PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 6, Fasc. 2 (1985), p. 109-120

DIFFERENTIABILITY OF LIKELIHOOD RATIOS WITH RATES

BY

WOLFGANG WEFELMEYER (COLOGNE)

Abstract. Call a function *differentiable at some rate* if the difference quotient approximates the derivative at this rate. Consider some root of the density of a one-parameter family of probability measures as a function of the parameter. We characterize differentiability of this root in some (not necessarily corresponding) mean at a certain rate by an appropriate differentiability of an arbitrary other root at the same rate. In particular, we characterize Hellinger differentiability at some rate in terms of differentiability of the densities. This allows us to compare Hellinger differentiability at some rate with a differentiability concept of Pfanzagl and Wefelmeyer [11], which is necessary and sufficient for local asymptotic normality in the i.i.d. case at a certain rate.

1. INTRODUCTION

Let P_t, $t \in \mathbf{R}$, be a family of mutually absolutely continuous probability measures on a measurable space $(X, \mathscr{A})$. Fix $P = P_0$, and let f_t denote a P-density of P_t. We want to compare several approaches to define differentiability of f_t at $t = 0$. Our aim is to decide which of these concepts is most appropriate for obtaining "higher order" local asymptotic normality in the i.i.d. case. More precisely, we are interested in the following particular form of the local asymptotic normality. For every bounded sequence u_n, $n \in \mathbf{N}$, the log-likelihood ratios can be written as

$$\sum_{\nu=1}^{n} \log f_{n^{-1/2}u_n}(x_\nu) = u_n n^{-1/2} \sum_{\nu=1}^{n} g(x_\nu) - \frac{1}{2} u_n^2 \sigma^2 + R_n(x) \tag{1.1}$$

with $n^{-1/2} \sum_{\nu=1}^{n} g(x_\nu)$ asymptotically normal $N(0, \sigma^2)$ under P^n, and $R_n \to 0$ in P^n-probability.

To examine higher order properties of statistical procedures, we require *rates* on $R_n \to 0$ of the form

$$P^n\{|R_n| > \varepsilon n^{-a/2}\} = o(n^{-b/2}) \quad \text{for every } \varepsilon > 0. \tag{1.2}$$

For $a \leqslant 1$ and $b = 2a$, such rates are needed by Bickel et al. [2] to prove that a test statistic is efficient of order $o(n^{-a})$ if it approximates the log-likelihood ratio up to $o(n^{-a/2})$.

For $a = b = 0$, condition (1.2) means that $R_n \to 0$ in P^n-probability. It is known (see [6] and [7]) that local asymptotic normality in the sense of (1.1) holds if and only if $f_t^{1/2}$ is *differentiable in quadratic mean* at $t = 0$. This means that there exists a *derivative* $\frac{1}{2}g$ such that

$$(1.3) \qquad f_t^{1/2} = 1 + t\tfrac{1}{2}g + tr_t \quad \text{with } P(r_t^2) = o(t^0).$$

(Here $P(f)$ stands for $\int f(x)P(dx)$.)

For arbitrary a, b, however (and under the assumptions $P(g) = 0$ and $P(|g|^{2(2+b)/(2-a)}) < \infty$), Pfanzagl and Wefelmeyer [11] found a somewhat different differentiability concept to be necessary and (nearly) sufficient for (1.2) (in the sense that it implies (1.2) for every $a' < a$ instead of a). This is the so-called DCC$_b$-*differentiability of* f_t *at a rate* $o(t^a)$, defined by

$$(1.4) \qquad f_t = 1 + tg + t^{1+a}r_t$$

with r_t fulfilling the following *degenerate convergence criterion* DCC$_b$:

$$\text{(i)} \qquad P\{|r_t| > \varepsilon t^{-1}\} = o(t^{2+b}) \quad \text{for every } \varepsilon > 0,$$

$$\text{(ii)} \qquad P(r_t\{|r_t| \leqslant t^{-1}\}) = o(t),$$

$$\text{(iii)} \qquad P(r_t^2\{|r_t| \leqslant t^{-1}\}) = o(t^0).$$

(For notational convenience, we have identified a set A with its indicator function.)

For $a = b = 0$, differentiability of f_t in this sense is equivalent to local asymptotic normality (1.1), (1.2) and, therefore, to differentiability of $f_t^{1/2}$ in quadratic mean. The question poses itself whether in general DCC$_b$-differentiability of f_t at a certain rate can be described by differentiability of an appropriate root $f_t^{1/c}$ *in c-mean at a rate* $o(t^a)$, say. The latter is the obvious generalization of (1.3) to

$$(1.5) \qquad f_t^{1/c} = 1 + t\frac{1}{c}g + t^{1+a}r_t \quad \text{with } P(|r_t|^c) = o(t^0).$$

To answer this question, we first show in Theorem 2.7 that for $0 \leqslant a < 1$ and $b, c > 0$, differentiability of a root $f_t^{1/c}$ in c-mean at a rate $o(t^a)$ can be described by differentiability of another root $f_t^{1/b}$ *in* (b, c)*-mean at a rate* $o(t^a)$. This means that

$$(1.6) \qquad f_t^{1/b} = 1 + t\frac{1}{b}g + t^{1+a}r_t$$

with r_t fulfilling the following *condition* $R_{1+a,b,c}$:

$$\text{(i)} \qquad P(|r_t|^b\{|r_t| > t^{-1-a}\}) = o(t^{(c-b)(1+a)}),$$

$$\text{(ii)} \qquad P(|r_t|^c\{|r_t| \leqslant t^{-1-a}\}) = o(t^0).$$

The case $b=1$, $c=2$ compares best with condition DCC. Differentiability of f_t in (1, 2)-mean at a rate $o(t^0)$ was introduced in [10] under the name *weak differentiability*. Theorem 2.7 shows that weak differentiability of f_t is equivalent to differentiability of $f_t^{1/2}$ in quadratic mean. A proof via equivalence with (1.1) is given by LeCam [7]. The equivalence remains true with rates $o(t^a)$, where $0 \leqslant a < 1$.

For the question of comparing DCC_b-differentiability with differentiability of roots in some mean with $o(t^a)$, however, the answer is less satisfactory. For $a>0$ we have only obtained the result that differentiability of $f_t^{1/2}$ in quadratic mean (or differentiability of f_t in (1, 2)-mean) at a rate $o(t^a)$, *together* with condition $DCC_{2a}(i)$ on the remainder, implies DCC_{2a}-differentiability of f_t at a rate $o(t^a)$. This follows from Theorems 2.7 and 2.12. Obviously, a condition $DCC_b(i)$ with $b>0$ is not entailed by (1, 2)-differentiability, whatever the rate imposed there. (The case $b=0$ is distinguished by the feature that (1, 2)-differentiability of f_t automatically implies $DCC_0(i)$ for the remainder.) As noted above, however, condition $DCC_b(i)$ is *necessary* for rates of the form (1.2). Hence for higher order considerations DCC_b-differentiability of f_t seems to be better suited than differentiability of roots in some mean.

2. RESULTS

Let P be a fixed probability measure on a measurable space $(X, \mathscr{A})$, and $V \subset (0, \infty)$ arbitrary. For $t \in V$ let P_t be a probability measure with P-density f_t. We think of P_t, $t \in V$, as a path converging to P as $t \to 0$. Such a concept is, of course, void unless 0 is an accumulation point of V.

A definition of differentiability will be based on the following convergence concept.

2.1. Definition. Let $a \geqslant 0$ and $b, c > 0$. For $t \in V$ let $r_t: X \to \mathbf{R}$ be measurable functions. We say that r_t, $t \in V$, *fulfills* $R_{a,b,c}$ if

(i) $$P(|r_t|^b \{|r_t| > t^{-a}\}) = o(t^{(c-b)a}),$$

(ii) $$P(|r_t|^c \{|r_t| \leqslant t^{-a}\}) = o(t^0).$$

Note that r_t fulfills $R_{a,b,b}$ if and only if $P(|r_t|^b) = o(t^0)$.

2.2. Remark. Let $a \geqslant 0$ and $b, c > 0$.

(i) If $b \leqslant c$, then $R_{a,b,c}$ implies $P(|r_t|^b) = o(t^0)$,
since

$$P(|r_t|^b) = P(|r_t|^b \{|r_t| \leqslant t^{-a}\}) + P(|r_t|^b \{|r_t| > t^{-a}\}) \leqslant P(|r_t|^c \{|r_t| \leqslant t^{-a}\})^{b/c} + o(t^{(c-b)a}) = o(t^0).$$

(ii) If $b \geqslant c$, then $P(|r_t|^b) = o(t^0)$ implies $R_{a,b,c}$, since

$$P(|r_t|^b \{|r_t| > t^{-a}\}) \leqslant P(|r_t|^b) = o(t^0) = o(t^{(c-b)a}),$$
$$P(|r_t|^c \{|r_t| < t^{-a}\}) \leqslant P(|r_t|^b)^{c/b} = o(t^0).$$

(iii) It is well known that $P(|r_t|^b) = o(t^0)$ implies $P(|r_t|^c) = o(t^0)$ for $b \geqslant c$; similarly, $R_{a,b,c}$ implies $R_{a,c,b}$.

With condition $R_{a,b,c}$ we can introduce the following differentiability concepts for functions h_t converging to the function $h_0 \equiv 1$ as $t \to 0$.

2.3. Definition. Let $a \geqslant 0$ and $b, c > 0$. For $t \in V$ let h_t: $X \to [0, \infty)$ be measurable functions. We call h_t, $t \in V$, *differentiable in b-mean* (resp., *in* (b, c)-*mean*) *at a rate* $o(t^a)$ *with derivative* g if

$$h_t = 1 + tg + t^{1+a} r_t$$

with r_t fulfilling $P(|r_t|^b) = o(t^0)$ (resp., condition $R_{1+a,b,c}$).

We have found it convenient to describe differentiability by an appropriate Taylor expansion. Of course, h_t is differentiable in b-mean at a rate $o(t^a)$ with derivative g if and only if

$$P(|t^{-1}(h_t - 1) - g|^b)^{1/b} = o(t^a). \tag{2.4}$$

2.5. Remark. For $h_t = f_t^{1/b}$, relation (2.4) writes

$$P(|t^{-1}(f_t^{1/b} - 1) - g|^b)^{1/b} = o(t^a). \tag{2.6}$$

For $a = 0$ and $b = 2$ this reduces to differentiability of $f_t^{1/2}$ in quadratic mean, introduced by LeCam [5]. (This concept was already used by Hájek [4] who refers to a preprint of LeCam's paper.)

2.7. Theorem. *Let* $0 \leqslant a < 1$ *and* $b, c > 0$, *and assume*

$$P(|g|^{c(1+a)} \{|g| \leqslant t^{-1}\}) = O(t^0),$$
$$P(|g|^c \{|g| > t^{-1}\}) = o(t^{ca}) \quad \text{if } b < c,$$
$$P(|g|^b \{|g| > t^{-1}\}) = o(t^{0 \vee (c(1+a)-b)}) \quad \text{if } b > c.$$

Then $h_t^{1/c}$ *is differentiable in c-mean at a rate* $o(t^a)$ *with derivative* $\frac{1}{c} g$ *if and only if* $h_t^{1/b}$ *is differentiable in* (b, c)-*mean at a rate* $o(t^a)$ *with derivative* $\frac{1}{b} g$.

In other words, the representation

$$h_t^{1/c} = 1 + t\frac{1}{c} g + t^{1+a} r_t$$

holds with r_t *fulfilling* $P(|r_t|^c) = o(t^0)$ *if and only if the representation*

$$h_t^{1/b} = 1 + t\frac{1}{b}g + t^{1+a}s_t$$

holds with s_t *fulfilling* $R_{1+a,b,c}$.

The conditions on g hold if $P(|g|^{b \vee c(1+a)})$ is finite. For $b < c$ the third condition follows from the second, for $b > c$ the converse holds. For $a = 0$ the three conditions reduce to $P(|g|^{b \vee c}) < \infty$.

2.8. Remark. Let $c \geqslant b$. If $h_t^{1/b}$ is differentiable in (b, c)-mean (at a rate $o(t^a)$), then, by Remark 2.2 (i), $h_t^{1/b}$ is differentiable in b-mean (at a rate $o(t^a)$). Hence by Theorem 2.7, differentiability of $h_t^{1/c}$ in c-mean implies differentiability of $h_t^{1/b}$ in b-mean (at the same rate). A direct proof of this consequence for $h_t = f_t$ and $a = 0$ is given by Pukelsheim ([12], Theorem 2).

2.9. Remark. We are mainly interested in expressing differentiability of the c-root $f_t^{1/c}$ in c-mean in terms of differentiability of f_t itself. The reader may wonder why we have formulated Definition 2.3 (ii) and Theorem 2.7 for arbitrary h_t instead of f_t. One reason is that some authors consider differentiability of f_t instead of $f_t^{1/c}$ in c-mean[1].

By Remark 2.2 (ii), differentiability of f_t in c-mean is stronger than differentiability of f_t in $(1, c)$-mean and hence, by Theorem 2.7, stronger than differentiability of $f_t^{1/c}$ in c-mean. Applying Theorem 2.7 for $a = 0$, $b = c^2$, and $h_t = f_t^c$, we obtain that f_t is differentiable in c-mean with derivative g if and only if $f_t^{1/c}$ is differentiable in (c^2, c)-mean with derivative $\frac{1}{c}g$. In particular, f_t is differentiable in quadratic mean with derivative g if and only if $f_t^{1/2}$ is differentiable in (4, 2)-mean with derivative $\frac{1}{2}g$.

To compare differentiability in (1, 2)-mean with DCC_b-differentiability, we derive below a relation between the conditions $R_{1+a,1,2}$ and DCC_{2a} defined in the Introduction.

We note first that under $\mathrm{DCC}_{2a}(i)$, the following two conditions are equivalent:

$$P(r_t\{|r_t| \leqslant dt^{-1-a}\}) = o(t^{1+a}), \tag{2.10}$$

$$P(r_t\{|r_t| \leqslant t^{-1}\}) = o(t^{1+a}). \tag{2.11}$$

This follows from

$$P(|r_t|\{t^{-1} < |r_t| \leqslant dt^{-1-a}\}) < dt^{-1-a}P\{|r_t| > t^{-1}\} = o(t^{1+a}).$$

Observe that (2.11) is stronger than DCC (*ii*).

This remark, together with the following theorem, implies that DCC_{2a}-

[1] See [1] (p. 487, Theorem 5) and, for $c = 2$, [10] (p. 23) and also [3] (p. 198).

differentiability of f_t at a rate $o(t^a)$ follows from (1, 2)-differentiability of f_t at a rate $o(t^a)$, augmented by condition $\mathrm{DCC}_{2a}(i)$ on the remainder.

2.12. THEOREM. *Let* $f_t = 1+tg+t^{1+a}r_t$ *with* $P(g) = 0$ *and*

$$P(g\{g > t^{-1}\}) = o(t^{1+2a}),$$

and r_t *fulfilling* $\mathrm{DCC}_{2a}(i)$. *Then* r_t *fulfills* $R_{1+a,1,2}$ *if and only if* r_t *fulfills* $\mathrm{DCC}(iii)$ *and* (2.11).

The proof is based on the following lemma.

2.13. LEMMA. *Let* $f_t = 1+tg+tr_t$ *with* $P(g) = 0$ *and*

$$P(g\{g > t^{-1}\}) = o(t^{1+b}),$$

$$P(r_t\{|r_t| \leqslant 2t^{-1}\}) = o(t^{1+b}).$$

Then $P(|r_t|\{|r_t| > 2t^{-1}\}) = o(t^{1+b})$.

2.14. Remark. Let $f_t^{1/2}$ be differentiable in quadratic mean at a rate $o(t^a)$ with derivative $\frac{1}{2}g$. Assume

$$P(|g|^{2(1+a)}\{|g| \leqslant t^{-1}\}) = o(t^0),$$

$$P(g^2\{|g| > t^{-1}\}) = o(t^{2a}).$$

Then, by Theorem 2.7, f_t is differentiable in (1, 2)-mean at a rate $o(t^a)$ with derivative g. If the remainder term fulfills, in addition, $\mathrm{DCC}_{2a}(i)$, then, by Theorem 2.12, f_t is also DCC_{2a}-differentiable at a rate $o(t^a)$. As indicated in the Introduction, this is sufficient for local asymptotic normality of the form

$$\sum_{\nu=1}^{n} \log f_{n^{-1/2}}(x_\nu) = n^{-1/2}\sum_{\nu=1}^{n} g(x_\nu) - \frac{1}{2}P(g^2) + R_n(x)$$

with $P^n\{|R_n| > \varepsilon n^{-a'/2}\} = o(n^{-a})$ for every $\varepsilon > 0$ and every $a' < a$. With a' replaced by a, this is the rate needed by Bickel et al. [2].

It turns out that for $a < 1/2$ the conditions on g are just sufficient for an appropriate normal convergence rate $o(n^{-a})$ of the log-likelihood ratios. For $\varepsilon_t \downarrow 0$ slowly enough we still have

$$P(g^2\{|g| > \varepsilon_t t^{-1}\}) = o(t^{2a}).$$

Hence by a theorem of Osipov [8] (see also [9], p. 118, Theorem 8), the distribution function of

$$n^{-1/2}\sum_{\nu=1}^{n} g(x_\nu)$$

under P^n converges uniformly to the normal distribution function with variance $P(g^2)$ at a rate $o(n^{-a})$ for $a < 1/2$ and $O(n^{-1/2})$ for $a = 1/2$. No better rates are obtained if the conditions on g are replaced by the sufficient condition that $P(|g|^{2(1+a)})$ is finite.

3. PROOFS

We need the following properties of $R_{a,b,c}$.

3.1. Remark. (i) Let $b \leqslant c$. If r_t fulfills $R_{a,b,c}$, then there exist $\varepsilon_t \downarrow 0$ and $e_t \uparrow \infty$ such that $\varepsilon_t \leqslant d_t \leqslant e_t$ implies

$$P(|r_t|^b \{|r_t| > d_t t^{-a}\}) = o(t^{(c-b)a}), \tag{3.2}$$

$$P(|r_t|^c \{|r_t| \leqslant d_t t^{-a}\}) = o(t^0). \tag{3.3}$$

Conversely, if (3.2), (3.3) hold for arbitrary d_t, then r_t fulfills $R_{a,b,c}$.

(ii) Let $b > c$. If r_t fulfills $R_{a,b,c}$, then (3.2), (3.3) hold for arbitrary d_t. Conversely, if (3.2), (3.3) hold for d_t bounded and bounded away from 0, then r_t fulfills $R_{a,b,c}$.

Proof. (i) Assertion (i) follows easily from the two inequalities

$$\begin{aligned} P(|r_t|^b \{|r_t| > \varepsilon t^{-a}\}) - P(|r_t|^b \{|r_t| > e t^{-a}\}) \\ = P(|r_t|^b \{e t^{-a} \geqslant |r_t| > \varepsilon t^{-a}\}) \\ \leqslant \varepsilon^{b-c} t^{(c-b)a} P(|r_t|^c \{|r_t| \leqslant e t^{-a}\}) \end{aligned}$$

and

$$\begin{aligned} P(|r_t|^c \{|r_t| \leqslant e t^{-a}\}) - P(|r_t|^c \{|r_t| \leqslant \varepsilon t^{-a}\}) \\ = P(|r_t|^c \{\varepsilon t^{-a} < |r_t| \leqslant e t^{-a}\}) \\ \leqslant e^{c-b} t^{(b-c)a} P(|r_t|^b \{|r_t| > \varepsilon t^{-a}\}). \end{aligned}$$

(ii) The case $b > c$ is similar to the case $b \leqslant c$.

3.4. Remark. $R_{a,b,c}$ is additive.

Proof. Let r_t and s_t fulfill $R_{a,b,c}$.

(i) We have

$$\begin{aligned} P(|r_t+s_t|^b \{|r_t+s_t| > 2t^{-a}\}) \leqslant 2^b P(|r_t|^b \{|r_t+s_t| > 2t^{-a}\}) + \\ + 2^b P(|s_t|^b \{|r_t+s_t| > 2t^{-a}\}) = o(t^{(c-b)a}), \end{aligned}$$

since

$$\begin{aligned} P(|r_t|^b \{|r_t+s_t| > 2t^{-a}\}) \leqslant P(|r_t|^b \{|r_t| \leqslant t^{-a}, |s_t| > t^{-a}\}) + P(|r_t|^b \{|r_t| > t^{-a}\}) \\ \leqslant t^{-ba} P\{|s_t| > t^{-a}\} + o(t^{(c-b)a}) \\ \leqslant t^{(c-b)a} P(|s_t|^b \{|s_t| > t^{-a}\}) + o(t^{(c-b)a}) = o(t^{(c-b)a}). \end{aligned}$$

(ii) We have

$$\begin{aligned} P(|r_t+s_t|^c \{|r_t+s_t| \leqslant 2t^{-a}\}) \leqslant P(|r_t+s_t|^c \{|r_t| \leqslant t^{-a}, |s_t| \leqslant t^{-a}\}) + \\ + P(|r_t+s_t|^c \{|r_t+s_t| \leqslant 2t^{-a}, |r_t| > t^{-a}\}) + \\ + P(|r_t+s_t|^c \{|r_t+s_t| \leqslant 2t^{-a}, |s_t| > t^{-a}\}) = o(t^0), \end{aligned}$$

since

$$P(|r_t+s_t|^c\{|r_t|\leqslant t^{-a},\ |s_t|\leqslant t^{-a}\})$$
$$\leqslant 2^c P(|r_t|^c\{|r_t|\leqslant t^{-a}\})+2^c P(|s_t|^c\{|s_t|\leqslant t^{-a}\})=o(t^0)$$

and

$$P(|r_t+s_t|^c\{|r_t+s_t|\leqslant 2t^{-a},\ |r_t|>t^{-a}\})\leqslant 2^c t^{-ca}P\{|r_t|>t^{-a}\}$$
$$\leqslant 2^c t^{(b-c)a}P(|r_t|^b\{|r_t|>t^{-a}\})=o(t^0).$$

(iii) Condition $R_{a,b,c}$ for r_t+s_t now follows from parts (i), (ii) of the proof and Remark 3.1.

3.5. Remark. If r_t fulfills $R_{a,b,c}$ and $|s_t|\leqslant |r_t|$, then s_t fulfills $R_{a,b,c}$.

Proof. Condition $R_{a,b,c}(i)$ is trivially fulfilled for s_t. For $b\leqslant c$, condition $R_{a,b,c}$ for s_t follows from

$$P(|s_t|^c\{|s_t|\leqslant t^{-a}\})\leqslant P(|s_t|^c\{|s_t|\leqslant t^{-a},\ |r_t|>t^{-a}\})+P(|s_t|^c\{|r_t|\leqslant t^{-a}\})$$
$$\leqslant t^{(b-c)a}P(|s_t|^b\{|r_t|>t^{-a}\})+P(|r_t|^c\{|r_t|\leqslant t^{-a}\})$$
$$\leqslant t^{(b-c)a}P(|r_t|^b\{|r_t|>t^{-a}\})+o(t^0)=o(t^0).$$

The case $b>c$ is treated similarly.

Proof of Theorem 2.7. We restrict ourselves to the case $b<c$. The case $b=c$ is trivial, and the case $b>c$ is treated as the case $b<c$.

(a) The following expansion will be used to prove both implications of the assertion. Let

$$h_t^{1/b}=1+t\frac{1}{b}g+t^{1+a}r_t. \tag{3.6}$$

Then

$$h_t^{1/c}=(1+t\frac{1}{b}g+t^{1+a}r_t)^{b/c}.$$

For z in a neighborhood of 0 we have the following Taylor expansion:

$$\left|(1+z)^{b/c}-\left(1+\frac{b}{c}z\right)\right|\leqslant dz^2.$$

Let

$$A_t:=\{|g|\leqslant \varepsilon_t t^{-1},\ |r_t|\leqslant \varepsilon t^{-1-a}\}$$

with $\varepsilon_t\downarrow 0$ sufficiently slowly and $\varepsilon>0$ sufficiently small. From the Taylor expansion we obtain

$$h_t^{1/c}=1+t\frac{1}{c}g+t^{1+a}s_t \tag{3.7}$$

with

$$(3.8)\qquad |s_t| \leqslant t^{-1-a} A_t^c + t^{-a}\frac{1}{c}|g|\, A_t^c + \frac{b}{c}|r_t|\, A_t +$$
$$+ t^{1-a} d\left(\frac{1}{b}g + t^a r_t\right)^2 A_t + t^{-1-a}\left|1 + t\frac{1}{b}g + t^{1+a} r_t\right|^{b/c} A_t^c.$$

(b) Assume that $h_t^{1/b}$ is differentiable in (b, c)-mean at a rate $o(t^a)$ with derivative $\frac{1}{b}g$. Then (3.6) holds with r_t fulfilling $R_{1+a,b,c}$. Hence (3.7) holds with s_t fulfilling (3.8). We have to prove $P(|s_t|^c) = o(t^0)$. We may treat each right-hand term in (3.8) separately as follows:

$$t^{-c(1+a)} P(A_t^c) \leqslant t^{-c(1+a)} P\{|g| > \varepsilon_t t^{-1}\} + t^{-c(1+a)} P\{|r_t| > \varepsilon t^{-1-a}\}$$
$$\leqslant \varepsilon_t^{-c} t^{-ca} P(|g|^c \{|g| > \varepsilon_t t^{-1}\}) +$$
$$+ \varepsilon^{-b} t^{(b-c)(1+a)} P(|r_t|^b \{|r_t| > \varepsilon t^{-1-a}\}) = o(t^0);$$

$$t^{-ca} P(|g|^c A_t^c) = t^{-ca} P(|g|^c \{|g| \leqslant \varepsilon_t t^{-1}, |r_t| > \varepsilon t^{-1-a}\}) +$$
$$+ t^{-ca} P(|g|^c \{|g| > \varepsilon_t t^{-1}\})$$
$$\leqslant \varepsilon_t^c t^{-c(1+a)} P\{|r_t| > \varepsilon t^{-1-a}\} + o(t^0)$$
$$\leqslant \varepsilon_t^c \varepsilon^{-b} t^{(b-c)(1+a)} P(|r_t|^b \{|r_t| > \varepsilon t^{-1-a}\}) + o(t^0) = o(t^0);$$

$$P(|r_t|^c A_t) \leqslant P(|r_t|^c \{|r_t| \leqslant \varepsilon t^{-1-a}\}) = o(t^0);$$
$$t^{c(1-a)} P(|g|^{2c} A_t) \leqslant \varepsilon_t^{c(1-a)} P(|g|^{c(1+a)} \{|g| \leqslant \varepsilon_t t^{-1}\}) = o(t^0);$$
$$t^{c(1-a)} t^{2ca} P(|r_t|^{2c} A_t) \leqslant \varepsilon^c P(|r_t|^c \{|r_t| \leqslant \varepsilon t^{-1-a}\}) = o(t^0);$$

$$t^{-c(1+a)} t^b P(|g|^b A_t^c) = t^{b-c(1+a)} P(|g|^b \{|g| \leqslant \varepsilon_t t^{-1}, |r_t| > \varepsilon t^{-1-a}\}) +$$
$$+ t^{b-c(1+a)} P(|g|^b \{|g| > \varepsilon_t t^{-1}\})$$
$$\leqslant \varepsilon_t^b t^{-c(1+a)} P\{|r_t| > \varepsilon t^{-1-a}\}$$
$$+ \varepsilon_t^{c-b} t^{-ca} P(|g|^c \{|g| > \varepsilon_t t^{-1}\})$$
$$\leqslant \varepsilon_t^b \varepsilon^{-b} t^{(b-c)(1+a)} P(|r_t|^b \{|r_t| > \varepsilon t^{-1-a}\}) + o(t^0) = o(t^0);$$

$$t^{-c(1+a)} t^{b(1+a)} P(|r_t|^b A_t^c) = t^{(b-c)(1+a)} P(|r_t|^b \{|r_t| \leqslant \varepsilon t^{-1-a}, |g| > \varepsilon_t t^{-1}\})$$
$$+ t^{(b-c)(1+a)} P(|r_t|^b \{|r_t| > \varepsilon t^{-1-a}\})$$
$$\leqslant \varepsilon^b t^{-c(1+a)} P\{|g| > \varepsilon_t t^{-1}\} + o(t^0)$$
$$\leqslant \varepsilon^b \varepsilon_t^{-c} t^{-ca} P(|g|^c \{|g| > \varepsilon_t t^{-1}\}) + o(t^0) = o(t^0).$$

(c) Assume that $h_t^{1/c}$ is differentiable in c-mean at a rate $o(t^a)$ with derivative $\frac{1}{c}g$. We apply part (a) of the proof with b, c and r_t, s_t interchanged. By assumption, (3.6) holds with s_t fulfilling $P(|s_t|^c) = o(t^0)$. Hence (3.7) holds with r_t fulfilling (3.8). We have to show that r_t fulfills $R_{1+a,b,c}$. By Remarks 3.4 and 3.5 it suffices to prove that each right-hand term in (3.8) fulfills $R_{1+a,b,c}$. Consider first $R_{1+a,b,c}(i)$. Choose $d \geqslant 1$ sufficiently large. Then

$$P(|r_t|^b \{|r_t| > dt^{-1-a}\}) = o(t^{(c-b)(1+a)})$$

is trivially fulfilled for those terms in (3.8) which are bounded by $o(t^{-1-a})$. The remaining terms are treated as follows:

$$t^{-ba} P(|g|^b A_t^c \{t^{-a}|g| A_t^c > dt^{-1-a}\}) \leqslant t^{-ba} P(|g|^b \{|g| > t^{-1}\})$$
$$\leqslant t^{c-b(1+a)} P(|g|^c \{|g| > t^{-1}\}) = o(t^{(c-b)(1+a)});$$
$$t^{-b(1+a)} t^c P(|g|^c A_t^c \{t^{-1-a} t^{c/b} |g|^{c/b} A_t^c > dt^{-1-a}\})$$
$$\leqslant t^{c-b(1+a)} P(|g|^c \{|g| > t^{-1}\}) = o(t^{(c-b)(1+a)});$$
$$t^{-b(1+a)} t^{c(1+a)} P(|s_t|^c A_t^c \{t^{-1-a} t^{c(1+a)/b} |s_t|^{c/b} A_t^c > dt^{-1-a}\})$$
$$\leqslant t^{(c-b)(1+a)} P(|s_t|^c) = o(t^{(c-b)(1+a)}).$$

Consider now $R_{1+a,b,c}(ii)$. We have to prove

$$P(|r_t|^c \{|r_t| \leqslant dt^{-1-a}\}) = o(t^0)$$

for the right-hand terms in (3.8). (Recall that b, c and r_t, s_t are now interchanged in (3.8).) This is done as follows:

$$t^{-c(1+a)} P(A_t^c) \leqslant t^{-c(1+a)} P\{|g| > \varepsilon_t t^{-1}\} + t^{-c(1+a)} P\{|s_t| > \varepsilon t^{-1-a}\}$$
$$\leqslant \varepsilon_t^{-c} t^{-ca} P(|g|^c \{|g| > \varepsilon_t t^{-1}\}) + \varepsilon^{-c} P(|s_t|^c) = o(t^0);$$
$$t^{-ca} P(|g|^c A_t^c \{t^{-a}|g| A_t^c \leqslant dt^{-1-a}\}) = t^{-ca} P(|g|^c A_t^c)$$
$$= t^{-ca} P(|g|^c \{|g| \leqslant \varepsilon_t t^{-1}, |s_t| > \varepsilon t^{-1-a}\}) +$$
$$+ t^{-ca} P(|g|^c \{|g| > \varepsilon_t t^{-1}\})$$
$$\leqslant \varepsilon_t^c t^{-c(1+a)} P\{|s_t| > \varepsilon t^{-1-a}\} + o(t^0)$$
$$\leqslant \varepsilon_t^c \varepsilon^{-c} P(|s_t|) + o(t^0) = o(t^0);$$
$$P(|s_t|^c A_t \{|s_t| A_t \leqslant dt^{-1-a}\}) \leqslant P(|s_t|^c) = o(t^0);$$
$$t^{c(1-a)} P(|g|^{2c} A_t \{t^{1-a} g^2 A_t \leqslant dt^{-1-a}\}) \leqslant t^{c(1-a)} P(|g|^{2c} A_t)$$
$$\leqslant \varepsilon_t^{c(1-a)} P(|g|^{c(1+a)} \{|g| \leqslant t^{-1}\}) = o(t^0);$$
$$t^{c(1-a)} t^{2ca} P(|s_t|^{2c} A_t \{t^{1-a} t^{2a} s_t^2 A_t \leqslant dt^{-1-a}\})$$
$$\leqslant t^{c(1+a)} P(|s_t|^{2c} \{|s_t| \leqslant \varepsilon t^{-1-a}\}) \leqslant \varepsilon^c P(|s_t|^c) = o(t^0);$$

$$t^{-c(1+a)} t^{c^2/b} P(|g|^{c^2/b} A_t^c \{t^{-1-a} t^{c/b} |g|^{c/b} A_t^c \leqslant dt^{-1-a}\})$$
$$\leqslant t^{c^2/b - c(1+a)} P(|g|^{c^2/b} \{|s_t| > \varepsilon t^{-1-a}, |g| \leqslant d^{b/c} t^{-1}\}) +$$
$$+ t^{c^2/b - c(1+a)} P(|g|^{c^2/b} \{\varepsilon_t t^{-1} < |g| \leqslant d^{b/c} t^{-1}\})$$
$$\leqslant d^c t^{-c(1+a)} P\{|s_t| > \varepsilon t^{-1-a}\} + d^{c-b} t^{-ca} P(|g|^c \{|g| > \varepsilon_t t^{-1}\})$$
$$\leqslant d^c \varepsilon^{-c} P(|s_t|^c) + o(t^0) = o(t^0);$$
$$t^{-c(1+a)} t^{c^2(1+a)/b} P(|s_t|^{c^2/b} A_t^c \{t^{-1-a} t^{c(1+a)/b} |s_t|^{c/b} A_t^c \leqslant dt^{-1-a}\})$$
$$\leqslant t^{c^2(1+a)/b - c(1+a)} P(|s_t|^{c^2/b} \{|s_t| \leqslant d^{b/c} t^{-1-a}\}) \leqslant d^{c-b} P(|s_t|^c) = o(t^0).$$

Proof of Theorem 2.12. (i) Assume that r_t fulfills DCC(*iii*) and (2.12). Then r_t also fulfills (2.11) with $d = 2$. By Lemma 2.13, applied for $b = 2a$ and $t^a r_t$ instead of r_t,

$$P(|r_t| \{|r_t| > 2t^{-1-a}\}) = o(t^{1+a}).$$

From DCC(*iii*) we obtain

$$P(r_t^2 \{|r_t| \leqslant 2t^{-1-a}\}) = P(r_t^2 \{t^{-1} < |r_t| \leqslant 2t^{-1-a}\}) + P(r_t^2 \{|r_t| \leqslant t^{-1}\})$$
$$\leqslant 2t^{-1-a} P(|r_t| \{|r_t| > t^{-1}\}) + o(t^0) = o(t^0).$$

Hence $R_{1+a,1,2}$ holds by Remark 3.1.

(ii) Assume that r_t fulfills $R_{1+a,1,2}$. Since $P(r_t) = 0$, we obtain

$$|P(r_t \{|r_t| \leqslant t^{-1-a}\})| = |P(r_t \{|r_t| > t^{-1-a}\})| \leqslant P(|r_t| \{|r_t| > t^{-1-a}\}) = o(t^{1+a}).$$

Hence (2.10) holds, which is equivalent to (2.11). Furthermore, DCC(*iii*) follows trivially from $R_{1+a,1,2}(ii)$.

Proof of Lemma 2.13. Since $f_t \geqslant 0$, we have $g + r_t \geqslant -t^{-1}$. Hence $r_t < -2t^{-1}$ implies

$$g \geqslant -t^{-1} - r_t > t^{-1}, \quad r_t \geqslant -t^{-1} - g > -2g.$$

We obtain

$$0 \leqslant -P(r_t \{r_t < -2t^{-1}\}) \leqslant 2P(g \{g > t^{-1}\}) = o(t^{1+b}).$$

By assumption, $P(r_t) = 0$. Hence

$$P(|r_t| \{|r_t| > 2t^{-1}\}) = -2P(r_t \{r_t < -2t^{-1}\}) - P(r_t \{|r_t| \leqslant 2t^{-1}\}) = o(t^{1+b}).$$

REFERENCES

[1] E. W. Barankin, *Locally best unbiased estimates*, Ann. Math. Statist. 20 (1949), p. 477-501.

[2] P. J. Bickel, D. M. Chibisov, and W. R. van Zwet, *On efficiency of first and second order*, Internat. Statist. Rev. 49 (1981), p. 169-175.

[3] V. Fabian and J. Hannan, *On the Cramér-Rao inequality*, Ann. Statist. 5 (1977), p. 197-205.

[4] J. Hájek, *Asymptotically most powerful rank-order test*, Ann. Math. Statist. 33 (1962), p. 1124-1147.

[5] L. LeCam, *Likelihood functions for large numbers of independent observations*. In: *Research Papers in Statistics*, Festschrift for J. Neyman (F. N. David, ed.), p. 167-187, Wiley, London 1966.

[6] – *Théorie asymptotique de la décision statistique*, Presses de l'Université de Montréal, 1969.

[7] – *Differentiability, tangent spaces and Gaussian auras*, Unpublished manuscript (1984).

[8] L. V. Osipov, *Refinement of Lindeberg's theorem*, Theory Probab. Appl. 11 (1966), p. 299-302.

[9] V. V. Petrov, *Sums of Independent Random Variables*, Springer-Verlag, New York 1975.

[10] J. Pfanzagl (with the assistance of W. Wefelmeyer), *Contributions to a General Asymptotic Statistical Theory*, Lecture Notes in Statistics 13, Springer-Verlag, New York 1982.

[11] – *Asymptotic Expansions for General Statistical Models*, Lecture Notes in Statistics 31, Springer-Verlag, Berlin 1985.

[12] F. Pukelsheim, *On L_p-differentiable distributions, exponential families, and the Cramér-Rao inequality*. Unpublished manuscript (1981).

Mathematical Institute
University of Cologne
Weyertal 86
5000 Cologne 41
FRG

Received on 22. 10. 1984