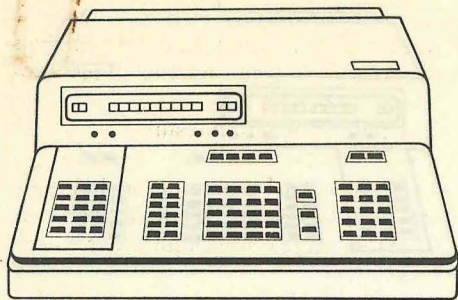




Tisch- und Mikrorechner- anwendungen

Statistischer Test auf programmierbaren Tischrechnern

Claus Kühnel

Zur Verallgemeinerung von Aussagen statistischer Untersuchungen ist für die vorliegenden Stichproben auf die Verteilungsfunktion der Grundgesamtheit, der sie entnommen wurden, zu schließen. Die als statistische Hypothese bezeichnete Annahme einer bestimmten Verteilung der Grundgesamtheit wird mittels statistischer Tests mit dem Ziel geprüft, eine Aussage über Annahme oder Ablehnung der Hypothese zu treffen. Beim Prüfen einer Hypothese H_0 muß stets auf alternative Hypothesen H_1 Bezug genommen werden. Da die

Entscheidung über Annahme oder Ablehnung einer Hypothese aufgrund einer Stichprobe getroffen wird, kann ein Irrtum vorliegen. Die Signifikanzwahrscheinlichkeit α beschreibt die Wahrscheinlichkeit eines Fehlers erster Art (die Nullhypothese wird abgelehnt, obwohl sie richtig ist), sagt jedoch wenig über Fehler zweiter Art aus (die Nullhypothese wird angenommen, obwohl sie falsch ist). Der χ^2 -Test ist eins der Prüfverfahren, welche es gestatten, Verteilungen von Zufallsgrößen ohne Kenntnis ihrer Form zu vergleichen. Für den hier entsprechend der Aufgabenstellung vorzunehmenden Test auf Normalverteilung der Grundgesamtheit einer Stichprobe lautet die Nullhypothese $F(x) = F_0(x)$, die Alternativhypothese $F(x) \neq F_0(x)$. $F(x)$ bezeichnet die unbekannte Verteilungsfunktion der Grundgesamtheit der Stichprobe, $F_0(x)$ die Verteilungsfunktion der Normalverteilung.

Einteilung der Stichprobe

Die Stichprobe soll in k Klassen eingeteilt werden, wodurch sich die Größen

$$y_i = \frac{h_i - n p_i}{n p_i} \quad (1)$$

bilden lassen. Die Anzahl der Stichprobenelemente sei mit n , die der i -ten Klasse entsprechende Häufigkeit mit h_i und die erwartete Wahrscheinlichkeit mit p_i bezeichnet. Als Prüfgröße dient dann

$$u^2 = \sum_{i=1}^k y_i^2 = \sum_{i=1}^k \frac{h_i^2}{n p_i} - n \quad (2)$$

die bei Zutreffen der Nullhypothese asymptotisch χ^2 -verteilt ist mit f Frei-

** DATA INPUT **

Zahl	Häufigkeit
158.00	1.00
161.00	3.00
164.00	4.00
167.00	21.00
170.00	16.00
173.00	29.00
176.00	31.00
179.00	21.00
182.00	18.00
185.00	9.00
188.00	5.00
191.00	2.00
0.00	0.00

Tab. 1 Protokoll der Dateneingabe

heitsgraden. Da bei der Normalverteilung anhand der Stichprobe zwei Parameter geschätzt werden (μ , σ), gilt für die Freiheitsgrade

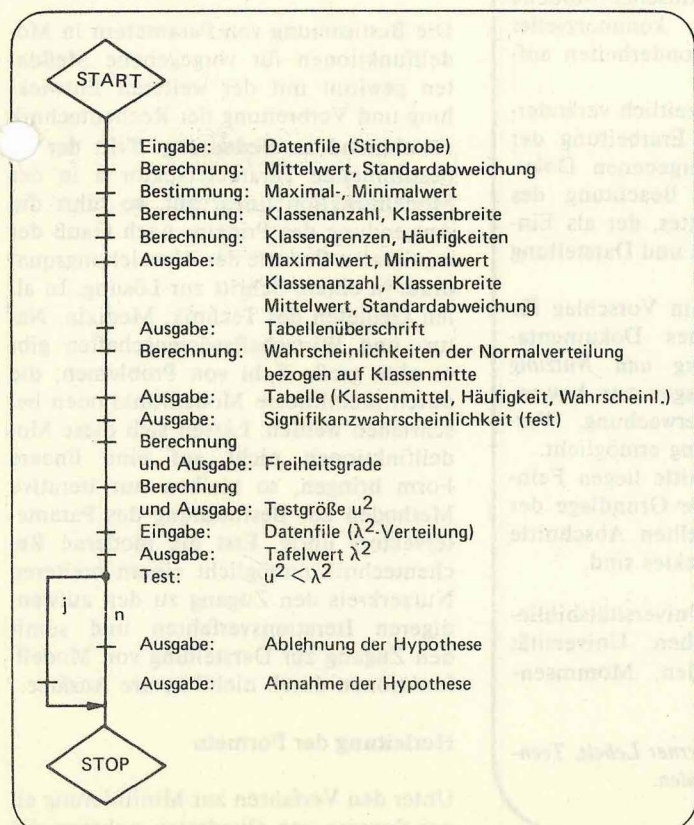
$$f = k - 3 \quad (3)$$

Bei der Festlegung der Anzahl der Klassen (k) soll nach einer Faustregel davon ausgegangen werden, daß die erwarteten Häufigkeiten größer als fünf sind $/1/$. Nach $/2/$ muß aber von Fall zu Fall über die Klassenbildung entschieden werden, da ein starres Festhalten an der Regel die Wirksamkeit des Testes herabsetzen kann.

Testverfahren

Zur effektiven Durchführung des beschriebenen Testverfahrens wurde für einen programmierbaren Tischrechner ein Rechenprogramm entworfen, welches

Abb. Programmablauf zum Test auf Normalverteilung der Grundgesamtheit einer Stichprobe



Tab. 2 Protokollausdruck zum Test auf Normalverteilung der Grundgesamtheit einer Stichprobe

*** TEST AUF NORMALVERTEILUNG EINER STICHPROBE ***

STICHPROBENUMFANG	=	160
MAXIMALWERT	=	191.0000
MINIMALWERT	=	158.0000
KLASSENANZAHL	=	11.0000
KLASSENBREITE	=	3.0000
MITTELWERT	=	175.1188
STANDARDABWEICHUNG	=	6.5034

KLASSEN MITTEL	HAEUFIGKEIT ABSOLUT	WAHRSCHEINLICHKEIT RELATIV	WAHRSCHEINLICHKEIT NORMALVERTEILUNG
159.50	1.000	0.006	0.005
162.50	3.000	0.019	0.017
165.50	4.000	0.025	0.042
168.50	21.000	0.131	0.084
171.50	16.000	0.100	0.136
174.50	29.000	0.181	0.177
177.50	31.000	0.194	0.185
180.50	21.000	0.131	0.155
183.50	18.000	0.113	0.104
186.50	9.000	0.056	0.056
189.50	7.000	0.044	0.024

SIGNIFIKANZWAHRSCHENLICHKEIT = 1 %
FREIHEITSGRADE : 8.00

TESTGROESSE U^2 = 12.24
TAFELWERT χ^2 = 20.10

HYPOTHESE AUF NORMALVERTEILUNG DER GRUNDGESAMTHEIT DER STICHPROBE WIRD NICHT ABGELEHNT

durch das in der Abb. gezeigte Flußdiagramm erklärt ist. Das in /3/ veröffentlichte BASIC-Programm zur Erstellung von Häufigkeitstabellen wurde hinsichtlich der Testdurchführung erweitert, so daß das Nachschlagen in einem Tafelwerk vermieden wird. Nach dem Einlesen der Elemente der Stichprobe von einem auf Magnetband abgespeicherten Datenfile werden die Stichprobe charakterisierende Werte wie Maximalwert, Minimalwert, Mittelwert, Standardabweichung sowie Klassenanzahl und Klassenbreite bestimmt. In einer Tabelle werden die absoluten und relativen Häufigkeiten in einer Klasse und die auf die Klassenmitte bezogene Wahrscheinlichkeit der Normalverteilung gegenübergestellt. Neben dem bereits möglichen Vergleich der beiden Verteilungen kann anhand dieser Tabelle das Problem der Klasseneinteilung begutachtet werden. Im folgenden wird aus der Klassenanzahl die Anzahl der Freiheitsgrade berechnet. Für die verschiedenen Untersuchungsgebiete haben sich typische Signifikanzwahrscheinlichkeiten als sinnvoll herausgestellt, so daß nur ein kleiner Teil der Tafel der χ^2 -Verteilung für den jeweiligen Test interessant sein wird. Im vorliegenden Beispiel wird stets mit einer Signifikanzwahrscheinlichkeit $\alpha = 1\%$ gearbeitet. Dieser Teil der Tafel der χ^2 -Verteilung ist ebenfalls auf einem Datenfile auf Magnetband abgelegt, welches nach der Berechnungsgröße u^2 geladen wird. Durch die Freiheitsgrade ist der zu vergleichende Wert der Tafel festgelegt, und je nach Vergleichsergebnis wird ausgegeben, ob die Hypothese angenommen wird oder nicht. Die Zahlenwerte für die Testgröße u^2 und den Vergleichswert χ^2 sind im Protokoll enthalten.

Zum Test des beschriebenen Verfahrens wurde ein in /4/ veröffentlichtes Beispiel berechnet. Tabelle 1 zeigt das Protokoll der Dateneingabe. Hierfür wurde ein gesondertes Programm zum Laden des Datenfiles aufgeschrieben, welches keine Besonderheiten in sich birgt. Um Eingabe zu sparen, wurden die Stichprobenelemente mit der entsprechenden Häufigkeit multipliziert und anschließend abgespeichert. Tabelle 2 zeigt den Protokollausdruck. Die 160 Elemente umfassende Stichprobe wird in 11 Klassen mit einer Klassenbreite von 3 eingeteilt. Die relativen Häufigkeiten stimmen recht gut mit den zugehörigen Wahrscheinlichkeiten überein, wodurch sich bereits auf Annahme der Hypothese schließen läßt. Entsprechend der 11 Klassen wird mit acht Freiheitsgraden weitergearbeitet. Die Testgröße u^2 liegt mit einem Zahlenwert von 12.24 weit unter dem zugehörigen Wert der χ^2 -Verteilung von 20.10, weshalb die Hypothese angenommen wird.

Mit dem vorliegenden Rechenprogramm kann der beschriebene Test effektiv durchgeführt werden. Die Genauigkeit des Ergebnisses ist gegenüber dem grafi-

schen Verfahren mit Wahrscheinlichkeitspapier weit überlegen. Modifikationen dieses Programmes ermöglichen eine Abarbeitung des beschriebenen Tests auch auf den programmierbaren Kleinstrechnern robotron K 1003, wobei aus gerätetechnischen Gründen auf das relativ komfortable Protokoll verzichtet werden müßte.

Literatur

- /1/ Weber, E.: Grundriß der biologischen Statistik. 8. Auflage. Gustav Fischer Verlag, Jena, 1980
- /2/ Cochran, W. G.: Some methods for strengthening the common chi square test. *Biometrics* 10 (1954) 417-451
- /3/ Strelocke, K.; Hoffmann, P.: Die Dialogprogrammiersprache BASIC. Verlag Technik Berlin, 1981
- /4/ Schindowski, E.; Schürz, O.: Statistische Qualitätskontrolle. 2. Auflage. Verlag Technik Berlin, 1965

wissenschaftliche arbeit

Horst Kasprzik

Dissertation A
Erfahrungen bei der Erarbeitung der Problemdokumentation mathematisch-ökonomischer Modelle und Schlußfolgerungen für die Erarbeitung einer nutzerbezogenen Dokumentation (1984)

Die durch den Verfasser gewonnenen Erfahrungen bei der Erarbeitung der Problemdokumentation nach der Rahmenmethodik der Datenverarbeitungsprojektierung werden an einem Beispielprojekt dargestellt.

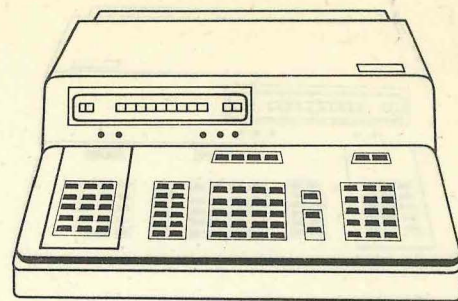
Es wird nachgewiesen, daß Projekte mathematisch-ökonomischer Modelle gegenüber Projekten kommerzieller Aufgaben einige Besonderheiten aufweisen.

Sie verlangen einen zeitlich veränderlichen Ablauf bei der Erarbeitung der dem Nutzer zu übergebenden Dokumentation und die Beachtung des pragmatischen Aspektes, der als Einheit von Form, Inhalt und Darstellung abgeleitet wird.

In der Arbeit wird ein Vorschlag für die Gliederung eines Dokumentationssteiles *Anwendung und Nutzung* entwickelt, der Aussagen zur Anwendung, Nutzung, Überwachung, Wartung und Nachnutzung ermöglicht. Für einzelne Abschnitte liegen Feindispositionen vor, die Grundlage der nachfolgend dargestellten Abschnitte des o. a. Beispielprojektes sind.

Ausleihbar in der Universitätsbibliothek der Technischen Universität Dresden, 8027 Dresden, Mommsenstraße 13.

Ausgewählt von Dr. Werner Lebelt, Technische Universität Dresden.



Tisch- und Mikrorechneranwendungen

Nichtlineare Ausgleichsrechnung für den Mikrorechner MC 80

Johannes Opfermann
Sektion Chemie der
Friedrich-Schiller-Universität Jena

Es wird ein nichtlineares Ausgleichsverfahren beschrieben, das auf einem modifizierten Marquardt/Levenberg-Algorithmus beruht. Das Programm ist primär für den Mikrorechner MC 80 (VEB Elektronik Gera) bestimmt und in der Sprache MC-80-BASIC geschrieben. Es gestattet das Ausgleichen für Modellfunktionen mit maximal 31 Parametern und maximal 255 Wertepaaren $\hat{Y}_i = \hat{Y}_i(x_{1i}, x_{2i}, x_{3i})$. Das Programm ermöglicht ebenfalls die Lösung von gekoppelten nichtlinearen Gleichungen.

Vorbemerkung

Die Bestimmung von Parametern in Modellfunktionen für vorgegebene Meßdaten gewinnt mit der weiteren Entwicklung und Verbreitung der Rechentechnik zunehmend an Bedeutung. Tritt der bestimmende Parametervektor a in der Modellfunktion linear auf, so führt die Anwendung des Prinzips nach Gauß der minimalen Summe der Abweichungsquadrate in einem Schritt zur Lösung. In allen Gebieten der Technik, Medizin, Natur- und Wirtschaftswissenschaften gibt es eine große Zahl von Problemen, die durch nichtlineare Modellfunktionen beschrieben werden. Lassen sich diese Modellfunktionen nicht auf eine lineare Form bringen, so bleiben nur iterative Methoden zur Bestimmung des Parametervektors übrig. Erst die moderne Rechentechnik ermöglicht einem breiteren Nutzerkreis den Zugang zu den aufwendigeren Iterationsverfahren und somit den Zugang zur Darstellung von Modellfunktionen durch nichtlineare Ansätze.

Herleitung der Formeln

Unter den Verfahren zur Minimierung einer Summe von Quadraten nehmen die