für ihre langjährigen Dienste in der gleichen Unternehmung den Dank aussprechen dürfen. Ich erwähne hier nur, dass von 1914 bis 1962 von unserem Verband in den alljährlich stattfindenden Jubilarfeiern insgesamt 35 Veteranen mit 50 Dienstjahren, 1750 Veteranen mit 40 Dienstjahren und weit über 8000 Jubilare mit 25 Dienstjahren geehrt werden konnten. Die sich in diesen Zahlen wiederspiegelnde Verbundenheit mit dem Arbeitgeber ist sicher auch darauf zurückzuführen, dass die sozialen Einrichtungen, wie sie bei den Elektrizitätswerken schon lange bestehen, von Arbeitern und Angestellten geschätzt werden. Eine dieser Institutionen, die Pensionskasse Schweizerischer Elektrizitätswerke (PKE), feiert übrigens dieser Tage das Jubiläum ihres 40jährigen Bestehens. Ich gratuliere hierzu herzlich und bitte ihren hier anwesenden Präsidenten, Herrn Dr. Zihlmann, der Verwaltung der PKE die besten Wünsche unseres Verbandes für eine weitere, gedeihliche Entwicklung der Kasse zu überbringen.

Das Problem des Gewässerschutzes ist zu einer dringenden nationalen Aufgabe geworden. Auf deren Bedeutung ist an der Jahresversammlung des Schweiz. Wasserwirtschaftsverbandes, die vor einer Woche in Zürich stattfand, in einem ausgezeichneten Referat von Herrn Dr. Schneider hingewiesen worden. Wenn auch die Elektrizitätswerke, wie an diesem Vortrag sehr richtig festgestellt wurde, an der Verschmutzung der Gewässer nicht direkt beteiligt sind, so schenken sie doch der Frage des Gewässerschutzes alle Aufmerksamkeit, und sie sind bereit, auch weiterhin an der Lösung dieses Problems mitzuwirken.

An der letztjährigen Generalversammlung in Montreux haben sie einstimmig der Beteiligung der Elektrizitätswerke an der Landesausstellung 1964 zugestimmt und zu deren Finanzierung die Erhebung eines Zusatzbeitrages während der Jahre 1962—1964 beschlossen.

Es ist sehr erfreulich, feststellen zu dürfen, dass unserem inzwischen an die Werke ergangenen Aufruf zur Leistung dieser Beiträge praktisch alle Unternehmungen nachgekommen sind. Die Vorarbeiten für die Elektrizitätsschau, die unter der Leitung von Herrn Dir. Vetsch stehen, sind heute im vollen Gange. Wir werden bei passender Gelegenheit näheres darüber mitteilen können. Allen Herren, die sich trotz starker anderweitiger Inanspruchnahme für das gute Gelingen dieser Ausstellung zur Verfügung gestellt haben, danke ich an dieser Stelle recht herzlich.

Wenn ich schon am danken bin, so möchte ich die Gelegenheit benützen, um auch meinen Kollegen vom Vorstand und den Präsidenten und Mitgliedern unserer Kommissionen für ihre grosse und uneigennützige Arbeit den herzlichsten Dank abzustatten. Unser Dank gilt auch unserem Sekretär und seinen Mitarbeitern. Unser Sekretariat erledigt auf den verschiedensten Gebieten mit einem kleinen Personalbestand eine Fülle von Arbeit, die für uns alle sehr wertvoll ist.

Damit, meine Herren und liebe Kollegen, erkläre ich unsere 71. Generalversammlung als eröffnet.

Anwendung der Methode der mehrfachen Regression für die Analyse von Belastungskurven

Von *Elie Védère*, Paris

(Fortsetzung aus Nr. 19, S. 924, Nr. 20, S. 961 und Nr. 21, S.1009)

Beilage II

Beispiel einer kritischen Prüfung der Wahl der Variablen bei einer Analyse mit mehrfacher Regression

von *Th. Franck*
Oberingenieur, Nordsjaellands Elektricitets og Sporvejs A/S

Diese Beilage zum Bericht der Arbeitsgruppe setzt sich zum Ziel, einige Probleme, die bei der Analyse einer Belastungskurve auftreten, zu beleuchten. Bei der Lösung solcher Probleme dürfen die Voraussetzungen für die Gültigkeit der Methode nicht ausser Acht gelassen werden und es ist manchmal schwer zu sagen, ob diese Voraussetzungen zutreffen.

Das Nichteinhalten der Gültigkeitsbedingungen schliesst das Risiko ein, dass die aus der Analyse gezogenen Schlussfolgerungen zu Fehlern führen, wenn sie für die Berechnung einer Belastungskurve mit einer von der für die Analyse verwendeten stark abweichenden Gruppierung der unabhängigen Variablen verwendet werden.

Um zu zeigen, wie das Problem angepackt werden kann, soll eine dem Statistikausschuss im August 1954 vorgelegte Studie von *Ch. Morel:* «Die Analyse der Belastungskurve. Notiz über eine neue, in der Schweiz entwickelte Methode» herangezogen werden.

Diese Studie betraf die Untersuchung der Belastungskurven von sieben je über eine Hochspannungsleitung belieferten Zonen. Als abhängige Variable wurde der in Ampère gemessene Strom einer Phase gewählt, bei einer Betriebsspannung von 12 kV.

Als unabhängige Variablen dienten die Anschlusswerte folgender drei Apparatekategorien:

x_1 Motoren
x_2 elektrische Kochherde
x_3 übrige thermische Anwendungen

wobei als Leistungseinheit 100 kW angenommen wurde.

Weiter sei festgehalten, dass es sich um eine Wohnzone ohne Industrieabnehmer handelte.

Die abhängige Variable wurde an fünf Tagen einer Woche, von Montag bis Freitag, alle Halbstunden während der Tageshelle gemessen.

Eine Hauptbedingung für die Durchführbarkeit der Analyse ist, dass die Annahme der Linearität zwischen der abhängigen Variablen und den gewählten unabhängigen Variablen auch zutrifft. In diesem Falle scheint diese vernünftig zu sein; lässt sich aber wirklich nichts dagegen einwenden?

Die thermischen Anwendungen x_2 und x_3 haben einen Leistungsfaktor gleich 1, während die Motoren x_1 einen kleineren Leistungsfaktor aufweisen. Der Gesamtstrom ergibt sich durch geometrische (vektorielle) Additionen; er ist also keine streng lineare Funktion der einzelnen Komponenten. Man kann einwenden, dass der hierdurch entstehende Fehler klein ist, wenn das Verhältnis $\frac{x_2 + x_3}{x_1}$ nicht stark variiert.

Nach der Theorie verlangt aber die Analyse nach der Regressionsmethode, dass die unabhängigen Variablen je einen möglichst grossen Variationsbereich aufweisen und voneinander unabhängig sind. Die erwähnte Schwierigkeit könnte dadurch behoben werden, dass für die abhängige Variable als Einheit anstelle des Stromes die Leistung in kW genommen würde.

Wie schon gesagt, sind als unabhängige Variablen die Anschlusswerte gewählt worden; eine weitere Möglichkeit würde darin bestehen, die Zahl der Apparate zu wählen, was vor allem für x_2 vorteilhafter sein könnte. In Dänemark ist festgestellt worden, dass der Anteil der Kochherde an der Belastung von ihrer Grösse praktisch unabhängig ist; dies bedeutet, dass es richtiger wäre, die Anzahl und nicht den Anschlusswert der Kochherde zu verwenden. Es kann eingewendet werden, dass diese Bemerkung nicht zutreffe falls das Verhältnis zwischen kleinen und grossen Kochherden in den sieben Zonen identisch bleibt. Die Zonen können aber verschieden sein, z. B. was die soziale Stellung der Abonnenten betrifft, so dass die durchschnittliche Grösse der elektrischen Kochherde merklich variieren könnte. In diesem Falle wäre es richtiger, als abhängige Variable, die Anzahl und nicht den Anschlusswert der Kochherde zu nehmen.

Zwecks der Regressionsanalyse ist die Schätzung der Regressionskoeffizienten b_1, b_2, b_3. Für die Kochherde gibt b_2 an, welcher Teil des Anschlusswertes im betreffenden Zeitpunkt an der gemessenen Belastung beteiligt ist. Im vorliegenden Falle hat b_2 die Dimension A/kW; ihr Wert kann, wie derjenige von b_1 und b_3, in der Zeit variieren.

Die Verwendbarkeit der Analyse setzt weiter voraus, dass die anteiligen Belastungen der Kochherde um den gleichen Mittelwert b_2 streuen. Man kann sich fragen, ob das immer der Fall sein wird. Die Antwort ist, dass es von den örtlichen Verhältnissen abhängig sein wird.

In Dänemark vorgenommene Untersuchungen zeigen, dass in einigen neuen Wohnvierteln von Kopenhagen eine in bezug auf ihre Klassifizierung nach Alter, Beschäftigung und Ernährungsgewohnheiten abnormal gleichmässige Verteilung der Abnehmer auftritt. Auch die Fahrpläne der öffentlichen Transportmittel führen zu einer starken Konzentration der Mahlzeiten auf eine kurze Zeitspanne.

Es ist augenscheinlich, dass der durchschnittliche Lastanteil einer solchen Kochherdgruppe von demjenigen einer heterogeneren Kochherdgruppe merklich abweichen kann. Sollte vermutlich diese Situation vorliegen, so muss man diese zwei Kochherdkategorien trennen und zwei Regressionskoeffizienten errechnen.

Die vorliegende Stichprobe weist drei unabhängige Variablen auf. Wahrscheinlich ist die Wahl unter der begründeten technischen Annahme erfolgt, dass diese drei Variablen genügen, um die abhängige Variable durch eine lineare Beziehung bis auf einen um Null normalverteilten stochastischen Fehler vollständig zu definieren.

Verwendet man in der Analyse eine unabhängige Variable, die in Wirklichkeit einen wesentlichen Anteil an der abhängigen Variablen hat, so wirkt sich diese Abhängigkeit auf die anderen Regressionskoeffizienten aus, und es kann ein konstantes Glied a auftreten, das von Null wesentlich verschieden ist.

Ein konstantes Glied hat hier kaum eine physikalische Bedeutung und sein Auftreten kann also als Zeichen dafür gedeutet werden, dass die Wahl der unabhängigen Variablen unzweckmässig ist.

Anderseits kann man es riskieren, unnötige unabhängige Variablen einzuführen. Unnötig in dem Sinne, dass sie die Berechnungen, deren Umfang mit der Zahl der Variablen wächst, erschweren, ohne dass dafür die Reststreuung der abhängigen Variablen merklich kleiner wird. Eine Korrelation zwischen zwei unabhängigen Variablen kann zur Folge haben, dass die Streuung der betreffenden Regressionskoeffizienten so gross wird, dass letztere unbrauchbar werden. In einem solchen Falle kann man eine der korrelierenden unabhängigen Variablen weglassen, ohne einen Verlust an Auskunft zu befürchten.

Infolgedessen sollte man immer die Streuung der Regressionskoeffizienten und die Korrelationskoeffizienten zwischen den unabhängigen Variablen berechnen. Die Streuung der Regressionskoeffizienten dient zur Prüfung mit Hilfe der Nullhypothese. Kann die Nullhypothese nicht verworfen werden, so ist die Gültigkeit der Regression zweifelhaft.

In dem von *Ch. Morel* zur Verfügung gestellten Beispiel ist einzig der Regressionskoeffizient b_1 gesichert von Null verschieden, während für b_2 und b_3 kein Grund besteht, die Nullhypothese zu verwerfen. Gleichzeitig beträgt der Korrelationskoeffizient zwischen x_2 und x_3 0,435. Diese Korrelation kann zur Folge haben, dass die Bestimmung von b_2 und b_3 ungenügend ist. Die Korrelation zwischen x_2 und x_3 kann hier nicht überraschen, weil der grösste Teil von x_3 wahrscheinlich aus Boilern in den Haushaltungen besteht, so dass in Wirklichkeit x_2 und x_3 nahezu proportional der Anzahl Haushaltungen sind. Dies führt konsequenterweise dazu, auf die Ermittlung des Anteils jeder der beiden Variablen an der Belastung zu verzichten. Man kann dies übrigens a priori tun, da die zur Verfügung stehenden Beobachtungsreihen nicht gestatten, eine detailierte Analyse vorzunehmen. Grundsätzlich kann die Berechnung durchgeführt werden, aber das Ergebnis ist wie gesagt mit einer starken Unsicherheit und infolgedessen mit einer grossen Fehlermöglichkeit behaftet.

Wie bereits erwähnt, muss sich eine Regressionsanalyse auf unabhängige Beobachtungsreihen stützen können. Es wäre also ein grundsätzlicher Fehler, im Beispiel die 7 Beobachtungen in einem bestimmten Zeitpunkt der 5 auf einander folgenden Werktage als 35 Beobachtungen zu betrachten. In Wirklichkeit handelt es sich um 5 Wiederholungen der Beobachtungen für einen Satz unabhängiger Variablen.

Dieser Umstand ist in den Berechnungen zu berücksichtigen. Die Rechnung hat für jeden Tag mit je 7 Beobachtungen oder aber für die Woche mit je 7 Durchschnittswerten aus den 5 Tagesbeobachtungen pro Kurvenpunkt zu erfolgen. Die letzte Art der Berechnung ist die einfachere und man könnte versucht sein, sie vorzuziehen. Überlegen wir aber noch kurz die Bedingungen für ihre Anwendung.

Wenn der Durchschnitt der Beobachtungen im gleichen Zeitpunkt während 5 verschiedener Tage eine Bedeutung haben soll, so muss man annehmen, dass von einem Tag zum andern keine systematische Änderung auftritt. Mit andern Worten, müssen die wahren Werte der Regressionskoeffizienten für alle Tage die gleichen sein, und die rechnerisch ermittelten Schätzungen dieser Werte dürfen nur zufällig davon abweichen. So müssen auch die Streuungen bis auf einen zufälligen Fehler identisch sein.

Ob diese Bedingungen erfüllt sind, kann nur die erste Berechnungsart zeigen. Wird sie für die Angaben des Beispiels angewendet, so erhält man 5 Reihen von Streuungen, welche, wie im Bericht der Arbeitsgruppe erwähnt, nach der Methode von Bartlett geprüft werden können. Es stellt sich dann heraus, dass die Streuung am Freitag sehr gross ist; so ist es unmöglich, die Hypothese aufrecht zu erhalten, alle Streuungen hätten den gleichen wahren Wert. Dies gilt unabhängig davon, ob die Regressionsanalyse mit einer, zwei oder drei unabhängigen Variablen durchgeführt wird.

Da die Messungen mehrere Jahre zurückliegen, ist es nicht mehr möglich herauszufinden, aus welchen Gründen die Streuung vom Freitag so stark divergiert. Es gibt dafür verschiedene Möglichkeiten: z. B. können die meteorologischen Verhältnisse abnormal gewesen sein.

Aus dem Vorangehenden ergibt sich auf alle Fälle, dass es nicht zulässig wäre, die zweite Berechnungsart anzuwenden, weil sonst Durchschnitte aus Werten gebildet würden, die systematisch verschiedenen Beobachtungsreihen angehören.

Dies sind meine Bemerkungen zu den von *Ch. Morel* mitgeteilten Erfahrungswerten. Wie bereits in der Einleitung erwähnt, war meine Absicht zu zeigen, dass die Regressionsanalyse viele Fehlermöglichkeiten einschliesst, weil immer wieder Annahmen getroffen werden müssen, von denen es selten sicher ist, dass sie zutreffen. Ich habe bewusst darauf verzichtet, Zahlen anzugeben, die meines Erachtens ermüdend sind. Für die Ausführung der Berechnungen verweise ich auf den Bericht. Mir lag es vor allem daran, einige der Probleme aufzuzeigen, die bei Arbeiten auf diesem interessanten Gebiet auftreten können. *D:Pf.*

Anhang III

Beispiel einer mathematischen Überprüfung der Gültigkeit der Ergebnisse einer Regressions-Analyse.

von *Georg Ott*, Berlin
c/o Berliner Kraft- und Licht (Bewag) -AG, Berlin.

Die nachfolgende Studie ist ein Beispiel der mathematischen Prüfung der Gültigkeitskriterien der mehrfachen Regressionsanalyse. Sie betrifft eine Stichproben-Analyse[1]) mit drei unabhängigen Variablen installierte Leistung (Licht, Kraft, Wärme) aus einer Gesamtheit von industriellen Abonnenten; die abhängige Variable ist der Verbrauch. Die betreffende Analyse umfasste insgesamt 110 Gleichungen (eine pro Abonnent).

Betrachten wir zunächst das Produkt der Matrizen $b_j = C_{ij} B$[2]), welches aus der Analyse hervorgeht und das den Energieverbrauch und die installierte Leistung (Licht, Kraft, Wärme) miteinander verbindet:

$$\begin{vmatrix} b_1 \\ b_2 \\ b_3 \end{vmatrix} = \begin{vmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{31} & C_{33} \end{vmatrix} \cdot \begin{vmatrix} \Sigma x_1 y \\ \Sigma x_2 y \\ \Sigma x_3 y \end{vmatrix} \qquad (1)$$

und das Ergebnis in Zahlen:

$$\begin{vmatrix} b_1 \\ b_2 \\ b_3 \end{vmatrix} = \begin{vmatrix} 0{,}002\,503\,58 & -\,0{,}000\,293\,72 & -\,0{,}000\,060\,17 \\ -\,0{,}000\,293\,72 & 0{,}000\,044\,74 & -\,0{,}000\,018\,55 \\ -\,0{,}000\,060\,00 & -\,0{,}000\,018\,58 & 0{,}001\,072\,28 \end{vmatrix} \cdot \begin{vmatrix} 5\,401\,444 \\ 41\,372\,886 \\ 1\,898\,900 \end{vmatrix} \qquad (1.1)$$

Mit Hilfe des Gleichungsystems (1.1) erhält man als jährlichen Energieverbrauch pro kW installierter Leistung[3])

b_1 = 1257 kWh pro Jahr für die Beleuchtung
b_2 = 229 kWh pro Jahr für motorische Zwecke
b_3 = 944 kWh pro Jahr für Wärmeanwendungen

Um die Resultate zu prüfen, können mathematisch-statistische Verfahren angewendet werden.

Zuerst muss überprüft werden, ob der Zusammenhang zwischen der gesuchten Grösse Y und den unabhängigen Variablen x_i zufällig oder gesichert ist. Das Bestimmtheitsmass R^2, auch durch die Buchstaben B oder Σ ausgedrückt, antwortet auf diese Frage. Es gibt an, welcher Anteil der Varianz Y auf die Veränderung der unabhängigen Variablen x_i zurückzuführen ist. Das Bestimmtheitsmass

$$R^2 = \frac{1}{\Sigma y^2} (b_1 \Sigma x_1 y + b_2 \Sigma x_2 y + b_3 \Sigma x_3 y) \qquad (2)$$

zeigt, wie eng sich die einzelnen Werte y_i um die Regressionsgerade scharen. Im vorliegenden Fall erhält man:

$$R^2 = \frac{1}{19\,348 \cdot 10^6} (1\,257 \cdot 5\,401\,444 + 229 \cdot 41\,372\,886 + 944 \cdot 1\,898\,900) = 0{,}93 \qquad (2.1)$$

Aus (2.1) geht hervor, dass 93% der Streuung von y_i — d. h. des Energieverbrauches — auf die Streuung der installierten Leistung (x_i) zurückzuführen sind.

---
[1]) Anhang I, Studie Nr. 3
[2]) B ist Spalten-Vektor des Produktes $x_i\, y$
[3]) oder als jährliche Gebrauchsdauer

Nun muss überprüft werden, ob das Bestimmtheitsmass R^2 (das sich auf die Stichprobe bezieht) für die Gesamtheit der Gruppe wesentlich von null abweicht. Ist R^2 nahezu null, so bedeutet dies, dass der Energieverbrauch nicht von der installierten Leistung abhängt. Wenn der Wert des Bestimmtheitsmasses R^2 im «Zufälligkeitsbereich» der F-Verteilung ($S < 95\%$) liegt, muss diese Hypothese angenommen werden. Befindet er sich in einem Bereich, der werder zufällig noch nahezu gesichert ist ($95\% < S < 99\%$), so wird man sie weder annehmen noch verwerfen, sondern eine andere Stichprobe auswählen.

Wenn R^2 im Bereich «beinahe gesichert» ($S > 99\%$) liegt, muss die Annahme, dass der Verbrauch nicht von der installierten Leistung abhängt, fallen gelassen und die Hypothese, dass ein Zusammenhang zwischen diesen Grössen wahrscheinlich besteht, angenommen werden. Wird F als Funktion von R^2 ausgedrückt, so erhält man:

$$F = \frac{R^2 (n-p)}{p(1-R^2)} \tag{3}$$

In diesem Ausdruck sind

$$\nu_1 = p = 3 \quad \text{und} \quad \nu_2 = n - p = 63 - 3 = 60$$

die Zahl der Freiheitsgrade [4]).

Die Bestimmtheitsmasse R^2, die den Wahrscheinlichkeiten $P = 0{,}05$ [5]) und $P = 0{,}01$ entsprechen, können mit Hilfe der Formel (3) errechnet oder Tafeln entnommen werden.

Um den Koeffizienten R^2 berechnen zu können, muss die Gleichung (3) umgewandelt werden:

$$R^2 = \frac{pF}{n - p(1-F)} \tag{3.1}$$

Für die Wahrscheinlichkeit $P = 0{,}05$ erhält man folgendes Resultat:

$$R^2{}_{0,05} = \frac{3 \cdot 2{,}76}{63 - 3(1 - 2{,}76)} = 0{,}14 \tag{3.11}$$

Desgleichen erhält man für $P = 0{,}01$, $R^2{}_{0,01} = 0{,}24$.

Das Bestimmtheitsmass $R^2 = 0{,}93$ ist also stark gesichert.

Während der Korrelationskoeffizient R, wie auch R^2 angeben, in welchem Masse die einzelnen Werte y_i der abhängigen Variablen mit der gewählten Funktion $y = f(x)$ [6]) übereinstimmen, kann mit Hilfe der Methode der Streuungsanalyse untersucht werden, ob zwischen den abhängigen Variablen y und den unabhängigen Variablen x_i ein funktioneller Zusammenhang besteht.

Zu diesem Zwecke prüft man, in welchem Masse die gesamte Streuung der abhängigen Variablen y durch die Regression Y verringert wird.

Verringerung durch die Regression:

$$s_1{}^2 = \frac{\Sigma y^2 - S}{p} = \frac{b_1 \Sigma x_1 y + b_2 \Sigma x_2 y + b_3 \Sigma x_3 y}{p} \tag{4}$$

Streuung um die Regressionsgrade:

$$s_2{}^2 = \frac{S}{n-p} = \frac{\Sigma y^2 - b_1 \Sigma x_1 y - b_2 \Sigma x_2 y - b_3 \Sigma x_3 y}{n-p} \tag{5}$$

Die Zahl der Freiheitsgrade beträgt dabei wiederum:

$$\nu_1 = p = 3 \quad \text{und} \quad \nu_2 = n - p = 63 - 3 = 60$$

Nun errechnen wir den Quotienten

$$\frac{s_1{}^2}{s_2{}^2} = \frac{6\,024\,247\,913}{22\,387\,959} = 269{,}8 \tag{6}$$

und vergleichen ihn mit dem Wert von F für die Wahrscheinlichkeitspunkte:

$$P(3/60; 0{,}05) = 2{,}76 \quad \text{und} \quad P(3/60; 0{,}01) = 4{,}13$$

Da der Streuungsquotient mehr als 43,7 mal grösser als der Wert F für $P = 0{,}001$ (6,17) ist, dürfte zwischen dem Verbrauch und der installierten Leistung ein gesicherter Zusammenhang bestehen.

Für die Prüfung haben wir bis jetzt die Regression als Ganzes betrachtet. Die Gültigkeit der einzelnen Regressionskoeffizienten b_j ist noch nicht bekannt. Um sie zu prüfen, muss man die Regressionskoeffizienten b_j der Stichprobe mit denjenigen der Gesamtheit der Gruppe (β_j) vergleichen.

Im allgemeinen gilt:

$$t_j = \frac{b_j - \beta_j}{s\sqrt{c_{jj}}} \tag{7}$$

s ist in diesem Ausdruck wiederum die mittlere quadratische Abweichung. Nach der Gleichung (6) erhält man:

$$s = \sqrt{s_2{}^2} = \sqrt{22\,387\,959} = 4\,731{,}6 \tag{5.1}$$

Wenn die Funktion $Y = f(x)$ ein konstantes Glied y enthält, wird die Zahl der Freiheitsgrade $\nu_2 = n - p$ um eine Einheit verringert und beträgt $\nu_2 - n - p - 1$.

Berechnen wir nun den Wert t für die Regressionskoeffizienten b_1. Zu diesem Zweck nimmt man an, dass die Regressionskoeffizienten β_j der Gesamtheit der Gruppe gleich null sind. Anders ausgedrückt: Man unterstellt, dass der Verbrauch nicht von der installierten Leistung abhängt. Man erhält dann folgenden t-Wert für die Regressionskoeffizienten b_1:

$$t_1 = \frac{b_1 - \beta_1}{s\sqrt{c_{11}}} = \frac{1\,257 - 0}{4\,731{,}6 \cdot \sqrt{0{,}002\,503\,58}} = 5{,}31 \tag{7.1}$$

Auf gleiche Weise erhält man:

$$t_2 = 7{,}46 \quad \text{und} \quad t_3 = 6{,}09$$

In den andern Ausdrücken führt man für t_j folgende Grössen ein:

$$C_{22} \text{ für } t_2 \quad \text{und} \quad C_{33} \text{ für } t_3$$

Man vergleicht diese empirischen t-Werte mit denjenigen der t-Verteilung für den Freiheitsgrad:

$$\nu_2 = n - p = 63 - 3 = 60$$

und für die Wahrscheinlichkeiten

$$P = 0{,}01 \quad \text{und} \quad P = 0{,}001$$

Den klassischen Tabellen entnehmen wir folgende Werte:

$$t_{0,01} = 2{,}660 \quad \text{und} \quad t_{0,001} = 3{,}460$$

Da die für alle Regressionskoeffizienten b_j gefundenen t-Werte grösser sind als die theoretischen Werte für $P = 0{,}001$ ($S = 99{,}9\%$), ist damit zu rechnen, dass

[4]) In Anbetracht der Tatsache, dass nicht alle Gleichungen unabhängige Variablen enthalten, wählen wir für die Berechnung der Grösse n das geometrische Mittel:

$$n_{\text{total}} = \sqrt[3]{100 \cdot 96 \cdot 24} = 63{,}3$$

[5]) $P = 0{,}05$ entspricht einer statistischen Wahrscheinlichkeit $S = 95\%$, $P = 0{,}01$ einer solchen von 99%.

[6]) Mit dem Ausdruck «Konkordanz» soll zum Ausdruck gebracht werden, dass sich die abhängige Variable im gleichen Sinn wie die unabhängige Variable ändert.

die Regressionskoeffizienten b_j beträchtlich von denjenigen der Gesamtheit der Gruppe ($\beta_j = 0$) abweichen.

Der Ausdruck (t) kann umgeformt werden und man erhält so den Vertrauensbereich für die Regressionskoeffizienten:

$$b_j \pm ts\sqrt{C_{jj}} \tag{8}$$

Hier muss die Grösse β_j nicht berücksichtigt werden, da sie gleich null ist.

Für $t_{0,05}$ findet man für den Regressionskoeffizienten b_1, der den Verbrauch mit der installierten Leistung (Licht) in Zusammenhang bringt, folgenden Vertrauensbereich:

$$b_1 \pm 2 \times 4\,731{,}6\sqrt{0{,}002\,503\,58} = 1\,257 \pm 473\,\text{kWh/Jahr}$$

Der Verbrauch von 95% der Konsumenten liegt also im Bereich

$$100\% \pm 38\%$$

Auf gleiche Weise erhält man für die Grösse b_2, welche den Verbrauch mit der installierten motorischen Leistung (Kraft) in Zusammenhang bringt, folgenden Vertrauensbereich:

$$229 \pm 63\,\text{kWh/Jahr oder } 100\% \pm 28\%$$

Der Vertrauensbereich für b_3 (Zusammenhang zwischen Verbrauch und installierter thermischer Leistung) beträgt:

$$944 \pm 310\,\text{kWh/Jahr oder } 100\% \pm 33\%$$

In Deutschland verwendet man anstelle des Bestimmtheitsmasses R^2 gelegentlich den Variationskoeffizienten A. Er gibt an, um wieviel die errechneten Werte der gesuchten Grösse Y relativ von den gemessenen Werten y_i abweichen. Die Grösse A geht aus folgendem Ausdruck hervor:

$$A = \frac{\sum \frac{Y_i}{y_i}}{n} - 1 \tag{9}$$

Multipliziert man das Resultat mit 100, so erhält man den Variationskoeffizienten in %. Die Grösse A ist ihrem Wesen nach geeigneter und signifikanter als R^2. *D: Kr.*

Literatur

1. **Allgemeine Studien über die Methode der mehrfachen Regression und ihrer Anwendung für die Analyse der Belastungskurven und der Verbrauchszahlen.**

Morel Ch.: Die modernen statistischen Methoden im Dienste der Elektrizitätswerke. Bull. SEV Bd. 45 (1954), Nr. 16, S. 667...677 und Nr. 17, S. 710...714.

Morel Ch.: Die Zerlegung der Belastungskurven mit Hilfe der mehrfachen Regression. Bul SEV Bd. 46 (1955), Nr. 11, S. 521...527.

Ott G.: Last- und Raumanalyse durch Stichproben. Elektr.-Wirtsch. Bd. 56 (1957), Nr. 15, S. 524...530 und Nr. 16, S. 556, 557.

Ott G.: Wirksames Stichprobenverfahren bei der Lastanalyse der Industrie. Elektr.-Wirtsch. Bd. 57 (1958), Nr. 1, S. 12...16 und Nr. 2, S. 33...37.

Puromäki A.: Die Analyse der Belastungskurven nach der Methode der mehrfachen Regression. Voima ja Valo (Kraft und Licht) (1959).

Puromäki A.: Un essai d'analyse des courbes de charge par la méthode de régression multiple. Kongress UNIPEDE 1958, Bericht VIII A, Beilage B.

Schiller P.: Fortschritte bei der Analyse von Belastungskurven. Electrical Review 1955, S. 1087.

Schiller P.: Méthodes d'analyse de la charge employées en Grande-Bretagne. Kongress UNIPEDE 1958, Bericht VIII-A1.

Védère E.: Méthode graphique pour l'analyse d'une courbe de charge en deux composantes. Kongress UNIPEDE 1958, Bericht VIII-A, Beilage C.

2. **Berichte über die nach der Methode der mehrfachen Regression gewonnenen Resultate.**

Davies M.: Beziehung zwischen den meteorologischen Verhältnissen und der Belastung des T.E.E.-Netzes. Monograph Nr. 3145 (1958).

Davies M.: Verbundbetrieb des «Grids» und die meteorologischen Verhältnisse Weathert. 15 (1960), nº 1, S. 3.

Schiller P. et Johnson N. L.: Beziehung zwischen der Tageshelle und der Belastung des E.R.A.-Netzes. Report K/T 115 (1955).

Electricity Council: (früher *Central Electricity Authority*): Utilisation Research Report:
- Eine neue Analyse der Resultate einer von der E.R.A. durchgeführten Stichprobe aus Haushaltabnehmern. 1950 Nr. 4.
- Untersuchung nach der Stichprobe 1955 aus den Haushaltabnehmern. 1958 Nr. 7.
- Untersuchung nach einer Stichprobe aus Gewerbeabnehmern. 1958 Nr. 8.

Belastungs- und Verbrauchs-Charakteristiken eines grossen Nachkriegswohnviertels: Report VRC/DS/R 24, 1955.

Verbandsmitteilungen

Protokoll
der 71. (ordentl.) Generalversammlung des VSE, Freitag, den 28. September 1962, 10.30 Uhr, in der Rathauslaube in Schaffhausen

Der **Vorsitzende**, P. Payot, Delegierter des Verwaltungsrates und technischer Direktor der Société Romande d'Electricité, Clarens, heisst die Gäste und Kollegen in Schaffhausen willkommen.

Der Wortlaut seiner Eröffnungsansprache findet sich in der vorliegenden Nummer des SEV-Bulletin (S. 1147...1149).

Zur Tagesordnung übergehend erklärt der Vorsitzende die 71. Generalversammlung des VSE als eröffnet. Er stellt fest, dass die Generalversammlung statutengemäss rechtzeitig einberufen worden ist, und zwar durch Publikation im Bulletin SEV, «Seiten des VSE», Nr. 17, vom 25. August 1962, in welcher Nummer auch die Traktandenliste, die Anträge des Vorstandes, die Rechnung und die Bilanz des VSE, die Rechnung, Bilanz und der Geschäftsbericht der Einkaufsabteilung veröffentlicht wurden. Der Bericht des Vorstandes des VSE an die Generalversammlung über das Geschäftsjahr 1961 ist im Bulletin SEV, «Seiten des VSE», Nr. 18, vom 8. September 1962, erschienen. Der Bericht der Rechnungsrevisoren befindet sich im Bulletin des SEV, «Seiten des VSE», Nr. 19, vom 22. September 1962.

Zur Traktandenliste werden keine Bemerkungen gemacht und zu keinem Traktandum wird Geheimabstimmung verlangt.

Trakt. 1:

Wahl zweier Stimmenzähler und des Protokollführers

Als Protokollführer der Generalversammlung wird Herr *Ch. Morel*, vom Sekretariat VSE, bestimmt, und als Stimmenzähler werden die Herren *E. Schneider*, Elektrizitätswerk Bischofszell, *P. Troller*, Elektrizitätswerk Basel, und *E. Schilling*, Elektrizitätswerk Biel, gewählt.

Trakt. 2:

Protokoll der 70. Generalversammlung vom 30. September 1961 in Montreux

Das Protokoll der 70. Generalversammlung vom 30. September 1961 in Montreux (veröffentlicht im Bull. SEV, «Seiten des VSE», Nr. 23 vom 18. November 1961) wird *genehmigt.*

Trakt. 3:

Bericht des Vorstandes und der Einkaufsabteilung über das Geschäftsjahr 1961

Der Bericht des Vorstandes und derjenige der Einkaufsabteilung über das Geschäftsjahr 1961 werden *genehmigt.*

Trakt. 4:

Verbandsrechnung über das Geschäftsjahr 1961; Rechnung der Einkaufsabteilung über das Geschäftsjahr 1961

Der **Vorsitzende** stellt fest, dass die Rechnung des Verbandes günstig abschliesst, so dass es wiederum möglich ist, eine kleine