5 Zentraler Grenzwertsatz von Lindeberg-Lévy

5.1 Charakteristische Funktionen

Definition
Es sei X Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Dann heißt

$$\phi_X(t) := Ee^{itX} = E\cos(tX) + iE\sin(tX)$$

*die **charakteristische Funktion** zu X.*

Bemerkung
Ist X diskret mit Werten $x_1, x_2, \ldots$, so gilt:

$$\phi_X(t) = \sum_{k=1}^{\infty} e^{itx_k} \cdot P(X = x_k)$$

Ist X absolutstetig mit Dichte f, so gilt:

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{itx} f(x) \mathrm{d}x \quad \text{(Fourier-Transformation)}$$

Beispiel 5.1

a) $X \sim B(n, p)$

$$\phi_X(t) = \sum_{k=0}^{n} e^{itk} \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} (pe^{it})^k (1-p)^{n-k} = (1 - p + pe^{it})^n$$

b) $X \sim U(0, 1)$
$\phi_X(0) = 1$ und für $t \neq 0$:

$$\phi_X(t) = \int_0^1 e^{itx} \cdot 1\mathrm{d}x = \int_0^1 \cos(tx)\mathrm{d}x + i\int_0^1 \sin(tx)\mathrm{d}x$$

$$= \frac{1}{t}\sin(t) - \frac{i}{t}\cos(t) + \frac{i}{t} = \frac{1}{it}(e^{it} - 1)$$

c) $X \sim N(0,1)$

$$\phi_X(t) = e^{-\frac{t^2}{2}} \quad \text{vgl. Stochastik 1}$$

Satz 5.1 *Sind X, Y unabhängige Zufallsvariablen mit charakteristischen Funktionen ϕ_X und ϕ_Y, so gilt für die charakteristische Funktion ϕ_{X+Y} der Faltung:*

$$\phi_{X+Y}(t) = \phi_X(t) \cdot \phi_Y(t) \quad \forall t \in \mathbb{R}$$

Beweis vgl. Stochastik 1, Satz 12.2. ∎

Lemma 5.1 *Für alle $m \in \mathbb{N}, t \in \mathbb{R}$ gilt:*

$$\left| e^{it} - \sum_{k=0}^{m-1} \frac{(it)^k}{k!} \right| \leq \min \left\{ \frac{|t|^m}{m!}, \frac{2|t|^{m-1}}{(m-1)!} \right\}$$

Beweis vgl. Stochastik 1, Satz 13.2. ∎

5.2 Umkehrsätze

Wir werden sehen, dass eine Verteilung eindeutig durch ihre charakteristische Funktion festgelegt ist. Hat man z.B. gezeigt, dass X die charakteristische Funktion $(1-p+pe^{it})^n$ hat, so ist $X \sim B(n,p)$.
Aus der Analysis ist die Integralsinusfunktion bekannt:

$$Si : \mathbb{R}_+ \to \mathbb{R}_+,\ Si(x) := \int_0^x \frac{\sin(y)}{y} \mathrm{d}y \quad \forall x > 0$$

Es gilt: $\lim_{x\to\infty}(Si(x)) = \frac{\pi}{2}$

Satz 5.2
Es sei X Zufallsvariable mit charakteristischer Funktion ϕ_X. Dann gilt für alle $-\infty < a < b < \infty$:

$$\frac{1}{2}P(X=a) + P(a < X < b) + \frac{1}{2}P(X=b) = \lim_{T\to\infty} \left(\frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \phi_X(t)\, dt \right)$$

Beweis
Sei $I(T) := \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita}-e^{-itb}}{it} \phi_X(t)\mathrm{d}t$. Definiere $\psi : \mathbb{R} \times [-T,T] \to \mathbb{C}$ durch

$$\psi(t,x) := \begin{cases} \frac{e^{-it(a-x)} - e^{-it(b-x)}}{it}, & t \neq 0 \\ b-a, & t = 0 \end{cases}$$

Mit Lemma 5.1 folgt, dass ψ stetig ist und wegen

$$\left| \frac{e^{-ita} - e^{-itb}}{it} \right| = \left| \int_a^b e^{ity} \mathrm{d}y \right| \leq b - a$$

ist $|\psi| \leq b-a$, also ist ψ $P^X \otimes \lambda_{[-T,T]}$-integrierbar. Mit Satz 3.1 (Fubini I) folgt:

$$I(T) = \frac{1}{2\pi}\int_{-T}^{T} \frac{e^{-ita}-e^{-itb}}{it}\left(\int e^{itx}P^X(\mathrm{d}x)\right)\mathrm{d}t$$

$$= \frac{1}{2\pi}\int \underbrace{\int_{-T}^{T} \frac{1}{it}\left(e^{-it(a-x)}-e^{-it(b-x)}\right)\mathrm{d}t}_{=:\psi_{a,b,T}(x)}\, P^X(\mathrm{d}x)$$

Inneres Integral:

Da $t \mapsto \frac{\cos(t(x-a))}{it}$ punktsymmetrisch ist, gilt:

$$\psi_{a,b,T}(x) = 2\cdot\int_0^T \frac{1}{t}\sin\left((x-a)\,t\right)\mathrm{d}t - 2\cdot\int_0^T \frac{1}{t}\sin\left((x-b)\,t\right)\mathrm{d}t$$

Es gilt weiterhin:

$$c\cdot\int_0^T \frac{1}{c\cdot t}\sin(ct)\mathrm{d}t = \mathrm{sgn}(c)\cdot Si(T|c|) \quad \text{mit}\,\mathrm{sgn}(c) = \begin{cases} 1, & c>0 \\ 0, & c=0 \\ -1, & c<0 \end{cases}$$

$$\implies \psi_{a,b,T}(x) = 2\cdot\mathrm{sgn}(x-a)Si(T|x-a|) - 2\cdot\mathrm{sgn}(x-b)Si(T|x-b|)$$

$$\implies \psi_{a,b}(x) := \lim_{T\to\infty}\left(\psi_{a,b,T}\left(x\right)\right) = \begin{cases} 0, & x<a \text{ oder } x>b \\ \pi, & x=a \text{ oder } x=b \\ 2\pi, & a<x<b \end{cases}$$

$\implies$ $(\psi_{a,b,T})_{T\geq 0}$ besitzt eine (konstante) integrierbare Majorante. Mit dem Satz über die majorisierte Konvergenz gilt:

$$\begin{aligned} \lim_{T\to\infty} I(T) &= \frac{1}{2\pi}\int \psi_{a,b}(x)P^X(\mathrm{d}x) \\ &= \frac{1}{2}P(X=a) + \frac{1}{2}P(X=b) + P(a<X<b) \end{aligned}$$

■

Korollar 5.1

Sind X und Y Zufallsvariablen mit derselben charakteristischen Funktion, so haben X und Y dieselbe Verteilung.

Beweis Sei $D = A(X)\cup A(Y)$ mit $A(X) = \{x\in\mathbb{R}|P(X=x)>0\}$, analog $A(Y)$. $A(X)$ ist abzählbar, da $A(X) = \bigcup_{n=1}^{\infty}\{x\in\mathbb{R}|P(X=x)\geq\frac{1}{n}\}$ und $\left|\{x\in\mathbb{R}|P(X=x)\geq\frac{1}{n}\}\right| \leq n \implies D$ abzählbar

$$\mathcal{D} := \{(a,b)| -\infty < a \leq b < \infty, a,b\notin D\}$$

ist ein durchschnittstabiles Erzeugendensystem von $\mathfrak{B}(\mathbb{R})$. $\overset{\text{Sa5.2}}{\Longrightarrow}$ P^X und P^Y stimmen auf $\mathcal{D}$ überein $\xRightarrow{\text{Eindeutigkeitssatz}}$ Behauptung. ■

Satz 5.3
Sei X eine Zufallsvariable mit charakteristischer Funktion ϕ. Gilt $\int |\phi(t)|dt < \infty$, so hat X eine stetige Dichte f, die gegeben ist durch

$$f(x) = \frac{1}{2\pi}\int e^{-itx}\phi(x)\,dt \quad \forall x \in \mathbb{R}$$

Beweis Wie in Beweis von Satz 5.2 gilt:

$$\left|\frac{e^{-ita} - e^{-itb}}{it}\right| \leq |b-a| \quad (*)$$

Da ϕ λ-integrierbar ist, ist $|b-a||\phi|$ eine integrierbare Majorante für diesen Ausdruck in Satz 5.2. Es folgt:

$$\frac{1}{2}P(X=a) + P(a < X < b) + \frac{1}{2}P(X=b) = \frac{1}{2\pi}\int \frac{e^{-ita} - e^{-itb}}{it}\phi(t)\mathrm{d}t$$

$$\begin{aligned}
\Longrightarrow \quad P(a < X < b) \quad &\leq \frac{1}{2\pi}|b-a| \underbrace{\int |\phi(t)|\mathrm{d}t}_{<\infty} \\
\Longrightarrow \quad P(X = x) \quad &= \lim_{n\to\infty} P(x - \frac{1}{n} < X < x + \frac{1}{n}) \\
&= 0
\end{aligned}$$

Ist F die Verteilungsfunktion von X, so gilt:

$$F(b) - F(a) = \frac{1}{2\pi}\int \frac{e^{-ita} - e^{-itb}}{it}\phi(t)\mathrm{d}t \quad \forall a < b$$

Wegen $(*)$ kann man den Satz von der majorisierten Konvergenz anwenden und bekommt:

$$\begin{aligned}
\lim_{h\downarrow 0}\frac{F(x+h) - F(x)}{h} \quad &= \quad \frac{1}{2\pi}\int e^{-itx}\lim_{h\downarrow 0}\frac{1 - e^{-ith}}{ith}\phi(t)\mathrm{d}t \\
&= \quad \frac{1}{2\pi}\int e^{-itx}\phi(t)\mathrm{d}t \\
&=: \quad f(x)
\end{aligned}$$

Außerdem folgt $x \mapsto f(x)$ ist stetig. ■

5.3 Verteilungskonvergenz

Definition

a) *Gegeben sei der messbare Raum* $(\mathbb{R}, \mathfrak{B})$ *mit Wahrscheinlichkeitsmaßen* $P, P_1, P_2, \ldots$ *und zugehörigen Verteilungsfunktionen* $F, F_1, F_2, \ldots$
P_n ***konvergiert schwach*** *gegen* P ($P_n \overset{w}{\to} P$), *wenn* $\lim_{n\to\infty} F_n(x) = F(x)\ \forall x \in \mathbb{R}$ *an denen* F *stetig ist.*

b) *Seien* $X, X_1, X_2, \ldots$ *Zufallsvariablen auf (unter Umständen verschiedenen) Wahrscheinlichkeitsräumen* $(\Omega, \mathcal{A}, P), (\Omega_1, \mathcal{A}_1, P_1), \ldots$
X_n ***konvergiert in Verteilung*** *gegen* X ($X_n \overset{d}{\to} X$), *wenn* $P^{X_n} \overset{w}{\to} P^X$.

Beispiel 5.2 Konvergenz in Verteilung bzw. schwache Konvergenz ist schwächer als f.s.-Kovergenz.
Sei z.B. $X \sim N(0,1)$ und $X_{2n} = X, X_{2n+1} = -X\ \forall n \in \mathbb{N}$. $\implies P^{X_n} \equiv P^X = N(0,1)$ und (X_n) konvergiert in Verteilung (gegen X) jedoch $X_n \overset{f.s.}{\not\to} X$

Jedoch gilt folgender nützlicher Satz:

Satz 5.4 (Darstellungssatz von Skorohod)
Es seien $X, X_1, X_2, \ldots$ *Zufallsvariablen mit* $X_n \overset{d}{\to} X$. *Dann existiert ein Wahrscheinlichkeitsraum* $(\Omega, \mathcal{A}, P)$ *und hierauf Zufallsvariablen* $X', X_1', X_2', \ldots$ *mit* $X' \overset{d}{=} X, X_n' \overset{d}{=} X_n\ \forall n \in \mathbb{N}$ *derart, dass* $X_n' \overset{f.s.}{\to} X'$.

Beweis Es seien $F, F_1, F_2, \ldots$ die Verteilungsfunktionen zu $X, X_1, X_2, \ldots$ und $(\Omega, \mathcal{A}, P) = \big((0,1), \mathfrak{B}_{(0,1)}, \lambda_{(0,1)}\big)$. Weiter sei $F^{-1} : (0,1) \to \mathbb{R}, F^{-1}(y) := \inf\{x \in \mathbb{R} | F(x) \geq y\}$ die Quantilsfunktion zu F, analog (Quantilsfunktion) $F_n^{-1}, n \in \mathbb{N}$. Setze $X' := F^{-1},\ X_n' := F_n^{-1}$.
Satz 5.7 (Stoch 1) $\implies X' \overset{d}{=} X, X_n' \overset{d}{=} X_n, n \in \mathbb{N}$ $(P(X' \leq x) = P(F^{-1}(\omega) \leq x) = \underbrace{P(\omega \leq F(x))}_{=\lambda(0,1)} = F(x)\ )$
Es bleibt also zu zeigen , dass für P-fast alle $\omega \in \Omega : \lim_{n\to\infty} X_n'(\omega) = X'(\omega)$.
Sei $\omega \in (0,1)$. Da X nur abzählbar viele Atome hat (vgl. Beweis von Korollar 5.1) existiert zu $\varepsilon > 0$ ein $x \in \mathbb{R}$ mit $X'(\omega) - \varepsilon < x < X'(\omega)$ und $P(X = x) = 0$.
Es gilt (Lemma 5.6, Stoch 1): $\forall y \in (0,1), x \in \mathbb{R}$:

$$y \leq F(x) \iff F^{-1}(y) \leq x$$

Hier: $\omega \leq F(x) \iff F^{-1}(\omega) = X'(\omega) \leq x$. Wegen $X'(\omega) > x$ folgt $F(x) < \omega$. Da $F_n(x) \to F(x)$ für $n \to \infty$ nach Voraussetzung, $\exists n_0 \in \mathbb{N}$, so dass $\forall n \geq n_0 : F_n(x) < \omega$. Also $X_n' > x$.
Mit $\varepsilon \downarrow 0$ folgt:

$$\liminf_{n\to\infty} X_n'(\omega) \geq X'(\omega) \quad \forall \omega \in \Omega.$$

Ist $\omega' > \omega$ und $\varepsilon > 0$, so $\exists$ ein x mit $X'(\omega') < x < X'(\omega') + \varepsilon$ und $P(X = x) = 0$. Da F rechtsseitig stetig, folgt $F(F^{-1}(y)) \geq y\ \forall y \in (0,1)$, also mit der Monotonie von $F : \omega < \omega' \leq F(X'(\omega')) \leq F(x)$.

Wegen $F_n(x) \to F(x)$ $(n \to \infty)$, $\exists n_0 \in \mathbb{N}$ sodass $\omega \leq F_n(x)$ (d.h. $X'_n(\omega) \leq x$) $\forall n \geq n_0$ gilt mit $\varepsilon \downarrow 0$ ergibt das

$$\limsup_{n\to\infty} X'_n(\omega) \leq X'(\omega') \quad \forall \omega' > \omega.$$

Satz 5.5 *Es sei $C_b(\mathbb{R})$ die Menge aller stetigen und beschränkten Funktionen, $h : \mathbb{R} \to \mathbb{R}$. Dann gilt:*

$$X_n \xrightarrow{d} X \iff Eh(X_n) \to Eh(X) \quad \forall h \in C_b(\mathbb{R})$$

Beweis
"$\Rightarrow$": Nach Satz 5.4 existieren ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und Zufallsvariablen $X' \overset{d}{=} X$, $X'_n \overset{d}{=} X_n \quad \forall n \in \mathbb{N}$ mit $X'_n \xrightarrow{f.s.} X'$. Es folgt:

$$\lim_{n\to\infty} \left(Eh\left(X_n\right)\right) = \lim_{n\to\infty} \left(Eh\left(X'_n\right)\right) \overset{\text{h stetig}, X'_n \xrightarrow{f.s.} X'}{=} Eh\left(X'\right) = Eh\left(X\right).$$

"$\Leftarrow$": Für $a, b \in \mathbb{R}, a < b$ sei $h_{a,b} : \mathbb{R} \to \mathbb{R}$ definiert durch

$$h_{a,b}\left(x\right) := \begin{cases} 1 & , x \leq a \\ \frac{b-x}{b-a} & , a < x < b \\ 0 & , x \geq b \end{cases}$$

$h_{a,b}$ ist stetig und beschränkt. Seien F, F_n die Verteilungsfunktionen zu $X, X_n \quad \forall n \in \mathbb{N}$. Dann gilt $\forall y > x$:

$$F_n\left(x\right) = E\left[\mathbf{1}_{(-\infty,x)}\left(X_n\right)\right] \leq E\left[h_{x,y}\left(X_n\right)\right] \overset{n\to\infty}{\to} E\left[h_{x,y}\left(X\right)\right],$$

$$E\left[h_{x,y}\left(X\right)\right] \leq E\left[\mathbf{1}_{(-\infty,y)}\left(X\right)\right] = F\left(y\right).$$

Also folgt da F rechtsseitig stetig ist mit $y \downarrow x$:

$$\limsup_{n\to\infty} \left(F_n\left(x\right)\right) \leq F\left(x\right) \quad \forall x \in \mathbb{R}$$

Analog erhält man für $y < x$:

$$F_n\left(x\right) \geq E\left[h_{y,x}\left(X_n\right)\right] \overset{n\to\infty}{\to} E\left[h_{y,x}\left(X\right)\right] \geq F\left(y\right)$$

Mit $y \uparrow x$: $\liminf_{n\to\infty}(F_n(x)) \geq F(x-) \quad \forall x \in \mathbb{R}$.
Ist F in x stetig, so gilt $F(x-) = F(x)$ und somit $F_n(X) \overset{n\to\infty}{\to} F(x)$. ■

Satz 5.6 ("Continuous Mapping Theorem")
Es seien $X, X_1, X_2, \ldots$ Zufallsvariablen mit $X_n \xrightarrow{d} X$. Weiter sei $f : \mathbb{R} \to \mathbb{R}$ eine Borel-messbare Funktion mit $P(X \in \{x \in \mathbb{R} \mid f \text{ nicht stetig in } x\}) = 0$.
Dann gilt auch $f(X_n) \xrightarrow{d} f(X)$ für $n \to \infty$.

Beweis Übung. ■

Satz 5.7 (Satz von Helly[1])
Zu jeder Folge $(F_n)_{n\in\mathbb{N}}$ *von Verteilungsfunktionen existieren eine Teilfolge* $(F_{n_k})_{k\in\mathbb{N}}$ *und eine schwach monoton wachsende, rechtsseitig stetige Funktion* $G : \mathbb{R} \to [0,1]$, *sodass* $\lim_{k\to\infty}(F_{n_k}(x)) = G(x) \quad \forall x \in \mathbb{R}$, *an denen* G *stetig ist.*

Beweis (Skizze)
Für $x \in \mathbb{R}$ ist $(F_n(x))_{n\in\mathbb{N}} \subset [0,1] \quad \xRightarrow{\text{Bolzano-Weierstraß}} \exists$ Häufungspunkt. Sei $(r_k)_{k\in\mathbb{N}}$ eine Abzählung von $\mathbb{Q}$. Wähle Teilfolgen $(F_{n_{k,j}})_{j\in\mathbb{N}}$ mit $F_{n_{k,j}} \overset{j\to\infty}{\to} G_0(r_k)$, wobei $(n_{k+1,j})_{j\in\mathbb{N}}$ eine Teilfolge von $(n_{k,j})_{j\in\mathbb{N}}$ ist. (Definition der Funktion f_0 auf $\mathbb{Q}$)
Für die Diagonalfolge $(n_{j,j})_{j\in\mathbb{N}}$ gil dann: $F_{n_{j,j}} \to G_0$ auf $\mathbb{Q}$. Sei G_0 auf ganz $\mathbb{R}$ durch $G(x) := \inf\{G_0(r) \mid r \in \mathbb{Q}, r > x\}$ fortgesetzt.
Rest: $\epsilon - \delta$-Argumente. ■

Bemerkung G aus Satz 5.7 muß keine Verteilungsfunktion sein.
Beispiel: $F_n := \mathbf{1}_{[n,\infty]} \implies G \equiv 0$

Definition *Eine Familie* $\mathcal{P}$ *von Wahrscheinlichkeitsmaßen auf* $(\mathbb{R}, \mathfrak{B})$ *heißt* ***straff****, wenn* $\forall \epsilon > 0\ \exists$ *kompaktes Intervall* $[a,b] \subset \mathbb{R}$ *mit:*

$$P([a,b]) \geq 1 - \epsilon \quad \forall P \in \mathcal{P}$$

Bemerkung

(i) Ist $\mathcal{P}$ straff, so auch jedes $\mathcal{P}' \subset \mathcal{P}$.

(ii) Sind alle $\mathcal{P}_i$ mit $i \in \{1, \ldots, , n\}$ straff, so auch $\bigcup_{i=1}^n \mathcal{P}_i$.

(iii) Ist $|\mathcal{P}| = 1$, so ist $\mathcal{P}$ straff.

Satz 5.8 *Ist* $\{P_n \mid n \in \mathbb{N}\}$ *eine straffe Familie von Wahrscheinlichkeitsmaßen auf* $(\mathbb{R}, \mathfrak{B})$, *so existieren eine Teilfolge* $(P_{n_k})_{k\in\mathbb{N}}$ *von* $(P_n)_{n\in\mathbb{N}}$ *und ein Wahrscheinlichkeitsmaß* P *derart, dass* $P_{n_k} \overset{w}{\to} P$ *für* $k \to \infty$.

Beweis Sei F_n die Verteilungsfunktion zu $P_n\ \forall n \in \mathbb{N}$.
$\xRightarrow{\text{Satz 5.7}} \exists$ Folge $(n_k)_{k\in\mathbb{N}}$ mit $F_{n_k}(x) \to G(x)$ für $k \to \infty\ \forall x \in \mathbb{R}$ mit G stetig in x; G ist wachsend und rechtsseitig stetig.
Bleibt zu zeigen: G ist Verteilungsfunktion, also $\lim_{x\to-\infty}(G(x)) = 0$ und $\lim_{x\to\infty}(G(x)) = 1$. Ist dann P das Wahrscheinlichkeitsmaß zu G, so folgt $P_{n_k} \overset{w}{\to} P$.
Sei also $\epsilon > 0$. Da $\{P_n \mid n \in \mathbb{N}\}$ straff ist $\implies \exists a, b \in \mathbb{R}$ mit $P_n([a,b]) \geq 1 - \epsilon\ \forall n \in \mathbb{N}$. $\implies F_n(a) \leq \epsilon\ \forall n \in \mathbb{N}$.
G hat höchstens abzählbar viele Unstetigkeitsstellen. $\implies \exists c < a$, in dem G stetig.
$\implies G(c) = \lim_{k\to\infty}(F_{n_k}(c)) \leq \epsilon \implies G(x) \leq \epsilon\ \forall x \leq c$.
Also: $\forall \epsilon > 0 \quad \exists c \in \mathbb{R} : \quad \forall x \leq c$ gilt $0 \leq G(x) \leq \epsilon \implies \lim_{x\to-\infty}(G(x)) = 0$ und $\lim_{x\to\infty}(G(x)) = 1$. ■

Satz 5.9 (Stetigkeitssatz für charakteristische Funktionen)
Es seien $X, X_1, X_2, \ldots$ Zufallsvariablen, $\phi, \phi_1, \phi_2, \ldots$ die zugehörigen charakteristischen Funktionen. Dann gilt:

$$X_n \xrightarrow{d} X \quad \Longleftrightarrow \quad \phi_n(t) \to \phi(t) \quad \forall t \in \mathbb{R}$$

Beweis
"$\Rightarrow$": Sei $t \in \mathbb{R}$. $x \mapsto \cos(tx),\ x \mapsto \sin(tx)$ sind stetig und beschränkt.
$\overset{\text{Satz 5.5}}{\Longrightarrow}$ $\phi_n(t) = E\cos(tX_n) + iE\sin(tX_n) \to E\cos(tX) + iE\sin(tX) = \phi(t)$.
"$\Leftarrow$": Wir zeigen zunächst: $\{P^{X_n}, n \in \mathbb{N}\}$ ist straff. $\mathbb{C}$-wertige Version von Fubini II liefert $\forall\, \delta > 0$.

$$\begin{aligned}
\frac{1}{\delta}\int_{-\delta}^{\delta}(1-\varphi_n(t))\mathrm{d}t &= \int(\frac{1}{\delta}\int_{-\delta}^{\delta}(1-e^{itx})\mathrm{d}t)P^{X_n}(\mathrm{d}x) \\
&= 2\int \underbrace{(1-\frac{\sin(\delta x)}{\delta x})}_{\geq 0} P^{X_n}(\mathrm{d}x) \\
&\geq 2\int_{|x|\geq\frac{2}{\delta}} \underbrace{(1-\frac{1}{|\delta x|})}_{\geq\frac{1}{2}} P^{X_n}(\mathrm{d}x) \\
&\geq P^{X_n}([-\frac{2}{\delta}, \frac{2}{\delta}]^C)
\end{aligned}$$

Sei $\varepsilon > 0$. Da φ in 0 stetig und $\varphi(0) = 1,\ \exists\, \delta > 0$:

$$|1-\varphi(t)| \leq \frac{\varepsilon}{4} \quad \forall\, |t| \leq \delta$$

$\Rightarrow |\frac{1}{\delta}\int_{-\delta}^{\delta}(1-\varphi(t))\mathrm{d}t| \leq \frac{1}{\delta}2\delta\frac{\varepsilon}{4} = \frac{\varepsilon}{2}$. Da $|\varphi_n| \leq 1$ folgt mit majorisierter Konvergenz:

$$\int_{-\delta}^{\delta}(1-\varphi_n(t))\mathrm{d}t \overset{n\to\infty}{\to} \int_{-\delta}^{\delta}(1-\varphi(t))\mathrm{d}t$$

$\Rightarrow \exists\, n_0 \in \mathbb{N}$, so dass $\frac{1}{\delta}\int_{-\delta}^{\delta}(1-\varphi_n(t))\mathrm{d}t \leq \varepsilon\ \forall n \geq n_0$. $\Rightarrow P^{X_n}([-\frac{2}{\delta}, \frac{2}{\delta}]) \geq 1-\varepsilon\ \forall n \geq n_0$.
Außerdem: $\forall\, n \in \{1, \ldots, n_0-1\}\ \exists\, a_n > 0$ mit $P^{X_n}([-a_n, a_n]) \geq 1-\varepsilon$ da $P^{X_n}([-m, m]) \to 1$ für $m \to \infty$.
Insgesammt: Sei $a := \max\{a_1, \ldots, a_{n_0-1}, \frac{2}{\delta}\} \Rightarrow P^{X_n}([-a, a]) \geq 1-\varepsilon\ \forall n \in \mathbb{N} \Rightarrow \{P^{X_n}, n \in N\}$ ist straff.

Annahme: $X_n \xrightarrow{d} X$ gilt nicht.
$\Rightarrow \exists\, x \in \mathbb{R}$ mit $P(X = x) = 0$ und $P(X_n \leq x) \not\to P(X \leq x), n \to \infty$.
d.h. $\exists\, \varepsilon > 0$ und eine Teilfolge $(X_{n_k})_{k\in\mathbb{N}}$ mit $|P(X_{n_k} \leq x) - P(X \leq x)| \geq \varepsilon\ \forall\, k \in \mathbb{N}$ $(*)$.
$\{P^{X_{n_k}}, k \in \mathbb{N}\}$ ist ebenfalls straff $\overset{S.5.8}{\Rightarrow}$ $\exists$ Teilfolge $(X_{n_{k_j}})_{j\in\mathbb{N}}$ und ein W'maß P_0 mit $P^{X_{n_{k_j}}} \xrightarrow{w} P_0$.

Sei φ_0 charakteristische Funktion zu P_0. Also folgt mit der Hinrichtung: $\varphi_{n_{k_j}}(t) \to \varphi_0(t) = \varphi(t) \overset{Kor. 5.1}{\Rightarrow} P_0 = P^X$, also $X_{n_{k_j}} \xrightarrow{d} X$ und damit $P(X_{n_{k_j}} \leq x) \to P(X \leq x)$. Wid zu $(*)$. ■

Wir benötigen noch folgendes technisches Hilfslemma:

Lemma 5.2
Für alle $z_1, \ldots, z_n, w_1, \ldots, w_n \in \{z \in \mathbb{C} | |z| \leq 1\}$ gilt:

$$|\prod_{k=1}^n z_k - \prod_{k=1}^n w_k| \leq \sum_{k=1}^n |z_k - w_k|$$

Beweis

$$\begin{aligned} |\prod_{k=1}^n z_k - \prod_{k=1}^n w_k| &\leq |\prod_{k=1}^n z_k - w_1 \prod_{k=2}^n z_k| + |w_1 \prod_{k=2}^n z_k - w_1 w_2 \prod_{k=3}^n z_k| + \cdots + |w_1 \ldots w_{n-1} z_n - \prod_{k=1}^n w_k| \\ &= |z_1 - w_1| \underbrace{|\prod_{k=2}^n z_k|}_{\leq 1} + |z_2 - w_2| \underbrace{|w_1 \prod_{k=3}^n z_k|}_{\leq 1} + \cdots + |z_n - w_n| \underbrace{|\prod_{k=1}^{n-1} w_k|}_{\leq 1} \end{aligned}$$ ■

Hauptsatz des Abschnitts:

Satz 5.10 (Zentraler Grenzwertsatz von Lindeberg-Lévy)
Für jedes $n \in \mathbb{N}$ seien $X_{nk}, k = 1, \ldots, r_n$ unabhängige Zufallsvariablen (nicht notwendig identisch verteilt) auf einem Wahrscheinlichkeitsraum $(\Omega_n, \mathcal{A}_n, P_n)$ mit $\mathrm{Var}(X_{nk}) = \sigma_{nk}^2 < \infty$ und $EX_{nk} = \mu_{nk} < \infty$. Es sei $s_n^2 := \sum_{k=1}^{r_n} \sigma_{nk}^2 > 0$. Ist dann für alle $\varepsilon > 0$ die ***Lindeberg-Bedingung***

$$(L) \quad \lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=1}^{r_n} \int_{|X_{nk}-\mu_{nk}|>\varepsilon s_n} (X_{nk} - \mu_{nk})^2 \, dP_n = 0$$

erfüllt, so gilt mit $n \to \infty$:

$$\frac{1}{s_n} \sum_{k=1}^{r_n} (X_{nk} - \mu_{nk}) \xrightarrow{d} Z \ , \ \ Z \sim N(0,1)$$

Bemerkung 5.1

1. Die Lindeberg-Bedingung schließt einen dominierenden Einfluss eines einzelnen Summanden X_{nk} auf die $X_{n1} + \cdots + X_{nr_n}$ aus. Insbesondere gilt:
$$\max\{\sigma_{nk}^2 | 1 \leq k \leq r_n\} = o(s_n^2) \text{ für } n \to \infty$$
2. Der ZGWS hat eine lange „Verbesserungsgeschichte“hinter sich. Gelegentlich ist die Lyapunov-Bedingung einfacher zu verwenden:
$$\lim_{n\to\infty} \frac{1}{s_n^{2+\delta}} \sum_{k=1}^{r_n} E(|X_{nk} - \mu_{nk}|^{2+\delta}) = 0 \text{ für ein } \delta > 0$$

3. Der Satz liefert eine Begründung für die „Allgegenwart“ der Normalverteilung.

Ein wichtiger Spezialfall ist

Satz 5.11 (ZGWS St. I)
Es seien $Y_1, Y_2, \dots$ *u.i.v. ZV mit* $EY_1 = \mu < \infty$ *und* $0 < \mathrm{Var}(Y_1) = \sigma^2 < \infty$. *Dann gilt:*

$$\frac{Y_1 + \cdots Y_n - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} Z \sim N(0,1)$$

Beweis Sei $X_{nk} := Y_k, r_n = n, (\Omega_n, \mathcal{A}_n, P_n) = (\Omega, \mathcal{A}, P)$. Es gilt: $s_n^2 = n\sigma^2$ und

$$\frac{1}{s_n^2} \sum_{k=1}^{r_n} \int_{|X_{nk}-\mu_{nk}|>\varepsilon s_n} (X_{nk} - \mu_{nk})^2 \mathrm{d}P_n = \frac{1}{\sigma^2} \int_{|Y_1-\mu_1|>\varepsilon\sqrt{n}\sigma} (Y_1 - \mu_1)^2 \mathrm{d}P =: I_n$$

Da $z_n := \mathbf{1}_{(\varepsilon\sqrt{n}\sigma,\infty)}(|Y_1 - \mu_1|)(Y_1 - \mu_1)^2 \leq (Y_1 - \mu_1)^2$ und $\lim_{n\to\infty} z_n = 0$ folgt mit majorisierter Konvergenz, dass $\lim_{n\to\infty} I_n = 0$. Also ist die Lindeberg-Bedingung erfüllt und die Behauptung folgt mit Satz 5.10. ∎

Beweis Beweis von Satz 5.10
O.B.d.A: $\mu_{nk} = 0$ und $s_n = 1$. Anderfalls ersetze X_{nk} durch $\frac{X_{nk}-\mu_{nk}}{s_n}$.
Idee: Verwende S.5.9: Sei φ_{nk} die charakteristische Funktion von X_{nk} und φ_{s_n} die von $\sum_{n=1}^{r_n} X_{nk} : \varphi_{s_n}(t) = \prod_{k=1}^{r_n} \varphi_{nk}(t) \to \varphi_z(t) = e^{-\frac{t^2}{2}}$
Zu zeigen:

$$\prod_{k=1}^{r_n} \phi_{n_k}(t) \to \phi_z(t) = e^{-\frac{t^2}{2}} \quad \forall t \in \mathbb{R}$$

Mit Lemma 5.1 ($m = 3$):

$$\begin{aligned} \left| e^{itx} - (1 + itx - \frac{1}{2}t^2x^2) \right| &\leq \min\{\frac{|tx|^3}{3!}, |tx|^2\} \quad \forall x \in \mathbb{R} \\ &\leq \min\{|tx|^3, |tx|^2\} \end{aligned}$$

Integral über x liefert (beachte: $EX = 0$)

$$\left| \phi_{n_k}(t) - (1 - \frac{1}{2}t^2\sigma_{n_k}^2) \right| \leq E \min\{|tX_{n_k}|^2, |tX_{n_k}|^3\} =: M_{n_k}$$

Sei $\varepsilon > 0$ beliebig. Es gilt

$$\begin{aligned} M_{n_k} &\leq \int_{|X_{n_k}|\leq\varepsilon} |tX_{n_k}|^3 \mathrm{d}P_n + \int_{|X_{n_k}|>\varepsilon} |tX_{n_k}|^2 \mathrm{d}P_n \\ &\leq |t|^3 \varepsilon \sigma_{n_k}^2 + t^2 \int_{|X_{n_k}|>\varepsilon} X_{n_k}^2 \mathrm{d}P_n \end{aligned}$$

$$\Longrightarrow \sum_{k=1}^{r_n} M_{n_k} \leq |t|^3 \varepsilon + t^2 \sum_{k=1}^{r_n} \int_{|X_{n_k}|>\varepsilon} X_{n_k}^2 \mathrm{d}P_n \overset{n\to\infty}{\to} \varepsilon |t|^3 + 0 \quad \text{folgt mit } (L)$$

Mit $\varepsilon \downarrow 0$ folgt:

$$\lim_{n\to\infty} \sum_{k=1}^{r_n} \left| \phi_{n_k}(t) - (1 - \frac{1}{2}t^2\sigma_{n_k}^2) \right| = 0 \quad \forall t \in \mathbb{R} \quad (1)$$

Behauptung: $\lim_{n\to\infty} \left| \prod_{k_n}^{r_n} \phi_{n_k}(t) - \prod_{k=1}^{r_n}(1 - \frac{1}{2}t^2\sigma_{n_k}^2) \right| = 0 \quad \forall t \in \mathbb{R} \quad (2)$
Beweis: $\forall \varepsilon > 0$ gilt:

$$\sigma_{n_k}^2 \le \int_{|X_{n_k}|>\varepsilon} X_{n_k}^2 \mathrm{d}P_n + \varepsilon^2$$

$$\begin{aligned} \implies \quad & \limsup_{n\to\infty} \max\{\sigma_{n_k}^2 | 1 \le k \le r_n\} \\ \le \quad & \lim_{n\to\infty} \left(\varepsilon^2 + \sum_{k=1}^{r_n} \int_{|X_{n_k}|>\varepsilon} X_{n_k}^2 \mathrm{d}P_n \right) \\ \overset{(L)}{=} \quad & \varepsilon^2 + 0 \end{aligned}$$

Mit $\varepsilon \downarrow 0$:

$$\lim_{n\to\infty} \max\{\sigma_{n_k}^2 | 1 \le k \le r_n\} = 0 \quad (3)$$

$\implies$ $\forall t \in \mathbb{R}, \exists n_0 \in \mathbb{N}$, so dass $\forall n \ge n_0 : |1 - \frac{1}{2}t^2\sigma_{n_k}^2| \le 1 \quad \forall k \in \{1, \ldots, r_n\}$
$\implies$ Für $n \ge n_0$ läßt sich das $\prod$ in (2) nach Lemma 5.2 durch die Summe in (1) abschätzen, d.h. (1) $\implies$ (2)
Es bleibt zu zeigen:

$$\lim_{n\to\infty} |\underbrace{\prod_{k=1}^{r_n} \exp(-\frac{1}{2}t^2\sigma_{n_k}^2)}_{=e^{-\frac{1}{2}t^2}} - \prod_{k=1}^{r_n}(1 - \frac{1}{2}t^2\sigma_{n_k}^2)| = 0 \quad \forall t \in \mathbb{R}$$

Behauptung fogt mit Lemma 5.2 falls

$$\lim_{n\to\infty} \sum_{k=1}^{r_n} \left| \exp(-\frac{1}{2}t^2\sigma_{n_k}^2) - 1 + \frac{1}{2}t^2\sigma_{n_k}^2 \right| = 0 \quad (4)$$

Für $x \in \mathbb{R}$ mit $|x| \le \frac{1}{2}$ gilt $|e^x - 1 - x| \le \frac{1}{2} \sum_{j=2}^{\infty} |x|^j \le x^2$

$$\implies \sum_{k=1}^{r_n} | \exp(\underbrace{-\frac{1}{2}t^2\sigma_{n_k}^2}_{=x}) - 1 + \frac{1}{2}t^2\sigma_{n_k}^2| \le \frac{1}{4}t^4 \sum_{k=1}^{r_n} \sigma_{n_k}^4$$

Wegen $\sum_{k=1}^{r_n} \sigma_{n_k}^4 \le \max\{\sigma_{n_k}^2 | 1 \le k \le r_n\} \cdot \underbrace{\sum_{k=1}^{r_n} \sigma_{n_k}^2}_{=1} \overset{n\to\infty,(3)}{\to} 0$

Also (3) $\implies$ (4) ∎

Beispiel 5.3 (Rekorde)
Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und $X_1, X_2, \dots$ eine Folge von unabhängigen identisch verteilten Zufallsvariablen darauf mit absolutstetiger Verteilungsfunktion F. Setze:

$$R_n := \begin{cases} 1, & \text{falls } X_n > X_i, i = 1, \dots, n-1 \\ 0, & \text{sonst} \end{cases}$$

$R_n = 1 \iff n$-ter Versuch ist ein Rekord. F stetig $\implies P(X_i = X_j) = 0 \ \forall i \neq j$

$$\begin{aligned} \implies \quad A \quad &:= \{\omega \in \Omega | \exists i \neq j, X_i(\omega) = X_j(\omega)\} \\ &= \bigcup_{i,j \in \mathbb{N}, i \neq j} \{X_i = X_j\} \\ \implies \quad P(A) \quad &= 0 \end{aligned}$$

Sei S_n die Menge der Permutationen der Zahlen $1, \dots, n$. Sei $\Psi_n : \Omega \to S_n$ gegeben durch

$$\Psi_n = \pi \iff X_{\pi(1)} < X_{\pi(2)} < \dots < X_{\pi(n)}$$

Ψ_n ist messbar, da $\Psi^{-1}(\{\pi\}) = \bigcap_{i=1}^{n} \{\underbrace{X_{\pi(i)} < X_{\pi(i+1)}}_{\in \mathcal{A}}\}$. Beispiel 3.5 $\implies (X_{\pi(1)}, \dots, X_{\pi(n)}) \stackrel{d}{=} (X_1, \dots, X_n) \ \forall \pi \in S_n$.
Ist $B := \{(x_1, \dots, x_n) \in \mathbb{R}^n | x_1 < \dots < x_n\}$ so gilt:

$$\begin{aligned} P(\Psi_n = \pi) &= P((X_{\pi(1)}, \dots, X_{\pi(n)}) \in B) \\ &= P((X_1, \dots, X_n) \in B) \\ &= P(\Psi_n = \mathbf{id}) \quad \text{unanhängig von } \pi \end{aligned}$$

$\implies \quad P(\Psi_n = \pi) = \frac{1}{n!} \ \forall \pi \in S_n$ und
$P(R_n = 1) = P(\Psi_n \in \{\pi \in S_n | \pi(n) = n\}) = \frac{1}{n}$
$\implies \quad R_n \sim B(1, \frac{1}{n})$ sind also nicht identisch verteilt
Wegen $\{R_{n+1} = 1\} \cap \{\Psi_n = \pi\} = \{\Psi_{n+1} = \tilde{\pi}\}$ mit

$$\tilde{\pi}(i) = \begin{cases} \pi(i) & , i \leq n \\ n+1 & , i = n+1 \end{cases}$$

folgt:

$$P(\Psi_n = \pi, R_{n+1} = 1) = \frac{1}{(n+1)!} = \underbrace{P(\Psi_n = \pi)}_{= \frac{1}{n!}} \underbrace{P(R_{n+1} = 1)}_{\frac{1}{n+1}} \quad \forall \pi \in S_n$$

$\implies \quad \Psi_n$ und R_{n+1} sind unabhängig
$\implies \quad$ Da $(R_1, \dots, R_n) = G(\Psi_n)$ sind R_{n+1} und $(R_1, \dots, R_n)$ unabhängig
$\implies \quad P(R_{i_1} = j_1, \dots, R_{i_n} = j_n) =$
$P(R_{i_1} = j_1, \dots, R_{i_{n-1}} = j_{n-1}) \cdot P(R_{i_n} = j_n) = \dots$
$P(R_{i_1} = j_1) \cdot \ldots \cdot P(R_{i_n} = j_n)$ für $i_1 < i_2 < \ldots < i_n, j_1, \dots, j_n \in \{0, 1\}$.
$\implies \quad$ die Zufallsvariablen $(R_n)_{n \in \mathbb{N}}$ sind unabhängig

Wie viele Rekorde gibt es unter den ersten n Versuchen?

$$S_n := \sum_{i=1}^{n} R_i$$

Es gilt:

$$ES_n = \sum_{k=1}^{n} ER_k = \sum_{k=1}^{n} \frac{1}{k}$$

$$\mathrm{Var}(S_n) = \sum_{k=1}^{n} \mathrm{Var}(R_k) = \sum_{k=1}^{n} \frac{1}{k}(1 - \frac{1}{k}) = \sum_{k=1}^{n} \frac{1}{k} - \sum_{k=1}^{n} \frac{1}{k^2}$$

Insbesondere:

$$\frac{ES_n}{\log n} \overset{n\to\infty}{\to} 1, \quad \frac{\mathrm{Var}(S_n)}{\log n} \overset{n\to\infty}{\to} 1$$

Mit dem zentralen Genzwertsatz (ZGWS) bekommen wir genauere Aussagen: Sei $X_{n_k} = R_k, r_n = n \implies s_n = (\mathrm{Var}(S_n))^{\frac{1}{2}}$
Überprüfen der Lyapunov-Bedingung ($\delta = 1$):

$$E|R_k - \underbrace{ER_k}_{=\frac{1}{k}}|^3 = \underbrace{P(R_k = 1)}_{=\frac{1}{k}}(1 - \frac{1}{k})^3 + \underbrace{P(R_k = 0)}_{=\frac{k-1}{k}} \left(\frac{1}{k}\right)^3 \leq \frac{2}{k}$$

$$\implies 0 \leq \frac{1}{s_n^3} \sum_{k=1}^{n} E|R_k - ER_k|^3 \leq \frac{1}{s_n^3} \sum_{k=1}^{n} \frac{2}{k} \to 0 \text{ für } n \to \infty$$

Der Zentrale Grenzwertsatz (Satz 5.10) liefert

$$\frac{S_n - ES_n}{\sqrt{Var(S_n)}} \overset{d}{\to} Z \sim N(0,1)$$

bzw.

$$\frac{S_n - \log(n)}{\sqrt{\log(n)}} \overset{d}{\to} Z \sim N(0,1)$$

Also für große n: $P(\log(n) - 1,96\sqrt{\log(n)} \leq S_n \leq \log(n) + 1,96\sqrt{\log(n)}) \approx P(-1,96 \leq Z \leq 1,96) = 0,9$.

Beispiel 5.4 (G. Polya, 1930: Eine Wahrscheinlichkeitsaufgabe zur Kundenwerbung, oder: "Coupon Collector's Problem")
Urne mit n verschiedenen Kugeln, Ziehen mit Zurücklegen
S_n = Anzahl der Züge, bis $r_n = [\phi \cdot n],\ 0 < \phi < 1$ verschiedene Kugeln gezogen werden. Es sei X_{nk} = Anzahl der bis zum Erhalt einer neuen Kugel nötigen Züge, wenn bereits $k-1$ verschiedene Kugeln gezogen (und zurückgelegt) wurden. $X_{n1} := 1$.
$X_{nk} \sim Geo(\frac{n-k+1}{n})$, d.h. $P(X_{nk} = j) = (\frac{k-1}{n})^{j-1} \cdot \frac{n-k+1}{n},\ j = 1, 2, \ldots$.
Falls $Y \sim Geo(p),\ p \in (0,1]$ mit $Y \equiv 1$ bei $p = 1$, gilt:
$EY = \frac{1}{p},\ \mathrm{Var}(Y) = \frac{1-p}{p^2},\ EY^4 \leq \frac{24}{p^4}$
Für $S_n = X_{n1} + \cdots + X_{nr_n}$ erhalten wir $\mu_n := ES_n = \sum_{k=1}^{r_n} \frac{n}{n-k+1}$ und $s_n^2 =$

$\mathrm{Var}(S_n) = \sum_{k=1}^{r_n} \frac{\frac{k-1}{n}}{(\frac{n-k+1}{n})^2} = n \sum_{k=1}^{r_n} \frac{k-1}{(n-k+1)^2}$.
Wir prüfen die Lyapunov-Bedingung mit $\delta = 2$: Für $Y \sim Geo(p)$ gilt:

$$E(Y - \frac{1}{p})^4 \leq E(\max\{Y, \frac{1}{p}\})^4 \leq EY^4 + \frac{1}{p^4} \leq \frac{25}{p^4}.$$

Insbesondere ist damit

$$\sum_{k=1}^{r_n} E|X_{nk} - \mu_{nk}|^4 \leq 25 \cdot \sum_{k=1}^{r_n} \frac{1}{(1 - \frac{(k-1)}{n})^4} \leq 25 \cdot [\phi \cdot n] \cdot \frac{1}{(1-\phi)^4} = O(n).$$

Wegen $s_n^2 \geq n \cdot \frac{1}{n^2} \sum_{k=1}^{r_n} (k-1) = \frac{1}{n} \cdot \frac{1}{2}(r_n - 1) \cdot r_n \geq \frac{1}{2n}(\phi n - 1)\phi n = \Theta(n)$ folgt damit

$$\lim_{n\to\infty} (\frac{1}{s_n^{2+2}} \sum_{k=1}^{r_n} E|X_{nk} - \mu_{nk}|^{2+2}) = 0.$$

Also folgt mit dem Zentralen Grenzwertsatz:

$$\frac{S_n - ES_n}{\sqrt{\mathrm{Var}(S_n)}} \xrightarrow{d} Z \sim N(0,1)$$

Weiter gilt: $\frac{1}{n} ES_n = \sum_{k=1}^{r_n} \frac{1}{n} \cdot \frac{1}{1 - \frac{k-1}{n}} = \int_0^{\frac{r_n}{n}} \frac{1}{1 - \frac{nx}{n}} \mathrm{d}x + O(\frac{1}{n}) = \int_0^{\frac{r_n}{n}} \frac{1}{1-x} \mathrm{d}x + O(\frac{1}{n}) = \int_0^{\phi} \frac{1}{1-x} \mathrm{d}x + O(\frac{1}{n}) = -\log(1-\phi) + O(\frac{1}{n})$.
Analog: $\lim_{n\to\infty} (\frac{1}{n} \mathrm{Var}(S_n^2)) = \int_0^{\phi} \frac{x}{(1-x)^2} \mathrm{d}x = \frac{\phi}{1-\phi} + \log(1-\phi)$.
Mit $a(\phi) = -\log(1-\phi),\ b(\phi) := \sqrt{\frac{\phi}{1-\phi} + \log(1-\phi)}$ folgt:

$$\frac{S_n - a(\phi)n}{b(\phi)\sqrt{n}} \xrightarrow{d} Z \sim N(0,1)$$

Beispiel (Numerisches Beispiel)
Wie groß muss ihr Bekanntenkreis sein, damit mit einer Wahrscheinlichkeit von mindestens 0,95 an 180 Tagen im Jahr Geburtstag gefeiert werden kann?
Also: $n = 365,\ \phi = \frac{180}{365},\ S_n \leq k \iff \underbrace{\frac{S_n - a(\phi)n}{b(\phi)\sqrt{n}}}_{\approx Z} \leq \frac{k - a(\phi)n}{(b(\phi)\sqrt{n}}$

$$\Phi(\underbrace{\frac{k - a(\phi)n}{b(\phi)\sqrt{n}}}_{=1{,}645}) \geq 0,95 \iff k \geq a(\phi)n + 1,645 \cdot b(\phi)\sqrt{n} \implies k \geq 266.$$

Für $\phi = 1$ kann man den Zentralen Grenzwertsatz nicht mehr anwenden:
$r_n = n,\ \mathrm{Var}(S_n) = n \sum_{k=1}^{n} \frac{k-1}{(n-k+1)^2} = n \sum_{k=1}^{n} \frac{n-k}{k^2} = n^2 \sum_{k=1}^{n} \frac{1}{k^2} - n \sum_{k=1}^{n} \frac{1}{k} = n^2 \cdot \frac{\pi^2}{6} + o(n^2)$.
$\mathrm{Var}(X_{n,n}) = \frac{1 - \frac{1}{n}}{\frac{1}{n^2}} = n^2 + o(n^2) \implies$ bei großem n steckt etwa $\frac{6}{\pi^2} \approx 0,61$ der Variabilität der Summe im letzten Summanden. Wir können jetzt eine andere Skalierung finden, allerdings ist die Grenzverteilung dann keine Normalverteilung mehr!
Sei $A_{m,i}$ das Ereignis, dass die Kugel i in den ersten m Ziehungen nicht auftaucht

$\implies \{S_n > m\} = \bigcup_{i=1}^n A_{m,i}$.
(S_n ist die Anzahl der Züge, bis alle n verschiedenen Kugeln mindestens einmal gezogen worden sind)
Mit der Siebformel:

$$P(S_n > m) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} P(\bigcap_{l=1}^{k} A_{m,i_l}) = \sum_{k=1}^{n-1} (-1)^{k+1} \binom{n}{k} (1 - \frac{k}{n})^m$$

Sei $c \in \mathbb{R}$ fest, $m_n = [n\log(n) + cn]$. Für $x > -1$ gilt $\log(1+x) \le x$. Damit: $\log\left(\binom{n}{k}\left(1-\frac{k}{n}\right)^{m_n}\right) \le k\log(n) - \log(k!) + \log\left(1-\frac{k}{n}\right)(n\log(n) + cn - 1) \overset{\text{s.o.}}{\le} k\log(n) - \log(k!) - \frac{k}{n}(n\log(n) + cn - 1) \le -ck + \frac{k}{n} - \log(k!)$

$$\implies \left|(-1)^{k+1}\binom{n}{k}\left(1-\frac{k}{n}\right)^{m_n}\right| \le \frac{1}{k!}\exp(\frac{k}{n} - ck) \quad \forall c \in \mathbb{R}.$$

$$\text{Ähnlich: } \lim_{n\to\infty} (-1)^{k+1}\binom{n}{k}\left(1-\frac{k}{n}\right)^{m_n} = (-1)^{k+1}\frac{1}{k!}e^{-ck} \quad \forall k \in \mathbb{N}.$$

Insgesamt:

$$\begin{aligned}
\lim_{n\to\infty}\left(P\left(\frac{S_n - n\log(n)}{n} > c\right)\right) &= \lim_{n\to\infty}\left(P\left(S_n > m_n\right)\right) \\
&= \lim_{n\to\infty}\left(\sum_{k=1}^{n-1}(-1)^{k+1}\binom{n}{k}\left(1-\frac{k}{n}\right)^{m_n}\right) \\
&= \lim_{n\to\infty}\left(\sum_{k=1}^{\infty}(-1)^{k+1}\binom{n}{k}\left(1-\frac{k}{n}\right)^{m_n}\right) \\
&\overset{\text{maj. Konv.}}{=} \sum_{k=1}^{\infty}\lim_{n\to\infty}\left((-1)^{k+1}\binom{n}{k}\left(1-\frac{k}{n}\right)^{m_n}\right) \\
&= \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k!}\left(e^{-c}\right)^k \\
&= 1 - e^{-e^{-c}}
\end{aligned}$$

Das Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathfrak{B})$ mit der Verteilungsfunktion $F(x) = e^{-e^{-x}}$ $\forall x \in \mathbb{R}$ heißt **Gumbel-Verteilung**.
Also gilt:

$$\frac{S_n - n\log(n)}{n} \xrightarrow{d} Z \sim Gumbel.$$

Beispiel (Variation des numerischen Beispiels von oben)
Der Bekanntenkreis soll jetzt so groß sein, dass mit einer Wahrscheinlichkeit von mindestens 0,95 täglich gefeiert werden kann.
$\implies k \ge 365 \cdot \log(365) \cdot 365 \cdot 2,97 \approx 3237,51$ (?)