Z. Wahrscheinlichkeitstheorie 3, 193—203 (1964)

Über den Zentralen Grenzwertsatz für abhängige Zufallsvariable

Von

F. Eicker

Zusammenfassung

Es wird eine Verallgemeinerung eines von Godwin und Zaremba (1961) angegebenen Zentralen Grenzwertsatzes für abhängige Zufallsvariable, der wegen zahlreicher Anwendungen in der mathematischen Statistik von Interesse ist und in dessen Zusammenhang der Begriff der Verkettung von Zufallsvariablen eingeführt wurde, hergeleitet. Der Beweis beruht, ebenso wie der des Satzes von Godwin und Zaremba, auf der Momentenmethode. Der hier vorgelegte Satz macht von dem Verkettungsbegriff keinen Gebrauch. In Abschnitt 3 wird ferner der bislang allgemeinste Zentrale Grenzwertsatz für m-abhängige Zufallsvariable (Orey 1958) in weiterhin leicht verallgemeinerter Form auf elementarem Wege bewiesen.

1.

Ein System disjunkter, endlicher, nicht leerer Mengen A, $B, \ldots$ von Zufallsveränderlichen (ZV), die alle auf demselben Wahrscheinlichkeitsraum definiert sind, heißt *unabhängig*, wenn jede Vereinigung endlich vieler endlicher Teilmengen $A' \subset A$, $B' \subset B, \ldots$ von der Vereinigung endlich vieler endlicher Teilmengen aus den übrigen Mengen des Systems unabhängig ist [*3*, Sc. 223—5]. Zwei disjunkte, endliche, nicht leere Mengen von ZV heißen unabhängig, wenn die gemeinsame Verteilungsfunktion (VF) der ZV der Vereinigungsmenge gleich dem Produkt der VF der ZV der beiden Mengen ist.

Eine nicht leere Menge von ZV heißt *irreduzibel*, wenn sie keine Zerlegung in unabhängige Teilmengen gestattet. Es ist leicht zu sehen, daß jede abzählbare Menge von ZV eine eindeutige Zerlegung in irreduzible Teilmengen besitzt (hinfort *I-Zerlegung* genannt) [*2*].

Wir betrachten Doppelfolgen $\{\xi_{nk}, k = 1, \ldots, k_n; n = 1, 2, \ldots\}$ von ZV, für die $k_n \to \infty$ gilt (dieser und alle folgenden Limites gelten für $n \to \infty$, soweit nichts anderes angegeben ist). Die Gesamtheit der irreduziblen Mengen von je s ZV $(\xi_{nj_1}, \ldots, \xi_{nj_s})$, die durch beliebige Kombination (mit Wiederholungen) von ZV aus der n-ten Zeile $(\xi_{n1}, \ldots, \xi_{nk_n})$ der Doppelfolge gebildet werden können, wird mit U_{ns} $(s = 2, 3, \ldots; n = 1, 2, \ldots)$ bezeichnet. Die Elemente von U_{ns} bezeichnen wir mit $u, v, \ldots$; die Indizes n und s an den Elementsymbolen $u, v, \ldots$ sind weggelassen, da Verwechselungen nicht zu befürchten sind. Die Elemente von U_{ns} lassen sich dann ausführlicher auch in der Form schreiben $(\xi_{nj_1(u)}, \ldots, \xi_{nj_s(u)})$, $u \in U_{ns}$. Indexvektoren, die aus dem Vektor $(j_1(u), \ldots, j_s(u))$ durch Permutation der Komponenten hervorgehen, liefern definitionsgemäß ein und dasselbe Element aus U_{ns}. Soll über alle Mengen $u \in U_{ns}$ summiert werden, so schreiben wir $\sum\limits_{U_{ns}}$.

Es gilt der folgende Zentrale Grenzwertsatz für abhängige ZV, in dem die Existenz sämtlicher Momente der ZV vorausgesetzt wird. Diese Voraussetzung wird jedoch in den späteren Sätzen beseitigt.

Satz 1. *Alle Momente der ZV einer Doppelfolge $\{\xi_{nk}\}$ mögen existieren, und es sei*

$$E\,\xi_{nk} = 0 \qquad \textit{für alle } n \textit{ und } k, \tag{1}$$

$$\sum_{j,k=1}^{k_n} E(\xi_{nj}\xi_{nk}) = 1 \qquad \textit{für alle } n. \tag{2}$$

Es gelte die Voraussetzung (I): *für* $s = 3, 4, \ldots$ *gelte*

$$\sum_{U_{ns}} \Big| E\Big(\prod_{t=1}^{s} \xi_{nj_t}\Big)\Big| \to 0. \tag{3}$$

Es sei ferner $\tau = (t_1, \ldots, t_w)$ *ein* w*-dimensionaler Vektor,* $w = 3, 4, \ldots$, *dessen Komponenten ganzzahlig sind und die Bedingungen*

α) $t_1 = 0 < t_2 < \cdots < t_w = s$,

β) $\min\limits_{v}(t_{v+1} - t_v) = 2$

erfüllen. $U_{ns}(\tau)$ *sei diejenige Teilmenge von* U_{ns}, *deren Elemente* $(\xi_{nj_1(u)}, \ldots, \xi_{nj_s(u)})$ *eine Zerlegung in* $w - 1$ *irreduzible (notwendig abhängige) Teilmengen mit je* t_2, $t_3 - t_2, \ldots, s - t_{w-1}$ *ZV gestatten. Diese Teilmengen seien mit* $\{\xi_{nj_t(u)},\ t_v + 1 \leqq t \leqq t_{v+1}\}$, $v = 1, \ldots, w-1$, *bezeichnet. Für jeden Vektor* τ, *der die Bedingungen* α) *und* β) *erfüllt, und alle Zahlen* $w = 3, 4, \ldots$ *sowie* $s = 3, 4, \ldots$ *gelte die Relation*

$$\sum_{u \in U_{ns}(\tau)} \Big|\prod_{v=1}^{w-1} E\Big(\prod_{t=t_v+1}^{t_{v+1}} \xi_{nj_t(u)}\Big)\Big| \to 0. \tag{4}$$

Unter diesen Voraussetzungen konvergieren die VF der ZV

$$\zeta_n = \sum_{k=1}^{k_n} \xi_{nk}$$

gegen N(0,1) *(= Standardnormalverteilung), und die* ξ_{nk} *sind infinitesimal, d.h.*

$$\lim_{n\to\infty} \max_{k} P(|\xi_{nk}| > \varepsilon) = 0 \qquad \textit{für alle } \varepsilon > 0.$$

Beweis: 1. Nach dem Momentenkonvergenzsatz (oder Zweiter Grenzwertsatz der Wahrscheinlichkeitsrechnung; vgl. etwa LOÉVE (1960), S. 185), ist für die Konvergenz einer Folge von VF gegen N(0,1) hinreichend, daß die Momente m-ter Ordnung ($m = 0, 1, 2, \ldots$) gegen die entsprechenden Momente der Normalverteilung

$$\mu_m = \begin{cases} m!/((m/2)!\,2^{m/2}), & m \text{ gerade}, \\ 0, & m \text{ ungerade}, \end{cases}$$

konvergieren. Wir werden zeigen, daß die Momente

$$M_{nm} = E(\zeta_n^m) = \sum_{j_1=1}^{k_n} \cdots \sum_{j_m=1}^{k_n} E\Big(\prod_{t=1}^{m} \xi_{nj_t}\Big) \tag{5}$$

für $n \to \infty$ gegen μ_m streben für $m = 1, 2, \ldots$. Da die folgenden Überlegungen jeweils für einen festgehaltenen Wert von n gelten, lassen wir diesen Index fort, soweit kein Irrtum entstehen kann.

2. Wir betrachten, jeweils für festgehaltene m, die $(k_n)^m$ Indexvektoren $(j_1, \ldots, j_m)$, die in der m-fachen Summe in (5) auftreten. Die Anordnung der Komponenten wird durch die Reihenfolge der Summationen in (5) festgelegt. Die Gesamtheit der Vektoren läßt sich durch die Gitterpunkte eines m-dimensionalen endlichen Raumes darstellen. Die Koordinatenachsen dieses Raumes numerieren wir mit $1, 2, \ldots, m$ durch. Es sei A_m die Gesamtheit der Zerlegungen α dieses Raumes in nichtleere, durch Achsen aufgespannte Unterräume. Die Nummern der diese Unterräume aufspannenden Achsen fassen wir zu den Mengen

$$T_1(\alpha), \quad T_2(\alpha), \ldots, T_{r(\alpha)}(\alpha)$$

zusammen; $r(\alpha)$ ist die Anzahl der Unterräume der Zerlegung α (wird α festgehalten, so schreiben wir auch nur r). Die Dimension des Teilraumes $T_i(\alpha)$ sei α_i.

Nun zerfällt jede der in (5) auftretenden geordneten Mengen $x = (\xi_{j_1}, \ldots, \xi_{j_m})$ in ein System irreduzibler Teilmengen, dem eine bestimmte Zerlegung des zugehörigen Indexvektors $(j_1, \ldots, j_m)$ [genauer: $(j_1(x), \ldots, j_m(x))$] und somit eine Zerlegung $\alpha \in A_m$ entspricht. Ist X_α die (von n abhängige, möglicherweise leere) Menge aller x, deren I-Zerlegung der Zerlegung α entspricht, so ist (5) identisch mit

$$M_m = \sum_{\alpha \in A_m} \sum_{x \in X_\alpha} E\Big(\prod_{t=1}^{m} \xi_{j_t(x)}\Big), \tag{6}$$

und weiter für $x \in X_\alpha$

$$E\Big(\prod_{t=1}^{m} \xi_{j_t(x)}\Big) = \prod_{i=1}^{r(\alpha)} E\Big(\prod_{t \in T_i(\alpha)} \xi_{j_t(x)}\Big). \tag{7}$$

Wir zeigen jetzt, daß die einzelnen Summen $\sum_{X_\alpha}$ für $n \to \infty$ konvergieren. Da die Menge A_m von n unabhängig ist, ergibt sich daraus das asymptotische Verhalten von M_m.

3. Ist für ein α $\alpha_{\min} \equiv \min \alpha_i = 1$, so verschwindet $\sum_{X_\alpha}$ für alle n identisch wegen (1). Es genügt daher, $\alpha_{\min} \geqq 2$ zu betrachten.

Es sei α eine Zerlegung, für die ein $\alpha_i \geqq 3$ ist. Wir betrachten die in $\sum_{X_\alpha}$ enthaltenen Teilsummen, in denen nur über die Indizes $j_t(x)$, $t \in T_i(\alpha)$, i fest mit $\alpha_i \geqq 3$, summiert wird und die übrigen Indizes $j_t(x)$, $t \notin T_i(\alpha)$, festgehalten werden. (Die Summationsbereiche der Teilsummen hängen im allgemeinen außer von n und α auch von den Werten der festgehaltenen Indizes ab.) Geht man in diesen Teilsummen zu den Beträgen über, so erhält man jeweils Teilsummen der gegen Null strebenden Summen über $U_{n\alpha_i}$ (oder von n unabhängiger Vielfacher dieser Summen) in der Beziehung (3). Diese letzteren bilden also Majoranten der betrachteten Teilsummen, die gleichmäßig in den $j_t(x)$, $t \notin T_i(\alpha)$, gegen Null streben.

Es folgt $\lim\limits_{n \to \infty} \sum_{X_\alpha} = 0$, wenn $\alpha_{\min} > 2$ ist. Dasselbe gilt für die Folge der Summen über die Beträge und für Summen über (für jedes n beliebige) Teilmengen der Mengen X_α.

4. Sei α eine Zerlegung, für die ein $\alpha_i = 2$ ist. Wir betrachten für ein solches i wieder wie in 3. die in $\sum_{X_\alpha}$ enthaltenen Teilsummen über die Indizes $j_t(x)$, $t \in T_i(\alpha)$, bei jeweils festgehaltenen Werten der übrigen Indizes, nehmen nun aber nicht die Betragssummen, sondern ergänzen die jeweiligen Teilsummen so, daß

beide Indizes j_t, $t \in T_i(\alpha)$, von 1 bis k_n laufen. Jede so ergänzte Teilsumme ergibt nach (2) den Wert Eins. Daher fallen in der bezüglich $T_i(\alpha)$ ergänzten Summe $\sum X_\alpha$ die Summationen über zwei Indizes heraus, und die ergänzte Summe schrumpft auf eine Teilsumme von $\sum X_{\alpha'}$ zusammen, wo $\alpha' \in A_{m-2}$ bis auf den wegzulassenden, durch $T_i(\alpha)$ gekennzeichneten Teilraum eine zu α analoge Zerlegung des $(m-2)$-dimensionalen Indexraumes ist.

In 5. wird gezeigt, daß die Summe der ergänzten Terme asymptotisch verschwindet.

Ist nun $\alpha'_{\min} > 2$, so folgt aus 3., daß die soeben betrachtete Teilsumme von $\sum X_\alpha$ und damit $\sum X_\alpha$ gegen Null geht, vorbehaltlich des Nachweises von Abschnitt 5.

Ist $\alpha'_{\min} = 2$, so kann man die betrachtete Teilsumme von $\sum X_\alpha$ wiederum in der oben beschriebenen Weise ergänzen. Der hierdurch gewonnene, sich möglicherweise vielfach verzweigende Reduktionsprozeß endet in jedem Zweig nach endlich vielen Schritten und liefert (vorbehaltlich 5.) $\sum X_\alpha \to 0$, wenn $\alpha_{\max} > 2$. Das gleiche gilt, wenn nur über einen Teilbereich von X_α summiert wird.

Ist m gerade und $\alpha_{\min} = \alpha_{\max} = 2$, so folgt $\sum X_\alpha \to 1$. Beachtet man noch, daß es $m!/((m/2)!\,2^{m/2})$ Zerlegungen α mit $\alpha_{\min} = \alpha_{\max} = 2$ gibt, so folgt

$$\lim_{n\to\infty} M_{nm} = \mu_m .$$

Vorbehaltlich 5. ist damit Satz 1 bewiesen.

5. Die bei der Ergänzung von $\sum X_\alpha$ in 4. hinzugefügten Mengen x gehören sämtlich zu Klassen X_{α^*} mit $\alpha^* \neq \alpha$ und $r(\alpha^*) < r(\alpha)$ (insoweit $\alpha^*_{\min} > 1$). Ein hinzugefügtes ZV-Paar $\xi_{j_t}, \xi_{j_{t'}}$, $t, t' \in T_i(\alpha)$, (soweit es irreduzibel ist) gehört notwendig einer irreduziblen Menge mit mindestens vier ZV-en an. Also muß $\alpha^*_{\max} \geqq 4$ gelten, und es folgt, wenn außer der Voraussetzung (3) nun auch Voraussetzung (4) herangezogen wird, das asymptotische Verschwinden der Summe der ergänzten Terme, wobei im Falle $\alpha^*_{\min} = 2$ das Reduktionsverfahren des Abschnitts 4. anzuwenden ist.

6. Aus (4) folgt $\sum\limits_k (E(\xi_{nk}^2))^2 \to 0$, somit $\lim\limits_{n\to\infty} \max\limits_k E(\xi_{nk}^2) = 0$. Die Tschebycheffsche Ungleichung liefert daraus schließlich die Infinitesimalität der ξ_{nk}, womit der Beweis abgeschlossen ist.

Bemerkung 1.1. Unter den Voraussetzungen von Satz 1 läßt sich in jeder Zeile der Doppelfolge $\{\xi_{nk}\}$ eine nicht leere Menge von unabhängigen ZV $\xi_{nj_1}, \ldots, \xi_{nj_u}$ finden mit der Eigenschaft, daß die Anzahl der ZV, $u = u(n)$, gegen ∞ strebt für $n \to \infty$.

Beweis: Da jede Zeile nur endlich viele ZV besitzt, gibt es stets eine Teilmenge von unabhängigen ZV mit maximaler Variablenzahl $u'(n) \geqq 1$. Angenommen, es gäbe eine unendliche Teilfolge $\{n_i\}$ der natürlichen Zahlen, so daß $\sup\limits_i u'(n_i) = u_0 < \infty$ ist. Dann gäbe es in den Zeilen mit $n = n_i$ keine $v\,(> u_0)$ untereinander unabhängige, irreduzible ZV-Paare. Folglich wären die im obigen Beweis auftretenden Klassen X_α für Zerlegungen α mit $\alpha_{\min} = \alpha_{\max} = 2$ für Momente der Ordnungen $m > 2u_0$ leer. Allein diese Klassen lieferten aber asymptotisch einen positiven Beitrag zu den Momenten (6).

2.

Die Voraussetzungen über die Existenz der Momente der ξ_{nk} lassen sich ohne Schwierigkeit durch andere ersetzen. Zwei Möglichkeiten, nach denen dies geschehen kann, fassen wir in Satz 2 zusammen. — Ist zunächst x eine beliebige ZV und d eine positive Konstante, so bedeutet x^d diejenige ZV, die gleich x ist für $|x| < d$ und $x = 0$ für $|x| \geqq d$.

Satz 2. *Sei* $\{x_{nk};\, k = 1, \ldots, k_n;\, n = 1, 2, \ldots\}$ *eine Doppelfolge von* ZV *mit* $k_n \to \infty$ *(alle Limesrelationen des Satzes gelten für* $n \to \infty$*). Es gelte:*

1. *für eine Folge* $\{d(n)\}$ *reeller positiver Zahlen und für jedes* $\delta > 0$ *strebe*

$$P|(\sum_k (x_{nk} - x_{nk}^{d(n)})| > \delta) \to 0 \tag{8}$$

(existieren sämtliche Momente aller x_{nk}*, so können alle* $d(n) = \infty$ *gewählt werden; Beziehung* (8) *ist dann identisch erfüllt). Anstelle von* (8) *kann die folgende stärkere Voraussetzung* 1'. *gemacht werden*:

1'. *es gebe ein* $t > 0$ *und eine Folge* $\{d(n)\}$, $d(n) > 0$, $n = 1, 2, \ldots$, *so daß die beiden Beziehungen*

$$\sum_k P(|x_{nk}| > t) \to 0 \tag{9}$$

und

$$\sum_k E(|x_{nk}^t - x_{nk}^{d(n)}|) \to 0 \tag{10}$$

gelten.

2. *für die* ZV

$$\xi_{nk} = x_{nk}^{d(n)} - E\, x_{nk}^{d(n)} \tag{11}$$

gelte die Beziehung (2) *und die Voraussetzung* (*I*) *von Satz 1.*

Ist 1. *und* 2. *erfüllt, so konvergieren die* VF *von*

$$\sum_k (x_{nk} - E(x_{nk}^{d(n)})) \tag{12}$$

gegen N(0,1), *und die* ξ_{nk} *sind infinitesimal.*

B e m e r k u n g: Da die VF der ξ_{nk} vollständig durch die der x_{nk} bestimmt sind, sind auch die Abhängigkeitsverhältnisse der ξ_{nk} durch die der x_{nk} bestimmt. Es genügt also, wenn die Voraussetzung (*I*) in den ξ_{nk} unter Zugrundelegung der Abhängigkeitsverhältnisse der x_{nk} erfüllt ist.

Beweis: 1. Gilt (8), so stimmen die Grenzverteilungen von $\sum_k \xi_{nk}$ und von (12) überein. Der Rest folgt aus Satz 1.

2. Um zu zeigen, daß 1. aus 1'. folgt, setzt man $x_{nk} = x_{nk}^t + y_{nk}$. Ist $\delta > 0$ beliebig, so erhält man durch wiederholte Anwendung der für hinreichend große k_n gültigen Ungleichung

$$\begin{aligned} P(|\sum_{k=1}^{k_n} y_{nk}| > \delta) &= P(|\sum_{k=1}^{k_n} y_{nk}| > \delta \;\&\; |y_{n1}| \leqq \delta/k_n) + \\ &+ P(|\sum_{k=1}^{k_n} y_{nk}| > \delta \;\&\; |y_{n1}| > \delta/k_n) \leqq \\ &\leqq P(|\sum_{k=2}^{k_n} y_{nk}| > \frac{k_n - 1}{k_n}\delta) + P(|x_{n1}| > t) \end{aligned} \tag{13}$$

die obere Schranke $\sum_k P(|x_{nk}| > t)$ für (13), die wegen (9) gegen Null strebt. Andererseits folgt aus (10)

$$P(|\sum_k x^t_{nk} - x^{d(n)}_{nk}| > \delta) \leqq \delta^{-1} \sum_k E(|x^t_{nk} - x^{d(n)}_{nk}|) \to 0 .$$

Die Voraussetzung (I) von Satz 1 ist in praktischen Fällen oft schwer zu handhaben. Sie läßt sich auf mannigfache Weise verschärfen und dabei vereinfachen. Dazu sei mit den Größen (11) von Satz 2 für ein beliebiges Paar abhängiger ZV (ξ_{nu}, ξ_{nv}) $(1 \leqq u,\ v \leqq k_n;\ n = 1, 2, \ldots)$ unter $T_{ns}(u, v)$ die Anzahl aller Mengen $\in U_{ns}$ (definiert in 1.) verstanden, die das Paar enthalten, und es sei

$$T_{ns} = \max_{1 \leqq u, v \leqq k_n} T_{ns}(u, v) .$$

Sei ferner für $s \geqq 3$ und $n = 1, 2, \ldots$ U'_{ns} eine Teilmenge von U_{ns} mit der Eigenschaft, daß alle ZV in den (irreduziblen) Mengen $(\xi_{nj_1}, \ldots, \xi_{nj_s}) \in U'_{ns}$ paarweise unabhängig sind. Dann gilt z.B.

Satz 3. *Es gelte die Voraussetzung* 1. *oder* 1'. *von Satz 2 für eine Folge* $\{d(n)\}$ *und*

$$T_{ns}\, d(n)^{s-2} \sum_{U_{n2}} E(|\xi_{nu}\xi_{nv}|) \to 0 \qquad \text{für } s \geqq 3 . \tag{14}$$

Ferner gelte die Beziehung (3) *der Voraussetzung* (I) *oder statt dessen*

$$\sum_{U'_{ns}} |E(\prod_{t=1}^{s} \xi_{nj_t})| \to 0 \qquad \text{für } s \geqq 3 . \tag{15}$$

Gilt schließlich die Beziehung (2), *so konvergieren die* VF *von* (12) *gegen* N(0,1).

Beweis: In jedem Summanden von (4) kommt wegen β) der Erwartungswert des Produkts der ZV eines irreduziblen Paares vor. Diejenige Teilsumme von (4), deren Terme alle ein bestimmtes solches Paar (ξ_{nu}, ξ_{nv}) enthalten, wird unter Beachtung der Definition der ξ_{nk} (siehe (11)) durch $2\,T_{ns} d(n)^{s-2} E(|\xi_{nu}\xi_{nv}|)$ majorisiert. Also folgt (4) aus (14). — Spaltet man auf (3) in die Summe über U'_{ns} und $U_{ns} - U'_{ns}$, so geht die erste Summe wegen (15) und die zweite wegen (14) gegen Null.

Bemerkung 3.1. Sind alle Voraussetzungen des Satzes 3 erfüllt, so folgt

$$\lim_{n \to \infty} d(n) = 0 , \tag{16}$$

denn es gilt stets $T_{ns} \geqq 1$, und die Summe in (14) ist nicht kleiner als die Summe (2). Ist ferner die Voraussetzung 1'. erfüllt, so folgt

$$\sum_k P(|x_{nk}| > \varepsilon) \to 0 \tag{17}$$

für alle $\varepsilon > 0$, wie man unter Beachtung von (16) leicht nachrechnet. Dann sind nicht nur die ξ_{nk}, sondern auch die x_{nk} infinitesimal.

Bemerkung 3.2. Der Summationsbereich in (15) läßt sich noch weiter verkleinern, in dem man Teilmengen $U''_{ns} \subset U'_{ns}$ mit der Eigenschaft einführt, daß je drei ZV in den Mengen $(\xi_{nj_1}, \ldots, \xi_{nj_s}) \in U''_{ns}$ untereinander unabhängig sind. Ist für ein beliebiges Tripel abhängiger, jedoch paarweise unabhängiger ZV $\{\xi_{nu}, \xi_{nv}, \xi_{nw}\}$ $T'_{ns}(u, v, w)$ die Anzahl aller Mengen aus $U'_{ns}(s \geqq 3)$, die das

Tripel enthalten, und

$$T'_{ns} = \max_{u,v,w} T'_{ns}(u, v, w),$$

so folgt (15) aus

$$\sum_{U''_{ns}} |E(\prod_{t=1}^{s} \xi_{nj_t})| \to 0$$

und

$$T'_{ns} d(n)^{s-3} \sum_{U'_{ns}-U''_{ns}} E(|\xi_{nu}\xi_{nv}\xi_{nw}|) \to 0 .$$

Entsprechende weitere Reduktionsschritte führen schließlich zu Mengen $U_{ns}^{(s-2)} \subset U_{ns}$, deren Elemente $(\xi_{nj_1}, \ldots, \xi_{nj_s})$ irreduzible Mengen von ZV sind derart, daß die ZV aller Teilmengen mit bis zu $s - 1$ Variablen untereinander unabhängig sind.

Folgerung. Sind die ZV x_{nk} in jeder Zeile der Doppelfolge $\{x_{nk}\}$ unabhängig, so erhält man aus Satz 3 (mit Voraussetzung 1'.) den klassischen Zentralen Grenzwertsatz in der Formulierung von LINDEBERG und FELLER (siehe z.B. [*3*], S. 316). Wegen (2) und Bemerkung 3.1 folgen die Voraussetzungen des letzteren aus denjenigen des Satzes 3. Wenn umgekehrt (17) für alle $\varepsilon > 0$ gilt, so gibt es eine Nullfolge $\{d(n)\}$, für die (17) gilt. Wegen $T_{ns} \equiv 1$ bei unabhängigen x_{nk} ist dann (14) mit diesen $d(n)$ erfüllt. Ferner läßt sich ein $t > 0$ finden, so daß (9) und (10) gilt. Schließlich folgt (15) aus der Infinitesimalität der ξ_{nk}. — GODWIN und ZAREMBA haben ebenfalls auf die Identität des von ihnen angegebenen Satzes (siehe Satz 4 weiter unten) mit dem klassischen Zentralen Grenzwertsatz im Falle unabhängiger x_{nk} hingewiesen.

Mit Hilfe des Begriffes der *Verkettung* von ZV (linkedness; [*2*]) läßt sich durch eine Verschärfung der Voraussetzung (I) erreichen, daß die Bedingungen der vorstehenden Sätze nur noch die gemischten zweiten absoluten Momente betreffen.

Definition. Sei M eine Menge von ZV. Eine symmetrische Funktion $L(x_i, x_k)$, die auf den Elementen $x_i, x_k \in M$ definiert ist und die Werte 0 und 1 annimmt, heißt eine *Verkettung*, wenn sie folgende Eigenschaft hat: sind M_1 und M_2 beliebige Untermengen von M, so folgt aus $L(x_i, x_k) = 0$ für alle $x_i \in M_1$ und alle $x_k \in M_2$ die Unabhängigkeit von M_1 und M_2. Ist $L(x_i, x_k) = 1$, so heißt das ZV-Paar x_i, x_k *verkettet*, andernfalls *unverkettet*.

Sind M_1 und M_2 abhängig, so gibt es mindestens ein verkettetes Paar $x_i \in M_1$, $x_k \in M_2$. Ein irreduzibles Paar ist demnach stets verkettet, insbesondere ist $L(x_i, x_i) = 1$. Im allgemeinen ist die Verkettung, die die Abhängigkeitsverhältnisse der ZV in einer Menge M beschreibt, nicht eindeutig bestimmt. Ist M endlich, so existiert jedoch eine Minimalverkettung, definiert durch $\sum_{i,k} L(x_i, x_k) = \min$. Doch auch eine Minimalverkettung einer Menge M braucht nicht eindeutig zu sein.

Besitzt eine Menge M die I-Zerlegung $M_\alpha, M_\beta, \ldots$, wo $\alpha, \beta, \ldots$ einer abstrakten Indexmenge A angehören, so gibt es eine Verkettung L mit $L(x_i, x_k) = 0$ für alle $x_i \in M_\alpha$, $x_k \notin M_\alpha$, $\alpha \in A$. Innerhalb einer irreduziblen Menge M_α ist jede ZV mit mindestens einer anderen verkettet. Besitzt M_α v Elemente, so gibt es zu jedem Paar $x_i, x_k \in M_\alpha$ eine Kette von ZV $x_i, x_{i_1}, \ldots, x_{i_{v-2}}, x_k \in M_n$, in der aufeinanderfolgende ZV verkettet sind; in der Kette braucht keine ZV doppelt aufzutreten.

Satz 4 [2]. *Mit den Bezeichnungen von Satz 2 sei für ein beliebiges verkettetes Paar* (x_{nu}, x_{nv}), $1 \leqq u$, $v \leqq k_n$, $n = 1, 2, \ldots$, $\Theta_{ns}(u, v)$ *die Anzahl derjenigen Mengen* $\in U_{ns}$, *die das Paar enthalten. Es gebe eine Funktion* $\Theta(n)$ *und eine Zahl* c, $2 \leqq c \leqq 3$, *so daß für* $s = 3, 4, \ldots$ *jeweils für* $n \to \infty$ *die Abschätzungen*

$$\max_{u,v} \Theta_{ns}(u, v) = O(\Theta(n)^{s-c}) \tag{18}$$

gelten. Es seien ferner die Beziehungen (9) *und* (10) *für ein* $t > 0$ *und für* $d(n) = \eta/\Theta(n)$ *bei beliebigem* $\eta > 0$ *erfüllt. Außerdem gelte für die durch* (11) *definierten* ξ_{nk} *die Beziehung* (2) *von Satz 1 sowie*

$$\Theta(n)^{2-c} \sum_{u,v} E(|\xi_{nu}\xi_{nv}|) = O(1), \tag{19}$$

wo die Summe über alle verketteten Paare aus der Menge $\xi_{n1}, \ldots, \xi_{nk_n}$ *zu erstrecken ist. Dann konvergieren die* VF *von*

$$\sum_{k=1}^{k_n} (x_{nk} - E(x^{1/\Theta(n)}))$$

gegen N(0,1).

Beweis: Es braucht nur gezeigt zu werden, daß die Voraussetzung (I) von Satz 1 aus (19) folgt. In jeder Menge $(\xi_{nj_1}, \ldots, \xi_{nj_s}) \in U_{ns}$, über die in (3) und (4) summiert wird, gibt es ein verkettetes Paar. Für (3) und (4) ergeben sich dann für $s = 3, 4, \ldots$ und beliebiges $\eta > 0$ die Abschätzungen, die jeweils für $n \to \infty$ gelten,

$$O(d(n)^{s-2}\Theta(n)^{s-c} \sum_{u,v} E(|\xi_{nu}\xi_{nv}|)) = \eta^{s-2} O(\Theta(n)^{2-c} \sum_{u,v} E(|\xi_{nu}\xi_{nv}|)) = \eta^{s-2} O(1).$$

Bemerkung 4.1. Da notwendig $d(n) \to 0$, gilt $\Theta(n) \to \infty$.

Die Einführung der Zahl c gestattet eine gewisse Freiheit in der Wahl der Funktion $\Theta(n)$. Dadurch wird der Satz flexibler gegenüber gewissen Anwendungen auf nichtparametrische Probleme (vgl. dazu [2], S. 684). Wegen weiterer Anwendungen in der Statistik, insbesondere auf Folgen m-abhängiger ZV (vgl. auch weiter unten), wird ebenfalls auf [2] sowie auf [5] verwiesen. Auf Grund der Bemerkung 1.1 lassen sich jedoch die verschiedentlich in der Literatur betrachteten Zentralen Grenzwertsätze für Summe von ZV ξ_k von der Gestalt

$$\xi_k = \operatorname*{l.\,i.\,m.}_{n\to\infty} \sum_{j=1}^{n} c_j x_{k+j}, \qquad k = 1, 2, \ldots,$$

wo die c_j reelle Konstante mit $\sum_j c_j^2 < \infty$ und die x_j unabhängige ZV mit $\sup_j E x_j^2 < \infty$ sind, nicht ohne weiteres aus den obigen Sätzen folgern (vgl. [1]).

Anwendung von Satz 4. Der Begriff der Verkettung erweist sich zur Herleitung eines Grenzwertsatzes für m-abhängige ZV als sehr vorteilhaft. Man nennt die ZV einer Doppelfolge $\{x_{nk}\}$ *m-abhängig*, wo m eine gegebene ganze nichtnegative Zahl ist, wenn für alle n und alle s, t mit $1 \leqq s < t \leqq k_n$, $t - s > m$ gilt: $(x_{n1}, \ldots, x_{ns})$ ist unabhängig von $(x_{nt}, \ldots, x_{nk_n})$. Offenbar genügt es, jede ZV einer Zeile nur mit den m vorangehenden und den m nachfolgenden ZV (soweit sie vorhanden sind) als verkettet anzusehen. Jede Menge $\in U_{ns}$ ist in einer Menge enthalten, die aus $m(s-1)+1$ aufeinanderfolgenden ZV einer Zeile besteht. Eine grobe Abschätzung von $\max\limits_{u,v} \Theta_{ns}(u,v)$ ist daher $O\{(2m(s-1))^{s-2}\}$, und in (18) kann $\Theta(n)$ als eine Konstante und c beliebig aus dem Intervall [2, 3] gewählt werden. Die Beziehung (10) braucht dann nur für $d(n)=\eta$, alle n, mit beliebigem $\eta>0$ erfüllt zu sein. Beziehung (19) ist äquivalent mit

$$\sum_{k=1}^{k_n} E(\xi_{nk}^2) = O(1). \tag{20}$$

Damit hat man die Voraussetzungen des Satzes von Orey (1958) für m-abhängige ZV gewonnen, dem bislang allgemeinsten Satz dieser Art. Die Bedingung (20) ist jedoch keinesfalls notwendig. In der Tat lassen sich Verallgemeinerungen des Oreyschen Satzes angeben, in denen sie abgeschwächt ist — in dem folgenden Satz 5 beispielsweise zu der Bedingung (25).

3.

Satz 5. *Es sei $\{x_{nk}\}$ eine Doppelfolge m-abhängiger ZV, wo m eine nichtnegative ganze Zahl ist. Es gebe eine bestimmte divergente Folge natürlicher Zahlen $\{m_n;\, n=1,2,\ldots\}$, so daß $k_n m_n^{-1}\to\infty$ und*

$$\sum_k P(|x_{nk}|>\varepsilon/m_n)\to 0 \tag{21}$$

für alle $\varepsilon>0$. (Alle Grenzübergänge gelten für $n\to\infty$.) Es gebe eine reelle Zahl $p>0$, für die unter Verwendung der in der Vorbemerkung zu Satz 2 eingeführten Bezeichnung x^p für die abgeschnittenen (truncated) ZV die folgenden Beziehungen bestehen:

$$\sum_k E(x_{nk}^p)\to\alpha, \tag{22}$$

$$\sum_{k,j}\operatorname{cov}(x_{nk}^p, x_{nj}^p)\to\sigma^2. \tag{23}$$

Es sei

$$z_{nv} = x_{n,v+1}^p + x_{n,v+2}^p + \cdots + x_{n,v+m} \tag{24}$$

$(v=0,1,\ldots,k_n-m)$, ferner $z_n(i)\in\{z_{vn};\ i\,m_n\leqq v<(i+1)\,m_n-2m\}$ so, daß

$$\operatorname{var}(z_n(i)) = \min_{im_n\leq v<(i+1)m_n-2m}\operatorname{var}(z_{vn}) \tag{24'}$$

$(i=0,1,\ldots,[(k_n-m)\,m_n^{-1}]\equiv D_n)$. Gilt dann noch

$$\sum_{i=0}^{D_n}\operatorname{var}(z_n(i))\to 0, \tag{25}$$

so folgt

$$\mathrm{VF}\Big(\sum_k x_{nk}\Big)\to N(\alpha,\sigma). \tag{26}$$

Der Satz läßt sich auf die folgende elementare Weise unabhängig von den vorangehenden Sätzen und dem ursprünglichen Oreyschen Beweis durch Rückführung auf den Lindeberg-Fellerschen Zentralen Grenzwertsatz ableiten.

Beweis: Analog der Abschätzung (13) im Beweis von Satz 2 gilt

$$P(|\sum_k (x_{nk} - x^p_{nk})| > \delta) \leqq \sum_k P(|x_{nk}| > p) \to 0$$

für alle $\delta > 0$ wegen (21). Es genügt also, (26) für die Doppelfolge $\{x^p_{nk}\}$ statt für $\{x_{nk}\}$ zu beweisen. Wegen (22) genügt es ferner,

$$\mathrm{VF}(\sum_k v_{nk}) \to \mathrm{N}(0, \sigma), \quad v_{nk} = x^p_{nk} - E\, x^p_{nk},$$

zu zeigen. Nun gilt wegen (25) und der Unabhängigkeit der $z_n(i)$

$$\sum_i (z_n(i) - E\, z_n(i)) \to 0 \tag{27}$$

nach Wahrscheinlichkeit.

Die Abschnitte der ZV in den Zeilen der Doppelfolge $\{x_{nk}\}$, die durch die Abschnitte der in den $z_n(i)$ aufaddierten ZV getrennt werden, seien $\{x_{nk}, k \in S_{ni}\}$, $i = 0, \ldots, D_n + 1$. Wegen (27) genügt es, damit (26) gilt,

$$\mathrm{VF}(\sum_i (\sum_{k \in S_{ni}} v_{nk})) \to \mathrm{N}(0, \sigma)$$

zu zeigen.

Da (21) für alle $\varepsilon > 0$ gilt, gibt es eine Folge $\{m'_n\}$ mit $m_n = o(m'_n)$, so daß

$$\sum_k P(|x_{nk}| > 1/m'_n) \to 0$$

gilt. Setzt man $t_{nk} = x^p_{nk} - x^{1/m'_n}_{nk}$, so folgt

$$\sum_k E(|t_{nk}|) \to 0 \tag{28}$$

und daraus (vgl. den Beweis von Satz 2)

$$P(|\sum_i (\sum_{k \in S_{ni}} t_{nk})| > \delta) \to 0$$

für alle $\delta > 0$. Es genügt mithin schließlich für (26)

$$\mathrm{VF}(\sum_i w_{ni}) \to \mathrm{N}(0, \sigma),$$

wo

$$w_{ni} = \sum_{k \in S_{ni}} v'_{nk}, \quad v'_{nk} = x^{1/m'_n}_{nk} - E(x^{1/m'_n}_{nk}).$$

gesetzt ist.

Die Doppelfolge $\{w_{ni}\}$ genügt nun dem klassischen Zentralen Grenzwertsatz (vgl. etwa [3], S. 316): die w_{ni} sind in jeder Zeile unabhängig, und es gilt $E\, w_{ni} = 0$ für alle n und i. Ferner gilt, wieder analog zu (13), für beliebiges $\varepsilon > 0$ und genügend große n unter Beachtung von (21) und der Tatsache, daß in den Mengen S_{ni} höchstens $2 m_n$ Indizes vorkommen können,

$$\sum_i P(|w_{ni}| > \varepsilon) \leqq \sum_i (\sum_{k \in S_{ni}} P(|v'_{nk}| > \varepsilon/2 m_n)) \leqq$$

$$\leqq \sum_{k=1}^{k_n} P(|x^{1/m'_n}_{nk}| > (\varepsilon/2 m_n) - (1/m'_n)) < \sum_{k=1}^{k_n} P(|x_{nk}| > \varepsilon/3 m_n) \to 0.$$

Schließlich erhält man mit $v_{nk} = v'_{nk} + (t_{nk} - E\,t_{nk})$ für die linke Seite von (23)

$$\sum_{k,j} \{\operatorname{cov}(v'_{nk}, v'_{nj}) + 2\operatorname{cov}(v'_{nk}, t_{nj}) + \operatorname{cov}(t_{nk}, t_{nj})\}; \tag{29}$$

bei festgehaltenem j braucht nur über $j - m \leqq k \leqq j + m$ summiert zu werden. Wegen $|\operatorname{cov}(v'_{nk}, t_{nj})| \leqq (4/m'_n) E(|t_{nj}|)$ und (28) strebt die Doppelsumme über die zweiten Terme gegen Null; ähnlich zeigt man, daß die Doppelsumme über die dritten Terme gegen Null strebt. Folglich besitzt (29) den Grenzwert σ^2, und der Beweis ist vollständig.

Bemerkung 5.1. Der Satz von Orey, in dem, wie zuvor bemerkt, statt (21)

$$\sum_k P(|x_{nk}| > \varepsilon) \to 0 \tag{30}$$

für alle $\varepsilon > 0$ und statt (25)

$$\sum_k \operatorname{var}(x^p_{nk}) = O(1) \tag{31}$$

sowie ferner (22) und (23) gefordert wird, folgt aus Satz 5. Erfüllen nämlich die m-abhängigen ZV $\{x_{nk}\}$ die Beziehungen (30) und (31), so gibt es stets eine Folge $\{m_n\}$ mit den in Satz 5 verlangten Eigenschaften. Wegen

$$\operatorname{var} z_{nv} \leqq \Big(\sum_{k=v+1}^{v+m} (\operatorname{var} x^p_{nk})^{1/2}\Big)^2 \leqq m^2 \sum_{k=v+1}^{v+m} \operatorname{var} x^p_{nk},$$

und wegen (24′) gilt

$$m_n \sum_{i=0}^{D_n} \operatorname{var}(z_n(i)) = O\Big(\sum_{k=1}^{k_n} \operatorname{var} x^p_{nk}\Big) = O(1),$$

woraus (25) folgt.

Den Herren Professoren Wassily Hoeffding und S. K. Zaremba bin ich für verschiedene Anregungen und kritische Bemerkungen im Zusammenhang mit der vorliegenden Arbeit zu besonderem Dank verpflichtet.

Literatur

[1] Eicker, F.: Ein Zentraler Grenzwertsatz für Summen von Variablen einer verallgemeinerten linearen Zufallsfolge. Z. Wahrscheinlichkeitstheorie (erscheint 1965).

[2] Godwin, H. J., and S. K. Zaremba: A central limit theorem for partly dependent variables. Ann. math. Statistics **32**, 677–686 (1961).

[3] Loève, M.: Probability theory. 2. Aufl. Princeton: Van Nostrand 1960.

[4] Orey, S.: A central limit theorem for m-dependent random variables. Duke math. J. **25**, 543–546 (1958).

[5] Zaremba, S. K.: Note on the Central Limit Theorem. Math. Zeitschrift **69**, 295–298 (1958).

Institut für Angewandte Mathematik
der Albert-Ludwigs-Universität
78 Freiburg i. Br.. Hebelstraße 40

(Eingegangen am 2. Juli 1963)