Kapitel 9

Funktionale Grenzwertsätze

Das fundamentale Resultat dieses Kapitels ist der *Funktionale Zentrale Grenzwertsatz*, welcher zuerst von Donsker 1951 bewiesen wurde. Er sagt, dass die Verteilung einer reskalierten Irrfahrt mit zentrierten iid-Zuwächsen von Varianz 1 in der Skalierung (7.1) schwach gegen das Wienermaß konvergiert. Gelegentlich bezeichnet man den Satz auch als *Invarianzprinzip von Donsker*, weil das behauptete asymptotische Verhalten der Irrfahrt weitgehend invariant ist gegenüber der Verteilung der Zuwächse. Die Aussage ist eine viel stärkere als die des klassischen Zentralen Grenzwertsatzes, der nur die Verteilung der Irrfahrt in einem festen Zeitpunkt zum Gegenstand hat. Im Zusammenhang mit der *Kolmogorov-Smirnov-Teststatistik* beweisen wir mithilfe des Invarianzprinzips die schwache Konvergenz des *uniformen empirischen Prozesses* gegen eine Brownsche Brücke. Der letzte Abschnitt widmet sich einem funktionalen Grenzwertsatz für den *empirischen Spektralprozess*.

Stochastische Gleichstetigkeit auf $[0, 1]$

Um Satz 8.15 zu benutzen, brauchen wir ein exaktes Verständnis des Begriffes Straffheit. Für eine Funktion $x \in \mathcal{C}[0,1]$ ist der *Stetigkeitsmodul* definiert durch

$$w_x(\delta) := \sup_{|t-s|\le\delta} \big|\, x(t) - x(s) \,\big| .$$

Jedes $x \in \mathcal{C}[0,1]$ ist gleichmäßig stetig und erfüllt $\lim_{\delta\to 0} w_x(\delta) = 0$. Wegen $|w_x(\delta) - w_y(\delta)| \le 2\|x-y\|_{\sup}$ ist $w_{(.)}(\delta)$ für festes δ stetig auf $\mathcal{C}[0,1]$ und damit messbar.

Satz 9.1. *Eine Folge $(\mathbb{P}_n)$ von W-Maßen über $\mathcal{B}(\mathcal{C}[0,1])$ ist straff genau dann, wenn gilt:*

$$\forall\, \eta > 0 \;\; \exists\, a, n_0, \quad \textit{so dass} \;\; \mathbb{P}_n\big(x : |x(0)| \ge a\big) \le \eta \;\; \forall\, n \ge n_0 \tag{9.1}$$

und

$$\forall\, \varepsilon, \eta > 0 \;\; \exists\, \delta, n_0 \;\; \textit{mit} \;\; 0 < \delta < 1 \;\; \textit{so dass} \;\; \mathbb{P}_n\big(x : w_x(\delta) \ge \varepsilon\big) \le \eta \;\; \forall\, n \ge n_0. \tag{9.2}$$

Damit ist eine Folge von Zufallsvariablen (X_n) mit Werten in $\mathcal{C}[0,1]$ straff, falls die Folge $(X(0)_n)$ straff ist und

$$\forall\, \varepsilon, \eta > 0 \;\; \exists\, \delta, n_0 \;\; \textit{mit} \;\; \mathbb{P}\Big(\sup_{|t-s|\le\delta} \big|\, X_n(s) - X_n(t) \,\big| \ge \varepsilon\Big) \le \eta \;\; \forall\, n \ge n_0. \tag{9.3}$$

Falls die letzte Bedingung erfüllt ist, heißt (X_n) auch gleichgradig stetig in Wahrscheinlichkeit. Zur Erinnerung sei der für den Beweis benötigte Satz von Arzelà und Ascoli zitiert. Er stammt aus der klassischen reellen Analysis und identifiziert die relativ kompakten Mengen im Raum $\mathcal{C}(K)$, den stetigen reellen Funktionen auf einem kompakten metrischen Raum K.

Satz 9.2. (Satz von Arzelà-Ascoli) *Sei (K, d) ein kompakter, metrischer Raum und $\mathcal{C}(K)$ der Raum der stetigen reellen Funktionen auf K, versehen mit der Topologie der gleichmäßigen Konvergenz. Eine Teilmenge $B \subset \mathcal{C}(K)$ ist relativ kompakt genau dann, wenn gilt:*

(i) Es existiert $x \in K$ mit $|f(x)| \leq c < \infty$ für alle $f \in B$ und ein $c > 0$;

(ii) B ist gleichmäßig gleichgradig stetig, d.h. $\lim_{\delta \searrow 0} \sup \big\{ w_\delta(f); f \in B \big\} = 0$.

Wegen (ii) kann man (i) durch die Aussage ersetzen: (i') B ist beschränkt, d.h. $\|f\|_{\sup} \leq c < \infty$ für alle $f \in B$ und ein $c > 0$.

BEWEIS (Satz 9.1). Angenommen, $(\mathbb{P}_n)$ sei straff. Zu gegebenem $\eta > 0$ wähle man eine kompakte Menge K, so dass $\mathbb{P}_n(K) > 1 - \eta$ für alle $n \in \mathbb{N}$. Nach dem Satz von Arzelà-Ascoli hat man $K \subset \{x : |x(0)| \leq a\}$ für hinreichend großes a und $K \subset \{x : w_x(\delta) \leq \varepsilon\}$ für hinreichend kleines δ, und somit folgen (9.1) und (9.2) mit $n_0 = 1$. Die Bedingungen sind also notwendig. Nach Satz 6.18 ist ein einziges W-Maß $\mathbb{P}$ auf $\mathcal{B}(\mathcal{C}[0,1])$ straff. Wegen der Notwendigkeit der Bedingungen (9.1) und (9.2) existiert zu gegebenem η ein a mit $\mathbb{P}(\{x : |x(0)| \geq a\}) \leq \eta$, und zu festem ε und η gibt es ein δ, so dass $\mathbb{P}(\{x : w_x(\delta) > \varepsilon\}) \leq \eta$. Falls $(\mathbb{P}_n)$ die beiden Bedingungen des Satzes erfüllt, können wir durch Vergrößern von a und δ (falls nötig) sicherstellen, dass $n_0 = 1$ gewählt werden kann. Erfülle $(\mathbb{P}_n)$ also (9.1) und (9.2) mit $n_0 = 1$. Zu gegebenem η wähle man a, so dass $\mathbb{P}_n(B) \geq 1 - \eta$ für alle $n \in \mathbb{N}$ mit $B := \{x : |x(0)| \leq a\}$. Dann wähle man δ_k so, dass für $B_k := \{x : w_x(\delta_k) < 1/k\}$ gilt $\mathbb{P}_n(B_k) \geq \eta/2^k$ für alle $n \in \mathbb{N}$. Falls K der Abschluss von $A := B \cap \big\{ \cap_{k=1}^{\infty} B_k \big\}$ ist, folgt $\mathbb{P}_n(K^c) \leq \mathbb{P}_n(A^c) \leq \eta + \sum_k \eta/2^k \leq 2\eta$ für alle $n \in \mathbb{N}$. Da A die beiden Bedingungen des Satzes von Arzelà-Ascoli erfüllt, ist K kompakt und $(\mathbb{P}_n)$ damit straff. $\square$

Die nächsten beiden Lemmata ergeben Abschätzungen, die den Nachweis der stochastischen Gleichstetigkeit in vielen Fällen erleichtern.

Lemma 9.3. *Sei X eine $\mathcal{D}[0,1]$-Zufallsvariable. Für beliebige $0 < \delta \leq 1$ und $\eta > 0$ ist*

$$\mathbb{P}\Big(w_X(\delta/2) > 2\eta \Big) \ \leq \ \frac{2}{\delta} \sup_{t \in [0,1)} \mathbb{P}\Big(\sup_{u \in [t, t+\delta] \cap [0,1]} \big| X(u) - X(t) \big| > \eta \Big).$$

BEWEIS. Für $0 \leq j \leq j(\delta)$ sei $I_j := [j\delta/2, (j+1)\delta/2] \cap [0,1]$, wobei $j(\delta)$ die größte ganze Zahl in $[0, 2/\delta)$ ist. Zu $u, v \in [0,1]$ mit $|u - v| < \delta$ existiert stets ein $j < j(\delta)$ mit

$u, v \in I_j \cup I_{j+1}$. Folglich ist

$$\begin{aligned}
\mathbb{P}\Big(w_X(\delta/2) > 2\eta\Big) &\le \sum_{j=0}^{j(\delta)-1} \mathbb{P}\Big(\sup_{u,v\in I_j\cup I_{j+1}} \big|\,X(u) - X(v)\,\big| > 2\eta\Big) \\
&\le \sum_{j=0}^{j(\delta)-1} \mathbb{P}\Big(\sup_{u\in I_j\cup I_{j+1}} \big|\,X(u) - X(j\delta/2)\,\big| > \eta\Big) \\
&\le \frac{2}{\delta} \sup_{t\in[0,1)} \mathbb{P}\Big(\sup_{u\in[t,t+\delta]\cap[0,1]} \big|\,X(u) - X(v)\,\big| > \eta\Big). \qquad \square
\end{aligned}$$

Die Frage ist nun, wie man die rechte Seite der Ungleichung in Lemma 9.3 abschätzen kann.

Lemma 9.4. (Lévy) *Sei X ein stochastischer Prozess auf $[a,b]$ mit rechtsseitig stetigen Pfaden und unabhängigen Zuwächsen; d.h. für $a \le t < u \le b$ sind $(X(s))_{s\in[a,t]}$ und $X(u) - X(t)$ stochastisch unabhängig. Für beliebige $\eta > 0$ ist*

$$\mathbb{P}\Big(\sup_{t\in[a,b]} |X(t)| > \eta\Big) \;\le\; \mathbb{P}\Big(|X(b)| > \eta/2\Big) \;+\; \sup_{t\in[a,b]} \mathbb{P}\Big(|X(b) - X(t)| > \eta/2\Big).$$

Falls zusätzlich die Verteilung von $(X(b) - X(t))$ symmetrisch um Null ist für alle $t \in [a,b]$, dann ist sogar

$$\mathbb{P}\Big(\sup_{t\in[a,b]} |X(t)| > \eta\Big) \;\le\; 2\mathbb{P}\Big(|X(b)| > \eta\Big).$$

BEWEIS. Wegen der rechtsseitigen Stetigkeit der Pfade von X genügt es, die Behauptung mit $\sup_{t\in T} |X(t)|$ anstelle von $\sup_{a\le t\le b} |X(t)|$ zu beweisen, wobei T eine beliebige endliche Teilmenge von $[a,b]$ ist, die $\{a\}$ und $\{b\}$ enthält. Die Behauptung folgt dann mithilfe des Satzes von der monotonen Konvergenz. Sei

$$\tau := \min\Big(\Big\{t \in T : |X(t)| > \eta\Big\} \cup \{\infty\}\Big).$$

Dann ist

$$\big\{\sup_{t\in T} |X(t)| > \eta\big\} = \big\{\tau \in T\big\} \subset \Big\{\tau \in T, |X(b)| \le \eta/2\Big\} \cup \Big\{|X(b)| > \eta/2\Big\},$$

und $\Big\{\tau \in T, |X(b)| \le \eta/2\Big\}$ ist gleich

$$\bigcup_{t\in T} \Big\{\tau = t, |X(b)| \le \eta/2\Big\} \subset \bigcup_{t\in T} \{\tau = t\} \cap \Big\{|X(b) - X(t)| > \eta/2\Big\},$$

denn $|X(t)| > \eta$ auf $\{\tau = t\}$. Folglich ist

$$\begin{aligned}
\mathbb{P}\Big(\sup_{t\in T} |X(t)| > \eta\Big) &\le \mathbb{P}\big(|X(b)| > \eta/2\big) + \mathbb{P}\big(\tau \in T, |X(b)| \le \eta/2\big) \\
&\le \mathbb{P}\big(|X(b)| > \eta/2\big) + \sum_{t\in T} \mathbb{P}(\tau = t)\mathbb{P}\big(|X(b) - X(t)| > \eta/2\big) \\
&\le \mathbb{P}\big(|X(b)| > \eta/2\big) + \sup_{t\in[a,b]} \mathbb{P}\big(|X(b) - X(t)| > \eta/2\big).
\end{aligned}$$

Im Falle symmetrisch verteilter Zuwächse $X(b) - X(t)$ ist $\mathbb{P}\big(X(b) - X(t) \geq 0\big) \geq 1/2$ und $\mathbb{P}\big(X(b) - X(t) \leq 0\big) \geq 1/2$, weshalb

$$\begin{aligned}
&\mathbb{P}(\tau = t) \\
&\quad = \; \mathbb{P}\big(\tau = t, X(t) > \eta\big) + \mathbb{P}\big(\tau = t, X(t) < -\eta\big) \\
&\quad \leq \; 2\mathbb{P}\Big(\tau = t, X(t) > \eta, X(b) - X(t) \geq 0\Big) + 2\mathbb{P}\Big(\tau = t, X(t) < -\eta, X(b) - X(t) \leq 0\Big) \\
&\quad \leq \; 2\mathbb{P}\big(\tau = t, X(b) > \eta\big) + 2\mathbb{P}\big(\tau = t, X(b) < -\eta\big) \\
&\quad = \; 2\mathbb{P}\big(\tau = t, |X(b)| > \eta\big),
\end{aligned}$$

und die gewünschte Ungleichung folgt durch Aufsummieren über alle $t \in T$. □

Lemma 9.4 nutzt entscheidend die Unabhängigkeit der Zuwächse aus. Ohne Beweis (da wir es nicht benutzen) zitieren wir an dieser Stelle ein einfaches Kriterium für die stochastische Gleichstetigkeit, was diese Voraussetzung nicht benötigt:

Satz 9.5. *Sei $(X_n)_n$ eine Folge von Zufallsvariablen mit Werten in $(\mathcal{C}[0,1]), \sup)$. Falls es $\gamma, \kappa > 0$, $K \in \mathbb{R}$ und $n_0 \in \mathbb{N}$ gibt, so dass für alle $0 \leq s, t \leq 1$ und alle $n \geq n_0$ gilt*

$$\mathbb{P}\big(|X_n(s) - X_n(t)| > \varepsilon\big) \;\leq\; K \cdot \frac{|s-t|^{1+\gamma}}{\varepsilon^\kappa}, \tag{9.4}$$

so ist die Folge $(X_n)_n$ gleichgradig stetig in Wahrscheinlichkeit, d.h. es gilt (9.3).

Bedingung (9.4) rechnet man zum Beispiel mit der Markov-Ungleichung nach.

Das Invarianzprinzip von Donsker

Wir kommen nun zum Hauptsatz dieses Kapitels. Seien $Y_{1,n}, ..., Y_{n,n}$ reellwertige Zufallsvariable. Dazu betrachten wir den Partialsummenprozess

$$t \mapsto S_n(t) \;:=\; \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} Y_{i,n}, \quad t \in [0,1].$$

Sei weiter $X_n(.)$ die stetig linear interpolierte Version von $S_n(.)$, d.h.

$$X_n(t) \;:=\; S_n(t) + (nt - [nt])n^{-1/2} Y_{[nt]+1,n}.$$

Satz 9.6. (Donsker) *Angenommen, für jedes n sind $Y_{1,n}, Y_{2,n}, ..., Y_{n,n}$ unabhängig und identisch verteilt mit*

$$\mathbb{E}Y_{n1} = 0 \quad \textit{und} \quad \mathit{Var}(Y_{1,n}) = 1.$$

Ferner sei

$$\lim_{n\to\infty} \mathbb{E}1_{\{Y_{1,n}^2 \geq \varepsilon n\}} Y_{1,n}^2 = 0 \;\; \textit{für beliebige } \varepsilon > 0.$$

Dann konvergiert $(X_n)_n$ schwach gegen eine Brownsche Bewegung auf $[0,1]$, d.h. $X_n(.) \to_{\mathcal{D}} B(.)$ in $\mathcal{B}(\mathcal{C}[0,1])$.

BEWEIS. Nach Satz 8.15 müssen wir die Konvergenz der fidis und die Straffheit zeigen:

(i) fidis: Aus dem Lindebergschen Zentralen Grenzwertsatz kann man wegen der Kompaktheit von $[0,1]^2$ folgern, dass für beliebige $f \in \mathcal{C}_b(\mathbb{R})$ gilt:

$$\lim_{n\to\infty} \sup_{0\le t-u\le 1} \Big| \mathbb{E}f\big(S_n(u)-S_n(t)\big) - \mathbb{E}f\big(\sqrt{u-t}Y\big) \Big| = 0, \tag{9.5}$$

wobei $Y \sim \mathcal{N}(0,1)$. Seien $0 = t_0 < t_1 < ... < t_m \le 1$, $m \in \mathbb{N}$, fest (aber beliebig). Wir möchten die asymptotische gemeinsame Verteilung der Zuwächse $X_n(t_i) - X_n(t_{i-1})$ bestimmen. Wegen der Chebychef-Ungleichung können wir die Interpolationsterme vernachlässigen und ohne Einschränkung die Zuwächse von S_n betrachten. Diese sind aber unabhängig, weshalb

$$\Big(X_n(t_i) - X_n(t_{i-1})\Big)_{1\le i\le m} \longrightarrow_{\mathcal{D}} \mathcal{N}\Big(0, \operatorname{diag}\Big((t_i - t_{i-1})_{1\le i\le m}\Big)\Big).$$

Zusammen mit $X_n(0) = 0$ folgt hieraus die schwache Konvergenz der endlichdimensionalen Verteilungen von X_n.

(ii) Straffheit: Nach Satz 9.1 ist die stochastische Gleichstetigkeit zu zeigen. Wegen $w_{X_n}(\delta) \le w_{S_n}(\delta)$ genügt nach den beiden vorangehenden Lemmata hierzu der Nachweis, dass

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \sup_{0<u-t\le\delta} \frac{1}{\delta}\mathbb{P}\Big(\big| S_n(u) - S_n(t) \big| > \varepsilon \Big) = 0$$

für alle $\varepsilon > 0$ ist. Doch aus (9.5) folgt, dass für $0 < \delta \le 1$ gilt:

$$\begin{aligned}
\lim_{n\to\infty} \sup_{0<u-t\le\delta} \frac{1}{\delta}\mathbb{P}\Big(\big| S_n(u) - S_n(t) \big| > \varepsilon \Big) &= \sup_{0<u-t\le\delta} \frac{1}{\delta}\mathbb{P}\Big(\sqrt{u-t}|Y| > \varepsilon\Big) \\
&= \frac{1}{\delta}\mathbb{P}\Big(|Y| > \varepsilon/\sqrt{\delta}\Big) \\
&\le \frac{1}{2\delta}\exp\Big(-\frac{\varepsilon^2}{2\delta}\Big) \\
&\to 0 \quad \text{für } \delta \searrow 0.
\end{aligned}$$

□

Der klassische empirische Prozess

Seien $X_1, ..., X_n \overset{iid}{\sim} \mathbb{P}$, mit einer (unbekannten) Wahrscheinlichkeitsverteilung $\mathbb{P}$ auf $\mathcal{B}(\mathbb{R})$. Bekanntermaßen wird $\mathbb{P}$ durch die Verteilungsfunktion F, punktweise gegeben durch

$$F(x) := \mathbb{P}((-\infty, x])$$

eindeutig bestimmt (die halboffenen Intervalle $(-\infty, x]$, $x \in \mathbb{R}$, bilden ein $\cap$-stabiles Erzeugendensystem). Ein naheliegender Schätzer für F ist die *empirische Verteilungsfunktion* $\widehat{F}_n$ mit

$$\widehat{F}_n(x) := \widehat{P}_n((-\infty, x]) = \frac{1}{n}\sum_{i=1}^n 1\{X_i \le x\}.$$

Angenommen, die wahre Verteilungsfunktion F ist unbekannt, und man möchte testen, ob F gleich einer hypothetischen Verteilungsfunktion F_0 ist. Eine mögliche Testgröße hierfür ist die *Kolmogorov-Smirnov-Teststatistik*

$$\| \widehat{F}_n - F_0 \|_{\sup}.$$

Was die Abhängigkeit der Verteilung obiger Statistik von F betrifft, kann man den Prozess $\widehat{F}_n$ und davon abgeleitete Objekte sehr elegant standardisieren. Das Hilfsmittel hierfür ist die *Quantiltransformation* oder *Quantilfunktion*

$$F^{-1}(u) \ := \ \min\{x \in \mathbb{R} : F(x) \geq u\}, \quad u \in (0,1),$$

von $\mathbb{P}$ bzw. F. Man kann leicht zeigen, dass F^{-1} monoton wachsend und linksseitig stetig ist. Ferner gilt für $u \in (0,1)$ und $x \in \mathbb{R}$:

$$\begin{aligned} & F(F^{-1}(u) \geq u \\ & F(F^{-1}(u) = u \ \text{ falls } \ \mathbb{P}(F^{-1}(u)) = 0, \\ & F^{-1}(u) \leq x \ \text{ genau dann, wenn } \ u \leq F(x). \end{aligned} \tag{9.6}$$

Eigenschaft (9.6) kann folgendermaßen genutzt werden: Seien $U_1, U_2, ..., U_n \overset{iid}{\sim} \mathcal{U}[0,1]$. Dann impliziert (9.6), dass

$$X_i \ := \ F^{-1}(U_i) \ \sim \ F,$$

denn $\mathbb{P}(F^{-1}(U_i) \leq x) = \mathbb{P}(U_i \leq F(x)) = F(x)$. Mithilfe dieser speziellen Konstruktion der Variablen X_i und der empirischen Verteilungsfunktion

$$\widehat{G}_n(t) \ := \ \frac{1}{n} \sum_{i=1}^{n} 1\{U_i \leq t\}$$

der Variablen $U_1, ..., U_n$ ist dann $\widehat{F}_n(x) \ = \ \widehat{G}_n(F(x))$. Insbesondere ist

$$\| \widehat{F}_n - F \|_{\sup} \ = \ \sup_{x \in F(\mathbb{R})} \big| \widehat{G}_n(x) - x \big| \ \leq \ \| \widehat{G}_n - id \|_{\sup},$$

$$\| \widehat{F}_n - F \|_{\sup} \ = \ \| \widehat{G}_n - id \|_{\sup} \ \text{ falls } F \text{ stetig ist.} \tag{9.7}$$

Aufgrund dieser Darstellung der empirischen Verteilungsfunktion genügt es im wesentlichen, den uniformen empirischen Prozess $(\widehat{G}_n(t) - t)_{t \in [0,1]}$ zu untersuchen.

Satz 9.7. (Donsker) *Seien $U_1, ..., U_n \overset{iid}{\sim} \mathcal{U}[0,1]$ und $t \mapsto Y_n(t)$, $t \in [0,1]$, die lineare Interpolation des zugehörigen uniformen empirischen Prozesses $(\widehat{G}_n(t) - t)_{t \in [0,1]}$ an den Punkten $U_1, ..., U_k$ mit $Y_n(0) = Y_n(1) = 0$. Dann konvergiert $\sqrt{n} \cdot Y_n$ schwach gegen eine Brownsche Brücke, d.h.*

$$\sqrt{n} \cdot Y_n(.) \ \to_{\mathcal{D}} \ B^0(.) \ \ in \ \ \mathcal{B}(\mathcal{C}[0,1]).$$

BEWEIS. Der Beweis ist in vier Schritte unterteilt.

1. Schritt: Nach Satz 5.3 (e) ist der Prozess

$$t \ \mapsto \ \sum_{i=1}^{n} 1\{U_i \le t\} = n \cdot \widehat{G}_n(t)$$

genauso verteilt wie ein Poissonprozess N_n in $[0,1]$ zum Intensitätsmaß $\lambda{\cdot}n$ gegeben $N_n(1) = n$, d.h.

$$\mathcal{L}\big(n \cdot \widehat{G}_n(.)\big) \ = \ \mathcal{L}\big(N_n(.)|N_n(1) = n\big).$$

(Für eine Zufallsvariable X bezeichne $\mathcal{L}(X)$ ihre Verteilung (Bildmaß) – $\mathcal{L}$ steht für *law*, engl.). Entsprechend gilt

$$\begin{aligned}\mathcal{L}\big(\sqrt{n}(\widehat{G}_n(t) - t)_{t\in[0,1]}\big) \ &= \ \mathcal{L}\Big(\frac{1}{\sqrt{n}}\big(N_n(t) - n \cdot t\big)_{t\in[0,1]} \,\Big|\, N_n(1) = n\Big) \\ &= \ \mathcal{L}\Big(\frac{1}{\sqrt{n}}\big(N(t \cdot n) - n \cdot t\big)_{t\in[0,1]} \,\Big|\, N(n) = n\Big),\end{aligned}$$

wobei N ein Poissonprozess zum Intensitätsmaß λ ist. Die Gleichheit dieser Verteilungsgesetze bleibt offenbar bestehen, wenn man zu den linear interpolierten Prozessen von $(\widehat{G}_n(t) - t)_{t\in[0,1]}$ und $\big(N_n(t) - n \cdot t\big)_{t\in[0,1]}$ übergeht.

Sei N ein Poissonprozess der Intensität 1, d.h. der *zentrierte* Zuwachs $N(t+1) - N(t)$ in einem Zeitschritt eine *zentrierte* Poissonvariable zum Parameter 1 ist:

$$\mathbb{P}(X_i = m - 1) = \frac{e^{-1}}{m!} \quad \text{für } m \in \mathbb{N}_0. \tag{9.8}$$

Für den Partialsummenprozess

$$S^{(n)}(k/n) \ := \ \frac{1}{\sqrt{n}} \sum_{i=1}^{k} \big(N(i) - N(i-1) - 1\big) \ = \ \frac{1}{\sqrt{n}}(N(k) - k), \quad 0 \le k \le n,$$

$S(.)$ affin dazwischen, hat man das Invarianzprinzip von Donsker, was auch für die zwischen den Sprungstellen stetig linear interpolierte Version $X^{(n)}$ von $n^{-1/2}(N(n \cdot t) - n \cdot t)$ wegen $\big\| S^{(n)} - X^{(n)} \big\|_{\text{sup}} = O(n^{-1/2})$ bestehen bleibt. Allerdings gilt es nur für die *unbedingte* Verteilung. Wir müssen jetzt aber den Grenzwert der *bedingten Verteilungen unter der Bedingung* $N(n) = n$ oder, äquivalent, $X^{(n)}(1) = 0$ ausrechnen. Für jedes $a \in \mathbb{R}$ seien

$$\mu_a^{(n)} \ = \ \mathcal{L}\big((X^{(n)}(t))_{t\in[0,1]} \,\big|\, X^{(n)}(1) = a\big)$$

und

$$\mu_a \ = \ \mathcal{L}\big((B^0(t) + a \cdot t)_{t\in[0,1]}\big)$$

mit einer Brownschen Brücke B^0.

Für je zwei reelle Zahlen $a < b$ hat man nach dem Invarianzprinzip von Donsker und Proposition 7.9

$$\sum_{\nu\cdot\sqrt{n}\in(a,b)\cap\mathbb{Z}} \mathbb{P}\big(X^{(n)}(1)=\nu\big)\cdot\mu_\nu^{(n)}(.) \;\Rightarrow\; \int_a^b \frac{1}{\sqrt{2\pi}}\exp(-\nu^2/2)\mu_\nu(.)d\nu. \tag{9.9}$$

(Für eine offene Menge O impliziert $\mathbb{P}_n \Rightarrow \mathbb{P}$ die Konvergenz $\mathbb{P}(.\cap O) \Rightarrow \mathbb{P}(.\cap O)$, denn für jede offene Menge $G \in \mathcal{B}$ gilt $\liminf_n \mathbb{P}_n(G\cap O) \geq \mathbb{P}(G\cap O)$ nach dem Portemanteau-Theorem – wieder nach dem Portemanteau-Theorem, welches übrigens auch für allgemeine endliche Maße gilt, folgt dann $\mathbb{P}_n(.\cap O) \Rightarrow \mathbb{P}(.\cap O)$.)

2. Schritt: *Die Maße μ_a hängen in der schwachen Topologie stetig von a ab:* Dies folgt aus Proposition 7.9 (und wurde bereits erwähnt in der darauffolgenden Bemerkung); man erhält einen Prozess, der nach μ_b verteilt ist, indem man zu einem Prozess mit Verteilung μ_a den linearen Prozess $t \mapsto t\cdot(b-a)$ addiert. Folglich haben wir

$$\int f(x)d\mu_b(x) = \int f\big(x+(b-a)\cdot id\big)d\mu_a(x),$$

womit für jede beschränkte Lipschitz-Funktion $f : \mathcal{C}[0,1] \to \mathbb{R}$ mit Lipschitzkonstante C gilt

$$\Big|\int f(x)d\mu_b(x) - \int f(x)d\mu_a(x)\Big| \;\leq\; C\cdot(b-a) \;\text{ für alle } n \text{ und } a\leq b.$$

Da die duale BL-Metrik die schwache Topologie metrisiert, folgt die Behauptung.

3. Schritt: *Die Maße $\mu_a^{(n)}$ hängen in der schwachen Topologie gleichmäßig stetig von a ab:* Wegen der Aussage (e) des Satzes 5.3 erhält man einen Poissonprozess N unter der Bedingung $N(n) = n+k$ aus einem Poissonprozess N unter der Bedingung $N(n) = n+j$ $(j<k)$, indem man unabhängig $k-j$ Punkte iid nach $\mathcal{U}[0,n]$ hinzufügt. $N(t)$ wird dabei größer, aber nicht mehr als $k-j$. Folglich kann man zwei Prozesse $Y_a^{(n)}$ und $Y_b^{(n)}$ auf ein und demselben W-Raum definieren, die nach $\mu_a^{(n)}$ bzw. $\mu_b^{(n)}$ verteilt sind, so dass gilt

$$\mathbb{P}\Big(Y_a^{(n)}(t) \leq Y_b^{(n)}(t) \leq Y_a^{(n)}(t)+b-a \;\text{ für alle } t\in[0,1]\Big) \;=\; 1.$$

Wie in Schritt 2 folgt

$$\Big|\int f(x)d\mu_b^{(n)}(x) - \int f(x)d\mu_a^{(n)}(x)\Big| \;\leq\; C\cdot(b-a) \;\text{ für alle } n \text{ und } a\leq b.$$

für jede beschränkte Lipschitzfunktion $f : \mathcal{C}[0,1] \to \mathbb{R}$ mit Lipschitzkonstante C.

4. Schritt: Wendet man die schwache Konvergenzaussage (9.9) auf eine beschränkte Lipschitzfunktion f (mit Lipschitzkonstante C) an, erhält man

$$\lim_{n\to\infty}\sum_{\nu\cdot\sqrt{n}\in(a,b)\cap\mathbb{Z}} \mathbb{P}(X^{(n)}(1)=\nu)\cdot\int f(x)d\mu_\nu^{(n)}(x) \;=\; \int_a^b\int \frac{1}{\sqrt{2\pi}}\exp(-\nu^2/2)f(x)d\mu_\nu(x)d\nu. \tag{9.10}$$

Da die Abbildungen $\nu \mapsto \int f(x)d\mu_\nu^{(n)}$ *gleichmäßig in* n beschränkt und Lipschitz-stetig sind (ebenso $\nu \mapsto \int f(x)d\mu_\nu(x)$), ist nach dem Grenzwertsatz für $X^{(n)}(1)$ die linke Seite von (9.10) gleich

$$\lim_n \left(\int_a^b \frac{1}{\sqrt{2\pi}} \exp(-\nu^2/2) \int f(x) d\mu_\nu^{(n)}(x) d\nu \right)$$
$$= \lim_{n\to\infty} \left(\left\{ \int f(x) d\mu_a^{(n)}(x) \right\} \int_a^b \frac{1}{\sqrt{2\pi}} \exp(-\nu^2/2) d\nu + R_{a,b}(n) \right)$$

mit $\lim_n |R_{a,b}(n)| \leq (\Phi(b)-\Phi(a))\cdot C\cdot(b-a)$ nach Schritt 3, wobei Φ die Verteilungsfunktion der Standardnormalverteilung bezeichnet; entsprechend hat man nach Schritt 2 für die rechte Seite von (9.10)

$$\left\{ \int f(x) d\mu_a(x) \right\} \int_a^b \frac{1}{\sqrt{2\pi}} \exp(-\nu^2/2) d\nu + R_{a,b} \quad \text{mit} \quad |R_{a,b}| \leq (\Phi(b) - \Phi(a)) \cdot C \cdot (b-a).$$

Teilt man nun links und rechts in (9.10) durch $\Phi(b) - \Phi(a)$, erhält man im Limes $b - a \to 0$

$$\lim_n \int f(x) d\mu_a^{(n)}(x) = \int f(x) d\mu_a(x)$$

für alle $a \in \mathbb{R}$, f beschränkt und Lipschitz-stetig. □

Bemerkung 9.8. Alternativ kann man Satz 9.7 auch mithilfe des entwickelten Konvergenzkonzeptes *Konvergenz der endlichdimensionalen Verteilungen + Straffheit* zeigen; etwas aufwendiger wird hierbei der Nachweis der stochastischen Gleichstetigkeit, da die Zuwächse des empirischen Prozesses nicht mehr unabhängig sind (man verwendet zum Beispiel ein Kriterium wie in Satz 9.5). Übrigens bleibt obige Argumentation auch gültig, wenn man ohne die lineare Interpolation des empirischen Prozesses die Konvergenzaussage von Satz 9.7 in der schwachen Topologie der Maße über $\mathcal{B}(\mathcal{D}[0,1])$ nachweisen möchte.

Als Anwendung von Satz 9.7 erhalten wir die asymptotische Verteilung der Kolmogorov-Smirnov-Statistik unter der Nullhypothese:

Korollar 9.9. *Seien $X_1, ..., X_n$ iid nach einer Lebesgue-stetigen Verteilung. Dann gilt*

$$\sqrt{n} \cdot \left\| \widehat{F}_n(.) - F(.) \right\| \ \to_{\mathcal{D}} \max_{0\leq t\leq 1} |B^0(t)|.$$

BEWEIS. Wegen der Stetigkeit von F reicht es nach (9.7) den uniformen empirischen Prozess zu betrachten. Ferner gilt

$$\left| \left\| \widehat{G}_n - id \right\|_{\sup} - \left\| \tilde{G}_n - id \right\|_{\sup} \right| \leq \frac{4}{n},$$

wobei $\tilde{G}_n(.)$ die stetig lineare Interpolation der empirischen Verteilungsfunktion ist. Die Aussage folgt dann aus Satz 9.7 mithilfe des continuous-mapping-Theorems, da Supremum und Betrag stetig sind. □

Der empirische Spektralprozess

Sei (X_n) ein stationärer Gaußprozess mit $\mathbb{E}X_1 = 0$ und Spektralmaß $F(.)$. Ausgehend von Beobachtungen $X_1, ..., X_n$ hatten wir in Kapitel 3 das empirische Spektralmaß als Spektralmaß zur Folge $\big(\widehat{c}_n(k)\big)$ der empirischen Kovarianzen definiert, welche positiv definit ist. Es ist Aussage von Satz 3.8, dass f.s. für $n \to \infty$ das empirische Spektralmaß schwach gegen das wahre Spektralmaß konvergiert. Für statistische Zwecke ist ein solches Resultat von limitiertem Nutzen – wichtiger im Zusammenhang mit Goodness-of-Fit-Tests ist die (asymptotische) Verteilung von Schätzern.

Wegen der Symmetrie um 0 ist das Spektralmaß auf Π bereits eindeutig festgelegt durch die Funktion F', gegeben durch

$$F'(\lambda) := \int_0^\lambda dF(\alpha), \quad \lambda \in [0, \pi].$$

Es erweist sich als technisch günstiger, die schwache Approximationstheorie für $(\widehat{F}_n' - F')$ anstelle von $(\widehat{F}_n - F)$ zu entwickeln, wobei

$$\widehat{F}_n'(\lambda) \ := \ \int_0^\lambda I_n(\alpha) d\alpha.$$

In Analogie zum klassischen empirischen Prozess definiert man den sogenannten *empirischen Spektralprozess*

$$\big(\widehat{E}_n(\lambda)\big)_{\lambda\in[0,\pi]} \ := \ \big(\widehat{F}_n'(\lambda) - F'(\lambda)\big)_{\lambda\in[0,\pi]}.$$

Satz 9.10. (Dahlhaus) *Sei (X_n) ein stationärer, zentrierter Gaußprozess mit Lipschitz-stetiger Spektraldichte f. Dann gilt*

$$\sqrt{n} \cdot \widehat{E}_n(.) \ \to_{\mathcal{D}} \ E(.)$$

in $\mathcal{B}(\mathcal{C}[0, \pi])$, wobei E ein stetiger, zentrierter Gaußprozess auf $[0, \pi]$ ist mit Kovarianzstruktur

$$\operatorname{cov}\big(E(\lambda), E(\mu)\big) \ = \ 2\pi \int_0^{\min(\lambda,\mu)} f(\alpha)^2 d\alpha.$$

Bemerkung 9.11. Im Gegensatz zum klassischen empirischen Prozess ist der schwache Grenzwert eine zeittransformierte Brownsche Bewegung, die der stochastischen Differentialgleichung

$$dE(\lambda) \ = \ \sqrt{2\pi} f(\lambda) \cdot dB(\lambda)$$

genügt (siehe nachfolgendes Kapitel).

Für einen Beweis verweisen wir auf

Dahlhaus, R. (1988). Empirical spectral processes and their applications to time series analysis, *Stochastic Processes and their Applications*, **30**, pp 69-83.