Kapitel 2.9
Der zentrale Grenzwertsatz

Vom SGGZ zum ZGWS

Satz 2.42 (Schwaches Gesetz großer Zahlen)

Sei $(X_i)_{i\geq 1}$ eine Folge unabhängiger, identisch verteilter Zufallsvariablen mit $E[X_1^2] < \infty$, dann gilt für jedes $\varepsilon > 0$

$$\lim_{n\to\infty} P\left[\left|\frac{1}{n}\sum_{i=1}^{n} X_i - E[X_1]\right| \geq \varepsilon\right] = 0.$$

Beweis (mit Tschebschev-Ungleichung) benutzt

$$\mathrm{Var}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i) = \frac{1}{n}\mathrm{Var}(X_1) \xrightarrow[n\to\infty]{} 0.$$

Folgerung: Verteilung von $\frac{1}{n}\sum_{i=1}^{n}(X_i - E[X_1])$ konzentriert sich für $n \to \infty$ immer mehr um 0 (im Grenzfall Einpunktmasse bei 0)

Vom SGGZ zum ZGWS

Varianz der zentrierten Summe ist

$$\mathrm{Var}\left[\sum_{i=1}^{n}(X_i - E[X_1])\right] = \sum_{i=1}^{n}\mathrm{Var}(X_i) = n \cdot \mathrm{Var}(X_1) =: n\sigma^2$$

Standardisierte Summe:

$$S_n^* := \frac{\sum_{i=1}^{n}(X_i - E[X_1])}{\sqrt{n\sigma^2}}, \qquad E[S_n^*] = 0, \quad \mathrm{Var}(S_n^*) = 1 \ \forall\, n \geq 1$$

Beachte im Vergleich zum SGGZ verschiedene Skalierung: $\sqrt{n}$ statt n im Nenner!

Frage: Wie verhält sich die Verteilung von S_n^* für $n \to \infty$?
Kann man $P(a \leq S_n^* \leq b)$ für große n (näherungsweise) berechnen?

Vom SGGZ zum ZGWS

Satz (Zentraler Grenzwertsatz)

Seien $(X_i)_{i\geq 1}$ unabhängige, identisch verteilte Zufallsvariablen mit $E[X_1^2] < \infty$, $0 < \mathrm{Var}(X_1) = \sigma^2$, dann gilt für die standardisierten Summen S_n^* und alle $-\infty \leq a < b \leq \infty$

$$\lim_{n\to\infty} P(a \leq S_n^* \leq b) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx.$$

Korollar (Satz von deMoivre-Laplace)

Seien $(X_i)_{i\geq 1}$ unabhängige, identisch verteilte Bernoulli-Variablen mit $P(X_1 = 1) = p = 1 - P(X_1 = 0)$, $S_n^* = \frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}}$ und $-\infty \leq a < b \leq \infty$, dann gilt

$$\lim_{n\to\infty} P(a \leq S_n^* \leq b) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx.$$

Vom SGGZ zum ZGWS

Problem: Die sich im Grenzwert ergebende Verteilung kann mit bisherigen Konzepten nicht beschrieben werden!

Beschreibung von S_∞ bzw. S_∞^* nur auf Grundraum $\Omega = \{0, 1\}^\infty$ möglich, dieser ist *überabzählbar!* ($|\Omega| = |\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$.).

Vollständige Behandlung des Grenzwertproblems erfordert Methoden der allgemeinen Wahrscheinlichkeitstheorie.

Einschub: Ausblick auf die allgemeine Wahrscheinlichkeitstheorie

Übersicht

Bisher	Künftig
Diskreter Ws-Raum $(\Omega, \mathfrak{A}, P)$	Allgemeiner Ws-Raum $(\Omega, \mathfrak{A}, P)$
Ω nichtleer, höchstens abzählbar $\mathfrak{A} = \mathcal{P}(\Omega)$ (Potenzmenge) $P : \mathcal{P}(\Omega) \to [0,1]$ Ws-Maß (σ-additiv)	Ω nichtleer, beliebig (evtl. überabzählbar) $\mathfrak{A} \subseteq \mathcal{P}(\Omega)$ σ-Algebra $P : \mathfrak{A} \to [0,1]$ Ws-Maß (σ-additiv)

Für überabzählbare Grundräume Ω ist $\mathcal{P}(\Omega)$ i.a. zu groß, um darauf ein Ws-Maß mit gewünschten Eigenschaften definieren zu können, dies ist nur auf einer kleineren σ-Algebra möglich. ($\longrightarrow$ Maßtheorie, Analysis III)

Typisches Beispiel für überabzählbaren Gundraum

$(\Omega, \mathfrak{A}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ reelle Zahlen mit Borel-σ-Algebra

Definiere Ws-Maß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ durch eine *Ws-Dichte* f, d.h. durch eine nicht-negative, messbare Funktion $f : \mathbb{R} \to \mathbb{R}_+$ mit $\int_{-\infty}^{+\infty} f(x)\, \lambda(\mathrm{d}x) = 1$, wobei $\int f(x)\, \lambda(\mathrm{d}x)$ das Lebesgue-Integral bezeichne.

Übersicht (Forts.)

$(\boldsymbol{\Omega}, \mathfrak{A}, \mathbf{P})$ **diskret**	$(\boldsymbol{\Omega}, \mathfrak{A}, \mathbf{P}) = (\mathbb{R}, \boldsymbol{\mathcal{B}}(\mathbb{R}), \mathbf{P})$
P definiert durch $p_i = P(\{\omega_i\})\ \forall \omega_i \in \Omega$, $A \in \mathfrak{P}(\Omega)$: $P(A) = \sum_{\omega_i \in A} P(\{\omega_i\})$	P definiert durch Ws-dichte f Für $A \in \mathcal{B}(\mathbb{R})$ ist $P(A) := \int_A f(x)\, \lambda(\mathrm{d}x)$
Sei Y ZV mit Verteilung P_Y	
$E[Y^r] = \sum_{\omega_i \in \Omega} Y(\omega_i)^r P(\{\omega_i\})$ $= \sum_{i \geq 1} y_i^r P_Y(\{y_i\})$	$E[Y^r] = \int_\Omega Y(\omega)^r \,\mathrm{d}P(\omega)$ $= \int_{\mathbb{R}} x^r f_Y(x)\, \lambda(\mathrm{d}x)$ (dabei sei f_Y die Dichte von P_Y)

Bemerkung: Für stetige Dichten f und $A = [a, b]$ entspricht

$$P([a,b]) = \int_{[a,b]} f(x)\, \lambda(\mathrm{d}x) = \int_{-\infty}^{\infty} \mathbb{1}_{[a,b]}(x)\, f(x)\, dx = \int_a^b f(x)\, \mathrm{d}x$$

dem bekannten Riemann-Integral.

Die Verteilungsfunktion einer ZV Y mit (stetiger) Dichte f_Y ist $F_Y(x) = P(Y \leq x) = P_Y([-\infty, x]) = \int_{(-\infty,x]} f_Y(y)\, \lambda(\mathrm{d}y) = \int_{-\infty}^{x} f_Y(y)\, \mathrm{d}y$.

$\Rightarrow P(a \leq Y \leq b) = F_Y(b) - F_Y(a) = \int_{-\infty}^{\infty} \mathbb{1}_{[a,b]}(y) f_Y(y) dy = E[\mathbb{1}_{[a,b]}(Y)]$.

Exponentialverteilung $\mathrm{Exp}(\lambda)$

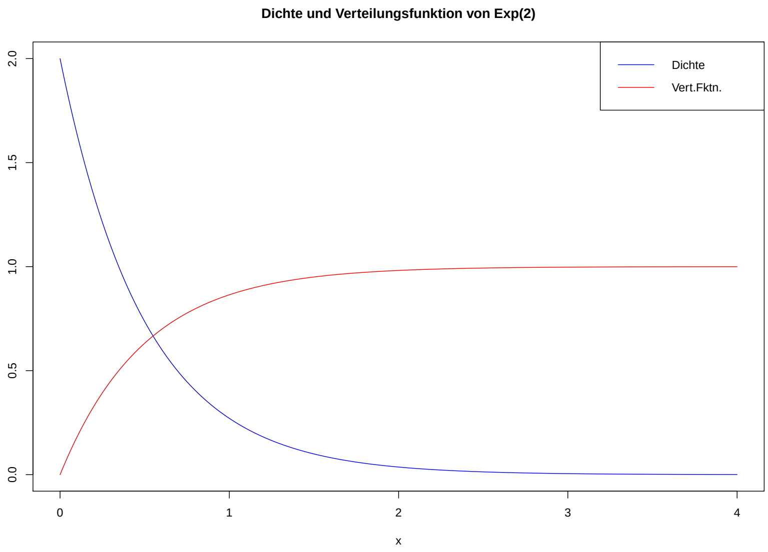
Die Exponentialverteilung $\mathrm{Exp}(\lambda)$ ist definiert durch ihre Dichte $f_{\mathrm{Exp}(\lambda)}(t) = \lambda e^{-\lambda t} \mathbb{1}_{[0,\infty)}(t)$ $(\lambda > 0)$ und hat die Verteilungsfunktion

$$F_{\mathrm{Exp}(\lambda)}(t) = \mathbb{1}_{[0,\infty)}(t) \int_0^t \lambda e^{-\lambda s} \, \mathrm{d}s = \left(1 - e^{-\lambda t}\right) \mathbb{1}_{[0,\infty)}(t)$$

Anwendung: Modell für Warte- oder Lebenszeiten

- Bei gleichartigen, unabhängig voneinander auftretenden Ereignissen: Wartezeit bis zum nächsten Ereignis (Anruf in einer Telefonzentrale, Zerfall eines Atoms in einer radioaktiven Materialprobe, Autounfall an verkehrsreicher Kreuzung, ...)
- Lebensdauer von Geräten, Glühbirnen, elektronischen Bauteilen (=Wartezeit bis zum (ersten) Ausfall)

In diesem Fall kann der Parameter λ als Ereignisrate (Anzahl Ereignisse pro Zeiteinheit) aufgefasst werden.


Dichte und Verteilungsfunktion von Exp(2)
Dichte
Vert.Fktn.
0.0
0.5
1.0
1.5
2.0
0
1
2
3
4
x

Exponentialverteilung $\mathrm{Exp}(\lambda)$

Sei T eine exponentialverteilte ZV mit $P_T = \mathrm{Exp}(\lambda)$, dann ist für $0 \leq s < t$ die Wahrscheinlichkeit $P(s \leq T \leq t) = e^{-\lambda s} - e^{-\lambda t}$ und

$$\begin{aligned} E[T] &= \int_0^\infty t \cdot \lambda\, e^{-\lambda t}\, \mathrm{d}t = -t \cdot e^{-\lambda t}\Big|_0^{+\infty} + \int_0^\infty e^{-\lambda t}\, \mathrm{d}t \\ &= -\frac{1}{\lambda} e^{-\lambda t}\Big|_0^{+\infty} = \frac{1}{\lambda} \end{aligned}$$

(je größer die Ereignisrate, desto kleiner die Wartezeit bis zum nächsten Ereignis) sowie

$$\begin{aligned} E[T^2] &= \int_0^\infty t^2 \cdot \lambda\, e^{-\lambda t}\, \mathrm{d}t = -t^2 \cdot e^{-\lambda t}\Big|_0^{+\infty} + \int_0^\infty 2t \cdot e^{-\lambda t}\, \mathrm{d}t \\ &= -\frac{2t}{\lambda} \cdot e^{-\lambda t}\Big|_0^{+\infty} + \frac{2}{\lambda}\int_0^\infty e^{-\lambda t}\, \mathrm{d}t = \frac{2}{\lambda} \cdot \frac{1}{\lambda}, \end{aligned}$$

also ist $\mathrm{Var}(T) = E[T^2] - E[T]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$.

Exponentialverteilung Exp(λ)

Die Exponentialverteilung kann auch als „stetige Version der geometrischen Verteilung“ aufgefasst werden:

Erinnerung: Geometrische Verteilung beschreibt Wartezeit bis zum ersten Erfolg bei unabhängigen Wiederholungen eines Experiments mit Erfolgswahrscheinlichkeit p: $P(\{n\}) = p \cdot (1-p)^{n-1}$.

Annahme: Experiment wird in kurzen Zeitintervallen Δt wiederholt mit Erfolgswahrscheinlichkeit $\lambda \Delta t$.

Sei T die Wartezeit bis zum ersten Erfolg. Wenn dieser zur Zeit t auftritt, hat man ungefähr $n \approx \frac{t}{\Delta t}$ Versuche benötigt, d.h.

$$P(T = t) \approx \lambda \Delta t (1 - \lambda \Delta t)^{\frac{t}{\Delta t} - 1} \quad \text{bzw.} \quad P(t < T \leq t + \Delta t) \approx \lambda \Delta t (1 - \lambda \Delta t)^{\frac{t}{\Delta t}}$$

Durch Grenzübergang $\Delta t \to dt$ (d.h. das Zeitintervall wird infinitesimal klein) erhält man (formal)

$$P(T \in (t, t + dt]) = \lambda e^{-\lambda t} dt \quad \text{und damit} \quad P(s \leq T \leq t) = \int_s^t \lambda e^{-\lambda r} \, \mathrm{d}r.$$

Normalverteilung $N(\mu, \sigma^2)$

Die (eindimensionale) Normalverteilung $N(\mu, \sigma^2)$ ist ein Ws-Maß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, definiert durch die Dichte $f_{N(\mu,\sigma^2)} : \mathbb{R} \to \mathbb{R}_+$ mit

$$f_{N(\mu,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad \mu \in \mathbb{R}, \sigma > 0.$$

Berechnung des Erwartungswertes und der Varianz

Für eine ZV X mit $P_X = N(\mu, \sigma^2)$ ist

$$\begin{aligned}
E[X] &= \int_{-\infty}^{+\infty} x \cdot f_{N(\mu,\sigma^2)}(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,\mathrm{d}x \\
&\underset{y=x-\mu}{=} \int_{-\infty}^{+\infty} (y+\mu) \cdot \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{y^2}{2\sigma^2}}\,\mathrm{d}y \\
&= \underbrace{\int_{-\infty}^{+\infty} y \cdot \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{y^2}{2\sigma^2}}\,\mathrm{d}y}_{=0 \text{ (Integrand ungerade: } g(-y)=-g(y))} + \mu \underbrace{\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{y^2}{2\sigma^2}}\,\mathrm{d}y}_{=1 \text{ (Dichteeigenschaft!)}} = \mu
\end{aligned}$$

Normalverteilung $N(\mu, \sigma^2)$

$$E[X^2] = \int_{-\infty}^{+\infty} x^2 \cdot \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, \mathrm{d}x \underset{y=x-\mu}{=} \int_{-\infty}^{+\infty} (y+\mu)^2 \cdot \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{y^2}{2\sigma^2}}\, \mathrm{d}y$$

$$= \int_{-\infty}^{+\infty} (y^2 + 2\mu y + \mu^2) \cdot \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{y^2}{2\sigma^2}}\, \mathrm{d}y$$

$$= \int_{-\infty}^{+\infty} y^2 \cdot \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{y^2}{2\sigma^2}}\, \mathrm{d}y + \mu^2$$

Partielle Integration liefert wegen $\left(e^{-\frac{y^2}{2\sigma^2}}\right)' = -\frac{y}{\sigma^2}\, e^{-\frac{y^2}{2\sigma^2}}$

$$\int_{-\infty}^{+\infty} y^2 \cdot \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{y^2}{2\sigma^2}}\, \mathrm{d}y = \int_{-\infty}^{+\infty} y \cdot \frac{y}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{y^2}{2\sigma^2}}\, \mathrm{d}y$$

$$= -y \cdot \frac{\sigma^2}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{y^2}{2\sigma^2}} \Bigg|_{-\infty}^{+\infty} + \sigma^2 \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{y^2}{2\sigma^2}}\, \mathrm{d}y$$

$$= 0 + \sigma^2$$

Normalverteilung $N(\mu, \sigma^2)$

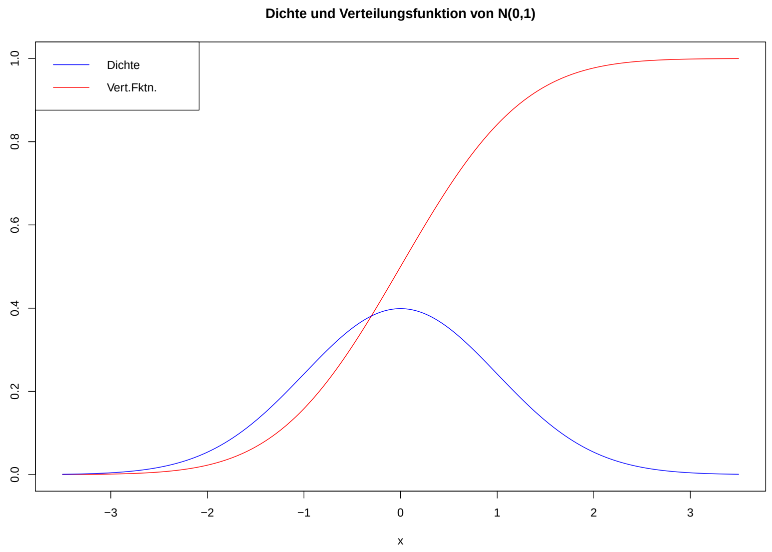
d.h. $E[X^2] = \sigma^2 + \mu^2$ und $\mathrm{Var}(X) = E[X^2] - E[X]^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$.

Mit $\mu = 0$, $\sigma = 1$ erhält man die **Standard-Normalverteilung $\mathbf{N(0,1)}$** mit Dichte $\varphi(x) := f_{N(0,1)}(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}$ und Verteilungsfunktion $\Phi(x) := \int_{-\infty}^{x} \varphi(x)\, \mathrm{d}x$. Diese liegt zumeist tabelliert vor.

Für eine ZV X mit $P_X = N(0,1)$ folgt aus der Symmetrie von $\varphi(x)$: $P(X \leq -x) = P(X \geq x) = 1 - P(X \leq x)$, d.h. $\Phi(-x) = 1 - \Phi(x)$ für alle $x \in \mathbb{R}$.

Falls Y eine ZV mit $P_Y = N(\mu, \sigma^2)$, hat $Y^* = \frac{Y-\mu}{\sigma}$ die Verteilung $P_{Y^*} = N(0,1)$ und somit

$$P(a \leq Y \leq b) = P\left(\frac{a-\mu}{\sigma} \leq \frac{Y-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$


Dichte und Verteilungsfunktion von N(0,1)
Dichte
Vert.Fktn.
0.0
0.2
0.4
0.6
0.8
1.0
−3
−2
−1
0
1
2
3
x

Beweis, dass $f_{N(\mu,\sigma^2)}(x)$ für alle $\mu \in \mathbb{R}$ und $\sigma^2 > 0$ eine Wahrscheinlichkeitsdichte ist:

Es genügt der Nachweis für $\varphi(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}$, denn sei $y = \sigma x + \mu$, so ist $x = \frac{y-\mu}{\sigma}$ und $dx = \sigma^{-1} dy$ und nach der Trafo-Regel für die Integration

$$1 \stackrel{!}{=} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\, dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(y-\mu)^2}{2\sigma^2}}\, dy.$$

Offensichtlich ist $\varphi(x)$ stetig, beschränkt auf $[-1, 1]$ und $\varphi(x) \leq \frac{1}{\sqrt{2\pi}}\, e^{-\frac{|x|}{2}}$ für $|x| \geq 1$, daher ist $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\, dx < \infty$.

Ferner $\int_{-\infty}^{\infty} \varphi(x)\, dx = 1 \quad \Longleftrightarrow \quad \left(\int_{-\infty}^{\infty} \varphi(x)\, dx\right) \cdot \left(\int_{-\infty}^{\infty} \varphi(x)\, dx\right) = 1.$

$$\begin{aligned}\left(\int_{-\infty}^{\infty} \varphi(x)\, dx\right) \cdot \left(\int_{-\infty}^{\infty} \varphi(x)\, dx\right) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi(x)\, \varphi(y)\, dx\, dy \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}}\, dx\, dy.\end{aligned}$$

Übergang zu Polarkoordinaten: $x = r \cdot \cos(\phi)$, $y = r \cdot \sin(\phi)$ mit $0 \leq r < \infty$, $0 \leq \phi < 2\pi$. Es gilt (Trafo-Satz) $dx\,dy = r\,dr\,d\phi$

$$\begin{aligned}\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}}\,dx\,dy &= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} e^{-\frac{r^2}{2}}\,r\,dr\,d\phi \\ &= \int_0^{\infty} e^{-\frac{r^2}{2}}\,r\,dr = -e^{-\frac{r^2}{2}}\Big|_0^{\infty} = 1.\end{aligned}$$

Satz (Unabhängigkeit und Produktdichten)

Seien $Y_1, \ldots, Y_n$ reellwertige Zufallsvariablen mit Dichten $f_{Y_1}, \ldots, f_{Y_n}$, dann sind äquivalent:

a) *$Y_1, \ldots, Y_n$ sind unabhängig.*

b) *Die Dichte der gemeinsamen Verteilung von $(Y_1, \ldots, Y_n)$ ist gegeben durch $f_{(Y_1,\ldots,Y_n)}(x_1, \ldots, x_n) = \prod_{i=1}^n f_{Y_i}(x_i)$.*

Der Beweis ist analog zu dem Fall diskreter Zufallsvariablen.

Definition 2.67

Seien $Z_1, \ldots, Z_n$ unabhängige, identisch standard-normalverteilte ZVn, dann heißt der Zufallsvektor $Z := (Z_1, \ldots, Z_n)$ **multivariat standard-normalverteilt** *und hat die Dichte*

$$f_Z(y) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{y_i^2}{2}} = \frac{1}{(2\pi)^{n/2}} e^{-\frac{\|y\|^2}{2}},$$

wobei $\|y\|^2 = y_1^2 + \cdots + y_n^2$ das Quadrat der Euklidischen Norm bezeichne.

Sei O eine orthogonale $n \times n$-Matrix, d.h. $O \cdot O^\top = I_n$, wobei I_n die $n \times n$-Einheitsmatrix bezeichne. Dann gilt $|\det(O)| = 1$ und

$$\|Oy\|^2 = \langle Oy, Oy \rangle = y^\top O^\top O y = y^\top y = \langle y, y \rangle = \|y\|^2 \qquad \forall\, y \in \mathbb{R}^n.$$

$\Longrightarrow$ Ist Z ein multivariat standard-normalverteilter Zufallsvektor und $X = OZ$ mit orthogonaler Matrix O, gilt

$$f_X(y) = \frac{1}{|\det(O)|} f_Z(Oy) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{\|Oy\|^2}{2}} = \frac{1}{(2\pi)^{n/2}} e^{-\frac{\|y\|^2}{2}} = f_Z(y),$$

d.h. auch $X = OZ$ ist multivariat standard-normalverteilt!

Da $X = OZ$ multivariat standard-normalverteilt, sind die Komponenten $X_i = O_i \cdot Z$ unabhängig (Produktdichte!), wobei O_i die i-te Zeile von O sei. Da O orthogonal, ist $\|O_i\|^2 = \sum_{j=1}^n O_{ij}^2 = 1$. Damit folgt

Korollar 2.68

Sind $X_1, \ldots, X_n$ unabhängige ZVn, so dass X_i nach $N(\mu_i, \sigma_i^2)$ verteilt ist für $i = 1, \ldots, n$, dann ist $Y = X_1 + \ldots + X_n$ nach $N\big(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\big)$ verteilt.

Beweis: Setze $\mu = \sum_{i=1}^n \mu_i, \sigma^2 = \sum_{i=1}^n \sigma_i^2$ und $\tau_i = \frac{\sigma_i}{\sigma}$. Dann sind

$$Z_i := \frac{X_i - \mu_i}{\sigma_i}$$

unabhängig und standard-normalverteilt. Außerdem gilt $\sum_{i=1}^n \tau_i^2 = 1$, d.h. die τ_i können als Zeile einer orthogonalen Matrix aufgefasst werden. Nach der Vorbemerkung ist dann auch $\tau_1 Z_1 + \cdots + \tau_n Z_n$ standard-normalverteilt, also ist

$$Y = X_1 + \cdots + X_n = \sigma(\tau_1 Z_1 + \cdots + \tau_n Z_n) + \mu$$

nach $N(\mu, \sigma^2)$ verteilt. □

Satz 2.69 (Zentraler Grenzwertsatz)

Seien $(X_i)_{i\geq 1}$ *unabhängige, identisch verteilte Zufallsvariablen mit* $E[X_1^2] < \infty$, $0 < \mathrm{Var}(X_1) = \sigma^2$, *dann gilt für die standardisierten Summen* $S_n^* = \frac{\sum_{i=1}^n (X_i - E[X_1])}{\sigma\sqrt{n}}$ *und alle* $-\infty \leq a < b \leq \infty$

$$\lim_{n\to\infty} P(a \leq S_n^* \leq b) = \int_a^b \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\, dx.$$

Beweisidee: O.B.d.A. gilt $E[X_1] = 0$ und $\mathrm{Var}(X_1) = \sigma^2 = 1$, da wir ansonsten zu den standardisierten Variablen $\frac{X_i - E[X_1]}{\sigma}$ übergehen können.

Falls die X_i standard-normalverteilt sind, ist nichts mehr zu beweisen, da nach Korollar 2.68 dann auch alle S_n^* standard-normalverteilt sind.

Seien nun $Z_1, Z_2, \ldots$ unabhängig und standard-normalverteilt und $\tilde{S}_n^* := \frac{\sum_{i=1}^n Z_i}{\sqrt{n}}$ die zugehörige, ebenfalls standard-normalverteilte standardisierte Summe. Zu zeigen ist dann, dass die sich durch sukzessive Ersetzung der Z_i in $\tilde{S}_n$ durch die X_i ergebende Abweichung von der Standard-Normalverteilung für $n \to \infty$ verschwindet.

Die technischen Details zum letzten Beweisschritt liefert das folgende

Lemma 2.70

Seien $(X_i)_{i\geq 1}$ unabhängige, identische verteilte und standardisierte ZVn, $h : \mathbb{R} \to \mathbb{R}$ eine dreimal stetig differenzierbare Funktion, deren erste drei Ableitungen beschränkt sind, und Z eine standard-normalverteilte Zufallsvariable. Dann gilt

$$E\left[h\left(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right)\right] \underset{n\to\infty}{\longrightarrow} E[h(Z)].$$

Beweis: Seien $Z_1, Z_2, ...$ unabhängige, identisch nach $N(0,1)$ verteilte ZVn sowie unabhängig von $X_1, X_2, ...$ Wir schreiben

$$U_i := \frac{X_1 + \cdots + X_{i-1} + Z_{i+1} + \cdots + Z_n}{\sqrt{n}}$$

und

$$h\left(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right) - h\left(\frac{Z_1 + \cdots + Z_n}{\sqrt{n}}\right) = \sum_{i=1}^{n} h\left(U_i + \frac{X_i}{\sqrt{n}}\right) - h\left(U_i + \frac{Z_i}{\sqrt{n}}\right)$$

Zweimalige Taylorentwicklung um U_i liefert

$$h\left(U_i + \frac{X_i}{\sqrt{n}}\right) - h\left(U_i + \frac{Z_i}{\sqrt{n}}\right) = h'(U_i)\frac{X_i - Z_i}{\sqrt{n}} + h''(U_i)\frac{X_i^2 - Z_i^2}{2n} + R_{in}$$

wobei der Restterm R_{in} sich nach der Taylor-Formel abschätzen lässt durch

$$\begin{aligned} |R_{in}| &\leq \left|h''(V_i) - h''(U_i)\right|\frac{X_i^2}{n} + \left|h''(W_i) - h''(U_i)\right|\frac{Z_i^2}{n} \\ &\leq c'''\frac{k^3}{n^{3/2}}\mathbb{1}_{\{|X_i|\leq k\}} + 2c''\frac{X_i^2}{n}\mathbb{1}_{\{|X_i|>k\}} + c'''\frac{|Z_i|^3}{n^{3/2}} \end{aligned}$$

mit ZVn V_i, W_i, so dass

$$|V_i - U_i| \leq \frac{|X_i|}{\sqrt{n}}, \qquad |W_i - U_i| \leq \frac{|Z_i|}{\sqrt{n}},$$

und Konstanten $c'' = \sup_{x\in\mathbb{R}} |h''(x)|, c''' := \sup_{x\in\mathbb{R}} |h'''(x)|$ und beliebigem $k > 0$.

Erinnerung: Taylor-Formel mit Restglied in Lagrange-Form

$$
\begin{aligned}
& h(x) \\
&= h(x_0) + (x-x_0)h'(x_0) + \frac{1}{2}(x-x_0)^2 h''(\xi_1) \quad (|\xi_1 - x_0| \le |x - x_0|) \\
&= h(x_0) + (x-x_0)h'(x_0) + \frac{1}{2}(x-x_0)^2 h''(x_0) + \frac{1}{2}(x-x_0)^2\big(h''(\xi_1) - h''(x_0)\big) \\
&= h(x_0) + (x-x_0)h'(x_0) + \frac{1}{2}(x-x_0)^2 h''(x_0) + \frac{1}{3!}(x-x_0)^3 h'''(\xi_2) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (|\xi_2 - x_0| \le |x - x_0|)
\end{aligned}
$$

Im hier betrachteten Fall ist $x_0 = U_i$, $x = U_i + \frac{X_i}{\sqrt{n}}$ und $\xi_1 = V_i$ bzw. $x = U_i + \frac{Z_i}{\sqrt{n}}$ und $\xi_1 = W_i$.
Ferner zerlegen wir $X_i = X_i \cdot \mathbb{1}_{\{|X_i| \le k\}} + X_i \cdot \mathbb{1}_{\{|X_i| > k\}}$ und benutzen für $X_i \cdot \mathbb{1}_{\{|X_i| \le k\}}$ die letzte Gleichung oben und für $X_i \cdot \mathbb{1}_{\{|X_i| > k\}}$ die vorletzte und verwenden die Beschränktheit der Ableitungen zur Abschätzung
$|h''(x_0)|, |h''(\xi_1)| \le \sup_{x \in \mathbb{R}} |h''(x)| = c''$,
$|h'''(\xi_2)| \le \sup_{x \in \mathbb{R}} |h'''(x)| = c'''$.

Da U_i, X_i, Z_i unabhängig sind und X_i, Z_i standardisiert (d.h. es gilt $E[X_i] = E[Z_i] = 0,\ E[X_i^2] = E[Z_i^2] = 1$), folgt

$$\begin{aligned}
&\left| E\left[h\left(U_i + \frac{X_i}{\sqrt{n}}\right) - h\left(U_i + \frac{Z_i}{\sqrt{n}}\right)\right] \right| \\
&= \left| E\left[h'(U_i)\frac{X_i - Z_i}{\sqrt{n}}\right] + E\left[h''(U_i)\frac{X_i^2 - Z_i^2}{2n}\right] + E[R_{in}] \right| \\
&= \left| E[h'(U_i)]\, E\left[\frac{X_i - Z_i}{\sqrt{n}}\right] + E[h''(U_i)]\, E\left[\frac{X_i^2 - Z_i^2}{2n}\right] + E[R_{in}] \right| \\
&= \left| E[R_{in}] \right| \leq E[|R_{in}|] \leq c''' \frac{k^3 + E[|Z_1|^3]}{n^{3/2}} + 2c'' \frac{E[X_1^2 \mathbb{1}_{\{|X_1| > k\}}]}{n}
\end{aligned}$$

Wegen

$$\begin{aligned}
&\left| E\left[h\left(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right) - h\left(\frac{Z_1 + \cdots + Z_n}{\sqrt{n}}\right)\right] \right| \\
&\qquad\qquad \leq \sum_{i=1}^{n} \left| E\left[h\left(U_i + \frac{X_i}{\sqrt{n}}\right) - h\left(U_i + \frac{Z_i}{\sqrt{n}}\right)\right] \right|
\end{aligned}$$

und da Z und $\frac{Z_1+\cdots+Z_n}{\sqrt{n}}$ dieselbe Verteilung haben (Korollar 2.68), folgt

$$\left| E\left[h\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\right)\right] - E\big[h(Z)\big]\right| \leq c''' \frac{k^3+E[|Z_1|^3]}{n^{1/2}} + 2c'' E\left[X_1^2 \mathbb{1}_{\{|X_1|>k\}}\right]$$

und daraus

$$\limsup_{n\to\infty} \left| E\left[h\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\right)\right] - E\big[h(Z)\big]\right| \leq 2c'' E\left[X_1^2 \mathbb{1}_{\{|X_1|>k\}}\right]$$

für beliebiges $k > 0$. Da $\mathrm{Var}(X_1) < \infty$ n.V., ist $E[X_1^2] < \infty$ und daher $\lim_{k\to\infty} E\left[X_1^2 \mathbb{1}_{\{|X_1|>k\}}\right] = 0$

(Diskreter Fall: $E\left[X_1^2 \mathbb{1}_{\{|X_1|>k\}}\right] = \sum_{|x_i|>k} x_i^2 \cdot P(X_1 = x_i) \underset{k\to\infty}{\longrightarrow} 0$).

Es folgt

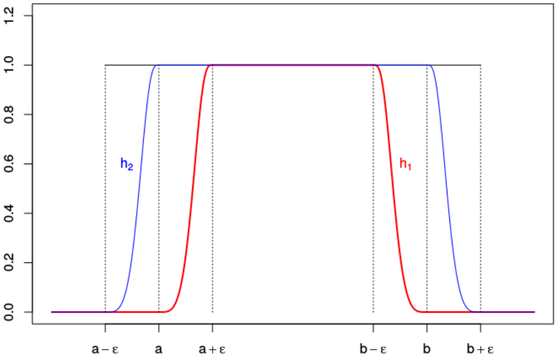
$$\lim_{n\to\infty} \left| E\left[h\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\right)\right] - E\big[h(Z)\big]\right| = 0$$

und damit die Behauptung. $\square$

Beweis des zentralen Grenzwertsatzes (Satz 2.69)

Seien $-\infty < a < b < \infty$ und $\varepsilon > 0$. Wähle h_1, h_2 dreimal stetig differenzierbar mit beschränkten ersten drei Ableitungen, so dass

$$\mathbb{1}_{[a+\varepsilon,b-\varepsilon]}(x) \leq h_1(x) \leq \mathbb{1}_{[a,b]}(x) \leq h_2(x) \leq \mathbb{1}_{[a-\varepsilon,b+\varepsilon]}(x).$$



Beweis des zentralen Grenzwertsatzes (Satz 2.69)

Sei $S_n^* = \frac{\sum_{i=1}^n X_i}{\sqrt{n}}$ und Z standard-normalverteilt, dann ist wegen der Monotonie des Erwartungswertes und Lemma 2.70

$$P(a+\varepsilon \le Z \le b-\varepsilon) = E[\mathbb{1}_{[a+\varepsilon,b-\varepsilon]}(Z)] \le E\left[h_1(Z)\right] = \liminf_{n\to\infty} E\left[h_1(S_n^*)\right]$$
$$\le \liminf_{n\to\infty} E[\mathbb{1}_{[a,b]}(S_n^*)] = \liminf_{n\to\infty} P(a \le S_n^* \le b) \le \limsup_{n\to\infty} P(a \le S_n^* \le b)$$
$$\le \limsup_{n\to\infty} E\left[h_2(S_n^*)\right] = E\left[h_2(Z)\right] \le E[\mathbb{1}_{[a-\varepsilon,b+\varepsilon]}(Z)] = P(a-\varepsilon \le Z \le b+\varepsilon).$$

Wegen

$$P(c \le Z \le d) = \int_c^d \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\, dx$$

folgt $\left|P(a \pm \varepsilon \le Z \le b \pm \varepsilon) - P(a \le Z \le b)\right| \xrightarrow{\varepsilon\to 0} 0$ und damit die Aussage des zentralen Grenzwertsatzes

$$\lim_{n\to\infty} P(a \le S_n^* \le b) = \int_a^b \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\, dx$$

für alle $-\infty < a < b < \infty$. Die Fälle $a = -\infty$ und $b = \infty$ können analog gezeigt werden. $\square$

Stirling-Formel

Satz 2.71

Für $n \to \infty$ gilt

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n (1 + o(1)) \quad \textit{oder äquivalent} \quad \lim_{n\to\infty} \frac{n!}{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}} = 1.$$

Beweis: Sei $t^+ = \max(t, 0)$ und $\varepsilon > 0$, dann gibt es dreimal stetig differenzierbare Funktionen $h_1(t), h_2(t)$ mit
$t^+ - \varepsilon \leq h_1(t) \leq t^+ \leq h_2(t) \leq t^+ + \varepsilon$
(z.B. kann man $h_2(t) = \sqrt{\varepsilon^2 + \frac{t^2}{4}} + \frac{t}{2}$ und $h_1(t) = h_2(t) - \varepsilon$ wählen).
Mit einem ähnlichen Approximationsargument wie im Beweis des ZGWS folgt mit Lemma 2.70

$$E\left[\left(\frac{X_1 + \cdots + X_n - n}{\sqrt{n}}\right)^+\right] \xrightarrow{n\to\infty} E[Z^+]$$

für unabhängige, identisch verteilte ZVn $X_1, X_2, \ldots$ mit $E[X_1] = \mathrm{Var}[X_1] = 1$ und standard-normalverteiltem Z.

Stirling-Formel

Es gilt

$$E[Z^+] = \frac{1}{\sqrt{2\pi}} \int_0^\infty z\, e^{-\frac{z^2}{2}}\, dz = -\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \Big|_0^\infty = \frac{1}{\sqrt{2\pi}}$$

Wähle X_1, X_2 als unabhängige, identisch poissonverteilte ZVn mit Parameter $\lambda = 1$, dann gilt $E[X_1] = \mathrm{Var}[X_1] = 1$, und $X_1 + \cdots + X_n$ ist poissonverteilt zum Parameter $\lambda = n$. Dann ist

$$\begin{aligned} E\left[\left(\frac{X_1 + \cdots + X_n - n}{\sqrt{n}}\right)^+\right] &= \frac{1}{\sqrt{n}} e^{-n} \sum_{k=n+1}^{\infty} (k-n) \frac{n^k}{k!} \\ &= \frac{e^{-n}}{\sqrt{n}} \sum_{k=n+1}^{\infty} \left(\frac{n^k}{(k-1)!} - \frac{n^{k+1}}{k!}\right) \\ &= \frac{e^{-n}}{\sqrt{n}} \frac{n^{n+1}}{n!} = \frac{\left(\frac{n}{e}\right)^n \sqrt{n}}{n!}. \end{aligned}$$

Also gilt $\lim_{n\to\infty} \frac{\left(\frac{n}{e}\right)^n \sqrt{n}}{n!} = \frac{1}{\sqrt{2\pi}}$ und damit die Behauptung. □