III. Anmerkungen zu einem Bootstrap - Grenzwertsatz

1. Einführung

Betrachtet wird ein sehr allgemein gehaltener Bootstrap-Grenzwertsatz von F.Strobl (1995) für eine Schätzfunktion, die mit einem statistischen Funktional T gebildet wird. Im III.Kapitel wird die Frage untersucht, unter welchen Voraussetzungen eine (in gewissem Sinne) erweiterte Form des Fréchet-Differentials von T die im Satz angegebene Integral-Darstellung besitzt, und wie diese aussieht.

Zunächst wird der Grenzwertsatz vorgestellt, wozu die Einführung eines relativ umfangreichen theoretischen Rahmens unumgänglich ist.

Sei $\mathcal{F}$ eine beliebige Verteilung in $\mathbb{R}$ mit der Verteilungsfunktion F, $F : \mathbb{R} \to \mathbb{R}$.

Mit $\mathbb{M}(\mathbb{R})$ wird der $\mathbb{R}$-Vektorraum der signierten finiten Maße in $\mathbb{R}$ bezeichnet.

$\mathbb{W} \subset \mathbb{M}(\mathbb{R})$ sei eine Menge von Wahrscheinlichkeitsmaßen, zu der neben $\mathcal{F}$ zumindest noch alle Verteilungen in $\mathbb{R}$ mit endlichem Träger gehören.

Ein statistisches Funktional T ist eine reellwertige Abbildung auf $\mathbb{W}$, $T : \mathbb{W} \to \mathbb{R}$.

Unter relativ wenig einschränkenden Annahmen sollen Informationen über einen gewissen Parameter $\theta = T(\mathcal{F})$ gewonnen werden.

Vom Funktional T wird keine Meßbarkeit vorausgesetzt, aber eine erweiterte Form der Fréchet-Differenzierbarkeit bezüglich einer geeigneten Halb-Norm, bezüglich welcher zugleich der Empirische $\mathbb{F}$–Prozeß $(\beta_n(f))_{f\in \mathbb{F}}$ stochastisch beschränkt ist in $\mathbb{R}^{\mathbb{F}}$.

Es ist $\beta_n(f) := \sqrt{n} \cdot \int_{\mathbb{R}} f \, d(\mathcal{F}_n - \mathcal{F})$ mit dem Empirischen Maß $\mathcal{F}_n$, $f \in \mathbb{F} \subset \mathcal{L}_1(\mathbb{R}, B, \mathcal{F})$ und $n \in \mathbb{N}$.

Es kann für Anwendungen sehr vorteilhaft sein, wenn ein manchmal relativ aufwendiger Meßbarkeitsnachweis für ein auf einem Maßraum definiertes Funktional nicht erbracht werden muß. Durch die Flexibilität des verwendeten Konzepts kann die benötigte Form der Fréchet-Differenzierbarkeit eines T in vielen Fällen durch entsprechende Anpassung erreicht werden. Die Flexibilität ergibt sich <u>einerseits</u> aus der Verwendung eines von R.M.Dudley (1990, S.65) vorgeschlagenen Typs von Halb-Normen, verbunden mit einem wählbaren Funktionenraum $\mathbb{F} \subset \mathcal{L}_1(\mathbb{R}, B, \mathcal{F})$ <u>andererseits</u> aus dem von F.Strobl konstruierten Definitionsbereich der Fréchet-Ableitung $T'_{\mathcal{F}}$, welcher über einen Parameter r variiert werden kann.

Die Schätzfunktion T_n für $\theta = T(\mathcal{F})$ wird mittels *plug in* - Prinzip gebildet. Das Empirische Maß $\mathcal{F}_n$ wird anstelle von $\mathcal{F}$ als Argument des Funktionals eingesetzt: $T_n := T(\mathcal{F}_n)$.
Zu untersuchen ist die Qualität dieser Schätzfunktion. Auskunft darüber gibt die Verteilung der Zufallsgröße $R_n = \sqrt{n} \cdot [\, T_n - \theta \,]$.
Die R_n-Verteilung H_n ist zwar selbst unbekannt, aber mit einer qualifizierten H_n-Schätzung können qualifizierte Aussagen über H_n und letztlich über θ gewonnen werden. Die verwendete H_n-Schätzfunktion soll zumindest in irgendeiner wohl definierten Form gegen ein Grenzelement H konvergieren. Hintergrund dieser Forderung ist, dass die punktweise Konvergenz der H_n -Schätzfunktion gegen ein H in jedem Stetigkeits-Argument von H bedeutet, dass R_n zumindest in Verteilung gegen eine Zufallsgröße R strebt. Daraus folgt die wichtige Beziehung (J.Shao / D.Tu 1996, S. 449): $T_n = \theta + O_P(n^{-1/2})$, $n \to +\infty$.
Als H_n-Schätzfunktion wird hier die Bootstrap-Verteilungsfunktion $H_{n,boot}$ verwendet.
$H_{n,boot}$ ist die Verteilungsfunktion der Bootstrap-Version R_n^* von R_n ,

$R_n^* = \sqrt{n} \cdot [\, T_n^* - T_n \,]$ mit $T_n^* = T(\mathcal{F}_n^*)$, $\mathcal{F}_n^*$ ist die Bootstrap-Version von $\mathcal{F}_n$.

Ähnlich wie die Empirische Verteilungsfunktion ist $H_{n,boot}(t)$ zu jedem festen $t \in \mathbb{R}$ eine Zufallsgröße - sie hängt von den Stichproben-Variablen der Original-Stichprobe ab. Daher macht es Sinn, eine bestimmte Form der Konvergenz von $H_{n,boot}$ gegen ein H zu definieren, die nur fast sicher gegeben ist. Anders formuliert:
Es ist eine Form der fast sicheren Verteilungskonvergenz von R_n^* gegen ein R zu definieren, R besitzt die Verteilungsfunktion H .
Wegen der für T nicht vorausgesetzten Meßbarkeit muß die gewöhnliche Verteilungskonvergenz verallgemeinert werden. F.Strobl erweitert zu diesem Zweck die Verteilungskonvergenz von Hoffmann-Jørgensen.
Der Bootstrap-Grenzwertsatz gibt die Voraussetzungen an, unter denen R_n^* in dem von F.Strobl definierten Sinne fast sicher in Verteilung gegen eine Zufallsgröße R konvergiert, die normalverteilt ist mit dem Erwartungswert Null und einer Varianz, die von einer gewissen Funktion $f_{\mathcal{F}}$ abhängt.
Eine der Voraussetzungen des Grenzwertsatzes ist die Darstellbarkeit (einer erweiterten Form) des Fréchet-Differentials von T an $\mathcal{F}$ als Integral der o.g. Abbildung $f_{\mathcal{F}}$ für geeignete $\mathcal{G} \in \mathbb{W}$:

$$T'_{\mathcal{F}}(\mathcal{G}-\mathcal{F}) = \int_{\mathbb{R}} f_{\mathcal{F}} \, d(\mathcal{G}-\mathcal{F}) \quad .$$

$T'_{\mathcal{F}}$ ist bezüglich einer Halb-Norm $\|\cdot\|_{\mathbb{F}}$ (mit einer Funktionenmenge $\mathbb{F}$) auf einem speziellen, über einen Parameter r variierbaren Vektorraum M_r definiert.

Sei $\mathcal{C}^b$ die Menge aller auf $\mathbb{R}$ definierten, reellwertigen, stetigen und beschränkten Funktionen.

Für den Fall $\mathbb{F}' = \mathcal{C}^b$ und $r = +\infty$ hat R.J.Serfling (1980, S. 222) mit $T_1(\mathcal{F}, \cdot)$ eine Funktion angegeben, so dass mit jedem in $\mathbb{R}$ definierten Wahrscheinlichkeitsmaß $\mathcal{G}$ die Integral-Darstellung $T'_{\mathcal{F}}(\mathcal{G}-\mathcal{F}) = \int_{\mathbb{R}} T_1(\mathcal{F}, .)\, d(\mathcal{G}-\mathcal{F})$ gültig ist.

Hier wird für $\mathbb{F}' = \mathcal{C}^b$ und beliebiges $r > 0$ untersucht, welche Wahrscheinlichkeitsmaße $\mathcal{G} \in \mathbb{W}$ die Integral-Darstellung unter gewissen Bedingungen erfüllen, und welche Form $f_{\mathcal{F}}$ besitzt. Dazu wird eine Zerlegung des Maßraumes $\mathbb{W}$ vorgenommen. Basis der Zerlegung von $\mathbb{W}$ ist die Menge $A_r(\mathcal{C}^b)$, $A_r(\mathcal{C}^b) := \left\{ t \in R \;:\; \sup_{f \in \mathcal{C}^b} \left| \int_{\mathbb{R}} f\, d(\delta_t - \mathcal{F}) \right| = \| \delta_t - \mathcal{F} \|_{\mathcal{C}^b} < r \right\}$.

$A_r(\mathcal{C}^b)$ ist die Menge aller $t \in \mathbb{R}$, für die $(\delta_t - \mathcal{F}) \in M_r$ gilt, wenn man $\mathbb{F}' = \mathcal{C}^b$ setzt.

Es wird gezeigt, dass $A_r(\mathcal{C}^b)$ für jedes $r > 0$ eine Borelmenge ist, und somit als Integrationsbereich für $f_{\mathcal{F}}$ bezüglich jedes Maßes aus $\mathbb{W}$ verwendet werden darf.

Für die einzelnen Komponenten der $\mathbb{W}$-Zerlegung werden (falls möglich) jeweils Bedingungen angegeben, unter denen die Integral-Darstellung von $T'_{\mathcal{F}}(\mathcal{G}-\mathcal{F})$ mit einem entsprechenden $\mathcal{G}$ existiert.

Speziell für $\mathcal{G}$ mit endlichem Träger, welcher nicht in $A_r(\mathcal{C}^b)$ liegt, befindet sich das Argument $(\mathcal{G} - \mathcal{F})$ nicht notwendig im Definitionsbereich M_r von $T'_{\mathcal{F}}$, so dass $T'_{\mathcal{F}}(\mathcal{G}-\mathcal{F})$ i.a. gar nicht definiert ist.

Für den globalen Sonderfall $A_r(\mathcal{C}^b) = \mathbb{R}$ (d.h. $r = +\infty$) stimmt $f_{\mathcal{F}}$ mit der bei R.J.Serfling angegebenen Funktion $T_1(\mathcal{F}, \cdot)$ überein.

2. Bootstrap-Variablen und Bootstrap-Verteilung

Zugrundegelegt wird ein Wahrscheinlichkeitsraum $(\Omega, \mathrm{A}, \mathcal{P})$.

Seien $X_1, \ldots, X_n$ ($n \in \mathbb{N}$) $\mathrm{A}, \mathcal{B}$ - meßbare, unabhängige und identisch verteilte Zufallsgrößen, $X_i : \Omega \to \mathbb{R}$; die Verteilung jedes X_i wird mit $\mathcal{F}$ bezeichnet, $i \in \{1, \ldots, n\}$.

$X_1^*, \ldots, X_n^*$ erhält man durch Ziehen mit Zurücklegen aus $M_n := \{ X_1, \ldots, X_n \}$.

$X_1^*, \ldots, X_n^*$ sind die bedingt unabhängigen Bootstrap-Variablen, gegeben M_n.

Es ist $P(X_i^* = X_k \mid M_n) = \frac{1}{n}$ für $k \in \{1, \ldots, n\}$ und $i \in \{1, \ldots, n\}$.

Anders als M_n ist $M_n^* := \{ X_1^*, \ldots, X_n^* \}$ keine Menge, sondern eine „ungeordnete Kollektion" (P.Hall 1992, S.2). Die Mengen-Schreibweise wird wie üblich trotzdem verwendet, und die Bezeichnung „ $A \in M_n^*$ " ist unmittelbar einsichtig.

Sei $\mathcal{X} := (X_k)_{k \in \mathbb{N}}$ die Folge der oben definierten unabhängigen und identisch gemäß $\mathcal{F}$ verteilten Zufallsgrößen. Mit $\mathcal{F}_n^{\mathcal{X}}$ wird das auf M_n basierende Empirische Maß bezeichnet.

Wenn die Folge $\mathcal{X}$ nicht hervorzuheben ist, wird kurz $\mathcal{F}_n := \mathcal{F}_n^{\mathcal{X}}$ gesetzt.

Betrachtet wird die von rechts stetige Version der Empirischen Verteilungsfunktion IF_n zum Stichprobenumfang $n \in \mathbb{N}$,

$$\mathrm{IF}_n(t) := \mathcal{F}_n^{\mathcal{X}}\big((-\infty\, ;\, t] \big) := \frac{1}{n} \cdot \sum_{i=1}^{n} 1_{(-\infty\,;\,t]}(X_i) \quad , \quad t \in \mathbb{R} . \tag{1}$$

Eine Realisation F_n von IF_n entsteht, wenn man in $\mathcal{F}_n^{\mathcal{X}}$ die Folge $\mathcal{X} = (X_i)_{i \in \mathbb{N}}$ ersetzt durch eine Zahlenfolge $x := (x_i)_{i \in N} \in \mathbb{R}^{\mathbb{N}}$:

$$F_n(t) := \mathcal{F}_n^{x}\big((-\infty\, ;\, t] \big) = \frac{1}{n} \cdot \sum_{i=1}^{n} 1_{(-\infty\,;\,t]}(x_i) \quad , \quad t \in \mathbb{R} . \tag{2}$$

$\mathcal{F}_n^{x}$ wird bezeichnet als das zu $x_1, \ldots, x_n$ gehörige Empirische Maß.

Entsprechend wird mit den Bootstrap-Variablen das auf M_n^* basierende Empirische Maß $\mathcal{F}_n^{*\mathcal{X}}$ sowie die zugehörige Verteilungsfunktion IF_n^* gebildet:

$$\mathrm{IF}_n^*(t) := \mathcal{F}_n^{*\mathcal{X}}\big((-\infty\, ;\, t] \big) := \frac{1}{n} \cdot \sum_{i=1}^{n} 1_{(-\infty\,;\,t]}(X_i^*) \quad , \quad t \in \mathbb{R} . \tag{3}$$

Wird für $\mathcal{X} = (X_k)_{k \in \mathbb{N}}$ eine Realisierung $x = (x_j)_{j \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ eingesetzt, erhält man

$$F_n^*(t) = \frac{1}{n} \cdot \sum_{i=1}^{n} 1_{(-\infty\,;\,t]}(x_i^*) \quad , \quad t \in \mathbb{R} \quad . \tag{4}$$

Nach Definition ist

$$R_n = \sqrt{n} \cdot [\, T_n - \theta \,] = \sqrt{n} \cdot [\, T(\mathcal{F}_n) - T(\mathcal{F}) \,] \quad . \tag{5}$$

Die Verteilungsfunktion $H_n(t)$ von R_n wird definiert als von rechts stetig:

$$H_n(t) := P\Big(\sqrt{n} \cdot \big[T(\mathcal{F}_n) - T(\mathcal{F}) \big] \le t \Big) \quad , \; t \in \mathbb{R} \, .$$

$$R_n^* := \sqrt{n} \cdot [\, T(\mathcal{F}_n^*) - T(\mathcal{F}_n) \,] \tag{6}$$

ist die mit dem *plug in* - Prinzip gewonnene Bootstrap-Version von R_n.
R_n^* besitzt die Verteilungsfunktion

$$H_{n\,,\,\mathrm{boot}}(t) := P\big(\sqrt{n} \cdot [\, T(\mathcal{F}_n^*) - T(\mathcal{F}_n) \,] \le t \;\big|\; M_n \big) \; , \; t \in \mathbb{R} \; . \tag{7}$$

$H_{n\,,\,\mathrm{boot}}$ wird als Bootstrap-Verteilung von R_n^* bezeichnet [P.Hall 1992, S. 83].
$H_{n\,,\,\mathrm{boot}}$ ist eine „Zufall-bedingte“ Verteilung, gegeben M_n. Zu jedem festen $t \in \mathbb{R}$ ist $H_{n\,,\,\mathrm{boot}}(t)$ eine Zufallsgröße. Daher macht es Sinn, eine Verteilungskonvergenz für R_n^* zu definieren, die nur fast sicher gültig ist.

3. Vorbereitungen für den Grenzwertsatz

3.1 Eine verallgemeinerte Verteilungs-Konvergenz

Als nächstes wird die Hoffmann-Jørgensen-Verteilungskonvergenz definiert.
Diese Konvergenz wurde eingeführt für eine Folge von Zufallselementen, welche selbst nicht meßbar zu sein brauchen, für die nur ein noch zu definierendes Grenzelement als meßbar vorausgesetzt wird (J. Hoffmann-Jørgensen 1991). Sie wird im Grenzwertsatz für die Konvergenz von R_n^* benötigt. Für Anwendungen kann es sehr vorteilhaft sein, wenn ein manchmal etwas aufwendiger Meßbarkeits-Nachweis für Funktionale auf einem Maßraum nicht erbracht zu werden braucht.
Die hier angegebene Definition ist ein Spezialfall jener von F.Strobl (1995, S. 6), indem man dort $E = \mathbb{R}$ setzt.

Die Situation H

Zugrunde liegen der oben angegebene Wahrscheinlichkeitsraum $(\Omega, A, \mathcal{P})$ und der meßbare Raum $(\mathbb{R}, \mathcal{B})$.

Sei $S \subset \mathbb{R}$ und $\mathcal{B}(S)$ die σ-Algebra der Borelmengen in S.

Definiert werden Zufallselemente folgender Art:

$W_n: \Omega \to \mathbb{R}$ für $n \geq 1$; jedes solche W_n ist ein Zufallselement, das nicht als $A, \mathcal{B}$ - meßbar vorausgesetzt ist.

$W_0: \Omega \to S$, W_0 ist $A, \mathcal{B}(S)$ - meßbar.

Um die verallgemeinerte Verteilungskonvergenz zu formulieren, wird noch eine Definition benötigt.

DEF : „Das äußere Integral eines Zufallselements“

Sei $(\Omega, A, \mathcal{G})$ ein Wahrscheinlichkeitsraum.

Sei $\overline{\mathcal{B}}$ die σ-Algebra der Borelmengen in $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ (siehe H.Bauer 1968, S. 43).

Das äußere Integral für ein nicht als meßbar vorausgesetztes Zufallselement Y, $Y: \Omega \to \overline{\mathbb{R}}$, wird folgendermaßen definiert:

$$\overline{E}(Y) := \inf\{\ E(Z) \mid Z: \Omega \to \overline{R},\ Z \text{ ist } A, \overline{\mathcal{B}} \text{ - meßbar und quasiintegrierbar, } Z \geq Y\ \}. \qquad (8)$$

DEF : „Verteilungskonvergenz in der Situation H “(kurz: H-Verteilungskonvergenz)

Mit den Größen in Situation H wird definiert:

Die Folge $(W_n)_{n \in \mathbb{N}}$ heißt H-verteilungskonvergent gegen W_0, falls gilt:

$\lim\limits_{n \to +\infty} \overline{E}(f \circ W_n) = E(f \circ W_0)$ für alle f, $f: \mathbb{R} \to \mathbb{R}$, f auf $\mathbb{R}$ beschränkt und auf S auch stetig.

Symbolisiert wird diese Konvergenz durch: $W_n \overset{d(H)}{\to} W_0$.

Bemerkung::

Ist $S = \mathbb{R}$ und sind die W_n $A, \mathcal{B}$ - meßbar ($n \in \mathbb{N}$), dann stimmt die H-Verteilungskonvergenz überein mit der Verteilungskonvergenz für eine Folge reellwertiger, meßbarer Zufallsgrößen (R.J.Serfling 1980, S. 8).

3.2 Die erweiterte Fréchet-Ableitung eines Funktionals

Es wird eine erweiterte Form der Fréchet-Ableitung einer Abbildung angegeben. Die gewöhnliche Definition (z.B. H.Heuser 2000, II S. 331) basiert auf Abbildungen zwischen Banachräumen - hier wird aber eine Abbildung auf einer Menge von Wahrscheinlichkeitsmaßen betrachtet, die Teilmenge eines i.a. nicht vollständigen und nur halb-normierten Vektorraumes ist.

DEF „(Erweiterte) Fréchet-Ableitung einer Abbildung" (F.Strobl 1995, S. 112)
Sei V ein Vektorraum über $\mathbb{R}$, $U \subset V$ ein linearer Unterraum und $\|\cdot\|$ eine endliche Halb-Norm auf U. Sei weiter $M \subset V$ eine beliebige Teilmenge von V.
Die Abbildung $T: M \to \mathbb{R}$ heißt bezüglich $\|\cdot\|$ Fréchet-differenzierbar an $b \in M$ innerhalb U, falls eine auf U definierte lineare Abbildung $T'_b : U \to \mathbb{R}$ existiert, so dass gilt:
Es ist $T(x) - T(b) = T'_b(x - b) + o(\|x - b\|)$ für alle $x \in M$ mit

$$(x - b) \in U \quad \text{und} \quad \|x - b\| \to 0$$

T'_b heißt Fréchet-Ableitung von T an b innerhalb U bezüglich $\|\cdot\|$.

Ergänzungen:
(a) Diese Definition kann man auch alternativ formulieren (F.Strobl 1995, S. 113):

$$\left| T(x) - T(b) - T'_b(x - b) \right| \leq \|x - b\| \cdot \epsilon(\|x - b\|) \tag{9}$$

für alle $x \in M$ mit $(x - b) \in U$ und $\|x - b\| \to 0$
sowie einer Abbildung $\epsilon: \mathbb{R}_+ \to \overline{\mathbb{R}}_+$, für die gilt: $\lim\limits_{t \downarrow 0} \epsilon(t) = 0 = \epsilon(0)$.

(b) T wird als Fréchet-differenzierbar an b bezüglich $\|\cdot\|$ bezeichnet, sofern T an b bezüglich $\|\cdot\|$ <u>innerhalb ganz V</u> Fréchet-differenzierbar ist.
Diese „globale" Definition der Fréchet-Ableitung T'_b kann man alternativ formulieren (F.Strobl 1995, S. 114). Es gilt für alle beschränkten, offenen Mengen $S \subset M$:

$$\sup_{x \in S} \left\{ \left| T(x + t\cdot(x-b)) - T(x) - t \cdot T'_b(x-b) \right| \right\} = o(t) \qquad \text{für } t \to 0$$

Im folgenden wird der Fall betrachtet, dass $V = \mathbb{M}(\mathbb{R})$, $M = \mathbb{W}$ und T ein statistisches Funktional ist, d.h. $T: \mathbb{W} \to \mathbb{R}$. Durch die Wahl einer geeigneten Halb-Norm (d.h. einer solchen, die eine hinreichend feine Topologie erzeugt, die also genügend große Werte für einen Abstand liefert) ist in vielen Fällen die Fréchet-Differenzierbarkeit von T an ein $\mathcal{Q} \in \mathbb{W}$

erreichbar. Bezüglich der in Anwendungen viel genutzten Supremumsnorm $\|\cdot\|_\infty$ ist die Fréchet-Differenzierbarkeit aber oft schon bei einfachen nicht-linearen T nicht mehr gegeben.

R.M.Dudley (1990, S. 64) zeigt, dass $T(\mathcal{F}) = \left[\int_\Omega X \, d\mathcal{F}\right]^2$ an $\mathcal{F}$ nicht bezüglich der Supremumsnorm, aber bezüglich einer anderen Norm an $\mathcal{F}$ Fréchet-differenzierbar ist, sofern für die Verteilung $\mathcal{F}$ ein Erwartungswert definiert ist.

Dagegen ist $T(\mathcal{F}) = \left[\int_\Omega X^2 d\mathcal{F}\right] - \left[\int_\Omega X \, d\mathcal{F}\right]^2$ an $\mathcal{F}$ bezüglich $\|\cdot\|_\infty$ Fréchet-differenzierbar, sofern für $\mathcal{F}$ die Varianz definiert ist (R.J.Serfling 1980, S. 220).

R.M.Dudley (1990, S. 65) betrachtet Halb-Normen eines bestimmten Typs auf einer Menge von Verteilungen.

DEF : Eine Halb-Norm auf $U \subset \mathbb{W}$.

Sei $\mathbb{F} \subset \bigcap_{\mu \in U} \mathcal{L}_1(\mathbb{R}, \mathcal{B}, \mu)$ eine Funktionenmenge.

Für $\mathcal{G} \in U$ wird bezüglich $\mathbb{F}$ folgende Halb-Norm definiert: $\|\mathcal{G}\|_{\mathbb{F}} := \sup_{f \in \mathbb{F}} \left|\int_{\mathbb{R}} f \, d\mathcal{G}\right|$.

F.Strobl (1995, S.118) zeigt, dass für $\|\cdot\|_{\mathbb{F}}$ insbesondere die Dreiecksungleichung gültig ist. Durch diese Halb-Norm wird für Maße $\mathcal{G}, \mathcal{D} \in U$ bezüglich $\mathbb{F}$ folgende Pseudo-Metrik induziert:

$$d(\mathcal{G}, \mathcal{D}) := \|\mathcal{G} - \mathcal{D}\|_{\mathbb{F}} = \sup_{f \in \mathbb{F}} \left|\int_{\mathbb{R}} f \, d(\mathcal{G} - \mathcal{D})\right| . \tag{10}$$

Dabei gilt (H.Bauer 1968, S.56) :

$f \in \mathcal{L}_1(\mathbb{R}, B, \mathcal{G}) \;\wedge\; f \in \mathcal{L}_1(\mathbb{R}, B, \mathcal{D})$

$\Rightarrow f \in \mathcal{L}_1(\mathbb{R}, B, \mathcal{G} \pm \mathcal{D})$ definiert durch $\int_{\mathbb{R}} f \, d(\mathcal{G} \pm \mathcal{D}) := \int_{\mathbb{R}} f \, d\mathcal{G} \pm \int_{\mathbb{R}} f \, d\mathcal{D}$.

Ein Beispiel für eine Metrik:

Sei $\mathbb{F} = \left\{ 1_{(-\infty; t]} \; ; \; t \in \mathbb{R} \right\}$, $U \subset \mathbb{W}$, G Verteilungsfunktion des Maßes $\mathcal{G} \in U$ und H Verteilungsfunktion des Maßes $\mathcal{H} \in U$. Dann gilt:

$\|\mathcal{G} - \mathcal{H}\|_{\mathcal{T}} = \sup_{t \in \mathbb{R}} | G(t) - H(t) | =: \|\mathcal{G} - \mathcal{H}\|_\infty$.

3.3 Die stochastische Beschränktheit des Empirischen $\mathbb{F}$ - Prozesses

Es wird definiert:

DEF : „Stochastische Beschränktheit des Empirischen $\mathbb{F}$-Prozesses“ (F.Strobl 1995, S. 119)

Sei $\mathbb{F} \subset \mathcal{L}_1(\mathbb{R}, \mathcal{B}, \mathcal{F})$ eine beliebige Teilmenge. $\mathbb{R}^{\mathbb{F}}$ ist die Menge aller Abbildungen von $\mathbb{F}$ nach $\mathbb{R}$, X_k ($k \in \mathbb{N}$) sind die im Abschnitt 3.2 definierten Zufallsgrößen der Folge $\mathcal{X}$.

Für jedes $f \in \mathbb{F}$, $n \in \mathbb{N}$ wird definiert:

$$\beta_n(f) := \sqrt{n} \cdot \left(\int_{\mathbb{R}} f \; d(\mathcal{F}_n - \mathcal{F}) \right) = \sqrt{n} \cdot \left(\int_{\mathbb{R}} f \, d\mathcal{F}_n - \int_{\mathbb{R}} f \, d\mathcal{F} \right)$$

$$= \sqrt{n} \cdot \left(\frac{1}{n} \cdot \sum_{k=1}^{n} f \circ X_k - \int_{\mathbb{R}} f \, d\mathcal{F} \right) . \tag{11}$$

$\beta_n := (\beta_n(f))_{f \in \mathbb{F}}$ wird als der Empirische $\mathbb{F}$- Prozeß zum Stichprobenumfang n bezeichnet. Wegen (10) und (11) gilt also:

$$\sup_{f \in \mathcal{F}} |\beta_n(f)| = \sqrt{n} \cdot \| \mathcal{F}_n - \mathcal{F} \|_{\mathbb{F}} . \tag{12}$$

$(\beta_n(f))_{f \in \mathbb{F}}$ heißt stochastisch beschränkt in $\mathbb{R}^{\mathbb{F}}$, falls gilt:

$$\lim_{t \to +\infty} \limsup_{n \to +\infty} \overline{\mathcal{P}} \left(\sup_{f \in \mathbb{F}} \left| \beta_n(f) \right| \geq t \right) = 0 . \tag{13}$$

Dabei ist $\overline{\mathcal{P}}$ das äußere Maß, definiert durch:

$\overline{\mathcal{P}}(M) := \inf \{ \mathcal{P}(N) : M \subset N \subset \Omega \}$ für $M \subset \Omega$.

Bemerkungen:

(a) $(\beta_n(f))_{f \in \mathbb{F}}$ kann man zu festem $n \in \mathbb{N}$ auffassen als Element von $\mathbb{R}^{\mathbb{F}}$, oder als signiertes finites Maß aus $\mathbb{M}(\mathbb{R})$.

(b) Hinreichende Bedingungen für die stochastische Beschränktheit von $(\beta_n(f))_{f \in \mathbb{F}}$ in $\mathbb{R}^{\mathbb{F}}$ findet man bei K.S.Alexander (1984) und M.Talagrand (1994).

(c) Verwendet man z.B. als Funktionenmenge $\mathbb{F} = \left\{ 1_{(-\infty; t]} \; ; \; t \in \mathbb{R} \right\}$, dann ist

$$\sup_{f \in \mathbb{F}} |\beta_n(f)| = \sqrt{n} \cdot \| \mathcal{F}_n - \mathcal{F} \|_{\mathbb{F}} = \sqrt{n} \cdot \mathbb{D}_n ,$$

$\mathbb{D}_n$ ist dabei die Kolmogorov / Smirnov - Distance. In diesem speziellen Fall folgt die stochastische Beschränktheit von $(\beta_n(f))_{f\in \mathbb{F}}$ aus der Ungleichung von Dvoretzki / Kiefer / Wolfowitz (R.J.Serfling 1980 S. 59).

(d) Verwendet man in $\mathbb{R}^{\mathbb{F}}$ die Supremumsnorm $\|\cdot\|_{\infty}$, und konvergiert $(\beta_n(f))_{f\in \mathbb{F}}$ mit $n \to +\infty$ in Verteilung gegen ein $\beta_0 \in \mathbb{R}^{\mathbb{F}}$, dann folgt die stochastische Beschränktheit von $(\beta_n(f))_{f\in \mathbb{F}}$ aus der Bemerkung 1.6 bei F.Strobl (1995, S. 11).

In der Literatur ist ebenfalls die „kompakte“ oder Hadamard-Differenzierbarkeit eines Funktionals gebräuchlich. Diese erhält man, wenn in (b) der Ergänzungen zur Definition der (erweiterten) Fréchet-Ableitung das Supremum nicht über die beschränkten, sondern über die kompakten Teilmengen von $M = \mathbb{W}$ erstreckt wird (R.M.Dudley 1990 , S. 65).
Bei der hier erweiterten Form der Fréchet-Ableitung sind Ausdrücke der Form

$\|\mathcal{G}\|_{\mathbb{F}} = \sup_{f\in\mathbb{F}} \left| \int_{\mathbb{R}} f \, d\mathcal{G} \right|$ mit einer wählbaren Funktionenmenge $\mathbb{F}$ zu betrachten, bei der Hadamard-Ableitung liegen kompakte Mengen des Typs $\sup_{\mathcal{G}\in \mathrm{IK}} \left| \int_{\mathbb{R}} f \, d\mathcal{G} \right| < +\infty$ mit stetigem, reellem f auf kompaktem Träger IK zugrunde (zu kompakten Mengen in Maßräumen siehe H.Bauer 1968, S.192).
Die (erweiterte) Fréchet-Differenzierbarkeit auf Basis des flexiblen (Halb-)Normen-Typs $\|\cdot\|_{\mathbb{F}}$ mit passendem $\mathbb{F}$ läßt sich mathematisch bequemer handhaben und in vielen wichtigen Fällen auch nachweisen.
Im unten stehenden Bootstrap-Grenzwertsatz wird die stochastische Beschränktheit des Empirischen $\mathbb{F}$- Prozesses $(\beta_n(f))_{f\in \mathbb{F}}$ vorausgesetzt, nicht dessen Konvergenz. Als weitere Voraussetzung wird die (erweiterte) Fréchet-Differenzierbarkeit von T an $\mathcal{F}$ benötigt. F.Strobl (1995, S. 152) zeigt, dass die Hadamard-Differenzierbarkeit von T nicht ausreicht, um die Satzaussage zu erhalten, wenn man von $(\beta_n(f))_{f\in \mathbb{F}}$ „nur“ die stochastische Beschränktheit fordert.
Es kann aber problematisch sein, eine Norm in einem Maßraum zu finden, die sowohl die (erweiterte) Fréchet-Differenzierbarkeit von T an $\mathcal{F}$ als auch die stochastische Beschränktheit von $(\beta_n(f))_{f\in \mathbb{F}}$ zuläßt. Zu dieser Problematik hat R.M.Dudley (1992) Ergebnisse erhalten.
Als letzte Vorbereitung für den Grenzwertsatz wird zu $\mathcal{F}$ und $r > 0$ ein halbnormierter linearer Unterraum $M_r := M_r(\mathcal{F})$ des $\mathbb{R}$-Vektorraums $\mathbb{M}(\mathbb{R})$ konzipiert, um die (erweiterte) Fréchet-Ableitung eines Funktionals T an $\mathcal{F}$ innerhalb $M_r(\mathcal{F})$ zu erklären.

DEF Der Definitionsbereich von $T'_{\mathcal{F}}$

Es sei $\mathbb{F} \subset \mathcal{L}_1(\mathbb{R}, \mathcal{B}, \mathcal{F})$; $\mathbb{W}$ ist ein Maßraum, der neben $\mathcal{F}$ zumindest noch alle Wahrscheinlichkeitsmaße in $\mathbb{R}$ mit endlichem Träger enthält.

$M_r := M_r(\mathcal{F})$ ist die Menge aller $\mathcal{H}$ mit folgenden Eigenschaften:

Es gibt ein $\mathcal{G} \in \mathbb{W}$ und ein $c \in \mathbb{R}$, so dass gilt:

(i) $\mathcal{H} = c \cdot (\mathcal{G} - \mathcal{F})$,

(ii) Es ist $\mathbb{F} \subset \mathcal{L}_1(\mathbb{R}, \mathcal{B}, \mathcal{G})$,

(iii) $\| \mathcal{G} - \mathcal{F} \|_{\mathbb{F}} < r$.

F.Strobl (1995, S. 117 f.) zeigt, dass M_r für jedes $r > 0$ zusammen mit $\| \cdot \|_{\mathbb{F}}$ ein halbnormierter (i.a. nicht vollständiger) linearer Unterraum von $\mathbb{M}(\mathbb{R})$ ist.

Mit der Wahl von r wird reguliert, welche Maße $\mathcal{G} \in \mathbb{W}$ für ein Element $(\mathcal{G} - \mathcal{F})$ der Basis des Vektorraums M_r infrage kommen.

Jedes $\mathcal{G} \in \mathbb{W}$ findet Eingang in ein Basis-Element $(\mathcal{G} - \mathcal{F})$ von M_r, wenn es nur „nahe genug" bei $\mathcal{F}$ liegt. Je größer r , um so mehr Maße aus $\mathbb{W}$ definieren Basis-Elemente von M_r. Je kleiner r , um so kleiner ist die Mächtigkeit der M_r - Basis.

4. Ein Bootstrap-Grenzwertsatz

Es gelten die obigen Bezeichnungen, insbesondere ist also $\mathcal{X} = (X_k)_{k \in N}$; $(\Omega, A, \mathcal{P})$ ist ein Wahrscheinlichkeits-Raum, X_k sind $A, \mathcal{B}$ - meßbare Zufallsgrößen, $k \in \mathbb{N}$.

Sei $IP = \bigotimes_{\mathbb{N}} \mathcal{P}$ das Maß auf der unendlichen Produkt - σ - Algebra $/A := \bigtimes_{\mathbb{N}} A$

(H.Bauer 1968, §33 ; hier mit der speziellen Indexmenge $I = \mathbb{N}$) .

Dann ist $IP\mathcal{X}^{-1}$ ein Wahrscheinlichkeitsmaß auf der σ-Algebra der Borelmengen des $\mathbb{R}^{\mathbb{N}}$.

Es gilt der folgende Satz.

Satz III.1

Sei $T : \mathbb{W} \to \mathbb{R}$ ein statistisches Funktional, das an $\mathcal{F}$ Fréchet-differenzierbar innerhalb M_r ist im Sinne obiger Definition, und zwar bezüglich der Halb-Norm $\| \cdot \|_{\mathbb{F}}$ mit einem $r > 0$ und $\mathbb{F} \subset \mathcal{L}_1(\mathbb{R}, \mathcal{B}, \mathcal{F})$.

$\mathbb{W}$ enthält neben $\mathcal{F}$ zumindest noch alle Wahrscheinlichkeitsmaße in $\mathbb{R}$ mit endlichem Träger.

Die Abbildung ϵ aus (9) besitze die folgende Konvergenzordnung:

$$\epsilon(t) = o(\ln(\max\{|\ln t|, e\})) \quad \text{für} \quad t \searrow 0 . \tag{14}$$

Der Empirische $\mathbb{F}$- Prozeß $(\beta_n(f))_{f\in\mathbb{F}}$ sei stochastisch beschränkt in $\mathbb{R}^{\mathbb{F}}$.

Es gelte für das äußere Integral $\bar{E}(\sup_{f\in\mathbb{F}} |f|^2) < +oo$.

Die (erweiterte) Fréchet-Ableitung $T'_{\mathcal{F}}$ besitze für geeignete $\mathcal{H} \in M_r$ die Darstellung

$$T'_{\mathcal{F}}(\mathcal{H}) = \int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{H} \tag{15}$$

mit einer Funktion $f_{\mathcal{F}}$, für die gilt: (16)

$f_{\mathcal{F}} \in \mathcal{L}_2(\mathbb{R}, \mathcal{B}, \mathcal{F})$ und $f_{\mathcal{F}} \in \mathcal{L}_1(\mathbb{R}, \mathcal{B}, \mathcal{G})$ für alle $\mathcal{G} \in \mathbb{W}$ mit $(\mathcal{G} - \mathcal{F}) \in M_r$.

Dann gilt für $IP\chi^{-1}$- fast alle Realisationen $x = (x_k)_{k\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ der Folge $\chi = (X_k)_{k\in N}$:

$$\sqrt{n}\cdot\left[T(\mathcal{F}_n^{*\chi}) - T(\mathcal{F}_n^{\chi}) \right] \xrightarrow{d(H)} R \tag{17}$$

$$\text{mit} \quad R \sim N(0 ; S^2(f_{\mathcal{F}})) \quad , \quad S^2(f_{\mathcal{F}}) := \int_{\mathbb{R}} f_{\mathcal{F}}^2\, d\mathcal{F} - \left(\int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{F}\right)^2 . \tag{18}$$

Beweis: F.Strobl (1995, S. 147 ff.)

4.1 Ergänzung zu Satz III.1

Mit $T''_{\mathcal{F}}$ wird die Fréchet-Ableitung 2.Ordnung bezeichnet; man erhält $T''_{\mathcal{F}}$ folgendermaßen:
Es wird $T'_{\mathcal{F}}$ anstelle von T in der Definition der (erweiterten) Fréchet-Ableitung eingesetzt
mit $U = M = M_r$ und $\|\kappa\| := \|\kappa\|_{\mathbb{F}}$ für $\kappa \in M_r$.
Existiert zusätzlich zu den Voraussetzungen des Satzes auch $T''_{\mathcal{F}}$ und ist $T''_{\mathcal{F}}$ sogar beschränkt, dann ist die für $\epsilon(t)$ geforderte Konvergenzordnung einfach nachweisbar (F.Strobl 1995, S.147).
Ohne eine solche zusätzliche Eigenschaft von T kann dieser Nachweis problematisch sein.

5. Anmerkungen zum Grenzwertsatz

Für die Funktionenmenge $\mathbb{F} = \mathcal{C}^b$ wird jetzt untersucht, welche Wahrscheinlichkeitsmaße $\mathcal{G} \in \mathbb{W}$ die Integral-Darstellung $T'_{\mathcal{F}}(\mathcal{G}-\mathcal{F}) = \int_{\mathbb{R}} f_{\mathcal{F}} \, d(\mathcal{G}-\mathcal{F})$ unter bestimmten Voraussetzungen erfüllen, und welche Form die Funktion $f_{\mathcal{F}}$ dann besitzt. Zu diesem Zweck wird $\mathbb{W}$ zerlegt. Basis dieser Zerlegung ist die im folgenden betrachtete Menge $A_r(\mathcal{C}^b)$.

5.1 Die Menge $A_r(\mathcal{C}^b)$

Sei C eine Menge, dann wird definiert: $\mathcal{C}^b(C) := \{ f \mid f : C \to \mathbb{R} \text{ stetig und beschränkt} \}$.
Im Falle $C = \mathbb{R}$ wird zur Abkürzung $\mathcal{C}^b := \mathcal{C}^b(\mathbb{R})$ geschrieben.
Für $r > 0$ wird die Menge $A_r(\mathcal{C}^b)$ definiert, die bei der Zerlegung des Maßraumes $\mathbb{W}$ eine wesentliche Rolle spielen wird.
Mit δ_t wird das Dirac'sche oder auch „Einpunkt-“ Maß bezeichnet, $t \in \mathbb{R}$.

DEF :

$$A_r := A_r(\mathcal{C}^b) := \left\{ t \in R \;:\; \sup_{f \in \mathcal{C}^b} \left| \int_{\mathbb{R}} f \, d(\delta_t - \mathcal{F}) \right| = \| \delta_t - \mathcal{F} \|_{\mathcal{C}^b} < r \right\} . \tag{19}$$

Das folgende Lemma zeigt, dass A_r zu jedem $r > 0$ eine Borelmenge in $\mathbb{R}$ ist. Diese Eigenschaft wird benötigt, um $f_{\mathcal{F}}$ über A_r und über $\mathbb{R} \setminus A_r$ i.f. Integrieren zu können.

Lemma III.1
Für jedes $r > 0$ ist A_r offen in $\mathbb{R}$.

Beweis:

Es wird definiert: $N(t) := \| \delta_t - \mathcal{F} \|_{\mathcal{C}^b}$, $t \in \mathbb{R}$.

Sei $t_0 \in \mathbb{R}$ beliebig und $(t_n)_{n \in \mathbb{N}}$ eine reelle Zahlenfolge mit Grenzwert t_0 .

Eine Anwendung der Vierecksungleichung (H.Heuser 2000, I S. 85) ergibt:

$$\left| N(t_n) - N(t_0) \right| = \left| \left\| \delta_{t_n} - \mathcal{F} \right\|_{\ell^b} - \left\| \delta_{t_0} - \mathcal{F} \right\|_{\ell^b} \right|$$

$$\leq \| \delta_{t_n} - \delta_{t_0} \|_{\ell^b} = \sup_{f \in \ell^b} \left| \int_{\mathbb{R}} f \, d(\delta_{t_n} - \delta_{t_0}) \right| = \sup_{f \in \ell^b} \left| f(t_n) - f(t_0) \right| .$$

Weil ℓ^b abgeschlossen in der Menge der beschränkten Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$ ist (H.Heuser 2000, II S. 703 f.), und weil jedes $f \in \ell^b$ beschränkt ist, gibt es ein $g \in \ell^b$, so dass gilt:

$$\sup_{f \in \ell^b} \left| f(t_n) - f(t_0) \right| = \left| g(t_n) - g(t_0) \right| .$$

Wegen der Stetigkeit von g folgt daraus:

$\lim_{n \to +\infty} \left| N(t_n) - N(t_0) \right| = 0$, also ist $N(t)$ stetig für alle $t \in \mathbb{R}$.

Mithilfe der Funktion $N(t)$ läßt sich $A_r(\ell^b)$ so schreiben: $A_r(\ell^b) = N^{-1}((0, r))$.
Aus der Stetigkeit von $N(t)$ in $\mathbb{R}$ folgt, dass $A_r(\ell^b)$ für jedes $r > 0$ eine offene Menge in $\mathbb{R}$ ist. Insbesondere ist $A_r(\ell^b)$ also eine Borelmenge in $\mathbb{R}$.

❑

5.2 Wahrscheinlichkeitsmaße mit endlichem Träger

Für jedes Wahrscheinlichkeitsmaß w wird mit supp(w) der Träger von w bezeichnet.
Gemäß Voraussetzung enthält der Definitionsbereich $\mathbb{W}$ von T neben $\mathcal{F}$ zumindest noch alle Wahrscheinlichkeitsmaße in $\mathbb{R}$, die einen endlichen Träger besitzen.
Die Menge der Wahrscheinlichkeitsmaße mit endlichem Träger wird mit Π bezeichnet.
Der Rand einer Menge M wird durch ∂M , der Abschluß von M durch $\overline{M} = M \cup \partial M$ bezeichnet.
Π liegt dicht in der Menge $WM(\mathbb{R})$ aller Wahrscheinlichkeitsmaße auf $\mathbb{R}$ bezüglich der schwachen Topologie (H.Bauer 1968, S. 186 ff.) , d.h. $WM(\mathbb{R}) = \Pi \cup \partial \Pi = \overline{\Pi}$.

5.3 Die Zerlegung von $\mathbb{W}$

Für jedes $\mathcal{G} \in \left\{ \mathbb{W} \cap \overline{\Pi} \right\}$ gibt es wegen $\mathbb{W} \subset \Pi$ eine der folgenden Darstellungen:

(i) $\mathcal{G} \in \Pi$:

$$\mathcal{G} = \sum_{k=1}^{m} \lambda_k \cdot \delta_{t_k} \quad \text{mit einem } m \in \mathbb{N}\,,\ t_k \in \mathbb{R} \text{ und } \lambda_k \in (0;1] \quad (k = 1 \ldots m),$$

$$\sum_{k=1}^{m} \lambda_k = 1\,,\ i \neq s \;\Rightarrow\; t_{i,n} \neq t_{s,n}\ .$$

(ii) $\mathcal{G} \in \{ \mathbb{W} \cap \partial \Pi \}$:

Es gibt eine Folge $(\epsilon_n)_{n \in \mathbb{N}} \subset \Pi$, als deren Grenzwert (bezüglich der schwachen Konvergenz) sich $\mathcal{G}$ erweist (H.Heuser 2000, II S. 62):

$$\int_{\mathbb{R}} f \; d\epsilon_n \;\to\; \int_{\mathbb{R}} f \; d\mathcal{G} \quad , \; n \to +\infty \quad \text{für jedes } f \in \ell^b\ .$$

Daraus folgt: $\sup\limits_{f \in \ell^b} \left| \int\limits_{\mathbb{R}} f \, d(\epsilon_n - \mathcal{G}) \right| = \left\| \epsilon_n - \mathcal{G} \right\|_{\ell^b} \to 0 \quad , \; n \to +\infty$.

Weil $(\epsilon_n)_{n \in \mathbb{N}} \subset \Pi$ ist, gilt für jedes ϵ_n ($n \in \mathbb{N}$) mit einem $m_n \in \mathbb{N}$:

$$\left.\begin{aligned} &\epsilon_n = \sum_{k=1}^{m_n} \lambda_{k,n} \cdot \delta_{t_{k,n}} \quad \text{mit } t_{k,n} \in \mathbb{R} \text{ und } \lambda_{k,n} \in (0;1] \quad (k \in \{1, \ldots, m_n\})\,, \\ &\sum_{k=1}^{m_n} \lambda_{k,n} = 1\,,\ i \neq s \;\Rightarrow\; t_{i,n} \neq t_{s,n} \quad (i, s \in \{1, \ldots, m_n\}). \end{aligned}\right| \tag{20}$$

Weil es Wahrscheinlichkeitsmaße in $\mathbb{R}$ gibt, deren Träger nicht endlich ist, ist Π nicht abgeschlossen in der Menge aller Wahrscheinlichkeitsmaße in $\mathbb{R}$.

Im folgenden werden die möglichen Fälle für die Darstellbarkeit $T'_{\mathcal{F}}(\mathcal{G} - \mathcal{F}) = \int\limits_{\mathbb{R}} f_{\mathcal{F}} \, d(\mathcal{G} - \mathcal{F})$ separat betrachtet, wenn man $\mathbb{F} = \ell^b$ als Funktionenmenge zugrundelegt.

Dazu wird mit dem Träger supp (w) des Wahrscheinlichkeitsmaßes w definiert:

$$\Pi_r := \left\{ w \in \Pi \; : \; \operatorname{supp}(w) \subset A_r(\ell^b) \right\} . \tag{21}$$

Π_r liegt dicht in $\overline{\Pi}_r$ bezüglich der schwachen Topologie.

Es wird von folgender Zerlegung von $\mathbb{W}$ ausgegangen:

$$\mathbb{W} = \left\{ \mathbb{W} \setminus \overline{\Pi}_r \right\} \cup \left\{ \mathbb{W} \cap \Pi_r \right\} \cup \left\{ \mathbb{W} \cap \partial \Pi_r \right\} . \tag{22}$$

5.4 Darstellungen des Fréchet-Differentials $T'_{\mathcal{F}}(\mathcal{G}-\mathcal{F})$

Sei im folgenden $r > 0$ eine beliebige, fest gewählte reelle Zahl.

Es werden gemäß Zerlegung (22) die nachfolgenden Fälle unterschieden.

5.4.1 Der Fall $\mathcal{G} \in \left\{ \mathbb{W} \setminus \overline{\Pi}_r \right\}$

Aus $\mathcal{G} \in \left\{ \mathbb{W} \setminus \overline{\Pi}_r \right\}$ folgt nicht $(\mathcal{G} - \mathcal{F}) \in M_r$, so dass $T'_{\mathcal{F}}(\mathcal{G} - \mathcal{F})$ i.a. gar nicht definiert ist. Das wird im folgenden begründet.

Da Π nicht abgeschlossen ist, ist $\Pi \neq \overline{\Pi}_r$; es gilt insbesondere $\Pi \setminus \Pi_r \neq \emptyset$.

Sei $\mathcal{G} \in \left\{ \Pi \setminus \overline{\Pi}_r \right\}$ und K die Menge aller t_k aus der Darstellung (20) mit $t_k \in A_r$.

Dann gilt wegen (21) die Darstellung:

$$\mathcal{G} - \mathcal{F} = \sum_{k=1}^{m} \lambda_k \cdot (\delta_{t_k} - \mathcal{F}) = \sum_{t_k \in K} \lambda_k \cdot (\delta_{t_k} - \mathcal{F}) + \sum_{t_k \notin K} \lambda_k \cdot (\delta_{t_k} - \mathcal{F})$$

$$=: \quad S_1 \quad + \quad S_2$$

Falls $K = \emptyset$ ist, gilt $0 = S_1 \in M_r$.

Wenn $K \neq \emptyset$ ist, liegt jeder Summand von S_1 in M_r, und weil M_r ein Vektorraum ist, gilt $S_1 \in M_r$.

Wegen $\mathcal{G} \notin \Pi_r$ nach Voraussetzung, ist $\mathbb{R} \setminus K \neq \emptyset$.

Da kein Summand von S_2 in M_r liegt, ist S_2 nur in Sonderfällen ein Element des Vektorraums M_r .

Daher folgt aus $\mathcal{G} \in \left\{ \mathbb{W} \setminus \overline{\Pi}_r \right\}$ nicht, dass $(\mathcal{G} - \mathcal{F}) \in M_r$ ist.

Beispiele:

(i) $x \notin A_r$, $\mathcal{F} \neq \delta_x =: \mathcal{G} \Rightarrow (\mathcal{G} - \mathcal{F}) \notin M_r$

(ii) $x_1 \notin A_r$, $x_2 \in A_r$, $\mathcal{F} \neq \delta_{x_1}$,

$\mathcal{G} := \lambda_1 \cdot \delta_{x_1} + \lambda_2 \cdot \delta_{x_2}$ mit $\lambda_1 + \lambda_2 = 1$; $\lambda_1 , \lambda_2 > 0$

$\Rightarrow (\mathcal{G} - \mathcal{F}) = \lambda_1 \cdot (\delta_{x_1} - \mathcal{F}) + \lambda_2 \cdot (\delta_{x_2} - \mathcal{F}) \notin M_r$.

5.4.2 Der Fall $\mathcal{G} \in \{ \mathbb{W} \cap \Pi_r \}$

Dieser Fall wird umfassend vom folgenden Lemma behandelt.

Lemma III.2

Es gelten die Bezeichnungen und Voraussetzungen von Satz III.1 , insbesondere ist also $T'_{\mathcal{F}} : M_r \to \mathbb{R}$ eine lineare Funktion (mit M_r wie im Satz angegeben), $r > 0$ beliebig vorgegeben.

Es wird definiert: $h_{\mathcal{F},T}(x) := T'_{\mathcal{F}}(\delta_x - \mathcal{F})$, $x \in A_r$.

Sei $h_{\mathcal{F},T} \in \mathcal{L}_1(A_r, \mathcal{B} \cap A_r, \mathcal{F})$.

Dann gilt die folgende Äquivalenz mit jedem $\mathcal{G} \in \{ \mathbb{W} \cap \Pi_r \}$:

$$f_{\mathcal{F}}(x) := f_{\mathcal{F},T}(x) = 1_{A_r}(x) \cdot \left\{ T'_{\mathcal{F}}(\delta_x - \mathcal{F}) + \int_{\mathbb{R}} f_{\mathcal{F}} \, d\mathcal{F} \right\} + 1_{R \setminus A_r}(x) \cdot g_{\mathcal{F},T}(x)$$

$$\text{für alle } x \in \mathbb{R} \quad \text{mit einem } g_{\mathcal{F},T} \in \mathcal{L}_1(\mathbb{R} \setminus A_r , B \cap \{ \mathbb{R} \setminus A_r \} , \mathcal{F})$$

<=>

$$T'_{\mathcal{F}}(\mathcal{G} - \mathcal{F}) = \int_{\mathbb{R}} f_{\mathcal{F}} \, d(\mathcal{G} - \mathcal{F})$$

Beweis:

Gemäß Lemma III.1 ist A_r eine offene Menge für jedes $r > 0$, so dass alle auftretenden Integrale über A_r und über $\mathbb{R} \setminus A_r$ erstreckt werden dürfen.

Sei $r > 0$ beliebig aber fest gewählt.

Wegen $\mathcal{G} \in \{ \mathbb{W} \cap \Pi_r \}$ gilt mit einem $m \in \mathbb{N}$:

$$\mathcal{G} = \sum_{k=1}^{m} \lambda_k \cdot \delta_{t_k} \quad \text{mit } t_k \in A_r \text{ und } \lambda_k \in [\, 0 \, ; 1 \,] \ (k \in \{1 , ..., m \}) ,$$

$$\sum_{k=1}^{m} \lambda_k = 1 , \quad i \neq s \Rightarrow t_i \neq t_s \ (i , s \in \{1 , ..., m \}).$$

Es werden beide Äquivalenz-Richtungen einzeln bewiesen.

“ <= “ :

Es gelte $T'_{\mathcal{F}}(\mathcal{G} - \mathcal{F}) = \int_{\mathbb{R}} f_{\mathcal{F}} \, d(\mathcal{G} - \mathcal{F})$.

Mit $m = 1$ ist $(\mathcal{G}-\mathcal{F}) = (\delta_{t_1}-\mathcal{F}) \in M_r$ und $t_1 \in A_r$.

Es folgt:

$$T'_{\mathcal{F}}(\mathcal{G}-\mathcal{F}) = T'_{\mathcal{F}}(\delta_{t_1}-\mathcal{F}) = \int_{\mathbb{R}} f_{\mathcal{F}}\, d(\delta_{t_1}-\mathcal{F}) = f_{\mathcal{F}}(t_1) - \int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{F}$$

$$\Rightarrow \quad f_{\mathcal{F}}(x) = T'_{\mathcal{F}}(\delta_x-\mathcal{F}) + \int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{F} \quad \text{für alle } x \in A_r$$

$$\Rightarrow \quad f_{\mathcal{F}}(x) = 1_{A_r}(x) \cdot \left\{ T'_{\mathcal{F}}(\delta_x-\mathcal{F}) + \int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{F} \right\} + 1_{\mathbb{R}\setminus A_r}(x) \cdot g_{\mathcal{F},T}(x)$$

$$\text{für alle } x \in \mathbb{R} \quad \text{mit einem } g_{\mathcal{F},T} \in \mathcal{L}_1(\mathbb{R}\setminus A_r, B \cap \{\mathbb{R}\setminus A_r\}, \mathcal{F})$$

“ => “ :

Jetzt wird ausgegangen von

$$f_{\mathcal{F}}(x) = 1_{A_r}(x) \cdot \left\{ T'_{\mathcal{F}}(\delta_x-\mathcal{F}) + \int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{F} \right\} + 1_{\mathbb{R}\setminus A_r}(x) \cdot g_{\mathcal{F}}(x)$$

$$\text{für alle } x \in \mathbb{R} \quad \text{mit einem } g_{\mathcal{F},T} \in \mathcal{L}_1(\mathbb{R}\setminus A_r, B \cap \{\mathbb{R}\setminus A_r\}, \mathcal{F})$$

Es folgt:

$$\int_{\mathbb{R}} f_{\mathcal{F}}\, d(\mathcal{G}-\mathcal{F}) = \int_{\mathbb{R}} f_{\mathcal{F}}\, d\left(\left(\sum_{k=1}^{m} \lambda_k \cdot \delta_{t_k} \right) - \mathcal{F} \right)$$

$$= \sum_{k=1}^{m} \lambda_k \cdot \int_{\mathbb{R}} f_{\mathcal{F}}\, d\delta_{t_k} - \int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{F}$$

$$= \sum_{k=1}^{m} \lambda_k \cdot f_{\mathcal{F}}(t_k) - \int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{F}$$

$$= \sum_{k=1}^{m} \lambda_k \cdot \left(1_{A_r}(t_k) \cdot \left\{ T'_{\mathcal{F}}(\delta_{t_k}-\mathcal{F}) + \int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{F} \right\} + 1_{\mathbb{R}\setminus A_r}(t_k) \cdot g_{\mathcal{F},T}(t_k) \right)$$

$$- \int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{F}$$

$$= \sum_{k=1}^{m} \lambda_k \cdot \left(T'_{\mathcal{F}}(\delta_{t_k}-\mathcal{F}) + \int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{F} \right) - \int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{F}$$

$$= \sum_{k=1}^{m} \lambda_k \cdot T'_{\mathcal{F}}(\delta_{t_k}-\mathcal{F}) .$$

Es ist $t_k \in A_r$, also $(\delta_{t_k} - \mathcal{F}) \in M_r$ für alle $k \in \{1, \dots, m\}$.

M_r ist Vektorraum, also folgt: $\sum_{k=1}^{m} \lambda_k \cdot \left(\delta_{t_k} - F\right) \in M_r$.

Damit gilt:

$$\sum_{k=1}^{m} \lambda_k \cdot T'_{\mathcal{F}}\left(\delta_{t_k} - \mathcal{F}\right) = T'_{\mathcal{F}}\left(\sum_{k=1}^{m} \lambda_k \cdot \left(\delta_{t_k} - \mathcal{F}\right)\right)$$

$$= T'_{\mathcal{F}}\left(\sum_{k=1}^{m} \lambda_k \cdot \delta_{t_k} - \mathcal{F} \right) = T'_{\mathcal{F}}(\mathcal{G} - \mathcal{F}) .$$

❑

Bemerkung zu Lemma III.2

Wegen $\int_R c \, d(\mathcal{G} - \mathcal{F}) = 0$ für jedes $c \in \mathbb{R}$ und für jedes $\mathcal{G} \in \mathbb{W}$ ist $f_{\mathcal{F}}$ durch

$T'_{\mathcal{F}}(\mathcal{G} - \mathcal{F}) = \int_{\mathbb{R}} f_{\mathcal{F}} \, d(\mathcal{G} - \mathcal{F})$ nur bis auf eine additive Konstante festgelegt.

$f_{\mathcal{F}}$ ist also unter den Voraussetzungen von Lemma III.2 zu jedem $r > 0$ auf A_r bis auf eine additive (von r abhängige) Konstante eindeutig bestimmt, auf $\mathbb{R} \setminus A_r$ hingegen lediglich an eine Integrierbarkeitsbedingung gebunden.

5.4.3 Der Fall $\mathcal{G} \in \left\{ \mathbb{W} \cap \partial \Pi_r \right\}$

Unter zusätzlichen Voraussetzungen zu denen von Satz III.1 erhält man die Identität

$T'_{\mathcal{F}}(\mathcal{G} - \mathcal{F}) = \int_{\mathbb{R}} f_{\mathcal{F}} \, d(\mathcal{G} - \mathcal{F})$ mit dem $f_{\mathcal{F}}$ aus Lemma III.2 für weitere $\mathcal{G} \in \mathbb{W}$.

Zur näheren Untersuchung wird vom nachfolgenden allgemeinen Fall ausgegangen.
Mit dem $f_{\mathcal{F}}$ aus Lemma III.2 und einer wie in (20) definierten Folge $(\epsilon_n)_{n \in \mathbb{N}} \subset \Pi$, die gegen ein $\mathcal{G} \in \left\{ \mathbb{W} \cap \overline{\Pi} \right\}$ schwach konvergiert, gilt für jedes $n \in \mathbb{N}$:

$$\int_{\mathbb{R}} f_{\mathcal{F}}\, d(\epsilon_n - \mathcal{F}) = \int_{\mathbb{R}} f_{\mathcal{F}}\, d\left(\left(\sum_{k=1}^{m_n} \lambda_{k,n} \cdot \delta_{t_{k,n}}\right) - \mathcal{F}\right)$$

$$= \sum_{k=1}^{m_n} \lambda_{k,n} \cdot \int_{\mathbb{R}} f_{\mathcal{F}}\, d\,\delta_{t_{k,n}} - \int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{F}$$

$$= \sum_{k=1}^{m_n} \lambda_{k,n} \cdot 1_{A_r}(t_{k,n}) \cdot T'_{\mathcal{F}}(\delta_{t_k} - \mathcal{F}) + \sum_{k=1}^{m_n} \lambda_{k,n} \cdot 1_{\mathbb{R}\setminus A_r}(t_{k,n}) \cdot g_{\mathcal{F}}(t_{k,n})$$

$$+ \sum_{k=1}^{m_n} \lambda_{k,n} \cdot 1_{A_r}(t_{k,n}) \cdot \int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{F} - \sum_{k=1}^{m_n} \lambda_{k,n} \cdot \int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{F}$$

$$\left.\begin{aligned} &= T'_{\mathcal{F}}\left(\sum_{k=1}^{m_n} \lambda_{k,n} \cdot 1_{A_r}(t_{k,n}) \cdot (\delta_{t_k} - \mathcal{F})\right) + \sum_{k=1}^{m_n} \lambda_{k,n} \cdot 1_{\mathbb{R}\setminus A_r}(t_{k,n}) \cdot g_F(t_{k,n}) \\ &\quad + \sum_{k=1}^{m_n} \lambda_{k,n} \cdot \left(1_{A_r}(t_{k,n}) - 1\right) \cdot \int_{\mathbb{R}} f_{\mathcal{F}}\, d\mathcal{F} \,. \end{aligned}\right] \qquad (23)$$

Im Falle $t_{k,n} \in A_r$ für alle $k \in \{1, ..., m_n\}$ stimmt dieser Ausdruck überein mit $T'_{\mathcal{F}}(\epsilon_n - \mathcal{F})$

Ansonsten ist die Identität von (23) mit $T'_{\mathcal{F}}(\epsilon_n - \mathcal{F})$ nur in Sonderfällen gegeben,

z.B. wenn $\lambda_{k,n} = 0$ für alle $k \in \{1, ..., m_n\}$ mit $t_{k,n} \notin A_r$.

Der Fall $\mathcal{G} \in \{ \mathbb{W} \cap \partial\Pi_r \}$ wird aufbauend auf der Darstellung (23) durch das folgende Lemma behandelt.

Lemma III.3

Es gelten die Bezeichnungen und Voraussetzungen von Lemma III.2 ; darüber hinaus wird von folgenden Bedingungen ausgegangen:

$h_{\mathcal{F},T} \in \ell^b(A_r)$, $g_{\mathcal{F},T} \in \ell^b(\mathbb{R} / A_r)$.

Außerdem wird vorausgesetzt, dass $\mathcal{G}(\partial A_r) = 0$ ist.

Dann gilt die Integral-Darstellung $T'_{\mathcal{F}}(\mathcal{G} - \mathcal{F}) = \int_{\mathbb{R}} f_{\mathcal{F}}\, d(\mathcal{G} - \mathcal{F})$

für jedes $\mathcal{G} \in \{ \mathbb{W} \cap \partial\Pi_r \}$.

Beweis:

Wegen der Voraussetzungen ist das in Lemma III.2 angegebene $f_{\mathcal{F}}$ beschränkt in $\mathbb{R}$, stetig in $\mathbb{R} \setminus \partial A_r$ und damit auch $\mathcal{B},\mathcal{B}$ - meßbar.

Weil Π_r dicht in $\overline{\Pi}_r$ liegt bezüglich der schwachen Topologie, darf von $\epsilon_n \in \Pi_r$ ausgegangen werden (für alle $n \in \mathbb{N}$). Für jedes ϵ_n $(n \in \mathbb{N})$ gilt also: $t_{k,n} \in A_r$ für alle $k \in \{1, \dots, m_n\}$.

Wegen (23) folgt daher: $\int_{\mathbb{R}} f_{\mathcal{F}} \, d(\epsilon_n - \mathcal{F}) = T'_{\mathcal{F}}(\epsilon_n - \mathcal{F})$ für alle $n \in \mathbb{N}$.

Es ist $\| \epsilon_n - \mathcal{G} \|_{\ell^b} \to 0$ für $n \to +\infty$, $f_{\mathcal{F}}$ ist $\mathcal{G}$-fast überall stetig und beschränkt in $\mathbb{R}$ sowie $\mathcal{B}, \mathcal{B}$ -meßbar.

Daraus folgt (H.Bauer 1968, S. 190): $\lim\limits_{n \to +\infty} \left(\int_{\mathbb{R}} f_{\mathcal{F}} \, d(\epsilon_n - \mathcal{F}) \right) = \int_{\mathbb{R}} f_{\mathcal{F}} \, d(\mathcal{G} - \mathcal{F})$.

Man erhält:

$$\int_{\mathbb{R}} f_{\mathcal{F}} \, d(\mathcal{G} - \mathcal{F}) = \lim_{n \to +\infty} T'_{\mathcal{F}}(\epsilon_n - \mathcal{F}) \; . \tag{24}$$

Abschließend wird gezeigt, dass $T'_{\mathcal{F}}(\mathcal{G} - \mathcal{F})$ der Grenzwert der rechten Seite von (24) ist.

Zunächst ist nachzuweisen, dass $(\mathcal{G} - \mathcal{F}) \in M_r$ ist.

Weil $t_{k,n} \in A_r$ für alle $k \in \{1, \dots, m_n\}$ ist, gilt:

$$\| \epsilon_n - \mathcal{F} \|_{\ell^b} = \left\| \sum_{k=1}^{m_n} \lambda_{k,n} \cdot \delta_{t_{k,n}} - \mathcal{F} \right\|_{\ell^b} = \left\| \sum_{k=1}^{m_n} \lambda_{k,n} \cdot \left(\delta_{t_{k,n}} - \mathcal{F} \right) \right\|_{\ell^b}$$

$$\leq \sum_{k=1}^{m_n} \lambda_{k,n} \cdot \left\| \delta_{t_{k,n}} - \mathcal{F} \right\|_{\ell^b} < \sum_{k=1}^{m_n} \lambda_{k,n} \cdot r = r \; .$$

Nach Voraussetzung ist $\lim\limits_{n \to +\infty} \| \epsilon_n - \mathcal{G} \|_{\ell^b} = 0$.

Daher gibt es zu jedem $\delta > 0$ ein $n_\delta \in \mathbb{N}$, so dass gilt:

$$\| \mathcal{G} - \mathcal{F} \|_{\ell^b} = \| (\epsilon_n - \mathcal{G}) - (\epsilon_n - \mathcal{F}) \|_{\ell^b} \leq \| \epsilon_n - \mathcal{G} \|_{\ell^b} + \| \epsilon_n - \mathcal{F} \|_{\ell^b}$$

$$< \delta + r \quad \text{für alle } n > n_\delta \; .$$

Weil $\delta > 0$ beliebig ist, folgt: $\| \mathcal{G} - \mathcal{F} \|_{\ell^b} < r$, also $(\mathcal{G} - \mathcal{F}) \in M_r$.

Die Definition der Stetigkeit einer Abbildung zwischen metrischen Räumen (H.Heuser 2000 II, S. 231) läßt sich auf den Fall erweitern, dass der Definitionsbereich „nur“ mit einer Pseudo-

Metrik ausgestattet ist.

Wegen $T'_{\mathcal{F}} \in \ell^b(M_r)$ gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass gilt:

$$\left\| \epsilon_n - \mathcal{G} \right\|_{\ell^b} = \left\| (\epsilon_n - \mathcal{F}) - (\mathcal{G} - \mathcal{F}) \right\|_{\ell^b} < \delta$$

$$\Rightarrow \ | T'_{\mathcal{F}}(\epsilon_n - \mathcal{F}) - T'_{\mathcal{F}}(\mathcal{G} - \mathcal{F}) | < \epsilon .$$

Daraus folgt $\lim\limits_{n \to +\infty} T'_{\mathcal{F}}(\epsilon_n - \mathcal{F}) = T'_{\mathcal{F}}(\mathcal{G} - \mathcal{F})$, und damit

$$\int_{\mathbb{R}} f_{\mathcal{F}} \, d(\mathcal{G} - \mathcal{F}) = T'_{\mathcal{F}}(\mathcal{G} - \mathcal{F}) .$$

❑

5.4.4 Bemerkungen zu den behandelten Fällen

(a) Damit das im Lemma III.2 angegebene $f_{\mathcal{F}}$ im obigen Grenzwertsatz verwendbar ist, muß natürlich zusätzlich $S^2(f_{\mathcal{F}}) := \int_R f_{\mathcal{F}}^2 \, d\mathcal{F} - \left(\int_R f_{\mathcal{F}} \, d\mathcal{F} \right)^2$ existieren.

Daher muß vorausgesetzt werden:

$h_{\mathcal{F},T} \in \mathcal{L}_2(A_r, \mathcal{B} \cap A_r, \mathcal{F})$, $g_{\mathcal{F},T} \in \mathcal{L}_2(\mathbb{R} \setminus A_r, \mathcal{B} \cap \{\mathbb{R} \setminus A_r\}, \mathcal{F})$.

(b) Angenommen, es gibt ein $r > 0$, so dass gilt:

$h_{\mathcal{F},T} \in \mathcal{L}_2(A_r, \mathcal{B} \cap A_r, \mathcal{F})$, $T''_{\mathcal{F}}$ existiert auf M_r und ist dort beschränkt.

Dann erfüllt T die Voraussetzungen von Satz III.1 (F.Strobl 1995, S. 147);

insbesondere ist die Integral-Darstellung $T'_{\mathcal{F}}(\mathcal{G} - \mathcal{F}) = \int_{\mathbb{R}} f_{\mathcal{F}} \, d(\mathcal{G} - \mathcal{F})$ für jedes $\mathcal{G} \in \{ \mathbb{W} \cap \Pi_r \}$ gültig (setze z.B. $g_{\mathcal{F},T} \equiv 0$).

(c) Im Sonderfall $A_r = \mathbb{R}$ (also $r = +oo$) stimmt $f_{\mathcal{F}}$ aus Lemma III.2 mit der bei R.J.Serfling (1980, S. 222) angegebenen Funktion $T_1(\mathcal{F}, \cdot)$ überein.

Im Falle $r = +oo$ gilt für jedes $x \in \mathbb{R}$:

$$f_{\mathcal{F}}(x) = T_1(\mathcal{F}, x) = T'_{\mathcal{F}}(\delta_x - \mathcal{F}) + \int_{\mathbb{R}} f_{\mathcal{F}} \, d\mathcal{F} .$$