

# Kapitel III

# Benachbarkeit

Reinhard Höpfner

Vorlesung Mathematische Statistik

Wintersemester 2004/2005 und Sommersemester 2008

Institut für Mathematik, Johannes Gutenberg Universität Mainz

17.12.04 und 29.05.08

## Übersicht zu Kapitel III :

### A. Le Cam's Erstes und Drittes Lemma

Likelihoodratios und Konventionen für  $\overline{\mathcal{R}}^d$ -wertige Zufallsvariable 3.1

Definition der Benachbarkeit 3.2

Benachbarkeit erhält Straffheit 3.3'

Benachbarkeit und  $\mathcal{R}$ -Straffheit von Likelihoodratios 3.3+3.4

Le Cam's Erstes Lemma (Formulierung und Interpretation) 3.5–3.5'

Le Cam's Drittes Lemma (Formulierung und Interpretation) 3.6-3.6'

Mittelwertshift im Limesmodell mit Normalverteilungen 3.6''

### B. Beweise und Varianten

$\varepsilon$ - $\delta$ -Charakterisierung der Benachbarkeit und Beweis von Satz 3.3' 3.7–3.7'

Bemerkungen zu  $\overline{\mathcal{R}}^d$ -wertigen ZV 3.8

Beweis von Satz 3.4 Teil a) 3.9

Gleichgradige Integrierbarkeit der Likelihoodratios 3.10

Le Cam's Erstes Lemma für einseitige Benachbarkeit 3.11

Benachbarkeit und 'kleine' Werte der Likelihoodratios 3.12

Beweis von Satz 3.4 Teil b) 3.13

Le Cam's Erstes Lemma für wechselseitige Benachbarkeit 3.14

Beweis von Satz 3.5 3.15

Le Cam's Drittes Lemma (einseitige Benachbarkeit) 3.16

Beweis von Satz 3.6 3.17

Beweis von Satz 3.6'' (Mittelwertshift) 3.18

In diesem Kapitel geht es um Likelihood Ratios und ihre Konvergenz. Wir stellen den Begriff der Benachbarkeit (contiguity) vor, der auf LeCam zurueckgeht (siehe LeCam 1969, Davies 1985, LeCam und Yang 1990) und in lokalsymptotischen Betrachtungen eine Schlüsselrolle spielt (siehe unten Kapitel VII). In Teilkapitel A formulieren wir die wichtigsten Aussagen über Benachbarkeit. Diese wurden nach Hájek und Sidak (1967) unter dem Namen 'Le Cam's Erstes Lemma' bzw. 'Le Cam's Drittes Lemma' bekannt. Danach bringen wir in Teilkapitel B alle Beweise, zusammen mit einer Reihe von Varianten zu den Aussagen aus Teilkapitel A.

## A. Le Cam's Erstes und Drittes Lemma

**3.1 Bezeichnungen und Voraussetzungen für dieses Kapitel:** Für jedes  $n \geq 1$  betrachten wir auf einem meßbaren Raum  $(\Omega_n, \mathcal{A}_n)$  zwei Wahrscheinlichkeitsmaße  $P_n$  und  $Q_n$ .

a) i) Ein *Dichtequotient (Likelihoodratio)*  $L_n$  von  $P_n$  bezüglich  $Q_n$  ist eine  $\mathcal{A}_n$ -meßbare Abbildung  $\Omega_n \rightarrow [0, \infty]$  mit der Eigenschaft, daß das Paar  $(L_n, \{L_n = +\infty\})$  eine Lebesgue-Zerlegung von  $P_n$  bezüglich  $Q_n$  liefert:

$$Q_n \{L_n = +\infty\} = 0, \quad P_n(A) = P_n(A \cap \{L_n = +\infty\}) + \int_A L_n dQ_n, \quad A \in \mathcal{A}_n.$$

Ein Dichtequotient  $L_n$  von  $P_n$  bezüglich  $Q_n$  ist  $(P_n + Q_n)$ -fast sicher eindeutig bestimmt. Siehe Witting (1985, Satz 1.110) und Strasser (1985, Kap. I).

Die *log-Likelihoodratio* von  $P_n$  bezüglich  $Q_n$  ist  $\Lambda_n := \log L_n$ , mit Werten in  $\overline{\mathbb{R}} := [-\infty, +\infty]$  (Konvention:  $\log(+\infty) = +\infty$  und  $\log(0) = -\infty$ ).

ii) Sei  $\nu_n$  auf  $(\Omega_n, \mathcal{A}_n)$  ein  $P_n$  und  $Q_n$  dominierendes  $\sigma$ -endliches Maß, schreibe  $p_n$  bzw.  $q_n$  für die  $\nu_n$ -Dichten von  $P_n$  bzw.  $Q_n$ . Dann liefert

$$L_n := \frac{p_n}{q_n} 1_{\{q_n > 0\}} + \infty 1_{\{q_n = 0\}}$$

eine Festlegung des Dichtequotienten von  $P_n$  bezüglich  $Q_n$ .

iii) Ist  $L_n$  eine Festlegung des Dichtequotienten  $L_n$  von  $P_n$  bezüglich  $Q_n$ , so liefert  $\frac{1}{L_n}$  (mit Konvention  $\frac{1}{0} = \infty$  und  $\frac{1}{\infty} = 0$ ) eine Festlegung des Dichtequotienten von  $Q_n$  bezüglich  $P_n$ .

b) Betrachte für jedes  $n \geq 1$  auf dem Raum  $(\Omega_n, \mathcal{A}_n, Q_n)$  eine  $\overline{\mathbb{R}}^d$ -wertige Zufallsvariable  $X_n$ ,  $\overline{\mathbb{R}} := [-\infty, +\infty]$ . Assoziiere für jedes  $n \geq 1$  zu  $X_n$  die  $\mathbb{R}^d$ -wertige ZV  $\hat{X}_n := X_n 1_{\{|X_n| < \infty\}}$ .

i) Die Familie  $(X_n)_n$  heißt  $\mathbb{R}^d$ -straff unter  $(Q_n)_n$  falls gilt:

$$(*) \quad \lim_{n \rightarrow \infty} Q_n(X_n \neq \widehat{X}_n) = 0 \quad , \quad \mathcal{L}(\widehat{X}_n | Q_n) \quad , \quad n \geq 1 \quad , \quad \text{ist straff in } \mathbb{R}^d.$$

ii) Gilt die erste Bedingung in (\*) und läßt sich für ein Wahrscheinlichkeitsmaß  $F$  auf  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  die zweite Bedingung in (\*) zu

$$(**) \quad \mathcal{L}(\widehat{X}_n | Q_n) \longrightarrow F \quad (\text{schwache Konvergenz in } \mathbb{R}^d, \text{ für } n \rightarrow \infty)$$

verschärfen, so spricht man von  $\mathbb{R}^d$ -schwacher Konvergenz der  $\overline{\mathbb{R}}^d$ -wertigen ZV  $(X_n)_n$  unter  $(Q_n)_n$  gegen  $F$  und benutzt die Schreibweise

$$\mathcal{L}(X_n | Q_n) \longrightarrow F \quad (\text{schwache Konvergenz in } \mathbb{R}^d, \text{ für } n \rightarrow \infty).$$

iii) Wir nennen die Familie  $(X_n)_n$  gleichgradig integrierbar unter  $(Q_n)_n$  falls

$$X_n = \widehat{X}_n \quad Q_n\text{-fast sicher} \quad , \quad n \geq 1 \quad , \quad \lim_{K \uparrow \infty} \sup_{n \geq 1} \int_{\{|\widehat{X}_n| > K\}} |\widehat{X}_n| dQ_n = 0 .$$

**3.2 Definition:** Die Folge  $(P_n)_n$  heißt *benachbart* zur Folge  $(Q_n)_n$  (Schreibweise  $(P_n)_n \triangleleft (Q_n)_n$ ) falls für beliebige Folgen von Ereignissen  $(A_n)_n$ ,  $A_n \in \mathcal{A}_n$ , für  $n \rightarrow \infty$  gilt:

$$Q_n(A_n) \rightarrow 0 \quad \implies \quad P_n(A_n) \rightarrow 0 .$$

$(P_n)_n$  und  $(Q_n)_n$  heißen *wechselseitig benachbart* (Schreibweise  $(P_n)_n \triangleleft \triangleright (Q_n)_n$ ) falls

$$(P_n)_n \triangleleft (Q_n)_n \quad \text{und} \quad (P_n)_n \triangleright (Q_n)_n .$$

*Bemerkung:* Benachbarkeit  $(P_n)_n \triangleleft (Q_n)_n$  ist ein asymptotisches Analogon zur klassischen Dominiertheit  $P \ll Q$  für Wahrscheinlichkeitsmaße  $P, Q$  auf einem festen Raum  $(\Omega, \mathcal{A})$ .

**3.3 Satz:** Stets ist  $(L_n)_n$  unter  $(Q_n)_n$   $\mathbb{R}$ -straff.

**Beweis:** Nach 3.1 1a) ist für jedes  $n$  das Ereignis  $\{L_n = +\infty\}$  eine  $Q_n$ -Nullmenge. Es bleibt zu zeigen, daß durch große Wahl von  $K$

$$\sup_{n \geq 1} Q_n(+\infty > L_n > K)$$

beliebig klein gemacht werden kann. Wählt man speziell  $\nu_n := P_n + Q_n$  als dominierendes Maß für  $P_n$  und  $Q_n$ , dann gilt für die zugehörigen Dichten  $p_n + q_n \equiv 1$ ,  $\nu_n$ -fast sicher auf  $\Omega_n$ ; es folgt

$$\begin{aligned} Q_n(+\infty > L_n > K) &= Q_n\left(q_n > 0, \frac{1 - q_n}{q_n} > K\right) = Q_n\left(0 < q_n < \frac{1}{K+1}\right) \\ &= \int_{\Omega_n} 1_{\{q_n < \frac{1}{K+1}\}} q_n d\nu_n \leq \frac{\nu_n(\Omega_n)}{K+1} = \frac{2}{K+1} \end{aligned}$$

nach 3.1 a)ii), wobei die rechte Seite der Ungleichungskette nicht mehr von  $n \geq 1$  abhängt.  $\square$

Nun formulieren und kommentieren wir die Hauptergebnisse dieses Kapitels. Alle Beweise werden in Teilkapitel B gegeben.

**3.3' Satz:** Seien auf  $(\Omega_n, \mathcal{A}_n)$  Zufallsvariable  $Y_n$  mit Werten in  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  gegeben,  $n \geq 1$ .

Aus

$$\mathcal{L}(Y_n | Q_n), n \geq 1 \text{ ist straff, und } (P_n)_n \triangleleft (Q_n)_n$$

folgt

$$\mathcal{L}(Y_n | P_n), n \geq 1 \text{ ist straff.}$$

**3.4 Satz:** a) Die folgenden Aussagen i) und ii) sind gleichwertig:

(i)  $(P_n)_n \triangleleft (Q_n)_n$

(ii)  $(L_n)_n$  unter  $(P_n)_n$  ist  $\mathbb{R}$ -straff.

b) Die folgenden Aussagen i) und ii) sind gleichwertig:

(i)  $(P_n)_n \triangleleft \triangleright (Q_n)_n$

(ii) die Folge  $(\Lambda_n)_n$  ist  $\mathbb{R}$ -straff unter  $(P_n)_n$  und unter  $(Q_n)_n$ .

Aus einer straffen Folge von Wahrscheinlichkeitsmaßen auf  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  kann man schwach konvergente Teilfolgen auswählen. Also ist es nahezu keine Einschränkung der Allgemeinheit, Straffheitsvoraussetzungen nach Übergang zu geeigneten Teilfolgen durch Voraussetzungen über schwache Konvergenz zu ersetzen.

**3.5 Le Cam's Erstes Lemma:** Gilt für ein Wahrscheinlichkeitsmaß  $F$  auf  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

$$\mathcal{L}(\Lambda_n | Q_n) \longrightarrow F \quad (\text{schwache Konvergenz in } \mathbb{R}, \text{ für } n \rightarrow \infty)$$

so hat man die Äquivalenz

$$(P_n)_n \triangleleft \triangleright (Q_n)_n \iff \int_{\mathbb{R}} e^\lambda F(d\lambda) = 1 .$$

**3.5' Bemerkung:** Wir unterlegen die Limesverteilung  $F$  auf  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  für die Folge der log-Likelihoodratios  $(\Lambda_n)_n$  unter  $(Q_n)_n$  durch eine reellwertige ZV  $\Lambda \sim F$ , definiert auf irgendeinem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, Q)$  (eine solche Wahl ist stets möglich, etwa durch  $\Lambda := id$  auf  $(\Omega, \mathcal{A}, Q) := (\mathbb{R}, \mathcal{B}(\mathbb{R}), F)$ ). Die Aussage  $\int_{\mathbb{R}} e^\lambda F(d\lambda) = 1$  bedeutet dann, daß man auf  $(\Omega, \mathcal{A})$  ein zweites Wahrscheinlichkeitsmaß  $P \sim Q$  definieren kann, indem man setzt

$$dP := e^\Lambda dQ .$$

Betrachtet man das binäre Experiment (genau zwei Wahrscheinlichkeitsmaße sind erklärt)

$$(\Omega, \mathcal{A}, \{P, Q\})$$

als *Limesexperiment* für  $n \rightarrow \infty$  zur Folge der binären Experimente

$$(\Omega_n, \mathcal{A}_n, \{P_n, Q_n\}) \quad , \quad n \geq 1 ,$$

so bedeutet wechselseitige Benachbarkeit, daß schwache Limiten von log-Likelihoodratios wieder eine log-Likelihoodratio sind, und zwar in einem Limesexperiment mit zwei *zueinander äquivalenten* Wahrscheinlichkeitsmaßen.

**3.6 Le Cam's Drittes Lemma:** Setze voraus

$$(P_n)_n \triangleleft \triangleright (Q_n)_n .$$

Zusätzlich gebe es auf  $(\Omega_n, \mathcal{A}_n)$  Zufallsvariable  $X_n$  mit Werten in  $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$  so daß

$$\mathcal{L}((\Lambda_n, X_n) | Q_n) \longrightarrow \tilde{F} \quad (\text{schwache Konvergenz in } \mathbb{R}^{1+k}, \text{ für } n \rightarrow \infty)$$

für ein geeignetes Wahrscheinlichkeitsmaß  $\tilde{F}$  auf  $(\mathbb{R}^{1+k}, \mathcal{B}(\mathbb{R}^{1+k}))$ . Dann ist auch

$$\tilde{G}(d\lambda, dx) := e^\lambda \tilde{F}(d\lambda, dx) , \quad \lambda \in \mathbb{R} , \quad x \in \mathbb{R}^k$$

ein Wahrscheinlichkeitsmaß auf  $(\mathbb{R}^{1+k}, \mathcal{B}(\mathbb{R}^{1+k}))$ , und es gilt

$$\mathcal{L}((\Lambda_n, X_n) \mid P_n) \longrightarrow \tilde{G} \quad (\text{schwache Konvergenz in } \mathbb{R}^{1+k}, \text{ für } n \rightarrow \infty).$$

**3.6' Bemerkung:** a) Auch hier unterlegen wir die Limesverteilung  $\tilde{F}$  durch eine kanonische Zufallsvariable  $(\Lambda, X)$ , definiert auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, Q)$ . Für das durch  $dP := e^\Lambda dQ$  definierte Limesexperiment  $(\Omega, \mathcal{A}, \{P, Q\})$  – im Vergleich zum Limesexperiment aus 3.5' ist hier zusätzlich die Existenz einer  $\mathbb{R}^k$ -wertigen ZV  $X$  auf  $(\Omega, \mathcal{A})$  sichergestellt – hat man dann

$$\tilde{F} = \mathcal{L}((\Lambda, X) \mid Q), \quad \tilde{G} = \mathcal{L}((\Lambda, X) \mid P).$$

Denken wir an  $(X_n)_n$  als eine Folge von Statistiken im Modell  $(\Omega_n, \mathcal{A}_n, \{P_n, Q_n\})$ , welche unter  $(Q_n)_n$  *gemeinsam mit den log-Likelihoodratios verteilungskonvergent* ist, so entspricht der Folge  $(X_n)_n$  – unter wechselseitiger Benachbarkeit von  $(P_n)_n$  und  $(Q_n)_n$  – die Statistik  $X$  im Limesmodell  $(\Omega, \mathcal{A}, \{P, Q\})$ .

b) Im Kontext von Minimum-Distanz-Schätzern mit Konvergenzgeschwindigkeit  $\sqrt{n}$ , die wie in Kapitel II D auf reellwertigen iid Beobachtungen basieren, denke man sich die Folge  $(Q_n)_n$  gegeben durch  $(P_\vartheta \mid \mathcal{F}_n)_n$  für ein festes  $\vartheta \in \Theta$ , und die Folge  $(P_n)_n$  gegeben durch  $(P_{\vartheta+h/\sqrt{n}} \mid \mathcal{F}_n)_n$  für ein festes  $h$  in  $\mathbb{R}^d$ ; dabei bezeichne  $\mathcal{F}_n$  die von den ersten  $n$  Beobachtungen erzeugte  $\sigma$ -Algebra,  $n \geq 1$ . Wir werden später (siehe unten 4.11 und 7.3) sehen, daß in diesem Fall wechselseitige Benachbarkeit  $(P_n)_n \triangleleft \triangleright (Q_n)_n$  unter schwachen Glattheitsvoraussetzungen an die Parametrisierung gilt. Die Folge  $(X_n)_n$  in 3.6 identifiziere man mit den reskalierten Schätzfehlern  $(\sqrt{n}(\vartheta_n^* - \vartheta))_n$  an der Stelle  $\vartheta$ . Sofern die reskalierten Schätzfehler in geeigneter Weise an die log-Likelihoodratios

$$\Lambda_n = \frac{d(P_{\vartheta+h/\sqrt{n}} \mid \mathcal{F}_n)}{d(P_\vartheta \mid \mathcal{F}_n)}$$

angekoppelt werden können, erlaubt LeCam's Drittes Lemma, aus der Kenntnis der Grenzverteilung im Fußpunkt  $\vartheta \in \Theta$

$$\mathcal{L}((\Lambda_n, \sqrt{n}(\vartheta_n^* - \vartheta)) \mid P_\vartheta) = \mathcal{L}((\Lambda_n, X_n) \mid Q_n) \longrightarrow \tilde{F}$$

die Grenzverteilung für die reskalierten Schätzfehler auch *in der Nähe von*  $\vartheta$  zu kennen:

$$\mathcal{L}((\Lambda_n, \sqrt{n}(\vartheta_n^* - \vartheta)) \mid P_{\vartheta+h/\sqrt{n}}) = \mathcal{L}((\Lambda_n, X_n) \mid P_n) \longrightarrow \tilde{G}.$$

Dabei gibt 3.6 für die Umrechnung von  $\tilde{F}$  nach  $\tilde{G}$  eine explizite Vorschrift. Nach Verschiebung des Arguments in der zweiten Komponente von  $\tilde{G}$  um  $h$  hat man damit die Grenzverteilung für

$$\mathcal{L}\left((\Lambda_n, \sqrt{n}(\vartheta_n^* - (\vartheta + h/\sqrt{n}))) \mid P_{\vartheta+h/\sqrt{n}}\right) = \mathcal{L}((\Lambda_n, X_n - h) \mid P_n)$$

für  $n \rightarrow \infty$  bei festem  $h$ . Damit verallgemeinert sich die Kenntnis von Grenzverteilungen für Schätzer aus einem festen Fußpunkt  $\vartheta$  heraus auf kleine Umgebungen von  $\vartheta$ .

Wir illustrieren den in LeCam's Drittem Lemma explizit gegebenen Übergang von der Grenzverteilung  $\tilde{F}$  für Paare  $(\Lambda_n, X_n)$  unter  $Q_n$  zur Grenzverteilung  $\tilde{G}$  für dieselben Paare unter  $P_n$  in der für viele Anwendungen wichtigsten Situation:

**3.6'' Satz:** (Mittelwertshift im Normalverteilungsmodell) Sei  $\tilde{F}$  in 3.6 eine Normalverteilung auf  $(\mathbb{R}^{1+k}, \mathcal{B}(\mathbb{R}^{1+k}))$  mit nichtdegenerierter erster Komponente. Dann hat  $\tilde{F}$  eine Gestalt

$$\tilde{F} = \mathcal{N}\left(\left(\begin{pmatrix} -\frac{1}{2}\sigma^2 \\ \mu \end{pmatrix}, \begin{pmatrix} \sigma^2 & \gamma^\top \\ \gamma & \Sigma \end{pmatrix}\right)\right)$$

mit  $\mu \in \mathbb{R}^k$ ,  $\gamma \in \mathbb{R}^k$ ,  $\sigma > 0$ , und mit einer symmetrischen und nichtnegativ definiten  $k \times k$ -Kovarianzmatrix  $\Sigma$ . Die Grenzverteilung  $\tilde{G}$  in 3.6 unter benachbarten Alternativen ist dann

$$\tilde{G} = \mathcal{N}\left(\left(\begin{pmatrix} +\frac{1}{2}\sigma^2 \\ \mu + \gamma \end{pmatrix}, \begin{pmatrix} \sigma^2 & \gamma^\top \\ \gamma & \Sigma \end{pmatrix}\right)\right).$$

Im Spezialfall gemeinsamer asymptotischer Normalität der  $(\Lambda_n, X_n)_n$  unter  $(Q_n)_n$  wird 3.6 damit zu einer einfach anwendbaren Aussage:  $\tilde{G}$  entsteht aus  $\tilde{F}$  durch Hinzufügen des Kovarianzterms

$$\gamma = \lim_{n \rightarrow \infty} \text{Cov}_{Q_n}(\Lambda_n, X_n)$$

zur  $x$ -Komponente des Mittelwerts, sowie durch Umklappen des Vorzeichens in der  $\lambda$ -Komponente.

## B. Beweise und Varianten

Benachbarkeit von  $(P_n)_n$  zu  $(Q_n)_n$  ist ein 'asymptotisches' Analogon zur Dominiertheit  $P \ll Q$  für zwei Wahrscheinlichkeitsmaße  $P, Q$  auf einem festen Raum  $(\Omega, \mathcal{A})$ . In der klassischen Situation hat man eine ' $\varepsilon$ - $\delta$ -Charakterisierung' (siehe Witting 1985, S. 109) der Dominiertheit. Auch

Benachbarkeit kann man so charakterisieren.

**3.7 Hilfssatz:** Die folgenden Aussagen i) und ii) sind gleichwertig:

- i) die Folge  $(P_n)_n$  ist benachbart zu  $(Q_n)_n$ ;
- ii) es gibt zu jedem  $\varepsilon > 0$  ein  $\delta = \delta(\varepsilon) > 0$  so daß gilt:

$$(+) \quad \text{für beliebige Folgen } (A_n)_n, A_n \in \mathcal{A}_n : \quad \limsup_{n \rightarrow \infty} Q_n(A_n) < \delta \implies \limsup_{n \rightarrow \infty} P_n(A_n) < \varepsilon .$$

**Beweis:** ii) $\implies$ i): Sei  $\varepsilon > 0$  beliebig, sei dazu  $\delta = \delta(\varepsilon) > 0$  wie in ii) gegeben. Sei wie in 3.2  $A_n \in \mathcal{A}_n$  eine beliebige Folge von Ereignissen mit  $\lim_{n \rightarrow \infty} Q_n(A_n) = 0$ . Trivialerweise gilt dann  $\limsup_{n \rightarrow \infty} Q_n(A_n) < \delta(\varepsilon)$ . Mit (+) erhält man  $\limsup_{n \rightarrow \infty} P_n(A_n) < \varepsilon$ . Dabei ist  $\varepsilon > 0$  beliebig, also gilt  $\lim_{n \rightarrow \infty} P_n(A_n) = 0$ .

i) $\implies$ ii): Setze voraus  $(P_n)_n \triangleleft (Q_n)_n$ . Angenommen, es gibt ein  $\varepsilon > 0$ , so daß für noch so kleines  $\delta > 0$  nicht (+) ausgesagt werden kann. Dann kann man insbesondere für jedes  $\delta$  der Form  $\delta = \frac{1}{k}$  eine Folge von Ereignissen  $(A_n^k)_n$  angeben mit

$$\limsup_{n \rightarrow \infty} Q_n(A_n^k) < \frac{1}{k} \quad , \quad \limsup_{n \rightarrow \infty} P_n(A_n^k) \geq \varepsilon .$$

Zu jedem  $k$  wählt man einen hinreichend großen Index  $n_k$  mit

$$P_{n_k}(A_{n_k}^k) > \frac{\varepsilon}{2} \quad , \quad Q_{n_k}(A_{n_k}^k) < \frac{1}{k} \quad , \quad k \geq 1 .$$

Die Folge von Ereignissen  $(A_{n_k}^k)_k$  – mit  $A_{n_k}^k \in \mathcal{A}_{n_k}$ ,  $k \geq 1$  – ergänzt man durch Einfügen der leeren Menge zu einer Folge  $(\tilde{A}_m)_m$  – mit  $\tilde{A}_m \in \mathcal{A}_m$ ,  $m \geq 1$  – und erhält

$$\lim_{m \rightarrow \infty} Q_m(\tilde{A}_m) = 0 \quad , \quad \limsup_{m \rightarrow \infty} P_m(\tilde{A}_m) \geq \frac{\varepsilon}{2}$$

im Widerspruch zur Voraussetzung  $(P_n)_n \triangleleft (Q_n)_n$ ; folglich war die Annahme absurd.  $\square$

Hiermit können wir Satz 3.3' beweisen: straffe Folgen  $\mathbb{R}^d$ -wertiger ZV bleiben beim Übergang zu einer benachbarten Folge von Wahrscheinlichkeitsmaßen straff.

**3.7' Beweis des Satzes 3.3':** Für  $n \geq 1$  betrachte Zufallsvariable  $Y_n$  auf  $(\Omega_n, \mathcal{A}_n, Q_n)$  mit Werten in  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ ; es gelte

$$\mathcal{L}(Y_n | Q_n) \quad , \quad n \geq 1 \quad , \quad \text{ist straff} \quad .$$

Sei  $(P_n)_n \triangleleft (Q_n)_n$ . Wähle nach 3.7 zu jedem  $\varepsilon > 0$  ein  $\delta = \delta(\varepsilon) > 0$  so daß (+) in 3.7 ii) gilt. Die Straffheitsvoraussetzung unter  $(Q_n)_n$  erlaubt, für Ereignisse  $A_n := \{|Y_n| > K\} \in \mathcal{A}_n$  durch hinreichend große Wahl einer Konstanten  $K < \infty$

$$\limsup_{n \rightarrow \infty} Q_n(|Y_n| > K) < \delta$$

zu erfüllen; mit (+) aus 3.7 ii) folgt

$$\limsup_{n \rightarrow \infty} P_n(|Y_n| > K) < \varepsilon .$$

Dabei ist  $\varepsilon > 0$  beliebig; also ist die Behauptung

$$\mathcal{L}(Y_n | P_n) , n \geq 1 , \quad \text{ist straff}$$

des Satzes 3.3' bewiesen. □

**3.8 Bemerkung:** Betrachte auf  $(\Omega_n, \mathcal{A}_n, Q_n)$   $\overline{\mathbb{R}}^d$ -wertige Zufallsvariable  $X_n$ ,  $n \geq 1$ , assoziiere zu diesen wie in 3.1 die  $\mathbb{R}^d$ -wertigen ZV  $\widehat{X}_n := X_n 1_{\{|X_n| < \infty\}}$ . Aus den Definitionen in 3.1 b) erhält man sofort die folgenden Aussagen.

a) Genau dann ist die Folge  $(X_n)_n$   $\mathbb{R}^d$ -straff unter  $(Q_n)_n$  falls gilt

$$\lim_{K \uparrow \infty} \limsup_{n \rightarrow \infty} Q_n(+\infty \geq |X_n| > K) = 0 ;$$

(hierbei darf – wegen (\*) in 3.1 b)i) – 'lim sup $_{n \rightarrow \infty}$ ' nicht durch 'sup $_{n \geq 1}$ ' ersetzt werden).

b) Sei  $F$  ein Wahrscheinlichkeitsmaß auf  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ . Dann ist  $\mathbb{R}^d$ -schwache Konvergenz der  $\overline{\mathbb{R}}^d$ -wertigen ZV  $(X_n)_n$  unter  $(Q_n)_n$  gemäß 3.1 b)ii) äquivalent zur Gültigkeit von ( $\times$ )

$$(\times) \quad \text{für jedes } f \in \mathcal{C}_b(\mathbb{R}^d) \text{ gilt } \lim_{n \rightarrow \infty} \int_{\Omega_n} f(X_n) 1_{\{|X_n| < \infty\}} dQ_n = \int_{\mathbb{R}^d} f dF .$$

Beachte hierbei: mit der speziellen Wahl  $f \equiv 1$  impliziert ( $\times$ ) die in 3.1 b)ii) benötigte Aussage

$$\lim_{n \rightarrow \infty} Q_n(|X_n| = \infty) = 0 ,$$

da  $F$  ein Wahrscheinlichkeitsmaß ist.

c) Gleichgradige Integrierbarkeit von  $(X_n)_n$  unter  $(Q_n)_n$  nach 3.1 b)iii) ist gleichwertig (dies zeigt man genau wie üblichen Fall reellwertiger ZV, die auf einem festen Wahrscheinlichkeitsraum leben) mit Gültigkeit der folgenden Aussagen i) und ii) zusammen:

i) es gilt

$$\sup_{n \geq 1} E_{Q_n}(|X_n|) < \infty ;$$

ii) für jedes  $\varepsilon > 0$  gibt es ein  $\delta = \delta(\varepsilon) > 0$  so daß

$$n \geq 1, A_n \in \mathcal{A}_n, Q_n(A_n) < \delta \implies \int_{A_n} |X_n| dQ_n < \varepsilon.$$

Nach 3.3 ist die Folge der Likelihood Ratios  $(L_n)_n$  stets  $\mathbb{R}$ -straff unter  $(Q_n)_n$ . Wir beweisen nun den ersten Teil des Satzes 3.4: Benachbarkeit  $(P_n)_n \triangleleft (Q_n)_n$  ist äquivalent zur  $\mathbb{R}$ -Straffheit von  $(L_n)_n$  auch unter  $(P_n)_n$ .

**3.9 Beweis von Satz 3.4 a):** 1) Wir zeigen die Richtung i) $\implies$ ii) in 3.4 a). Sei  $(P_n)_n \triangleleft (Q_n)_n$ . Zu zeigen ist  $\mathbb{R}$ -Straffheit der Folge der Likelihood Ratios  $(L_n)_n$  unter  $(P_n)_n$ , also gemäß 3.8 a)

$$\lim_{K \uparrow \infty} \limsup_{n \rightarrow \infty} P_n(L_n \in [K, +\infty]) = 0.$$

Angenommen, dies gelte nicht. Dann gibt es ein  $\varepsilon > 0$ , eine Folge  $K_j \uparrow \infty$  und eine Teilfolge  $(n_j)_j$  der natürlichen Zahlen so daß

$$P_{n_j}(L_{n_j} \in [K_j, +\infty]) > \varepsilon \quad \text{für alle } j \geq 1.$$

Da nach 3.3  $(L_n)_n$  unter  $(Q_n)_n$  stets  $\mathbb{R}$ -straff ist, muß gelten

$$\lim_{j \rightarrow \infty} \limsup_{n \rightarrow \infty} Q_n(L_n \in [K_j, +\infty]) = 0.$$

Definiert man sich nun eine Folge  $A_m \in \mathcal{A}_m$  durch  $A_m := \{L_{n_j} \in [K_j, +\infty]\}$  falls  $m = n_j$ , und  $A_m := \emptyset$  sonst, so ist eine Folge von Ereignissen gefunden mit

$$\lim_{m \rightarrow \infty} Q_m(A_m) = 0, \quad \limsup_{m \rightarrow \infty} P_m(A_m) \geq \varepsilon,$$

im Widerspruch zur Voraussetzung  $(P_n)_n \triangleleft (Q_n)_n$ . Also war die Annahme absurd.

2) Wir zeigen ii) $\implies$ i) in 3.4 a): aus  $\mathbb{R}$ -Straffheit von  $(L_n)_n$  unter  $(P_n)_n$  folgt  $(P_n)_n \triangleleft (Q_n)_n$ . Betrachte dazu eine Folge von Ereignissen  $A_n \in \mathcal{A}_n$  mit  $\lim_{n \rightarrow \infty} Q_n(A_n) = 0$  und schreibe

$$\begin{aligned} P_n(A_n) &= P_n(A_n \cap \{+\infty \geq L_n \geq K\}) + P_n(A_n \cap \{L_n \leq K\}) \\ &\leq P_n(L_n \geq K) + \int_{A_n \cap \{L_n \leq K\}} L_n dQ_n \\ &\leq \sup_{m \geq n} P_m(L_m \geq K) + K Q_n(A_n). \end{aligned}$$

Wegen  $\mathbb{R}$ -Straffheit der  $\overline{\mathbb{R}}$ -wertigen Zufallsvariablen  $(L_n)_n$  unter  $(P_n)_n$  nach Voraussetzung kann die rechte Seite der Ungleichungskette im Limes für  $n \rightarrow \infty$  durch große Wahl von  $K$  beliebig

klein gemacht werden. Also gilt  $\lim_{n \rightarrow \infty} P_n(A_n) = 0$ .  $\square$

Wir bereiten den Beweis von Satz 3.4 b) (siehe unten 3.13) vor.

**3.10 Satz:** Die folgenden drei Aussagen sind gleichwertig:

- i)  $(P_n)_n \triangleleft (Q_n)_n$  ;
- ii) Die Folge  $(L_n)_n$  unter  $(Q_n)_n$  ist gleichgradig integrierbar, und  $\lim_{n \rightarrow \infty} P_n(L_n = +\infty) = 0$  ;
- iii) Die Folge  $(L_n)_n$  unter  $(Q_n)_n$  ist gleichgradig integrierbar, und es gilt  $\lim_{n \rightarrow \infty} E_{Q_n}(L_n) = 1$ .

**Beweis:** 1) Beachte, dass  $P_n(L_n = +\infty)$  das Gewicht des  $Q_n$ -singulären Anteils von  $P_n$  bezeichnet, nach 3.1 a)i), während  $E_{Q_n}(L_n)$  die Gesamtmasse des  $Q_n$ -absolutstetigen Anteils von  $P_n$  liefert. Folglich gilt

$$P_n(L_n = +\infty) + E_{Q_n}(L_n) = 1 \quad \text{für alle } n \geq 1$$

und ii) und iii) sind äquivalent.

2) Für jede Konstante  $0 < K < \infty$  zerlegt man

$$\begin{aligned} P_n(K < L_n \leq +\infty) &= P_n(L_n = +\infty) + \int 1_{\{K < L_n < +\infty\}} L_n dQ_n \\ &= (1 - E_{Q_n}(L_n)) + \int 1_{\{K < L_n < +\infty\}} L_n dQ_n . \end{aligned}$$

Unter der Zusatzbedingung  $\lim_{n \rightarrow \infty} P_n(L_n = +\infty) = 0$  ist also die gleichgradige Integrierbarkeit von  $(L_n)_n$  unter  $(Q_n)_n$  – siehe 3.1 b)iii) – äquivalent zur  $\mathbb{R}$ -Straffheit von  $(L_n)_n$  unter  $(P_n)_n$ , und damit wegen 3.4 a) (oder 3.9) äquivalent zur Benachbarkeit  $(P_n)_n \triangleleft (Q_n)_n$ .  $\square$

Nach 3.3 ist die Folge der Likelihoodratios  $(L_n)_n$  unter  $(Q_n)_n$  stets  $\mathbb{R}$ -straff. Also kann man aus jeder Teilfolge  $(n_k)_k$  von  $\mathbb{N}$  eine weitere Teilfolge  $(n_{k_l})_l$  auswählen, welche schwach in  $\mathbb{R}$  entlang  $(n_{k_l})_l$  gegen ein Wahrscheinlichkeitsmaß  $\check{F}$  auf  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  konvergiert. Bis auf den Übergang zu solchen Teilfolgen, den die gewählte Schreibweise nicht zeigt, ist also eine Voraussetzung wie im nächsten Satz fast selbstverständlich. Wir beweisen zuerst eine Fassung von Le Cam's Erstem Lemma für einseitige Benachbarkeit.

**3.11 Satz:** Sei  $F$  ein Wahrscheinlichkeitsmaß auf  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  so daß

$$\mathcal{L}(L_n | Q_n) \longrightarrow \check{F} \quad (\text{schwache Konvergenz in } \mathbb{R}, \text{ für } n \rightarrow \infty)$$

(notwendig ist dabei  $\check{F}$  konzentriert auf  $[0, \infty)$ , möglich ist  $\check{F}(\{0\}) > 0$ ); dann gilt

$$(P_n)_n \triangleleft (Q_n)_n \iff \int_0^\infty \ell \check{F}(d\ell) = 1.$$

*Bemerkung:* Unterlegt man  $\check{F}$  mit einer ZV  $L$  auf einem Raum  $(\Omega, \mathcal{A}, Q)$ , so definiert  $L$  wegen  $\int y \check{F}(dy) = 1$  ein Wahrscheinlichkeitsmaß  $P$  auf  $(\Omega, \mathcal{A})$  via  $dP := LdQ$ . Unter einseitiger Benachbarkeit ist also der schwache Limes von Likelihoodratios stets eine *Dichte*  $\frac{dP}{dQ}$  in einem binären Limesmodell  $(\Omega, \mathcal{A}, \{P, Q\})$ .

**Beweis:** Da Likelihoodratios nichtnegativ sind, da die Folge  $(L_n)_n$  unter  $(Q_n)_n$  nach 3.3 stets  $\mathbb{R}$ -straff ist, muß  $\check{F}$  ein auf  $[0, \infty)$  konzentriertes Wahrscheinlichkeitsmaß sein.

1) Für  $K > 0$  definiere eine 'trunkierte Identität'  $g_K(\cdot) \in \mathcal{C}_b(\mathbb{R})$  durch

$$g_K(x) = 0 \vee x \wedge K, \quad x \in \mathbb{R}.$$

Wegen monotoner Konvergenz gilt sicher

$$\lim_{K \uparrow \infty} \int_0^\infty g_K d\check{F} = \int_0^\infty x \check{F}(dx) \leq +\infty$$

und zugleich

$$(+) \quad 0 \leq E_{Q_n}(L_n) - \int g_K dQ_n^{L_n} = \int 1_{\{K < L_n < +\infty\}} L_n dQ_n - K Q_n^{L_n}((K, \infty)).$$

Dabei gilt für  $n \rightarrow \infty$  in (+) nach Voraussetzung

$$(++) \quad \lim_{n \rightarrow \infty} \int g_K dQ_n^{L_n} = \int g_K d\check{F}.$$

2) Wir zeigen '⇒'. Unter der Voraussetzung  $(P_n)_n \triangleleft (Q_n)_n$  gilt nach 3.10

$$\lim_{n \rightarrow \infty} E_{Q_n}(L_n) = 1, \quad \lim_{K \uparrow \infty} \limsup_{n \rightarrow \infty} \int 1_{\{K < L_n < +\infty\}} L_n dQ_n = 0$$

und damit erst recht  $\lim_{K \uparrow \infty} K Q_n^{L_n}((K, \infty)) = 0$ . Also zeigt (+)

$$\lim_{K \uparrow \infty} \lim_{n \rightarrow \infty} \int g_K dQ_n^{L_n} = 1.$$

Aus (++) erhält man daher  $\int x \check{F}(dx) = 1$ .

3) Wir zeigen '⇐'. Unter der Voraussetzung  $\int x \check{F}(dx) = 1$  hat man in (+)

$$\lim_{K \uparrow \infty} \lim_{n \rightarrow \infty} \int g_K dQ_n^{L_n} = \lim_{K \uparrow \infty} \int g_K d\check{F} = 1, \quad \lim_{n \rightarrow \infty} E_{Q_n}(L_n) = 1.$$

Wählt man nun  $K$  als Stetigkeitsstelle der Verteilungsfunktion  $x \rightarrow \check{F}([0, x])$  mit der Eigenschaft

$$K \check{F}((K, \infty)) \leq \int_K^\infty x \check{F}(dx) < \varepsilon,$$

so ergeben sich aus (+) für hinreichend großes  $K$  die Abschätzungen

$$\lim_{n \rightarrow \infty} K Q_n^{L_n}((K, \infty)) < \varepsilon, \quad \limsup_{n \rightarrow \infty} \int 1_{\{K < L_n < +\infty\}} L_n dQ_n < 2\varepsilon.$$

Dabei ist  $\varepsilon > 0$  beliebig, und wegen 3.10 erhält man  $(P_n)_n \triangleleft (Q_n)_n$ . □

**3.12 Satz:** Gleichwertig sind

- i)  $(P_n)_n \triangleright (Q_n)_n$  ;
- ii)  $\lim_{c \downarrow 0} \limsup_{n \rightarrow \infty} Q_n(L_n \leq c) = 0$  .

**Beweis:** Nach 3.1 a)iii) ist die  $[0, \infty]$ -wertige ZV  $\frac{1}{L_n}$  eine Festlegung der Likelihoodratio von  $Q_n$  bezüglich  $P_n$ . Die Aussage ii) bedeutet  $\mathcal{R}$ -Straffheit der Folge  $(\frac{1}{L_n})_n$  unter  $(Q_n)_n$ . Also ist die Äquivalenz i)  $\iff$  ii) eine Anwendung von 3.4 a) (bzw. 3.9), wobei  $(P_n)_n$  und  $(Q_n)_n$  die Plätze getauscht haben. □

**3.13 Beweis von Teil b) des Satzes 3.4:** Zu zeigen ist, daß wechselseitige Benachbarkeit  $(P_n)_n \triangleleft \triangleright (Q_n)_n$  äquivalent ist zur  $\mathcal{R}$ -Straffheit der Folge  $(\Lambda_n)_n$  sowohl unter  $(P_n)_n$  als auch unter  $(Q_n)_n$ .

1) Hat man wechselseitige Benachbarkeit, so liefern 3.3 und 3.12 zusammen

$$(+) \quad \lim_{K \uparrow \infty} \limsup_{n \rightarrow \infty} Q_n \left( L_n \notin \left[ \frac{1}{K}, K \right] \right) = 0.$$

Für  $K > 1$  gilt aber

$$P_n \left( L_n < \frac{1}{K} \right) < \int 1_{\{L_n < \frac{1}{K}\}} L_n dQ_n < Q_n \left( L_n < \frac{1}{K} \right)$$

so daß aus 3.4 a) (bzw. 3.9) kombiniert mit (+) sofort folgt

$$(++) \quad \lim_{K \uparrow \infty} \limsup_{n \rightarrow \infty} P_n \left( L_n \notin \left[ \frac{1}{K}, K \right] \right) = 0.$$

Wegen  $L_n = e^{\Lambda_n}$  zeigen (+) und (++) die  $\mathcal{R}$ -Straffheit von  $(\Lambda_n)_n$  unter  $(P_n)_n$  und unter  $(Q_n)_n$ .

2) Aus der  $\mathcal{R}$ -Straffheit von  $(\Lambda_n)_n$  unter  $(P_n)_n$  und unter  $(Q_n)_n$  folgt – wegen 3.4 a) (bzw. 3.9) ii)  $\implies$  i) für große Werte von  $L_n$  unter  $(P_n)_n$ , und wegen 3.12 ii)  $\implies$  i) für kleine Werte von  $L_n$

unter  $(Q_n)_n$  – sofort die wechselseitige Benachbarkeit  $(P_n)_n \triangleleft \triangleright (Q_n)_n$ .  $\square$

Satz 3.4 ist damit vollständig bewiesen. Aus 3.4 b) und 3.11 folgt nun eine Formulierung von LeCam's Erstem Lemma unter wechselseitiger Benachbarkeit.

**3.14 Satz:** Sei  $F$  ein Wahrscheinlichkeitsmaß auf  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  so daß gilt

$$\mathcal{L}(L_n | Q_n) \longrightarrow \check{F} \quad (\text{schwache Konvergenz in } \mathbb{R}, \text{ für } n \rightarrow \infty).$$

Dann gilt

$$(P_n)_n \triangleleft \triangleright (Q_n)_n \iff \check{F}(0) = 0 \text{ und } \int_0^\infty \ell \check{F}(d\ell) = 1.$$

**Beweis:** Für Stetigkeitsstellen  $c$  der Verteilungsfunktion  $y \rightarrow \check{F}([0, y])$  gilt

$$\lim_{n \rightarrow \infty} Q_n(L_n \leq c) = \check{F}([0, c]).$$

Also ergibt sich die Behauptung aus 3.11 kombiniert mit 3.12.  $\square$

*Bemerkung:* Unterlegt man  $\check{F}$  mit einer ZV  $L$  auf einem Raum  $(\Omega, \mathcal{A}, Q)$ , so definiert  $L$  wegen  $\check{F}(0) = 0$  und  $\int y \check{F}(dy) = 1$  ein Wahrscheinlichkeitsmaß  $P \sim Q$  auf  $(\Omega, \mathcal{A})$  via  $dP := LdQ$ . Unter wechselseitiger Benachbarkeit ist also der schwache Limes von Likelihoodratios stets eine *strikt positive Dichte*, d.h. das binäre Limesexperiment  $(\Omega, \mathcal{A}, \{P, Q\})$  besteht aus äquivalenten Wahrscheinlichkeitsmaßen.

LeCam's Erstes Lemma in der in Satz 3.5 gegebenen Formulierung ergibt sich nun sofort aus 3.14, indem statt Likelihoodratios unter  $(Q_n)_n$ , deren Grenzverteilung auf  $(0, \infty)$  konzentriert ist, unter leichter Verschärfung der Voraussetzung log-Likelihoodratios betrachtet werden:

**3.15 Beweis von Satz 3.5:** Die in 3.5 gemachte Voraussetzung ( $\mathbb{R}$ -schwache Konvergenz der  $(\Lambda_n)_n$  unter  $(Q_n)_n$  gegen ein Wahrscheinlichkeitsmaß  $F$  auf  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ) impliziert nach dem continuous mapping theorem

$$\mathcal{L}(L_n | Q_n) \longrightarrow \check{F} \text{ mit } \check{F}(\{0\}) = 0 \quad (\text{schwache Konvergenz in } \mathbb{R}, \text{ für } n \rightarrow \infty).$$

wobei  $\check{F}$  als Bild des aus 3.5 gegebenen  $F$  unter der Abbildung  $\lambda \rightarrow e^\lambda$  definiert wird. Danach sind die Aussagen von 3.5 und 3.14 äquivalent.  $\square$

Wir bereiten den Beweis von Le Cam's Drittem Lemma in der Formulierung 3.6 vor. Nach 3.3 ist  $\mathcal{L}(L_n|Q_n)$ ,  $n \geq 1$ , stets  $\mathbb{R}$ -straff. Hat man Straffheit von  $\mathcal{L}(X_n|Q_n)$ ,  $n \geq 1$ , in  $\mathbb{R}^d$ , so hat man auch Straffheit der Paare  $\mathcal{L}((L_n, X_n) | Q_n)$ ,  $n \geq 1$ , in  $\mathbb{R}^{1+d}$ . Dann ist es möglich, Teilfolgen  $(n_l)_l$  und Wahrscheinlichkeitsmaße  $\check{F}$  auszuwählen, so daß  $(L_{n_l}, X_{n_l})$  unter  $(Q_{n_l})_{n_l}$  schwach in  $\mathbb{R}^{1+d}$  gegen  $\check{F}$  konvergieren. Dies ist – bis auf Auswahl von Teilfolgen – die Situation der folgenden Version von Le Cam's Drittem Lemma unter einseitiger Benachbarkeit.

**3.16 Satz:** Sei  $(P_n)_n \triangleleft (Q_n)_n$ . Betrachte eine Schar von ZV  $X_n : (\Omega_n, \mathcal{A}_n) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ ,  $n \geq 1$ , und ein Wahrscheinlichkeitsmaß  $\check{F}$  auf  $(\mathbb{R}^{1+d}, \mathcal{B}(\mathbb{R}^{1+d}))$  so daß

$$\mathcal{L}((L_n, X_n) | Q_n) \longrightarrow \check{F} \quad (\text{schwache Konvergenz in } \mathbb{R}^{1+d}, \text{ für } n \rightarrow \infty).$$

Definiere auf  $(\mathbb{R}^{1+d}, \mathcal{B}(\mathbb{R}^{1+d}))$

$$\check{G}(dl, dx) := \int \ell \mathbf{1}_{\{\ell \geq 0\}} \check{F}(dl, dx), \quad \ell \in \mathbb{R}, x \in \mathbb{R}^d.$$

Dann ist  $\check{G}$  ein Wahrscheinlichkeitsmaß, und es gilt

$$\mathcal{L}((L_n, X_n) | P_n) \longrightarrow \check{G} \quad (\text{schwache Konvergenz in } \mathbb{R}^{1+d}, \text{ für } n \rightarrow \infty).$$

**Beweis:** 1)  $\check{G}$  ist ein Wahrscheinlichkeitsmaß auf  $(\mathbb{R}^{1+d}, \mathcal{B}(\mathbb{R}^{1+d}))$ : es genügt,  $\check{F}(dl)$  für die erste Marginalie von  $\check{F}(dl, dx)$  zu schreiben, und 3.11 anzuwenden:

$$(P_n)_n \triangleleft (Q_n)_n \implies \int \ell \mathbf{1}_{\{\ell \geq 0\}} \check{F}(dl) = 1.$$

2) Schreibe wieder für die trunkierte Identität auf  $\mathbb{R}$

$$g_K(x) := 0 \vee x \wedge K, \quad x \in \mathbb{R}.$$

Sei  $f \in \mathcal{C}_b(\mathbb{R}^{1+d})$  beliebig,  $M := \sup |f|$ . Zu  $\varepsilon > 0$  beliebig wähle  $K = K(\varepsilon)$  so daß

$$(+) \quad \int_{\mathbb{R}^{1+d}} \ell \mathbf{1}_{\{\ell > K\}} \check{F}(dl, dx) = \int_K^\infty \ell \check{F}(dl) < \varepsilon$$

nach 3.11, und

$$(++) \quad \sup_{n \geq 1} \int L_n \mathbf{1}_{\{L_n > K\}} dQ_n < \varepsilon$$

wegen 3.10.

3) Betrachte für  $n \rightarrow \infty$  Zerlegungen

$$\begin{aligned} \int f(L_n, X_n) 1_{\{L_n < \infty\}} dP_n &= \int f(L_n, X_n) L_n dQ_n \\ &= \int f(L_n, X_n) g_K(L_n) dQ_n + \int f(L_n, X_n) (L_n - g_K(L_n)) dQ_n. \end{aligned}$$

Die Funktion  $(l, x) \rightarrow f(l, x) g_K(l)$  ist in  $\mathcal{C}_b(\mathbb{R}^{1+d})$ , also gilt nach Voraussetzung

$$(\diamond) \quad \lim_{n \rightarrow \infty} \int f(L_n, X_n) g_K(L_n) dQ_n = \int f(l, x) g_K(l) \tilde{F}(dl, dx).$$

Für den zweiten Term auf der rechten Seite der Gleichungskette gilt aber nach  $(++)$

$$(\diamond\diamond) \quad \sup_{n \geq 1} \int f(L_n, X_n) (L_n - g_K(L_n)) dQ_n < M \varepsilon.$$

Im Grenzübergang  $n \rightarrow \infty$  entspricht  $(\diamond\diamond)$  wegen  $(+)$  die Aussage

$$(\diamond\diamond\diamond) \quad \int f(l, x) (l - g_K(l)) \tilde{F}(dl, dx) < M \varepsilon$$

und den obenstehenden Zerlegungen entspricht

$$\begin{aligned} &\int f(l, x) g_K(l) \tilde{F}(dl, dx) + \int f(l, x) (l - g_K(l)) \tilde{F}(dl, dx) \\ &= \int f(l, x) l 1_{\{l \geq 0\}} \tilde{F}(dl, dx) = \int f(l, x) \tilde{G}(dl, dx). \end{aligned}$$

Da  $\varepsilon > 0$  beliebig war, hat man mit  $(\diamond)$ – $(\diamond\diamond\diamond)$

$$\lim_{n \rightarrow \infty} \int f(L_n, X_n) 1_{\{L_n < \infty\}} dP_n = \int f(l, x) \tilde{G}(dl, dx).$$

bewiesen. Gemäß  $(\times)$  in 3.8 b) ist damit die gewünschte Aussage

$$\mathcal{L}((L_n, X_n) | P_n) \longrightarrow \tilde{G} \quad (\text{schwache Konvergenz in } \mathbb{R}^{1+d}, \text{ für } n \rightarrow \infty)$$

bewiesen. □

Die in 3.6 gegebene Formulierung von Le Cam's Drittem Lemma unter wechselseitiger Benachbarkeit benutzt log-Likelihoods und braucht daher – siehe 3.12 oder 3.14 – eine stärkere Voraussetzung als 3.16. Der Beweis geht analog zu dem von 3.16:

**3.17 Beweis des Satzes 3.6:** 1) Mit 3.5 sieht man sofort, daß das in 3.6 definierte  $\tilde{G}$  ein Wahrscheinlichkeitsmaß auf  $(\mathbb{R}^{1+d}, \mathcal{B}(\mathbb{R}^{1+d}))$  ist.

2) Definiere  $\tilde{F}$  als Bild des Wahrscheinlichkeitsmaßes  $\tilde{F}$  aus 3.6 unter der Abbildung  $(\lambda, x) \rightarrow (e^\lambda, x)$ . Betrachte eine trunkierte Identität

$$h_K(y) := e^{-K} \vee y \wedge e^{+K}, \quad y \in \mathbb{R}.$$

Wegen 3.5 verschärft sich (+) aus dem Beweis von 3.16 zu

$$(+) \quad \int_{\mathbb{R}^{1+d}} |l - h_K(l)| \tilde{F}(dl, dx) \leq e^{-K} + \int_{\mathbb{R}^{1+d}} l 1_{\{l > e^{+K}\}} \tilde{F}(dl, dx) < \varepsilon$$

für hinreichend große Wahl von  $K$ , und (++) dort verschärft sich wegen 3.14 zu

$$(++) \quad \sup_{n \geq 1} \int |L_n - h_K(L_n)| dQ_n \leq e^{-K} + \sup_{n \geq 1} \int L_n 1_{\{L_n > e^{+K}\}} dQ_n < \varepsilon.$$

3) Nun betrachtet man mit den Modifikationen aus 2) für  $f \in \mathcal{C}_b(\mathbb{R}^{1+d})$  Zerlegungen

$$\begin{aligned} \int f(\Lambda_n, X_n) 1_{\{0 < L_n < \infty\}} dP_n &= \int f(\Lambda_n, X_n) L_n dQ_n \\ &= \int f(\Lambda_n, X_n) h_K(e^{\Lambda_n}) dQ_n + \int f(\Lambda_n, X_n) (L_n - h_K(L_n)) dQ_n \end{aligned}$$

was (bis auf Fehlerterme der Größe  $2M\varepsilon$ , gleichmässig in  $n \geq 1$ , genau wie in  $(\diamond\diamond)$  und  $(\diamond\diamond\diamond)$  im Beweis von 3.16) der Zerlegung

$$\begin{aligned} \int f(\lambda, x) h_K(e^\lambda) \tilde{F}(d\lambda, dx) + \int f(\lambda, x) (e^\lambda - h_K(e^\lambda)) \tilde{F}(d\lambda, dx) \\ = \int f(\lambda, x) e^\lambda \tilde{F}(d\lambda, dx) = \int f(\lambda, x) \tilde{G}(d\lambda, dx) \end{aligned}$$

entspricht. Damit ist die Aussage

$$\lim_{n \rightarrow \infty} \int f(\Lambda_n, X_n) 1_{\{|\Lambda_n| < \infty\}} dP_n = \int f(\lambda, x) \tilde{G}(d\lambda, dx)$$

bewiesen. Diese gilt für beliebiges  $f \in \mathcal{C}_b(\mathbb{R}^{1+d})$ . Gemäß  $(\times)$  in 3.8 b) schließt dies den Beweis des Satzes 3.6 ab.  $\square$

Zum Abschluß betrachten wir Mittelwertshifts in Normalverteilungsmodellen als Spezialfall von Le Cam's Drittem Lemma.

**3.18 Beweis von Satz 3.6'':** Wir betrachten  $\tilde{F}, \tilde{G}$  in Le Cam's Drittem Lemma 3.6 in folgendem Spezialfall:  $\tilde{F}$  ist eine Normalverteilung, deren erste Marginalie – die Grenzverteilung der  $(\Lambda_n)_n$  unter  $(Q_n)_n$  – nicht deterministisch ist.

1) Zuerst ist zu zeigen, daß in diesem Fall  $\tilde{F}$  notwendig eine Gestalt

$$\tilde{F} = \mathcal{N} \left( \left( \begin{pmatrix} -\frac{1}{2}\sigma^2 \\ \mu \end{pmatrix}, \begin{pmatrix} \sigma^2 & \gamma^\top \\ \gamma & \Sigma \end{pmatrix} \right) \right)$$

(mit  $\mu, \gamma \in \mathbb{R}^k$ ,  $\sigma^2 > 0$ , und  $\Sigma$  symmetrisch und nichtnegativ definit) hat. Nach 3.5 weiß man  $\int e^{\lambda \tilde{F}}(d\lambda, dx) = 1$ , also muß ein möglicher Mittelwert  $\nu \in \mathbb{R}$  in der ersten Komponente von  $\tilde{F}$

$$1 = \int e^{\lambda \mathcal{N}(\nu, \sigma^2)}(d\lambda) = \int \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(\lambda^2 - 2\lambda(\nu + \sigma^2) + \nu^2)} d\lambda$$

erfüllen. Nach Einschieben von  $\pm(\nu + \sigma^2)^2$  im Exponent auf der rechten Seite erzwingt dies  $(\nu + \sigma^2)^2 = \nu^2$ , und damit – beachte  $\sigma^2 > 0$  nach Voraussetzung –  $\nu = -\frac{1}{2}\sigma^2$ .

2) Wir beweisen die in 3.6" behauptete Gestalt

$$\tilde{G} = \mathcal{N} \left( \left( \begin{pmatrix} +\frac{1}{2}\sigma^2 \\ \mu + \gamma \end{pmatrix}, \begin{pmatrix} \sigma^2 & \gamma^\top \\ \gamma & \Sigma \end{pmatrix} \right) \right)$$

des Wahrscheinlichkeitsmaßes  $\tilde{G}(d\lambda, dx) := e^{\lambda \tilde{F}}(d\lambda, dx)$ . Setze an  $\tilde{\phi} : \mathbb{R}^{1+d} \rightarrow \mathbb{R}$  für die Laplacetransformierte von  $\tilde{G}$ . Schreibt man das Argument von  $\tilde{\phi}$  in der Form  $\tilde{z} = (\zeta_0, \zeta)$  mit  $\zeta_0 \in \mathbb{R}$  und  $\zeta = (\zeta_1, \dots, \zeta_d) \in \mathbb{R}^d$ , so nach Definition von  $\tilde{G}$

$$\begin{aligned} \tilde{\phi}(\tilde{z}) &= \int_{\mathbb{R}^{1+d}} e^{-\langle \tilde{z}, (\lambda, x) \rangle} \tilde{G}(d\lambda, dx) \\ &= \int e^{-(\zeta_0 \lambda + \zeta^\top x)} \tilde{G}(d\lambda, dx) = \int e^{-((\zeta_0 - 1)\lambda + \zeta^\top x)} \tilde{F}(d\lambda, dx) \end{aligned}$$

(in dieser Gleichung wurde die Schreibweise  $\langle \cdot, \cdot \rangle$  für das Skalarprodukt in  $\mathbb{R}^{1+d}$  und  $\cdot^\top$  für das Skalarprodukt in  $\mathbb{R}^d$  benutzt). Eine Normalverteilung  $\mathcal{N}(m, \Lambda)$  auf  $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$  besitzt die Laplacetransformierte

$$\mathbb{R}^k \ni x \longrightarrow \exp \left\{ -x^\top m + \frac{1}{2} x^\top \Lambda x \right\} \in \mathbb{R}.$$

Folglich erhält man aus der in 1) gegebenen Gestalt von  $\tilde{F}$

$$\tilde{\phi}(\tilde{z}) = \exp \left\{ -(\zeta_0 - 1)\left(-\frac{1}{2}\sigma^2\right) - \zeta^\top \mu + \frac{1}{2} z^\top \begin{pmatrix} \sigma^2 & \gamma^\top \\ \gamma & \Sigma \end{pmatrix} z \right\}, \quad z := \begin{pmatrix} \zeta_0 - 1 \\ \zeta \end{pmatrix}.$$

Eine einfache Umrechnung in einigen zusätzlichen Zeilen bringt dies in die Form

$$\tilde{\phi}(\tilde{z}) = \exp \left\{ -\tilde{z}^\top \tilde{\mu} + \frac{1}{2} \tilde{z}^\top \begin{pmatrix} \sigma^2 & \gamma^\top \\ \gamma & \Sigma \end{pmatrix} \tilde{z} \right\}, \quad \tilde{\mu} := \begin{pmatrix} +\frac{1}{2}\sigma^2 \\ \mu + \gamma \end{pmatrix}$$

und beweist damit die behauptete Gestalt von  $\tilde{G}$ . □