Kapitel V
Verteilungskonvergenz und der zentrale Grenzwertsatz

Neben der Konvergenz in Wahrscheinlichkeit, die wir in Abschnitt 9 im Zusammenhang mit dem schwachen Gesetz der großen Zahlen kennengelernt hatten, und der dort nur am Rande erwähnten fast sicheren Konvergenz spielen noch andere Konvergenzarten in der Stochastik eine Rolle. Hierzu gehört die sogenannte *Verteilungskonvergenz*, die im folgenden eingeführt und später zur Formulierung eines herausragend wichtigen Resultats der Stochastik, dem *zentralen Grenzwertsatz*, benötigt wird. Obgleich wir nunmehr zu diskreten Zufallsexperimenten zurückkehren, tauchen regulär stetige Verteilungen und speziell Normalverteilungen weiterhin auf, weil die Limiten diskreter Verteilungen nicht notwendig diskret zu sein brauchen.

16. Verteilungskonvergenz von Zufallsgrößen

Für diskrete W-Verteilungen $Q, Q_1, Q_2, \ldots$ über $\mathbb{R}$ ist es naheliegend, von einer Konvergenz der Q_n gegen Q zu sprechen, wenn

$$\lim_{n\to\infty} Q_n(\{x\}) \;=\; Q(\{x\}) \tag{16.1}$$

für alle $x \in \mathbb{R}$, wenn also die Zähldichten der Q_n gegen die Zähldichte von Q konvergieren. Man kann zeigen, daß dies für die Verteilungsfunktionen F_n, F von Q_n, Q

$$\lim_{n\to\infty} F_n(x) \;=\; F(x) \tag{16.2}$$

für alle $x \in \mathbb{R}$ impliziert. Hier ist ein wichtiges Beispiel für eine derartige Konvergenzaussage:

16.1. Poissonscher Grenzwertsatz. *Sei $(\theta_n)_{n\ge1}$ eine Folge in $(0,1)$ mit der Eigenschaft*

$$\lim_{n\to\infty} n\theta_n \;=\; \mu \tag{16.3}$$

für ein $\mu \in (0,\infty)$. Dann gilt

$$\lim_{n\to\infty} B(n,\theta_n)(\{k\}) \;=\; Poi(\mu)(\{k\}) \tag{16.4}$$

für alle $k \ge 0$.

Beweis: Da $\lim_{n\to\infty}\left(1-\frac{x_n}{n}\right)^n = e^{-x}$ für jedes $x \in \mathbb{R}$ und jede gegen x konvergente Folge $(x_n)_{n\ge1}$, ergibt sich die Behauptung leicht aus der Darstellung

$$\begin{aligned} B(n,\theta_n)(\{k\}) &= \binom{n}{k}\theta_n^k(1-\theta_n)^{n-k} \\ &= \frac{1}{k!}(n\theta_n)^k\left(1-\frac{n\theta_n}{n}\right)^{n-k}\prod_{i=1}^{k-1}\left(1-\frac{i}{n}\right) \end{aligned}$$

für alle $k \ge 0$ und $n \ge k$. $\diamondsuit$

Der Poissonsche Grenzwertsatz besagt also, daß für großes n und kleines $\theta \approx \frac{\mu}{n}$

$$B(n,\theta)(\{k\}) \;\approx\; Poi(\mu)(\{k\}) \;=\; e^{-\mu}\frac{\mu^k}{k!}$$

für $k = 0,...,n$. Dies ist schon deswegen nützlich, weil sich Poisson-Wahrscheinlichkeiten bei großem n viel einfacher als Binomialwahrscheinlichkeiten berechnen lassen.

Daß die oben vorgestellte Konvergenzart aber in vielen Situationen zu restriktiv ist, sollen die nächsten beiden Beispiele belegen:

16.2. Beispiel. Betrachte die Folge $X_n = \frac{1}{n}$, $n \ge 1$, konstanter Zufallsgrößen mit zugehörigen (Dirac)-Verteilungen $Q_n = \delta_{1/n}$. Da $X_n \to X = 0$ für $n \to \infty$, wird man auch $\delta_{1/n} \to \delta_0$ erwarten. Im Sinne der obigen Konvergenzart ist dies allerdings nicht der Fall, denn für $x = 0$ erhalten wir

$$\delta_{1/n}(\{0\}) \;=\; 0 \;\not\to\; 1 \;=\; \delta_0(\{0\}) \quad (n \to \infty)$$

und entsprechend für die Verteilungsfunktionen $F_n = \mathbf{1}_{[1/n,\infty)}$, $F = \mathbf{1}_{[0,\infty)}$ von δ_n, δ

$$F_n(0) \;=\; 0 \;\not\to\; 1 \;=\; F(0) \quad (n \to \infty).$$

16.3. Beispiel. Für jedes $n \ge 1$ sei X_n eine Laplace-verteilte Zufallsgröße auf der Menge $\mathcal{X}_n \stackrel{\text{def}}{=} \{\frac{k}{n} : 0 \le k \le n\}$. Da die Punkte von $\mathcal{X}_n$ mit wachsendem n eine immer feinere "Approximation" des Einheitsintervalls bilden und jeder dieser Punkte unter P^{X_n} die gleiche Masse besitzt, kann man intuitiv erwarten, daß P^{X_n} gegen $R(0,1)$, die Gleichverteilung auf $[0,1]$ konvergiert. Da $R(0,1)$ jedoch eine regulär stetige Verteilung darstellt, kann diese Konvergenz nicht im obigen Sinne (16.1) gelten. Andererseits sieht man leicht, daß (16.2) für alle $x \in \mathbb{R}$ erfüllt ist, denn: Bezeichnet F die Verteilungsfunktion der $R(0,1)$-Verteilung und F_n die Verteilungsfunktion von X_n (bzw. P^{X_n}), so folgt offenbar

$$F_n(x) \;=\; \frac{1}{n+1}\sum_{k=0}^{n}\mathbf{1}_{[k/n,\infty)}(x) \;\to\; F(x) \;=\; (x\wedge 1)\mathbf{1}_{[0,\infty)}(x), \quad n \to \infty.$$

Die Schwierigkeit im ersten Beispiel besteht offenkundig darin, daß der Punkt 0 unter δ_0 die Masse 1 besitzt und somit eine Unstetigkeits(Sprung-)stelle von F ist, während er unter den Q_n immer nur Wahrscheinlichkeit 0 besitzt. Im zweiten Beispiel zeigt sich, daß der naheliegende Limes diskreter W-Verteilungen P^{X_n} nicht notwendig wieder diskret sein muß.

In der anschließenden Definition von *Verteilungskonvergenz* fordert man deshalb nur

$$\lim_{n\to\infty} F_n(x) \;=\; F(x)$$

für alle $x \in \mathbb{R}$, in denen F stetig ist, und läßt außerdem als Grenzwert Funktionen F zu, die nicht mehr notwendig Verteilungsfunktionen diskreter W-Verteilungen Q sind, sondern nur noch die Eigenschaften

(1) Isotonie ($F(x) \leq F(y)$ für $x \leq y$),

(2) rechtsseitige Stetigkeit,

(3) $\lim_{x\to\infty} F(x) = 1$ und $\lim_{x\to-\infty} F(x) = 0$

haben, die auch Verteilungsfunktionen regulär stetiger Verteilungen besitzen. In der W-Theorie wird in der Tat gezeigt, daß eine Funktion mit den Eigenschaften (1)–(3) *immer* die Verteilungsfunktion genau einer W-Verteilung auf $\mathbb{R}$ bildet. Wir bezeichnen deshalb hiernach jedes derartige F als Verteilungsfunktion und mit $\mathcal{C}(F)$ die Menge ihrer Stetigkeitspunkte.

16.4. Definition. Es seien $Q_1, Q_2, \ldots$ diskrete W-Verteilungen über $\mathbb{R}$ mit Verteilungsfunktionen $F_1, F_2, \ldots$ Die Q_n heißen *schwach konvergent*, wenn es eine Verteilungsfunktion F gibt, so daß

$$\lim_{n\to\infty} F_n(x) \;=\; F(x) \tag{16.5}$$

für alle $x \in \mathcal{C}(F)$. Die zu F gehörende W-Verteilung Q heißt Grenzverteilung, und wir schreiben $Q_n \xrightarrow{w} Q$. Sind $X, X_1, \ldots$ Zufallsgrößen mit Verteilungen $Q, Q_1, \ldots$, so heißen die X_n *konvergent in Verteilung gegen* X, kurz $X_n \xrightarrow{d} X$.

16.5. Bemerkungen. (a) Es gilt also Verteilungskonvergenz von X_n gegen eine Zufallsgröße X genau dann, wenn die Verteilung von X_n schwach gegen die von X konvergiert:

$$X_n \xrightarrow{d} X \quad \Leftrightarrow \quad P^{X_n} \xrightarrow{w} P^X. \tag{16.6}$$

(b) Der Leser beachte, daß aus $X_n \xrightarrow{d} X$ auch $X_n \xrightarrow{d} Y$ für jede andere Zufallsgröße Y mit derselben Verteilung wie X folgt. Der Limes einer verteilungskonvergenten Folge von Zufallsgrößen ist demnach nicht wie gewöhnlich eindeutig bestimmt, sondern nur dessen Verteilung. Aus diesem Grund schreibt man statt $X_n \xrightarrow{d} X$ bisweilen auch $X_n \xrightarrow{d} Q$, wenn Q die Verteilung von X angibt ($Q = P^X$).

(c) In 16.2 gilt offenbar

$$\lim_{n\to\infty} F_n(x) \; = \; \lim_{n\to\infty} \mathbf{1}_{[1/n,\infty)}(x) \; = \; \mathbf{1}_{[0,\infty)}(x) \; = \; F(x)$$

für alle $x \neq 0$ und folglich $\delta_{1/n} \xrightarrow{w} \delta_0$.

Verteilungskonvergenz ist schwächer als Konvergenz in Wahrscheinlichkeit.

16.6. Satz. *$X, X_1, X_2, \ldots$ seien Zufallsgrößen mit $X_n \xrightarrow{P} X$. Dann konvergieren die X_n auch in Verteilung gegen X, d.h. $X_n \xrightarrow{d} X$ bzw. $P^{X_n} \xrightarrow{w} P^X$.*

Beweis: $F^X, F^{X_1}, F^{X_2}, \ldots$ mögen die Verteilungsfunktionen zu $X, X_1, X_2, \ldots$ bezeichnen. Zu jeder Stetigkeitsstelle x von F^X $(x \in \mathcal{C}(F^X))$ und jedem $\varepsilon > 0$ existiert dann ein $\eta > 0$, so daß

$$F^X(x) - \varepsilon \; \leq \; F^X(x-\eta) \; \leq \; F^X(x+\eta) \; \leq \; F^X(x) + \varepsilon.$$

Nun gilt aber

$$\begin{aligned} \{X \leq x - \eta\} \; &= \; \{X \leq x-\eta, X_n \leq x\} \cup \{X \leq x - \eta, X_n > x\} \\ &\subset \; \{X_n \leq x\} \cup \{|X_n - X| > \eta\} \end{aligned}$$

für alle $n \geq 1$ und führt zu der Abschätzung

$$F^X(x) \; \leq \; F^X(x-\eta) + \varepsilon \; \leq \; F^{X_n}(x) + P(|X_n - X| > \eta) + \varepsilon.$$

Entsprechend gilt für alle $n \geq 1$

$$\{X_n \leq x\} \; \subset \; \{X \leq x + \eta\} \cup \{|X_n - X| > \eta\}$$

und somit

$$F^{X_n}(x) \; \leq \; F^X(x+\eta) + P(|X_n - X| > \eta) \; \leq \; F^X(x) + P(|X_n - X| > \eta) + \varepsilon.$$

Insgesamt gilt demnach

$$|F^{X_n}(x) - F^X(x)| \; \leq \; \varepsilon + P(|X_n - X| > \eta),$$

Da der letzte Term für $n \to \infty$ gegen 0 strebt, konvergiert $F^{X_n}(x)$ gegen $F^X(x)$. $\diamondsuit$

Die Umkehrung des Satzes gilt nicht, wie das folgende einfache Beispiel bestätigt:

16.7. Beispiel. $X, X_1, X_2, \ldots$ seien stochastisch unabhängige und jeweils $B(1, \frac{1}{2})$-verteilte Zufallsgrößen. Dann gilt trivialerweise $P^{X_n} \xrightarrow{w} P^X$, wegen

$$P(|X_n - X| \geq 1) \; = \; P(X_n = 1, X = 0) + P(X_n = 0, X = 1) \; = \; \frac{1}{2}$$

für alle $n \geq 1$ aber nicht $X_n \overset{P}{\to} X$.

Es gibt allerdings eine wichtige Ausnahme:

16.8. Satz. *$X_1, X_2, ...$ seien Zufallsgrößen mit $X_n \overset{d}{\to} X \equiv a$, also $P^{X_n} \overset{w}{\to} \delta_a$ für ein $a \in \mathbb{R}$. Dann gilt $X_n \overset{P}{\to} a$.*

Beweis: $X \equiv a$ hat die Verteilungsfunktion $F(x) = \mathbf{1}_{[a,\infty)}(x)$, die in allen $x \neq a$ stetig ist. Daher sichert die Voraussetzung

$$\lim_{n\to\infty} P(X_n \leq a - \varepsilon) \;=\; F(a-\varepsilon) \;=\; 0$$

sowie

$$\lim_{n\to\infty} P(X_n > a + \varepsilon) \;=\; 1 - F(a+\varepsilon) \;=\; 0$$

für alle $\varepsilon > 0$. Wegen $\{|X_n - a| > \varepsilon\} \subset \{X_n \leq a - \varepsilon\} + \{X_n > a + \varepsilon\}$ ergibt dies

$$\lim_{n\to\infty} P(|X_n - a| > \varepsilon) \;=\; 0$$

für alle $\varepsilon > 0$ und somit $X_n \overset{P}{\to} a$. ♢

17. Der zentrale Grenzwertsatz

Nach diesem Satz läßt sich das schwache Gesetz der großen Zahlen für paarweise unabhängige, identisch verteilte Zufallsgrößen $X_1, X_2, ...$ mit Erwartungswert μ und Varianz σ^2 (☞ Korollar 12.10) auch in der Form

$$P^{\frac{1}{n}\sum_{i=1}^n (X_i-\mu)} \;\overset{w}{\to}\; \delta_0$$

schreiben. Es stellt sich nun die Frage, ob man bei Division von $\sum_{i=1}^n (X_i - \mu)$ durch weniger als n eine Verteilungskonvergenz gegen eine echte Zufallsgröße X erhält. Um einen Hinweis darauf zu bekommen, wie dies möglich sein könnte, betrachten wir die Varianz von $\overline{X}_n - \mu$, wobei an $\overline{X}_n = \frac{1}{n}(X_1 + ... + X_n)$ erinnert sei. Aufgrund der paarweisen Unabhängigkeit und damit paarweisen Unkorreliertheit gilt

$$\mathrm{Var}(\overline{X}_n - \mu) \;=\; n^{-2}\sum_{i=1}^n \mathrm{Var}X_i \;=\; \frac{\sigma^2}{n}.$$

Dagegen bleibt die Varianz konstant, wenn man $\sum_{i=1}^n (X_i - \mu)$ durch $\sigma n^{1/2}$ dividiert:

$$\mathrm{Var}\left(\frac{1}{\sigma n^{1/2}}\sum_{i=1}^n (X_i-\mu)\right) \;=\; \frac{1}{n\sigma^2}\sum_{i=1}^n \mathrm{Var}X_i \;=\; 1.$$

Wir betrachten daraufhin die *standardisierten Summen*

$$\frac{1}{\sigma n^{1/2}} \sum_{i=1}^{n} (X_i - \mu) \;=\; \frac{n^{1/2}}{\sigma} (\overline{X}_n - \mu)$$

und werden zeigen, daß diese tatsächlich verteilungskonvergent sind, und zwar gegen eine Standard-Normalverteilung. Dabei beschränken wir uns auf diskrete Zufallsgrößen bzw. W-Verteilungen, obwohl das Ergebnis auch ohne diese Beschränkung gilt. Als Hilfsmittel benötigen wir ein hinreichendes Kriterium für die schwache Konvergenz diskreter W-Verteilungen gegen eine regulär stetige Grenzverteilung.

17.1. Satz. *Es seien Q_n diskrete W-Verteilungen über $\mathbb{R}$ und Q eine regulär stetige Verteilung mit Dichte f. Gilt für jede stetige Funktion $g : \mathbb{R} \to \mathbb{R}$, für die $\lim_{x\to\infty} g(x)$ und $\lim_{x\to-\infty} g(x)$ existieren,*

$$\lim_{n\to\infty} \sum_{x\in\mathbb{R}} g(x) Q_n(\{x\}) \;=\; \int_{-\infty}^{\infty} g(x) f(x)\, dx, \tag{17.1}$$

so folgt $Q_n \xrightarrow{w} Q$.

Bezeichnen $Y, Y_1, Y_2, \ldots$ Zufallsgrößen mit Verteilungen $Q, Q_1, Q_2, \ldots$, so läßt sich (17.1) unter Hinweis auf Satz 15.7 auch in der Form

$$\lim_{n\to\infty} Eg(Y_n) \;=\; Eg(Y) \tag{17.2}$$

schreiben, und "$Q_n \xrightarrow{w} Q$" durch die gleichbedeutende Aussage "$Y_n \xrightarrow{d} Y$" ersetzen. Nach Definition gilt letzteres gerade, wenn (17.1) bzw. (17.2) für die Funktionen $g = \mathbf{1}_{(-\infty,x]}$, $x \in \mathbb{R}$, erfüllt ist, die zwar für $x \to \pm\infty$ einen Limes besitzen, aber *nicht* stetig sind. Der Beweis des obigen Satzes basiert folglich darauf zu zeigen, daß jedes $g = \mathbf{1}_{(-\infty,x]}$ beliebig genau durch ein stetiges g wie im Satz gefordert approximiert werden kann. Wir weisen noch darauf hin, daß jedes solche g beschränkt ist.

Beweis: Es seien $Y, Y_1, Y_2, \ldots$ wie gerade angegeben. Wir definieren weiter

$$g(x) \stackrel{\text{def}}{=} \mathbf{1}_{(-\infty,0]}(x) + (1-x)\mathbf{1}_{(0,1)}(x) \quad \text{und} \quad g_k(x) \stackrel{\text{def}}{=} g(kx)$$

für $k \in \mathbb{N}$, $x \in \mathbb{R}$ und stellen fest, daß g_k stetig ist mit $\lim_{x\to-\infty} g_k(x) = 1$, $\lim_{x\to\infty} g_k(x) = 0$ sowie

$$1_{(-\infty,0]}(x) \;\le\; g_k(x) \;\le\; \mathbf{1}_{(-\infty,1/k]}(x), \quad x \in \mathbb{R},$$

für alle $k \ge 1$. Für die zu den Q_n gehörenden Verteilungsfunktionen F_n gilt dann für alle $x \in \mathbb{R}$

und $k \geq 1$ zum einen

$$\begin{aligned}
\limsup_{n\to\infty} F_n(x) &= \limsup_{n\to\infty} E\mathbf{1}_{(-\infty,x]}(Y_n) \\
&= \limsup_{n\to\infty} E\mathbf{1}_{(-\infty,0]}(Y_n - x) \\
&\leq \limsup_{n\to\infty} Eg_k(Y_n - x) \\
&= Eg_k(Y - x) && (*) \\
&\leq E\mathbf{1}_{(-\infty,1/k]}(Y - x) \\
&= F(x + 1/k),
\end{aligned}$$

wobei F die Verteilungsfunktion von Q angibt, und zum anderen

$$\begin{aligned}
\liminf_{n\to\infty} F_n(x) &= \liminf_{n\to\infty} E\mathbf{1}_{(-\infty,x]}(Y_n) \\
&= \liminf_{n\to\infty} E\mathbf{1}_{(-\infty,1/k]}(Y_n - x + 1/k) \\
&\geq \liminf_{n\to\infty} Eg_k(Y_n - x + 1/k) \\
&= Eg_k(Y - x + 1/k) && (*) \\
&\geq E\mathbf{1}_{(-\infty,0]}(Y - x + 1/k) \\
&= E\mathbf{1}_{(-\infty,x-1/k]}(Y) \\
&= F(x - 1/k),
\end{aligned}$$

also insgesamt

$$F(x - 1/k) \;\leq\; \liminf_{n\to\infty} F_n(x) \;\leq\; \limsup_{n\to\infty} F_n(x) \;\leq\; F(x + 1/k)$$

für alle $k \geq 1$. Dabei beachte man, daß mit g_k natürlich auch $g_k(\cdot - x)$ für jedes $x \in \mathbb{R}$ die geforderten Eigenschaften des Satzes besitzt und somit (17.1) in den mit $(*)$ gekennzeichneten Zeilen verwendet werden konnte. Da F stetig ist, folgt außerdem

$$\lim_{k\to\infty} F(x \pm 1/k) \;=\; F(x),$$

insgesamt also $\lim_{n\to\infty} F_n(x) = F(x)$ für alle $x \in \mathbb{R}$. $\diamondsuit$

17.2. Zentraler Grenzwertsatz (von Lindeberg-Levy). *Gegeben seien stochastisch unabhängige und identisch verteilte diskrete Zufallsgrößen $X_1, X_2, \ldots$, für die der Erwartungswert μ und die Varianz σ^2 existieren. Dann konvergieren die zugehörigen standardisierten Summen für $n \to \infty$ in Verteilung gegen eine Standard-Normalverteilung, d.h.*

$$\frac{1}{\sigma n^{1/2}} \sum_{i=1}^{n} (X_i - \mu) \;\xrightarrow{d}\; N(0,1) \tag{17.3}$$

beziehungsweise

$$\lim_{n\to\infty} P\left(\frac{1}{n^{1/2}}\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)\le t\right) \;=\; \Phi(t) \;\stackrel{\text{def}}{=}\; \int_{-\infty}^{t}\frac{1}{(2\pi)^{1/2}}e^{-x^2/2}\,dx \tag{17.4}$$

für alle $t\in\mathbb{R}$ *(beachte* $\mathcal{C}(\Phi)=\mathbb{R}$*).*

Beweis: Zur Vereinfachung der Bezeichnungen setzen wir $Y_n \stackrel{\text{def}}{=} (X_n-\mu)/\sigma$ für $n\ge 1$ und notieren, daß diese Zufallsgrößen wiederum stochastisch unabhängig (nach Satz 12.13) und identisch verteilt sind. Ferner gilt $EY_n=0$ und $\mathit{Var}Y_n=1$, d.h. die Y_n sind standardisiert, sowie

$$n^{1/2}\overline{Y}_n \;=\; \frac{1}{n^{1/2}}\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)$$

für alle $n\ge 1$.

Sei $\varphi(x) = (2\pi)^{-1/2}e^{-x^2/2} = \Phi'(x)$ die Dichte der Standard-Normalverteilung. Wir zeigen im folgenden, daß für jede stetige Funktion g, für welche die Grenzwerte $\lim_{x\to\pm\infty} g(x)$ existieren,

$$\lim_{n\to\infty}\sum_{y\in\mathbb{R}} g(y)P(n^{1/2}\overline{Y}_n=y) \;=\; \int_{-\infty}^{\infty} g(x)\varphi(x)\,dx \;\stackrel{\text{def}}{=}\; I(g) \tag{17.5}$$

gilt, wobei die linke Seite auch in der Form $\lim_{n\to\infty} Eg(n^{1/2}\overline{Y}_n)$ geschrieben werden kann. Nach Satz 17.1 folgt dann die Behauptung des Satzes.

Sei also ein beliebiges g der geforderten Art gewählt und o.E. $I(g)=0$; andernfalls gehe man zu $g-I(g)$ über. Wir betrachten nun

$$h(x) \;\stackrel{\text{def}}{=}\; \varphi(x)^{-1}\int_{-\infty}^{x} g(y)\varphi(y)\,dy.$$

Da g aufgrund der Voraussetzungen sogar gleichmäßig stetig und beschränkt ist, ist h wohldefiniert und als Quotient differenzierbarer Funktionen (bei strikt positivem Nenner) ebenfalls differenzierbar. Aus

$$\varphi'(x) \;=\; (2\pi)^{-1/2}(-x)e^{-x^2/2} \;=\; -x\varphi(x)$$

folgt für die Ableitung h'

$$h'(x) \;=\; \frac{g(x)\varphi(x)^2-\varphi'(x)\int_{-\infty}^{x} g(y)\varphi(y)\,dy}{\varphi(x)^2} \;=\; g(x)+xh(x) \tag{17.6}$$

für alle $x\in\mathbb{R}$. Nach der Regel von L'Hospital gilt ferner

$$\begin{aligned}\lim_{x\to\pm\infty} xh(x) \;&=\; \lim_{x\to\pm\infty}\frac{\int_{-\infty}^{x} g(y)\varphi(y)\,dy}{\varphi(x)/x}\\ &=\; \lim_{x\to\pm\infty}\frac{g(x)\varphi(x)}{-\varphi(x)(1+x^{-2})} \;=\; \lim_{x\to\pm\infty} g(x).\end{aligned}$$

Auch h' ist somit gleichmäßig stetig. Nunmehr erhalten wir wegen (17.6)

$$\begin{aligned} Eg(n^{1/2}\overline{Y}_n) &= Eh'(n^{1/2}\overline{Y}_n) - E\Big(n^{1/2}\overline{Y}_n h(n^{1/2}\overline{Y}_n)\Big) \\ &= Eh'(n^{1/2}\overline{Y}_n) - n^{1/2}\frac{1}{n}\sum_{i=1}^{n} E(Y_i h(n^{1/2}\overline{Y}_n)). \end{aligned}$$

Da die Y_i stochastisch unabhängig und identisch verteilt sind, ist dieser Ausdruck gleich

$$Eh'(n^{1/2}\overline{Y}_n) - n^{1/2}E(Y_1 h(n^{1/2}\overline{Y}_n)).$$

Eine Taylor-Entwicklung von h um $Z_n \overset{\text{def}}{=} n^{-1/2}\sum_{i=2}^{n} Y_i$ liefert

$$h(n^{1/2}\overline{Y}_n) = h(Z_n) + h'(Z_n)\frac{Y_1}{n^{1/2}} + \frac{Y_1}{n^{1/2}}R_n$$

mit $R_n = h'(Z_n + \theta Y_1/n^{1/2}) - h'(Z_n)$, $\theta = \theta(Y_1, ..., Y_n) \in [0,1)$. Da aber Y_1 und Z_n (als Funktion von $Y_2, ..., Y_n$) unabhängig sind, folgt daraus

$$\begin{aligned} E(Y_1 h(n^{1/2}\overline{Y}_n)) &= EY_1\, Eh(Z_n) + EY_1^2\, Eh'(Z_n)/n^{1/2} + E(Y_1^2 R_n)/n^{1/2} \\ &= n^{-1/2}\Big(Eh'(Z_n) + E(Y_1^2 R_n)\Big). \end{aligned}$$

Insgesamt erhalten wir also

$$\begin{aligned} Eg(n^{1/2}\overline{Y}_n) &= Eh'(n^{1/2}\overline{Y}_n) - n^{1/2}E(Y_1 h(n^{1/2}\overline{Y}_n)) \\ &= E(h'(Z_n + Y_1/n^{1/2}) - h'(Z_n)) - E(Y_1^2\{h'(Z_n) + \theta Y_1/n^{1/2}) - h'(Z_n)\}). \end{aligned}$$

Da h' gleichmäßig stetig ist, konvergieren bei festem $\omega \in \Omega$ wegen $\lim_{n\to\infty} n^{-1/2}Y_1(\omega) = 0$ die Ausdrücke unter beiden Erwartungswerten punktweise gegen 0 und sind ferner beschränkt durch eine Konstante bzw. ein Vielfaches von Y_1^2, weil h' beschränkt ist. Nach dem im Anschluß formulierten Satz von der majorisierten Konvergenz, den wir nur angeben können, konvergieren dann auch die Erwartungswerte gegen 0, und wir erhalten die gewünschte Konvergenzaussage (17.5). ♢

17.3. Satz von der majorisierten Konvergenz. *Sind $Y, X, X_1, X_2, ...$ integrierbare Zufallsgrößen mit den Eigenschaften*

$$\sup_{n\geq 0}|X_n| \leq Y \quad \textit{und} \quad \lim_{n\to\infty} X_n = X,$$

so gilt auch

$$\lim_{n\to\infty} E|X_n - X| = 0 \quad \textit{und} \quad \lim_{n\to\infty} EX_n = EX.$$

17.4. Beispiel. Seien $X_1, X_2, ...$ stochastisch unabhängige $B(1,\theta)$-verteilte Zufallsgrößen, wobei $\theta \in (0,1)$. Dann besitzt $S_n = X_1 + ... + X_n$ eine Binomialverteilung mit Parametern

n und θ (☞ Beispiel 12.11), kurz $B(n,\theta)$-Verteilung. Es folgt für $n \geq 1$ und $t \in [0,n]$

$$P(S_n \leq t) \;=\; \sum_{k=0}^{t} \binom{n}{k} \theta^k (1-\theta)^{n-k}.$$

Für große Werte von n und t ist diese Summe sehr unhandlich zu berechnen. Mittels des zentralen Grenzwertsatzes erhalten wir unter Benutzung von $\mu = EX_1 = \theta$ und $\sigma^2 = \mathrm{Var}X_1 = \theta(1-\theta)$ und mit $t_n \stackrel{\text{def}}{=} \frac{t-n\theta}{(\theta(1-\theta)n)^{1/2}}$

$$P(S_n \leq t) \;=\; P\left(\frac{S_n - n\theta}{(\theta(1-\theta)n)^{1/2}} \leq t_n\right) \;\approx\; \Phi(t_n) \;=\; \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{t_n} e^{-x^2/2}\, dx.$$

Man nennt $\Phi(t_n)$ auch die *Normalapproximation* von $P(S_n \leq t)$. Allerdings hängt ihre Güte sowohl von n (klar) als auch von θ ab. Ist $|\theta - \frac{1}{2}|$ nahe bei 1/2, so ist die Konvergenz im zentralen Grenzwertsatz langsamer als für θ nahe bei 1/2.

Ein wichtiger Spezialfall der gerade beschriebenen Situation ergibt sich bei Betrachtung *relativer Häufigkeiten* (☞ Beispiel 9.6). Bezeichnen $Y_1, Y_2, \ldots$ stochastisch unabhängige, identisch verteilte Zufallsgrößen, z.B. die Ergebnisse beim wiederholten Würfeln, und ist A ein Ereignis von Interesse, so definieren

$$X_n \stackrel{\text{def}}{=} \mathbf{1}_A(Y_n), \quad n \geq 1$$

stochastisch unabhängige, identisch $B(1,\theta_A)$-verteilte Zufallsgrößen, wobei $\theta_A \stackrel{\text{def}}{=} P(Y_1 \in A)$.