenen Umständen nach Aussagen über die Mutmasslichkeit der Realisation verschiedener möglicher Werte von ζ gesucht werden. Solche Angaben werden zwar in Einzelfällen zu Fehlurteilen führen, sollten aber in gewissem – noch näher zu bezeichnendem Sinne – als optimale Schlussfolgerungen angesehen werden können.

Es braucht wohl nicht besonders betont zu werden, dass es sich bei diesem Problem um eine für die praktischen Anwendungen bedeutsame Fragestellung handelt. Sehr oft kommt man in die Lage, dass

man auf Grund zweier Zufallsstichproben Aussagen über die Differenz der Mittelwerte der zugehörigen Phänomene machen soll. Die Forderung, dass die Grundgesamtheiten normalverteilt sein sollen, entspricht zwar nicht immer, aber vielfach den praktischen Gegebenheiten. Als besonders realistisch erweist es sich, dass man keine restriktiven Voraussetzungen über die Parameter macht und auch nicht fordert, dass die Umfänge der beiden Stichproben gleich gross sind.

Die letztgenannte Tatsache ist deshalb bemerkenswert, weil in vielen Lehrbüchern der Statistik eine Lösung des vorliegenden Problems unter der Annahme hergeleitet wird, dass die beiden Streuungen σ_{10}^2 und σ_{20}^2 zwar unbekannt aber gleich gross seien. Unter dieser Einschränkung vereinfacht sich das Problem wesentlich, und es stehen dann bekannte Resultate, die auf dem t-Test oder auf Konfidenzintervallen der t-Verteilung beruhen, zur Verfügung (vgl. etwa [27] und [17]). Es ist aber offensichtlich, dass die Forderung nach Gleichheit der beiden Streuungen oft nicht realistisch ist und vor allem zur Vereinfachung der mathematischen Schwierigkeiten erhoben wird. Wir werden deshalb diese Bedingung nicht aufstellen und zulassen, dass $\sigma_{10}^2 \neq \sigma_{20}^2$.

Das Behrens-Fisher-Problem lässt sich als Spezialfall der in § 1 beschriebenen allgemeinen Fragestellung unterordnen. Hierbei besteht $\underset{\sim}{x}$ aus den Beobachtungsdaten $x_{11}, \ldots, x_{1n_1}$, und $x_{21}, \ldots, x_{2n_2}$; für $\mathfrak{X}$ ist $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ ($\mathbb{R}^k$ bezeichnet den k-dimensionalen euklidischen Raum und $\times$ die cartesische Produktbildung) und für $\mathfrak{U}$ die Klasse der Borelschen Mengen von $\mathfrak{X}$ zu wählen.

F bestimmt sich aus der Forderung, dass x_{pl} normalverteilt mit Mittelwert μ_p und Streuung σ_p^2 sein soll. $\underset{\sim}{\theta}$ schliesslich setzt sich aus den vier Angaben $(\mu_1, \mu_2, \sigma_1, \sigma_2)$ oder aus der äquivalenten Charakterisierung $(\zeta, \mu_2, \sigma_1^2, \sigma_2^2)$ zusammen $[\mu_1, \mu_2, \zeta \in \mathbb{R};\ \sigma_1, \sigma_2 > 0]$.

Das Behrens-Fisher-Problem wurde – wie der Name andeutet – zuerst von *W. V. Behrens* [2] untersucht. Auf Grund dieser Studien hat *R. A. Fisher* eine Lösung der Aufgabe vorgeschlagen [7]. Fishers Resultat ist in Form einer Wahrscheinlichkeitsverteilung für ζ gegeben. Bei dieser Lösung kann man also die Wahrscheinlichkeit berechnen, dass der effektive Wert von ζ in einer vorgegebenen Menge von reellen Zahlen liegt. Das Resultat von Fisher lässt sich folgendermassen charakterisieren:

$$\zeta = \mu_1 - \mu_2 = \bar{x}_1 - \bar{x}_2 - \left(\frac{s_{x_1}}{\sqrt{n_1}} t_{n_1-1} - \frac{s_{x_2}}{\sqrt{n_2}} t_{n_2-1} \right) \quad [n_1, n_2 \geqq 2]. \tag{1}$$

Hier und bei ähnlichen Situationen im folgenden repräsentieren $\bar{x}_p = \frac{1}{n_p} \sum_{l=1}^{n_p} x_{pl}$ den Mittelwert und $s_{x_p}^2 = \frac{1}{n_p - 1} \sum_{l=1}^{n_p} (x_{pl} - \bar{x}_p)^2$ die Streuung der p-ten Stichprobe, während t_ν eine Zufallsvariable mit der t-Verteilung von Student und ν Freiheitsgraden bezeichnet. Die Verteilung von ζ ist demnach durch einen linearen Ausdruck, in welchem zwei voneinander unabhängige t-Variablen auftreten, gegeben.

Zur praktischen Auswertung der Formel (1) wurden kurz nach Auffindung des Resultats Tabellenwerke geschaffen [9, 30]. In der statistischen Praxis wurde das Ergebnis verschiedentlich mit Erfolg angewandt. Auf die Relation (1) wird in vielen Lehrbüchern der Statistik hingewiesen, auch in solchen, die sich vorwiegend mit den statistischen Anwendungen befassen.

Dessen ungeachtet wurde bald einmal Kritik an Fishers Herleitung der Relation (1) laut. Fisher erhielt das Ergebnis mit Hilfe der sogenannten «fiduziellen Methode». Diese Rückschlusstechnik wird heute nicht mehr allgemein anerkannt. Es ist zweifelhaft, ob das Verfahren mit den anerkannten Regeln der Wahrscheinlichkeitsrechnung in Einklang gebracht werden kann. Wir brauchen uns aber mit dieser Problematik nicht auseinanderzusetzen. Im Jahre 1957 gelang *J. W. Tukey* [31] nämlich ein einwandfreier Beweis, dass Fishers Herleitung nicht stichhaltig ist. Es stellt sich heraus, dass selbst wenn man die Zweckmässigkeit der fiduziellen Methode nicht in Frage zieht, Schwierigkeiten auftreten. Die von Fisher vorgeschlagene Lösung (1) ist nämlich nicht eindeutig. Vielmehr können nach demselben Verfahren eine Vielzahl anderer Verteilungen für ζ hergeleitet werden, ohne dass die Technik einen Hinweis dafür liefern würde, welche dieser Verteilungen die richtige ist.

Nach Bekanntwerden dieser Unzulänglichkeiten, die in analoger Weise auch bei andern mehrparametrigen Anwendungen des Verfahrens aufgedeckt worden sind, versuchte man durch Modifikation der fiduziellen Methode die gewünschte Eindeutigkeit zu erzwingen. Diese Versuche sind aber bis anhin erfolglos geblieben. Man sah sich somit vor die groteske Lage gestellt, dass für eine Relation, die sich in praktischen Anwendungen bewährt hat, keine theoretische Begründung bekannt war. Angesichts dieser Sachlage ist es verständlich, dass man andere zuverlässigere statistische Rückschlussmethoden zur Behandlung des Behrens-Fisher-

Problems beanspruchte. Über die dabei erzielten Resultate wird im restlichen Teil der Arbeit ausführlich berichtet.

§ 3 Die Methode der parametrischen Schätzung

Allgemeine Problemlage: Aus der Klasse $\mathfrak{M} = \{M_\theta\}$ soll dasjenige wahrscheinlichkeitstheoretische Modell M_θ ausgewählt werden, das sich auf Grund der Beobachtungswerte $\underset{\sim}{x}$ am realistischsten erweist. Dies erfolgt durch geeignete Wahl einer Stichprobenfunktion T. Hierbei ist T eine Kollektion $\{T_n\}$ von Abbildungen T_n – für jeden möglichen Stichprobenumfang n eine – welche den Beobachtungswerten $\underset{\sim}{x} = (x_1, \ldots, x_n)$ einen Wert $T_n(x) \in \Omega$ des Parameterraums zuweist. Bei der Anwendung der Schätztheorie stellen sich zwei Hauptprobleme. Einerseits werden Verfahren zur Konstruktion von Schätzfunktionen benötigt, andrerseits ist die Güte der Schätzfunktion zu untersuchen. Die Eignung oder Güte der Stichprobenfunktion wird auf Grund gewisser Kriterien, die sich auf die für T_n induzierte Verteilung beziehen, beurteilt.

Im Spezialfall des Behrens-Fisher-Problems besteht das Schätzproblem darin, jeder Kombination von beobachteten Werten $x_{11}, \ldots, x_{1n_1}$ und $x_{21}, \ldots, x_{2n_2}$ auf eine von den Parametern unabhängige Weise vier Zahlen zuzuordnen, welche dem wahren Wert $\underset{\sim}{\theta}_0$ von $\underset{\sim}{\theta} = (\zeta, \mu_2, \sigma_1^2, \sigma_2^2)$ möglichst nahekommen.

Konstruktion der Schätzfunktion: Was die Konstruktion von Schätzfunktionen betrifft, so soll hier lediglich das gebräuchlichste Verfahren, die Bestimmung einer Schätzfunktion nach der Maximum-Likelihood-Methode, erwähnt werden. Bei der Anwendung dieser Technik bestimmt man denjenigen Wert $\hat{\theta}$ des Parameterraumes Ω, für welchen die Wahrscheinlichkeit bzw. die Wahrscheinlichkeitsdichte für die Realisierung der registrierten Beobachtungswerte am grössten wird. Als Schätzfunktion für θ wird dann $\hat{\theta}$ oder eine einfache Funktion von $\hat{\theta}$ verwendet.

Wollen wir die Maximum-Likelihood-Schätzfunktion für den Parameter $(\zeta_1, \mu_2, \sigma_1^2, \sigma_2^2)$ beim Behrens-Fisher-Problem bestimmen, so müssen wir also die Wahrscheinlichkeitsdichte dafür, die Werte $x_{11}, \ldots, x_{1n_1}$ und $x_{21}, \ldots, x_{2n_2}$, zu beobachten in bezug auf die unbekannten Grössen maximieren. Diese Dichte ist

$$L[x_1,\ x_2 : \zeta_1,\ \mu_2,\ \sigma_1^2,\ \sigma_2^2]$$

$$= \frac{1}{(2\pi)^{\frac{n_1+n_2}{2}}\ \sigma_1^{n_1}\ \sigma_2^{n_2}} \exp\left[-\frac{1}{2}\left(\sum_{i=1}^{n_1}\frac{(x_{1i}-\zeta-\mu_2)^2}{\sigma_1^2}+\sum_{j=1}^{n_2}\frac{(x_{2j}-\mu_2)^2}{\sigma_2^2}\right)\right].$$

Statt das Maximum von L können wir dasjenige von $ln\ L$ bilden. Wir erhalten also $\hat{\underset{\sim}{\theta}}$ indem wir das Gleichungssystem

$$\frac{\partial l n L}{\partial \zeta} = \sum_{i=1}^{n_1}\left(\frac{x_{1i}-\zeta-\mu_2}{\sigma_1^2}\right) = 0$$

$$\frac{\partial l n L}{\partial \mu_2} = \sum_{i=1}^{n_1}\left(\frac{x_{1i}-\zeta-\mu_2}{\sigma_1^2}\right) + \sum_{j=1}^{n_2}\left(\frac{x_{2j}-\mu_2}{\sigma_2^2}\right) = 0$$

$$\frac{\partial l n L}{\partial \sigma_1^2} = -\frac{n_1}{2\sigma_1^2} + \frac{1}{2}\sum_{i=1}^{n_1}\left(\frac{x_{1i}-\zeta-\mu_2}{\sigma_1^2}\right)^2 = 0$$

$$\frac{\partial l n L}{\partial \sigma_2^2} = -\frac{n_2}{2\sigma_2^2} + \frac{1}{2}\sum_{j=1}^{n_2}\left(\frac{x_{2j}-\mu_2}{\sigma_2^2}\right)^2 = 0$$

nach den unbekannten Parametern auflösen. Man überzeugt sich leicht davon, dass

$$\hat{\underset{\sim}{\theta}} = \left[\bar{x}_1-\bar{x}_2,\ \bar{x}_2,\ \frac{1}{n_1}\sum_{i=1}^{n_1}(x_{1i}-\bar{x}_1)^2,\ \frac{1}{n_2}\sum_{j=1}^{n_2}(x_{2j}-\bar{x}_2)^2\right]$$

die Maximum-Likelihood-Schätzfunktion für $\underset{\sim}{\theta}$ ist. Wie ersichtlich, soll für die Schätzung der uns besonders interessierenden Komponente ζ die Differenz der Mittelwerte der beiden Stichproben verwendet werden. Dies ist ein Ansatz, der aus anschaulichen Gründen unmittelbar einleuchtet.

Güte der Schätzung: Die für die gesamte Theorie massgebenden Begriffsbildungen stammen von *R. A. Fisher* [5, 6]. Im folgenden werden einige bedeutsame Kriterien zur Beurteilung der Güte einer Schätzfunktion zusammengestellt und es wird untersucht, was für Schlussfolgerungen sich daraus in Bezug auf das Behrens-Fisher-Problem ergeben.

1. Man kontrolliert, ob eine Schätzfunktion erwartungstreu ist, um zu überprüfen ob ihre Verteilung in gewissem Sinne richtig zentriert ist. Eine Schätzfunktion T für θ heisst *erwartungstreu* für θ, falls für jeden Stichprobenumfang n gilt $E[T_n:\theta]=\theta\,[\theta\in\Omega;\ n=1,2,\ldots]$. Sofern – wie in unserem Beispiel – die Schätzfunktion mehrdimensional ist, bezieht sich die Forderung auf jede Komponente gesondert, also $E[T_n^{(q)}:\underset{\sim}{\theta}]=\theta^{(q)}$ $[q=1,\ldots,L;\ L$ Dimension von $\underset{\sim}{\theta}]$.

Offensichtlich ist in unserem Beispiel die Maximum-Likelihood-Schätzfunktion $\hat{\underset{\sim}{\theta}}$ nicht erwartungstreu für den gesuchten Parameter $\underset{\sim}{\theta}$. Zwar zeigen die beiden ersten Komponenten das gewünschte Verhalten $E[\bar{x}_1-\bar{x}_2]=\mu_1-\mu_2$, $E[\bar{x}_2]=\mu_2$, aber bekanntlich gilt

$$E\left[\frac{1}{n_1}\sum_{i=1}^{n_1}(x_{1i}-\bar{x}_1)^2\right]=\frac{n_1-1}{n_1}\sigma_1^2 \quad\text{und}\quad E\left[\frac{1}{n_2}\sum_{j=1}^{n_2}(x_{2j}-\bar{x}_2)^2\right]=\frac{n_2-1}{n_2}\sigma_2^2.$$

Dem üblichen Vorgehen entsprechend ersetzen wir $\hat{\underset{\sim}{\theta}}$ daher durch die modifizierte Maximum-Likelihood-Schätzfunktion

$$\underset{\sim}{T}=\underset{\sim}{T}(\underset{\sim}{x}_1,\underset{\sim}{x}_2)=[\bar{x}_1-\bar{x}_2,\ \bar{x}_2,\ s_{x_1}^2,\ s_{x_2}^2] \quad\text{mit}\quad s_{x_p}^2=\frac{1}{n_p-1}\sum_{i=1}^{n_p}(x_{pi}-\bar{x}_p)^2.$$

Wir stellen nun fest: Die Schätzfunktion $\underset{\sim}{T}$ ist erwartungstreu für $\underset{\sim}{\theta}$.

2. Das asymptotische Verhalten einer Schätzfunktion kann beispielsweise folgendermassen beurteilt werden: Eine Schätzfunktion $\underset{\sim}{T}$ für $\underset{\sim}{\theta}$ heisst *konsistent*, falls $\underset{\sim}{T}_n$ mit wachsender Zahl der Beobachtungswerte stochastisch gegen $\underset{\sim}{\theta}$ konvergiert, d.h. sofern für jedes $\varepsilon>0$ der Grenzwert der Wahrscheinlichkeit $\lim\limits_{n\to\infty} Pr(|T_n^{(q)}-\theta^{(q)}|>\varepsilon)=0$. Hierbei wird vorausgesetzt, dass Ω ein metrischer Raum ist [10, p. 384].

Um zu überprüfen, ob die oben eingeführte Schätzfunktion $\underset{\sim}{T}(\underset{\sim}{x}_1,\underset{\sim}{x}_2)$ für die Parameterwerte des Behrens-Fisher-Problems konsistent ist, erinnern wir uns an die bekannte Tatsache, dass die Stichprobennullmomente für die Nullmomente und bei einer Stichprobe mit n Elementen ausserdem die Masszahlen $\frac{1}{n}\sum\limits_{i=1}^{n}(x_i-\bar{x})^k$ und somit auch $\frac{1}{n-1}\sum\limits_{i=1}^{n}(x_i-\bar{x})^k$ für die Hauptmomente der Verteilung der Grund-

gesamtheit konsistent sind [10, p. 306]. Daraus ergibt sich, dass die letzten drei Komponenten von $\underset{\sim}{T}$ die gewünschte Eigenschaft besitzen, falls beide Stichprobenumfänge n_1 und n_2 gegen unendlich streben. Weil die Differenz zweier unabhängiger konsistenter Schätzfunktionen für μ_1 bzw. μ_2 konsistent ist für $\mu_1 - \mu_2$, strebt auch $\bar{x}_1 - \bar{x}_2$ unter diesen genannten Voraussetzungen stochastisch nach ζ, und daraus folgt, dass in unserem Beispiel $\underset{\sim}{T}(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$ konsistent ist für $\underset{\sim}{\theta}$, sofern $n_1, n_2 \to \infty$.

3. Das nachstehende Kriterium kann verwendet werden, um zu beurteilen, ob der Informationsgehalt einer Schätzfunktion optimal ist: Eine Schätzfunktion T für θ heisst *erschöpfend*, falls für jede von T verschiedene Schätzfunktion T^* die bedingte Verteilung von T^* bei beliebigem, bekanntem Wert von T von θ unabhängig ist.

Dass die Schätzfunktion $\underset{\sim}{T}(\underset{\sim}{x}_1, \underset{\sim}{x}_2)$ für die Parameter des Behrens-Fisher-Problems erschöpfend ist, ergibt sich unmittelbar aus einem im wesentlichen auf *J. Neyman* zurückgehenden Faktorisationstheorem [20, p. 168; 15, p. 219] in Anbetracht der Tatsache, dass sich die Dichte $L[\underset{\sim}{x}_1, \underset{\sim}{x}_2 : \zeta_1, \mu_2, \sigma_1^2, \sigma_2^2]$ folgendermassen darstellen lässt:

$$L = \frac{1}{(2\pi)^{\frac{n_1+n_2}{2}} \sigma_1^{n_1} \sigma_2^{n_2}} \exp\left[\frac{n_1-1}{2\sigma_1^2} s_{x_1}^2 - \frac{n_2-1}{2\sigma_2^2} s_{x_2}^2 - \frac{n_1(\bar{x}_1-\zeta-\mu_2)^2}{2\sigma_1^2}\right] \cdot$$

$$\exp\left[\frac{n_2(\bar{x}_2-\mu_2)^2}{2\sigma_2^2}\right].$$

4. Es liegt auf der Hand, dass bei erwartungstreuen Schätzfunktionen eine starke Konzentration ihrer Verteilung um den Mittelwert als vorteilhaft erachtet wird. Wir nennen eine erwartungstreue reellwertige Schätzfunktion T für θ deren Varianz endlich ist und stets nicht grösser ausfällt als die Varianz jeder anderen erwartungstreuen Schätzfunktion T^* für θ, für die also $\mathrm{Var}[T : \underset{\sim}{\theta}] \leqq \mathrm{Var}[T^* : \underset{\sim}{\theta}]$, *wirksamst* für θ[1]). Es kann bewiesen werden, dass im allgemeinen die Varianz einer erwartungstreuen Schätzfunktion T für eine vorgegebene Parameter-

---

[1]) In der Literatur wird diese Bezeichnung für verschiedenartige Tatbestände benützt. Wir übernehmen hier die Begriffsbildung von *S. Wilks* [35, p. 351].

funktion $\tau(\underset{\sim}{\theta})$ eine aus $\tau(\underset{\sim}{\theta})$ und der Klasse $\{F(\underset{\sim}{x}:\underset{\sim}{\theta})\}$ berechenbare, positive Schranke nicht unterschreiten kann. In der Tat gilt unter schwachen Regularitätsbedingungen die Cramér-Rao-Ungleichung [17, p. 16]:

$$\mathrm{Var}[T:\underset{\sim}{\theta}] \geqq \sum_{l=1}^{L} \sum_{m=1}^{L} \frac{\partial \tau(\underset{\sim}{\theta})}{\partial \theta^{(l)}} \cdot \frac{\partial \tau(\underset{\sim}{\theta})}{\partial \theta^{(m)}} I_{lm}^{-1}$$

mit
$$I_{lm} = E\left[\frac{1}{L}\frac{\partial L}{\partial \theta^{(l)}} \cdot \frac{1}{L}\frac{dL}{\partial \theta^{(m)}} : \underset{\sim}{\theta}\right]$$

und
$$L = L[\underset{\sim}{x}:\underset{\sim}{\theta}] : \text{Dichte von } \underset{\sim}{x} \text{ im Modell } M_{\underset{\sim}{\theta}}.$$

Eine wirksamste Schätzfunktion, deren Varianz überdies stets gleich der unteren Schranke der Cramér-Rao-Ungleichung ist, heisst *MVB–Schätzfunktion* (minimum variance bound estimator).

Durch direkte Verifikation kann man feststellen, dass beim Behrens-Fisher-Problem $\bar{x}_1 - \bar{x}_2$ eine MVB-Schätzfunktion und daher auch wirksamste Schätzfunktion für den unbekannten Parameter ζ ist. Setzt man nämlich $\tau(\mu_1, \mu_2, \sigma_1, \sigma_2) = \zeta = \mu_1 - \mu_2$, so erhält man für die rechte Seite der Cramér-Rao-Ungleichung – unter Berücksichtigung von $\frac{\partial \tau}{\partial \mu_1} = -\frac{\partial \tau}{\partial \mu_2} = 1$, $\frac{\partial \tau}{\partial \sigma_1} = \frac{\partial \tau}{\partial \sigma_2} = 0$, $I_{11} = \frac{n_1}{\sigma_1^2}$, $I_{12} = I_{21} = 0$ und $I_{22} = \frac{n_2}{\sigma^2}$ – den Ausdruck $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$, der mit der Varianz von $\bar{x}_1 - \bar{x}_2$ übereinstimmt.

Zusammenfassend darf festgestellt werden, dass beim Behrens-Fisher-Problem $\bar{x}_1 - \bar{x}_2$ eine Schätzfunktion für ζ ist, die in mancher Beziehung als besonders geeignet bezeichnet werden darf.

§ 4 Konfidenzmengen

Allgemeine Problemlage: Bei der Methode der Konfidenzmengen wird unter gewissen Einschränkungen aus der Klasse $\mathfrak{M} = \{M_\theta[\theta \in \Omega]\}$ der in Betracht gezogenen Modelle M_θ eine Unterklasse $\{M_\theta[\theta \in \Omega'(\underset{\sim}{x})]\}$ ausgewählt, die sich – anhand der Beobachtungswerte beurteilt – am

besten zur Beschreibung des Zufallsvorganges eignen. Genauer gesagt, handelt es sich darum, zu einer vorgegebenen Zahl α $[0 \leqq \alpha \leqq 1]$ und zu jeder möglichen Kombination von Beobachtungswerten $x_1, \ldots, x_n$ eine Teilmenge $\Omega'(\underset{\sim}{x}; 1-\alpha)$, genannt Konfidenzmenge von Ω, so zu bestimmen, dass die Wahrscheinlichkeit dafür, dass $\Omega'(\underset{\sim}{x}; 1-\alpha)$ θ überdeckt, im Modell M_θ gleich $1-\alpha$ ist, d.h. $Pr(\theta \in \Omega(\underset{\sim}{x}; 1-\alpha) : \theta) = 1-\alpha$ für alle $\theta \in \Omega$. Ist Ω eine Teilmenge des L-dimensionalen reellen Zahlenraumes R^L, so nennen wir $\Omega'(\underset{\sim}{x}; 1-\alpha)$, falls diese Menge zusammenhängend ist, einen Konfidenzbereich mit Konfidenzkoeffizienten $1-\alpha$. Für $L = 1$ nennen wir $\Omega'(\underset{\sim}{x}; 1-\alpha)$ unter diesen Bedingungen einen Konfidenzintervall für θ. $\Omega'(\underset{\sim}{x}; 1-\alpha)$ lässt sich in diesem Fall durch eine Bedingung der Form $\theta_1(\underset{\sim}{x}; 1-\alpha) \leqq \theta \leqq \theta_2(\underset{\sim}{x}; 1-\alpha)$ charakterisieren, wobei θ_1 und θ_2 nur von $\underset{\sim}{x}$ und α abhängig sind. Eine systematische Begründung dieser Rückschlussmethode verdankt man *J. Neyman* [21].

Anwendungen der Theorie auf das Behrens-Fisher-Problem: Es handelt sich hier darum, zu vorgegebenem α auf Grund der Beobachtungswerte $x_{11}, \ldots x_{1n_1}$ und $x_{21}, \ldots, x_{2n_2}$ ein Konfidenzintervall für ζ zu ermitteln. *S. Wilks* [34] hat im Jahre 1940 festgestellt, dass im allgemeinen – also abgesehen von einigen Spezialfällen wie z.B. $n_1 = n_2$ – keine Lösung des Behrens-Fisher-Problems in dieser Form existiert.

Es können aber approximative Lösungen angegeben werden [1, 3, 4, 25, 26, 33]. Die bekanntesten unter ihnen beruhen auf der Idee, durch unterschiedliche Berücksichtigung der Beobachtungswerte der grösseren Stichprobe die Schwierigkeiten, die sich aus den ungleichen Stichprobenumfängen ergeben, zu eliminieren. *Bartlett* schlug vor, aus der grösseren Stichprobe $(\max.(n_1, n_2) - \min.(n_1, n_2))$ Beobachtungswerte zufällig auszulesen und wegzulassen. Demgegenüber verbesserte *H. Scheffé* [25, 26] das Verfahren, indem er alle Beobachtungswerte in die Rechnung einbezieht, aber ungleich bewertet. Die Gewichtung wird so durchgeführt, dass der Erwartungswert der Länge der auf der t-Verteilung beruhenden Konfidenzintervalle minimal ausfällt. Diese Technik wird in der Literatur manchmal «exaktes Verfahren von Scheffé» genannt. Die Bezeichnungsweise ist insofern etwas irreführend, als nicht alle beobachteten Werte gleichmässig in Rechnung gestellt werden.

Kürzlich wurden auch Versuche unternommen, auf sequentielle Weise Konfidenzintervalle zu konstruieren [24, 28]. Hierbei werden die

Anzahlen n_1 und n_2 der Beobachtungswerte aus den beiden Stichproben nicht im voraus festgelegt, sondern sind von den Ergebnissen des Experimentes abhängig. Auf ein solches Verfahren, das gewisse wünschenswerte Eigenschaften besitzt, verwies *Srivastava* [29].

§ 5 Statistische Testverfahren

Allgemeine Problemlage: Auf Grund der beobachteten Werte soll entschieden werden, ob eine im voraus getroffene Annahme H_0 über die Klasse $\mathfrak{M}$ der Wahrscheinlichkeitsräume M_θ als richtig oder falsch anzusehen ist. Wir bezeichnen die Menge der Parameter θ, für welche H_0 richtig ist, mit ω. Die erforderliche Entscheidung hat auf Grund eines statistischen Tests, d.h. durch Angabe der Teilmenge S des Stichprobenraumes $\mathfrak{X}$, für welche H_0 verworfen wird, zu erfolgen. Hierbei sei das Maximum bezüglich $\theta \in \omega$ der Wahrscheinlichkeit dafür, dass H_0 fälschlicherweise in M_θ verworfen wird, gleich einer vorgeschriebenen Zahl α (Sicherheitsschwelle des Tests). Ein Test heisst ähnlich bezüglich $\mathfrak{X}$, falls diese Wahrscheinlichkeit für alle $\theta \in \omega$ gleich α ist.

Anwendungen auf das Behrens-Fisher-Problem: Im Rahmen der Theorie der statistischen Tests ist es zum Beispiel sinnvoll, nach einem Test für die Hypothese $\zeta = 0$ beim Behrens-Fisher-Problem zu suchen. Es scheint noch keine zufriedenstellende Lösung dieser Fragestellung gefunden worden zu sein [18, p. 230]. Die von R. A. Fisher vorgeschlagene Verteilung (1) beruht nicht auf dem Prinzip wiederholter Stichprobenbildung bezüglich $\mathfrak{X}$ und kann daher nicht verwendet werden [22]. Grundlegende Untersuchungen über die Existenz ähnlicher Tests beim Behrens-Fisher-Problem stammen von *Linnik* [19].

§ 6 Das Verfahren von Bayes

Allgemeine Problemlage: Vielfach stehen bei statistischen Untersuchungen auf Grund früherer Experimente oder theoretischer Betrachtungen abgesehen von den genannten Modellvoraussetzungen noch zusätzliche Kenntnisse über den zu untersuchenden Zufallsvorgang zur Verfügung. Für die Anwendung der Methode von Bayes ist erforderlich, dass der unbekannte Parameter θ als Zufallsvariable, dessen Wahr-

scheinlichkeitsverteilung bekannt ist, interpretiert werden kann. Wir haben es dann mit einem speziellen statistischen Modell

$$\mathfrak{M}^{(B)} = \{M_\vartheta(\mathfrak{X}, \mathfrak{U}, F(\underset{\sim}{x}:\theta))[\theta \in \Omega]; G(\theta)\}; \quad \underset{\sim}{x}$$

zu tun, bei dem abgesehen von den üblichen Angaben noch eine Verteilungsfunktion $G(\theta)$ gegeben ist, welche charakterisiert, mit welchen Wahrscheinlichkeiten die einzelnen Modelle M_θ realisiert werden. Sind diese Voraussetzungen erfüllt, so kann man nach Durchführung des Experimentes diese Vorkenntnisse mit der Information, welche durch die Beobachtungswerte geliefert wird, kombinieren. Auf Grund des Theorems von Bayes gewinnt man aus der sogenannten A-priori-Verteilung $G(\theta)$ die A-posteriori-Verteilung $G^{(B)}(\theta | x_1, \ldots, x_n)$, welche die dem Statistiker nach Abschluss des Versuches zur Verfügung stehenden Kenntnisse wiederspiegelt. Sind sowohl $G(\theta)$ als auch $F(\underset{\sim}{x}:\theta)$ absolut stetige Verteilungsfunktionen mit Dichtefunktionen $g(\theta)$ und $f(\underset{\sim}{x}:\theta)$ so nimmt die der Analyse von Bayes zugrunde liegende Beziehung die folgende Gestalt an

$$dG^{(B)}(\theta:\underset{\sim}{x}) = g^{(B)}(\theta:\underset{\sim}{x})\,d\theta = \frac{f(x:\theta)\,g(\theta)}{\int\limits_{\Omega} f(x:\theta)\,g(\theta)\,d\theta}\,d\theta\,,$$

wobei mit $g^{(B)}(\theta:\underset{\sim}{x})$ die Dichtefunktion der A-posteriori-Verteilung bezeichnet wird.

Bei praktischen Anwendungen ist die Voraussetzung, dass die A-priori-Verteilung von θ bekannt sein soll, vielfach nicht erfüllt. Gewisse Anhänger der Bayesschen Methode empfehlen in diesem Fall die Bestimmung von $G(\theta)$ auf Grund von theoretischen Überlegungen. Ein häufig gebrauchtes Verfahren ist das Prinzip der gleichmässigen Unkenntnis (Law of equal ignorance). Es basiert auf der Idee, dass bei völliger Unkenntnis der A-priori-Verteilung die Gleichverteilung für alle Parameterwerte die Wahrscheinlichkeitsverhältnisse am besten wiedergibt. Demzufolge ist jeder Wert des Parameterraumes gleichwertig in Rechnung zu stellen. Die Anhänger der subjektiven Wahrscheinlichkeitstheorie befürworten demgegenüber die Bestimmung einer persönlichen A-priori-Verteilung, die nicht auf objektive Weise verifizierbar zu sein braucht.

Anwendungen auf das Behrens-Fisher-Problem: Sofern eine A-priori-Verteilung für den Parameter $(\mu_1, \mu_2, \sigma_1, \sigma_2)$ resp. $(\zeta, \mu_2, \sigma_1^2, \sigma_2^2)$ objektiverweise vorliegt, so sollte die Methode von Bayes unbedingt angewandt werden, weil gemäss eines allgemein anerkannten Grundsatzes alle zur Verfügung stehende Information bei statistischen Rückschlüssen in Rechnung zu stellen ist [8, p. 55]. Eine Nichtberücksichtigung der Kenntnis von $G(\theta)$ könnte zu verfälschten Resultaten führen. Über die Anwendung des Verfahrens in allen andern Fällen lässt sich in Unkenntnis der beabsichtigten Nutzanwendung nichts Allgemeinverbindliches sagen. Es sei hier lediglich noch erwähnt, dass die von Fisher gefundene Reziehung (1) unter Anwendung der Analyse von Bayes hergeleitet werden kann, falls speziell

$$dG(\mu_1, \mu_2, \sigma_1, \sigma_2) = d\mu_1\, d\mu_2 \frac{d\sigma_1}{\sigma_1} \frac{d\sigma_2}{\sigma_2}$$

als Wahrscheinlichkeitselement für den Parameter gewählt wird (vgl. [17, p. 151]). Gründe für die Wahl von $d\mu$ bei Lageparametern und $\frac{d\sigma}{\sigma}$ bei Streuungsparametern als Wahrscheinlichkeitselement der A-priori-Verteilung, sofern diese nicht empirisch bestimmt werden kann, werden in [16] angeführt.

§ 7 Strukturelle Rückschlüsse

Einführendes Beispiel (Formulierung des Behrens-Fisher-Problems im Rahmen eines strukturellen Modells): Bei den bisher beschriebenen Lösungsversuchen des Behrens-Fisher-Problems werden die Beobachtungswerte $x_{11}, \ldots, x_{1n_1}$ und $x_{21}, \ldots, x_{2n_2}$ als die primären Grössen des statistischen Experimentes angesehen und angenommen, dass die Werte der ersten Stichprobe von einer normalverteilten Grundgesamtheit mit unbekanntem Mittelwert $\mu_1 = \mu_{10}$ und unbekannter Streuung $\sigma_1^2 = \sigma_{10}^2$ und die Werte der zweiten Stichprobe von einer normalverteilten Grundgesamtheit mit unbekanntem Mittelwert $\mu_2 = \mu_{20}$ und unbekannter Streuung $\sigma_2^2 = \sigma_{20}^2$ stammen. Demgegenüber bilden bei der strukturellen Rückschlussmethode die auf die Beobachtungswerte wirkenden Zufallsschwankungen den Ausgangspunkt der Betrachtun-

gen und es wird angenommen, dass diese durch eine Zufallsvariable (zufällige Fehlergrösse) erzeugt werden. Bei dieser Deutung entsprechen den oben angegebenen Beobachtungswerten die Fehlerwerte $e_{11}, \ldots, e_{1n_1}$ und $e_{21}, \ldots, e_{2n_2}$. e_{pl} gibt hierbei den Einfluss des Zufalls auf die Beobachtung x_{pl} an. Im Hinblick auf die spezifische Problemlage setzen wir voraus, dass e_{pl} eine von den übrigen e-Werten unabhängige Realisation einer standartisiert normalverteilten Zufallsvariablen ist. Natürlich besteht eine Bindung zwischen e_{pl} und x_{pl}. Es ist naheliegend und sinnvoll, zu fordern, dass der Zusammenhang zwischen x_{pl} und e_{pl} durch die Beziehung (sogenannte strukturelle Gleichung)

$$x_{pl} = \mu_p + \sigma_p e_{pl} := [\mu_p, \sigma_p] o\, e_{pl}$$

$$[p = 1, 2;\ \ l = 1, \ldots, n_p;\ \ \mu_1, \mu_2 \in \mathbb{R};\ \ \sigma_1, \sigma_2 > 0]$$

gegeben ist. Die Werte der ersten bzw. zweiten Stichprobe ergeben sich somit, indem man die zugehörigen Fehlerwerte mit der unbekannten Standartabweichung der Grundgesamtheit multipliziert und den unbekannten Mittelwert der Grundgesamtheit addiert. Bemerkenswert ist, dass man den unbekannten Parameterwerten in eineindeutiger Weise ein Element einer Gruppe von Transformationen des Stichprobenraumes $\mathbb{R}^{n_1+n_2}$ sämtlicher Beobachtungen auf sich selbst zuordnen kann. In der Tat entspricht den Angaben $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ das Paar der sich auf $\mathbb{R}^{n_1}$ bzw. $\mathbb{R}^{n_2}$ beziehenden linearen Transformationen $\langle[\mu_1, \sigma_1], [\mu_2, \sigma_2]\rangle$, welches zur Gruppe $G = \{\langle[\alpha_1, \beta_1], [\alpha_2, \beta_2]\rangle;\ \alpha_1, \alpha_2 \in \mathbb{R};\ \beta_1, \beta_2 > 0\}$ gehört. In der obigen Anschrift und im folgenden kennzeichnet das Zeichen o die Ausübung einer Transformation. In [12] wird gezeigt, dass bei den meisten wichtigen Rückschlussproblemen (wie z. B. bei der Regressions-, Varianz- und Faktoranalyse) eine derartige eineindeutige Entsprechung zwischen den unbekannten Parametern und den Elementen einer auf den Stichprobenraum wirkenden Transformationsgruppe aufgedeckt werden kann.

Man beachte, dass auch bei dieser Deutung der Problemlage die Beobachtungswerte x_{pl} normalverteilt mit Durchschnitt μ_p und Streuung $\sigma_p^2 [p = 1, 2]$ sind. An der Grundkonzeption hat sich also nichts geändert. Durch die Interpretation im Sinne der strukturellen Methode erhält man indessen zusätzliche Information über die Art und Weise, wie die Beobachtungswerte entstehen, was mit Hilfe der strukturellen Analyse verwertet werden kann (vergleiche hierzu auch [11]).

Allgemeine strukturelle Modelle und strukturelle Analyse: Gegenüber dem üblichen statistischen Modell

$$\mathfrak{M} = \{M_\theta = (\mathfrak{X}, \mathfrak{U}, F(\underset{\sim}{x}:\theta))\ [\theta \in \Omega]\};\ \underset{\sim}{x}$$

ist ein strukturelles Modell durch folgende Spezifikationen gekennzeichnet:

$$\mathfrak{M}^{(S)} = \{M_\theta^{(S)} = (\mathfrak{X}, \mathfrak{U}, X = \theta\, o\, E,\ F(E))\ [\theta \in G]\};\ \underset{\sim}{x}\,.$$

Strukturelle Modelle sind also spezielle statistische Modelle, bei denen der unbekannte Parameter als Element einer Transformationsgruppe G interpretiert werden kann und ausserdem die Verteilung $F(\underset{\sim}{x}:\theta)$ des Modells $M_\theta^{(S)}$ implizite durch Angabe einer strukturellen Gleichung $X = \theta o E$ und der Verteilung der Fehlergrössen E festgelegt wird. Vielfach sind hierbei – wie etwa im Beispiel des Behrens-Fisher-Problems – die Grössen X – welche die Beobachtungswerte repräsentiert – θ und E durch mehrere Zahlenangaben charakterisiert. Beim Behrens-Fisher-Problem ist $\mathfrak{X} = \mathbb{R}^{n_1+n_2}$, $\mathfrak{U}$ die σ-Algebra der Borelmengen von $\mathbb{R}^{n_1+n_2}$ und $F(E)$ durch die Forderung bestimmt, dass die Fehlergrössen voneinander unabhängig und standartisiert normalverteilt sein sollen. Die massgebende Deutung der Parameter als Gruppenelemente und die für dieses Beispiel zuständige strukturelle Gleichung wurden bereits besprochen.

Ohne auf die technischen Einzelheiten einzutreten soll hier kurz die Hauptidee erläutert werden, die den strukturellen Rückschlüssen zugrunde liegt. Wir gehen von der Feststellung aus, dass die Beobachtungsdaten X nicht von einem beliebigen Element E von $\mathfrak{X}$ herrühren können. Bei bekanntem X fallen nur diejenigen Werte für E in Betracht, welche durch Anwendung einer Transformation der Gruppe G in X übergeführt werden können. Mit andern Worten muss gelten: $E \in \{g o X[g \in G]\}$ oder – wie man auch sagt –, E muss zum G-Orbit von X gehören. Bei allen Anwendungen der strukturellen Theorie ist diese Menge stets eine echte Untermenge des Stichprobenraumes $\mathfrak{X}$. Betrachten wir die Verhältnisse beim Behrens-Fisher-Problem: $\mathfrak{X}$ besteht hier aus der Zusammenfassung von $\mathbb{R}^{n_1}$ mit $\mathbb{R}^{n_2}$. Wenn man bei gegebenem X, also gegebenem $\underset{\sim}{x}_1$ und $\underset{\sim}{x}_2$, den G-Orbit von X bestimmt, indem man σ-Vielfache von $\underset{\sim}{x}_1$ bzw. $\underset{\sim}{x}_2$ bildet und μ-Vielfache des $\underset{\sim}{1}$-Vektors in $\mathbb{R}^{n_1}$ und $\mathbb{R}^{n_2}$ bildet, so erhält man das cartesische Produkt einer Halb-

ebene in $\mathbb{R}^{n_1}$ und $\mathbb{R}^{n_2}$, also eine Teilmenge von $\mathfrak{X}$. Aus dem bisher Gesagten folgt, dass nach Bekanntwerden von X der G-Orbit, auf welchem das zugehörige E liegt, festgelegt ist. Für die Beurteilung der wahrscheinlichkeitstheoretischen Situation ist dann nicht die Verteilung $F(E)$ über dem gesamten Stichprobenraum relevant; von Interesse ist vielmehr die Verteilung von E über dem G-Orbit von X. In der strukturellen Analyse ist es üblich, jeden Wert E durch zwei Angaben $\langle D(E), [E]\rangle$ zu kennzeichnen. Hierbei ist $D(E)$ ein passend gewählter Bezugspunkt auf dem Orbit von E, welcher diesen Orbit charakterisiert, und $[E]$ bezeichnet dasjenige Element der Transformationsgruppe, welches $D(E)$ in E überführt. In der allgemeinen Theorie wird – unter schwachen Voraussetzungen – aus dem bekannten allgemeinen Wahrscheinlichkeitselement $dF(E)$ das Wahrscheinlichkeitselement $dG([E]:D)$ der bedingten Verteilung der Position $[E]$ von E auf dem Orbit, bei bekanntem Orbit, also bekanntem D hergeleitet. Nun interessiert uns aber nicht primär die Verteilung dieser Fehlercharakteristik. Vielmehr möchten wir Aufschluss über den unbekannten Parameter θ erhalten. Aus der strukturellen Gleichung $X = \theta o E$ folgt aber unmittelbar $[X] = \theta \cdot [E]$ und $[E] = \theta^{-1} \cdot [X]$ (Das Symbol $\cdot$ deutet hier Gruppenmultiplikation an.). Wir können also durch eine einfache Transformation aus der Fehlerverteilung die Verteilung für den Parameter θ erhalten. Diese sogenannte «strukturelle Verteilung» des Parameters ist durch ein Wahrscheinlichkeitselement $dG^*(\theta : X)$ gegeben. $dG^*(\theta : X)$ gibt das Wahrscheinlichkeitselement dafür an, dass der Parameter gleich θ ist unter Berücksichtigung der durch die Beobachtungsdaten X gelieferten Information. Die strukturelle Analyse liefert uns also ein natürliches Restriktionsprinzip, welches uns gestattet, den Übergang von $dF(E)$ zu dem für die Fragestellung bedeutsameren $dG^*(\theta : X)$ zu machen. G^* repräsentiert unsere Kenntnisse über den wahren Wert θ_0 von θ nach Durchführung des Experiments. Mit diesen Angaben ist im wesentlichen dargelegt, wie man zu strukturellen Rückschlüssen gelangt. Alle erforderlichen Details der Theorie findet man in [12].

Strukturelle Analyse des Behrens-Fisher-Problems: In einer kürzlich fertiggestellten Arbeit [14] konnten Prof. *D.A.S.Fraser* (Universität Toronto) und der Autor zeigen, dass die Methode der strukturellen Rückschlüsse zur Lösung des Behrens-Fisher-Problems und einer Vielzahl von Verallgemeinerungen dieser Fragestellung herangezogen wer-

den kann (vgl. auch [13]). Wir begnügen uns an dieser Stelle damit, kurz darzulegen, dass man mit Hilfe dieses Verfahrens als Lösung der von uns betrachteten urspünglichen Version des Behrens-Fisher-Problems die von R. A. Fisher vorgeschlagene Beziehung (1) erhält. Auf Grund der strukturellen Analyse des zu Beginn dieses Paragraphen besprochenen Modells findet man durch Spezialisation einer allgemeinen Formel aus der Theorie [12, p. 64] für die strukturelle Dichtefunktion der Parameter den Ausdruck

$$g^*_{\mu_1, \mu_2, \sigma_1, \sigma_2}(\mu_1, \mu_2, \sigma_1, \sigma_2 : X)\, d\mu_1\, d\mu_2\, d\sigma_1\, d\sigma_2 =$$

$$= \prod_{p=1}^{2} \left(c_p \exp\left[-\frac{1}{2\sigma_p^2} \sum_{l=1}^{n_p} (x_{pl} - \mu_p)^2 \right] \frac{s_{x_p}^{n_p-1}}{\sigma_p^{n_p+1}}\, d\sigma_p\, d\mu_p \right)$$

$$[-\infty < \mu_1, \mu_2 < +\infty;\ \sigma_1, \sigma_2 > 0;\ n_1, n_2 \geqq 2]$$

mit
$$c_p = \sqrt{\frac{n_p}{2\pi}}\, \frac{(n_p - 1)^{\frac{n_p-1}{2}}}{\Gamma\left(\frac{n_p-1}{2}\right) 2^{\frac{n_p-3}{2}}} \qquad [p = 1, 2].$$

Γ bezeichnet hier die Gammafunktion. Die Randverteilung der Lageparameter μ_1 und μ_2 erhält man, indem man über alle möglichen Werte von σ_1 und σ_2 integriert, wobei sich die Transformation

$u_p = \frac{1}{2\sigma_p^2} \sum (x_{pl} - \mu_p)^2$ als nützlich erweist.

$$g^*_{\mu_1, \mu_2}(\mu_1, \mu_2 : X)\, d\mu_1\, d\mu_2 =$$

$$= \prod_{p=1}^{2} \left[\left(\int_0^\infty c_p \exp\left[-\frac{1}{2\sigma_p^2} \sum_{l=1}^{n_p} (x_{pl} - \mu_p)^2 \right] \frac{s_{x_p}^{n_p-1}}{\sigma_p^{n_p+1}}\, d\sigma_p \right) \cdot d\mu_p \right]$$

$$= \prod_{p=1}^{2} \left[\frac{c_p 2^{\frac{n_p}{2}-1} s_{x_p}^{n_p-1}}{\left(\sum (x_{pl}-\mu_p)^2\right)^{\frac{n_p}{2}}} \left(\int_0^\infty e^{-u} u^{\frac{n_p}{2}-1} du \right) d\mu_p \right]$$

$$= \prod_{p=1}^{2} \left[\frac{c_p \Gamma\left(\frac{n_p}{2}\right) 2^{\frac{n_p}{2}-1} s_{x_p}^{n_p-1}}{\left((n_p-1)\, s_{x_p}^2 + n_p(\bar{x}_p-\mu_p)^2\right)^{\frac{n_p}{2}}} d\mu_p \right]$$

$$= \prod_{p=1}^{2} \left[\frac{\sqrt{n_p}}{B\left[\frac{1}{2}, \frac{n_p-1}{2}\right] s_{x_p} \sqrt{n_p-1}} \cdot \frac{1}{\left(1 + \frac{1}{n_p-1}\left(\frac{\sqrt{n_p}\,(\mu_p-\bar{x}_p)}{s_{x_p}}\right)^2\right)^{\frac{n_p}{2}}} d\mu_p \right]$$

Es zeigt sich also, dass die Grössen μ_1 und μ_2 voneinander unabhängig und im wesentlichen t-verteilt sind. Es gilt

$$\mu_p = \bar{x}_p + \frac{s_{x_p}}{\sqrt{n_p}} t_{n_p-1} \qquad [p = 1, 2].$$

Die uns insbesondere interessierende strukturelle Verteilung von $\zeta = \mu_1 - \mu_2$ ist durch

$$g_\zeta^*(\zeta : X)\, d\zeta = \left(\int_{-\infty}^{+\infty} g_{\mu_1, \mu_2}^*(\mu_1, \mu_1 - \zeta)\, d\mu_1 \right) d\zeta$$

gegeben. Dieses Resultat kann auf äquivalente Weise auch folgendermassen dargestellt werden:

$$\zeta = \mu_1 - \mu_2 = \bar{x}_1 - \bar{x}_2 - \left(\frac{s_{x_1}}{\sqrt{n_1}} t_{n_1-1} - \frac{s_{x_2}}{\sqrt{n_2}} t_{n_2-1} \right) \qquad [n_1, n_2 \geqq 2].$$

Dies ist die ursprünglich von Fisher vorgeschlagene Lösung (1) des Behrens-Fisher-Problems. Damit haben wir diese Relation im Rahmen der strukturellen Theorie neu begründet und auf widerspruchlose Weise hergeleitet.

Literaturverzeichnis

[1] *Banerji, S.K.:* Approximate Confidence Interval for Linear Functions of Means of K Populations when the Population Variances are not Equal. Sankhyā *22*, 357–358, 1960.

[2] *Behrens, W.V.:* Ein Beitrag zur Fehlerberechnung bei wenigen Beobachtungen. Landwirtschaftliche Jahresberichte *68*, 807–837, 1929.

[3] *Chand, U.:* Distributions Related to Comparison of Two Means and Two Regression Coefficients. Ann. Math. Stat.*21*, 507–522, 1950.

[4] *Chernoff, H.:* Asymptotic Studentization in Testing of Hypotheses. Ann. Math. Stat. *20*, 268–278, 1949.

[5] *Fisher, R.A.:* On The Mathematical Foundations of Theoretical Statistics. Phil. Trans. Roy. Soc. London, *Ser. A 222*, 309–368, 1922.

[6] – Theory of Statistical Estimation. Proc. Camb. Phil. Soc. *22*, 700–725, 1925.

[7] – The Fiducial Argument in Statistical Inference. Ann. of Eugenics *6*, 391–398, 1935.

[8] – Statistical Methods and Scientific Inference. Oliver and Boyd, Edinburgh 1956.

[9] *Fisher, R.A., Yates, F.:* Statistical Tables for Biological, Agricultural and Medical Research. Oliver and Boyd, Edinburgh 1963.

[10] *Fisz, M.:* Wahrscheinlichkeitsrechnung und Mathematische Statistik. Deutscher Verlag der Wissenschaften, Berlin 1965.

[11] *Fraser, D.A.S.:* A Black Box or a Comprehensive Model. Technometrics *10*, 219–229, 1968.

[12] – The Structure of Inference. J. Wiley, New York 1968.

[13] *Fraser, D.A.S., Streit, F.:* Some Applications of the Method of Structural Inference Involving Several Unknown Parameters. Abstract. Annals of Mathematical Statistics *40*, 1871, 1969.

[14] – On the Behrens-Fisher-Problem. Zur Veröffentlichung eingereicht, 1970.

[15] *Hogg, R.V., Craig, A.T.:* Introduction to Mathematical Statistics. Macmillan Co., London 1970.

[16] *Jeffreys, H.:* Theory of Probability. Clarendon Press. Oxford 1961.

[17] *Kendall, M.G., Stuart, A.:* The Advanced Theory of Statistics. Vol. II. Ch. Griffin Co., London 1961.

[18] *Lehmann, E.L.:* Testing Statistical Hypotheses. J. Wiley, New York 1959.

[19] *Linnik, Y.W.:* On the Behrens-Fisher-Problem. Bulletin of the International Statistical Institute, Vol. *15*, 833–841, 1963.

[20] *Mood, A.M., Graybill, F.A.:* Introduction to the Theory of Statistics. Mc Graw Hill Co., New York 1963.

[21] *Neyman, J.:* Outline of a Theory of Statistical Estimation based on the Classical Theory of Probability. Phil. Trans. Roy. Soc. London, *Ser. A, 236*, 333–380, 1937.

[22] – Fiducial Argument and the Theory of Confidence Intervals. Biometrika *32*, 128–150, 1941.

[23] *Rao, C.R.:* Linear Statistical Inference and Its Applications. J.Wiley, New York 1967.

[24] *Robbins, H., Simons, G., Starr, N.:* A Sequential Analogue of the Behrens-Fisher Problem. Ann. Math. Stat. *38*, 1384–1391, 1967.

[25] *Scheffé, H.:* On Solutions of the Behrens-Fisher Problem, Based on the *t*-Distribution. Ann. Math. Stat. *14*, 35–44, 1943.

[26] – A Note on the Behrens-Fisher Problem. Ann. Math. Stat. *15*, 430–432,1944.

[27] *Schmetterer, L.:* Einführung in die Mathematische Statistik. Springer, Wien 1956.

[28] *Srivastava, M.S.:* On the Asymptotic Theory of Sequential Confidence Intervals. Abstract. Ann. Math. Stat. *37*, 313, 1966.

[29] – On a Sequential Analogue of the Behrens-Fisher Problem. J.R.S.S. *B32*, 144–148, 1970.

[30] *Sukhatme, P.V.:* On Fisher and Behrens Test of Significance for the Difference in Means of two Normal Samples. Sankhyā *4*, 39–48, 1938.

[31] *Tukey, J.W.:* Some Examples with Fiducial Relevance. Ann. Math. Stat. *28*, 687–695, 1957.

[32] – Discussion, Fiducial Probability. Bulletin of the International Statistical Institute, Vol. *15*, p. 921–924, 1963.

[33] *Wallace, D.L.:* Asymptotic Approximations to Distributions. Ann. Math. Stat. *29*, 635–654, 1958.

[34] *Wilks, S.:* On the Problem of two Samples from Normal Populations with Unequal Variances. Ann. Math. Stat. *11*, 475–476, 1940.

[35] – Mathematical Statistics. J.Wiley, New York 1962.

Zusammenfassung

In dieser Arbeit wird nach einer Beschreibung des Grundproblems der Theorie ein Überblick über die wichtigsten statistischen Rückschlussverfahren vermittelt. Diese Ausführungen werden am Beispiel des Behrens-Fisher-Problems illustriert, bei dem es sich darum handelt, auf Grund von Stichprobenwerten Aussagen über die Differenz der Mittelwerte zweier normalverteilter Grundgesamtheiten mit eventuell verschiedenen Streuungen zu machen.

Summary

In this article a survey on some of the most important methods of statistical inference is given after a description of the basic problem of the theory. The applications of these procedures to the Behrens-Fisher problem, which is concerned in making inference about the difference of the means of two normal populations with possibly different variances, is demonstrated.

Résumé

Dans la présente note après une brève description du problème général de la théorie un résumé sur quelques-unes des plus importantes méthodes pour l'inférence statistique est présenté. L'application de ces idées au problème de Behrens-Fisher, qui concerne l'inférence à propos de la différence des moyens de deux populations avec distributions normales, ayant possiblement des variances différentes, est démontrée.