The Annals of Probability
1981, Vol. 9, No. 6, 971–987

LARGE DEVIATIONS OF GOODNESS OF FIT STATISTICS AND LINEAR COMBINATIONS OF ORDER STATISTICS

BY PIET GROENEBOOM[1] AND GALEN R. SHORACK[2]

University of Washington

Asymptotic behavior of large deviations of empirical distribution functions (df's) is considered. Borovkov (1967) and Hoadley (1967) obtained results for functionals continuous in the sup norm topology on the set of df's. Groeneboom, Oosterhoff, and Ruymgaart (1979) extended this to functionals continuous in a stronger τ-topology. This result is now extended to functionals that are τ-continuous only on a particular useful subset of df's. Applications to the Anderson-Darling statistic and linear combinations of order statistics are considered. We begin by correcting the work of Abrahamson (1967); from this the role of the key weight function $\psi(t) = -\log t(1-t)$ is discovered. It is then exploited to the end indicated above, and it is considered as a weight function in tests of fit.

1. Introduction and summary. Let $X_1, \cdots, X_n$ be iid F with empirical df $\hat{F}_n$. We consider weight functions ψ that are finite, positive and continuous on (0, 1). The *weighted Kolmogorov-Smirnov statistics* with weight function ψ are defined by

$$K^+ \equiv K^+_{\psi,n} \equiv \sup_x n^{1/2}[\hat{F}_n(x) - F(x)]\psi(F(x))$$

$$K^- \equiv K^-_{\psi,n} \equiv \sup_x n^{1/2}[F(x) - \hat{F}_n(x)]\psi(F(x)) \tag{1.1}$$

$$K \equiv K_{\psi,n} \equiv K^+_{\psi,n} \vee K^-_{\psi,n}.$$

and the *weighted Cramér-von Mises statistic* is

$$C \equiv C_{\psi,n} \equiv \int_{-\infty}^{\infty} n[\hat{F}_n(x) - F(x)]^2\psi^2(F(x))\, dF(x). \tag{1.2}$$

These statistics are often used to test the null hypothesis that F is the true df of the X_i's.

If ξ is uniform (0, 1), then $X \equiv F^{-1}(\xi)$ (where $F^{-1}(t) \equiv \inf\{x : F(x) \geq t\}$) has df F. By this *inverse transformation* we may suppose that

$$K^+ = \sup_x U_n(F(x))\psi(F(x)) \tag{1.1$'$}$$

$$C = \int_{-\infty}^{\infty} U_n^2(F(x))\psi^2(F(x))\, dF(x) \tag{1.2$'$}$$

where $\hat{\Gamma}_n$ is the empirical df of n independent uniform (0, 1) rv's $\xi_1, \cdots, \xi_n$ and

$$U_n(t) = n^{1/2}[\hat{\Gamma}_n(t) - t] \quad \text{for} \quad 0 \leq t \leq 1$$

is the associated empirical process. We note that

Received May 19, 1980; revised October, 1980.

[1] This work was done while the first author was on leave from the Mathematical Centre Amsterdam as a Visiting Professor in Statistics at the University of Washington.

[2] Research supported by the National Science Foundation under grant MCS 78-09858.

AMS 1970 *subject classifications.* Primary 60F10, secondary 62G20, 62G30.

Key words and phrases. Large deviations, Sanov problem, empirical measures, τ-topology, Kolmogorov-Smirnov tests, linear combinations of order statistics, Bahadur efficiency, local efficiency.

(1.1″) $$K^+ = \sup_{0<t<1} U_n(t)\psi(t) \qquad \text{if } F \text{ is continuous}$$

(1.2″) $$C = \int_0^1 U_n^2(t)\psi^2(t)\,dt \qquad \text{if } F \text{ is continuous.}$$

We now assume that F is continuous.

The evaluation of Bahadur efficiency of tests based on such statistics requires a large deviation result for (1.1″) and (1.2″). It is closely linked to Chernoff's theorem for independent Bernoulli trials; that is,

(1.3) $$n^{-1}\log P(\hat{\Gamma}_n(t) \geq t + a) \to -f(a, t) \text{ from below as } n \to \infty,$$

for $0 \leq t \leq 1$, $a \geq 0$ where (in the notation of Bahadur (1971))

(1.4) $$f(a, t) = \begin{cases} (a + t)\log\dfrac{a + t}{t} + (1 - a - t)\log\dfrac{1 - a - t}{1 - t} & \text{if } 0 \leq a \leq 1 - t \\ \infty & \text{if } a > 1 - t. \end{cases}$$

This function is given in Table 1. We now define

(1.5) $$g_\psi(a) \equiv \inf_{0<t<1} f(a/\psi(t), t).$$

We consider g_1 and g_2 in Table 2, where

(1.6) $$g_i \equiv g_{\psi_i} \quad \text{with} \quad \psi_1(t) \equiv 1 \quad \text{and} \quad \psi_2(t) = -\log[t(1 - t)].$$

THEOREM 1. *Let $X_1, X_2, \cdots$ be* iid *with continuous* df F. *Let ψ denote a finite, positive and continuous function on* (0, 1) *that is symmetric about* $t = \frac{1}{2}$ *and for which* $\lim_{t\to 0} \psi(t)$ *exists in* $[0, \infty]$. *Letting $K^{\#}$ denote any of K^+, K^- or K, we have*

(1.7) $$\lim n^{-1}\log P(K^{\#}_{\psi,n} \geq n^{1/2}a) = -g_\psi(a)$$

for each $a \geq 0$. Moreover, (*when ψ puts too much weight in the tails*)

(1.8) $$\liminf_{t\downarrow 0}\, (\log 1/t)/\psi(t) = 0 \quad \textit{implies} \quad g_\psi(a) = 0 \quad \textit{for all} \quad a \geq 0.$$

REMARK 1. The Theorem (1.7) still holds if the assumption of symmetry is dropped provided we replace g_ψ by $g^{\#}_\psi$ where $g^{\#}_\psi$ is any of

(1.9) $$g^+_\psi = g_\psi \text{ of (1.5)}, \qquad g^-_\psi = \inf_{0<t<1} f(a/\psi(t), 1 - t), \qquad g_\psi = g^+_\psi \wedge g^-_\psi.$$

REMARK 2. The function $\psi_2(t) \equiv -\log(t(1 - t))$ of (1.6) satisfies Theorem 1. Moreover

(1.10) $$f(a/\psi_2(t), t) \to a \quad \text{as } t \downarrow 0 \text{ or as } t \uparrow 1,$$

and the positive minimum of this function will be attained in (0, 1). In fact

(1.11) $$g_{\psi_2}(a) \sim e^2a^2/8 \quad \text{as } a \downarrow 0,$$

while the point of minimization t_a converges to a solution of $t(1 - t) = \exp(-2)$ as $a \downarrow 0$. Moreover, $g_{\psi_2}(a) \to \infty$ as $a \to \infty$ (see Section 4). The function $\psi_1(t) \equiv 1$ of (1.6) also satisfies Theorem 1. It is known from Lemma 5.1 of Bahadur (1971) that

(1.12) $$g_{\psi_1}(a) \sim 2a^2 \quad \text{as } a \downarrow 0$$

and that this function approaches ∞ as $a \uparrow 1$.

REMARK 3. The functions $\psi_\delta(t) \equiv [t(1 - t)]^{-\delta}$ with $\delta > 0$, satisfy Theorem 1 and (1.8); thus their g functions are identically zero. Intuitively, these functions are too severe because they put too much weight on the extreme order statistics; we now demonstrate

TABLE 1.

$$f(a, t) = (a + t) \log \frac{a + t}{t} + (1 - a - t) \log \frac{1 - a - t}{1 - t}.$$

.90	.0167	.1054										
.80	.0084	.0367	.2231									
.70	.0062	.0254	.1163	.357								
.60	.0053	.0216	.0915	.226	.511							
.50	.0050	.0201	.0823	.193	.368	.693						
.40	.0051	.0204	.0811	.184	.335	.551	.916					
.30	.0058	.0226	.0872	.192	.339	.534	.794	1.20				
.20	.0074	.0282	.1046	.223	.382	.583	.832	1.15	1.61			
.10	.0122	.0444	.1537	.311	.511	.751	1.033	1.36	1.76	2.30		
.05	.0206	.0702	.2251	.434	.688	.983	1.318	1.70	2.13	2.65	3.00	
.01	.0588	.1690	.4611	.815	1.217	1.661	2.144	2.67	3.25	3.89	4.25	4.61
t/a	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99

TABLE 2.

$\hat{g}_1(a) = 2a^2$; $\hat{g}_2(a) = e^2a^2/8 = .9236a^2$

Interpolate linearly in columns. See (1.6) for definition of g_1 and g_2.

a	$g_1(a)/\hat{g}_1(a)$	$g_2(a)/\hat{g}_2(a)$
.00	1.0000	1.0000
.01	1.0000	.9917
.02	1.0001	.9835
.03	1.0002	.9755
.05	1.0006	.9596
.07	1.0011	.9442
.10	1.0022	.9219
.12	1.0032	.9076
.15	1.0051	.8868
.20	1.0091	.8541
.25	1.0144	.8237
.30	1.0210	.7953
.40	1.0390	.7440
.50	1.0646	.6989
.60	1.1009	.6591
.70	1.1540	.6236
.80	1.2394	.5918
.90	1.4211	.5632
.97	1.9342	—
1.00		.5373
1.50		.4368
2.00		.3679
3.00		.2790

this. Let $\xi_{n:1} \le \cdots \le \xi_{n:n}$ denote the order statistics of uniform (0, 1) rv's $\xi_1, \cdots, \xi_n$. Then

$$\begin{aligned}[K^+_{\psi_\delta,n} \ge n^{1/2}a] &\supset [1/n \ge a(\xi_{n:1}(1 - \xi_{n:1}))^\delta + \xi_{n:1}] \\ &\supset [1/n \ge a\xi^\delta_{n:1} + \xi_{n:1} \quad \text{and} \quad \xi_{n:1} \le 1/(2n)] \\ &\supset [\xi_{n:1} \le (2n)^{-1} \wedge (2na)^{-1/\delta}] \supset [\xi_1 \le (2n)^{-1} \wedge (2na)^{-1/\delta}],\end{aligned}$$

and even the probability of the smallest event does not go to zero exponentially fast since $\lim n^{-1}\log((2n)^{-1} \wedge (2na)^{-1/\delta}) = 0$.

Asymptotic distributions of truncated or restandardized forms of $K_{\psi_{1/2},n}$ are computed in Jaeschke (1979) and Eicker (1979). Again, truncation or standardization is needed to avoid too heavy weights on the extreme order statistics.

COROLLARY 1. *If* $X_1, X_2, \cdots$ *are* iid *with continuous* df G *and if*

$$K^{\#}_{\psi,n}/n^{1/2} \to \phi(G) \text{ a.s.} \quad \text{as } n \to \infty \tag{1.13}$$

for some $\phi(G) \ge 0$, *then the exact Bahadur slope of the* $K^{\#}_{\psi,n}$*-test for testing against the alternative* G *is*

$$2g_\psi(\phi(G)). \tag{1.14}$$

REMARK 4. Abrahamson (1967) states a similar theorem. However, her proof fails (the part when $i = 1$ is incorrect) in case $\psi(t) \to \infty$, as $t \downarrow 0$ or as $t \uparrow 1$. Moreover, her conclusion is incorrect, because (as in paragraph two of her proof) she claims that $f(a/\psi(t), t) \to \infty$ as $t \downarrow 0$ or $t \uparrow 1$; whereas in the case $\psi(t) = (t(1-t))^{-1/2}$ she considered, $f(a/\psi(t), t) \to 0$ as $t \downarrow 0$ or $t \uparrow 1$. Thus, her basic expansion of her log $\rho^{*}_{\psi}(\epsilon)$ following her (3.16) is incorrect, and also her Table 1 is incorrect.

We will now relate Theorem 1 to some papers in the literature that consider the Sanov problem, and we will present our second theorem along these lines. To this end, we introduce the following notation. Let D denote the set of all df's on the real line $\mathbb{R}$. For F, $G \in D$ the Kullback-Leibler number $K(G, F)$ of G with respect to F is defined by

$$K(G, F) = \begin{cases} \int_{\mathbb{R}} \log\left(\frac{dG}{dF}\right) dG & \text{if } G \ll F \\ \infty & \text{otherwise.} \end{cases} \tag{1.15}$$

For a set Ω of df's we define

$$K(\Omega, F) = \inf\{K(G, F) : G \in \Omega\}. \tag{1.16}$$

REMARK 5. For ψ as in Theorem 1 or Remark 1, define $T_F : D \to \mathbb{R}$ by

$$T_F(G) \equiv \sup_{x \in \mathbb{R}} |G(x) - F(x)| \psi(F(x)). \tag{1.17}$$

For $a \ge 0$ let

$$\Omega_a = \{G : T_F(G) \ge a\}. \tag{1.18}$$

In Section 4, it is shown that if ψ is bounded, then

$$\lim n^{-1} \log P(K^{\#}_{\psi,n} \ge n^{1/2}a) = \lim n^{-1} \log P(T_F(\hat{F}_n) \ge a) = -K(\Omega_a, F) = -g_\psi(a) \tag{1.19}$$

follows from Theorem 3.1 of Groeneboom, Oosterhoff and Ruymgaart (1979) (henceforth referred to as GOR(1979)). This also follows from Hoadley's (1967) Theorem 1. That the conclusion (1.19) holds more broadly follows from (1.7) and (1.25) below.

REMARK 6. As a matter of curiosity, we note that Theorem 1 implies that Sievers' (1976) conditions are satisfied. In fact, the proof of Theorem 1 shows that the event $\{\hat{F}_n \in \Omega_a\}$ is contained in the union of events $\cup_{i=1}^{k_n} \{\hat{F}_n \in U_{i,n}\}$, where $k_n = o(n)$, $U_{i,n}$ is a *convex* set of df's and $\liminf_{n\to\infty} K(\cup_{i,n}, F) \ge K(\Omega_a, F)$. Using this fact, one can show that Sievers' (1976) Condition II is satisfied. However, it seems difficult to prove this directly (i.e. without essentially repeating the proof of Theorem 1).

Hoadley (1967) established the Sanov result (1.18) for certain functionals $T : D \to \mathbb{R}$ uniformly continuous in the sup norm topology on D. GOR (1979) extended this to certain functionals continuous in a stronger τ-topology. This result is now extended to certain functionals that are τ-continuous only on particular subsets of D. (The τ-topology is described in detail in Section 4 just after the proof of Corollary 1.)

THEOREM 2. *Let F be a continuous* df *on $\mathbb{R}$ and let $T:D \to \mathbb{R}$ be τ-continuous on*

(1.20) $D_m \equiv \{G \in D: d_F(G, F) \le m\}$, *where* $d_F(G, F) \equiv \sup_{x\in\mathbb{R}} |G(x) - F(x)|\psi_2(F(x))$

for each $m > 0$. Let

$$\Omega_a \equiv \{G \in D: T(G) \ge a\}, \tag{1.21}$$

and suppose $K(\Omega_a, F) < \infty$. Then

$$\limsup_{n\to\infty} n^{-1} \log P(T(\hat{F}_n) \ge a) \le -K(\Omega_a, F). \tag{1.22}$$

Moreover, if the function $t \to K(\Omega_t, F)$ is right continuous at $t = a$ and $u_n \to 0$ then

$$\lim_{n\to\infty} n^{-1} \log P(T(\hat{F}_n) \ge a + u_n) = -K(\Omega_a, F). \tag{1.23}$$

There are several important facts that prove useful in the applications and throughout the proofs. To avoid the impression of circular arguments, we prove these in this section.

PROPOSITION 1. *Let D_m and d_F be defined as in* (1.20).

(i) *For $m \equiv m_{F,\Omega}$ sufficiently large*

(1.24) $K(\Omega \cap D_m, F) = K(\Omega, F)$ *for any F and Ω having $K(\Omega, F) < \infty$.*

(ii) *For continuous F any ψ positive on* $(0, 1)$,

(1.25) $K(\Omega_a, F) = g_\psi(a)$ *where Ω_a is defined as in* (1.18).

(iii) *Suppose F is continuous and G_r are arbitrary* df's; *then*

(1.26) $\limsup_{r\downarrow 0} r^{-2}K(G_r, F) \le B$ *implies* $\limsup_{r\downarrow 0} r^{-1}d_F(G_r, F) \le (8B/e^2)^{1/2}$.

PROOF. (i) Since $t\psi_2(t)$ is bounded for $t \le ½$ and likewise $(1 - t)\psi_2(t)$ is bounded for $t \ge ½$, we have for all large m that

$$D_m^c \subset \{G \in D: \sup_{F(x)\le 1/2} G(x)\psi_2(F(x)) \ge m/2\} \cup \{G \in D: \sup_{F(x)\ge 1/2}(1 - G(x))\psi_2(F(x)) \ge m/2\}.$$

Also for m sufficiently large, we have

$$\sup_{F(x)\le 1/2} G(x)\, \psi_2(F(x)) \ge m/2 \quad \text{implies} \quad (K(G, F)) \ge m/4$$

(and a similar implication with $1 - G(x)$) since

$$K(G, F) \ge G(x)\log(G(x)/F(x)) + (1 - G(x)) \log(1 - G(x))/(1 - F(x)).$$

Thus for some large m_1 we have that

$$K(G, F) \ge m_1/4 \ge K(\Omega, F) + 1 \quad \text{whenever} \quad d_F(G, F) \ge m_1.$$

(ii) Since $F \circ F^{-1}$ is the identity for continuous df's F, we can define the df G_t on $\mathbb{R}$ so that the df $\bar{G}_t \equiv G_t \circ F^{-1}$ on the unit interval has the uniform on intervals density

$$\bar{g}_t(u) = \begin{cases} [t + a/\psi(t)]/t & \text{if } 0 < u < t \\ [(1 - t) - a/\psi(t)]/(1 - t) & \text{if } t < u < 1. \end{cases} \tag{a}$$

Note that $G_t \in \Omega_a$ since $T_F(G_t) = a$, and note that

$$K(G_t, F) = f(a/\psi(t), t). \tag{b}$$

Thus

$$K(\Omega_a, F) \le \inf_t K(G_t, F) = g_\psi(a). \tag{c}$$

To obtain the reverse inequality, we note that for $\epsilon > 0$ there exists a df $G_a^* \in \Omega_a$ for which

$$K(G_a^*, F) < K(\Omega_a, F) + \epsilon. \tag{d}$$

Now $G_a^* \in \Omega_a$ implies $|\bar{G}_a^*(t_a) - t_a| \geq a/\psi(t_a)$ for some t_a in (0, 1). But the df $\bar{G}_a^\circ$ that is linear from (0, 0) to $(t_a, \bar{G}_a^*(t_a))$ to (1, 1) has the smallest Kullback-Leibler number, when compared to the Uniform (0, 1) df, among all df's that pass through the point $(t_a, \bar{G}_a^*(t_a))$; and its information number is $\geq K(G_{t_a}, F)$ (note that $\bar{G}_a^\circ = \bar{G}_{t_a}$ when $\bar{G}_a^*(t_a) - t_a = a/\psi(t_a)$). Thus for $\epsilon > 0$

(e) $$K(\Omega_a, F) + \epsilon \geq K(G_a^*, F) \geq K(G_{t_a}, F) = f(a/\psi(t_a), t_a) \geq g_\psi(a).$$

(iii) Assume there exist df's G_r contradicting (1.26), so that $c_r \equiv r^{-1} d_F(G_r, F)$ satisfies $\limsup c_r > (8B/e^2)^{1/2}$. Then with Ω_r as in (1.18) we have

(1.27) $$\begin{aligned} B &\geq \limsup r^{-2} K(G_r, F) \geq \limsup r^{-2} K(\Omega_{rc_r}, F) \\ &= \limsup r^{-2} g_\psi(rc_r) \quad \text{by (1.25) for any such } \psi \\ &= \limsup r^{-2} e^2 r^2 c_r^2/8 \quad \text{for } g_2 \text{ by (1.11)} \\ &> \mathrm{B} \end{aligned}$$

providing the necessary contradiction. □

Results analogous to (1.11) for other ψ's produce, in conjunction with (1.27), results analogous to (1.26) for other ψ's; equation (1.12) allows this for ψ_1. Thus, if F is continuous and G_r are arbitrary df's, then

(1.28) $$\limsup_{r\downarrow 0} r^{-2} K(G_r, F) \leq B \quad \text{implies } \limsup_{r\downarrow 0} r^{-1} \sup_x |G(x) - F(x)| \leq (B/2)^{1/2}.$$

2. Applications to tests of fit. Three statistics will be considered in this section: the two-sided Kolmogorov-Smirnov statistics, call them K_1 and K_2, of (1.1) with weight functions $\psi_1(t) \equiv 1$ and $\psi_2(t) \equiv -\log(t(1-t))$, and the Anderson-Darling statistic A that corresponds to (1.2) with weight function $[t(1-t)]^{-1/2}$.

The key results for K_1 and K_2 are contained in (1.7), (1.11), and (1.12). The corollary below presents the corresponding results for A.

EXAMPLE 1. The Anderson-Darling statistic $A \equiv A_n = T(\hat{\Gamma}_n)$, where $T: D \to \mathbb{R} \cup \{\infty\}$ is defined via (1.2″) by

(2.1) $$T(G) = \int_0^1 [G(t) - t]^2/[t(1-t)]\, dt \qquad \text{for} \quad G \in D.$$

We will now show that this functional is finite and τ-continuous on each subset

(2.2) $$D_m \equiv \{G \in D: \sup_{t\in(0,1)} |G(t) - t| \psi_2(t) \leq m\}$$

with $m > 0$. To see this, notice that the integrals over $(0, \eta]$ and $[1-\eta, 1)$ in (2.1) tend to zero as $\eta \downarrow 0$, uniformly in $G \in D_m$ for each fixed $m > 0$. Also, for fixed $\eta \in (0, 1/2)$, the functional $G \to \int_\eta^{1-\eta} [G(t) - t]^2/[t(1-t)]\, dt$ is obviously τ-continuous (even continuous with respect to the topology of weak convergence) on D. From these two statements the τ-continuity of T on D_m follows. (However, T is not continuous with respect to the τ-topology on the whole set of df's.) From Theorem 2, we obtain (2.3) below.

COROLLARY 2. *Let F denote the uniform* (0, 1) *distribution. Let $\Omega_a = \{G \in D: T(G) \geq a\}$ for each $a \geq 0$, with T as in* (2.1). *Then*:

(2.3) $$\lim_{n\to\infty} n^{-1} \log P(A_n \geq a) = -K(\Omega_a, F) \quad \text{for each} \quad a \geq 0.$$

(2.4) *The function* $a \to K(\Omega_a, F)$, $a \geq 0$, *is continuous.*

(2.5) $$K(\Omega_a, F) = a + o(a) \quad \text{as} \quad a \downarrow 0.$$

REMARK 7. Result (2.5) of Corollary 2 appears in Nikitin (1976), where it is stated that it can be derived using the theory of branching of solutions of non-linear equations in Banach spaces in Vainberg and Trenogin (1974). Some of the statements in Nikitin (1976) are proved along these lines in Nikitin (1979) and a proof of (2.5) is announced as a part of a forthcoming paper, Nikitin (1980). A more straightforward proof of (2.5) is given in Section 4 below. Result (2.3) of Corollary 2 is used implicitly in Nikitin's paper; he only remarks that asymptotic probabilities of large deviations of functionals of empirical distribution functions can be computed by finding the Kullback-Leibler numbers with respect to the corresponding set of distribution functions. It is clear, however, that the functionals, and also the Kullback-Leibler numbers of the sets of distribution functions, have to satisfy certain regularity conditions (for examples where the relation between Kullback-Leibler numbers and large deviation limits does not hold, see e.g. Bahadur and Zabell (1979) and GOR (1979)).

Local Bahadur efficiencies of the Anderson-Darling test with respect to other well-known goodness of fit tests are computed in Nikitin (1976), indicating that the Anderson-Darling test is to be preferred for normal and logistic location alternatives and that Watson's test is best for Cauchy location alternatives.

EXAMPLE 2. (Logistic scale alternatives). Let the df $F(x) = 1/(1 + \exp(-x))$ have density f; note that $F(1 - F) = f$. For $\theta > 0$ let $F_\theta = F(\cdot/\theta)$. Then two times

$$\int_R [F(x/\theta) - F(x)]^2/[F(x)(1 - F)]f(x)\,dx$$

behaves as

$$2(1-\theta)^2 \int_R x^2 f^2(x)\,dx = .430(1-\theta)^2 \quad \text{as} \quad \theta \to 1. \tag{2.6}$$

Thus, by (2.5), the exact slope of the Anderson-Darling statistic for testing $\theta = 1$ against $\theta \neq 1$ also behaves as (2.6). The analogous result for the K_2-test comes from

$$\begin{aligned} &\sup_{x\in R} |F(x/\theta) - F(x)|\,\psi_2(F(x)) \\ &\quad \sim |(1-\theta) \sup_{x\in R} xf(x) \log(1/f(x))| = .476\,|1-\theta| \quad \text{as} \quad \theta \to 1 \end{aligned} \tag{2.7}$$

and (1.11); so that the exact slope of the K_2-statistic behaves as

$$(2e^2/8)(.476(1-\theta))^2 = .419(1-\theta)^2 \quad \text{as} \quad \theta \to 1. \tag{2.8}$$

(The maximum of .476 is attained at approximately $x \approx 2.168$). Analogously, the K_1-statistic has exact slope

$$4(.224(1-\theta))^2 = .201(1-\theta)^2 \quad \text{as} \quad \theta \to 1 \tag{2.9}$$

(since $xf(x)$ attains a maximum of $\approx .224$ at $x \approx 1.543$). Thus the Bahadur efficiencies of the A and K_2 tests with respect to the K_1 test converge to 2.14 and 2.08 respectively as $\theta \to 1$; see Table 3 for the rate of convergence to 2.08.

Exact Bahadur efficiencies of K_1 and K_2-tests for any fixed alternative can be computed from (1.13) via Table 2. This was done for logistic and normal location and scale alternatives; the results are given in Table 3. We note that the weight function ψ_2 outperforms ψ_1 for scale alternatives with a reversal for location alternatives.

3. Applications to linear combinations of order statistics.

Theorem 2 can also be used to obtain large deviation results (and hence Bahadur efficiencies) for linear combinations of order statistics. We consider linear combinations of order statistics of the form

$$T(\hat{F}_n) = \int_0^1 J(u)\hat{F}_n^{-1}(u)\,du, \tag{3.1}$$

TABLE 3.
$a_i \equiv \sup_x[F(x) - F((x - L)/S)]\psi_i(F(x))$ for $i = 1, 2$. $e_{2,1}$ is the Bahadur efficiency of the $K^-_{\psi_2}$-test with respect to the $K^-_{\psi_1}$-test; here we are testing F while $F((.-L)/S)$ is actually true.

	L	S	a_1	a_2	$e_{2,1}$
Logistic scale	0.00	2.00	.150	.259	1.13
	"	1.50	.0897	.167	1.38
	"	1.10	.0213	.0439	1.88
	"	1.01	.00223	.00472	2.07
Normal Scale	0.00	2.00	.161	.282	1.13
	"	1.50	.0968	.185	1.46
	"	1.10	.0231	.0498	2.07
	"	1.01	.00241	.00540	2.31
Logistic location	1.00	1.00	.245	.378	.820
	.50	"	.124	.178	.821
	.10	"	.0250	.0347	.863
	.01	"	.00250	.00347	.885
Normal location	1.00	1.00	.3829	.6983	.926
	.50	"	.1974	.3048	.865
	.10	"	.03988	.05582	.865
	.01	"	.00399	.00553	.884

where J is an L_1-function (i.e. $\int_0^1 |J(u)|\,du < \infty$) and for a df $G \in D$, the inverse G^{-1} is defined by $G^{-1}(u) = \inf\{x \in \mathbb{R} : G(x) \geq u\}$.

In GOR (1979), large deviation results for statistics of the form (3.1) with score functions having support contained in $[\alpha, \beta] \subset (0, 1)$ are given. By Theorem 2, we can obtain an extension of these results to statistics with score functions J, whose support is not necessarily restricted to a closed subinterval of (0, 1). This is shown in the next corollary.

COROLLARY 3. *Suppose that F is a continuous* df *and that J is an L_1-function satisfying*

(i)
$$\int_{1/2}^1 |J(u)F^{-1}(1 - \exp(-c/(1 - u)))|\,du < \infty \quad \textit{and}$$
$$\int_0^{1/2} |J(u)F^{-1}(\exp(-c/u))|\,du < \infty \quad \textit{for each} \quad c > 0.$$

(ii)
$$J \geq 0 \textit{ on an interval } (\gamma, \delta) \subset (0, 1) \textit{ and } \int_\gamma^\delta J(u)\,du > 0.$$

Let $\Omega_r = \{G \in D : r \leq \int_0^1 J(u)G^{-1}(u)\,du < \infty\}$. *Then, if $\hat{F}_n$ is the empirical* df *of independent random variables $X_1, \cdots, X_n$ with* df F,

(3.2) $$\lim_{n\to\infty} n^{-1}\log P(T(\hat{F}_n) \geq r) = -K(\Omega_r, F).$$

(3.3) *The function $t \to K(\Omega_t, F)$ is a continuous mapping from $\mathbb{R}$ into $[0, \infty]$, where $[0, \infty]$ is endowed with the topology of the extended real line.*

REMARK 8. It is shown in Oosterhoff (1978) that if J is monotonically non-decreasing, a condition like (i) is not needed to ensure (3.2).

EXAMPLE 3. Let $J(u) = u(1 - u)$, $u \in (0, 1)$, and let F be the logistic df $F(x) = (1 + \exp(-x))^{-1}$. It can be shown that Corollary 3 applies, and that

(3.4) $$K(\Omega_r, F) \sim 6r^2 \quad \text{as} \quad r \downarrow 0.$$

Technical details can be found in Groeneboom and Shorack (1980).

Consider location alternatives $F_\theta(x) = F(x - \theta)$, $\theta > 0$. Then $\int_0^1 J(u)F_\theta^{-1}\,du = \theta \int_0^1 J(u)\,du = (1/6)\theta$ and hence by (3.4) the Bahadur slope of the sequence $\{\int_0^1 J(u)\hat{F}_n^{-1}(u)\,du\}$ at θ is asymptotically equivalent to $12(\int_0^1 J(u)F^{-1}(u)\,du)^2 = (1/3)\theta^2$, as $\theta \downarrow 0$. The Bahadur slope of the most powerful test for testing $\mathscr{H}_0: \theta = 0$ against an alternative $\theta > 0$ is given by

$$c(\theta) = 2\{\log 2 - \log(1 + e^{2\theta}) + 2e^\theta(e^{2\theta} - 1)^{-1}(\theta e^\theta - \arctan(1/2(e^\theta - e^{-\theta})))\}$$

(see Groeneboom and Oosterhoff (1977), page 21). By Taylor series expansion it is seen that $c(\theta) \sim (1/3)\theta^2$, as $\theta \downarrow 0$. Hence the Bahadur slope of a test based on the sequence of test statistics $\{\int_0^1 J(u)\hat{F}_n^{-1}(u)\,du\}$ coincides locally with the Bahadur slope of the most powerful test, as $\theta \downarrow 0$. It follows from Corollary 2 that the Anderson-Darling test also has this property.

EXAMPLE 4. If Φ is the standard normal df, then

$$\Phi^{-1}(\exp(-c/u)) \sim -\sqrt{2c/u}, \quad \text{as} \quad u \downarrow 0 \quad \text{and}$$

$$\Phi^{-1}(1 - \exp(-c/(1 - u))) \sim \sqrt{\frac{2c}{1 - u}}, \quad \text{as} \quad u \uparrow 1.$$

Hence, in this case, Corollary 3 allows us even to take score functions J which tend to infinity as $u \downarrow 0$ and $u \uparrow 1$, for example score functions of type $J(u) = \{u(1 - u)\}^{-(1/2)+\epsilon}$ with $u \in (0, 1)$ and $\epsilon > 0$.

4. Proofs.

PROOF OF THEOREM 1. We first consider K^+. Fix $a \geq 0$. For any fixed t we have from (1.3) that

(a) $$n^{-1}\log P(K^+ \geq n^{1/2}a) \geq n^{-1}\log P(\hat{\Gamma}_n(t) \geq t + a/\psi(t)) \to -f(a/\psi(t), t).$$

Hence

(b) $$\liminf n^{-1}\log P(K^+ \geq n^{1/2}a) \geq -\inf_{t\in(0,1)} f(a/\psi(t), t) = -g_\psi(a).$$

We now turn to the other inequality. Note first that

(c) $$\lim_{t\downarrow 0} \psi(t) = \infty \quad \text{implies} \quad f(a/\psi(t), t) = (a/\psi(t))(\log 1/t) + o(1)$$

where $|o(1)| \leq \epsilon$ for all $0 < t \leq$ some δ_ϵ.

If the condition of (1.8) holds, it follows immediately from (c) that $g_\psi(a) = 0$. Thus (1.7) follows trivially from (b) in this case. The conclusion (1.7) also holds trivially when $a = 0$. Thus we now assume that $a > 0$ and

(d) $$\liminf_{t\downarrow 0} (\log 1/t)/\psi(t) > 0.$$

For each n we choose numbers $t_{n,1} < \cdots < t_{n,m(n)}$ such that

(e) $$1 - 2/n < t_{n,i}/t_{n,i+1} < 1 - 1/n \quad \text{for} \quad 1 \leq i < m(n) \quad \text{and}$$

(f) $$t_{n,1} = 1 - t_{n,m(n)} = \exp(-n^2).$$

For $1 \leq i < m(n)$, we let $\psi_{n,i} \equiv \max\{\psi(t): t_{n,i} \leq t \leq t_{n,i+1}\}$. Then,

$$P(\sup_{t\in[t_{n,i},t_{n,i+1}]} [\hat{\Gamma}_n(t) - t]\psi(t) \geq a) \leq P(\hat{\Gamma}_n(t_{n,i+1}) - t_{n,i} \geq a/\psi_{n,i})$$

(g) $$\leq \exp(-nf(a/\psi_{n,i} + t_{n,i} - t_{n,i+1}, t_{n,i+1})) \quad \text{by (1.3)}$$

$$\leq \exp(-n(g_\psi(a) - \epsilon)) \quad \text{for all } i \text{ and large } n$$

by continuity and an argument analogous to (c). Note that

(h) $$m(n) = o(n^3)$$

(solve $(1 - 1/n)^m = \exp(-n^2)$ for m). Also

(i)
$$\begin{aligned} P(\sup_{t\le t_{n,1}} |\hat\Gamma_n(t) - t|\psi(t) \ge a) &= P(\sup_{t\ge t_{n,m(n)}} |\hat\Gamma_n(t) - t|\psi(t) \ge a) \\ &= P(\sup_{t\le t_{n,1}} [\hat\Gamma_n(t) - t]\psi(t) \ge a) \qquad \text{for all large } n \\ &\le P(\hat\Gamma_n(t_{n,1}) > 0) \le nP(\xi_1 \le t_{n,1}) = n\exp(-n^2) \\ &\le \exp(-n^2/2) \qquad \text{for all large } n \end{aligned}$$

Hence (g), (h) and (i) give, for arbitrary $\epsilon > 0$,

(j)
$$\begin{aligned} &\limsup n^{-1}\log P(K^+ \ge n^{1/2}a) \\ &\qquad \le \limsup n^{-1}\log\{m(n)\exp(-n(g_\psi(a) - \epsilon)) + 2\exp(-n^2/2)\} = -g_\psi(a) + \epsilon. \end{aligned}$$

Thus (b) and (j) complete the proof for K^+.

Finally, K^- is completely analogous to K^+; while

(k)
$$\begin{aligned} &P(K^+ \ge n^{1/2}a) \le P(K \ge n^{1/2}a) \\ &\qquad \le P(K^+ \ge n^{1/2}a) + P(K^- \ge n^{1/2}a) \le 2P(K^+ \ge n^{1/2}a). \qquad \square \end{aligned}$$

PROOF OF (1.11). Let $\phi \equiv \phi_a(t) \equiv t + a/\psi_2(t) \equiv t + a/\psi$ for $t \in (0, 1)$. Then for a, t such that $\phi < 1$ we have

(a) $$f \equiv f(a/\psi, t) = \phi\log(\phi/t) + (1 - \phi)\log((1 - \phi)/(1 - t)).$$

Letting

(b) $$\phi' \equiv (d/dt)\phi_a(t) = 1 + \frac{a(1 - 2t)}{t(1 - t)\psi^2},$$

we note that

(c)
$$\begin{aligned} f' \equiv (d/dt)f(a/\psi, t) &= \phi'[\log(\phi/t) - \log((1 - \phi)/(1 - t))] \\ &\quad - (\phi/t) + ((1 - \phi)/(1 - t)) \\ &= \phi'[\log(1 + a/(t\psi)) - \log(1 - a/((1 - t)\psi))] - a/(t(1 - t)\psi). \end{aligned}$$

We first fix $0 < \eta < 1$, and then fix $\epsilon > 0$ so that $(1 - 2\epsilon)x - \log(1 + x) > 0$ for $x \ge \eta$. Let $T \equiv T_a(t) \equiv (1/\psi)\log(1 + a/(t\psi))$. We will now choose δ so small that (for small a) the function $f(a/\psi, t)$ does not reach a minimum on $[0, \delta]$. We will do this by showing that when a is sufficiently small and $t \le$ some δ we have

(i) $f' < 0$ if $a/(t\psi) \le \eta$,

(ii) $f' < 0$ if $a/(t\psi) > \eta$ and $T \le \epsilon$, and

(iii) $f > \inf_r f(a/\psi(r), r)$ if $T > \epsilon$.

It follows by analogy that $f(a/\psi, t)$ does not reach a minimum on $[1 - \delta, 1]$ when a is small. It is trivial that

(d) $$f = a^2/[2t(1 - t)\psi^2] + o(a^3) \quad \text{uniformly for} \quad t \in [\delta, 1 - \delta];$$

while differentiating the main term in (d) yields a minimum of $e^2a^2/8$ whenever t solves $t(1 - t) = e^{-2}$. Thus the result holds.

It remains only to prove (i), (ii) and (iii):

(i) In equation (c) we expand $\log(1 + a/(t\psi))$ to three terms, expand $\log(1 - a/((1 - t)\psi))$ in an infinite series, and plug in (b) to obtain

(e) $$f' \le \frac{a^2}{[t(1 - t)\psi]^2}\left\{\frac{1 - 2t}{\psi} + \frac{t^2 - (1 - t)^2}{2} + \frac{\phi'}{3}\left\{\left(\frac{a}{t\psi}\right)^3 + \sum_{j=3}^{\infty}\frac{3}{j}\left(\frac{a}{(1 - t)\psi}\right)^j\right\}\right\}$$

whenever $a/[t(1-t)\psi] < \eta$. For $t <$ some δ_1 and $a/(t\psi) \le \eta$ the rhs of (e) is dominated by the term involving $-(1-t)^2/2$.

(ii) In equation (c) we have, for $t <$ some δ_2, that

$$\begin{aligned} f' &\le [1 + (a/(t\psi))/\psi][\log(1 + a/(t\psi)) - \log(1 - a/((1-t)\psi))] - a/[t(1-t)\psi] \\ &\le \log(1 + a/(t\psi)) + (a/(t\psi))\epsilon - [1 + a/(t\psi^2)]\log(1 - a/((1-t)\psi)) - a/(t\psi) \\ &\le (1 - 2\epsilon)(a/(t\psi)) + (a/(t\psi))\epsilon - a/(t\psi) - [1 + a/(t\psi^2)]\log(1 - a/((1-t)\psi)) \\ &< -\epsilon a/(t\psi) + [1 + a/(t\psi^2)]2at/(\psi t(1-t)) \\ &= -[\epsilon - 2(1 + a/(t\psi^2))(t/(1-t))]a/(t\psi) < 0 \end{aligned}$$

using $T \le \epsilon$ in the second inequality, the definition of ϵ in the third inequality and $a \le$ some a_0.

(iii) From (a) we have for $t <$ some δ_3

$$\begin{aligned} f &\ge (a/\psi)\log(1 + a/(t\psi)) + (1 - \phi)\log(1 - a/((1-t)\psi)) \\ &\ge a\epsilon = (a\epsilon/2) + a\epsilon/2 \end{aligned}$$

using $T > \epsilon$. But from (d) we have $f \le$ (some K) a^2 on $[\delta, 1-\delta]$ with $\delta \equiv \delta_1 \wedge \delta_2 \wedge \delta_3$. Thus for small a, $f(t) > \inf_r f(a/\psi(r), r)$ if $T > \epsilon$ and $t \le \delta$.

Note that (1.10), (1.4), and a slight modification of the above argument yield $g_\psi(a) \to \infty$ as $a \to \infty$. □

PROOF OF COROLLARY 1. Use Theorem 1 and Theorem 7.2 of Bahadur (1971). □

We now define the *τ-topology*. Let $\mathscr{P} = \{B_1, \cdots, B_m\}$ denote a partition of $\mathbb{R}$ into Borel measurable sets. Let

$$K_{\mathscr{P}}(G, F) \equiv \textstyle\sum_{i=1}^m P_G(B_i) \log P_G(B_i)/P_F(B_i) \tag{4.1}$$

where P_F (or P_G) is the probability distribution corresponding to F (or G). For a set of df's Ω let

$$K_{\mathscr{P}}(\Omega, F) = \inf\{K_{\mathscr{P}}(G, F): G \in \Omega\}. \tag{4.2}$$

(We use the conventions $0 \log \infty = 0$ and $a \log 0 = -\infty$ for $a > 0$.) Consider also the pseudometric $d_{\mathscr{P}}$ on D given by

$$d_{\mathscr{P}}(G, F) \equiv \max_{1 \le i \le m} |P_F(B_i) - P_G(B_i)|. \tag{4.3}$$

The topology on D generated by all such $d_{\mathscr{P}}$ will be denoted by τ (cf. GOR (1979)); thus τ is the smallest topology such that the sets $\{G \in D: d_{\mathscr{P}}(G, F) < \epsilon\}$ are open for each $\epsilon > 0$, each $F \in D$ and each finite partition $\mathscr{P}$. In fact, these sets form a subbase for the topology.

PROOF OF (1.19). For bounded ψ, the function T_F of (1.17) is clearly τ-continuous. Thus, the second equality of (1.19) follows from Theorem 3.1 of GOR (1979). The third equality of (1.19) follows from (1.25). □

REMARK 9. If ψ is unbounded, then the proof of (1.19) fails since T_F is not τ-continuous (hence a fortiori not continuous with respect to the topology on D induced by the supremum metric, cf. GOR (1979) Lemma 2.1). Even the weaker condition

$$K(\Omega_a, F) = \sup\{K_{\mathscr{P}}(\Omega_a, F): \mathscr{P} \text{ is a finite partition of } \mathbb{R}\}$$

fails, if $K(\Omega_a, F) > 0$. This condition is used to obtain upper bounds for the probabilities $P(\hat{F}_n \in \Omega_a)$ in Stone (1974) and also in GOR (1979). To show that the condition is not satisfied we suppose for simplicity that F is the uniform df on (0, 1). If $\limsup_{t\downarrow 0} \psi(t) = \infty$, there exists a decreasing sequence of points $t_n \in (0, 1)$ such that $\lim_{n\to\infty} t_n = 0$ and $\lim_{n\to\infty} \psi(t_n) = \infty$. Define the df $G_n \in \Omega_a$ by the density

$$g_n(u) = \begin{cases} 1 + a/(t_n\psi(t_n)),\ 0 < u \le t_n \\ 1 - a/\{(1-t_n)\psi(t_n)\},\ t_n < u < 1 \\ 0,\ u \notin (0, 1). \end{cases}$$

For each Borel measurable set B we have

$$\lim_{n\to\infty} \int_B dG_n(x) = \int_B 1_{(0,1)}\, du,$$

and hence $K_{\mathscr{P}}(\Omega_a, F) = 0$ for each partition $\mathscr{P} = \{B_1, \cdots, B_m\}$. Note, however, that although the usual proofs fail, we still have (by Theorem 1)

$$\lim_{n\to\infty} n^{-1}\log P(\hat{F}_n \in \Omega_a) = -K(\Omega_a, F),$$

since $K(\Omega_a, F) = g_\psi(a)$.

Proof of Theorem 2. We will show below that, for all $m \ge$ some m_1, we have both

(a) $$K(\Omega_a, F) = K(\Omega_a \cap D_m, F) \quad \text{and}$$

(b) $$K(\Omega_a \cap D_m, F) = K(cl_\tau(\Omega_a \cap D_m), F).$$

Then, using Lemmas 2.4 and 3.1 of GOR (1979), and (b) at inequality (d) we have

$$\limsup_{n\to\infty} n^{-1}\log P(T(\hat{F}_n) \ge a) = \limsup_{n\to\infty} n^{-1}\log P(\hat{F}_n \in \Omega_a)$$

(c) $$= \limsup_{n\to\infty} n^{-1}\log P(\hat{F}_n \in \Omega_a \cap D_m) \quad \text{for sufficiently large } m$$

(d) $$\le -K(\Omega_a \cap D_m, F)$$

$$= -K(\Omega_a, F) \quad \text{by } (a);$$

equation (c) holds because Theorem 1 implies

(e) $$\limsup_{n\to\infty} n^{-1}\log P(\hat{F}_n \in D_m^c) \to -\infty \quad \text{as} \quad m \to \infty$$

since g_{ψ_2} (a) $\to \infty$ as $a \to \infty$ (see Remark 2).

Since (a) is merely (1.24), it remains only to prove (b). To establish (b), it suffices to show that

(f) $$G \in cl_\tau(\Omega_a \cap D_m) \quad \text{with} \quad K(G, F) \le K(\Omega_a, F) + 1$$

implies $G \in \Omega_a \cap D_m$. We suppose (f). We will first establish that $G \in D_m$. Now for some constants C_i we have

$$\begin{aligned} d_F(G, F) &\le \sup_{F(x)\le 1/2} G \log 2/F + \sup_{F(x)\ge 1/2} (1 - G)\log 2/(1 - F) \\ &\quad + \sup_x 2F(1 - F)\log 1/(F(1 - F)) \\ &\le \sup_x [G \log 1/F + (1 - G)\log 1/(1 - F)] + C_1 \\ &\le \sup_x [G \log G/F + (1 - G)\log(1 - G)/(1 - F)] + C_2 \\ &\le K(G, F) + C_2 \\ &\le K(\Omega_a, F) + C_3 \qquad \text{by} \quad (f) \\ &\le m \qquad \text{for } m \text{ sufficiently large.} \end{aligned}$$

Thus $G \in D_m$. From the first half of (f) we thus have $G \in cl_\tau(\Omega_a \cap D_m) \cap D_m$. Since T is τ-continuous on D_m, $\Omega_a \cap D_m$ is closed in the relative τ-topology on D_m; thus $\Omega_a \cap D_m = cl_\tau(\Omega_a \cap D_m) \cap D_m$. Thus $G \in \Omega_a \cap D_m$ as we set out to show. Thus (b) holds. Using the τ-continuity of T on D_m, one can prove (1.23) along the same lines as Theorem 3.2 in GOR (1979). □

PROOF OF COROLLARY 2. It is shown in Example 2.1 that the function T defined by (2.1) is τ-continuous on each subset D_m, with $m > 0$. Hence, according to Theorem 2, result (2.3) holds if the function $a \to K(\Omega_a, F)$ is continuous from the right.

First of all, $K(\Omega_a, F)$ is finite for all $a \in (0, \infty)$ (and also trivially for $a \le 0$). To see this, consider the df's G_x, $x \in (0, 1)$, defined by

$$G_x(t) = x^{-1}t1_{[0,x)}(t) + 1_{[x,\infty)}(t).$$

Then

$$T(G_x) = \int_0^1 \frac{(G_x - t)^2}{t(1-t)}\,dt = \left(\frac{1}{x} - 1\right)^2 \int_0^x (t/(1-t))\,dt + \int_x^1 ((1-t)/t)\,dt.$$

Hence $T(G_x) \to \infty$, as $x \downarrow 0$, but $K(G_x, F) = x(1/x)\log(1/x) < \infty$, all $x \in (0, 1)$.

Fix $a > 0$. By (1.24), there exists an $m > 0$ such that

$$K(\Omega_a, F) = K(\Omega_a \cap D_m, F).$$

Since $\Omega_a \cap D_m$ is τ-closed, there exists by Lemma 3.2 in GOR (1979) a df $G \in \Omega_a \cap D_m$ such that $K(G, F) = K(\Omega_a, F)$. Because $K(G, F) < \infty$, we have $G \ll F$. Hence G has a density g. Since $G \in \Omega_a$ and $a > 0$, there exists $x \in (0, 1)$ such that $G(x) \ne x$, say $G(x) > x$. Without loss of generality we may assume $G(x) < 1$. Let $y \le \inf\{t \ge x : G(t) \le t\}$ have $x < y < 1$. Define the df G_δ by the density

$$g_\delta(t) = \begin{cases} g(t) + \delta, \ t \in [x, x + \delta] \\ g(t) - (y - x - \delta)^{-1}\delta^2, \ t \in (x + \delta, y] \\ g(t), \qquad \text{elsewhere,} \end{cases}$$

where $\delta > 0$ is a suitably chosen small number. Then $G_\delta(t) > G(t)$ for $t \in (x, y)$ and $G_\delta(t) = G(t)$, elsewhere.

Hence $T(G_\delta) \downarrow T(G)$. It is clear that the Kullback-Liebler number $K(G_\delta, F)$ tends to $K(G, F)$ as $\delta \downarrow 0$. This implies that the function $a \to K(\Omega_a, F)$ is continuous from the right. Thus, part (2.3) of the corollary follows.

By the continuity of T on D_m and Lemma 3.3 in GOR (1979), the function $a \to K(\Omega_a, F)$ is also continuous from the left, implying (2.4).

Finally, to prove (2.5), we first observe that the df's $\bar{G}_a$ defined by $\bar{G}_a(t) = t + t(1 - t)\sqrt{6a}$, $t \in [0, 1]$, $6a \le 1$ have $K(\bar{G}_a, F) \sim a$, as $a \downarrow 0$. Hence

(a) $$K(\Omega_a, F) \le a + o(a), \qquad \text{as} \quad a \downarrow 0.$$

Fix $\epsilon > 0$. We shall prove

(b) $$\liminf_{a \downarrow 0} a^{-1}K(\Omega_a, F) \ge 1 - \epsilon.$$

Since ϵ is arbitrary, (2.5) will follow from (a) and (b).

Suppose that the collection of df's $\{G_a\}$ satisfies $K(G_a, F) = K(\Omega_a, F)$ and $T(G_a) = a$ (by the first part of the proof such a collection exists). By (a) we have $K(G_a, F) \le a + o(a)$, as $a \downarrow 0$. Hence, by (1.26),

(c) $$d_F(G_a, F) \le k\sqrt{a},$$

uniformly in a, where $k > 0$ is some fixed number.

Let $\eta \in (0, 1/2)$, then $\int_0^\eta (G_a(t) - t)^2(t(1-t))^{-1}\,dt < k^2 a \int_0^\eta \{t(1-t)\log^2(t(1-t))\}^{-1}\,dt$. Thus by choosing $\eta > 0$ sufficiently small, we obtain $\int_0^\eta (G_a(t) - t)^2\,dt \le \epsilon a/2$, uniformly in a. By the same argument, we may choose $\eta > 0$ such that $\int_{1-\eta}^1 (G_a(t) - t)^2(t(1-t))^{-1}\,dt \le \epsilon a/2$.

Define the df H_a by

$$H_a(t) = \begin{cases} \eta^{-1} t G_a(\eta), & t \in [0, \eta) \\ G_a(t), & t \in [\eta, 1-\eta] \\ G_a(1-\eta) + \eta^{-1}(t - 1 + \eta)(1 - G_a(1-\eta)), & t \in (1-\eta, 1]. \end{cases}$$

Then $K(H_a, F) \le K(G_a, F)$ and $H_a \in \Omega_{(1-\epsilon)a}$. The df's H_a have the property

$$\text{(d)} \qquad \sup_{t\in(0,1)} \frac{|H_a(t) - t|}{t(1-t)} \le M\sqrt{a}$$

where M is some fixed number independent of a, since $|H_a(t) - t| = |G_a(t) - t| \le K\sqrt{a}/\psi(\eta)$ by (c) if $t \in [\eta, 1-\eta]$ and

$$|H_a(t) - t| < \frac{t(1-t)}{\eta(1-\eta)} \max\{|G_a(\eta) - \eta|, |G_a(1-\eta) - (1-\eta)|\}, \qquad \text{if} \quad t \in [\eta, 1-\eta].$$

Using (d), we now prove that $K(H_a, F) \ge (1-\epsilon)a + o(a)$, as $a \downarrow 0$. Define the function $\phi(t, y, z)$ by

$$\phi(t, y, z) = (1+z)\log(1+z) - y^2/(t/(1-t))$$

where $t \in (0, 1)$, $y \in [-1, 1]$ and $z \ge -1$. We shall prove that

$$\int_0^1 \phi(t, H_a(t) - t, h_a(t) - 1)\,dt \ge o(a) \quad \text{as} \quad a \downarrow 0,$$

where h_a is the density of the df H_a. Let $p(t, y) = y(1-2t)/(t(1-t))$, $t \in (0, 1)$ and $y \in [-1, 1]$, and let

$$\text{(e)} \qquad I(H_a) = \int_0^1 \{\phi(t, y_a, p) - p\phi_z(t, y_a, p) + z_a p_z(t, y_a, p)\}\,dt$$

where

$$\phi_z(t, y, z) = \frac{\partial}{\partial z}\phi(t, y, z),\ y_a = y_a(t) = H_a(t) - t,\ z_a = z_a(t) = H_a(t) - 1$$

and

$$p = p(t, y_a(t)).$$

(It is true, but not needed in this proof, that the functions $p(t, y)$ correspond to an "approximate" field of extremals for the function $\phi(t, y, z)$, since they correspond to approximate solutions $f(t) = \alpha t(1-t)$ of the differential equation

$$y'' + \frac{2}{t(1-t)}\{y + yy'\} = 0,$$

which is the Euler-Lagrange equation for the minimization of the functional $\int_0^1 \phi(t, y(t), y'(t))\,dt$.)

Expansion of $\log(1 + p(t, y_a(t)))$ and integration by parts twice yields

$$\int_0^1 z_a(t)\phi_z(t, y_a, p)\,dt = \int_0^1 z_a(t)\log(1 + p(t, y_a(t)))\,dt$$

$$= \int_0^1 z_a(t)y_a(t)\frac{1-2t}{t(1-t)}\,dt - \frac{1}{2}\int_0^1 z_a(t)y_a^2(t)\left(\frac{1-2t}{t(1-t)}\right)^2 dt + O(a^{3/2})$$

$$= \frac{1}{2}\int_0^1 p^2(t, y_a(t))\,dt + \int_0^1 \frac{y_a^2}{t(1-t)}\,dt + O(a^{3/2}), \qquad \text{as} \quad a \downarrow 0.$$

Here we used (d). Moreover, again by expanding $\log(1+p)$ it is seen that

$$\int_0^1 \{\phi(t, y_a, p) - p\phi_z(t, y_a, p)\}\,dt = \int_0^1 [\log(1+p) - p - y_a^2/(t(1-t))]\,dt$$

$$= -\frac{1}{2}\int_0^1 p^2(t, y_a(t))\,dt - \int_0^1 \frac{y_a^2(t)}{t(1-t)}\,dt + O(a^{3/2}).$$

Hence $I(H_a) = O(a^{3/2})$. By the mean value theorem we have

$$\phi(t, y_a, z_a) - \phi(t, y_a, p) - (z_a - p)\phi_z(t, y_a, p) = ½\,(z_a - p)^2\phi_{zz}(t, y_a, \theta),$$

where θ is a point between $z_a = z_a(t)$ and $p = p(t, y_a(t))$, if $z_a \neq p$, and may be taken arbitrarily (> -1) otherwise (as usual we denote $(\partial/\partial z^2)\,\phi$ by ϕ_{zz}). Thus $\int_0^1 \phi(t, y_a(t), z_a(t))\,dt - I(H_a) \geq 0$, implying $\int_0^1 \phi(t, y_a(t), z_a(t))\,dt \geq o(a)$, as $a \downarrow 0$. Since

$$\int_0^1 \frac{y_a^2(t)}{t(1-t)}\,dt = \int_0^1 \frac{(H_a(t) - t)^2}{t(1-t)}\,dt \geq (1-\epsilon)a,$$

this implies $K(H_a, F) \geq (1-\epsilon)a + o(a)$, $a \downarrow 0$. Hence we may conclude $\liminf_{a\downarrow 0} a^{-1}K(G_a, F) \geq \liminf_{a\downarrow 0} a^{-1}K(H_a, F) > 1 - \epsilon$, which we set out to prove. □

REMARK 10. In the last part of the proof of Corollary 2, it was actually shown that Weierstrass' sufficiency conditions for a strong minimum of the functional $\int_0^1 \phi(t, G(t) - t, g(t) - 1)\,dt$, g the density of G, are "approximately" satisfied for the df's $\bar{G}_a(t) = t + t(1 - t)\sqrt{6a}$. Note that we only have to consider values of the function $\phi(t, y, z)$ with $z \geq -1$, since $g(t) - 1 \geq -1$ for each density g. For Weierstrass' sufficiency conditions, see e.g. Gelfand and Fomin (1963), page 148.

PROOF OF COROLLARY 3. Let the sets D_m be defined as in Theorem 2. We first show that the integrals

$$\int_0^1 J(u)G^{-1}(u)\,du \quad \text{are well-defined for} \quad G \in D_m,\ m > 0.$$

Fix $m > 0$, $G \in D_m$ and let $\psi = \psi_2$ (see (1.6)). Then

$$\sup_{t\in(0,1)} (t - G(F^{-1}(t)))\psi(t) < m.$$

There exists $\delta > 0$ such that $t - m/\psi(t) \geq 1 - 2m/\psi(t)$, if $t \in [1-\delta, 1)$. Hence, for these values of t

$$G^{-1}(1 - 2m/\psi(t)) \leq G^{-1}\left(t - \frac{m}{\psi(t)}\right) \leq F^{-1}(t).$$

Putting $u = u(t) = 1 - 2m/\psi(t)$ (so $1 - t \geq t(1-t) = \exp(-2m/(1-u))$, we obtain for all $u(t)$ such that $t \in [1-\delta, 1)$ (this is, for all u in some interval $[1-\delta', 1)$)

(a) $$G^{-1}(u) \leq F^{-1}(1 - \exp(-2m/(1-u))).$$

Likewise there exists $\delta'' > 0$ such that

(b) $$G^{-1}(u) \geq F^{-1}(\exp(-2m/u))$$

for all $u \in (0, \delta'']$.

Hence, by (a) and (b) and conditions (i) and (ii):

$$\int_0^1 |J(u)G^{-1}(u)|\, du < \infty.$$

Moreover, because of a common dominating function, the above argument shows that the integrals of JG^{-1} over $(0, \eta]$ and $[1-\eta, 1)$ tend to zero as $\eta \downarrow 0$, uniformly in $G \in D_m$.

For fixed $\eta \in (0, 1/2)$ the function $T_\eta: D \to \mathbb{R}$ defined by

$$T_\eta(G) = \int_\eta^{1-\eta} J(u)G^{-1}(u)\, du, \qquad G \in D,$$

is τ-continuous on D (cf. Section 6 of GOR (1979)). This fact, together with the uniform integrability of $J(u)G^{-1}(u)$ for $G \in D_m$, implies that the function $T: D \to \mathbb{R}$ defined by

$$T(G) = \begin{cases} \int_0^1 J(u)G^{-1}(u)\, du, & \text{if } G \in D_m \quad \text{for some} \quad m \geq 0 \\ 0, \text{ otherwise} & \end{cases}$$

is τ-continuous on each set D_m.

Thus (1.22) of Theorem 2 holds, if $K(\Omega_r, F) < \infty$. It remains to show in this case only that $t \to K(\Omega_t, F)$ is a continuous mapping. If $K(\Omega_r, F) < \infty$, then the continuity from the right at $t = r$ can be proved by the argument of "displacement of mass" used in GOR (1979), page 579. Since, if $K(\Omega_r, F) < \infty$, we have by (1.24) that $K(\Omega_r, F) = K(\Omega_r \cap D_m, F)$ for m sufficiently large, and moreover T is τ-continuous on D_m, the left continuity of the mapping $t \to K(\Omega_t, F)$ in $t = r$ follows from Lemma 3.3 in GOR (1979). If $K(\Omega_r, F) = \infty$, then (3.2) must still be established. In this case this mapping is certainly continuous from the right at $t = r$ (by monotonicity). It is continuous from the left since trivially $K(\Omega_r \cap D_m, F) = K(\Omega_r, F)$ for each $m > 0$, and then apply Lemma 3.3 of GOR (1979) to get left continuity of $K(\Omega_r \cap D_m, F)$.

Let $r_0 = \inf\{t: K(\Omega_t, F) = \infty\}$ (possibly $r_0 = \infty$). Then $\lim_{t \uparrow r_0} K(\Omega_t, F) = \infty$ and hence, by the last paragraph, we have

$$\begin{aligned} \limsup_{n\to\infty} n^{-1}\log P(T(\hat{F}_n) \geq r) &\leq \lim_{t\uparrow r_0} \limsup_{n\to\infty} n^{-1}\log P(T(\hat{F}_n) \geq t) \\ &\leq \lim_{t\uparrow r_0} (-K(\Omega_t, F)) = -\infty \end{aligned}$$

using the case already proved. This completes the proof. □

Acknowledgement. We wish to thank the referee for pointing out relevant references [14] and [15].

REFERENCES

[1] Abrahamson, I. G. (1967). The exact Bahadur efficiencies for the Kolmogorov-Smirnov and Kuiper one- and two-sample statistics. *Ann. Math. Statist.* **38** 1475–1490.

[2] Bahadur, R. R. (1971). *Some Limit Theorems in Statistics*. SIAM, Philadelphia.

[3] Bahadur, R. R. and Zabell, S. L. (1979). Large deviations of the sample mean in general vector spaces. *Ann. Probability* **7** 587–621.

[4] Borovkov, A. A. (1967). Boundary value problems for random walks and large deviations in function spaces. *Theor. Probability Appl.* **12** 575–595.

[5] EICKER, F. (1979). The asymptotic distribution of the suprema of the standardized empirical processes. *Ann. Statist.* **7** 116–138.
[6] GELFAND, I. M. and FOMIN, S. V. (1963). *Calculus of Variations.* Prentice-Hall, New York.
[7] GROENEBOOM, P. and OOSTERHOFF, J. (1977). Bahadur efficiency and probabilities of large deviations. *Statist. Neerlandica* **31** 1-24.
[8] GROENEBOOM, P., OOSTERHOFF, J. and RUYMGAART, F. H. (1979). Large deviation theorems for empirical probability measures. *Ann. Probability* **7** 553-586.
[9] GROENEBOOM, P. and SHORACK, G. R. (1980). Large deviations of goodness of fit statistics and linear combinations of order statistics. Tech. Report No. 1, Department of Statistics, Univ. of Washington, Seattle.
[10] HOADLEY, A. B. (1967). On the probability of large deviations of functions of several empirical cdf's. *Ann. Math. Statist.* **38** 360–381.
[11] JAESCHKE, D. (1979). The asymptotic distribution of the supremum of the standardized empirical distribution function on subintervals. *Ann. Statist.* **7** 108–115.
[12] MOGUL'SKII, A. A. (1977). Remarks on large deviations for the ω^2 statistic. *Theor. Probability Appl.* **12** 166–171.
[13] NIKITIN, J. J. (1976). Bahadur relative asymptotic efficiency of statistics based on the empirical distribution function. *Soviet Math. Dokl.* **17** 1645-1649.
[14] NIKITIN, J. J. (1979). Large deviations and asymptotic efficiency of integral-type statistics I. *Zap. Naučn. Sem. Leningrad Otdel. Math. Inst. Steklov LOMI* **85** 175-187 (in Russian).
[15] NIKITIN, J. J. (1980). Large deviations and asymptotic efficiency of integral-type statistics II. To appear in *Zap. Naučn. Sem. Leningrad Otdel. Math. Inst. Steklov LOMI* **86** (in Russian).
[16] OOSTERHOFF, J. (1978). Large deviations of multivariate L-estimators with monotone weight functions. Report SW 63-78, Mathematical Centre, Amsterdam.
[17] SANOV, I. N. (1957). On the probability of large deviations of random variables. *Mat. Sb.* **42** 11–44 (English Translation in *Sel. Transl. Math. Statist. Prob.* **1** (1961), 213–244).
[18] SIEVERS, G. L. (1976). Probabilities of large deviations for empirical measures. *Ann. Statist.* **4** 766–770.
[19] STONE, M. (1974). Large deviations of empirical probability measures. *Ann. Statist.* **2** 362–366.
[20] VAINBERG, M. M. and TRENOGIN, V. A. (1974). *The Theory of Branching of Solutions of Non-Linear Equations.* Wolters-Noordhoff, Groningen.

DEPARTMENT OF STATISTICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195