Kapitel VIII

Der zentrale Grenzwertsatz

Reinhard Höpfner

Vorlesung Stochastik I

Wintersemester 2003/2004 , Sommer/Winter 2006/2007, Sommersemester 2010

Institut für Mathematik, Johannes Gutenberg Universität Mainz

25.11.03, 27.10.06, 22.07.10

Übersicht zu Kapitel VIII :

8.1 Definition : Ein *Dreiecksschema* $(X_{n,j})_{j=1,\ldots,k_n\,;\,n=1,2,\ldots}$ *von zeilenweise unabhängigen* L^2-*Variablen* ist ein System

$$\begin{array}{llllll} X_{1,1} & \ldots & X_{1,k_1} & & & \\ X_{2,1} & \ldots & \ldots & X_{2,k_2} & & \\ \vdots & & & \ddots & & \\ X_{n,1} & \ldots & \ldots & \ldots & \ldots & X_{n,k_n} \\ \vdots & & & & & \ddots \end{array}$$

von Zufallsvariablen auf $(\Omega, \mathcal{A}, P)$, mit Werten in $(I\!R, \mathcal{B}(I\!R))$, mit den folgenden Eigenschaften:
i) für die Folge $(k_n)_n$ der Zeilenlängen gilt $k_n \uparrow \infty$,
ii) in jeder Zeile $n \in I\!N$ sind die Zufallsvariablen $X_{n,1}, \ldots, X_{n,k_n}$ unabhängig unter P,
iii) es gilt $X_{n,j} \in L^2(P)$ für alle $1 \leq j \leq k_n$, alle $n \in I\!N$.
In einem Dreiecksschema zeilenweise unabhängiger L^2-Variablen bezeichne

$$\begin{aligned} \mu_{n,j} &:= E(X_{n,j}) \,, \quad \sigma^2_{n,j} := \mathrm{Var}(X_{n,j}) \,, \quad 1 \leq j \leq k_n \,,\ n \geq 1 \,, \\ s_n &:= \sqrt{\mathrm{Var}(X_{n,1} + \ldots + X_{n,k_n})} \;=\; \sqrt{\sigma^2_{n,1} + \ldots + \sigma^2_{n,k_n}} \,,\ n \geq 1 \,. \end{aligned}$$

Ein Dreiecksschema mit

$$\mu_{n,j} \;=\; 0 \,, \quad 1 \leq j \leq k_n \,,\ n \geq 1$$

nennt man *zentriert*; es heißt *standardisiert* falls

$$s_n^2 \;=\; 1 \quad \text{für alle } n \geq 1.$$

8.2 Beispiel: Sei $(Y_j)_{j\geq 1}$ eine Folge von iid Zufallsvariablen mit $E(Y_1^2) < \infty$, reellwertig, definiert auf irgendeinem $(\Omega, \mathcal{A}, P)$.
a) $(Y_j)_{j\geq 1}$ liefert via

$$\begin{array}{llll} Y_1 & & & \\ Y_1 & Y_2 & & \\ \vdots & & \ddots & \\ Y_1 & Y_2 & \ldots & Y_n \\ \vdots & & & \ddots \end{array}$$

ein Dreiecksschema zeilenweise unabhängiger L^2-Variabler.

b) Wende mit $\mu := E(Y_1)$ und $\sigma^2 := Var(Y_1)$ auf die n-te Zeile des Dreiecksschemas die lineare Transformation $y \to \frac{1}{\sqrt{n}} \frac{y-\mu}{\sigma}$ an und definiere

$$X_{n,j} := \frac{1}{\sqrt{n}} \frac{Y_j - \mu}{\sigma} , \quad 1 \leq j \leq n , \; n \geq 1 .$$

Das so entstandene Dreiecksschema $(X_{n,j})_{j=1,\dots,n\,;\,n=1,2,\dots}$ (mit $k_n = n$) ist zentriert und standardisiert.

c) In der 'Einführung in die Stochastik' lernt man den Satz von de Moivre-Laplace (entstanden um 1733) kennen, der im Spezialfall binomialverteilter ZV $(Y_j)_{j\geq 1}$ iid $\sim \mathcal{B}(1,p)$ zeigt

$$P\left(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \frac{Y_j - p}{\sqrt{p(1-p)}} \leq b \right) \longrightarrow \Phi(b) , \quad n \to \infty$$

für jedes $b \in I\!R$; dabei bezeichnet $\Phi(t) = \mathcal{N}(0,1)((-\infty, t])$ die Verteilungsfunktion der eindimensionalen Standardnormalverteilung. Wegen 6.12 ist dies – im Spezialfall binomialverteilter iid ZV – eine Aussage über *Verteilungskonvergenz von Zeilensummen* in einem zentrierten und standardisierten Dreiecksschema. □

Ziel dieses Kapitels ist es, hinreichende (und notwendige) Bedingungen anzugeben, unter denen Zeilensummen in zentrierten und standardisierten Dreiecksschemata aus zeilenweise unabhängigen L^2-Variablen für $n \to \infty$ schwach in $I\!R$ gegen die Standardnormalverteilung konvergieren. Sätze dieser Bauart – wie sie auch in viel allgemeineren Situationen, etwa in stochastischen Prozessen, formuliert werden können – bezeichnet man mit dem Namen *zentraler Grenzwertsatz.*

8.3 Definition: Betrachte ein Dreiecksschema $(X_{n,j})_{j=1,\dots,k_n\,;\,n=1,2,\dots}$ aus zeilenweise unabhängigen L^2-Variablen, und mögliche Eigenschaften dieses Schemas.

i) Die *asymptotische Vernachlässigbarkeitsbedingung* (AV) verlangt

$$\forall\, \varepsilon > 0 : \quad \max_{1\leq j \leq k_n} P\left(\left| \frac{X_{n,j} - \mu_{n,j}}{s_n} \right| > \varepsilon \right) \longrightarrow 0 , \quad n \to \infty .$$

ii) Die *Feller-Bedingung* (FE) verlangt

$$\max_{1\leq j\leq k_n} \frac{\sigma^2_{n,j}}{s^2_n} \longrightarrow 0 , \quad n \to \infty .$$

iii) Die *Lindeberg-Bedingung* (LIN) verlangt

$$\forall\, \varepsilon > 0 \text{ fest:} \quad L_n(\varepsilon) := \sum_{j=1}^{k_n} E\left(\left(\frac{X_{n,j} - \mu_{n,j}}{s_n} \right)^2 1_{\{|\frac{X_{n,j}-\mu_{n,j}}{s_n}| > \varepsilon\}} \right) \quad \stackrel{n\to\infty}{\longrightarrow} \quad 0 .$$

iv) Die *Lyapunov-Bedingung* (LY) verlangt

$$\begin{cases} \exists\, \delta > 0 \quad \text{so daß} \quad X_{n,j} \in L^{2+\delta}(P) \quad \text{für alle } j,n, \text{ und es gilt} \\ \sum\limits_{j=1}^{k_n} E\left(\left|\frac{X_{n,j}-\mu_{n,j}}{s_n}\right|^{2+\delta}\right) \longrightarrow 0 \quad \text{für } n \to \infty\,. \end{cases}$$

Beachte: alle diese Bedingungen sind invariant unter linearen Transformationen, die auf alle Variablen *einer ganzen Zeile* des Dreiecksschemas angewandt werden.

8.4 Hilfssatz: In einem Dreiecksschema $(X_{n,j})_{j=1,\ldots,k_n\,;\,n=1,2,\ldots}$ aus zeilenweise unabhängigen L^2-Variablen gilt für die in 8.3 definierten Eigenschaften:

a) (LIN) $\Longrightarrow$ (FE) $\Longrightarrow$ (AV);

b) falls sogar $X_{n,j} \in L^{2+\delta}(P)$ für alle j,n: (LY) $\Longrightarrow$ (LIN).

Beweis: Da $X_{n,j} \in L^2(P)$ für alle j,n, gilt für jedes feste $\varepsilon > 0$

$$\begin{aligned} \sigma^2_{n,j} &= \mathrm{Var}(X_{n,j}) \;=\; s_n^2\, E\left(\left(\frac{X_{n,j}-\mu_{n,j}}{s_n}\right)^2\right) \\ &\le s_n^2 \left[\varepsilon^2 + E\left(\left(\frac{X_{n,j}-\mu_{n,j}}{s_n}\right)^2 1_{\{|\frac{X_{n,j}-\mu_{n,j}}{s_n}|>\varepsilon\}}\right)\right] \end{aligned}$$

und damit

$$\max_{1\le j\le k_n} \frac{\sigma^2_{n,j}}{s_n^2} \;\le\; \varepsilon^2 + L_n(\varepsilon)\,.$$

Dies zeigt (LIN) $\Longrightarrow$ (FE). Die elementare Chebychev-Ungleichung

$$P(|X_{n,j}-\mu_{n,j}| > \varepsilon s_n) \;\le\; \frac{\sigma^2_{n,j}}{\varepsilon^2 s_n^2}$$

liefert (FE) $\Longrightarrow$ (AV). Unter der Zusatzvoraussetzung $X_{n,j} \in L^{2+\delta}(P)$ für alle j,n folgt aus

$$|y|^2\, 1_{\{|y|>\varepsilon\}} \;\le\; \mathrm{cst}(\delta,\varepsilon)\, |y|^{2+\delta}\,, \quad y \in I\!R$$

die Implikation (LY) $\Longrightarrow$ (LIN). □

8.5 Beispiel: Betrachte eine Folge $(Y_j)_{j\ge1}$ von iid ZV mit $E(Y_1^2) < \infty$ wie in 8.2 als Dreiecksschema zeilenweise unabhängiger L^2-Variablen, entweder in der Form aus 8.2 a) oder in der zentrierten und standardisierten Form aus 8.2 b):

$$
\begin{array}{llllcllll}
Y_1 & & & & \qquad & X_{1,1} & & & \\
Y_1 & Y_2 & & & & X_{2,1} & X_{2,2} & & \\
\vdots & & \ddots & & & \vdots & & \ddots & \\
Y_1 & Y_2 & \dots & Y_n & & X_{n,1} & X_{n,2} & \dots & X_{n,n} \\
\vdots & & & \ddots & & \vdots & & & \ddots
\end{array}
$$

mit

$$X_{n,j} \;:=\; \frac{1}{\sqrt{n}} \frac{Y_j - \mu}{\sigma} \;, \quad \mu := EY_1 \;,\; \sigma^2 := \mathrm{Var}(Y_1) \;.$$

a) Dann gilt die Lindeberg-Bedingung (LIN) (und damit nach 8.4 auch (FE) und (AV)):
Mit den Bezeichnungen aus 8.1 für $(X_{n,j})_{j=1,2,\dots\,;\,n=1,2,\dots}$ hat man

$$s_n = 1 \text{ für alle } n, \;, \mu_{n,j} = 0 \text{ für alle } n,\, j,$$

also schreibt sich der in der Lindeberg-Bedingung zu betrachtende Ausdruck in jeder der Formen

$$
\begin{aligned}
L_n(\varepsilon) &= \sum_{j=1}^{n} E\left(\left(\frac{X_{n,j}-\mu_{n,j}}{s_n}\right)^2 1_{\{|X_{n,j}-\mu_{n,j}|>\varepsilon s_n\}}\right) = \sum_{j=1}^{n} E\left((X_{n,j})^2\, 1_{\{|X_{n,j}|>\varepsilon\}}\right) \\
&= \sum_{j=1}^{n} E\left(\left(\frac{Y_j-\mu}{\sqrt{n}\,\sigma}\right)^2 1_{\{|\frac{Y_j-\mu}{\sqrt{n}\,\sigma}|>\varepsilon\}}\right) = \frac{1}{\sigma^2}\, E\left((Y_1-\mu)^2\, 1_{\{\,|Y_1-\mu|>\varepsilon\sigma\sqrt{n}\,\}} \right) \;.
\end{aligned}
$$

Da $Y_1 \in L^2(P)$, strebt der letzte Ausdruck gegen 0 für $n \to \infty$: damit ist (LIN) nachgewiesen.
b) Gilt sogar $Y_1 \in L^{2+\delta}(P)$ für ein $\delta > 0$, gilt auch die Lyapunov-Bedingung (LY) wegen

$$\sum_{j=1}^{n} E\left(\left|\frac{X_{n,j}-\mu_{n,j}}{s_n}\right|^{2+\delta}\right) \;=\; n\cdot E\left(\left|\frac{Y_1-\mu}{\sqrt{n}\,\sigma}\right|^{2+\delta}\right) \;=\; n^{-\delta/2}\cdot \mathrm{cst} \quad\longrightarrow\quad 0$$

für $n \to \infty$. □

Nun geben wir die erste Formulierung des 'zentralen Grenzwertsatzes':

8.6 Hauptsatz: Sei $(X_{n,j})_{j=1,\dots,k_n\,;\,n=1,2,\dots}$ ein Dreiecksschema aus zeilenweise unabhängigen L^2-Variablen. Erfüllt dieses Dreiecksschema die Lindeberg-Bedingung (LIN), so gilt

$$\mathcal{L}\left(\sum_{j=1}^{k_n} \frac{X_{n,j}-\mu_{n,j}}{s_n} \mid P\right) \quad \xrightarrow[n\to\infty]{w} \mathcal{N}(0,1)$$

mit $\mu_{n,j} = E(X_{n,j})$ und $s_n^2 = \sum\limits_{j=1}^{k_n} \mathrm{Var}(X_{n,j})$.

Beweis: 1) Wir betrachten o.E. ein zentriertes und standardisiertes Dreiecksschema $(X_{n,j})_{n,j}$. Die Lindeberg-Bedingung (und die anderen in 8.4 betrachteten Bedingungen) sind invariant unter dieser Umformung. Im Fall

$$(*) \qquad \mu_{n,j} = 0 \quad \forall\, n, j, \quad s_n^2 = \sigma_{n,1}^2 + \ldots \sigma_{n,k_n}^2 = 1 \quad \forall\, n \geq 1$$

beweisen wir dann asymptotische Normalität der Zeilensummen

$$\sum_{j=1}^{k_n} X_{n,j} \xrightarrow[n\to\infty]{\mathcal{L}} \xi\ , \quad \xi \sim \mathcal{N}(0,1)$$

unter der Voraussetzung (LIN):

$$L_n(\varepsilon) := \sum_{j=1}^{k_n} E\left(X_{n,j}^2 1_{\{|X_{n,j}|>\varepsilon\}}\right) \stackrel{n\to\infty}{\longrightarrow} 0\ , \quad \forall\, \varepsilon > 0\ .$$

Schreibe $Q_{n,j} = \mathcal{L}(X_{n,j}|P)$, sei $\varphi_{n,j}$ die zugehörige charakteristische Funktion.
2) Wegen $X_{n,j} \in L^2(P)$ ist $\varphi_{n,j}$ zweimal stetig differenzierbar an der Stelle 0:

$$\varphi_{n,j}^{(r)}(t) = \int (ix)^r e^{itx} Q_{n,j}(dx)\ , \quad \varphi_{n,j}^{(r)}(0) = i^r E(X_{n,j}^r)\ , \quad r = 1, 2\ ,$$

siehe 7.15. Wegen der Standardisierung $(*)$

$$\varphi'_{n,j}(0) = 0, \quad \varphi''_{n,j}(0) = -\sigma_{nj}^2$$

liefert das eine Entwicklung der charakteristischen Funktion $\varphi_{n,j}$ an der Stelle 0

$$\varphi_{n,j}(t) = 1 - \sigma_{n,j}^2 \cdot \frac{t^2}{2} + r_{n,j}(t) \cdot \frac{t^2}{2}, \quad t \in \mathbb{R}$$

mit Resttermen (siehe 7.16)

$$|r_{n,j}(t)| \;\leq\; \sup_{\eta\in[0,1]} |\varphi''_{n,j}(\eta t) - \varphi''_{n,j}(0)| \;\leq\; \sup_{\eta\in[0,1]} \int x^2\, |e^{i\eta tx} - 1|\, Q_{n,j}(dx)\ .$$

Für festes $t \in \mathbb{R}$ gibt es zu $\varepsilon > 0$ ein $\delta = \delta(t,\varepsilon) > 0$, so daß

$$|x| \leq \delta \implies \sup_{\eta\in[0,1]} |e^{i\eta tx} - 1| < \varepsilon\ ,$$

damit gilt

$$|r_{n,j}(t)| \;\leq\; \varepsilon\sigma_{n,j}^2 + 2\int_{\{|x|>\delta\}} x^2 Q_{n,j}(dx)$$

und damit wegen Standardisierung $(*)$

$$\sum_{j=1}^{k_n} |r_{n,j}(t)| \;\leq\; \varepsilon \sum_{j=1}^{k_n} \sigma_{n,j}^2 + 2\sum_{j=1}^{k_n} \int x^2 1_{\{|x|>\delta\}} Q_{n,j}(dx) \;=\; \varepsilon + 2L_n(\delta)\ , \quad \delta = \delta(t,\varepsilon)\ .$$

Hierbei ist $\varepsilon > 0$ beliebig. Wegen (LIN) gilt $\lim_{n\to\infty} L_n(\delta) = 0$ für jedes $\delta > 0$. Also ist gezeigt

$$(+) \qquad \sum_{j=1}^{k_n} |r_{n,j}(t)| \stackrel{n\to\infty}{\longrightarrow} 0\ , \quad \forall\, t \in \mathbb{R}\ .$$

3) Wir assoziieren zu dem gegebenen Dreiecksschema $(X_{n,j})_{j=1,\dots,k_n\,;\,n=1,2,\dots}$ ein neues Dreiecksschema $(Y_{n,j})_{j=1,\dots,k_n\,;\,n=1,2,\dots}$ aus normalverteilten zeilenweise unabhängigen ZV, so daß jeweils die ersten beiden Momente übereinstimmen:

$$\begin{cases} Y_{n,1},\dots,Y_{n,k_n} & \text{sind unabhängig für jedes } n \\ Y_{n,j} = \mathcal{N}(0,\sigma^2_{n,j})\ , & 1 \le j \le k_n\ ,\ n = 1,2,\dots \end{cases}$$

mit zugehörigen charakteristischen Funktionen $\psi_{n,j}(t) := \int e^{itx}\mathcal{N}(0,\sigma^2_{n,j})(dx)$. Wie in 2) gilt

$$\begin{aligned} \psi_{n,j}(t) &= 1 - \sigma^2_{n,j}\frac{t^2}{2} + \widetilde{r}_{n,j}(t)\frac{t^2}{2}, \quad t \in \mathbb{R} \\ |\widetilde{r}_{n,j}(t)| &\le \varepsilon\sigma^2_{n,j} + 2\int_{\{|x|>\delta\}} x^2\mathcal{N}(0,\sigma^2_{n,j})(dx)\ , \end{aligned}$$

wobei die übliche Skalierungseigenschaft für Normalverteilungen zeigt

$$\begin{aligned} \sum_{j=1}^{k_n}\int_{\{|x|>\delta\}} x^2\,\mathcal{N}(0,\sigma^2_{n,j})(dx) &= \sum_{j=1}^{k_n}\sigma^2_{n,j}\int_{\{|x|>\delta/\sigma_{nj}\}} x^2\,\mathcal{N}(0,1)(dx) \\ &\le 1\cdot\int_{\{|x|>\delta/(\max\limits_{j=1,\dots,k_n}\sigma_{n,j})\}} x^2\mathcal{N}(0,1)(dx)\ . \end{aligned}$$

Nach 8.4 gilt aber (LIN) $\Longrightarrow$ (FE), und die Feller-Bedingung mit Normierung $(*)$ zeigt

$$\max_{j=1,\dots,k_n}\sigma_{n,j} \longrightarrow 0 \quad \text{für } n\to\infty\ .$$

Also verschwindet das letzte Integral für $n\to\infty$. Zusammen erhält man also

$$(++) \qquad \sum_{j=1}^{k_n} |\widetilde{r}_{n,j}(t)| \stackrel{n\to\infty}{\longrightarrow} 0\ , \quad \forall\, t \in \mathbb{R}\ .$$

4) Der Vergleich beider Dreiecksschemata schließt nun den Beweis ab. Sei $t \in \mathbb{R}$ fest. Wegen $(*)$ und wegen der Faltungseigenschaft von Normalverteilungen gilt

$$e^{-\frac{1}{2}t^2} = \varphi_{\mathcal{N}(0,1)}(t) = \varphi_{\mathcal{N}(0,\sum_{j=1}^{k_n}\sigma^2_{n,j})}(t) = \varphi_{\sum\limits_{j=1}^{k_n} Y_{n,j}}(t) = \prod_{j=1}^{k_n}\psi_{n,j}(t)\ .$$

Wegen zeilenweise Unabhängigkeit im Dreiecksschema $(X_{n,j})_{j=1,\dots,k_n\,;\,n=1,2,\dots}$ hat man

$$\varphi_{\sum\limits_{j=1}^{k_n} X_{n,j}}(t) = \prod_{j=1}^{k_n}\varphi_{n,j}(t)\ .$$

Für Punkte $z_1, \dots, z_k, \widetilde{z}_1, \dots, \widetilde{z}_k$ in der abgeschlossenen Einheitsscheibe in $\mathbb{C}$ gilt die Ungleichung

$$|\prod_{j=1}^{k_n} z_j - \prod_{j=1}^{k_n} \widetilde{z}_j| \;\le\; \sum_{j=1}^{k_n} |z_j - \widetilde{z}_j| \,,$$

woraus man sofort schließt

$$|e^{-\frac{1}{2}t^2} - \varphi_{\sum\limits_{j=1}^{k_n} X_{n,j}}(t)| \;\le\; \sum_{j=1}^{k_n} |\varphi_{n,j}(t) - \psi_{n,j}(t)| \;\le\; \frac{t^2}{2}\left(\sum_{j=1}^{k_n} |r_{n,j}(t)| + \sum_{j=1}^{k_n} |\widetilde{r}_{n,j}(t)| \right) \,.$$

Die rechte Seite dieser Ungleichung verschwindet aber für $n \to \infty$ nach (+) und (++). Man hat also punktweise Konvergenz auf $\mathbb{R}$ der charakteristischen Funktionen $\varphi_{\sum\limits_{j=1}^{k_n} X_{n,j}}(\cdot)$ gegen die charakteristische Funktion der Standardnormalverteilung. Nach Stetigkeitssatz 7.10 ist damit die in Schritt 1) behauptete asymptotische Normalität nachgewiesen. □

8.6' Bemerkung: Zu einer anschaulichen Interpretation der Lindeberg-Bedingung kommt man, wenn man die Variablen der n-ten Zeile eines (zentrierten und standardisierten) Dreiecksschemas 8.1 als Sprunghöhen in einem zeitskalierten Random Walk

$$t \;\longrightarrow\; S_t^{(n)} := \sum_{j=1}^{\lfloor t \cdot k_n \rfloor} X_{n,j} \,, \quad 0 \le t \le 1$$

auffaßt: nach Voraussetzung sind alle Sprunghöhen L^2-Variablen, und (LIN) garantiert, daß in der Asymptotik $n \to \infty$ der Einfluß 'großer' Sprünge im Pfad von $S^{(n)}$

$$X_{n,k} 1_{\{|X_{n,k}| > \varepsilon\}} \,, \quad 1 \le k \le k_n$$

(Sprünge mit Sprunghöhe $|\cdot| > \varepsilon$ für $\varepsilon > 0$ beliebig aber fest) in einem L^2-Sinn verschwindet. □

Die einfachste Art, Satz 8.6 – unter Benutzung von 8.5 a) – anzuwenden, ist die folgende:

8.7 Korollar: Für reellwertige iid ZV $(Y_j)_j$ mit endlichen zweiten Momenten gilt

$$\mathcal{L}\left(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \frac{Y_j - \mu}{\sigma} \mid P \right) \quad \xrightarrow[n\to\infty]{\mathcal{L}} \quad \mathcal{N}(0,1)$$

wobei $\mu = E(Y_1)$ und $\sigma^2 = \mathrm{Var}(Y_1)$.

Der Satz von deMoivre und Laplace war ein Vorläufer dieser Aussage, vgl. 8.2 c) und 6.12. Eine weitere nützliche Anwendung von 8.6 –unter einer zusätzlichen Voraussetzung– ist:

8.7' Korollar: Sei $(X_{n,j})_{j=1,\ldots,k_n\,;\,n=1,2,\ldots}$ ein Dreiecksschema aus zeilenweise unabhängigen L^2-Variablen. Dieses erfülle die Lindeberg-Bedingung (LIN), und es existiere

$$(+) \qquad \widetilde{\mu} \;:=\; \lim_{n\to\infty} \sum_{j=1}^{k_n} \mu_{n,j} \quad \text{in } I\!R \quad , \quad \widetilde{\sigma^2} \;:=\; \lim_{n\to\infty} \sum_{j=1}^{k_n} \sigma^2_{n,j} \quad \text{in } (0,\infty) \; .$$

Dann gilt schwache Konvergenz in $I\!R$

$$\mathcal{L}\left(\sum_{j=1}^{k_n} X_{n,j} \mid P \right) \;\longrightarrow\; \mathcal{N}\left(\widetilde{\mu}\,,\,\widetilde{\sigma^2} \right) \quad , \quad n\to\infty \; .$$

Beweis: Bezeichne Z eine nach $\mathcal{N}(0,1)$ verteilte Zufallsvariable auf $(\Omega,\mathcal{A},P)$ (o.E. kann stets angenommen werden, daß eine solche auf demselben Wahrscheinlichkeitsraum existiert, auf dem die $X_{n,j}$ leben: gegebenenfalls erweitert man dazu den Wahrscheinlichkeitsraum, indem man ein Produkt aus dem ursprünglichen $(\Omega,\mathcal{A},P)$ mit $(I\!R,\mathcal{B}(I\!R),\mathcal{N}(0,1))$ bildet und alle Zufallsvariablen in natürlicher Weise auf dem Produktraum redefiniert). Nach 8.6 gilt dank (LIN)

$$Z_n \;:=\; \sum_{j=1}^{k_n} \frac{X_{n,j}-\mu_{n,j}}{s_n} \quad \xrightarrow{\mathcal{L}} \quad Z \quad , \quad n\to\infty \; .$$

Die Konvergenzvoraussetzung (+) erlaubt, hieraus auf Straffheit der Folge

$$(++) \qquad \mathcal{L}\left(\sum_{j=1}^{k_n} X_{n,j} \mid P \right) \quad , \quad n\geq 1$$

in $I\!R$ zu schließen. Aufgrund der Straffheit von (++) sind die Folgen $(Z_n)_n$ und $(\widetilde{Z}_n)_n$

$$\widetilde{Z}_n \;:=\; \frac{1}{\widetilde{\sigma}} \left[\sum_{j=1}^{k_n} X_{n,j} - \widetilde{\mu} \right] \quad , \quad n\geq 1$$

stochastisch äquivalent für $n\to\infty$, siehe 6.5 c). Mit 6.6' konvergiert dann auch die Folge $(\widetilde{Z}_n)_n$ schwach in $I\!R$ gegen Z, für $n\to\infty$. Skalieren (dies ist eine stetige Abbildung) liefert mit 6.3 oder 6.4' dann die Behauptung des Korollars. □

Man beachte, daß die Zusatzvoraussetzung (+) in Korollar 8.7' keineswegs eine Lindeberg-Bedingung impliziert (dies darf nicht mit 8.5 a) verwechselt werden), wie das Beispiel

$$X_{n,1} \;\sim\; \Gamma(a,p) \; , \; X_{n,2} \;=\; \ldots \;=\; X_{n,k_n} \;=\; 0 \quad , \quad n\geq 1$$

sofort deutlich macht.

Im nächsten Satz betrachten wir die Bedingungen für asymptotische Normalität der Zeilensummen in Satz 8.6 genauer. Ohne Zweifel kann asymptotische Normalität der Zeilensummen auch ohne Lindeberg-Bedingung gelten, wie das triviale Beipiel

$$\forall\, n \geq 1: \quad X_{n,1} \sim \mathcal{N}(0,1), \quad X_{n,2} = \ldots = X_{n,k_n} = 0 \quad P\text{-fs}$$

zeigt. In Dreiecksschemata aus 'gleichmäßig kleinen' zeilenweise unabhängigen L^2-Variablen ist (LIN) jedoch notwendig und hinreichend für asymptotische Normalität der Zeilensummen.

8.8 Satz von Lindeberg-Feller: Sei $(X_{n,j})_{j=1,\ldots,k_n\,;\,n=1,2,\ldots}$ ein Dreiecksschema aus zeilenweise unabhängigen L^2-Variablen. Betrachte die Eigenschaft (ZSK)

$$\text{(ZSK)} \qquad \text{es gilt für } n \to \infty: \quad \mathcal{L}\left(\sum_{j=1}^{k_n} \frac{X_{n,j} - \mu_{n,j}}{s_n} \;\middle|\; P \right) \quad \xrightarrow{w} \quad \mathcal{N}(0,1)$$

mit $\mu_{n,j}$ und s_n wie in 8.1. Dann sind folgende Aussagen gleichwertig:

i) es gilt (LIN),

ii) es gilt (ZSK) und (FE),

iii) es gilt (ZSK) und (AV).

Beweis: Nach Satz 8.6 und Hilfssatz 8.4 ist nur noch die Implikation iii) $\Longrightarrow$ i) zu zeigen. Die ZV $X_{n,j}$ habe Verteilung $Q_{n,j} = \mathcal{L}(X_{n,j}|P)$ und charakteristische Funktion $\varphi_{n,j}$, für alle n, j; wir arbeiten wie im Beweis von 8.6 mit der Standardisierung

$$(*) \qquad \mu_{n,j} = 0 \quad \forall\, n, j, \quad s_n^2 = \sigma_{n,1}^2 + \ldots \sigma_{n,k_n}^2 = 1 \quad \forall\, n \geq 1\,.$$

1) Schreibe $\varphi_{n,j}(t) = \int \cos(tx) Q_{n,j}(dx) + i \int \sin(tx) Q_{n,j}(dx)$ und betrachte für $t > 0$

$$\widetilde{L}_n(t) \;:=\; 1 - \frac{2}{t^2} \sum_{j=1}^{k_n} \int (1 - \cos tx)\, Q_{n,j}(dx) \;=\; 1 - \frac{2}{t^2} \sum_{j=1}^{k_n} \mathrm{Re}(1 - \varphi_{n,j}(t))\,.$$

Der Beweisplan ist der folgende. In Schritt 3) unten werden wir zeigen, daß (AV) und (ZSK) zusammen die Aussage

$$(+) \qquad \lim_{n \to \infty} \widetilde{L}_n(N) \;=\; 0 \quad \text{für jedes feste } N \in I\!N$$

implizieren. Zugleich wird aber gelten (siehe Schritt 2) unten)

$$(++) \qquad \text{für beliebiges } \varepsilon > 0,\ N \in I\!N \ : \quad L_n(\varepsilon) \ \leq \ \widetilde{L}_n(N) \ + \ 4\frac{1}{\varepsilon^2 N^2} \ .$$

Zusammen ergeben (+) und (++) die gewünschte Aussage

$$(\text{AV}) + (\text{ZSK}) \quad \Longrightarrow \quad (\text{LIN}) \ .$$

2) Der Beweis von (++) ist elementar. Man hat wegen der Standardisierung $(*)$

$$\begin{aligned}
L_n(\varepsilon) &= \sum_{j=1}^{k_n} E\left(|X_{n,j}|^2 \, 1_{\{|X_{n,j}|>\varepsilon\}}\right) \ = \ \sum_{j=1}^{k_n} \int_{\{|x|>\varepsilon\}} x^2 \, Q_{n,j}(dx) \\
&= \sum_{j=1}^{k_n} \sigma_{n,j}^2 - \sum_{j=1}^{k_n} \int_{\{|x|\leq\varepsilon\}} x^2 \, Q_{n,j}(dx) \\
&= 1 \ - \ \sum_{j=1}^{k_n} \int_{\{|x|\leq\varepsilon\}} x^2 \, Q_{n,j}(dx) \\
&\leq 1 \ - \ \frac{2}{N^2} \sum_{j=1}^{k_n} \int_{\{|x|\leq\varepsilon\}} (1-\cos Nx) \, Q_{n,j}(dx) \\
&= \widetilde{L}_n(N) \ + \ \frac{2}{N^2} \sum_{j=1}^{k_n} \int_{\{|x|>\varepsilon\}} (1-\cos Nx) \, Q_{n,j}(dx)
\end{aligned}$$

wobei in der vorletzten Abschätzung Hilfssatz 7.3 a) benutzt wurde:

$$(1-\cos tx) = -\text{Re}\left(e^{itx} - 1 - itx\right) \leq |e^{itx} - 1 - itx| \leq \frac{|tx|^2}{2} \ , \quad t, x \in I\!R \ .$$

Mit Hilfe der trivialen Ungleichung

$$(1-\cos Nx) \, 1_{\{|x|>\varepsilon\}} \ \leq \ 2 \, 1_{\{|x|>\varepsilon\}} \ \leq \ \frac{2}{\varepsilon^2} x^2 \ , \quad x \in I\!R$$

schätzt man in der Ungleichungskette oben weiter ab

$$\frac{2}{N^2} \sum_{j=1}^{k_n} \int_{\{|x|>\varepsilon\}} (1-\cos Nx) \, Q_{n,j}(dx) \ \leq \ \frac{4}{N^2\varepsilon^2} \sum_{j=1}^{k_n} \sigma_{n,j}^2 \ = \ \frac{4}{N^2\varepsilon^2}$$

und erhält (++).

3) Als Vorbereitung auf den Beweis von (+) zeigen wir zuerst, daß unter (AV) und $(*)$ gilt

$$(\circ) \qquad \sum_{j=1}^{k_n} |\varphi_{n,j}(t) - 1|^2 \quad \overset{n\to\infty}{\longrightarrow} \quad 0 \quad \text{gleichmäßig auf kompakten } t\text{-Intervallen.}$$

Mit Hilfssatz 7.3 a) startet man dabei von der Abschätzung

$$|\varphi_{n,j}(t) - 1| \ \leq \ \int |e^{itx} - 1| Q_{n,j}(dx) \ \leq \ 2 \cdot Q_{n,j}(\{|x| > \delta\}) + \delta \, t$$

für beliebig kleines $\delta > 0$, in der unter (AV) und $(*)$

$$\max_{1\le j\le k_n} Q_{n,j}(\{|x|>\delta\}) \;=\; \max_{1\le j\le k_n} E\left(\left|\frac{X_{n,j}-\mu_{n,j}}{s_n}\right| > \delta\right) \;\longrightarrow\; 0 \quad \text{für } n\to\infty\,.$$

Also hat man unter (AV)

$$(\circ\circ) \qquad \max_{1\le j\le k_n} |\varphi_{n,j}(t)-1| \quad \overset{n\to\infty}{\longrightarrow} \quad 0 \quad \text{gleichmäßig auf kompakten } t\text{-Intervallen.}$$

Als nächstes gilt wegen Standardisierung $(*)$ insbesondere $\int x\,Q_{n,j}(dx)=0$, daher

$$\varphi_{n,j}(t)-1 \;=\; \int (e^{itx}-1-itx)\,Q_{n,j}(dx)$$

woraus wieder mit Hilfssatz 7.3 a) folgt

$$\sum_{j=1}^{k_n} |\varphi_{n,j}(t)-1| \;\le\; \frac{t^2}{2}\sum_{j=1}^{k_n}\int x^2 Q_{n,j}(dx) \;=\; \frac{t^2}{2}\sum_{j=1}^{k_n}\sigma_{n,j}^2 \;=\; \frac{t^2}{2}\,.$$

Das bedeutet

$$(\circ\circ\circ) \quad \sum_{j=1}^{k_n} |\varphi_{n,j}(t)-1| \quad \text{ist beschränkt für } n\to\infty, \text{ gleichmäßig auf kompakten } t\text{-Intervallen.}$$

Mit Hilfe der trivialen Abschätzung

$$\sum_{j=1}^{k_n} |\varphi_{nj}(t)-1|^2 \;\le\; \left(\max_{1\le j\le k_n}|\varphi_{nj}(t)-1|\right)\sum_{j=1}^{k_n}|\varphi_{nj}(t)-1|$$

ergibt sich $(\circ)$ aus $(\circ\circ)$ und $(\circ\circ\circ)$.

4) Wir zeigen $(+)$. Starte von der Reihenentwicklung für den (Hauptast des) Logarithmus

$$\log(1+z) \;=\; z + \sum_{r=2}^{\infty}\frac{(-1)^{r-1}}{r}\,z^r\,, \quad \forall\, z\in\mathbb{C} \quad \text{mit } |z|<\frac{1}{2}$$

in $\mathbb{C}$, woraus mit Hilfe der geometrischen Reihe die Abschätzung

$$|\log(1+z)-z| \;\le\; |z|^2\,, \quad z\in\mathbb{C},\ |z|<\frac{1}{2}$$

folgt. Auf jedem kompakten t-Intervall gilt $|\varphi_{n,j}(\cdot)-1|<\frac{1}{2}$, $1\le j\le k_n$, für schließlich alle n, wegen $(\circ\circ)$ in Schritt 3). Also hat man eine Entwicklung

$$\log(\varphi_{n,j}(t)) \;=\; \log(1+(\varphi_{n,j}(t)-1)) \;=\; (\varphi_{n,j}(t)-1)+R_{n,j}(t)$$

zusammen mit einer Abschätzung der Restterme für hinreichend großes n

$$\sum_{j=1}^{k_n}|R_{n,j}(t)| \;\le\; \sum_{j=1}^{k_n}|\varphi_{n,j}(t)-1|^2 \quad \overset{n\to\infty}{\longrightarrow} \quad 0$$

gleichmäßig auf kompakten t-Intervallen, nach ($\circ$) in Schritt 3).

Die Voraussetzung (ZSK) bedeutet nach Stetigkeitssatz von P. Lévy (7.10)

$$\prod_{j=1}^{k_n} \varphi_{n,j}(t) \;=\; \varphi_{\sum_{j=1}^{k_n} X_{n,j}}(t) \quad \stackrel{n\to\infty}{\longrightarrow} \quad \varphi_{\mathcal{N}(0,1)}(t) \;=\; e^{-\frac{1}{2}t^2}$$

für jedes $t \in I\!R$. Die logarithmische Entwicklung oben erlaubt, hieraus auf Konvergenz in $C\!\!\!\!I$

$$e^{\sum_{j=1}^{k_n}(\varphi_{n,j}(t)-1)} \quad \longrightarrow \quad e^{-\frac{1}{2}t^2} \quad \text{für } n \to \infty$$

zu schließen. Damit konvergieren aber auch die Beträge

$$|\, e^{\sum_{j=1}^{k_n}(\varphi_{n,j}(t)-1)} \,| \;=\; e^{\mathrm{Re}(\sum_{j=1}^{k_n}(\varphi_{nj}(t)-1))} \quad \longrightarrow \quad e^{-\frac{1}{2}t^2}$$

und man erhält

$$\sum_{j=1}^{k_n} \mathrm{Re}(1-\varphi_{nj}(t)) \quad \longrightarrow \quad +\frac{1}{2}t^2 \quad \text{für } n \to \infty.$$

Nach Definition von $\widetilde{L}_n(\cdot)$ bedeutet das aber

$$\lim_{n\to\infty} \widetilde{L}_n(t) \;=\; 0 \quad \text{für jedes } t \neq 0 \;,$$

und die Aussage (+) ist bewiesen. Damit ist der Beweis von Satz 8.8 abgeschlossen. $\square$

8.9 Bemerkung: a) Die in diesem Kapitel durchgängig gemachte Voraussetzung endlicher zweiter Momente kann im iid Fall noch einmal minimal abgeschwächt werden, siehe Bingham, Goldie und Teugels (1987, S. 346).

b) Ohne Lindeberg-Bedingung können in Dreiecksschemata aus zeilenweise unabhängigen asymptotisch vernachlässigbaren L^2-Variablen andere Typen von Grenzverteilungen entstehen. Dies wird in Kapitel IX genauer betrachtet werden.

Ein einfaches Beispiel: sei $(p_n)_n$ eine Folge in $(0,1)$ mit der Eigenschaft $\lim_{n\to\infty} n\,p_n =: \lambda$, $0 < \lambda < \infty$. Der aus der 'Einführung in die Stochastik' bekannte Poissonsche Grenzwertsatz zeigt schwache Konvergenz in $I\!R$

$$\mathcal{B}(n,p_n) \quad \stackrel{w}{\longrightarrow} \quad \mathcal{P}(\lambda)\;, \quad n \to \infty$$

(für diskrete Verteilungen reicht es nach 6.12, die Konvergenz der Massen dieser Verteilungen auf den Einpunktmengen $\{k\}$, $k \in I\!N_0$, nachzurechnen). Betrachtet man nun das Dreiecksschema

$$\text{für festes } n \geq 1: \quad X_{n,1},....,X_{n,n} \;\text{ sind iid} \sim \mathcal{B}(1,p_n)$$

aus zeilenweise unabhängigen L^2-Variablen, die nur die Werte 0 oder 1 annehmen, so sieht man leicht, daß (AV) und (FE) erfüllt sind, (LIN) aber verletzt ist. Der schwache Limes für zentrierte und standardisierte Zeilensummen in diesem Dreiecksschema ist eine um λ verschobene und um den Faktor $\frac{1}{\sqrt{\lambda}}$ gestreckte Poissonvariable mit Parameter λ. □

Aus dem 'eindimensionalen' zentralen Grenzwertsatz 8.6 (bzw. aus Korollar 8.7') und Cramér-Wold 7.12 erhält man abschließend einen zentralen Grenzwertsatz für Zufallsvariable mit Werten in $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ ohne große zusätzliche Arbeit:

8.10 Zentraler Grenzwertsatz in $\mathbb{R}^d$: Betrachte ein Dreiecksschema aus zeilenweise unabhängigen L^2-Variablen mit Werten in $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$

$$(*) \qquad (Y_{n,j})_{j=1,\ldots,k_n\,,\,n=1,2,\ldots}$$

(dies bedeutet, daß alle Komponenten $(Y^{(r)}_{n,j})$, $1 \le r \le d$, Zufallsvariable in $L^2(P)$ sind). Schreibe

$$\nu_{n,j} := E(Y_{n,j}) \quad , \quad \Lambda_{n,j} := Cov(Y_{n,j}) = E\left((Y_{n,j} - \nu_{n,j})(Y_{n,j} - \nu_{n,j})^\top\right)$$

und setze voraus: es gelte Konvergenz der ersten beiden Momente

$$\begin{cases} \text{es gebe } \widetilde{\nu} = (\widetilde{\nu}^{(r)})_{1\le r\le d} \text{ in } \mathbb{R}^d \text{ und } \widetilde{\Lambda} = (\widetilde{\Lambda}^{(r,s)})_{1\le r,s\le d} \text{ in } \mathbb{R}^{d\times d} \\ \text{so daß } \quad \lim\limits_{n\to\infty} \sum_{j=1}^{k_n} \nu^{(r)}_{n,j} \;=\; \widetilde{\nu}^{(r)} \,, \quad \lim\limits_{n\to\infty} \sum_{j=1}^{k_n} \Lambda^{(r,s)}_{n,j} \;=\; \widetilde{\Lambda}^{(r,s)} \,, \quad 1 \le r,s \le d \,, \end{cases}$$

(notwendig ist dann $\widetilde{\Lambda}$ symmetrisch und nichtnegativ definit), und das Dreiecksschema $(*)$ erfülle eine Lindeberg-Bedingung in der folgenden Form

$$\lim_{n\to\infty} \sum_{j=1}^{k_n} E\left(|Y_{n,j} - \nu_{n,j}|^2 1_{\{\,|Y_{n,j}-\nu_{n,j}|>\varepsilon\,\}}\right) \;=\; 0 \quad \text{für jedes } \varepsilon > 0 \,.$$

Dann sind die Zeilensummen im Dreiecksschema $(*)$ asymptotisch normalverteilt: man hat

$$\sum_{j=1}^{k_n} Y_{n,j} \quad \xrightarrow{\mathcal{L}} \quad \mathcal{N}\left(\widetilde{\nu}\,,\,\widetilde{\Lambda}\right)$$

(schwache Konvergenz in $\mathbb{R}^d$) für $n \to \infty$.

Beweis: Nach Cramér-Wold 7.12 reicht es, die Schar schwacher Konvergenzen in $\mathbb{R}$

$$(\diamond) \qquad \text{für jedes feste } \alpha \in \mathbb{R}^d \;:\quad \alpha^\top \sum_{j=1}^{k_n} Y_{n,j} \quad \xrightarrow{\mathcal{L}} \quad \mathcal{N}\left(\alpha^\top \widetilde{\nu}\,,\,\alpha^\top \widetilde{\Lambda}\,\alpha\right)$$

für $n \to \infty$ nachzuweisen. Weiter reicht es, hierbei den zentrierten Fall zu betrachten:

$$\nu_{n,j} = 0 \ , \ \widetilde{\nu} = 0 \ , \ \Lambda_{n,j} = E(Y_{n,j}Y_{n,j}^\top) \ .$$

Wir schreiben bezüglich eines festen $\alpha \in I\!R^d$

$$X_{n.j} := \alpha^\top Y_{n,j} \ , \quad \mu_{n,j} := E(X_{n,j}) = 0 \ , \quad \sigma^2_{n,j} := E(X^2_{n,j}) = \alpha^\top \Lambda_{n,j}\,\alpha$$

und betrachten das Dreiecksschema aus zentrierten $I\!R$-wertigen L^2-Variablen

$$(**) \qquad (\, X_{n,j}\,)_{j=1,\ldots,k_n\,,\,n\geq 1} \ .$$

<u>Fall 1:</u> Gilt $\alpha^\top \widetilde{\Lambda}\,\alpha = 0$, so verschwindet $\sum_{j=1}^{k_n} X_{n,j}$ für $n \to \infty$ in $L^2(P)$ (aufgrund der Konvergenzbedingung an die ersten beiden Momente), damit auch P-stochastisch, also auch in Verteilung. Gleichzeitig ist in diesem Fall aber die Normalverteilung auf der rechten Seite von $(\diamond)$ aber ein Diracmass in 0. Also gilt $(\diamond)$.

<u>Fall 2:</u> Sei $\alpha^\top \widetilde{\Lambda}\,\alpha > 0$. Dann ist mit $\widetilde{\mu} = 0$ und $\widetilde{\sigma}^2 = \alpha^\top \widetilde{\Lambda}\,\alpha > 0$ die Momentenbedingung (+) aus Korollar 8.7' erfüllt. Zum Beweis von $(\diamond)$ bleibt als nur nachzuweisen, daß das Schema $(**)$ die Lindebergbedingung (LIN) gemäß 8.3

$$\forall\, \varepsilon > 0 \text{ fest:} \quad L_n(\varepsilon) \ = \ \frac{1}{s_n^2} \sum_{j=1}^{k_n} E\left(X^2_{n,j}\; 1_{\{|X_{n,j}|>\varepsilon\cdot s_n\}} \right) \quad \stackrel{n\to\infty}{\longrightarrow} \quad 0$$

mit $s_n^2 = \sum_{j=1}^{k_n} \sigma^2_{n,j}$ erfüllt. Aufgrund der vorausgesetzten Konvergenz $s_n^2 \to \widetilde{\sigma}^2 \in (0,\infty)$ ist diese aber äquivalent zur einfacheren Bedingung

$$(\circ) \qquad \lim_{n\to\infty} \sum_{j=1}^{k_n} E\left(X^2_{n,j}\; 1_{\{|X_{n,j}|>\varepsilon\}} \right) \ = \ 0 \quad \text{für jedes } \varepsilon > 0$$

die wir jetzt nachzuprüfen: man hat (notwendig gilt $|\alpha| \in (0,\infty)$, wegen Fall 2)

$$|X_{n,j}| = |\alpha^\top Y_{n,j}| \ \leq \ |\alpha||Y_{n,j}| \quad \text{und} \quad \left\{ |X_{n,j}| = |\alpha^\top Y_{n,j}| \ > \ \varepsilon' \right\} \subset \left\{ |Y_{n,j}| > \frac{\varepsilon'}{|\alpha|} \right\}$$

so daß $(\circ)$ sofort aus der in der Formulierung des Satzes gemachten Voraussetzung

$$\lim_{n\to\infty} \sum_{j=1}^{k_n} E\left(|Y_{n,j}|^2 1_{\{\,|Y_{n,j}|>\varepsilon\,\}} \right) \ = \ 0 \quad \text{für jedes } \varepsilon > 0$$

folgt. Damit ist Satz 8.10 bewiesen. □