HANDOUT #4
STOCHASTISCHES LANDAU-SYMBOL

FOLKMAR BORNEMANN

Maßkonzentrationsphänomene führen auf Tailabschätzungen etwa der Art

$$\mathbb{P}(|X_n - x| \geqslant t/\sqrt{n}) \leqslant 2e^{-2t^2} \qquad (t \geqslant 0, n \in \mathbb{N}).$$

Talagrand schlug 1998 die Intuition vor, dass sich X_n dann „im wesentlichen wie die Konstante x verhalte". Ich möchte das mit dem stochastischen Landau-Symbol

$$X_n = x + \mathbb{O}(n^{-1/2})$$

präzisieren und zeigen, dass hierfür im wesentlichen die gleichen Rechenregeln wie für das deterministische Landau-Symbol O gelten.

1. Tailklassen und Definition des stochastischen Landau-Symbols

Das stochastische Landau-Symbol beruht auf der speziellen Klasse

$$\mathcal{T} = \{\psi : [0,\infty) \to [0,\infty) : \psi(t) = Ce^{-ct^\gamma} \text{ mit } \gamma, c > 0 \text{ und } C \geqslant 1\}$$

subexponentieller Tailabschätzungen, für die man leicht zeigt ($\phi, \psi \in \mathcal{T}$):

(i) ψ ist integrierbar;

(ii) ψ ist stetig und monoton fallend mit $\psi(0) \geqslant 1$ und $\lim_{t\to\infty} \psi(t) = 0$;

(iii) $\psi(t/2)$, $\psi(\sqrt{t})$ und $\phi(t) + \psi(t)$ werden in $\mathcal{T}$ dominiert.[1]

Aus (ii) und (iii) folgt, dass für $\alpha, \beta, \gamma > 0$ auch $\alpha\psi(\beta t^\gamma)$ in $\mathcal{T}$ dominiert wird.

Definition 1. Es seien X_n und Y_n Folgen komplexer Zufallsvariablen. Wird

$$\sup_{n\in\mathbb{N}} \mathbb{P}(|X_n| \geqslant t|Y_n|)$$

als Funktion von t in $\mathcal{T}$ dominiert, so schreiben wir kurz

$$X_n = \mathbb{O}(Y_n).$$

(Sprich: „X_n ist mit überwältigender Wahrscheinlichkeit $O(Y_n)$.") Ist X_n' eine weitere Folge komplexer Zufallsvariablen so steht

$$X_n = X_n' + \mathbb{O}(Y_n) \quad \text{für} \quad X_n - X_n' = \mathbb{O}(Y_n).$$

Bemerkung 1. In der Literatur (siehe etwa A. W. van der Vaart, *Asymptotic Statistics*, Cambridge University Press 1998, §2.2) findet sich eine weniger restriktive Definition des stochastischen Landau-Symbols. Dort bedeutet $X_n = O_\mathbb{P}(1)$, dass die

Date: 16. Juni 2010.

[1] Eine reelle Funktion ϕ auf $[0,\infty)$ wird in $\mathcal{T}$ dominiert, wenn es ein $\psi \in \mathcal{T}$ mit $\phi \leqslant \psi$ gibt.

Folge der Zufallsvariablen X_n stochastisch gleichmäßig beschränkt ist, d.h. eine gleichmäßig *straffe* Folge von Zufallsvariablen darstellt:

$$\lim_{t\to\infty} \sup_{n\in\mathbb{N}} \mathbb{P}(|X_n| \geqslant t) = 0.$$

Da $\psi(t) = \sup_{n\in\mathbb{N}} \mathbb{P}(|X_n| \geqslant t)$ eine monoton fallende Funktion mit $\psi(0) = 1$ ist, und sich jede solche Funktion mit $\psi(t) \to 0$ durch eine stetige Funktion mit all diesen Eigenschaften dominieren lässt, entspricht diese Definition des stochastischen Landau-Symbols der Wahl der folgenden Klasse von Tailabschätzungen:

$$\mathcal{T}'' = \{\psi : [0,\infty) \to [0,\infty) : \psi \text{ stetig, monoton fallend, } \psi(0) \geqslant 1,\ \lim_{t\to\infty} \psi(t) = 0\}.$$

Eigenschaft (ii) auf der vorigen Seite wird hier also zur Definition; Eigenschaft (iii) folgt dann sofort. Fügt man Eigenschaft (i) hinzu (also die Integrierbarkeit), so bleibt Eigenschaft (iii) erhalten:

$$\mathcal{T}' = \{\psi \in \mathcal{T}'' : \psi \text{ ist integrierbar}\}.$$

Es gilt $\mathcal{T} \subset \mathcal{T}' \subset \mathcal{T}''$. Da wir im folgenden nur die Eigenschaften (i), (ii) und (iii) benutzen werden, hätten wir das $\mathbb{O}$-Symbol auch auf Basis von Tailabschätzungen in $\mathcal{T}'$ definieren können. In §2 geht Eigenschaft (i) nur in Lemma 3(i) und (iii) ein, so dass alles andere auch für das $O_{\mathbb{P}}$-Symbol richtig bleibt.

2. Das Kalkül des stochastischen Landau-Symbols

In diesem Abschnitt formuliere und beweise ich die grundlegenden Eigenschaften des stochastischen Landau-Symbols. Im folgenden seien X_n, X'_n, Y_n, Y'_n Folgen komplexer Zufallsvariablen. Ungleichungen in den Beweisen gelten gleichmäßig für sämtliche Werte der freien Variablen.

Lemma 1 (Arithmetik)**.** *Es gilt:*

(i) $X_n = \mathbb{O}(Y_n) \;\Leftrightarrow\; X_n = \mathbb{O}(|Y_n|),\ -X_n = \mathbb{O}(|Y_n|)$ *und* $|X_n| = \mathbb{O}(|Y_n|)$;

(ii) $X_n = \mathbb{O}(Y_n)$ *und* $X'_n = \mathbb{O}(Y'_n) \;\Rightarrow\; X_n + X'_n = \mathbb{O}(|Y_n| + |Y'_n|)$;

(iii) $X_n = \mathbb{O}(Y_n)$ *und* $X'_n = \mathbb{O}(Y'_n) \;\Rightarrow\; X_n \cdot X'_n = \mathbb{O}(Y_n \cdot Y'_n)$;

(iv) $X_n = \mathbb{O}(Y_n)$ *und* $\kappa > 0 \;\Rightarrow\; |X_n|^\kappa = \mathbb{O}(|Y_n|^\kappa)$.

Beweis. Teil (i) folgt unmittelbar aus der Definition, da in diese nur $|X_n|$ und $|Y_n|$ eingeht. Für (ii) wähle nach Definition $\psi, \psi' \in \mathcal{T}$, so dass

$$\mathbb{P}(|X_n| \geqslant t|Y_n|) \leqslant \psi(t), \qquad \mathbb{P}(|X'_n| \geqslant t|Y'_n|) \leqslant \psi'(t).$$

Wegen der Dreiecksungleichung zieht $|X_n + X'_n| \geqslant t(|Y_n| + |Y'_n|)$ wenigstens eines der beiden Ereignisse $|X_n| \geqslant t|Y_n|$ oder $|X'_n| \geqslant t|Y'_n|$ nach sich. Damit gilt

$$\mathbb{P}(|X_n + X'_n| \geqslant t(|Y_n| + |Y'_n|)) \leqslant \mathbb{P}(|X_n| \geqslant t|Y_n|) + \mathbb{P}(|X'_n| \geqslant t|Y'_n|) \leqslant \psi(t) + \psi'(t),$$

eine in $\mathcal{T}$ dominierte Tailabschätzung. In (iii) zieht $|X_n \cdot X'_n| \geqslant t|Y_n \cdot Y'_n|$ wenigstens eines der beiden Ereignisse $|X_n| \geqslant \sqrt{t}|Y_n|$ oder $|X'_n| \geqslant \sqrt{t}|Y'_n|$ nach sich. Dann ist

$$\begin{aligned}&\mathbb{P}(|X_n \cdot X'_n| \geqslant t|Y_n \cdot Y'_n|)\\ &\qquad\qquad \leqslant \mathbb{P}(|X_n| \geqslant \sqrt{t}|Y_n|) + \mathbb{P}(|X'_n| \geqslant \sqrt{t}|Y'_n|) \leqslant \psi(\sqrt{t}) + \psi'(\sqrt{t});\end{aligned}$$

auch diese Tailabschätzung wird in $\mathcal{T}$ dominiert. In (iv) gilt

$$\mathbb{P}(|X_n|^\kappa \geqslant t|Y_n|^\kappa) = \mathbb{P}(|X_n| \geqslant t^{1/\kappa}|Y_n|) \leqslant \psi(t^{1/\kappa}),$$

eine in $\mathcal{T}$ dominierte Tailabschätzung. □

Lemma 2 (Kompatibilität)**.** *Es gilt:*

(i) $X_n = O(Y_n) \;\Rightarrow\; X_n = \mathbb{O}(Y_n)$;

(ii) $X_n = \mathbb{O}(X_n')$ *und* $X_n' = \mathbb{O}(Y_n) \;\Rightarrow\; X_n = \mathbb{O}(Y_n)$.

Beweis. In (i) gibt es nach Voraussetzung eine Konstante $c > 0$ mit $|X_n| \leqslant c|Y_n|$. Damit ist also $\mathbb{P}(|X_n| \geqslant t|Y_n|) = 0$ für $t > c$ und sonst $\leqslant 1$. Für jedes $\psi \in \mathcal{T}$ gibt es aus Stetigkeitsgründen ein $t_* > 0$ mit $\psi(t_*) > 0$, so dass wegen der Monotonie

$$\mathbb{P}(|X_n| \geqslant t|Y_n|) \leqslant \psi(t_*)^{-1}\psi(c^{-1}t \cdot t_*);$$

dies ist eine in in $\mathcal{T}$ dominierte Tailabschätzung. Für (ii) wähle $\psi, \psi' \in \mathcal{T}$ mit

$$\mathbb{P}(|X_n| \geqslant t|X_n'|) \leqslant \psi(t), \qquad \mathbb{P}(|X_n'| \geqslant t|Y_n|) \leqslant \psi'(t).$$

Da $|X_n| \geqslant t|Y_n|$ wenigstens eines der beiden Ereignisse $|X_n| \geqslant \sqrt{t}|X_n'|$ oder $|X_n'| \geqslant \sqrt{t}|Y_n|$ nach sich zieht, gilt

$$\mathbb{P}(|X_n| \geqslant t|Y_n|) \leqslant \mathbb{P}(|X_n| \geqslant \sqrt{t}|X_n'|) + \mathbb{P}(|X_n'| \geqslant \sqrt{t}|Y_n|) \leqslant \psi(\sqrt{t}) + \psi'(\sqrt{t}),$$

ebenfalls eine in $\mathcal{T}$ dominierte Tailabschätzung. □

Lemma 3 (Extraktion)**.** *Es gilt:*

(i) $X_n = \mathbb{O}(1) \;\Rightarrow\; \mathbb{E}X_n = O(1)$;

(ii) X_n *reellwertig und* $X_n = \mathbb{O}(1) \;\Rightarrow\; \mathbb{M}X_n = O(1)$;

(iii) $X_n = \mathbb{O}(n^{-\kappa})$ *und* $\kappa > 0 \;\Rightarrow\; X_n \overset{\text{a.s.}}{\to} 0$.

Beweis. Für (i) wähle $\psi \in \mathcal{T}$ mit $\mathbb{P}(|X_n| \geqslant t) \leqslant \psi(t)$, so dass

$$|\mathbb{E}X_n| \leqslant \mathbb{E}|X_n| = \int_0^\infty \mathbb{P}(|X_n| \geqslant t)\,dt \leqslant \int_0^\infty \psi(t)\,dt < \infty.$$

In (ii) gibt es (wegen $\psi(0) \geqslant 1$, der Stetigkeit und des monotonen Abfalls von $\psi(t)$ gegen Null für $t \to \infty$) ein *maximales* t_* mit $\psi(t_*) \geqslant 1/2$. Damit gilt für $t > t_*$, dass (anderenfalls wäre t_* nicht maximal)

$$\max(\mathbb{P}(X_n \leqslant -t), \mathbb{P}(X_n \geqslant t)) \leqslant \mathbb{P}(|X_n| \geqslant t) \leqslant \psi(t) < 1/2 \leqslant \psi(t_*).$$

Für (jeden) Median $\mathbb{M}X_n$ gilt also $-t_* \leqslant \mathbb{M}X_n \leqslant t_*$. In (iii) wähle $\psi \in \mathcal{T}$ mit

$$\mathbb{P}(|X_n| \geqslant tn^{-\kappa}) \leqslant \psi(t).$$

Damit gilt wegen der Monotonie von ψ für $\epsilon > 0$, dass

$$\sum_{n=1}^\infty \mathbb{P}(|X_n| \geqslant \epsilon) \leqslant \sum_{n=1}^\infty \psi(\epsilon n^\kappa) \leqslant \int_0^\infty \psi(\epsilon t^\kappa)\,dt < \infty.$$

Nach dem Lemma von Borel–Cantelli ist deshalb $X_n \overset{\text{a.s.}}{\to} 0$. □

3. Beispiele für die Verwendung des Kalküls

Es seien X_n, X'_n, X''_n und Y_n, Y'_n, Y''_n Folgen komplexer Zufallsvariablen und θ_n eine deterministische Folge positiver Zahlen.

Korollar 1. *Es gilt*

$$X_n = X'_n + \mathbb{O}(Y_n) \text{ und } X'_n = X''_n + \mathbb{O}(Y'_n) \;\Rightarrow\; X_n = X''_n + \mathbb{O}(|Y_n| + |Y'_n|).$$

Beweis. Addiere $X_n - X'_n = \mathbb{O}(Y_n)$ und $X'_n - X''_n = \mathbb{O}(Y'_n)$. □

Korollar 2. *Es gilt für $X_n = \mathbb{O}(Y_n)$ und $X'_n = \mathbb{O}(Y'_n)$, dass*

$$X_n = X'_n + \mathbb{O}(Y''_n) \;\Rightarrow\; X_n^2 = X_n'^2 + \mathbb{O}((|Y_n| + |Y'_n|)Y''_n).$$

Beweis. Multiplikation liefert $X_n^2 = X_n X'_n + \mathbb{O}(Y_n Y''_n)$ und $X_n X'_n = X_n'^2 + \mathbb{O}(Y'_n Y''_n)$. □

Korollar 3. *Es gilt für $X_n = \mathbb{O}(1)$*

$$X_n = \mathbb{E}X_n + \mathbb{O}(\theta_n) \;\Rightarrow\; X_n^2 = \mathbb{E}X_n^2 + \mathbb{O}(\theta_n) \text{ und } \operatorname{var} X_n = O(\theta_n).$$

Beweis. Extraktion liefert $\mathbb{E}X_n = O(1)$, so dass $X_n^2 = (\mathbb{E}X_n)^2 + \mathbb{O}(\theta_n)$. Eine weitere Extraktion ergibt $\mathbb{E}X_n^2 = (\mathbb{E}X_n)^2 + O(\theta_n)$. □

Korollar 4. *Für reellwertige X_n gilt*

$$X_n = \mathbb{E}X_n + \mathbb{O}(1) \;\Leftrightarrow\; X_n = \mathbb{M}X_n + \mathbb{O}(1).$$

Beweis. Extraktion sowie die Beziehungen $\mathbb{M}\mathbb{E}X_n = \mathbb{E}X_n$ und $\mathbb{E}\mathbb{M}X_n = \mathbb{M}X_n$ liefern für jede der beiden Prämissen die Beziehung $\mathbb{E}X_n = \mathbb{M}X_n + O(1)$. □

Korollar 5. *Für Lipschitz-stetiges $f_n : \mathbb{C} \to \mathbb{R}$ mit $\operatorname{Lip} f_n = O(1)$ gilt*

$$X_n = X'_n + \mathbb{O}(Y_n) \;\Rightarrow\; f(X_n) = f(X'_n) + \mathbb{O}(Y_n).$$

Beweis. Verwende $f(X_n) = f(X'_n) + O(|X_n - X'_n|)$ und $|X_n - X'_n| = \mathbb{O}(Y_n)$. □

Korollar 6. *Für Lipschitz-stetiges $f_n : \mathbb{C} \to \mathbb{R}$ mit $\operatorname{Lip} f_n = O(1)$ gilt*

$$X_n = \mathbb{E}X_n + \mathbb{O}(\theta_n) \;\Rightarrow\; \mathbb{E}f(X_n) = f(\mathbb{E}X_n) + O(\theta_n).$$

Beweis. Aus Korollar 5 folgt $f(X_n) = f(\mathbb{E}X_n) + \mathbb{O}(\theta_n)$; Extraktion liefert schließlich $\mathbb{E}f(X_n) = f(\mathbb{E}X_n) + O(\theta_n)$. □

4. Stochastische Landau-Symbole aus Masskonzentration

Aus den klassischen Maßkonzentrations-Ungleichungen können ganz unmittelbar sehr effektive Formeln der Bauart

$$X_n = \mathbb{E}X_n + \mathbb{O}(\theta_n)$$

hergeleitet werden (die dann als Ausgangspunkt für allerlei Rechnungen entlang der Lemmata und Korollare aus §§2–3 dienen können); dabei ist θ_n signifikant kleiner als die maximale Skala von X_n. Derartige Formeln treten auf, wenn X_n Funktion einer Vielzahl von einzelnen, unabhängigen Zufallselementen ist.

Wir geben drei für unsere Vorlesung wichtige Beispiele. (Die O-Terme in den Voraussetzungen verstehen wir als *gleichmäßig* bezüglich aller freien Variablen.)

Satz 1 (Hoeffding 1963). *Für jedes n seien die Zufallsvariablen $X_1^{(n)}, \dots, X_{m_n}^{(n)}$ unabhängig und es gelte $X_j^{(n)} = O(1)$. Dann ist*

$$S_n = \mathbb{E}S_n + \mathbb{O}(\sqrt{m_n}),$$

wobei $S_n = X_1^{(n)} + \cdots + X_{m_n}^{(n)}$.

Satz 2 (McDiarmid 1989). *Für jedes n seien die Zufallselemente $X_j^{(n)} \in A_j^{(n)}$ ($j = 1, \dots, m_n$) unabhängig. Es seien $F_n : A_1^{(n)} \times \cdots \times A_{m_n}^{(n)} \to \mathbb{C}$ messbare Abbildungen, die bezüglich Änderungen im j-ten Slot maximal um eine Konstante variieren:*

$$F_n(\dots, x_j, \dots) - F(\dots, x_j', \dots) = O(1) \qquad (x_j, x_j' \in A_j^{(n)}).$$

Dann gilt

$$F_n = \mathbb{E}F_n + \mathbb{O}(\sqrt{m_n}),$$

wobei $F_n = F_n(X_1^{(n)}, \dots, X_{m_n}^{(n)})$.

Satz 3 (Talagrand 1995). *Für jedes n seien die Zufallsvariablen $X_1^{(n)}, \dots, X_{m_n}^{(n)}$ unabhängig und es gelte $X_j^{(n)} = O(1)$. Es seien $F_n : \mathbb{C}^{m_n} \to \mathbb{R}$ konvexe und Lipschitz-stetige Funktionen mit $\mathrm{Lip}(F_n) = O(1)$. Dann ist*

$$F_n(X^{(n)}) = \mathbb{E}F_n(X^{(n)}) + \mathbb{O}(1),$$

wobei $F_n(X^{(n)}) = F_n(X_1^{(n)}, \dots, X_{m_n}^{(n)})$.

Aufgrund von Korollar 4 kann der Erwartungswert (im reellwertigen Fall) stets durch den Median ersetzt werden .

5. Beispielanwendungen für Zufallsvektoren und Zufallsmatrizen

Beispiel 1 (Euklidische Norm). Es sei $X_n \in \mathbb{C}^n$ eine Folge von Zufallsvektoren mit gleichmäßig beschränkten Komponenten, die für gegebenes n unabhängig seien. Es gelte $\mathbb{E}|(X_n)_j|^2 = 1$ für alle j und n. Die Hoeffding'sche Ungleichung (Satz 1) liefert unmittelbar

$$\|X_n\|^2 = \mathbb{E}\|(X_n)\|^2 + \mathbb{O}(\sqrt{n}) = n + \mathbb{O}(\sqrt{n}).$$

Aus der einfachen Abschätzung

$$|\|X_n\|^2 - n| = |\|X_n\| - \sqrt{n}| \cdot (\|X_n\| + \sqrt{n}) \geqslant \sqrt{n} \cdot |\|X_n\| - \sqrt{n}|$$

erhalten wir deshalb

$$\boxed{\|X_n\| = \sqrt{n} + \mathbb{O}(1).}$$

Beispiel 2 (Quadratische Formen)**.** Es sei $X_n \in \mathbb{C}^n$ eine Folge von Zufallsvektoren mit gleichmäßig beschränkten und zentrierten Komponenten der Varianz 1, die für gegebenes n unabhängig seien. Ferner sei $A_n \in C^{n\times n}$ eine Folge komplexer Zufallsmatrizen mit gleichmäßig beschränkter Spektralnorm $\|A_n\| = O(1)$; für gegebenes n seien A_n und X_n unabhängig voneinander. Wir studieren die Asymptotik der quadratischen Form $X_n^H A_n X_n$.

1. *Schritt*. A_n sei deterministisch und hermitesch positiv semi-definit. Die reellwertige Funktion $f_n(x) = \sqrt{x^H A_n x} = \|A_n^{1/2} x\|$ ist auf C^n konvex und Lipschitz-stetig mit $\text{Lip} f_n = O(1)$. Die Talagrand'sche Ungleichung (Satz 3) liefert

$$\sqrt{X_n^H A_n X_n} = \mathbb{E}\sqrt{X_n^H A_n X_n} + \mathbb{O}(1);$$

aus Beispiel 1 erhalten wir

$$\sqrt{X_n^H A_n X_n} = \|A_n^{1/2} X_n\| \leqslant \|A_n^{1/2}\| \cdot \|X_n\| = O(1) \cdot \mathbb{O}(\sqrt{n}) = \mathbb{O}(\sqrt{n}).$$

Korollar 3 führt daher auf $X_n^H A_n X_n = \mathbb{E} X_n^H A_n X_n + \mathbb{O}(\sqrt{n})$.

2. *Schritt*. A_n sei deterministisch. Wir zerlegen $A_n = A_n' + iA_n''$ mit $A_n' = (A_n + A_n^H)/2$ und $A_n'' = (A_n - A_n^H)/2i$ hermitesch; A_n' und A_n'' schreiben wir jeweils als Differenz zweier positiv semi-definiter Matrizen. Alle Bestandteile besitzen jetzt eine durch $\|A_n\|$ (also gleichmäßig) beschränkte Spektralnorm. Wegen Linearität bleibt das Resultat aus Schritt 1 daher weiterhin gültig. Wir berechnen zudem aufgrund der Unabhängigkeit, Zentrierung und Normierung der Komponenten von X_n, dass

$$\begin{aligned} \mathbb{E} X_n^H A_n X_n &= \sum_{j,k} \mathbb{E}\overline{(X_n)_j}(A_n)_{jk}(X_n)_k \\ &= \sum_{j\neq k} \mathbb{E}\overline{(X_n)_j}(A_n)_{jk}\mathbb{E}(X_n)_k + \sum_j (A_n)_{jj}\mathbb{E}|(X_n)_j|^2 = \sum_j (A_n)_{jj} = \text{tr} A_n. \end{aligned}$$

3. *Schritt*. A_n sei schließlich stochastisch. Das Resultat aus Schritt 2 lässt sich wegen der Unabhängigkeit von A_n und X_n in der Form einer bedingten Wahrscheinlichkeit schreiben: Es gibt ein $\psi \in \mathcal{T}$, so dass

$$\mathbb{P}(|X_n^H A_n X_n - \text{tr} A_n| \geqslant t\sqrt{n} \; | A_n) \leqslant \psi(t)$$

gleichmäßig in den Realisierungen von A_n. Integration über die Wahrscheinlichkeitsverteilung der A_n ergibt daher

$$\boxed{X_n^H A_n X_n = \mathbb{O}(n) \quad \text{und} \quad X_n^H A_n X_n = \text{tr} A_n + \mathbb{O}(\sqrt{n}).}$$

Bemerkung 2. Nach Borell/Sudakov/Tsirel´son ist Satz 3 auch für standard-normalverteilte Zufallsvariablen richtig ist (sogar ohne die Voraussetzung der Konvexität von F_n). Deshalb bleiben die Resultate beider Beispiele auch dann richtig, wenn wir die Voraussetzung der gleichmäßigen Beschränkung der Komponenten der Zufallsvektoren X_n durch die Bedingung ersetzen, dass jene Komponenten standard-normalverteilt sind.

Wir bemerken im übrigen, dass jede Folge ξ_n normalverteilter Zufallsvariablen $\xi_n = \mathbb{E}\xi_n + \mathbb{O}(\text{var}\,\xi_n)^{1/2}$ erfüllt, weil $\mathbb{P}(|\xi_n - \mathbb{E}\xi_n| \geqslant t(\text{var}\,\xi_n)^{1/2}) \leqslant e^{-t^2/2}$ gilt.