The Annals of Statistics
1987, Vol. 15, No. 4, 1401–1420

ON LOCAL AND NONLOCAL MEASURES OF EFFICIENCY

BY WILBERT C. M. KALLENBERG AND TERESA LEDWINA[1]

Twente University of Technology and Technical University of Wrocław

General results on the limiting equivalence of local and nonlocal measures of efficiency are obtained. Why equivalence occurs in so many testing and estimation problems is clarified. Uniformity of the convergence is a key point. The concepts of Fréchet- and Hadamard-type differentiability, which imply uniformity, play an important role. The theory is applied to tests based on linear rank statistics, showing equivalence of the local limit of *exact* Bahadur efficiency and Pitman efficiency. As a second application, the relation between the inaccuracy rate and the asymptotic variance of L-estimators is investigated.

1. Introduction. Both in testing theory and in estimation theory asymptotic comparison of competitive procedures can be made in a local and a nonlocal way. In testing theory the most familiar measures are Pitman efficiency and Bahadur efficiency, while their counterparts in estimation theory are based on asymptotic variances and inaccuracy rates. In typical cases the local limit of the Bahadur efficiency equals the limiting Pitman efficiency, where in the latter case the limit is taken with respect to levels of significance tending to zero. A similar phenomenon appears in estimation theory: In typical cases the local limit of the inaccuracy rate equals $(2\sigma^2)^{-1}$, where σ^2 denotes the asymptotic variance. In Jurečková and Kallenberg (1987) it is shown that both issues are strongly related.

Under mild conditions, Wieand (1976) has shown that the limiting *approximate* Bahadur efficiency equals the limiting Pitman efficiency. Approximate Bahadur efficiency, however, is in itself of little value as a measure of performance of tests since monotone transformations of a test statistic may lead to entirely different approximate Bahadur slopes [cf. Groeneboom and Oosterhoff (1977)]. Following the same line of argument in estimation theory as Wieand did in testing theory, it is shown in Jurečková and Kallenberg (1987) that the local limit of the approximate inaccuracy rate trivially equals $(2\sigma^2)^{-1}$.

It seems to be more natural to consider *exact* Bahadur efficiency and *exact* inaccuracy rates, because they are meaningful measures of nonlocal performance by itself. Coincidence of the limits of exact Bahadur efficiency and Pitman efficiency requires an inversion in the order of taking limits (cf. Section 2). This is permitted when we have *uniform* convergence. (With approximate measures of comparison, one effectively ignores this problem by changing the original definition of the nonlocal measure.) The phenomenon of equivalence in the limit of the

Received May 1986; revised December 1986.
AMS 1980 *subject classifications*. Primary 62G20; secondary 60F10.
Key words and phrases. Exact Bahadur efficiency, Pitman efficiency, inaccuracy rate, asymptotic variance, large deviations, Fréchet-type differentiability, Hadamard-type differentiability, linear rank statistics, L-estimators.

exact nonlocal and the local measure has been mentioned by several authors and is established in a lot of special cases. However, no general results are available, specifying what is meant by "typical" cases and clarifying why the equivalence occurs in so many testing and estimation problems. It is the purpose of this paper to present a general theory to give more insight into the nature of the problem.

Here we consider estimators and test statistics of the form $T_n = T(\hat{P}_n)$, where T is a *fixed* functional and $\hat{P}_n$ is the empirical probability measure based on n observations. The concepts of Fréchet- and Hadamard-type differentiability of T, which imply uniformity, play an important role. It has to be emphasized that statistics, which are locally very close to each other, may be very different in their nonlocal behaviour. A typical example is given by Basu (1955) [cf. also Jurečková and Kallenberg (1987), Section 5(a)]. Therefore, local properties (like Fréchet- and Hadamard-type differentiability) have to be used after having obtained the required large deviation probabilities. By doing that, the complicated large deviation result is approximated by a large deviation result corresponding to a sum of i.i.d. random variables [cf. also Remarks 2.1 and 2.4]. Restriction to fixed functionals and uniformity in the differentiability concepts are the two elements by which the inversion in order of taking limits is obtained.

In view of the preceding we use the following approach: First we invoke a general large deviation theorem for statistical functionals and then we use the local properties of our statistical functional in order to reduce the complicated expression obtained from the large deviation theorem to a manageable quantity.

The equivalence of local and nonlocal measures is described in detail in terms of estimation theory. The analogous results on exact Bahadur and Pitman efficiency in testing theory can be obtained in a similar way.

In Section 2 the main results are presented and in Section 3 the theory is applied both on testing and on estimation problems. The local behaviour of the exact Bahadur slope of linear rank tests is obtained; this, e.g., generalizes results of Kremer (1981). As a second application, the inaccuracy rate of L-estimators is investigated. Previous results of Fu (1980) are clarified and generalized. The proofs are given in Section 4.

2. Main results. Let S be a Hausdorff space and let $\mathscr{B}$ be the σ-field of Borel sets in S. The set of all probability measures on $\mathscr{B}$ is denoted by Λ. Let $X_1, X_2 \dots$, be a sequence of i.i.d. random variables taking values in S according to a probability measure $P \in \Lambda$. For each positive integer n, the empirical probability measure based on $X_1, \dots, X_n$ is denoted by $\hat{P}_n$. So $\hat{P}_n(B)$ is the fraction of X_i's, $1 \le i \le n$, with values in the set $B \in \mathscr{B}$. Consider a real valued functional T: $\Lambda \to \mathbb{R}$. The statistic $T_n = T(\hat{P}_n)$ is applied as an estimator of the unknown parameter $T(P) = \theta$, say, that has to be estimated. Denote by

$$(2.1) \qquad a_n(\varepsilon, T) = P\{|T_n - T(P)| > \varepsilon\}$$

the probability dispersion of the estimator outside an ε-neighbourhood of $T(P)$.

The inaccuracy rate of $\{T_n\}$ at P is defined to be

$$(2.2)\qquad \begin{aligned} e(\varepsilon, T) &= -\lim_{n\to\infty} n^{-1}\log P\{|T_n - T(P)| > \varepsilon\} \\ &= -\lim_{n\to\infty} n^{-1}\log a_n(\varepsilon, T), \end{aligned}$$

provided that the limit exists. If $\{T_n\}$ is asymptotically normal, i.e.,

$$(2.3)\qquad \mathscr{L}\big(n^{1/2}\{T_n - T(P)\}|P\big) \to N\big(0, \sigma^2(P)\big),$$

then for each $c > 0$,

$$\lim_{n\to\infty} a_n(n^{-1/2}c, T) = 2\Phi\big(-c\sigma^{-1}(P)\big),$$

where Φ denotes the standard normal distribution function. Hence, we obtain under (2.3),

$$(2.4)\qquad -\lim_{c\to\infty}\lim_{n\to\infty} c^{-2}\log a_n(n^{-1/2}c, T) = \{2\sigma^2(P)\}^{-1}.$$

Noting that $n^{-1/2}c \to 0$ for fixed $c > 0$ as $n \to \infty$, the argument of $a_n(\cdot, T)$ in (2.4) tends to zero, while the factor before the logarithm is kept fixed as $n \to \infty$. More or less inverting the order of taking limits on the left-hand side of (2.4) yields

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty} (n\varepsilon^2)^{-1}\log a_n(\varepsilon, T),$$

where the argument of $a_n(\cdot, T)$ is kept fixed and the factor before the logarithm tends to zero when taking the first limit. In typical cases, indeed, we have

$$(2.5)\qquad -\lim_{\varepsilon\to 0}\lim_{n\to\infty} (n\varepsilon^2)^{-1}\log a_n(\varepsilon, T) = \{2\sigma^2(P)\}^{-1}.$$

To prove (2.5) rigourously, we first consider for fixed $\varepsilon > 0$ the large deviation probability (2.2). A very general large deviation theorem for statistical functionals is presented in Groeneboom, Oosterhoff and Ruymgaart (1979). To formulate this basic theorem we introduce the τ-topology on Λ. This topology τ of convergence on all Borel sets is the coarsest topology for which the map $Q \to Q(B)$, $Q \in \Lambda$, is continuous for all $B \in \mathscr{B}$. In this topology a sequence of probability measures $\{Q_n\}$ in Λ converges to a probability measure $Q \in \Lambda$ iff $\lim_{n\to\infty}\int_S f\,dQ_n = \int_S f\,dQ$ for each bounded $\mathscr{B}$-measurable function $f\colon S \to \mathbb{R}$. For more details about the choice of this topology on Λ, we refer to Groeneboom, Oosterhoff and Ruymgaart (1979).

Further, we need the concept of the Kullback–Leibler information number $K(Q, P)$, which is defined by

$$(2.6)\qquad K(Q, P) = \begin{cases} E_Q\log\dfrac{dQ}{dP}(X), & \text{if } Q \ll P, \\ \infty, & \text{otherwise.} \end{cases}$$

If Ω is a subset of Λ and $P \in \Lambda$, we define

$$K(\Omega, P) = \inf_{Q \in \Omega} K(Q, P). \tag{2.7}$$

The inaccuracy rate can now be given in terms of Kullback–Leibler numbers.

PROPOSITION 2.1 (Groeneboom, Oosterhoff and Ruymgaart). *Let $P \in \Lambda$ and let $T\colon \Lambda \to \mathbb{R}$ be a functional which is τ-continuous at each $Q \in \Gamma = \{R \in \Lambda\colon K(R, P) < \infty\}$. Define*

$$\Omega_\varepsilon = \{Q \in \Lambda\colon |T(Q) - T(P)| \geq \varepsilon\}. \tag{2.8}$$

Then, if the function $t \to K(\Omega_t, P)$, $t \in \mathbb{R}$, is continuous from the right at $t = \varepsilon$ and if $\{u_n\}$ is a sequence of real numbers such that $\lim_{n \to \infty} u_n = 0$,

$$-\lim_{n \to \infty} n^{-1} \log P\{|T(\hat{P}_n) - T(P)| \geq \varepsilon + u_n\} = K(\Omega_\varepsilon, P). \tag{2.9}$$

In view of the preceding proposition, to obtain (2.5), we need general conditions ensuring

$$K(\Omega_\varepsilon, P) = \{2\sigma^2(P)\}^{-1}\varepsilon^2 + o(\varepsilon^2) \quad \text{as } \varepsilon \to 0. \tag{2.10}$$

Such a result is given in the following theorem and its corollary.

THEOREM 2.2. *Let ψ be a function satisfying*

$$\int e^{r\psi(x)}\, dP(x) < \infty \quad \text{for some } r > 0 \text{ and } E_P(\psi - E_P\psi)^2 > 0. \tag{2.11}$$

Let $g_i(\varepsilon) = \varepsilon + o(\varepsilon)$ as $\varepsilon \to 0$, $i = 1, 2, 3$. Then

$$\begin{aligned} &\inf\left\{K(Q, P)\colon \int \psi(x)\, dQ(x) - \int \psi(x)\, dP(x) \geq g_1(\varepsilon)\right\} \\ &\quad = \left\{2E_P(\psi - E_P\psi)^2\right\}^{-1}\varepsilon^2 + o(\varepsilon^2) \quad \text{as } \varepsilon \to 0. \end{aligned} \tag{2.12}$$

Hence, if

$$\begin{aligned} &\inf\{K(Q, P)\colon T(Q) - T(P) \geq g_2(\varepsilon)\} \\ &\quad = \inf\left\{K(Q, P)\colon \int \psi(x)\, dQ(x) - \int \psi(x)\, dP(x) \geq g_3(\varepsilon)\right\}, \end{aligned} \tag{2.13}$$

with ψ satisfying (2.11), then

$$\begin{aligned} &\inf\{K(Q, P)\colon T(Q) - T(P) \geq g_2(\varepsilon)\} \\ &\quad = \left\{2E_P(\psi - E_P\psi)^2\right\}^{-1}\varepsilon^2 + o(\varepsilon^2) \quad \text{as } \varepsilon \to 0. \end{aligned} \tag{2.14}$$

COROLLARY 2.3. *Let $g_i(\varepsilon) = \varepsilon + o(\varepsilon)$ as $\varepsilon \to 0$, $i = 1, 2$. Suppose that*

$$\begin{aligned} &\inf\{K(Q, P)\colon |T(Q) - T(P)| \geq g_1(\varepsilon)\} \\ &\quad = \inf\left\{K(Q, P)\colon \left|\int \psi(x)\, dQ(x) - \int \psi(x)\, dP(x)\right| \geq g_2(\varepsilon)\right\} \end{aligned} \tag{2.15}$$

with ψ *satisfying*

$$\int e^{r\psi(x)}\,dP(x) < \infty, \tag{2.16}$$
$$\int e^{-r\psi(x)}\,dP(x) < \infty \quad \textit{for some } r > 0 \textit{ and } E_P(\psi - E_P\psi)^2 > 0.$$

Then

$$K\big(\Omega_{g_1(\varepsilon)}, P\big) = \big\{2E_P(\psi - E_P\psi)^2\big\}^{-1}\varepsilon^2 + o(\varepsilon^2) \quad \textit{as } \varepsilon \to 0. \tag{2.17}$$

In particular, if $g_1(\varepsilon) = \varepsilon$ *and* (2.15) *and* (2.16) *hold, then*

$$K(\Omega_\varepsilon, P) = \big\{2E_P(\psi - E_P\psi)^2\big\}^{-1}\varepsilon^2 + o(\varepsilon^2) \quad \textit{as } \varepsilon \to 0. \tag{2.18}$$

REMARK 2.1. Note that the right-hand sides of (2.13) and (2.15) correspond to large deviation probabilities of $\Sigma_{i=1}^{n}\psi(X_i)$ (cf. also Remark 2.4).

REMARK 2.2. It is seen from Lemma 4.1 that under (2.11), $K(Q, P) = \infty$ for all Q for which $\int\psi\,dQ$ does not exist. Therefore, in (2.12), (2.13) and (2.15) we may or may not include, in the set of Q's over which the infimum has to be taken, those Q's for which $\int\psi\,dQ$ does not exist.

The general ideal behind Theorem 2.2 and Corollary 2.3 is that if T is differentiable in a suitable way with influence curve ψ, then $T(Q) - T(P) \approx \int\psi\,dQ - \int\psi\,dP$, implying (2.13) and (2.15). Moreover, in that case, the asymptotic variance of $T(\hat{P}_n)$ will be $E_P(\psi - E_P\psi)^2$ and, hence, (2.18) then yields (2.10).

Next we consider two forms of differentiability, Fréchet- and Hadamard-type differentiability, to make the preceding ideas more precise. Fréchet-type differentiability is used, e.g., by Boos (1979). Let V be the set of signed measures λ on $\mathscr{B}$, which are absolutely continuous with respect to P and satisfy $\lambda(S) = 0$. With the usual addition and scalar multiplication of measures, V is a real linear space. Defining $\|\cdot\|$: $V \to [0, \infty)$ by $\|\lambda\| = |\lambda|(S)$ for $\lambda \in V$, where $|\lambda|$ denotes total variation, it is easily seen that $\|\cdot\|$ is a norm on V. Associating $f \in W = \{f \in L_1(S, \mathscr{B}, P)\colon \int f\,dP = 0\}$, with $\lambda \in V$ by

$$\lambda(E) = \int_E f\,dP, \tag{2.19}$$

we obtain a norm-isomorphism between V and W. Using the Hahn–Banach extension theorem and the correspondence between the conjugate space of L_1 and L_∞, it is seen by this identification that a continuous linear functional T' on V can be represented by a bounded function ψ, i.e.,

$$T'(\lambda) = \int f\psi\,dP = \int\psi\,d\lambda, \qquad \lambda \in V. \tag{2.20}$$

Now we consider the first form of differentiability which will be discussed. A functional T: $\Lambda \to \mathbb{R}$ is called *Fréchet-type differentiable* at $P \in \Lambda$ if there exists a continuous linear functional T_P': $V \to \mathbb{R}$ or, equivalently, a bounded function ψ, such that

$$\lim_{\|Q-P\|\to 0} \frac{T(Q) - T(P) - T_P'(Q-P)}{\|Q-P\|} = 0 \tag{2.21}$$

or

$$\lim_{\|Q-P\|\to 0} \frac{T(Q) - T(P) - \int\psi\, d(Q-P)}{\|Q-P\|} = 0, \tag{2.22}$$

uniformly for $Q \in \{\tilde{Q} \in \Lambda: \tilde{Q} \ll P\}$.

As a second notion of differentiability we consider Hadamard-type differentiability. A functional T: $\Lambda \to \mathbb{R}$ is called *Hadamard-type differentiable* at $P \in \Lambda$ if there exists a continuous linear functional T_P': $V \to \mathbb{R}$ or, equivalently, a bounded function ψ such that for any *compact subset* C of V

$$\lim_{\|Q-P\|\to 0} \frac{T(Q) - T(P) - T_P'(Q-P)}{\|Q-P\|} = 0 \tag{2.23}$$

or

$$\lim_{\|Q-P\|\to 0} \frac{T(Q) - T(P) - \int\psi\, d(Q-P)}{\|Q-P\|} = 0 \tag{2.24}$$

uniformly for $Q \in \{Q \in \Lambda: Q \ll P, (Q-P)/\|Q-P\| \in C\}$.

The following result shows that a slightly weaker condition than Fréchet-type differentiability (ψ may be unbounded) yields the required expansions.

THEOREM 2.4. *Suppose that*

$$\lim_{\|Q-P\|\to 0} \frac{T(Q) - T(P) - \int\psi\, d(Q-P)}{\|Q-P\|} = 0 \tag{2.25}$$

uniformly for $Q \in \{Q \in \Lambda: Q \ll P, \int|\psi|\, dQ < \infty\}$ *for some* ψ *satisfying* (2.16). *Then* (2.14), (2.17) *and* (2.18) *hold for each*

$$g_i(\varepsilon) = \varepsilon + o(\varepsilon) \quad \textit{as } \varepsilon \to 0,\ i = 1,2.$$

If T is Fréchet-type differentiable, then ψ is bounded and, hence, ψ satisfies (2.16) unless ψ is P-a.e. constant.

COROLLARY 2.5. *If T is Fréchet-type differentiable with nonconstant influence curve $\psi[P]$, then* (2.14), (2.17) *and* (2.18) *hold for each*

$$g_i(\varepsilon) = \varepsilon + o(\varepsilon) \quad \textit{as } \varepsilon \to 0,\ i = 1,2.$$

Hadamard-type differentiability is not quite enough to obtain the required expansions. We need some small extra conditions (cf. also Remark 2.3). Again the influence curve may be unbounded.

THEOREM 2.6. *Suppose that for some compact subset C_1 of V, independent of ε, we have*

$$\begin{aligned}&\inf\{K(Q,P)\colon T(Q)-T(P)\geq\varepsilon\}\\ &\quad=\inf\{K(Q,P)\colon (Q-P)/\|Q-P\|\in C_1,\\ &\qquad\qquad T(Q)-T(P)\geq\varepsilon\}(1+o(1))\end{aligned}\tag{2.26}$$

and for some compact subset C_2 of V, independent of ε,

$$\begin{aligned}&\inf\{K(Q,P)\colon T(Q)-T(P)\leq-\varepsilon\}\\ &\quad=\inf\{K(Q,P)\colon (Q-P)/\|Q-P\|\in C_2,\\ &\qquad\qquad T(Q)-T(P)\leq-\varepsilon\}(1+o(1))\end{aligned}\tag{2.27}$$

as $\varepsilon\to 0$. Further, suppose that for all compact subsets C of V,

$$\lim_{\|Q-P\|\to 0}\frac{T(Q)-T(P)-\int\psi\,d(Q-P)}{\|Q-P\|}=0,\tag{2.28}$$

uniformly for $Q\in\{Q\in\Lambda\colon Q\ll P,\ (Q-P)/\|Q-P\|\in C,\ \int|\psi|\,dQ<\infty\}$ for some ψ satisfying (2.16). *Then* (2.14), (2.17) *and* (2.18) *hold for each $g_i(\varepsilon)=\varepsilon+o(\varepsilon)$ as $\varepsilon\to 0$, $i=1,2$.*

COROLLARY 2.7. *If T is Hadamard-type differentiable with nonconstant influence curve $\psi[P]$ and if* (2.26) *and* (2.27) *are satisfied, then* (2.14), (2.17) *and* (2.18) *hold for each $g_i(\varepsilon)=\varepsilon+o(\varepsilon)$ as $\varepsilon\to 0$, $i=1,2$.*

Theorems 2.4, 2.6 and Corollaries 2.5 and 2.7 state that for a large class of statistics, the complicated expression $K(\Omega_\varepsilon,P)$ behaves locally for $\varepsilon\to 0$ as $\{2\sigma^2(P)\}^{-1}\varepsilon^2$. As mentioned earlier, uniformity is essential here, since we are dealing with exact measures of performance. Therefore, the concepts of Fréchet-type and Hadamard-type differentiability [or the conditions (2.25) and (2.28)], which imply uniformity, are very natural in this context.

REMARK 2.3. Suppose that the first infimum in (2.26) is attained by some probability measure Q_ε for $0<\varepsilon\leq\varepsilon_0$. [This may be shown in many cases by applying Lemma 3.2 of Groeneboom, Oosterhoff and Ruymgaart (1979).] Further suppose that $\|(Q_\varepsilon-P)/\|Q_\varepsilon-P\|-\lambda_0\|\to 0$ as $\varepsilon\to 0$ for some $\lambda_0\in V$. Then, condition (2.26) may be omitted, which can be seen by a slight modification of the proof of Theorem 2.6 (cf. Remark 4.4). Usually in this case the set $C_1=\{\lambda_0\}\cup\{(Q_\varepsilon-P)/\|Q_\varepsilon-P\|\colon 0<\varepsilon\leq\varepsilon_0\}$ is compact and then Theorem 2.6 can be applied directly. Of course a similar statement holds w.r.t. (2.27).

REMARK 2.4. In local theory one usually has the following approach. A statistical functional is approximated by a linear statistical functional leading, when applied to the empirical probability measure, to a sum of i.i.d. random variables. Then it is shown that the remainder terms are small (in probability) and, hence, the distribution function of the statistic is approximated by the distribution function of a sum of i.i.d. random variables, which can be handled. In large deviation theory such an approach fails, since the remainder terms have to be shown to be exponentially small [c.f. the example in Section 5(a) of Jurečková and Kallenberg (1987)]. However, the same *idea* as in local theory can be used here. The large deviation probability of the statistic (and not the statistic itself) is approximated locally by the large deviation probability of a sum of i.i.d. random variables $\psi(X_i)$, which can be handled [cf. also (2.13), (2.15) and Remark 2.1].

3. Examples and applications. In this section, we present some direct applications of the theory developed in Section 2. It is not our aim in this paper to give the most general formulations for those applications or to give an exhaustive overview of the applicability of the theory. We only show that some results previously obtained by rather hard calculus now easily follow using the differential approach.

Thus far the theory is described in an estimation framework. As mentioned before, analogous results hold in testing theory. To show this, we start this section with an application of our results in testing theory, obtaining the local behaviour of the exact Bahadur slopes of linear rank statistics [cf. Woodworth (1970)].

Let $R_1, \dots, R_n$ be the ranks of n random variables $Z_1, \dots, Z_n$. Consider the null hypothesis that $(R_1, \dots, R_n)$ is equally likely to be any of the $n!$ permutations of $(1, \dots, n)$. A linear rank statistic is one of the form

$$(3.1) \qquad T_n = \sum_{i=1}^{n} a_n(R_i/(n+1),\, i/(n+1)),$$

where $a_n(u, v)$ is a function on the unit square. Let

$$(3.2) \qquad \mathscr{H} = \left\{h\colon h \geq 0,\, \int h(u,v)\,du = 1 = \int h(u,v)\,dv\right\}$$

be the set of all bivariate densities on the unit square with uniform marginals. It will be assumed that the sequence $\{a_n\}$ determining T_n in (3.1) satisfies Woodworth's (1970) property A, i.e.,

(i) for each n, a_n is constant over the rectangles $\{i - 1 \leq nu < i,\ j - 1 \leq nv < j\}$, $1 \leq i,\ j \leq n$;

(ii) there exists a function a over the unit square such that

$$\sup_{h \in \mathscr{H}} \left| \iint (a_n - a)h \right| \to 0 \quad \text{as } n \to \infty.$$

Without loss of generality we assume

$$(3.3) \qquad \iint a(u,v)\,du\,dv = 0.$$

PROPOSITION 3.1 (Woodworth). *Let $\{T_n\}$ be a sequence of linear rank statistics satisfying property* A *and let $\{\varepsilon_n\}$ be a sequence of real numbers with*

$\lim_{n\to\infty}\varepsilon_n = \varepsilon$. *Then*

$$\lim_{n\to\infty} -n^{-1}\log P\{T_n \geq n\varepsilon_n\} = I(\varepsilon, a), \qquad 0 < \varepsilon < \varepsilon(a), \tag{3.4}$$

where

$$\varepsilon(a) = \sup\left\{\iint ah \colon h \in \mathscr{H}\right\}$$

and for $0 < \varepsilon < \varepsilon(a)$,

$$I(\varepsilon, a) = \inf\left\{\iint h\log h \colon \iint ah \geq \varepsilon,\ h \in \mathscr{H}\right\}. \tag{3.5}$$

The main point in evaluating the local behaviour of the exact Bahadur slope of $\{T_n\}$ is to investigate $I(\varepsilon, a)$ as $\varepsilon \downarrow 0$. In the following theorem the local behaviour of $I(\varepsilon, a)$ is established.

THEOREM 3.2. *Let*

$$\tilde{a}(u,v) = a(u,v) - \int a(u,y)\,dy - \int a(x,v)\,dx + \iint a(x,y)\,dx\,dy$$

and let

$$c = \iint \tilde{a}^2(u,v)\,du\,dv. \tag{3.6}$$

Suppose that

$$\iint e^{r\tilde{a}(u,v)}\,du\,dv < \infty \quad \textit{for some } r > 0 \textit{ and } c > 0. \tag{3.7}$$

Then

$$I(\varepsilon, a) = \tfrac{1}{2}\varepsilon^2/c + o(\varepsilon^2) \quad \textit{as } \varepsilon \downarrow 0. \tag{3.8}$$

The proof of Theorem 3.2, which is based on Theorem 2.2, is in Section 4. Woodworth [(1970), pages 261–262] stated a sufficient condition for the validity of (3.8). However, this condition is rather impracticable. For the independence problem, which is a special case of the preceding setup, Kremer (1981) proved (3.8) under conditions which exclude, e.g., unbounded score-generating functions. It is obvious that condition (3.7) holds for *all* bounded functions; moreover, it also holds, for instance, for $a(u,v) = \Phi^{-1}(u)\Phi^{-1}(v)$ corresponding to the normal-scores correlation coefficient. To illustrate more explicitly how Theorem 3.2 can be used in deriving the local behaviour of the exact Bahadur slope, we consider the following example.

EXAMPLE 3.1. Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be i.i.d. random vectors each distributed as $H(x, y)$, where H is a continuous distribution function with marginals F and G, respectively. For every n, let $(R_{n1}, \dots, R_{nn})$ and $(S_{n1}, \dots, S_{nn})$ be the ranks in $(X_1, \dots, X_n)$ and $(Y_1, \dots, Y_n)$, respectively. The testing problem is $H(x, y) = F(x)G(y)$ for all x, y against $H(x, y) \geq F(x)G(y)$ for all x, y with at least one inequality strict. Consider the rank statistic

$$T_n = \sum_{i=1}^{n} B_{1n}(R_{ni}) B_{2n}(S_{ni}), \tag{3.9}$$

where $B_{in}(1 + [nu]) = b_{ni}(u)$ and $\int(b_{in} - b_i)^2 \, du \to 0$, $i = 1, 2$. ($[x]$ denotes the largest integer less than or equal to x.) Moreover, $\int b_i \, du = 0$, $\int b_i^2 \, du = 1$ and $\int\int \exp\{r b_1 b_2\} \, du \, dv < \infty$ for some $r > 0$. Let $H_j(x, y)$ be a sequence of alternatives tending to the null hypothesis in the sense that

$$\sup_{x, y} |H_j(x, y) - H_0(x, y)| \to 0 \quad \text{as } j \to \infty$$

for some $H_0(x, y) = F_0(x) G_0(y)$ in the null hypothesis. Since for each fixed j, as $n \to \infty$,

$$T_n / n \to_{H_j} \int\int b_1(F_j(x)) b_2(G_j(y)) \, dH_j(x, y) = \varepsilon_j, \quad \text{say},$$

where F_j and G_j are the marginals of H_j, it follows that the exact Bahadur slope of T_n at H_j equals $2I(\varepsilon_j, a)$ with $a = b_1 b_2$. Application of Theorem 3.2 now yields that the exact Bahadur slope $2I(\varepsilon_j, a)$ satisfies

$$2I(\varepsilon_j, a) = \left\{ \int\int b_1(F_j(x)) b_2(G_j(y)) \, dH_j(x, y) \right\}^2 + o(\varepsilon_j^2) \quad \text{as } j \to \infty. \tag{3.10}$$

For instance, if H_j is the distribution function of a bivariate normal random variable with correlation coefficient ρ_j tending to zero if $j \to \infty$, then the exact Bahadur slope of T_n satisfies

$$2I(\varepsilon_j, a) = \left\{ \rho_j \int\int b_1(u) b_2(v) \Phi^{-1}(u) \Phi^{-1}(v) \, du \, dv \right\}^2 + o(\rho_j^2)$$

as $j \to \infty$. Hence, the limit of the exact Bahadur slope of T_n divided by ρ_j^2 equals the Pitman efficacy of T_n, implying local equivalence of both concepts of efficiency. Similar results can be obtained for general linear rank tests of the one-sample symmetry problem and the k-sample problem.

For a second class of applications, we return to estimation theory, especially to L-estimators. Let $X_1, X_2, \ldots$ be real valued i.i.d. random variables with distribution function F. For a distribution function G, its inverse G^{-1} is defined in the usual way by $G^{-1}(s) = \inf\{x \in \mathbb{R}: G(x) \geq s\}$. Suppose $J: [0, 1] \to \mathbb{R}$ is a Lebesgue-integrable function, i.e.,

$$\int_0^1 |J(s)| \, ds < \infty.$$

We consider linear combinations of order statistics of the form $T_n = T(\hat{P}_n)$ with

$$T(Q) = \int_0^1 J(s) G^{-1}(s) \, ds, \tag{3.11}$$

where G denotes the distribution function corresponding to Q. The influence function of T is derived in Huber (1981), for example. For these statistics, the following large deviation result is obtained by a slight modification of Theorem 6.1 in Groeneboom, Oosterhoff and Ruymgaart (1979).

PROPOSITION 3.3 (Groeneboom, Oosterhoff and Ruymgaart). *Let P be a probability measure with continuous distribution function F and let $J \in L_1[0,1]$ have support in $[\alpha, 1-\alpha]$ for some $\alpha > 0$. Further assume that $J \geq 0$ on an interval (γ, δ) and $\int_\gamma^\delta J(s)\,ds > 0$. Then for each $\varepsilon > 0$,*

$$e(\varepsilon, T) = K(\Omega_\varepsilon, P), \tag{3.12}$$

where

$$\Omega_\varepsilon = \{Q \in \Lambda \colon |T(Q) - T(P)| \geq \varepsilon\}. \tag{3.13}$$

To obtain Fréchet-type differentiability of T w.r.t. the uniform topology and, hence, also w.r.t. the total variation metric, we apply Theorem 1 in Boos (1979).

PROPOSITION 3.4 (Boos). *Let P be a probability measure on $\mathbb{R}$ with distribution function F. Suppose J has support in $[\alpha, 1-\alpha]$ for some $\alpha > 0$. Further assume that J is bounded and continuous a.e. Lebesgue and a.e. F^{-1}. Then T is Fréchet-type differentiable at P w.r.t. the uniform topology with differential*

$$T_P'(Q - P) = -\int_{-\infty}^{\infty} \{G(x) - F(x)\} J(F(x))\,dx, \tag{3.14}$$

where G denotes the distribution function of Q. Hence, (2.25) *holds with*

$$\psi(x) = \int_{-\infty}^{x} J(F(y))\,dy - \int_{-\infty}^{\infty} \{1 - F(y)\} J(F(y))\,dy, \tag{3.15}$$

which is a bounded function.

Combination of Theorem 2.4, Proposition 3.3 and Proposition 3.4 now yields

THEOREM 3.5. *Let P be a probability measure with a continuous distribution function F. Let J have support in $[\alpha, 1-\alpha]$ for some $\alpha > 0$. Further assume that J is bounded and continuous a.e. Lebesgue and a.e. F^{-1} and that $J \geq 0$ on an interval (γ, δ) and $\int_\gamma^\delta J(s)\,ds > 0$. Then* (2.5) *holds, i.e.,*

$$-\lim_{\varepsilon \to 0} \lim_{n \to \infty} (n\varepsilon^2)^{-1} \log a_n(\varepsilon, T) = \{2\sigma^2(P)\}^{-1}, \tag{3.16}$$

where

$$\sigma^2(P) = E_P(\psi - E_P\psi)^2$$

with ψ given by (3.15).

This result generalizes Fu's (1980) Theorem 4.2. Moreover, the proof presented here is much shorter and places the result in a more general context.

EXAMPLE 3.2. The α-trimmed mean for $0 < \alpha < \frac{1}{2}$ is an L-estimator with $J(s) = (1 - 2\alpha)^{-1}$ for $s \in (\alpha, 1-\alpha)$ and $J(s) = 0$ elsewhere. If the distribution function F is continuous and has unique quantiles $F^{-1}(\alpha)$ and $F^{-1}(1-\alpha)$, then writing $X_{(1)} < \cdots < X_{(n)}$ for the order statistics,

$$-\lim_{\varepsilon\to 0}\lim_{n\to\infty}(n\varepsilon^2)^{-1}\log P\left\{\left|\frac{1}{n-2[\alpha n]}\sum_{i=[\alpha n]+1}^{n-[\alpha n]}X_{(i)} - \int_\alpha^{1-\alpha}\frac{F^{-1}(s)}{1-2\alpha}\,ds\right| > \varepsilon\right\}$$
$$= -\frac{1}{2\sigma^2}$$

with

$$\sigma^2 = E_P(\psi - E_P\psi)^2,$$

where ψ denotes the influence curve of the α-trimmed mean.

Similar applications can be made, for instance, for M-estimators. Some results on Fréchet-type differentials of M-estimators are given in Boos and Serfling (1980).

4. Proofs. Before starting the proofs of the main results of Section 2 and Theorem 3.2, we present the following lemma, which may be of independent interest.

LEMMA 4.1. *Let ψ be a function satisfying*

$$\int e^{r\psi(x)}\,dP(x) < \infty \quad \textit{for some } r > 0. \tag{4.1}$$

Then

$$\int \psi^+(x)\,dQ(x) = \infty \quad \Rightarrow \quad K(Q,P) = \infty, \tag{4.2}$$

where $\psi^+(x) = \max(0, \psi(x))$.

PROOF. Since $K(Q,P) = \infty$ if Q is not absolutely continuous w.r.t. P, we may assume $Q \ll P$. Noting that for $r > 0$,

$$e^{r\psi(x)} \le e^{r\psi^+(x)} \le e^{r\psi(x)} + 1,$$

hence,

$$\int e^{r\psi(x)}\,dP(x) < \infty \quad \Leftrightarrow \quad \int e^{r\psi^+(x)}\,dP(x) < \infty.$$

Writing $f(x) = dQ/dP(x)$, we have by (4.1) for some $r > 0$,

$$\infty > \int e^{r\psi^+(x)}\,dP(x) \ge \int_{f>0} e^{r\psi^+}\,dP = \int_{f>0} e^{r\psi^+ - \log f}\,dQ$$
$$\ge \int_{f>0}(1 + r\psi^+ - \log f)\,dQ = \int (1 + r\psi^+ - \log f)\,dQ.$$

Therefore,

$$\int \psi^+ \, dQ(x) = \infty \quad \Rightarrow \quad \int -\log f \, dQ = -K(Q, P) = -\infty. \qquad \square$$

As an immediate consequence of Lemma 4.1 we obtain Corollary 4.2.

COROLLARY 4.2. *Let ψ be a function satisfying*

$$\int e^{r\psi(x)} \, dP(x) < \infty \quad \textit{and} \quad \int e^{-r\psi(x)} \, dP(x) < \infty \quad \textit{for some } r > 0.$$

Then

$$\int |\psi(x)| \, dQ(x) = \infty \quad \Rightarrow \quad K(Q, P) = \infty.$$

PROOF OF THEOREM 2.2. Let $g_1(\varepsilon) = \varepsilon + o(\varepsilon)$ as $\varepsilon \to 0$ and let ψ satisfy (2.11). Without loss of generality assume $\int \psi \, dP = 0$. By Theorem 4.2 in Bahadur (1971) and Lemma 4.1, we have

$$\inf\left\{K(Q, P): \int \psi(x) \, dQ(x) \geq g_1(\varepsilon)\right\} = -\log \inf\{\phi_\varepsilon(t): t \geq 0\}, \tag{4.3}$$

where

$$\phi_\varepsilon(t) = E_P \exp(t\{\psi - g_1(\varepsilon)\}) \tag{4.4}$$

and $\inf\{K(Q, P): Q \in \varnothing\} = \infty$. Writing $\beta = \sup\{t: \phi_\varepsilon(t) < \infty\}$, (2.11) implies

$$0 < r \leq \beta \leq \infty, \qquad \phi_\varepsilon'(0+) = E_P\{\psi - g_1(\varepsilon)\} = -g_1(\varepsilon) < 0 \tag{4.5}$$

for sufficiently small $\varepsilon > 0$, and for all $0 < b < \beta$,

$$\phi_\varepsilon'(b) = \phi_\varepsilon'(0+) + b\phi_\varepsilon''(\xi_{b,\varepsilon}) = -g_1(\varepsilon) + b\phi_\varepsilon''(\xi_{b,\varepsilon}) \tag{4.6}$$

for some $0 < \xi_{b,\varepsilon} < b$. By the dominated convergence theorem,

$$\lim_{\varepsilon \downarrow 0} \phi_\varepsilon''(t_\varepsilon) = E_P\psi^2 > 0 \quad \text{if } t_\varepsilon \downarrow 0 \text{ as } \varepsilon \downarrow 0. \tag{4.7}$$

Hence, by (4.6) and (4.7),

$$\phi_\varepsilon'\left(2\{E_P\psi^2\}^{-1}\varepsilon\right) > 0 \tag{4.8}$$

for sufficiently small $\varepsilon > 0$. In view of (4.5) and (4.8), the standard conditions of Bahadur [(1971), pages 3–4] are satisfied for sufficiently small $\varepsilon > 0$, implying

$$\inf\{\phi_\varepsilon(t): t \geq 0\} = \phi_\varepsilon(\tau_\varepsilon), \tag{4.9}$$

where τ_ε is defined as the unique solution of $\phi_\varepsilon'(\tau_\varepsilon) = 0$. By (4.6) and (4.7), we obtain

$$\tau_\varepsilon = \varepsilon\{E_P\psi^2\}^{-1} + o(\varepsilon) \quad \text{as } \varepsilon \to 0. \tag{4.10}$$

Hence, for some $0 < \xi_\varepsilon < \tau_\varepsilon$,

$$\begin{aligned}\phi_\varepsilon(\tau_\varepsilon) &= \phi_\varepsilon(0) + \tau_\varepsilon\phi_\varepsilon'(0+) + \tfrac{1}{2}\tau_\varepsilon^2\phi_\varepsilon''(\xi_\varepsilon)\\ &= 1 - \tfrac{1}{2}\varepsilon^2\{E_P\psi^2\}^{-1} + o(\varepsilon^2)\end{aligned} \tag{4.11}$$

and, therefore, by (4.3), (4.9) and (4.11),

$$\inf\left\{K(Q,P): \int\psi(x)\,dQ(x) \geq g_1(\varepsilon)\right\} = \tfrac{1}{2}\varepsilon^2\{E_P\psi^2\}^{-1} + o(\varepsilon^2)$$

as $\varepsilon \to 0$. □

REMARK 4.1. It is seen in the preceding proof that for sufficiently small $\varepsilon > 0$ the infimum in (2.12) is attained by the probability measure

$$dQ_\varepsilon(x) = \frac{\exp(\tau_\varepsilon\{\psi(x) - E_P\psi - g_1(\varepsilon)\})\,dP(x)}{E_P\exp(\tau_\varepsilon\{\psi - E_P\psi - g_1(\varepsilon)\})}, \tag{4.12}$$

where τ_ε is the unique solution of

$$E_P[\{\psi - E_P\psi - g_1(\varepsilon)\}\exp(\tau_\varepsilon\{\psi - E_P\psi - g_1(\varepsilon)\})] = 0 \tag{4.13}$$

[cf. also Hoeffding (1965), Lemma 1, and Csiszár (1975), Theorem 3.1]. Furthermore, it is seen in (4.10) that

$$\tau_\varepsilon = \varepsilon\{E_P\psi^2\}^{-1} + o(\varepsilon) \quad \text{as } \varepsilon \to 0,$$

a result which may be of independent interest.

PROOF OF COROLLARY 2.3. By (2.7), (2.8) and (2.15) we have

$$\begin{aligned}K(\Omega_{g_1(\varepsilon)}, P) &= \inf\left\{K(Q,P): \left|\int\psi\,dQ - \int\psi\,dP\right| \geq g_2(\varepsilon)\right\}\\ &= \min\left[\inf\left\{K(Q,P): \int\psi\,dQ - \int\psi\,dP \geq g_2(\varepsilon)\right\},\right.\\ &\qquad \left.\inf\left\{K(Q,P): \int(-\psi)\,dQ - \int(-\psi)\,dP \geq g_2(\varepsilon)\right\}\right].\end{aligned}$$

Application of (2.12) now yields the result. □

PROOF OF THEOREM 2.4. Let T satisfy (2.25) for some ψ satisfying (2.16). Without loss of generality, assume $\int\psi\,dP = 0$. We will first prove (2.14). Let $g_2(\varepsilon) = \varepsilon + o(\varepsilon)$ as $\varepsilon \to 0$. Define

$$\begin{aligned}g_3(\varepsilon) = g_2(\varepsilon) - \inf\Big\{&T(Q) - T(P) - \int\psi\,dQ:\\ &\int|\psi|\,dQ < \infty \text{ and } \|Q - P\| \leq 2\varepsilon\{E_P\psi^2\}^{-1/2}\Big\}.\end{aligned} \tag{4.14}$$

Since the convergence in (2.25) is required to be *uniform*, it follows that

$$g_3(\varepsilon) = \varepsilon + o(\varepsilon) \quad \text{as } \varepsilon \to 0.$$

Define the measure Q_ε by

$$dQ_\varepsilon(x) = \exp(\tau_\varepsilon\{\psi(x) - g_3(\varepsilon)\})\, dP(x)/E_P\exp(\tau_\varepsilon\{\psi - g_3(\varepsilon)\}),$$

where τ_ε is the unique solution of

$$E_P\{\psi - g_3(\varepsilon)\}\exp(\tau_\varepsilon\{\psi - g_3(\varepsilon)\}) = 0. \tag{4.15}$$

It is seen from the proof of Theorem 2.2 that Q_ε exists for sufficiently small $\varepsilon > 0$ (cf. also Remark 4.1). Moreover,

$$K(Q_\varepsilon, P) = \tfrac{1}{2}\varepsilon^2\{E_P\psi^2\}^{-1} + o(\varepsilon^2) \quad \text{as } \varepsilon \to 0. \tag{4.16}$$

Using the inequality

$$\|Q - P\| \le \{2K(Q, P)\}^{1/2} \tag{4.17}$$

[cf., e.g., Kemperman (1969)], it follows that for sufficiently small $\varepsilon > 0$,

$$\|Q_\varepsilon - P\| \le 2\varepsilon\{E_P\psi^2\}^{-1/2} \tag{4.18}$$

and, hence, by (4.14) and (4.15),

$$T(Q_\varepsilon) - T(P) \ge \int \psi\, dQ_\varepsilon + g_2(\varepsilon) - g_3(\varepsilon) = g_2(\varepsilon).$$

Therefore [cf.(4.16)],

$$\inf\{K(Q, P): T(Q) - T(P) \ge g_2(\varepsilon)\} \le \tfrac{1}{2}\varepsilon^2\{E_P\psi^2\}^{-1} + o(\varepsilon^2) \tag{4.19}$$

as $\varepsilon \to 0$.

Next we consider measures $\tilde{Q}_\varepsilon$ such that

$$T(\tilde{Q}_\varepsilon) - T(P) \ge g_2(\varepsilon)$$

and

$$0 \le K(\tilde{Q}_\varepsilon, P) - \inf\{K(Q, P): T(Q) - T(P) \ge g_2(\varepsilon)\} \le o(\varepsilon^2) \quad \text{as } \varepsilon \to 0. \tag{4.20}$$

By (4.17) and (4.19), we have for sufficiently small $\varepsilon > 0$,

$$\|\tilde{Q}_\varepsilon - P\| \le 2\varepsilon\{E_P\psi^2\}^{-1/2}.$$

Moreover, $\int|\psi|\, d\tilde{Q}_\varepsilon < \infty$ by Corollary 4.2. Hence, writing

$$g_4(\varepsilon) = g_2(\varepsilon) + \inf\left\{\int \psi\, dQ - T(Q) + T(P): \int|\psi|\, dQ < \infty \text{ and } \|Q - P\| \le 2\varepsilon\{E_P\psi^2\}^{-1/2}\right\}, \tag{4.21}$$

we have for sufficiently small $\varepsilon > 0$,

$$\int \psi\, d\tilde{Q}_\varepsilon \ge g_4(\varepsilon).$$

Since the convergence in (2.25) is required to be *uniform* and $g_2(\varepsilon) = \varepsilon + o(\varepsilon)$, it follows that

$$g_4(\varepsilon) = \varepsilon + o(\varepsilon) \quad \text{as } \varepsilon \to 0.$$

By (2.12) we obtain

$$K\left(\tilde{Q}_\varepsilon, P\right) \geq \tfrac{1}{2}\varepsilon^2\left\{E_P\psi^2\right\}^{-1} + o(\varepsilon^2) \quad \text{as } \varepsilon \to 0. \tag{4.22}$$

Combination of (4.19), (4.20) and (4.22) yields (2.14).

Let $g_1(\varepsilon) = \varepsilon + o(\varepsilon)$. If (2.25) holds for (T, ψ), it also holds for $(-T, -\psi)$ with ψ or $-\psi$ satisfying (2.16). Hence,

$$\inf\{K(Q, P) \colon T(Q) - T(P) \leq -g_1(\varepsilon)\} = \left\{2E_P(\psi - E_P\psi)^2\right\}^{-1}\varepsilon^2 + o(\varepsilon^2)$$

as $\varepsilon \to 0$. Since

$$\begin{aligned} K\left(\Omega_{g_1(\varepsilon)}, P\right) = \min\bigl[&\inf\{K(Q, P) \colon T(Q) - T(P) \geq g_1(\varepsilon)\}, \\ &\inf\{K(Q, P) \colon T(Q) - T(P) \leq -g_1(\varepsilon)\}\bigr], \end{aligned}$$

(2.17) and (2.18) are easily obtained. □

PROOF OF THEOREM 2.6. Let T satisfy (2.28) for some ψ satisfying (2.16). Without loss of generality assume $\int\psi\, dP = 0$. Let $g_2(\varepsilon) = \varepsilon + o(\varepsilon)$ as $\varepsilon \to 0$. We will first prove

$$\limsup_{\varepsilon \to 0} \varepsilon^{-2}\inf\{K(Q, P) \colon T(Q) - T(P) \geq g_2(\varepsilon)\} \leq \tfrac{1}{2}\left(E_P\psi^2\right)^{-1}. \tag{4.23}$$

Choose $\delta > 0$. Define for sufficiently small $\varepsilon > 0$ the probability measure Q_ε by

$$dQ_\varepsilon(x) = \exp\left((1 + \delta)\varepsilon\psi(x)/E_P\psi^2\right) dP(x)/\gamma(\varepsilon),$$

with

$$\gamma(\varepsilon) = \int \exp\left((1 + \delta)\varepsilon\psi(x)/E_P\psi^2\right) dP(x).$$

Note that this can be done because (2.16) holds. Define the signed measure $\lambda \in V$ by

$$d\lambda = \psi\left(E_P|\psi|\right)^{-1} dP.$$

By dominated convergence, we have

$$\lim_{\varepsilon \to 0} \frac{\gamma(\varepsilon) - 1}{\varepsilon^2} = \frac{1}{2}(1 + \delta)^2\left(E_P\psi^2\right)^{-1},$$

$$\lim_{\varepsilon \to 0} \frac{\|Q_\varepsilon - P\|}{\varepsilon} = (1 + \delta)E_P|\psi|\left(E_P\psi^2\right)^{-1}$$

and

$$\lim_{\varepsilon \to 0} \left\| \frac{(Q_\varepsilon - P)}{\|Q_\varepsilon - P\|} - \lambda \right\| = 0.$$

Therefore, for sufficiently small $\varepsilon_0 > 0$, the set

$$C = \{\lambda\} \cup \left\{ \frac{(Q_\varepsilon - P)}{\|Q_\varepsilon - P\|} : 0 < \varepsilon \le \varepsilon_0 \right\} \quad (4.24)$$

is a compact subset of V. Again by dominated convergence we have

$$\lim_{\varepsilon \to 0} \varepsilon^{-2} K(Q_\varepsilon, P) = \lim_{\varepsilon \to 0} \varepsilon^{-2} \int \left(\frac{dQ_\varepsilon}{dP} \log \frac{dQ_\varepsilon}{dP} \right) dP = \frac{1}{2}(1 + \delta^2)\left(E_P \psi^2\right)^{-1}.$$

Application of (2.28) yields

$$T(Q_\varepsilon) - T(P) = \int \psi \, dQ_\varepsilon + o(\|Q_\varepsilon - P\|) = (1 + \delta)\varepsilon + o(\varepsilon) \ge g_2(\varepsilon)$$

for sufficiently small $\varepsilon > 0$. Hence,

$$\limsup_{\varepsilon \to 0} \varepsilon^{-2} \inf\{K(Q, P) : T(Q) - T(P) \ge g_2(\varepsilon)\} \le \lim_{\varepsilon \to 0} \varepsilon^{-2} K(Q_\varepsilon, P) = \tfrac{1}{2}(1 + \delta)^2 \left(E_P \psi^2\right)^{-1}. \quad (4.25)$$

Since $\delta > 0$ is arbitrarily chosen, (4.23) follows from (4.25). Similarly we obtain, replacing (T, ψ) by $(-T, -\psi)$,

$$\limsup_{\varepsilon \to 0} \varepsilon^{-2} \inf\{K(Q, P) : T(Q) - T(P) \le -g_2(\varepsilon)\} \le \tfrac{1}{2}\left(E_P \psi^2\right)^{-1}. \quad (4.26)$$

Now suppose that T also satisfies (2.26) for some compact subset C_1 of V. We have $\lim_{\varepsilon \to 0} g_2(\varepsilon) = 0$ and, hence,

$$\lim_{\varepsilon \to 0} \frac{\inf\{K(Q, P) : T(Q) - T(P) \ge g_2(\varepsilon), (Q - P)/\|Q - P\| \in C_1\}}{\inf\{K(Q, P) : T(Q) - T(P) \ge g_2(\varepsilon)\}} = 1.$$

Consider measures $\tilde{Q}_\varepsilon$ such that

$$(\tilde{Q}_\varepsilon - P)/\|\tilde{Q}_\varepsilon - P\| \in C_1,$$
$$T(\tilde{Q}_\varepsilon) - T(P) \ge g_2(\varepsilon)$$

and

$$0 \le K(\tilde{Q}_\varepsilon, P) - \inf\{K(Q, P) : T(Q) - T(P) \ge g_2(\varepsilon)\} \le o(\varepsilon^2)$$

as $\varepsilon \to 0$. Following the same line of argument as in the proof of Theorem 2.4, with (4.21) replaced by

$$g_4(\varepsilon) = g_2(\varepsilon) + \inf\left\{ \int \psi \, dQ - T(Q) + T(P) : \int |\psi| \, dQ < \infty, \right.$$
$$\left. \|Q - P\| \le 2\varepsilon\left(E_P \psi^2\right)^{-1/2} \text{ and } (Q - P)/\|Q - P\| \in C_1 \right\},$$

we obtain

$$\begin{aligned}(4.27)\qquad &\liminf_{\varepsilon\to 0}\varepsilon^{-2}\inf\{K(Q,P): T(Q)-T(P)\geq g_2(\varepsilon)\}\\ &\qquad\geq \liminf_{\varepsilon\to 0}\varepsilon^{-2}K\big(\tilde{Q}_\varepsilon, P\big)\geq \tfrac{1}{2}\big(E_P\psi^2\big)^{-1}.\end{aligned}$$

Combination of (4.23) and (4.27) yields (2.14).

Assuming that T also satisfies (2.27) for some compact subset C_2 of V, it similarly follows that

$$(4.28)\quad \liminf_{\varepsilon\to 0}\varepsilon^{-2}\inf\{K(Q,P): T(Q)-T(P)\leq -g_2(\varepsilon)\}\geq \tfrac{1}{2}\big(E_P\psi^2\big)^{-1}.$$

Combination of (4.25)–(4.28) now yields (2.17) and (2.18). □

REMARK 4.2. It is seen in the preceding proof that for proving (4.25) and (4.26) we only need condition (2.28). This can be used, e.g., if $E_P(\psi - E_P\psi)^2$ equals the Fisher-information, in which case the inequalities (4.27) and (4.28) may be obtained by optimality considerations [cf. Bahadur (1960)] and, hence, conditions (2.26) and (2.27) can be skipped.

REMARK 4.3. Note that we need (2.28) only for the following compact subsets of V: C as defined in (4.24), C_1 and C_2.

REMARK 4.4. Suppose that the first infimum in (2.26) is attained by some probability measure Q_ε for $0 < \varepsilon \leq \varepsilon_0$ and that

$$\|(Q_\varepsilon - P)/\|Q_\varepsilon - P\| - \lambda_0\| \to 0 \quad \text{as } \varepsilon\to 0$$

for some $\lambda_0 \in V$. To obtain (4.27), we now only need to show $\int\psi\, dQ_\varepsilon - T(Q_\varepsilon) + T(P) = o(\varepsilon)$ as $\varepsilon \to 0$. This is established by application of (2.28) with $C = \{\lambda_0\} \cup \{(Q_{\varepsilon_n} - P)/\|Q_{\varepsilon_n} - P\|\}$ for a suitable sequence $\{\varepsilon_n\}$ with $\varepsilon_n \to 0$ as $n\to\infty$.

PROOF OF THEOREM 3.2. Without loss of generality, assume

$$(4.29)\qquad \int a(u,v)\,du = 0 = \int a(u,v)\,dv,\qquad \iint a^2(u,v)\,du\,dv = 1.$$

By (2.12) we have, writing P for the uniform distribution on the unit square,

$$\begin{aligned}I(\varepsilon,a) &= \inf\left\{\iint h\log h: \iint ah\geq\varepsilon,\ h\in\mathscr{H}\right\}\\ (4.30)\qquad &\geq \inf\left\{\iint h\log h: \iint ah\geq\varepsilon,\ h\geq 0,\ \iint h = 1\right\}\\ &= \inf\left\{K(Q,P): \iint a\,dQ\geq\varepsilon\right\} = \tfrac{1}{2}\varepsilon + o(\varepsilon^2)\quad \text{as } \varepsilon\downarrow 0.\end{aligned}$$

In view of Remark 4.1, $\inf\{K(Q,P): \iint a\,dQ \geq \varepsilon\}$ is attained by the probability measure Q_ε with density

$$q_\varepsilon(u,v) = \exp[\tau_\varepsilon\{a(u,v)-\varepsilon\}] \Big/ \iint \exp[\tau_\varepsilon\{a(u,v)-\varepsilon\}]\,du\,dv,$$

where τ_ε is the unique solution of

$$E_P[(a-\varepsilon)\exp\{\tau_\varepsilon(a-\varepsilon)\}] = 0.$$

By (4.10), we have

$$\tau_\varepsilon = \varepsilon\{E_P a^2\}^{-1} + o(\varepsilon) = \varepsilon + o(\varepsilon) \quad \text{as } \varepsilon \downarrow 0,$$

and, hence,

$$q_\varepsilon \approx 1 + \varepsilon a(u,v). \tag{4.31}$$

This implies that at least approximately $q_\varepsilon \in \mathscr{H}$ and, therefore, in (4.30) equality is obtained in the limit as $\varepsilon \downarrow 0$. To make this argument more precise, let $\delta > 0$ and define

$$h_\varepsilon(u,v) = 1 + \varepsilon a_\varepsilon(u,v)(1+\delta),$$

where

$$a_\varepsilon(u,v) = a_\varepsilon^*(u,v) - \int a_\varepsilon^*(u,y)\,dy - \int a_\varepsilon^*(x,v)\,dx + \iint a_\varepsilon^*(x,y)\,dx\,dy$$

and

$$a_\varepsilon^*(u,v) = \begin{cases} a(u,v), & \text{if } |a(u,v)| < \dfrac{1}{4\varepsilon(1+\delta)}, \\ 0, & \text{if } |a(u,v)| \geq \dfrac{1}{4\varepsilon(1+\delta)}. \end{cases}$$

Then we have $h_\varepsilon \in \mathscr{H}$, $|h_\varepsilon - 1| < 1$ since $|a_\varepsilon^*|\varepsilon(1+\delta) < \frac{1}{4}$ and $|a_\varepsilon^*| \leq |a|$. Hence,

$$\int |a_\varepsilon^*(u,y)|\,dy \leq \int |a(u,y)|\,dy \leq \left\{\int a^2(u,y)\,dy\right\}$$

and, therefore,

$$a_\varepsilon^2(u,v) \leq 4a^2(u,v) + 4\int a^2(u,y)\,dy + 4\int a^2(x,v)\,dx + 4,$$

which is integrable. Application of the dominated convergence theorem yields

$$\lim_{\varepsilon\downarrow 0} \iint \frac{ah_\varepsilon}{\varepsilon} = (1+\delta)\lim_{\varepsilon\downarrow 0}\iint aa_\varepsilon = 1+\delta > 1,$$

implying $\iint ah_\varepsilon > \varepsilon$ for ε sufficiently small. Since $|h_\varepsilon - 1| < 1$ and, hence,

$$\left|\frac{h_\varepsilon \log h_\varepsilon - (h_\varepsilon - 1)}{\varepsilon^2}\right| \leq \frac{(h_\varepsilon - 1)^2}{\varepsilon^2} \sup_{|x|<1}\left|\frac{(1+x)\log(1+x) - x}{x^2}\right|,$$

it follows again by dominated convergence that

$$\lim_{\varepsilon\downarrow 0}\frac{\int\int h_\varepsilon \log h_\varepsilon}{\varepsilon^2}=\lim_{\varepsilon\downarrow 0}\frac{\int\int h_\varepsilon \log h_\varepsilon-(h_\varepsilon-1)}{\varepsilon^2}=\frac{1}{2}(1+\delta)^2.$$

and, therefore,

$$I(\varepsilon, a)\le \tfrac{1}{2}(1+\delta)^2\varepsilon^2+o(\varepsilon^2)\quad \text{as } \varepsilon\downarrow 0. \tag{4.32}$$

Since $\delta > 0$ was arbitrarily chosen, combination of (4.30) and (4.32) yields the result. □

REFERENCES

BAHADUR, R. R. (1960). Asymptotic efficiency of tests and estimates. *Sankhyā* **22** 229–252.

BAHADUR, R. R. (1971). *Some Limit Theorems in Statistics*. SIAM, Philadelphia.

BASU, D. (1955). The concept of asymptotic efficiency. *Sankhyā* **17** 193–196.

BOOS, D. D. (1979). A differential for *L*-statistics. *Ann. Statist.* **7** 955–959.

BOOS, D. D. and SERFLING, R. J. (1980). A note on differentials and the CLT and LIL for statistical functions with applications to *M*-estimates. *Ann. Statist.* **8** 618–624.

CSISZÁR, I. (1975). *I*-divergence geometry of probability distributions and minimization problems. *Ann. Probab.* **3** 146–158.

FU, J. C. (1980). Large deviations, Bahadur efficiency, and the rate of convergence of linear function of order statistics. *Bull. Inst. Math. Acad. Sinica* **8** 15–37.

GROENEBOOM, P. and OOSTERHOFF, J. (1977). Bahadur efficiency and probabilities of large deviations. *Statist. Neerlandica* **31** 1–24.

GROENEBOOM, P., OOSTERHOFF, J. and RUYMGAART, F. H. (1979). Large deviation theorems for empirical probability measures. *Ann. Probab.* **7** 553–586.

HOEFFDING, W. (1965). On probabilities of large deviations. *Proc. Fifth Berkeley Symp. Math. Statist. Probab.* **1** 203–219. Univ. California Press.

HUBER, P. J. (1981). *Robust Statistics*. Wiley, New York.

JUREČKOVÁ, J. and KALLENBERG, W. C. M. (1987). On local inaccuracy rates and asymptotic variances. *Statist. Decisions* **5** 139–158.

KEMPERMAN, J. H. B. (1969). On the optimum rate of transmitting information. *Probability and Information Theory. Lecture Notes in Math.* **89** 126–169. Springer, Berlin.

KREMER, E. (1981). Local Bahadur efficiency of rank tests for the independence problem. *J. Multivariate Anal.* **11** 532–543.

WIEAND, H. S. (1976). A condition under which the Pitman and Bahadur approaches to efficiency coincide. *Ann. Statist.* **4** 1003–1011.

WOODWORTH, G. G. (1970). Large deviations and Bahadur efficiency of linear rank statistics. *Ann. Math. Statist.* **41** 251–283.

DEPARTMENT OF APPLIED MATHEMATICS
TWENTE UNIVERSITY OF TECHNOLOGY
P.O. BOX 217
7500 AE ENSCHEDE
NETHERLANDS

INSTITUTE OF MATHEMATICS
TECHNICAL UNIVERSITY OF WROCŁAW
50-370 WROCŁAW
WYBRZEZE WYSPIAŃSKIEGO 27
POLAND