Kapitel VII

Lokale Asymptotik

Reinhard Höpfner

Vorlesung Mathematische Statistik im Winter 2004/2005 und im Sommer 2005

Institut für Mathematik, Fachbereich 08, Johannes Gutenberg Universität Mainz

12.05.05

Übersicht zu Kapitel VII :

A. Lokal asymptotische Normalität / gemischte Normalität / Quadratizität

B. Le Cam's Ein-Schritt-Modifikation von Schätzern

A. Lokal asymptotische Normalität / gemischte Normalität / Quadratizität

7.1 Definition: Sei $\Theta \subset I\!R^d$ offen, betrachte eine Folge statistischer Experimente

$$\mathcal{E}_n \; := \; \left(\Omega_n \, , \, \mathcal{A}_n \, , \, \{P_{n,\vartheta} : \vartheta \in \Theta\}\right) \; , \quad n \geq 1 \; ,$$

und fixiere einen Referenzpunkt $\vartheta \in \Theta$.

a) $(\mathcal{E}_n)_n$ heißt *lokal asymptotisch quadratisch (LAQ) an der Stelle* ϑ falls eine Folge

$$\delta_n = \delta_n(\vartheta) \downarrow 0 \; , \quad n \to \infty$$

und Paare von Statistiken

$$(S_n, J_n) \; = \; (S_n(\vartheta), J_n(\vartheta)) \; : \; \Omega_n \longrightarrow I\!R^d \times D \quad \mathcal{A}_n\text{-meßbar} \; , \quad n \geq 1$$

existieren ($D \in \mathcal{B}(I\!R^{d\times d})$ bezeichnet die Menge der symmetrischen und strikt positiv definiten $d\times d$-Matrizen, cf. 6.1), so daß die folgenden Aussagen i) und ii) erfüllt sind:

i) für in $I\!R^d$ beschränkte Folgen $(h_n)_n$ gilt

$$\log \frac{d\,P_{n,\vartheta+\delta_n h_n}}{d\,P_{n,\vartheta}} \quad = \quad h_n^\top S_n \; - \; \frac{1}{2} h_n^\top J_n \, h_n \; + \; o_{(P_{n,\vartheta})_n}(1) \; , \quad n \to \infty \; ;$$

ii) es gibt ein quadratisches Experiment

$$\mathcal{E}_\infty \; = \; \mathcal{E}(S,J) \; = \; \left(\Omega \, , \, \mathcal{A} \, , \, \{P_h : h \in I\!R^d\}\right) \; , \quad \mathcal{E}_\infty = \mathcal{E}_\infty(\vartheta) \; , \; (S,J) = (S(\vartheta), J(\vartheta))$$

wie in 6.1 so daß gilt

$$\mathcal{L}\left(S_n, J_n \mid P_{n,\vartheta}\right) \; \longrightarrow \; \mathcal{L}\left(S, J \mid P_0\right) \quad \text{(schwache Konvergenz in } I\!R^d \times I\!R^{d\times d}\text{, für } n \to \infty) \; .$$

b) Ist $\mathcal{E}_\infty(\vartheta)$ in a)ii) gemischt normal (vgl. 6.2), so heißt die Folge von Experimenten $(\mathcal{E}_n)_n$ *lokal asymptotisch gemischt normal (LAMN) an der Stelle* ϑ.

b) Ist $\mathcal{E}_\infty(\vartheta)$ in a)ii) ein Gauß-Shift Experiment, d.h. ist $J(\vartheta)$ in a)ii) deterministisch (vgl. 5.2), so heißt die Folge $(\mathcal{E}_n)_n$ *lokal asymptotisch normal (LAN) an der Stelle* ϑ. In diesem Fall können alle J_n, $n \geq 1$, in a) durch das deterministische $J = J(\vartheta)$ ersetzt werden.

7.1' Bezeichnungen: a) Wir schreiben $L^{h/0}_{n,\vartheta} := \frac{d\,P_{n,\vartheta+\delta_n h}}{d\,P_{n,\vartheta}}$ und $\Lambda^{h/0}_{n,\vartheta} = \log L^{h/0}_{n,\vartheta}$ in $\mathcal{E}_n$, $n \geq 1$;

$$\left(\Omega_n, \mathcal{A}_n, \left\{P_{n,\vartheta+\delta_n(\vartheta)\,h} : h \in \Theta_{\vartheta,n}\right\}\right) \quad , \quad \Theta_{\vartheta,n} \; := \; \{h \in I\!R^d : h \text{ so daß } \vartheta + \delta(\vartheta)\,h \in \Theta\}$$

heißt *lokales Modell an der Stelle* ϑ, mit *lokalem Parameter* $h \in \Theta_{\vartheta,n}$ und *lokalem Maßstab* $\delta_n(\vartheta)$.

b) $\mathcal{E}_\infty = \mathcal{E}_\infty(\vartheta)$ in a)ii) heißt *Limesexperiment an der Stelle* ϑ.

Wir schreiben $L^{h/0} := \frac{dP_h}{dP_0}$ und $\Lambda^{h/0} = \log L^{h/0}$ in $\mathcal{E}_\infty$.

c) Die Folge $(Z_n)_n$

$$Z_n = Z_n(\vartheta) \ , \quad Z_n \ := \ J_n^{-1} S_n \ , \quad n \geq 1$$

heißt *zentrale Folge in* $(\mathcal{E}_n)_n$ *an der Stelle* ϑ.

Definition 7.1 setzt lokale Eigenschaften des Modells $\mathcal{E}_n$ in schrumpfenden δ_n-Umgebungen eines festen Referenzpunktes $\vartheta \in \Theta$ in Beziehung zu den Eigenschaften eines quadratischen Limesmodells $\mathcal{E}_\infty = \mathcal{E}(S, J)$. Die Parametrisierung des Limesmodells entspricht dem lokalen Parameter h im lokalen Modell an der Stelle ϑ. Der lokale Maßstab $\delta_n(\vartheta)$ kann in Abhängigkeit von dem betrachteten Fußpunkt ϑ der lokalen Asymptotik variieren. Hauptreferenzen für dieses Kapitel sind Le Cam (1968), Davies (1985), Jeganathan (1988/1995), Le Cam und Yang (1990).

7.1" Beispiel: Hinreichende Bedingungen für LAN oder LAMN oder LAQ präsentiert man häufig unter dem Namen eines 'Zweiten Le Camschen Lemmas'. Insbesondere hat Le Cam's Zweites Lemma 4.11 gezeigt:

bei n-facher Versuchswiederholung, ausgehend von einem an jeder Stelle $\vartheta \in \Theta$ L^2-differenzierbaren Experiment $\mathcal{E}$, gilt lokal asymptotische Normalität in ϑ mit lokalem Maßstab $\delta_n = 1/\sqrt{n}$, unabhängig von ϑ. Als Limesexperiment für die lokalen Modelle an der Stelle ϑ erhält man Gauß-Shift Experimente $\mathcal{E}_\infty(\vartheta) = \mathcal{E}(J_\vartheta)$ in der Schreibweise aus 5.2, bestimmt durch die Fisher-Information J_ϑ im Einzelexperiment $\mathcal{E}$ an der Stelle ϑ.

7.2 Hilfssatz: Unter LAQ an der Stelle ϑ wie in 7.1 gilt für konvergente Folgen $h_n \to h$

$$\begin{array}{lcl} \Lambda_{n,\vartheta}^{h_n/0} \ - \ \Lambda_{n,\vartheta}^{h/0} & \longrightarrow & 0 \ , \quad n \to \infty \ , \quad (P_{n,\vartheta})_n\text{-stochastisch,} \\ L_{n,\vartheta}^{h_n/0} \ - \ L_{n,\vartheta}^{h/0} & \longrightarrow & 0 \ , \quad n \to \infty \ , \quad (P_{n,\vartheta})_n\text{-stochastisch.} \end{array}$$

Beweis: 1) Wegen Straffheit der Paare $(S_n, J_n)_n$ unter $(P_{n,\vartheta})_n$ nach 7.1 a)ii) hat man für konvergente Folgen $h_n \to h$

$$(+) \qquad \left(h_n^\top S_n \ - \ \frac{1}{2} h_n^\top J_n \, h_n \right) - \left(h^\top S_n \ - \ \frac{1}{2} h^\top J_n \, h \right) \ = \ o_{(P_{n,\vartheta})_n}(1)$$

und für beschränkte Folgen $(h_n)_n$

$$(++) \qquad \mathcal{L}\left(h_n^\top S_n \;-\; \frac{1}{2}\, h_n^\top J_n\, h_n \mid P_{n,\vartheta}\right) \;,\; n \geq 1 \;, \quad \text{ist straff in } I\!R.$$

2) Aus (+) und der Darstellung der log-Likelihoods in 7.1 a)i) ergibt sich sofort die erste Behauptung des Hilfssatzes.

3) Für beschränkte Folgen $(h_n)_n$ ist $\left(\Lambda_{n,\vartheta}^{h_n/0}\right)_n$ eine Folge $\overline{I\!R}$-wertiger ZV. Aus (++) und der Darstellung der log-Likelihoods in 7.1 a)i) ergibt sich im Sinne von 3.1 b)

$$(*) \qquad \left(\Lambda_{n,\vartheta}^{h_n/0}\right)_n \quad \text{ist } I\!R\text{-straff unter } \left(P_{n,\vartheta}\right)_n.$$

4) In Einschränkung auf ein beliebig großes Kompaktum $K \subset I\!R^d$ ist die Funktion $\exp(\cdot)$ gleichmäßig stetig. Wegen der Straffheitsaussagen (++) und $(*)$ ergibt sich für konvergente Folgen $h_n \to h$ nun die zweite Behauptung des Hilfssatzes aus der ersten. □

7.3 Satz: Unter LAQ an der Stelle ϑ wie in 7.1 gilt wechselseitige Benachbartheit

$$(\diamond) \qquad \left(P_{n,\vartheta+\delta_n(\vartheta)h_n}\right)_n \;\lhd\rhd\; \left(P_{n,\vartheta}\right)_n$$

für beschränkte Folgen $(h_n)_n$, und schwache Konvergenz

$$\mathcal{L}\left(\Lambda_{n,\vartheta}^{h_n/0}, S_n, J_n \mid P_{n,\vartheta+\delta_n(\vartheta)h_n}\right) \;\longrightarrow\; \mathcal{L}\left(\Lambda^{h/0}, S, J \mid P_h\right) \quad (\text{in } I\!R \times I\!R^d \times I\!R^{d\times d}, \text{ für } n \to \infty)$$

für konvergente Folgen $h_n \to h$.

Beweis: 1) Nach Auswahl von Teilfolgen reicht es, die Behauptung $(\diamond)$ für konvergente Folgen nachzuweisen. Für $h_n \to h$ gilt nach 7.2

$$\mathcal{L}\left(\Lambda_{n,\vartheta}^{h_n/0} \mid P_{n,\vartheta}\right) \;\longrightarrow\; \mathcal{L}\left(h^\top S \;-\; \frac{1}{2}\, h^\top J\, h \mid P_0\right) \quad (\text{schwach in } I\!R, \text{ für } n \to \infty)$$

(beachte 3.1 b)ii)). Rechts erscheint die Verteilung der log-Likelihoodratio $\Lambda^{h/0}$ unter P_0 im Limesexperiment $\mathcal{E}(S,J)$. Nach Le Cam's Erstem Lemma 3.5 folgt wechselseitige Benachbartheit $(\diamond)$ für konvergente Folgen $h_n \to h$. Damit ist die erste Behauptung des Satzes bewiesen.

2) Schwache Konvergenz der Paare

$$\mathcal{L}\left(S_n, J_n \mid P_{n,\vartheta}\right) \;\longrightarrow\; \mathcal{L}\left(S, J \mid P_0\right) \quad (\text{in } I\!R^d \times I\!R^{d\times d}, \text{ für } n \to \infty)$$

nach 7.1 impliziert nach dem Satz über stetige Abbildungen und nach Darstellung der log-Likelihoodratios in 7.1

$$\mathcal{L}\left(\Lambda_{n,\vartheta}^{h/0}, S_n, J_n \mid P_{n,\vartheta}\right) \;\longrightarrow\; \mathcal{L}\left(\Lambda^{h/0}, S, J \mid P_0\right) \quad (\text{schwach in } I\!R \times I\!R^d \times I\!R^{d\times d}, \text{ für } n \to \infty)$$

für jedes $h \in I\!R^d$, woraus für konvergente Folgen $h_n \to h$ nach 7.2 folgt

$$\mathcal{L}\left(\Lambda_{n,\vartheta}^{h_n/0}, S_n, J_n \mid P_{n,\vartheta}\right) \;\longrightarrow\; \mathcal{L}\left(\Lambda^{h/0}, S, J \mid P_0\right) \quad (\text{schwach in } I\!R \times I\!R^d \times I\!R^{d\times d}, \text{ für } n \to \infty)$$

(beachte wieder 3.1 b)ii)). Wechselseitige Benachbartheit ($\diamond$) und Le Cam's Drittes Lemma 3.6 transformieren die letzte Aussage in

$$\mathcal{L}\left(\Lambda_{n,\vartheta}^{h_n/0}, S_n, J_n \mid P_{n,\vartheta+\delta_n h_n}\right) \;\longrightarrow\; \mathcal{L}\left(\Lambda^{h/0}, S, J \mid P_h\right)$$

(schwach in $I\!R \times I\!R^d \times I\!R^{d\times d}$, für $n \to \infty$); dies ist die zweite Behauptung des Satzes. □

7.4 Bemerkung: Die Grundidee für die folgenden Resultate in diesem Kapitel über asymptotisch beste Schätzfolgen für den unbekannten Parameter in der Folge von Experimenten $(\mathcal{E}_n)_n$ aus 7.1 ist einfach. Fixiere $\vartheta \in \Theta$. Es gelte LAQ in ϑ; sei T_n ein Schätzer für den unbekannten Parameter in $\mathcal{E}_n$, $n \geq 1$. Schreibe

$$U_n \;=\; U_n(\vartheta) \;:=\; \delta_n^{-1}(\vartheta)\,(T_n - \vartheta)$$

für den reskalierten Schätzfehler an der Stelle ϑ, dann liefert

$$U_n - h \;=\; \delta_n^{-1}\,(T_n - (\vartheta + \delta_n h)) \;, \quad h \in \Theta_{\vartheta,n}$$

die reskalierten Schätzfehler im lokalen Modell $\{P_{n,\vartheta+\delta_n h} : h \in \Theta_{\vartheta,n}\}$ an der Stelle ϑ.
Kann man der Folge $(U_n)_n$ in einem hinreichend starken Sinn ein 'Limesobjekt' U im Limesmodell $\mathcal{E}_\infty = \mathcal{E}(S,J)$ an der Stelle ϑ zuordnen, so wird man zuerst (mit den entsprechenden Sätzen der Kapitel V und VI) die Eigenschaften von U als Schätzer für den unbekannten Parameter $h \in I\!R^d$ im Limesexperiment studieren, und dann asymptotische Eigenschaften der Folge $(U_n)_n$ für $n \to \infty$ aus den Eigenschaften von U herzuleiten suchen.

Technisch beruht dieser Ansatz auf dem folgenden Resultat.

7.5 Satz: Es gelte LAQ an der Stelle ϑ wie in 7.1.
Für jede Folge $\mathcal{A}_n$-meßbarer Abbildungen $U_n = U_n(\vartheta) : \Omega_n \to I\!R^k$ mit der Eigenschaft

$$\mathcal{L}\left(U_n \mid P_{n,\vartheta}\right) \;, \quad n \geq 1 \;, \quad \text{ist straff in } I\!R^k$$

gibt es eine Teilfolge $(n_l)_l$ und eine $\mathcal{A}$-meßbare Abbildung $U : \Omega \to I\!R^k$ im (gegebenenfalls Markov- erweiterten) Limesexperiment $\mathcal{E}_\infty = \mathcal{E}(S,J)$ so daß für jede konvergente Folge $h_l \to h$

$$\mathcal{L}\left(S_{n_l}, J_{n_l}, U_{n_l} \mid P_{n_l,\vartheta+\delta_{n_l}h_l}\right) \;\longrightarrow\; \mathcal{L}\left(S, J, U \mid P_h\right) \;, \quad l \to \infty$$

(schwache Konvergenz in $I\!R^d \times I\!R^{d\times d} \times I\!R^k$).

Beweis: 1) Wegen 7.1 ist $\mathcal{L}(S_n, J_n \mid P_{n,\vartheta})$, $n \geq 1$, insbesondere straff, also gilt wegen der Voraussetzung an $(U_n)_n$ auch

$$\mathcal{L}\left(S_n, J_n, U_n \mid P_{n,\vartheta}\right) \ , \quad n \geq 1 \ , \quad \text{ist straff in } I\!R^d \times I\!R^{d\times d} \times I\!R^k.$$

Also gibt es eine Teilfolge $(n_l)_l$ und ein Wahrscheinlichkeitsmaß μ_0 auf $I\!R^d \times I\!R^{d\times d} \times I\!R^k$ so daß

$$\mathcal{L}\left(S_{n_l}, J_{n_l}, U_{n_l} \mid P_{n,\vartheta}\right) \longrightarrow \mu_0 \ , \quad l \to \infty \ .$$

Für beliebige konvergente Folgen $h_l \to h$ folgt daraus wie im Beweisschritt 2) von 7.3

$$(+) \qquad \mathcal{L}\left(\Lambda^{h_l/0}_{n_l,\vartheta}, S_{n_l}, J_{n_l}, U_{n_l} \mid P_{n,\vartheta}\right) \longrightarrow \widetilde{\mu}_0 \ , \quad l \to \infty$$

wobei das Wahrscheinlichkeitsmaß $\widetilde{\mu}_0$ auf $I\!R \times I\!R^d \times I\!R^{d\times d} \times I\!R^k$ aus μ_0 als Bildmaß unter

$$(s, j, u) \longrightarrow \left(h^\top s - \frac{1}{2} h^\top j \, h \, , \, s, j, u\right)$$

entsteht. Wechselseitige Benachbartheit ($\diamond$) in 7.3 und Le Cam's Drittes Lemma 3.6 transformieren (+) in schwache Konvergenz unter $(P_{n_l,\vartheta+\delta_{n_l}h_l})_l$

$$\mathcal{L}\left(\Lambda^{h_l/0}_{n_l,\vartheta}, S_{n_l}, J_{n_l}, U_{n_l} \mid P_{n_l,\vartheta+\delta_{n_l}h_l}\right) \longrightarrow: \ \widetilde{\mu}_h \ , \quad l \to \infty$$

mit einer Grenzverteilung $\widetilde{\mu}_h$, die durch Le Cam's Drittes Lemma 3.6 explizit bekannt ist:

$$\widetilde{\mu}_h(d\lambda, ds, dj, du) \ = \ e^\lambda \, \widetilde{\mu}_0(d\lambda, ds, dj, du) \quad \text{auf } I\!R \times I\!R^d \times I\!R^{d\times d} \times I\!R^k.$$

In Projektion auf die Komponenten (s, j, u) heißt das für $h_l \to h$

$$(++) \qquad \mathcal{L}\left(S_{n_l}, J_{n_l}, U_{n_l} \mid P_{n_l,\vartheta+\delta_{n_l}h_l}\right) \longrightarrow \mu_h \ , \quad l \to \infty$$

$$(+++) \qquad \mu_h(ds, dj, du) \ := \ e^{h^\top s - \frac{1}{2} h^\top j \, h} \, \mu_0(ds, dj, du) \quad \text{auf } I\!R^d \times I\!R^{d\times d} \times I\!R^k.$$

2) Nun sucht man eine Statistik U im Limesexperiment $\mathcal{E}(S, J)$, die es erlaubt, μ_h in (++) als $\mathcal{L}(S, J, U \mid P_h)$ zu identifizieren, $h \in I\!R^d$. Dazu muß man im allgemeinen zu einer Markov-Erweiterung von $\mathcal{E}(S, J)$ übergehen.

Schreibe $\check{s}$, $\check{j}$, $\check{u}$ für die Koordinatenprojektionen im Produktraum $I\!R^d \times I\!R^{d\times d} \times I\!R^k$. In dem durch (++) und (+++) entstandenen statistischen Modell $\{\mu_h : h \in I\!R^d\}$ ist $(\check{s}, \check{j})$ eine suffiziente Statistik, also können bedingte Verteilungen

$$(*) \qquad \mu_\bullet^{\check{u} | (\check{s},\check{j})=(s,j)}(du) \ =: \ K((s,j), du)$$

unabhängig vom Parameter $h \in I\!R^d$ festgelegt werden; $K(\cdot,\cdot)$ ist eine Übergangswahrscheinlichkeit von $I\!R^d \times I\!R^{d\times d}$ nach $I\!R^k$. Nun setzt man

$$\widehat{\Omega} := \Omega \times I\!R^k \ , \quad \widehat{\mathcal{A}} := \mathcal{A} \otimes \mathcal{B}(I\!R^k) \ , \quad \widehat{P}_h(d\omega, du) := P_h(d\omega) K((S,J)(\omega), du) \ ,$$

liftet die im Limesexperiment $\mathcal{E}(S,J)$ bereits vorhandenen Statistiken (S,J) auf $(\widehat{\Omega}, \widehat{\mathcal{A}})$ durch $S(\omega,u) := S(\omega)$, $J(\omega,u) := J(\omega)$, und schreibt U für die Zufallsvariable $U(\omega,u) := u$ auf $(\widehat{\Omega}, \widehat{\mathcal{A}})$. Damit ist ein Experiment

$$\left(\widehat{\Omega} \, , \, \widehat{\mathcal{A}} \, , \, \left\{ \widehat{P}_h : h \in I\!R^d \right\} \right)$$

gefunden, welches sich *als statistisches Experiment nicht von* $\mathcal{E}(S,J)$ *unterscheidet*, denn es gilt

$$(\diamond) \qquad \mathcal{L}\left(S, J \mid \widehat{P}_h \right) \ = \ \mathcal{L}\left(S, J \mid P_h \right) \quad \text{für alle } h \in I\!R^d \ ,$$

$$(\diamond\diamond) \qquad \frac{d\widehat{P}_h}{d\widehat{P}_0} \ = \ e^{h^\top S \,-\, \frac{1}{2} h^\top J\, h} \quad \text{auf } \widehat{\Omega}, \text{ für alle } h \in I\!R^d \ .$$

Dies ist die *Markov-Erweiterung* von $\mathcal{E}(S,J)$.

3) Auf der Markov-Erweiterung von $\mathcal{E}(S,J)$ gibt es die Statistik U, deren bedingte Verteilung gegeben (S,J) nach Konstruktion nicht vom Parameter h abhängt:

$$(**) \qquad \widehat{P}_\bullet^{U|(S,J)=(s,j)}(du) \ = \ K((s,j), du) \ .$$

Weil für konvergente Folgen $h_l \to h$ nach 7.3 und $(\diamond)$ in Schritt 2) gilt

$$\mathcal{L}\left(S_{n_l}, J_{n_l} \mid P_{n_l, \vartheta + \delta_{n_l} h_l} \right) \ \longrightarrow \ \mathcal{L}\left(S, J \mid \widehat{P}_h \right) \ , \quad l \to \infty$$

und weil man nach $(++)$ insbesondere hat

$$\mathcal{L}\left(S_{n_l}, J_{n_l} \mid P_{n_l, \vartheta + \delta_{n_l} h_l} \right) \ \longrightarrow \ \mu_h^{(\check{s}, \check{j})} \ , \quad h \in I\!R^d \ ,$$

ist gezeigt:

$$\mathcal{L}\left(S, J \mid \widehat{P}_h \right) \ = \ \mu_h^{(\check{s}, \check{j})} \ , \quad h \in I\!R^d \ .$$

Mit $(*)$ und $(**)$ erhält man daraus

$$\mu_h(ds, dj, du) \ = \ \mathcal{L}\left(S, J, U \mid \widehat{P}_h \right) \ , \quad h \in I\!R^d \ .$$

Dies zusammen mit $(++)$ schließt den Beweis des Satzes ab. $\square$

Der Vergleich zwischen den lokalen Experimenten an der Stelle ϑ mit dem Limesexperiment an der Stelle ϑ motiviert die folgende Definition, die der starken Äquivarianz von Schätzern (siehe 6.5) in einem gemischt normalen Experiment entspricht.

7.6 Definition: Es gelte LAMN in ϑ. Betrachte eine Folge $\kappa_n : (\Omega_n, \mathcal{A}_n) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ von Schätzern für den unbekannten Parameter in $\mathcal{E}_n$, $n \geq 1$.
Gibt es ein Wahrscheinlichkeitsmaß $\widetilde{F} = \widetilde{F}(\vartheta)$ auf $\mathbb{R}^d \times \mathbb{R}^{d \times d}$ so daß für jedes feste h

$$\mathcal{L}\left(\delta_n^{-1}\left(\kappa_n - (\vartheta + \delta_n h)\right), J_n \mid P_{n,\vartheta+\delta_n h}\right) \longrightarrow \widetilde{F}\,, \quad n \to \infty$$

(schwache Konvergenz in $\mathbb{R}^d \times \mathbb{R}^{d \times d}$) gilt, *wobei die Grenzverteilung $\widetilde{F}$ nicht vom Wert des lokalen Parameters $h \in \mathbb{R}^d$ abhängt*, so heißt die Folge $(\kappa_n)_n$ *regulär an der Stelle ϑ*.

Regularität bedeutet, daß eine Schätzfolge *asymptotisch an allen Punkten in schrumpfenden Umgebungen von ϑ* gleich gut arbeitet, wobei 'schrumpfende Umgebungen' im Sinn des lokalen Maßstabs $(\delta_n(\vartheta))_n$ zu verstehen ist. Dem Faltungssatz 6.6 im gemischt normalen Limesexperiment entspricht in der lokalen Asymptotik das folgende Resultat.

7.7 Faltungssatz: (Jeganathan 1982) Es gelte LAMN in ϑ wie in 7.1. Für eine an der Stelle ϑ reguläre Schätzfolge $(\kappa_n)_n$ hat die Grenzverteilung $\widetilde{F}$ in 7.6 die Gestalt

$$\widetilde{F}(A) \;=\; \int\!\!\int P_\bullet^J(dj)\ \left[\mathcal{N}(0, j^{-1}) \star Q^j\right](du)\ 1_A(u,j)\,, \quad A \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R}^{d \times d})$$

für eine Familie von Störverteilungen $\{Q^j : j \in \mathbb{R}^{d \times d}\}$ wie in 6.6.

Beweis: Schreibt man $U_n := \delta_n^{-1}(\kappa_n - \vartheta)$, dann liefert die Regularitätsvoraussetzung

$$\mathcal{L}\left(U_n - h, J_n \mid P_{n,\vartheta+\delta_n h}\right) \longrightarrow \widetilde{F}\,, \quad n \to \infty$$

(schwache Konvergenz in $\mathbb{R}^d \times \mathbb{R}^{d \times d}$), mit $\widetilde{F}$ unabhängig vom Wert des lokalen Parameters h. Durch Auswahl von Teilfolgen nach 7.5 bekommt man für die Grenzverteilung $\widetilde{F}$ die Darstellung

$$(+) \qquad \widetilde{F} \;=\; \mathcal{L}\left(U - h, J \mid P_h\right) \quad \text{unabhängig von } h \in \mathbb{R}^d$$

mit einer Statistik U im (gegebenenfalls Markov-erweiterten) Limesexperiment $\mathcal{E}(S, J)$.
U ist wegen (+) ein stark äquivarianter Schätzer für den unbekannten Parameter $h \in \mathbb{R}^d$ im Limesexperiment $\mathcal{E}(S, J)$. Dieses ist nach Voraussetzung gemischt normal, und die Aussage des

Satzes 7.7 folgt aus dem Faltungssatz 6.6 im Limesexperiment. □

Im Spezialfall lokal asymptotischer Normalität ist 7.7 der berühmte Faltungssatz von Hájek. Die Formulierung (insbesondere der Regularitätsbedingung) vereinfacht sich hier, da in diesem Spezialfall alle J_n, $n \geq 1$ in 7.1 durch die deterministische Matrix $J = J(\vartheta)$ ersetzt werden dürfen:

7.8 Faltungssatz: (Hájek 1970) Es gelte LAN in ϑ. Für jede an der Stelle ϑ reguläre Folge $(\kappa_n)_n$ von Schätzern für den unbekannten Parameter in $(\mathcal{E}_n)_n$, d.h.

$$\mathcal{L}\left(\delta_n^{-1}\left(\kappa_n - (\vartheta + \delta_n h)\right) \mid P_{n,\vartheta+\delta_n h}\right) \longrightarrow F\ , \quad n \to \infty$$

(schwache Konvergenz in $\mathbb{R}^d$), wobei die Grenzverteilung F nicht vom Wert des lokalen Parameters $h \in \mathbb{R}^d$ abhängt, gibt es eine Darstellung

$$F \ = \ \mathcal{N}\left(0, J^{-1}\right) \star Q$$

mit einem geeigneten Wahrscheinlichkeitsmaß Q auf $\mathbb{R}^d$.

Nachden Sätzen aus den Kapiteln V und VI ist nun klar, daß unter LAMN/LAN in ϑ eine optimale Konzentration der Grenzverteilungen in 7.7 und 7.8 für in ϑ reguläre Schätzfolgen genau dann erreicht wird, wenn die Störverteilungen Q^j, $j \in \mathbb{R}^{d\times d}$, Q verschwinden. Das ist insbesondere dann der Fall, wenn die Statistik U im Limesexperiment $\mathcal{E}(S,J)$, die in 7.5 und im Beweis von 7.7 entsteht, mit der zentralen Statistik $Z = J^{-1}S$, $Z = Z(\vartheta)$, zusammenfällt.

7.9 Definition: Es gelte LAMN in ϑ. Eine an der Stelle ϑ reguläre Schätzfolge $(\kappa_n)_n$ für den unbekannten Parameter in $(\mathcal{E}_n)_n$ heißt *effizient in* ϑ wenn gilt:

$$\text{die Grenzverteilung } \widetilde{F} \text{ in 7.6 ist gegeben durch } \widetilde{F} = \mathcal{L}\left(Z, J \mid P_0\right)\ .$$

Unter LAN in ϑ vereinfacht sich diese Bedingung zu $F = \mathcal{L}\left(Z \mid P_0\right)$ in 7.8.

Bevor wir die effizienten Schätzfolgen unter LAMN charakterisieren, unterbrechen wir die Optimalitätsbetrachtungen (diese werden in 7.12 unten und im lokal asymptotischen Minimaxsatz 7.13 wiederaufgenommen) und betrachten den allgemeinen LAQ-Fall.

7.10 Satz: Unter LAQ in ϑ betrachte eine Schätzfolge $(\eta_n)_n$ in $(\mathcal{E}_n)_n$ mit der Eigenschaft

$$(*) \qquad \delta_n^{-1}(\vartheta)(\eta_n - \vartheta) \;=\; Z_n(\vartheta) \;+\; o_{(P_{n,\vartheta})_n}(1)\,, \quad n \to \infty\,.$$

a) Unter $(*)$ gilt für beliebige konvergente Folgen $h_n \to h$:

$$\mathcal{L}\left(\,\delta_n^{-1}\,(\eta_n - (\vartheta + \delta_n h_n))\,, J_n \mid P_{n,\vartheta+\delta_n h_n}\right) \;\longrightarrow\; \mathcal{L}\,(Z - h, J \mid P_h)\;, \quad n \to \infty$$

(schwache Konvergenz in $\mathbb{R}^d \times \mathbb{R}^{d\times d}$).

b) Unter $(*)$ gilt für subkonvexe Verlustfunktionen $\ell(\cdot)$ in $\mathcal{C}_b(\mathbb{R}^d)$ und $C < \infty$

$$\sup_{|h|\le C} \left| E_{n,\vartheta+\delta_n h}\left(\,\ell\left(\delta_n^{-1}\,(\eta_n - (\vartheta + \delta_n h))\right)\right) \;-\; E_h\,(\ell(Z-h))\right| \quad\longrightarrow\quad 0\,, \quad n \to \infty\,.$$

Beweis: 1) Nach Satz 7.3 gilt unter LAQ in ϑ für konvergente Folgen $h_n \to h$

$$\mathcal{L}\left(S_n, J_n \mid P_{n,\vartheta+\delta_n h_n}\right) \;\longrightarrow\; \mathcal{L}\,(S, J \mid P_h)$$

woraus sich mit dem Satz über stetige Abbildungen und wegen $h_n \to h$ sofort ergibt

$$\mathcal{L}\left(Z_n - h_n, J_n \mid P_{n,\vartheta+\delta_n h_n}\right) \;\longrightarrow\; \mathcal{L}\,(Z - h, J \mid P_h)\;, \quad n \to \infty\,.$$

Setze $U_n := \delta_n^{-1}(\eta_n - \vartheta)$. Nach Satz 7.3 gilt Benachbartheit $\left(P_{n,\vartheta+\delta_n^{-1} h_n}\right)_n \;\lhd\rhd\; \left(P_{n,\vartheta}\right)_n$, daher liefert $(*)$ auch

$$U_n \;=\; Z_n \;+\; o_{(P_{n,\vartheta+\delta_n h_n})_n}(1)\,, \quad n \to \infty\,.$$

Aus den letzten beiden Formelzeilen zusammen folgt

$$\mathcal{L}\left(U_n - h_n, J_n \mid P_{n,\vartheta+\delta_n h_n}\right) \;\longrightarrow\; \mathcal{L}\,(Z - h, J \mid P_h)\;, \quad n \to \infty\,.$$

Damit ist a) bewiesen.

2) Im Limesexperiment $\mathcal{E}_\infty(\vartheta) = \mathcal{E}(S,J)$ ist wegen 6.1 die Familie $\left\{L^{h_n/0}, n \ge 1, L^{h/0}\right\}$ gleichgradig integrierbar unter P_0, für konvergente Folgen $h_n \to h$.

Für festes $f \in \mathcal{C}_b(\mathbb{R}^d)$ impliziert dies für $n \to \infty$ die Aussage

$$(\circ) \quad E_{h_n}\,(f(Z - h_n)) \;=\; E_0\left(L^{h_n/0}\; f(Z - h_n)\right) \;\longrightarrow\; E_0\left(L^{h/0}\; f(Z - h)\right) \;=\; E_h\,(f(Z-h))\;;$$

zum Nachweis von $(\circ)$ benutzt zuerst die gleichgradige Integrierbarkeit, um sich auf

$$E_0\left(\phi(L^{h_n/0})\; f(Z - h_n)\right) \;\longrightarrow\; E_0\left(\phi(L^{h/0})\; f(Z-h)\right)$$

für geeignetes $\phi(\cdot) \in \mathcal{C}_b(I\!R^d)$, welches auf einem großen Kompaktum mit der Identität übereinstimmt, zurückzuziehen; die letzte Aussage gilt aber für $h_n \to h$ wegen dominierter Konvergenz.
3) Betrachtet man eine subkonvexe Verlustfunktion $\ell(\cdot) \in \mathcal{C}_b(I\!R^d)$, so impliziert Schritt 1)

$$E_{n,\vartheta+\delta_n h_n}\left(\,\ell\left(U_n - h_n\right)\right) \;\longrightarrow\; E_h\left(\ell(Z-h)\right)\;,$$

was wegen ($\circ$) in Schritt 2) zu

$$(+) \qquad E_{n,\vartheta+\delta_n h_n}\left(\,\ell\left(U_n - h_n\right)\right) \;-\; E_{h_n}\left(\ell(Z-h_n)\right) \;=\; o(1)\;, \quad n \to \infty$$

umgeschrieben werden kann. (+) gilt für beliebige konvergente Folgen $h_n \to h$. Um aus (+) für beliebiges $C < \infty$ die Behauptung

$$(++) \qquad \sup_{|h|\le C} \left| E_{n,\vartheta+\delta_n h}\left(\,\ell\left(U_n - h\right)\right) \;-\; E_h\left(\ell(Z-h)\right)\right| \quad\longrightarrow\quad 0\;, \quad n \to \infty$$

herzuleiten, genügt es, beschränkte Folgen $(h_n)_n$ in der abgeschlossenen Kugel $\{|h| \le C\}$ so zu wählen, daß

$$E_{n,\vartheta+\delta_n h_n}\left(\,\ell\left(U_n - h_n\right)\right) \;-\; E_{h_n}\left(\ell(Z-h_n)\right)$$

das Supremum in (++) bis auf einen Fehler von höchstens 2^{-n} approximiert, aus $(h_n)_n$ konvergente Teilfolgen $h_{n_l} \to h$ auszuwählen, und (+) auf konvergente Folgen $(\widetilde{h}_m)_m$ der Art

$$\widetilde{h}_m \;:=\; h_{n_l} \quad \text{falls } \; n_{l-1} < m \le n_l,\;\; l \ge 1, \text{ mit } n_0 := 0$$

anzuwenden. Damit ist auch b) bewiesen. □

7.11 Bemerkung: a) Unter LAQ in ϑ arbeitet eine Schätzfolge $(\eta_n)_n$ in $(\mathcal{E}_n)_n$ mit der Eigenschaft ($*$) aus 7.10 *im lokalen Modell an der Stelle* ϑ *so gut wie der Maximum-Likelihood-Schätzer im Limesmodell*, genauer: die reskalierten Schätzfehler von η_n im lokalen Modell an der Stelle ϑ entsprechen den Schätzfehlern des Maximum-Likelihood Schätzers $Z = J^{-1}S$ im Limesexperiment $\mathcal{E}(S,J)$. Im allgemeinen LAQ-Fall hängt die Verteilung der Schätzfehler des Maximum-Likelihood Schätzers im Limesexperiment vom Parameter ab.
b) Die Aussage in a) ist kein wirklicher Effizienzbegriff für Schätzfolgen, wie man ihn im LAN- und LAMN-Fall dank der Faltungssätze 7.8 und 7.6 (und des Minimaxsatzes 7.13 unten) zur Verfügung hat, aber mehr scheint im allgemeinen LAQ-Fall (mit Ausnahme der Resultate von Gushchin 1996) nicht möglich zu sein.

Wegen der quadratischen Gestalt des Feldes von log-Likelihood Ratios im Limesmodell

$$\left(h^\top S(\omega) \;-\; \frac{1}{2}\, h^\top J(\omega)\, h \right)_{h \in \mathbb{R}^d} \quad , \quad \omega \in \Omega \text{ so daß } J(\omega) \in D$$

ist es zwar plausibel, daß die Maximalstelle $Z(\omega) = J^{-1}(\omega) S(\omega)$ dieses Feldes einen brauchbaren Schätzer für den unbekannten Parameter h abgeben sollte, aber Einwände wie in Kapitel I B können auf der Ebene des Limesexperiments nicht ausgeschlossen werden.

Nun beweisen wir die angekündigte Charakterisierung effizienter Schätzer unter LAMN:

7.12 Satz: Es gelte LAMN in ϑ. Für jede Folge $(\eta_n)_n$ von Schätzern für den unbekannten Parameter in $(\mathcal{E}_n)_n$ sind folgende Aussagen gleichwertig:

i) die Folge $(\eta_n)_n$ ist regulär und effizient in ϑ;

ii) es gilt $(*)$ aus 7.10:

$$\delta_n^{-1}(\vartheta)(\eta_n - \vartheta) \;=\; Z_n(\vartheta) \;+\; o_{(P_{n,\vartheta})_n}(1) \;, \quad n \to \infty \;.$$

Beweis: Setze $U_n := \delta_n^{-1}(\eta_n - \vartheta)$.

1) Wegen LAMN in ϑ liefert 7.10 für Folgen $(\eta_n)_n$ mit (*)

$$\mathcal{L}\left(U_n - h, J_n \mid P_{n,\vartheta+\delta_n h} \right) \;\longrightarrow\; \mathcal{L}\left(Z - h, J \mid P_h \right) \;=\; \mathcal{L}\left(Z, J \mid P_0 \right) \quad (\text{unabhängig von } h \in \mathbb{R}^d).$$

Das bedeutet Regularität und Effizienz von $(\eta_n)_n$ in ϑ. Also gilt ii)$\Longrightarrow$ i).

2) Sei $(\eta_n)_n$ regulär an der Stelle ϑ. Dann gilt für jedes h Verteilungskonvergenz

$$(\circ) \qquad \mathcal{L}\left(U_n - h, J_n \mid P_{\vartheta+\delta_n h} \right) \;\longrightarrow\; \widetilde{F} \;, \quad n \to \infty$$

mit einer von h unabhängigen Grenzverteilung $\widetilde{F}$.

Fixiere eine beliebige Teilfolge der natürlichen Zahlen. Wendet man Satz 7.5 entlang dieser Teilfolge an, so erhält man eine weitere Teilfolge $(n_l)_l$ und eine Statistik U im (eventuell Markoverweiterten) Limesexperiment $\mathcal{E}(S, J)$ an der Stelle ϑ so daß

$$\mathcal{L}\left(U_{n_l}, S_{n_l}, J_{n_l} \mid P_{\vartheta+\delta_{n_l} h} \right) \;\longrightarrow\; \mathcal{L}\left(U, S, J \mid P_h \right) \;, \quad l \to \infty$$

und damit auch

$$(\circ\circ) \qquad \mathcal{L}\left(U_{n_l} - h, Z_{n_l} - h, J_{n_l} \mid P_{\vartheta+\delta_{n_l} h} \right) \;\longrightarrow\; \mathcal{L}\left(U - h, Z - h, J \mid P_h \right) \;, \quad l \to \infty$$

für jedes feste h. Wegen Regularität von $(\eta_n)_n$ an der Stelle ϑ hängt dabei

$$\widetilde{F} \;=\; \mathcal{L}\,(U-h, J \mid P_h) \;=\; \mathcal{L}\,(U, J \mid P_0)$$

nicht von h ab, und der Faltungssatz 6.6 im gemischt normalen Limesexperiment, anzuwenden auf den stark äquivarianten Schätzer U, zeigt

$$\widetilde{F}(A) \;=\; \int P_\bullet^J(dj) \int [\mathcal{N}(0, j^{-1}) \star Q^j](du)\; 1_A(u,j)\;, \quad A \in \mathcal{B}(I\!R^d \times I\!R^{d\times d})\;.$$

Ist $(\eta_n)_n$ effizient an der Stelle ϑ, so müssen wegen $(\circ)$ und wegen Anderson's Lemma 5.7 die Störverteilungen verschwinden:

$$Q^j \;=\; \epsilon_0 \quad \text{für } P_\bullet^J\text{-fast alle } j \in I\!R^{d\times d}\;.$$

Im Beweisschritt 5) des Faltungssatzes 6.6 war aber gezeigt worden

$$(j, B) \;\longrightarrow\; Q^j(B) \quad \text{ist eine reguläre Version der bedingten Verteilung } P_0^{U-Z|J=j}(\cdot)\;.$$

Kombiniert ergeben die beiden letzten Aussagen

$$U \;=\; Z \quad P_0\text{-fast sicher.}$$

Damit folgt aus $(\circ\circ)$ wieder mit dem Satz über stetige Abbildungen

$$\mathcal{L}\left(U_{n_l} - Z_{n_l} \mid P_{n_l,\vartheta}\right) \;\longrightarrow\; \mathcal{L}\,(U - Z \mid P_0) \;=\; \epsilon_0\;, \quad l \to \infty\;.$$

Verteilungskonvergenz gegen eine Konstante ist aber gleichwertig mit stochastischer Konvergenz, also ist bewiesen

$$(+) \qquad U_{n_l} \;=\; Z_{n_l} \;+\; o_{(P_{n_l,\vartheta})_l}(1)\;, \quad l \to \infty\;.$$

Zusammen hat man gezeigt, daß aus jeder Teilfolge der natürlichen Zahlen eine weitere Teilfolge $(n_l)_l$ ausgewählt werden kann, welche die Eigenschaft $(+)$ besitzt. Das aber bedeutet

$$U_n \;=\; Z_n \;+\; o_{(P_{n,\vartheta})_n}(1)\;, \quad n \to \infty$$

und damit ii). Der Beweis von 7.12 ist nun vollständig. □

In Teilkapitel B wird – unter moderaten zusätzlichen Voraussetzungen – ein Verfahren beschrieben, wie eine auf Le Cam zurückgehende 'Ein-Schritt-Korrektur' jede gegebene Schätzfolge, die

an der Stelle ϑ die richtige Konvergenzgeschwindigkeit besitzt, in einfacher und expliziter Weise so abändern kann, daß eine (reguläre und) effiziente Folge entsteht.
Zum Abschluß dieses Teilkapitels formulieren wir den lokalasymptotischen Minimaxsatz. Er erlaubt, unter LAMN in ϑ *alle* Schätzfolgen $(\eta_n)_n$ für den unbekannten Parameter, sofern sie in ϑ die richtige Konvergenzgeschwindigkeit haben, zu vergleichen.

7.13 Minimaxsatz: Es gelte LAMN in ϑ. Sei $\ell(\cdot)$ eine stetige, beschränkte, subkonvexe Verlustfunktion.
a) Für jede Schätzfolge $(\eta_n)_n$ in $(\mathcal{E}_n)_n$, welche an der Stelle ϑ der Straffheitsbedingung

$$\mathcal{L}\left(\delta_n^{-1}(\eta_n - \vartheta) \mid P_{n,\vartheta}\right) \ , \quad n \geq 1 \ , \quad \text{ist straff}$$

genügt, gilt die *lokalasymptotische Minimaxschranke*:

$$\sup_{c<\infty} \liminf_{n\to\infty} \sup_{|h|\leq c} E_{n,\vartheta+\delta_n h}\left(\ell\left(\delta_n^{-1}\left(\eta_n - (\vartheta + \delta_n h)\right)\right)\right) \quad \geq \quad E_0\left(\ell(Z)\right) \ .$$

b) Schätzfolgen $(\eta_n)_n$, die die Bedingung $(*)$ aus 7.10

$$\delta^{-1}(\vartheta)(\eta_n - \vartheta) \ = \ Z_n(\vartheta) \ + \ o_{(P_{n,\vartheta})_n}(1) \ , \quad n \to \infty$$

erfüllen, erreichen die lokalasymptotische Minimaxschranke: für sie gilt

$$\lim_{n\to\infty} \sup_{|h|\leq c} E_{n,\vartheta+\delta_n h}\left(\ell\left(\delta_n^{-1}\left(\eta_n - (\vartheta + \delta_n h)\right)\right)\right) \quad = \quad E_0\left(\ell(Z)\right)$$

für jedes $0 < c < \infty$.

Beweis: 1) Schreibe $U_n := \delta^{-1}(\vartheta)(\eta_n - \vartheta)$. Sei $c \in I\!N$ fest. Da $\ell(\cdot)$ nichtnegativ und beschränkt, existiert

$$\liminf_{n\to\infty} \sup_{|h|\leq c} E_{n,\vartheta+\delta_n h}\left(\ell\left(U_n - h\right)\right) \ .$$

Wähle eine Teilfolge der natürlichen Zahlen, entlang der 'liminf' in der letzten Zeile durch 'lim' ersetzt werden kann, danach aus dieser mit 7.5 eine weitere Teilfolge $(n_l)_l$ und eine Statistik U im (gegebenenfalls Markov-erweiterten) Limesexperiment $\mathcal{E}(S,J)$ so daß gilt

$$\mathcal{L}\left(U_{n_l} - h_l \mid P_{n_l,\vartheta+\delta_{n_l}h_l}\right) \ \longrightarrow \ \mathcal{L}\left(U - h \mid P_h\right) \ , \quad l \to \infty$$

für jede konvergente Folge $h_l \to h$. Für $\ell \in \mathcal{C}_b(I\!R^d)$ und $h \in I\!R^d$ fest hat man insbesondere

$$(+) \qquad E_{n_l,\vartheta+\delta_{n_l}h}\left(\ell\left(U_{n_l} - h\right)\right) \ \longrightarrow \ E_h\left(\ell\left(U - h\right)\right) \ , \quad l \to \infty \ .$$

Sei wie im Beweis von 6.8 π_c die Gleichverteilung auf der abgeschlossenen Kugel in $I\!R^d$ mit Radius c. Dann gilt

$$\begin{aligned} \sup_{|h|\le c} E_{n_l,\vartheta+\delta_{n_l}h}\left(\,\ell\,(U_{n_l}-h)\right) \quad &\ge \quad \int \pi_c(dh)\; E_{n_l,\vartheta+\delta_{n_l}h}\left(\,\ell\,(U_{n_l}-h)\right) \\ &\longrightarrow \quad \int \pi_c(dh)\; E_h\left(\,\ell\,(U-h)\right) \quad , \quad l\to\infty \end{aligned}$$

wegen (+) und dominierter Konvergenz.

2) Mit dem letzten Ausdruck ist man aber im gemischt normalen Limesexperiment $\mathcal{E}(S,J)$, in dem jeder Schätzer unter sehr diffuser Vorbewertung asymptotisch 'fast äquivariant' ist, vgl. 6.7. In diesem Sinn kopiert man den Beweis des Minimaxsatzes 6.8 im Limesexperiment

$$\int \pi_c(dh)\; E_h\left(\,\ell\,(U-h)\right) \;=\; \int P_\bullet^J(dj) \int \left[\mathcal{N}(0,j^{-1})\star Q_c^j\right](du)\; l(u) \;+\; \varrho(c)$$

mit Resttermen wie im Beweis von 6.8 bzw. von 6.7 betrachtet:

$$\varrho(c)\;, \quad c\in I\!N\;, \quad \text{mit} \quad \lim_{c\to\infty} \varrho(c) \;=\; 0\;.$$

Mit 5.7 schätzt man wie im Beweis von 6.8 nach unten ab

$$\int \pi_c(dh)\; E_h\left(\,\ell\,(U-h)\right) \;\ge\; \int P_\bullet^J(dj) \int \mathcal{N}(0,j^{-1})(du)\; l(u) \;+\; \varrho(c)\;.$$

Das Wahrscheinlichkeitsmaß auf der rechten Seite ist $\mathcal{L}\left(Z \mid P_0\right)$.

3) Zusammen zeigen die Beweisschritte 1) und 2): für jedes feste c gilt

$$\begin{aligned} \liminf_{n\to\infty} \sup_{|h|\le c} E_{n,\vartheta+\delta_n h}\left(\,\ell\,(U_n-h)\right) \quad &= \quad \lim_{l\to\infty} \sup_{|h|\le c} E_{n_l,\vartheta+\delta_{n_l}h}\left(\,\ell\,(U_{n_l}-h)\right) \\ &\ge \quad E_0\left(\,\ell(Z)\right) \;+\; \varrho(c)\;, \end{aligned}$$

wobei die Restterme $\varrho(c)$ für $c\to\infty$ verschwinden. Das aber ergibt die Aussage a) des Satzes.

b) Für Schätzfolgen $(\eta_n)_n$, die die Bedingung $(*)$ aus 7.10 erfüllen, gilt nach 7.10 b) sogar

$$\sup_{|h|\le c} \left| E_{n,\vartheta+\delta_n h}\left(\,\ell\,(U_n-h)\right) \;-\; E_h\left(\,\ell(Z-h)\right)\right| \quad \longrightarrow \quad 0\;, \quad n\to\infty$$

für jedes feste c. Wegen Äquivarianz von Z im gemischt normalen Limesexperiment folgt sofort

$$\lim_{n\to\infty} \sup_{|h|\le c} E_{n,\vartheta+\delta_n h}\left(\,\ell\,(U_n-h)\right) \quad = \quad E_0\left(\,\ell(Z)\right)$$

für beliebiges c. Damit ist auch Aussage b) von Satz 7.13 bewiesen. □

B. Le Cam's Ein-Schritt-Modifikation von Schätzern

Ist eine Folge $(\mathcal{E}_n)_n$ statistischer Modelle mit Parameterraum Θ an jeder Stelle $\vartheta \in \Theta$ lokal asymptotisch quadratisch, so kann (unter milden Zusatzannahmen) jede gegebene Schätzfolge $(g_n)_n$ für den unbekannten Parameter, sofern sie die richtige Konvergenzgeschwindigkeit besitzt, zu einer Schätzfolge $(\eta_n)_n$ umgebaut werden, die an jeder Stelle $\vartheta \in \Theta$ die Eigenschaft $(*)$ aus 7.10 besitzt. Dieser Umbau durch eine 'Ein-Schritt-Korrektur' ist vollkommen explizit und sehr einfach; die hier gegebene Darstellung folgt Davies (1985).

7.14 Voraussetzungen für Teilkapitel B: Man betrachtet eine Folge von Experimenten

$$\mathcal{E}_n \;=\; (\Omega_n, \mathcal{A}_n, \{P_{n,\vartheta} : \vartheta \in \Theta\}) \;, \quad n \geq 1 \;, \quad \Theta \subset \mathbb{R}^d \text{ offen}$$

mit den folgenden Eigenschaften A) - D).

A) An jeder Stelle $\vartheta \in \Theta$ gilt LAQ wie in 7.1 a), mit Folgen

$$(\delta_n(\vartheta))_n \;, \quad (S_n(\vartheta))_n \;, \quad (J_n(\vartheta))_n$$

und Limesexperiment

$$\mathcal{E}_\infty(\vartheta) \;=\; \mathcal{E}\left(S(\vartheta), J(\vartheta)\right) \;.$$

B) Für jedes feste $n \geq 1$ ist die Funktion $\delta_n(\cdot) : \Theta \to (0, \infty)$ meßbar, o.E. beschränkt durch 1, und an jeder Stelle $\vartheta \in \Theta$ gilt

$$(*) \qquad \sup_{|h| \leq c} \left| \frac{\delta_n(\vartheta + \delta_n(\vartheta)\, h)}{\delta_n(\vartheta)} - 1 \right| \quad \longrightarrow \quad 0 \;, \quad n \to \infty$$

für beliebiges $c < \infty$.

C) Mit $S := \{\alpha 2^{-k} : k \in \mathbb{N}_0, \alpha \in \mathbb{Z}^d\}$ abzählbar dicht in $\mathbb{R}^d$ gilt an jeder Stelle $\vartheta \in \Theta$

i) $$\sup_{\xi \in S \cap \Theta \,,\, |\xi - \vartheta| \leq c\, \delta_n(\vartheta)} \left| S_n(\xi) - \left\{ S_n(\vartheta) - J_n(\vartheta) \left[\delta_n^{-1}(\vartheta)(\xi - \vartheta) \right] \right\} \right| \;=\; o_{(P_{n,\vartheta})_n}(1) \;, \quad n \to \infty$$

ii) $$\sup_{\xi \in S \cap \Theta \,,\, |\xi - \vartheta| \leq c\, \delta_n(\vartheta)} |J_n(\xi) - J_n(\vartheta)| \;=\; o_{(P_{n,\vartheta})_n}(1) \;, \quad n \to \infty$$

für beliebiges $c < \infty$.

D) Es gibt eine Schätzfolge $(g_n)_n$ in $(\mathcal{E}_n)_n$ für den unbekannten Parameter mit der Eigenschaft

$$\text{für jedes } \vartheta \in \Theta : \quad \mathcal{L}\left(\delta_n^{-1}(\vartheta)\, (g_n - \vartheta) \mid P_{n,\vartheta} \right) \;,\; n \geq 1 \;, \quad \text{ist straff in } \mathbb{R}^d \;.$$

7.15 Bemerkungen: a) Wir kommentieren die Voraussetzungen 7.14 C) im Fall von iid Beobachtungen. Das Einzelexperiment $(\Omega, \mathcal{A}, \{P_\vartheta : \vartheta \in \Theta\})$ – mit $\Theta \subset I\!R^d$ offen – sei an jeder Stelle $\vartheta \in \Theta$ L^2-differenzierbar mit Ableitung V_ϑ, siehe 4.10, und Fisher-Information $J_\vartheta = E_\vartheta(V_\vartheta V_\vartheta^\top)$. Seien zusätzlich die Dichten $f(\vartheta, \omega)$ im Einzelexperiment strikt positiv und hinreichend glatt im Parameter, dann gilt $V_\vartheta = \nabla \log f_\vartheta$ für alle $\vartheta \in \Theta$, wie in 4.1".

Betrachte die Folge $(\mathcal{E}_n)_n$ der n-fachen Produktexperimente , vgl. 4.10. Nach 7.1 und Le Cam's Zweitem Lemma 4.11 gilt LAN an jeder Stelle $\vartheta \in \Theta$, mit Folgen

$$\begin{aligned}
\delta_n(\vartheta) &= \frac{1}{\sqrt{n}} \quad \text{unabhängig von } \vartheta \\
J_n(\vartheta)(\omega) &\equiv J_\vartheta \;=\; E_\vartheta(V_\vartheta V_\vartheta^\top) \quad \text{deterministisch und unabhängig von } n \\
S_n(\vartheta)(\omega) &= \frac{1}{\sqrt{n}} \sum_{i=1}^n V_\vartheta(\omega_i)\ , \quad \omega = (\omega_1, \ldots, \omega_n) \in \mathop{\times}_{i=1}^n \Omega
\end{aligned}$$

und – als Konsequenz des zentralen Grenzwertsatzes – Gauß-Shift-Limesexperiment

$$\mathcal{E}_\infty(\vartheta) \;=\; \mathcal{E}(J_\vartheta)$$

in der Notation aus 5.2. Voraussetzung 7.14 C)ii) ist dann nichts anderes als Stetigkeit

$$\Theta \ni \vartheta \longrightarrow J_\vartheta \in I\!R^{d \times d}$$

der Fisher-Information im Einzelversuch. Voraussetzung 7.14 C)i) bedeutet Gültigkeit – lokal gleichmäßig in $h \in I\!R^d$, und an jeder Stelle $\vartheta \in \Theta$ – einer guten Approximation

$$\begin{aligned}
\left(S_n(\vartheta + h/\sqrt{n}) - S_n(\vartheta)\right)_k(\omega) &\approx \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\sum_{i=j}^n \frac{\partial}{\partial \vartheta_j} \frac{\partial}{\partial \vartheta_k} \log f_\vartheta(\omega_i) \cdot \frac{h_j}{\sqrt{n}} \right] \\
&= \frac{1}{n} \sum_{i=1}^n \left(\left(\nabla \nabla^\top \log f_\vartheta \right) \ h \ \right)_k (\omega_i)
\end{aligned}$$

in jeder Komponente $1 \le k \le d$, zusammen – wegen des starken Gesetzes der großen Zahlen – mit Gültigkeit der Vertauschungsbedingung (1.14)

$$E_\vartheta\left(\nabla \nabla^\top \log f_\vartheta\right) \;=\; -E_\vartheta\left(V_\vartheta V_\vartheta^\top\right) \;=\; -J_\vartheta\ .$$

b) In Voraussetzung 7.14 C) zieht man sich auf S abzählbar dicht in $I\!R^d$ zurück, um Glattheitsvoraussetzungen im Scharparameter der Familien von ZV $\{S_n(\xi) : \xi \in \Theta\}$ und $\{J_n(\xi) : \xi \in \Theta\}$ (für jedes endliche n) vermeiden zu können.

7.16 Hilfssatz: Es gibt einen *lokalen Maßstab mit geschätztem Parameter*

$$D_n \;:=\; \delta_n(g_n)\,, \quad n \geq 1$$

mit der Eigenschaft

$$\text{für jedes } \vartheta \in \Theta : \quad \frac{D_n}{\delta_n(\vartheta)} \;=\; 1 \;+\; o_{(P_{n,\vartheta})_n}(1)\,, \quad n \to \infty\,.$$

Beweis: Wegen 7.14 B) ist D_n eine Zufallsvariable auf $(\Omega_n, \mathcal{A}_n)$ mit Werten in $(0,1]$. Setze $U_n(\vartheta) := \delta_n^{-1}(\vartheta)(g_n - \vartheta)$; nach 7.14 D) ist die Familie $\mathcal{L}(U_n(\vartheta) \mid P_{n,\vartheta})$, $n \geq 1$, straff in $\mathbb{R}^d$. Aus

$$D_n \;=\; \delta_n(g_n) \;=\; \delta_n\left(\vartheta \;+\; \delta_n(\vartheta)\, U_n(\vartheta)\right)$$

und $(*)$ in 7.14 B) folgt nun die Behauptung. □

7.17 Hilfssatz: Überdecke für jedes $k \in \mathbb{N}_0$ den $\mathbb{R}^d$ mit halboffenen Würfeln

$$C(k,\alpha) \;:=\; \mathop{\times}_{i=1}^{d} \left[\alpha_i 2^{-k}, (\alpha_i+1)2^{-k}\right[\,, \quad \alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{Z}^d\,,$$

und setze

$$Z(k) \;:=\; \left\{\alpha \in \mathbb{Z}^d : C(k,\alpha) \subset \Theta\right\}\,.$$

Sei ϑ_0 in $\Theta \cap S$ beliebig. Dann definiert $(G_n)_n$

$$G_n \;:=\; \sum_{k=0}^{\infty} 1_{]2^{-(k+1)}, 2^{-k}]}(D_n) \left(\vartheta_0 + \sum_{\alpha \in Z(k)} (\alpha 2^{-k} - \vartheta_0)\, 1_{C(k,\alpha)}\right)(g_n)$$

eine Folge von Schätzern für den unbekannten Parameter in $(\mathcal{E}_n)_n$ mit den Eigenschaften:

i) für jedes n nimmt G_n alle seine Werte in der abzählbaren Menge $\Theta \cap S$ an ;

ii) für jedes $\vartheta \in \Theta$ gilt $\quad G_n \;=\; g_n \;+\; O_{(P_{n,\vartheta})_n}(\,\delta_n(\vartheta)\,)\,, \quad n \to \infty\,.$

iii) für jedes $\vartheta \in \Theta$:

$$\mathcal{L}\left(\,\delta_n^{-1}(\vartheta)\,(G_n - \vartheta) \mid P_{n,\vartheta}\right)\,,\; n \geq 1\,, \quad \text{ist straff in } \mathbb{R}^d\,.$$

Man nennt $(G_n)_n$ eine *Diskretisierung* der Schätzfolge $(g_n)_n$.

Beweis: Nach Konstruktion gilt i). Zum Nachweis von ii) fixiere $\vartheta \in \Theta$, und wähle $\varepsilon > 0$ mit $B_{(1+2\sqrt{d})\,\varepsilon} \subset \Theta$. Auf dem Ereignis $\{g_n \in B_\varepsilon(\vartheta)\}$ ist man wegen $diam(C(k,\alpha)) = \sqrt{d} \cdot 2^{-k}$ sicher, daß für hinreichend großes k derjenige Würfel $C(k,\alpha)$, in welchen g_n fällt, ganz in Θ enthalten ist, und folglich G_n durch den 'linken unteren Eckpunkt' von $C(k,\alpha)$ gegeben ist. Genauer gilt

$$\begin{array}{l} \sqrt{d}\cdot 2^{-k} < \sqrt{d}\cdot 2D_n < 2\sqrt{d}\cdot\varepsilon \quad \text{und} \quad |G_n - g_n| \;\leq\; \sqrt{d}\cdot 2^{-k} \\ \text{auf} \quad \{\, g_n \in B_\varepsilon(\vartheta)\,,\, D_n < \varepsilon \,\} \cap \left\{ D_n \in]2^{-(k+1)}, 2^{-k}] \right\} \end{array}$$

und insbesondere

$$(+) \qquad |G_n - g_n| \;<\; \sqrt{d}\, 2\, D_n \quad \text{auf} \;\; \{\, g_n \in B_\varepsilon(\vartheta)\,,\, D_n < \varepsilon \,\}\;.$$

Die Konsistenz der Schätzfolge $(g_n)_n$ kombiniert mit 7.16 garantiert aber

$$P_{n,\vartheta}\,(\, g_n \in B_\varepsilon(\vartheta)\,,\, D_n < \varepsilon \,) \;\longrightarrow\; 1\;, \quad n\to\infty\;,$$

und – wieder mit 1.16 – folgt Behauptung ii) aus (+). Danach ergibt sich Behauptung iii) sofort aus der Voraussetzung 7.14 D). □

7.18 Hilfssatz: Es gibt eine *Information mit geschätztem Parameter*

$$\widehat{J}_n \;:=\; J_n(G_n)\;, \quad n\geq 1$$

(wohldefiniert als ZV auf $(\Omega_n,\mathcal{A}_n)$, $n\geq 1$, mit Werten in der Menge D der invertierbaren $d{\times}d$-Matrizen), die für jedes $\vartheta\in\Theta$ erfüllt

$$\text{i)} \qquad \widehat{J}_n \;=\; J_n(\vartheta) \;+\; o_{(P_{n,\vartheta})_n}(1)\;, \quad n\to\infty\;,$$

$$\text{ii)} \qquad \widehat{J}_n^{-1} \;=\; J_n^{-1}(\vartheta) \;+\; o_{(P_{n,\vartheta})_n}(1)\;, \quad \widehat{J}_n^{-1} J_n(\vartheta) \;=\; I_d \;+\; o_{(P_{n,\vartheta})_n}(1)\;, \quad n\to\infty\;.$$

Beweis: 1) Da G_n nur abzählbar viele Werte annimmt, liefert 'Einsetzen' des Schätzers G_n anstelle des Scharparameters ϑ in der Familie von Zufallsvariablen $\{J_n(\vartheta):\vartheta\in\Theta\}$ eine wohldefinierte Zufallsvariable:

$$\widehat{J}_n(\omega) \;=\; \sum_{\xi\,\in\,\Theta\cap S} \left(J_n(\xi)\right)(\omega)\; 1_{\{G_n=\xi\}}(\omega)\;, \quad \omega\in\Omega_n\;.$$

2) Mit $U_n(\vartheta) := \delta_n^{-1}(\vartheta)(G_n-\vartheta)$ schreibt man

$$\widehat{J}_n \;:=\; J_n(G_n) \;=\; J_n(\,\vartheta + \delta_n(\vartheta)U_n(\vartheta)\,)$$

Wegen 7.17 iii) ist für jedes $\vartheta \in \Theta$ die Familie $\mathcal{L}\left(U_n(\vartheta) \mid P_{n,\vartheta}\right)$, $n \geq 1$, straff in $I\!R^d$, G_n ist S-wertig, daher liefert Voraussetzung 7.14 C)ii) sofort die Behauptung i).

2) Auf der Menge D aller invertierbaren $d{\times}d$-Matrizen ist die Abbildung $A \to A^{-1}$ stetig. Wegen LAQ in ϑ hat man nach 7.1 und Schritt 1) oben

$$\mathcal{L}\left(\widehat{J}_n, J_n(\vartheta) \mid P_{n,\vartheta}\right) \longrightarrow \mathcal{L}\left(J(\vartheta), J(\vartheta) \mid P_\vartheta\right)$$

(schwache Konvergenz in $I\!R^{d\times d}{\times}I\!R^{d\times d}$, für $n \to \infty$). Wegen 6.1 ist $\mathcal{L}\left(J(\vartheta) \mid P_\vartheta\right)$ konzentriert auf D. Mit dem Satz über stetige Abbildungen folgt daher

$$\mathcal{L}\left(J_n^{-1}(\vartheta), \widehat{J}_n^{-1}, J_n(\vartheta) \mid P_{n,\vartheta}\right) \longrightarrow \mathcal{L}\left(J^{-1}(\vartheta), J^{-1}(\vartheta), J(\vartheta) \mid P_\vartheta\right)$$

(schwache Konvergenz in $I\!R^{d\times d}{\times}I\!R^{d\times d}{\times}I\!R^{d\times d}$, für $n \to \infty$). Eine weitere Anwendung desselben Satzes liefert Verteilungskonvergenz von $\widehat{J}_n^{-1} - J_n^{-1}(\vartheta)$ und $\widehat{J}_n^{-1} J_n(\vartheta)$ unter $(P_{n,\vartheta})_n$ gegen die deterministischen Matrizen 0_d und I_d , damit also $(P_{n,\vartheta})_n$-stochastische Konvergenz wie in ii). □

7.19 Hilfssatz: Es gibt einen *Score mit geschätztem Parameter*

$$\widehat{S}_n \;:=\; S_n(G_n)\;, \quad n \geq 1\;,$$

und dieser erfüllt für jedes $\vartheta \in \Theta$

$$\widehat{S}_n \;=\; S_n(\vartheta) \;-\; J_n(\vartheta)\left[\delta_n^{-1}(\vartheta)(G_n - \vartheta)\right] \;+\; o_{(P_{n,\vartheta})_n}(1)\;, \quad n \to \infty\;.$$

Beweis: Analog zu Schritten 1) und 2) des Beweises von 7.18, unter Benutzung der Voraussetzung 7.14 C)i) anstelle von 7.14 C)ii) dort. □

7.20 Hauptsatz: Unter den Voraussetzungen 7.14 gilt:

a) Mit den Bezeichnungen aus 7.14–7.19 liefert die *Ein-Schritt-Korrektur*

$$\eta_n \;:=\; G_n \;+\; D_n\, \widehat{J}_n^{-1}\, \widehat{S}_n\;, \quad n \geq 1$$

eine neue Schätzfolge $(\eta_n)_n$ in $(\mathcal{E}_n)_n$, die für jeden Wert des unbekannten Parameters die Eigenschaft $(*)$ aus 7.10 besitzt:

$$\text{für jedes } \vartheta \in \Theta: \quad \delta_n^{-1}(\vartheta)(\eta_n - \vartheta) \;=\; Z_n(\vartheta) \;+\; o_{(P_{n,\vartheta})_n}(1)\;, \quad n \to \infty\;.$$

b) Im lokalen Modell an jeder festen Stelle $\vartheta \in \Theta$ arbeitet $(\eta_n)_n$ asymptotisch so gut wie der Maximum-Likelihood-Schätzer im Limesmodell $\mathcal{E}_\infty(S(\vartheta), J(\vartheta))$.

c) Gilt für ein $\vartheta \in \Theta$ sogar LAMN in ϑ, so folgt:

i) $(\eta_n)_n$ ist regulär und effizient an dieser Stelle ϑ;

ii) $(\eta_n)_n$ erreicht die lokalasymptotische Minimaxschranke in ϑ.

Beweis: Für jedes feste $\vartheta \in \Theta$ folgt mit 7.16–7.19

$$\begin{aligned}
\delta_n^{-1}(\vartheta)(\eta_n - \vartheta) &= \delta_n^{-1}(\vartheta)(G_n - \vartheta) \ + \ \frac{D_n}{\delta_n(\vartheta)} \, \widehat{J}_n^{-1} \, \widehat{S}_n \\
&= \delta_n^{-1}(\vartheta)(G_n - \vartheta) \ + \ \Big(1 + o_{(P_{n,\vartheta})_n}(1)\Big) \, J_n^{-1}(\vartheta) \, \widehat{S}_n \\
&= \Big(1 + o_{(P_{n,\vartheta})_n}(1)\Big) \Big(J_n^{-1}(\vartheta) \, S_n(\vartheta) \ + \ o_{(P_{n,\vartheta})_n}(1)\Big) \\
&= Z_n(\vartheta) \ + \ o_{(P_{n,\vartheta})_n}(1) \ , \quad n \to \infty \ .
\end{aligned}$$

Damit ist Aussage a) des Satzes bewiesen.

Die Aussagen b) und c) folgen wegen a) aus 7.11, 7.12, 7.13. □