

UNTERSUCHUNG DER ZUSAMMENHÄNGE IN STOCHASTISCHEN SYSTEMEN BEI DER ANALYSE VON VERKEHRSUNFALLURSACHEN

Von

G. GYULAI

Institut für Verkehrstechnik und Organisation, Technische Universität, Budapest

Eingegangen am 29. Juni 1976

Vorgelegt von Prof. Dr. I. TURÁNYI

Zweck des Verkehrs ist eine *Ortsbewegung*, die *Zeitaufwand* erfordert; daher stellt die Erhöhung der Reise- und Fahrgeschwindigkeiten eine vorrangige qualitative Forderung dar. Bei der hochgradigen Motorisierung und dem Urbanisationsvorgang unserer Tage sind die Fahrbahnen überfüllt und es darf nicht gestattet werden, daß sich — sei es durch die hohen Geschwindigkeiten, sei es durch die Überfülltheit der Verkehrsstraßen — Zahl und Schwere der Unfälle der wachsenden Zahl der Gefahrsituationen verhältnismäßig zunehmen.

Der Unfallverhütung kommt heute neben dem Schutz von Leben und Eigentum auch in der Verminderung von *Verkehrshindernissen* eine entscheidende Bedeutung zu, u. zw. vor allem im Massenverkehr, weil Fahrer und Fahrzeuge stets den ganzen Tag über unterwegs sind, weil die Fahrer für die Unverletztheit großer Fahrgästemassen verantwortlich gemacht werden und die Fahrzeuge größerer Masse im allgemeinen auch längere Bremswege haben.

Um Verkehrssicherheit zu gewährleisten, müssen die mit dieser zusammenhängenden Erscheinungen in Raum und Zeit vielfach beobachtet und gemessen, in geeigneter Weise zahlenmäßig gekennzeichnet *und analysiert*, die *Ursachen und Zusammenhänge* geklärt werden, damit in erster Reihe das Vorkommen von Unfällen *verhütet* bzw. die Folgen der eingetretenen Unfälle rasch beseitigt werden. Die Aufdeckung von Unfallursachen zeigt den Weg zur Lösung, Art und Ausmaß *ihrer kausalen Wirkung* deuten meistens auf die Ebene hin, auf der die notwendigen Maßnahmen zu treffen sind.

Das System ist eine Gesamtheit mit bestimmten Beziehungen. Aus der Sicht der Systemtheorie als *Wissenszweig* stellt das *Verkehrssystem* eine aus der Gesamtheit von Menschen, Mitteln und Einrichtungen womöglich optimal herausgebildete Apparatur dar, die geschaffen wird, um die Verkehrsansprüche der Gesellschaft zu befriedigen. Was die Teilsysteme desselben anbelangt, wird die Verkehrssicherheit durch

- den Menschen
- das zum Verkehr in primärer Beziehung stehende Fahrzeug

- dessen Fahrbahn und die Umweltelemente sowie
- durch die Verkehrsleitung

beeinflußt.

Für diese komplexe Gemeinschaft sind zweckdienliche Bedingungen zu gewährleisten, wobei unterstrichen wird, daß

1. die Unfälle letzten Endes *überwiegend auf menschliches Versagen* zurückzuführen sind, u. zw. in weitem Umfang — direkt — auf die Kenntnisse und das Verhalten der aktiven oder passiven Verkehrsteilnehmer;

2. in der Unfallverhütung auch der *technische Dienst* bei der Wahl der Vielzahl der *Straßenparameter* und -kennwerte, bei den technischen Mitteln der Verkehrsleitung sowie in der gründlichen und regelmäßigen Instandhaltung der Fahrzeuge: im allgemeinen in der Erkenntnis der Bedeutung der *Infrastruktur* eine große Rolle spielt;

3. der Verkehrsregelung (StVO) eine entscheidende Bedeutung zukommt. Alle diese Aufgaben lassen sich durch richtige *Analyse und ausgedehnte Klärung der Ursachen* auf der Basis einer ausführlichen Datenlieferung lösen.

In einem anderen Aspekt der Systemtheorie wird das System — nach dem mathematischen Gesichtspunkt von Zadeh — als eine Sammlung mathematischer Modelle aufgefaßt, in der für die Beschreibung der Beziehungen deterministische und stochastische Zeichenbeschreibungen und Funktionen angewandt werden. Im Allgemeinfall stellen die Ein- und Ausgangsgrößen des mathematischen Systems und in diesem der Informationsstrom einen stochastischen Prozeß dar, durch den in diesem Falle in der Input-Output-Datenpaarmenge des Unfallgeschehens die Wechselwirkungen geschaltet werden. In dieser mathematischen Struktur sollen zuerst die Annahme der hinreichenden und entsprechenden Input-Veränderlichen, dann das Modellieren der Input-Output-Beziehungen behandelt werden.

I

Jede Unfallanalyse beruht gerade im Interesse der späteren ausführlichen Klärung der Unfallursache auf umsichtiger Tatbestandsaufnahme auf dem Unfallsort — was auch wegen des Schadenersatzverfahrens notwendig ist — sowie auf einer richtigen Auswertungstätigkeit. Vor der eigentlichen Behandlung des Themas möchten wir einführend in diesem Zusammenhang einige Bemerkungen und Vorschläge machen.

1. Die Richtigkeit der Schlußfolgerungen wird durch die Brauchbarkeit der Grunddaten entscheidend beeinflusst; bei der Erhebung am Unfallsort müssen weiterhin auch Daten aufgezeichnet werden, aus denen auf *verkehrs-technische Ursachen* geschlossen und erreicht werden kann, daß von den Entwürfsinstituten *zugleich* Unfallverhütungsmaßnahmen *eingepplant werden*.

2—3. In den Städten Ungarns werden die Daten des Unfallgeschehens

verschiedenartig erfaßt; es wäre eine *internationale* Nomenklatur erforderlich. Es gibt auch Städte, wo zwar viele Daten erfaßt werden, die *Analyse jedoch beschränkt ist*. Die Einheitlichkeit ist auch eine Voraussetzung für die *Codierung* der Daten, da die Datenverarbeitung und mannigfaltige Gruppierung typische rechentechnische Aufgaben sind, besonders mit Rücksicht auf die Aktualität der Maßnahmen, aber auch weil sie die Voraussetzung des Aufstellens einer *Datenbank* bilden. Das Unfallgeschehen soll durch eine *Rechenanlage* registriert werden, die wiederholte Wegstrecken ausschreibt und die Wirksamkeit der Eingriffe auswertet, was die ganze Verarbeitung erleichtert und beschleunigt.

4. Die gute Datenerhebung ist mit der womöglich raschen Wiederherstellung des gestörten Verkehrs abzustimmen. Eine *Beschleunigung* der Datenerhebungen im Interesse des Verkehrs läßt sich in erster Reihe durch Abkürzung der Besichtigungsdauer — was sich durch Dezentralisation erreichen läßt — und durch Vereinfachung der Schadenaufnahme erreichen; statt übertriebener Aufnahme überflüssiger Einzelheiten sollen »diktierte« Durchschnittszahlen auch verwendet werden.

5. Mit Rücksicht auf die Einordnung nach mehreren Gesichtspunkten, ist es zweckmäßig, die verschiedenartigen Daten in sog. *kombinative statistische Tafeln* zusammenzufassen. In den Spalten und Zeilen derselben werden Ausgang (Schwere) und Art (z. B. Zusammenstoß oder Entgleisung) des Unfalls miteinander und mit Verkehrsdichte, Straßenparametern, Fahrzeugart und Verkehrsträger, mit Zeitpunkt und Witterungsverhältnissen sowie mit Lebensalter, Qualifizierung des Fahrers in Beziehung gesetzt. Von Dr. Rozgonyi wird vorgeschlagen, in einer *dreidimensionalen Matrix* senkrecht auf die Achse mit der Anzahl der Unfälle den Unfallausgang (Verletzte, Schaden) und über dieser Ebene die Ursachen aufzutragen. Der Unfallausgang kann auch der Summierung und dadurch der Vergleichbarkeit halber durch Gewichtsfaktoren gekennzeichnet werden, wie z. B. mit folgenden Multiplikatoren: leichte Verletzung = 1, schwere Verletzung = 15, tödliche = 30. Auch Verkehrsmenge (Kfz/Tag) sowie Straßenparameter und -kennwerte können im Vergleich zueinander *abgewogen* werden.

6. Ein zeitlicher und gegenseitiger Vergleich ermöglicht, Gesetzmäßigkeiten und Zusammenhänge zu erschließen, in Form von — durch den Vergleich der Daten je nach dem Charakter der Unfälle in geeigneter Weise — konstruierten Kennwerten. Die Erfahrung zeigt, daß im allgemeinen wenig *spezifische Kennwerte* (Verhältniszahlen) gebildet werden: von dem Budapester Verkehrsbetrieb wird nur auf 1 Million Wagenkilometer Bezug genommen; von Sándor Koller wird in seiner Arbeit »Stadtverkehrstechnik« empfohlen, sich auf den Kraftfahrzeugbestand, im Lehrfach »Verkehrsbetriebslehre« auf 1000 Fahrzeuge, 100 000 Einwohner zu beziehen. Wir meinen, daß der Bereich des Vergleichs *mit der Leistung* erweitert werden sollte, weil es Unfallarten gibt, die

neben dem Fahrzeugbestand

- mit dem Fahrzeugkilometer
- mit dem Platzkilometer (Wegen der Wagengrößen)
- mit dem Zugkilometer (z. B. bei Zusammenstößen) zu teilen,
- auf Verkehrsdauer, Wagenstunden,
- Fahrgastzahl, ja sogar auf Einwohnerzahlen zu beziehen sind.

7. Es ist ein schwerer Fehler, daß die Verteilung der Zeitdauer der durch Unfälle verursachten Verkehrshindernisse (auf Schiene und Straße) nicht untersucht wird, obwohl für die Bildung eines guten und einhaltbaren *Fahrplans* und für die Planung des Reservefahrzeugbedarfs die Kenntnis der Zeitdauer je Ereignis und von deren Korrelation unerläßlich ist. Glücklicherweise sind Aufzeichnungen über die Zeitpunkte der Meldungen von Unfällen und der Wiederaufnahme des Verkehrs vorhanden, die eine mathematische Analyse ermöglichen.

8. Unter den Unfallursachen sind solche, die in erster Betrachtung die Erscheinung unmittelbar zu erklären scheinen, hinter denen sich jedoch bei näherer Prüfung Ursachen verbergen, die als sekundär bezeichnet werden dürften, auf die sich — als *einige gemeinsame Grundlagen* — mehrere Unfälle zurückführen lassen. Das sind die »Faktoren«, auf deren Berechnung in Abschnitt III näher eingegangen werden soll. Die Feststellung der *direkten* Verantwortung, laut der der Fahrer das Fahrzeug »nicht den Straßenverhältnissen entsprechend« fuhr, ist nicht immer befriedigend.

Durch diese mannigfaltig wirkenden Ursachen — besonders durch die verborgenen, indirekten Ursachen — ist die Analyse derselben mittels exakter mathematischer Methoden begründet, u. zw. — da im Verkehr *zufallsbedingte Ereignisse massenhaft* vorkommen, — unter Anwendung der modernen Disziplinen der mathematischen Statistik. Die mathematische Statistik befaßt sich nämlich mit der statistischen Beobachtung von Gesamtheiten, charakterisiert durch Zufallsgrößen, mit der Verarbeitung, und der Analyse dieser empirischen Daten. Die Unfallanalyse kann von den Hauptabschnitten der mathematischen Statistik (Momente, Varianz, Verteilungen usw.) die Möglichkeiten der *statistischen Zusammenhangsprüfungen* am nützlichsten anwenden.

Verfasser setzt sich das Ziel, von den mathematischen Systemen im Rahmen der stochastischen Systeme einen Überblick der Möglichkeiten der Korrelationsrechnung zu geben, u. zw. aus der Sicht, welches Untersystem in Abhängigkeit von der Zahl der Veränderlichen und von dem Zweck der Untersuchung — über die rein diagrammartige Darstellung der Zusammenhänge hinaus — für die Unfallanalyse empfohlen werden kann.

II

Einführend sei noch bemerkt, daß für die Kennzeichnung der Veränderlichen durch nur eine einzige Größe die sog. Momente zur Verfügung stehen.

Im Vergleich zum Mittelwert bietet der *dichteste Wert* jedenfalls mehr, da dieser als häufigster Wert wirklich vorkommt. Bei der Betrachtung nimmt man dagegen auf der Häufigkeitskurve sofort wahr, daß — in prozentualer Verteilung — die meisten Unfälle a) durch welche Ursachen oder b) durch welche Streckenabschnitte und c) wann herbeigeführt werden. Die Ursachen werden auf die Horizontalachse des Koordinatensystems, das Verhältnis von deren Vorkommen in Prozenten aller Ereignisse auf die Vertikalachse aufgetragen.

Durch die *Varianz* (Mittelwert der quadratischen Abweichungen von dem Durchschnitt) wird auch — als Moment zweiter Ordnung — die Streuung des Vorkommens gekennzeichnet; all das gibt jedoch jeweils nur auf den Verlauf je eines einzigen Kriteriums Aufschluß. Durch die Zerlegung des Streuungsquadrats σ^2 auf zwei Komponenten ermöglicht die Varianzanalyse, die vollständige Varianz der Grundmenge auf zwei Komponenten (z. B. Ursachen außerhalb und innerhalb von Unfallgruppen) zu zerlegen, die Wirkung der Einflußfaktoren zu trennen und $\sigma^2 = \sigma_k^2 + \sigma_B^2$ zu bestimmen. Für unsere Analyse ist es wichtig, daß — ist $F = \left(\frac{\sigma}{\sigma_B}\right)^2$ signifikant — auch eine maßgebende äußere Ursache vorliegt.

Mit der Analyse der Zusammenhänge zwischen den quantitativen Kriterien beschäftigt sich der Abschnitt der mathematischen Statistik über *Korrelationsrechnung*, deren Möglichkeiten nun in der Unfallanalyse behandelt werden sollen. Das ist bereits eine mathematisch-statistische Untersuchungsmethode der *stochastischen Beziehung zwischen mehreren — vor allem zwei — Kriterien*, bei der versucht wird, die Wirkung des Einflußfaktors auf die Zielgröße: auf die Erfolgsveränderliche zu kennzeichnen.

1. Die *Rangkorrelation* z. B., bei der die Glieder einer Menge nach der Häufigkeit der Wirkungstypen zweier Kriterien in eine *Reihenfolge* geordnet werden, weist nur auf das Vorhandensein einer Beziehung hin, klärt deren quantitative Elemente auf, wobei der Spearmansche Beiwert das Ausmaß der Übereinstimmung der Häufigkeitsreihenfolge anzeigt. Die Rangkorrelation hat bei psychologischen und Wirtschaftlichkeitsuntersuchungen Bedeutung, wo sich eine genaue Messung der Elemente schwer ermöglichen läßt, und das gilt auch für die Unfallerhebungen. Es können in dieser Weise die *verkehrstechnische Güte von Knotenpunkten* in Städten beurteilt oder die Häufigkeit der Straßunfälle verglichen und auf dieser Grundlage deren Reihenfolge bestimmt werden. Beispiel für die Knotenpunkte 1 und 2:

Unfallart	Unfallziffer / Verkehrsmenge		Reihenfolge nach Unfallziffern		Differenzen und Quadrate der Folgezeiger	
	1	2	1	2	1-2	(1-2) ²
A	128	192	1	1	0	0
B	56	20	2	4	2	4
C	32	4	3	2	-1	1
D	24	16	4	3	-1	1
Insgesamt			4			6

$r = 1 - \frac{6 \cdot 6}{4(16-1)} = 0,4$. Das ist eine lockere Beziehung: die beiden Lösungen sind sichtlich abweichend, erfordern besondere Untersuchung.

2. Die *Kovarianz* bzw. der Assoziationskoeffizient gibt ebenfalls über die Beziehung zweier *qualitativer* Kriterien Aufschluß. Ein interessantes Anwendungsbeispiel ist hier die Folgerung auf den Zusammenhang zwischen der *Vorbestraftheit von Kraftfahrern* und der *Verlässlichkeit* im Betrieb. Bei Unabhängigkeit ist der Yulesche Beiwert, und auch sein Zähler = 0 und bei Beziehungen verschiedener Straffheit $-1 < r < 1$, zum Beispiel:

	Von unbescholtene[m] Vorleben	Vorbestraft
Verlässlich	35	14
Weniger verlässlich	28	24
Insgesamt	63	38

$$r = \frac{\frac{35}{63} \cdot \frac{24}{38} - \frac{28}{63} \cdot \frac{14}{38}}{\frac{35}{63} \cdot \frac{24}{38} + \frac{28}{63} \cdot \frac{14}{38}} = 0,363$$

Beispiel für die Unabhängigkeit: Welche Zahl würde z. B. an die Stelle von 24 kommen?

$35 \cdot x - 28 \cdot 14 = 0$, $x = \frac{28 \cdot 14}{35} = 11$, d. h. unter den Vorbestraften wären die weniger verlässlichen in geringerer Zahl ($11 < 14$), was unwahrscheinlich ist.

3. Durch die *Regressionsanalyse* können auch die *Richtung r* und die *Steilheit b* der Straffheit des stochastischen Zusammenhangs zwischen der abhängigen Veränderlichen Y (Ziel) und der unabhängigen Veränderlichen x durch die Minimalisierung der Quadratsumme der Abweichungen von der

den Zusammenhang (bei linearer Beziehung) annähernden Regressionsgeraden berechnet werden. Der lineare Korrelationskoeffizient, ferner der lineare Regressionskoeffizient sind nämlich

$$r = \frac{\Sigma dx \cdot dy}{\sqrt{\Sigma dx^2 \cdot dy^2}} \quad \text{und} \quad b = \frac{\Sigma dx \cdot dy}{\Sigma dx^2}$$

(dx und dy sind die Abweichungen und $r \max = 1$).

Wieder ein Beispiel aus dem Bereich der Straßenverkehrsunfallverhütung (Koller):

Y Unfälle/Millionen Wagenkm = $0,254 \cdot x_1$ (Verkehrsmenge/Tag) + 1,28 oder $Y = 0,13$. (Längsneigung der Straße in %) + 1,61. Ebenso läßt sich feststellen, daß die Unfallziffern je Straßenkm mit der Verkehrsmenge linear zunehmen, die Unfallziffern auf Nebenlinien 3–4mal höher als auf Hauptverkehrsstraßen sind, daß abends etwa doppelt soviel Unfälle vorkommen als morgens, und nachts 2–3mal mehr als am Tage, selbstverständlich spezifisch.

4. Dehnen wir die Untersuchung auf n Veränderliche aus, von denen $(n-1)$ auf x_1 wirken. Um die Anzahl der Veränderlichen optimal zu vermindern, können vor allem jene *beseitigt* werden, die nicht quantitativer Art sind und keine wichtige Rolle spielen, ferner jene, die in vollen funktionellen Zusammenhängen einander ersetzen können. Beeinflussen sich die $(n-1)$ Veränderlichen gegenseitig nicht, so können die Wirkung von $(n-2)$ Veränderlichen nach dem »ceteris paribus«-Prinzip als Konstanten ausgeschaltet und dadurch die Beziehungen zwischen je zwei Veränderlichen mit Hilfe der *partiellen Korrelationskoeffizienten* untersucht werden. Dieses Verfahren läßt sich bei der Analyse der *von mehreren Ursachen abhängigen Unfälle* ausgezeichnet anwenden, wenn diese Ursachen voneinander *unabhängig* sind.

Im Falle von *drei Veränderlichen*: Nimmt die Zahl der Kraftfahrzeuge dauernd zu, vermindert sich im allgemeinen die spezifische Zahl der Unfälle, weil sich die Fahrerroutine von Jahr zu Jahr entwickelt. Die Zahl der tödlichen Unfälle je Kraftfahrzeug beträgt $D/N = 0,0003 \left(\frac{N}{P} \frac{\text{Kraftfahrzeuge}}{\text{Einwohner}} \right)^{2/3}$, mit den drei Veränderlichen D , N und P .

a) Auch hier wird die Straffheit der Beziehungen zwischen der *Wirkung und den einzelnen Ursachen* durch den Korrelationskoeffizienten gekennzeichnet, auch die Steilheit der — linearen — Beziehung wird durch die Regressionskoeffizienten b , gekennzeichnet, die ganz ähnlich wie die vorigen, jedoch in Matrixform angeschrieben werden können. Die Formel der Korrelation zwischen einer Zielgröße Y und der wirksamen Veränderlichen x_1 lautet, nachdem die Wirkung von x_2 *ausgeschaltet* wurde:

$$r_{Y1,2} = \frac{r_{Y1} - r_{Y2} \cdot r_{12}}{(1 - r_{Y2}^2) \cdot (1 - r_{12}^2)}.$$

Wird zum Beispiel die Straffheit (r) der Beziehung zwischen den Veränderlichen, durch die Matrix

	Y	x_1	x_2
Y	1		
x_1	0,85	1	
x_2	0,73	0,76	1

gekennzeichnet, dann erhält man z. B. nach Ausschalten der Wirkung der Veränderlichen Nr. 2 (x_2): $r_{Y1,2} = 0,66$, jedoch nur $r_{Y2,1} = 0,25$. Daraus ist zu erkennen, daß zwischen der Zielgröße Y und der Veränderlichen Nr. 2 eine viel lockerere, fast gar keine Beziehung besteht. Dieser Gedankengang ist übrigens ähnlich der »Standardisierung« der einen Verhältniszahl bei dem Vergleich von Intensitäts-Verhältniszahlen und wird in den Indexgleichungen mit *partiellen Elastizitäten* als Exponenten angewandt.

b) Die von Prof. Kádas beschriebene Indexgleichung ermöglicht durch die Berücksichtigung der Veränderung von vier Faktoren eine schätzungs-mäßige Voraussage der Fahrgastzahlerwartungen im Budapester Massenverkehr; diese vier Faktoren sind Einwohnerzahl, Netzlänge, Indexzahl der Reallohnzunahme und die negative Wirkung des Realtarifs. Das Verhalten der Exponenten wieder unter Anwendung des »ceteris paribus«-Prinzips bestimmt, lassen sich aus einer solchen Indexgleichung auch für den gemeinsamen Verlauf der Unfälle mehrere zusammengesetzte Wirkungen nachweisen. Sind z. B. die Elastizitäten von vier unabhängigen Unfallursachen 1,2; 0,7; 0,1 (gering) und $-0,1$ (Verbesserungsmaßnahme), ergeben sich aus diesen Wirkungen Zunahmeindexe der Unfallziffer von $I_b = c \cdot O_1^{1,2} \cdot O_2^{0,7} \cdot O_3^{0,1} \cdot O_4^{-0,1}$, wo z. B. $c = 10^{-2}$.

c) Sollen z. B. in zwei Bezirken die von den Kraftfahrern verursachten Unfälle verglichen werden, werden die Unfallzahlen selbstverständlich zunächst spezifisch auf die Zahl der Fahrer bezogen. Das genügt jedoch nicht, da die Zusammensetzung nach Altersgruppen — von der bekanntlich die Verursachung von Unfällen abhängig ist — unterschiedlich sein kann. Um diese Störwirkung bei der Untersuchung zu beseitigen, werden *im Laufe der Standardisierung* die spezifischen Unfallziffern je Altersgruppen beider Bezirke mit der gleichen Altersgruppenverteilung (z. B. $N_2\%$) multipliziert, und damit ist die Gestaltung der Unfälle von der Wirkung eines der Einflußfaktoren befreit.

Altersgruppen der Fahrer (Jahre)	Verteilung ihrer Anzahl		Spezifische Unfallziffer in %	
	N ₁	N ₂	B ₁	B ₂
16—25	16,1	15,3	3,6	4,6
26—34	17,4	16,6	0,5	0,7
35—43	16,4	12,5	1,0	1,3
44—52	15,5	14,1	1,7	2,0
53—61	22,6	24,8	6,0	6,2
62—	12,0	16,7	47,7	47,9
Durchschnitt:			8,2	10,8

Mit der *eigenen* Altersklassenverteilung betrug die spezifische Unfallziffer für den Bezirk 1:

$$0,153 \cdot 4,6 + 0,166 \cdot 0,7 + \dots + 0,167 \cdot 47,9 = 10,8,$$

auch hier mit der Altersklassenverteilung des Bezirks 2 gerechnet, würde sich

$$0,153 \cdot 3,6 + 0,166 \cdot 0,5 + \dots + 0,167 \cdot 47,7 = 10,5$$

ergeben.

Wird nun dieser Wert mit dem Durchschnitt 10,8 des Bezirks 2 verglichen, erhält man schon ein weniger günstiges Bild; der niedrige Durchschnittswert 8,2 in der Tabelle ist großenteils die Folge der für das Fahren günstigeren Alterszusammensetzung.

5. Für die partiellen Korrelationsrechnungen ist auch die — bereits erwähnte — Untersuchung der Unabhängigkeit von je zwei Faktoren zu unternehmen. Bei der *Unabhängigkeitsprüfung* wird die Annahme der Unabhängigkeit zweier Veränderlichen (mit s bzw. mit t Varianten), die sog. Nullhypothese dann abgelehnt, wenn der berechnete Wert höher ist als die Tabellenwerte des chi-Quadrattests mit $DF = (t-1) \cdot (s-1)$ Freiheitsgrad.

Es ist bekannt, daß nach dem *Bayesschen Satz* der »inversen Wahrscheinlichkeit« — der Berechnung der bedingten und der vollen Wahrscheinlichkeit — zwei Ereignisse *dann unabhängig sind*, wenn die Wahrscheinlichkeit des *gemeinsamen* Eintreffens gleich dem Produkt der Wahrscheinlichkeiten der beiden Ereignisse ist. Für unsere Zwecke ist es wichtig, daß der Bayessche Satz ermöglicht, die Wahrscheinlichkeit der Folgen nicht aus der primären Wahrscheinlichkeit der Ursachen zu berechnen, sondern in Kenntnis des Ausgangs des Ereignisses auf die wirkliche »a posteriori« Wahrscheinlichkeit *einen Rückschluß zu ziehen*, der z. B. in Verbindung mit der Klärung grundlegender Unfallursachen bei der für Gerichtsverhandlungen erforderlichen statistischen Analyse benutzt werden kann.

III

Auf den Fall der n Veränderlichen zurückkommend, können diese im Laufe der Analyse kaum mit der Begründung, daß sie konstant seien, festgelegt werden, da die Änderung einer der Veränderlichen auch auf die anderen auszuwirken pflegt, und sich daher die einseitigen Beziehungen in *gegenseitige* verwandeln; auch eine *gemeinsame Ursache* (das ist der »Faktor«) kann den gleichzeitigen Verlauf *mehrerer* Veränderlichen herbeiführen. Damit sind wir bei der mehrfachen Korrelation des Zusammenhanges mit mehreren Veränderlichen angelangt, deren Untersuchung zu der zeitgemäßen Theorie der *Faktoranalyse* gehört.

Wird eine Erfolgsveränderliche (Zielgröße) durch viele, auch *voneinander abhängige* Veränderliche beeinflusst, werden — um die über die beobachteten Veränderlichen erhaltenen Informationen zu erklären, — von der Faktoranalyse *eine geringere Zahl gemeinsamer, voneinander unabhängiger* (orthogonaler) Faktoren eingeführt, die dazu dienen, die Veränderlichen mit der wichtigsten Korrelation zu ermitteln und so zu gruppieren, daß sie *im Grunde* genommen den überwiegenden Teil der Streuung der ursprünglichen Veränderlichen erklären und der Rangordnung nach einstufen. Diese ausgedehnte Untersuchung der Zusammenhänge, welche bisher vor allem als für die Analyse von gesellschaftlich-wirtschaftlichen Erscheinungen geeignet betrachtet wurde, die sich schwer numerisch ausdrücken lassen (die Form der partiellen Korrelation und die mit zwei Veränderlichen ist mehr bekannt), wird im angeführten Buch von Jahn-Vahle auf ein Straßeninvestitionsbeispiel angewandt — das durch die Verkehrszunahme begründet war — und es ergab sich, daß die wirksamen 8 Veränderlichen *grundlegend* durch 3 Faktoren beeinflusst werden: von der Einwohnerzahl, dem Ausmaß der Beschäftigung und (negativ) von der Transportentfernung.

Wir setzten uns zum Ziel, diese Methode auch auf die Unfallanalyse auszudehnen, die die Ereignisse *grundsätzlich erklärenden und wesentlichen* Zusammenhänge zu verfolgen und die einzelnen Schritte zu zeigen. Man muß davon ausgehen, daß das einzelne Unfallereignis nicht von einer einzigen und scheinbar unmittelbaren Ursache abhängt, ferner die Ursachen gewiß auch zueinander in Beziehung stehen. Daher müssen nach der Methode der Faktoranalyse die wichtigsten erforscht, und durch Gruppierung die grundlegenden Ursachen gefunden werden, da nur in Kenntnis dieser wirksame Vorbeugungsmaßnahmen getroffen werden können. Es läßt sich z. B. klären, ob für den Zusammenstoß oder Zusammenrutsch, den Sturz im Wagen infolge des Abschwenkens des Lenkrades unmittelbar der Fahrer oder im Grunde genommen der das Schwenken herbeiführende — von dem Gehsteig plötzlich auf die Fahrbahn tretende — Fußgänger verantwortlich war.

1. Die Untersuchung der Ereignisse zeigt überwiegend, daß die Unfallziffer (die nach Schwere und Art des Unfalls aufgeschlüsselt werden kann) als *Zielgröße* von folgenden Veränderlichen abhängt (die grobenteils auch voneinander abhängig sind):
2. Fahrereigenschaften (Fachkenntnis, Lebensalter, Gewissenhaftigkeit)
3. Fahrzeugart
4. Fahrzeug-Instandhaltung
5. Fahrbahnzustand
6. Straßenparameter
7. Verkehrsmenge
8. Zeitpunkt
9. Witterungsverhältnisse
10. Verkehrsleitung.

a) Die Faktoranalyse geht von der alle r gegenseitigen Beziehungen der — in der Aufzählung gemessenen Veränderlichen $x_1 - x_{10}$ oder standardisierten Veränderlichen $z_1 - z_{10}$ darstellenden — Korrelationsmatrix R aus (r_{ij} Korrelationskoeffizienten), aus der die Faktoren bzw. deren Gewichte a später abgeschätzt werden können.

b) Die Faktorengewichte a bedeuten nämlich die Beteiligung eines grundlegenden Faktors F an den einzelnen beobachteten Eigenschaften (den Grad der Zusammenhänge), und ihre Rolle ist in der »Modellgleichung« der Faktoranalyse gut zu erkennen, die die standardisierten Meßergebnisse z_j als mit diesen Faktorengewichten gewogene Summe der gesuchten Faktoren F_m angibt: $z_j = \sum a_{jm} \cdot F_m$. Diese Wirkungen a_j je eines solchen Faktors auf die Veränderlichen, können zusammen in eine Matrix A der Faktorengewichte gefaßt werden, die die *Straffheit der Korrelationen* zwischen den Spalten der m Faktoren und den Zeilen der genannten zehn Veränderlichen darstellt.

c) Die genannten Korrelationskoeffizienten der Veränderlichen lassen sich auf die Produktensumme der Faktorengewichte a zurückführen; in Matrixform geschrieben ist $R = 1/N \cdot Z^x \cdot Z = A \cdot A^x$ (transponiert); diese Gleichung wird *die Grundgleichung der Faktoranalyse* genannt (nach Überla: Fundamentaltheorem). Hierin bedeuten N die Zahl der untersuchten Individuen, m die Zahl der gemeinsamen Faktoren in den zehn beobachteten Kriterien.

d) Nebenbei bemerkt, müssen zuerst für die Diagonale der Korrelationsmatrix R mit der Näherung $h^2 = \max(r_j)$ die sog. h^2 -Kommunalitäten, d. h. die Anteile der Veränderlichen z an dem durch die gemeinsamen Faktoren erschließbaren, vollen Streuungsquadrat (Varianz), abgeschätzt werden.

e) Darauf folgend können aus der Korrelationsmatrix R die Faktorengewichte mit Hilfe der didaktisch einfacheren Thurstoneschen Zentroid-Spiegelung (oder nach der den Einsatz einer Rechenanlage erfordernden Hotellingschen Hauptfaktormethode) bestimmt werden; damit dem Ziel der

Faktoranalyse entsprechend alle Faktoren das maximale Streuungsquadrat vertreten — d. h. die Restmatrizen R_1 minimal und die Spaltensummen positiv sind — müssen die Spalten mit einer großen Anzahl negativer Korrelationen mit -1 multipliziert werden.

f) So kommt man nach *Weglassen der Faktoren mit geringen Zusammenhängen* zu der Matrix A der »extrahierten« Faktorengewichte und zu dem — aus für die Erklärung minimal erforderlichen Faktoren der Anzahl m aufgetragenen — Koordinatensystem,

g) das zwecks *eindeutiger* Interpretation rotiert (transformiert, verdreht) werden muß (damit nicht ein einziger Faktor sämtliche große Faktorengewichte enthält), ferner um den mit allen Veränderlichen hoch korrelierenden, also nichtssagenden sog. allgemeinen Faktor nach der Thurstone-Regel zu beseitigen. $A' = A \cdot T$, wo T Kreisfunktionsmatrix des Rotationswinkels φ ist.

h) Auf die Zielgröße üben jene Faktoren eine *grundlegende Wirkung aus, in denen viele hochwertige* Korrelationen vorkommen (0,8; 0,9), und die durch diese vertretenen Kriterien sind für die einzelnen Faktoren *kennzeichnend*. In der Unfallursachenforschung erhält man als Ergebnis mindestens 3 Faktoren als grundlegende, *gemeinsame* Faktoren mit größerem Einfluß, das sind die »human factors« F_1 (z. B. auch in der Instandhaltung), die Fahrbahn F_2 (die auch in den Punkten 3, 5, 6, 9 vorkommt) und die Verkehrsmenge F_3 .

i) Schließlich läßt sich auch die *Rangordnung der Einflüsse* der auf die Zielgröße wirkenden Faktoren feststellen, wenn die Wirkung der anderen Faktoren auf die Zielgröße sehr klein ist.

Die Rechenanlage ermöglicht die *Simulation* von komplizierten Berechnungen, wenn sie die reellen Verhältnisse durch Versuche annähert und nach einem Iterationsverfahren bestimmt, unter welchen Bedingungen die aus Stichproben der Veränderlichen erhaltenen Faktoren (also auch aus deren Korrelationen) den Eigenschaften der Grundgesamtheit entsprechen.

Faktoren sind in der Regel wenig sinnennahe, abstrakte Begriffe, jedoch kann dem Umstand, zu welcher Art von Veränderlichen sie hauptsächlich in Beziehung stehen, der Charakter der einzelnen Faktoren entnommen werden. Um dies zu veranschaulichen, weisen wir wieder auf die Ergebnisse in den Beispielen von Jahn-Vahle hin, nach denen die Wirkung der die *Arbeitsproduktivität* unmittelbar beeinflussenden (miteinander verbundenen) wenigstens 13 Veränderlichen auf die Wirkung von *zwei* Hauptfaktoren zurückgeführt werden kann, die finanziellen und Arbeitsmengenbeziehungen, Einflüsse vereinen, also diese am meisten beeinflussen. Die etwa 8 Kennwerte der landwirtschaftlichen Produktion werden am meisten durch die für tierische Produkte, Erträge kennzeichnenden Kriterien bestimmt. Die Haushalts-Dienstleistungen werden durch einen mehr als die Hälfte der Veränderlichen bewegend, also allgemeinen Faktor sowie negativ durch die Wirkung der Reservevorräte »aus dem Hintergrund« gelenkt.

Zusammenfassend darf ausgesagt werden, daß es zweckmäßig ist, in der Unfallforschung die Disziplinen der mathematischen Statistik anzuwenden, da im Verkehr Zufallsereignisse massenhaft vorkommen, u. zw. für die Analyse der Unfallursachen vor allem die zeitgemäßen Methoden zur Prüfung der statistischen Zusammenhänge. Es muß zur Kenntnis genommen werden, daß einerseits viele Unfälle voneinander nicht unabhängig sind, andererseits daß sie auch gemeinsame, grundlegende Ursachen haben. Daher wurden hier die Schritte der Faktoranalyse gezeigt, mit deren Hilfe man zu einer *geringeren Zahl grundlegender* Ursachen gelangt, und damit wirksamere Unfallverhütungsmaßnahmen ermöglicht werden. Für all das sind auch die in Abschnitt I vorgeschlagenen Aufgaben für eine bessere Datenlieferung unerlässlich.

Zusammenfassung

Die Methoden der Unfallverhütung werden durch die Unfallursachen, ihr org. Niveau durch die Größe der Unfallziffern bestimmt. Die Rechentechnik ist die Voraussetzung der modernen Verarbeitung ausführlicher Datenerhebungen, der Simulation und der Verwirklichung einer internationalen Datenbank. Es müssen mehr Kennwerte als bisher eingeführt werden, um die Verkehrsunfälle vor allem zu den Leistungen in Verhältnis zu stellen; und es ist auch die Verteilung der Dauer der Hindernisse infolge von Unfällen zu prüfen.

Die Methoden der Korrelationsrechnung beschäftigen sich mit der stochastischen Beziehung zwischen zwei und mehreren Veränderlichen. Hängen diese Veränderlichen auch miteinander zusammen, ist die moderne Methode der *Faktoranalyse* heranzuziehen. Bei der Faktoranalyse werden eine geringe Zahl voneinander unabhängiger gemeinsamer *Faktoren* eingeführt, die gewissermaßen hinter den Veränderlichen stehen und die grundlegenden Ursachen von deren Entstehung mit einer bestimmten Genauigkeit erklären. Daher spielt die Faktoranalyse in der Erforschung der Unfallursachen, bei der Klarlegung der mittelbaren, sekundären Ursachen eine wichtige Rolle.

Im Beitrag werden Anwendungsbeispiele der Korrelationsmethoden in der Unfallanalyse sowie die Schritte der Faktoranalyse gezeigt.

Literatur

1. KOLLER, S.: Stadtverkehrstechnik,* Tankönyvkiadó, Budapest, 1968.
2. KÁDAS, K.: Methoden der Verkehrsstatistik,* Tankönyvkiadó, Budapest, 1974.
3. KÖVES P.—PÁRNICZKY, G.: Allgemeine Statistik,* Közgazdasági és Jogi Könyvkiadó, Budapest, 1973.
4. JAHN, W.—VAHLE, H.: Die Faktoranalyse und ihre Anwendung,* Übersetzt von M. Basa und L. Ferich. Közgazdasági és Jogi Könyvkiadó, Budapest, 1974.
5. ÜBERLA, K.: Faktoranalyse. Springer Verlag, Berlin, 1971.
6. GYULAI, G.: Városi közlekedés. 15, p. 224 (1975).
7. Informationsschrift über die Gestaltung der Verkehrsunfälle.* Manuskript Budapester Verkehrsunternehmen, 1974.
8. 3. Symposium für Sicherheit im Stadtverkehr*, Közdoc, Budapest, 1972.
9. Sicherheit im Massenverkehr.* Studie Nr. 11(A)13, Manuskript TU Budapest, 1974.

Dr. Géza GYULAI, 1117 Budapest, Schönherz Z. u. 25.

* In ungarischer Sprache.