stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE STATISTIEK SW 47/76 JULI
(DEPARTMENT OF MATHEMATICAL STATISTICS)

P. GROENEBOOM & J. OOSTERHOFF

BAHADUR EFFICIENCY AND PROBABILITIES OF LARGE DEVIATIONS

Prepublication

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

AMS(MOS) subject classification scheme (1970): 62G20, 60F10

Bahadur efficiency and probabilities of large deviations*)

by

P. Groeneboom & J. Oosterhoff **)

SUMMARY

The relative performance of two statistical tests of a hypothesis for large sample sizes is often investigated by means of asymptotic relative efficiencies in the sense of Pitman or Bahadur. Pitman efficiencies are directed at local alternatives, Bahadur efficiencies at fixed (non-local) alternatives. The present paper reviews some methods and results on Bahadur efficiency with special attention to probabilities of large deviations, which play a key role in the computation of Bahadur efficiencies.

KEY WORDS & PHRASES: *Bahadur efficiency, large deviation probabilities*

---

*) This report will be submitted for publication elsewhere.

**) J. Oosterhoff - Vrije Universiteit

Let $\Theta_0 \subset \Theta$ and suppose H: $\theta \in \Theta_0$ is a hypothesis to be tested against $\theta \in \Theta_1 = \Theta\backslash\Theta_0$. If n observations $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ are available and $\underline{t}_n$ is a test statistic based on these observations, $\underline{t}_n = t_n(\underline{s})$ depends on $\underline{s}$ only through its first n coordinates. Without essential loss of generality we assume that H is to be rejected for large values of $\underline{t}_n$. For each real c we obtain a test of H by considering the critical region $\{s: t_n(s) \geq c\}$. The *power function* of such a test is $P_\theta(\underline{t}_n \geq c)$ as a function of θ and its *size* is defined by

$$\sup \{P_\theta(\underline{t}_n \geq c): \theta \in \Theta_0\}.$$

For $0 < \beta < 1$ and $\theta \in \Theta_1$ let $c_n = c_n(\beta,\theta)$ be defined by

$$(1.1) \qquad P_\theta(\underline{t}_n > c_n) \leq \beta \leq P_\theta(\underline{t}_n \geq c_n).$$

Then

$$(1.2) \qquad \alpha_n(\beta,\theta) := \sup \{P_{\theta'}(\underline{t}_n \geq c_n): \theta' \in \Theta_0\}$$

is the minimal size of a one-sided test based on $\underline{t}_n$ for which the power at θ is at least equal to β.

So far we have considered a fixed sample size n. Suppose that for each $n \in \mathbb{N}$ a test statistic $\underline{t}_n$ is given. By an appropriate choice of critical values tests of H may be obtained for each sample size at any desired significance level α. For a given sequence $t = \{\underline{t}_n\}$, $0 < \alpha < \beta < 1$ and $\theta \in \Theta_1$ we define

$$(1.3) \qquad N_t(\alpha,\beta,\theta) := \min \{n: \alpha_m(\beta,\theta) \leq \alpha \text{ for all } m \geq n\},$$

i.e. the minimal required sample size of level-α tests based on $\{\underline{t}_n\}$ with power $\geq \beta$ at θ.

Suppose we have two sequences of test statistics, $\{\underline{t}_n\}$ and $\{\underline{\tilde{t}}_n\}$, for testing H. The *relative efficiency* of $\{\underline{\tilde{t}}_n\}$ with respect to $\{\underline{t}_n\}$ is then defined as

(1.4) $$e_{\tilde{t},t}(\alpha,\beta,\theta) := N_t(\alpha,\beta,\theta)/N_{\tilde{t}}(\alpha,\beta,\theta),$$

where $0 < \alpha < \beta < 1$ and $\theta \in \Theta_1$. A value of the relative efficiency (1.4) larger than one indicates that, for the given α,β,θ, the test statistics $\{\tilde{\underline{t}}_n\}$ are to be prefered to $\{\underline{t}_n\}$, since with $\{\tilde{\underline{t}}_n\}$ fewer observations are needed to attain a power β at θ for the given level α than with $\{\underline{t}_n\}$.

Although (1.4) is a good measure for comparing the relative merits of two sequences of test statistics, it is not very well suited for practical use because of its dependence on three arguments. Moreover, it is often hard to compute. To avoid these difficulties, various limiting procedures have been suggested to obtain more manageable measures. All these limiting procedures have in common that the sample sizes N appearing in (1.4) tend to infinity. We mention three such procedures:

(i) keeping α and β fixed, let $\theta \to \theta_0 \in \Theta_0$;
(ii) keeping β and θ fixed, let $\alpha \downarrow 0$;
(iii) keeping α and θ fixed, let $\beta \uparrow 1$.

To each of these limiting procedures corresponds a concept of asymptotic relative efficiency.

DEFINITION 1.1. Let $\{\underline{t}_n\}$ and $\{\tilde{\underline{t}}_n\}$ be two sequences of test statistics for testing H, where $\underline{t}_n$ and $\tilde{\underline{t}}_n$ are based on n observations (n=1,2,...).

(i) If $\theta_0 \in \Theta_0$ is a boundary point of Θ_0 (in some topology on Θ),

$$e^{P}_{\tilde{t},t}(\alpha,\beta,\theta_0) := \lim_{\theta\to\theta_0} e_{\tilde{t},t}(\alpha,\beta,\theta) \qquad (0<\alpha<\beta<1)$$

is the *Pitman efficiency* of $\{\tilde{\underline{t}}_n\}$ with respect to $\{\underline{t}_n\}$,

(ii) $$e^{B}_{\tilde{t},t}(\beta,\theta) := \lim_{\alpha\downarrow 0} e_{\tilde{t},t}(\alpha,\beta,\theta) \qquad (0<\beta<1, \theta\in\Theta_1)$$

is called the *Bahadur efficiency* of $\{\tilde{\underline{t}}_n\}$ with respect to $\{\underline{t}_n\}$, and

(iii) $$e^{HL}_{\tilde{t},t}(\alpha,\theta) := \lim_{\beta\uparrow 1} e_{\tilde{t},t}(\alpha,\beta,\theta) \qquad (0<\alpha<1, \theta\in\Theta_1)$$

is called the *Hodges-Lehmann efficiency* of $\{\tilde{\underline{t}}_n\}$ with respect to $\{\underline{t}_n\}$, provided these limits exist.

These asymptotic relative efficiencies are not as difficult to compute as the efficiencies defined by (1.4). Since Pitman efficiency usually does not depend on α or β and Bahadur efficiency as a rule is independent of β, these efficiencies are also easier to interpret. Little experience has been had with Hodges-Lehmann efficiency and we will therefore not discuss it further.

Pitman efficiency, introduced by E.J.G. Pitman in 1949 [25], is a classical tool for comparison of the power of two tests at near alternatives for large sample sizes. On the other hand Bahadur efficiency, first proposed by R.R. Bahadur in 1960 [2], can serve as a yardstick to compare the performance of two tests at fixed alternatives for large samples (and small significance levels).

There is extensive numerical evidence that in many problems the (asymptotic) Pitman efficiency is also a good measure of the behaviour of the respective powers for moderate sample sizes. A less satisfactory feature of Pitman efficiency is its insensitivity to small power differences. As a result it often happens that two different tests for the same testing problem have Pitman efficiency one with respect to each other. One way of overcoming this difficulty is the introduction of deficiencies; in a recent paper [1] in this journal W. Albers has given an exposition of this concept and some of its results. An alternative approach is to compute Bahadur efficiencies in the hope that they may discriminate between the two tests.

It is a definite advantage of Bahadur efficiency that it is often more sensitive to differences in power than Pitman efficiency. On the other hand, since Bahadur efficiency usually is a non-constant function of θ, it is less convenient to work with. Another feature of Bahadur efficiency is that it is strongly influenced by the extreme tails of the distribution of the test statistic. Unfortunately, there is little numerical evidence so far that Bahadur efficiency is close to (1.4) for moderate values of α (and hence for moderate sample sizes).

Now let us introduce some notation and turn to the actual computation of asymptotic relative efficiencies. Throughout this paper Φ denotes the distribution function (df) of the standard-normal distribution, ϕ its density function and Φ^{-1} the inverse function of Φ. The indicator function of a set A is denoted by 1_A (i.e. $1_A(x) = 1$ if $x \in A$ and $= 0$ otherwise).

By $a \sim b$ we mean that the ratio of a and b tends to one under specified conditions.

In many testing problems Pitman efficiencies can be found with the aid of the following well-known result (cf. PITMAN [25]).

THEOREM 1.1. (Pitman). *Let* $\Theta = \mathbb{R}$, $\Theta_0 = (-\infty,\theta_0]$. *Let* $\{\underset{\sim}{t}_n\}$ *be a sequence of test statistics such that* $P_\theta(\underset{\sim}{t}_n \geq a)$ *is non-decreasing in* θ *for each* $a \in \mathbb{R}$ *and*

$$(1.5) \qquad \lim_{n\to\infty} P_{\theta_n}(n^{\frac{1}{2}}(\underset{\sim}{t}_n - \mu(\theta_n))/\sigma(\theta_n) \leq x) = \Phi(x)$$

for $x \in \mathbb{R}$ *and all* $\theta_n = \theta_0 + kn^{-\frac{1}{2}}$, $k \geq 0$, *where* $\mu(\theta)$ *is a function with a right-hand derivative* $\mu'(\theta_0) > 0$ *and* $\sigma(\theta)$ *is a function continuous from the right at* θ_0. *If* $\{\tilde{\underset{\sim}{t}}_n\}$ *is a second sequence of test statistics satisfying the same conditions, but with* μ *and* σ *replaced by* $\tilde{\mu}$ *and* $\tilde{\sigma}$, *then*

$$(1.6) \qquad e^P_{\tilde{t},t}(\alpha,\beta,\theta_0) = e^P_{\tilde{t},t}(\theta_0) = \left\{\frac{\tilde{\mu}'(\theta_0)}{\tilde{\sigma}(\theta_0)} \Big/ \frac{\mu'(\theta_0)}{\sigma(\theta_0)}\right\}^2$$

for all $0 < \alpha < \beta < 1$.

Bahadur efficiencies may be derived by way of the next theorem.

THEOREM 1.2. (Bahadur). *If the sequence of test statistics* $\{\underset{\sim}{t}_n\}$ *satisfies*

$$(1.7) \qquad \lim_{n\to\infty} -\frac{1}{n} \log \alpha_n(\beta,\theta) = c^B_t(\beta,\theta) > 0$$

for some function c^B_t *of* β *and* θ $(0<\beta<1, \theta\in\Theta_1)$, *then*

$$(1.8) \qquad N_t(\alpha,\beta,\theta) \sim -(\log \alpha)/c^B_t(\beta,\theta) \qquad \textit{for } \alpha \downarrow 0.$$

Hence, if $\{\tilde{\underset{\sim}{t}}_n\}$ *satisfies the same condition, but with* c^B_t *replaced by* $c^B_{\tilde{t}}$,

$$(1.9) \qquad e^B_{\tilde{t},t}(\beta,\theta) = c^B_{\tilde{t}}(\beta,\theta)/c^B_t(\beta,\theta)$$

for all $0 < \beta < 1$ *and* $\theta \in \Theta_1$.

PROOF. Since (1.9) is an immediate consequence of (1.8), it suffices to prove (1.8). Write $N_\alpha = N_t(\alpha,\beta,\theta)$. We first show that $N_\alpha \to \infty$ as $\alpha \downarrow 0$. By (1.7) $-n^{-1} \log \alpha_n(\beta,\theta) < 2c_t^B(\beta,\theta)$, or equivalently, $\alpha_n(\beta,\theta) > \exp(-2nc_t^B(\beta,\theta))$ for all sufficiently large n. Hence, for any such n, $\alpha < \exp(-2nc_t^B(\beta,\theta))$ entails $N_\alpha \geq n$ and $N_\alpha \to \infty$ follows.

The definition of N_α implies $\alpha_{N_\alpha}(\beta,\theta) \leq \alpha < \alpha_{N_\alpha - 1}(\beta,\theta)$ or

$$-\frac{1}{N_\alpha} \log \alpha_{N_\alpha - 1}(\beta,\theta) < -\frac{1}{N_\alpha} \log \alpha \leq -\frac{1}{N_\alpha} \log \alpha_{N_\alpha}(\beta,\theta).$$

By (1.7) the extreme members tend to $c_t^B(\beta,\theta)$ as $\alpha \downarrow 0$, whence (1.8). □

As an illustration we consider a typical example.

EXAMPLE 1.1. Let $\underline{x}_1, \underline{x}_2, \underline{x}_3, \ldots$ be i.i.d. random variables with normal $N(\theta,1)$ distributions and suppose H: $\theta \leq 0$ is to be tested against $\theta > 0$. The UMP Gauss test based on n observations rejects H for large values of $\underline{\bar{x}}_n = n^{-1}\sum_{i=1}^n \underline{x}_i$. The Student t-test, which would be employed if the variance were unknown, rejects H for large values of $\underline{\tilde{t}}_n = \underline{\bar{x}}_n/\underline{s}_n$, where $\underline{s}_n^2$ is the sample variance.

Note that $\{\underline{\bar{x}}_n\}$ satisfies (1.5) with $\mu(\theta) \equiv \theta$ and $\sigma(\theta) \equiv 1$. Since $n^{\frac{1}{2}}(\underline{\tilde{t}}_n - \underline{\bar{x}}_n) = n^{\frac{1}{2}}\, \underline{\bar{x}}_n(1-\underline{s}_n)/\underline{s}_n \to 0$ in P_{θ_n}-probability for each sequence $\{\theta_n\} = \{kn^{-\frac{1}{2}}\}$ with $k \geq 0$, $\{\underline{\tilde{t}}_n\}$ also satisfies (1.5) for the same functions μ and σ. Hence, by theorem 1.1, the Pitman efficiency is

$$e^P_{\tilde{t},\bar{x}}(0) = 1,$$

although of course the Gauss test is more powerful than the t-test.

To find the Bahadur efficiency we first determine $c^B_{\bar{x}}(\beta,\theta)$. If c_n satisfies $P_\theta(\underline{\bar{x}}_n \geq c_n) = \beta$ (cf. (1.1)), then $\lim_{n\to\infty} c_n = \theta$ since $\underline{\bar{x}}_n \to \theta$ in P_θ-probability. Application of

$$1 - \Phi(y) \sim y^{-1}\phi(y) \quad \text{for } y \to \infty$$

yields

$$\alpha_n(\beta,\theta) = P_0(\underline{\bar{x}}_n \geq c_n) = 1 - \Phi(n^{\frac{1}{2}} c_n) \sim n^{-\frac{1}{2}}\theta^{-1}\phi(n^{\frac{1}{2}} c_n)$$

for $n \to \infty$. Hence

$$c^B_{\bar{x}}(\beta,\theta) = \lim_{n\to\infty} - n^{-1} \log \alpha_n(\beta,\theta) = \tfrac{1}{2}\theta^2. \tag{1.10}$$

Next we turn to the Student test. Again $\underline{\tilde{t}}_n \to \theta$ in P_θ-probability. It follows that $\tilde{c}_n$, determined by $P_\theta(\underline{\tilde{t}}_n \geq \tilde{c}_n) = \beta$, satisfies $\lim_{n\to\infty} \tilde{c}_n = \theta$. Writing $f_n(t)$ for the density of the t-distribution with n degrees of freedom and making use of the relation

$$\int_y^\infty f_n(t)dt \sim y^{-1}(1+n^{-1}y^2)\, f_n(y) \quad \text{for } y \to \infty,$$

we find

$$\tilde{\alpha}_n(\beta,\theta) = P_0(\underline{\tilde{t}}_n \geq \tilde{c}_n) = \int_{n^{\frac{1}{2}}\tilde{c}_n}^\infty f_{n-1}(t)dt \sim n^{-\frac{1}{2}}\theta^{-1}(1+\theta^2) f_{n-1}(n^{\frac{1}{2}}\tilde{c}_n)$$

for $n \to \infty$. Therefore

$$c^B_{\tilde{t}}(\beta,\theta) = \lim_{n\to\infty} - n^{-1} \log \tilde{\alpha}_n(\beta,\theta) = \tfrac{1}{2}\log(1+\theta^2).$$

Combining this result with (1.10) we obtain from theorem 1.2 that the Bahadur efficiency is equal to

$$e^B_{\tilde{t},\bar{x}}(\beta,\theta) = \theta^{-2} \log(1+\theta^2),$$

which is smaller than one for all $\theta > 0$, but tends to one for $\theta \downarrow 0$.

2. BAHADUR SLOPE

There is an intimate relationship between tail probabilities and Bahadur efficiency, which sheds new light on the whole concept of Bahadur efficiency. We start by exploring this relationship in some detail.

Consider the same setup as in section 1: the hypothesis H: $\theta \in \Theta_0$ is to be tested against $\theta \in \Theta_1$ on the basis of $\underline{t}_n = t_n(\underline{s})$, where t_n depends on the first n coordinates of $\underline{s} = (\underline{x}_1, \underline{x}_2, \underline{x}_3, \ldots)$, $n = 1,2,\ldots$, and H is to be rejected for large values of $\underline{t}_n$. For $n = 1,2,\ldots$ let

$$(2.1) \qquad G_n(t) := \inf\{P_\theta(\underline{t}_n < t) : \theta \in \Theta_0\}, \quad t \in \mathbb{R},$$

i.e. $1 - G_n(t) = \sup\{P_\theta(\underline{t}_n \geq t) : \theta \in \Theta_0\}$. The *tail probability* of the test based on $\underline{t}_n$ is defined as the random variable

$$(2.2) \qquad \underline{L}_n = L_n(\underline{s}) := 1 - G_n(t_n(\underline{s})).$$

Note that $\underline{L}_n$ only depends on $\underline{x}_1, \ldots, \underline{x}_n$. In practical statistical work tail probabilities are often computed for the actual values $x_1, \ldots, x_n$ realized in the experiment. If $L_n(s)$ is smaller than the given significance level α, the hypothesis H is rejected. Generally speaking, the smaller the value of $L_n(s)$, the more untenable is the hypothesis H in the light of the observations.

Under the hypothesis H the random variable $\underline{L}_n$ is often (approximately) uniformly distributed on (0,1). For $\theta \in \Theta_1$ the distribution of $\underline{L}_n$ is more concentrated near zero, especially for large n. In fact, it will often be true that, for $n \to \infty$ and $\theta \in \Theta_1$,

$$(2.3) \qquad -\frac{1}{n} \log \underline{L}_n \to c(\theta) \text{ in } P_\theta\text{-probability},$$

where $c(\theta)$ is a positive number. If (2.3) holds true, $2c(\theta)$ is called the *(exact) Bahadur slope* of the sequence $\{\underline{t}_n\}$. The factor 2 has an historical origin going back to K. Pearson and R.A. Fisher: under H the random variable $-2 \log \underline{L}_n$ is often approximately distributed as χ^2_2.

<u>THEOREM 2.1</u>. (Bahadur). *If for* $\theta \in \Theta_1$ (2.3) *holds for some* $c(\theta) > 0$,

$$(2.4) \qquad \lim_{n\to\infty} -n^{-1} \log \alpha_n(\beta,\theta) = c(\theta) \quad \textit{for all } 0 < \beta < 1.$$

<u>PROOF</u>. Fix $\theta \in \Theta_1$ and let (2.3) be satisfied. Suppose (2.4) does not hold

for some $\beta \in (0,1)$. Let c_n be defined by (1.1) for this β. Then there exist an increasing subsequence $\{n_k\}$ and an $\varepsilon > 0$ such that either

(a) $-n_k^{-1} \log \alpha_{n_k}(\beta,\theta) < c(\theta) - \varepsilon$ or

(b) $-n_k^{-1} \log \alpha_{n_k}(\beta,\theta) > c(\theta) + \varepsilon$ for all k.

In case (a)

$$\begin{aligned} P_\theta(-n_k^{-1}\log \underline{L}_{n_k} > c(\theta) - \varepsilon) &\le P_\theta(-n_k^{-1} \log \underline{L}_{n_k} > -n_k^{-1} \log \alpha_{n_k}(\beta,\theta)) \\ &= P_\theta(\underline{L}_{n_k} < \alpha_{n_k}(\beta,\theta)) \\ &= P_\theta(1 - G_{n_k}(\underline{t}_{n_k}) < 1 - G_{n_k}(c_{n_k})) \\ &\le P_\theta(\underline{t}_{n_k} > c_{n_k}) \le \beta \end{aligned}$$

for all k. But for $k \to \infty$ the left hand member of this string of inequalities tends to one in contradiction to $0 < \beta < 1$. Similarly a contradiction is obtained in case (b) and the proof is complete. □

Consider the function c_t^B appearing in theorem 1.2. If (2.3) is satisfied, theorem 2.1 implies that (1.7) is also satisfied and that $2c_t^B(\beta,\theta)$ is the Bahadur slope of the test statistics $\{\underline{t}_n\}$. Suppose a second sequence of test statistics $\{\underline{\tilde{t}}_n\}$ has Bahadur slope $\tilde{c}(\theta) > 0$. It follows that the Bahadur efficiency of $\{\underline{\tilde{t}}_n\}$ with respect to $\{\underline{t}_n\}$ is the ratio of their Bahadur slopes:

$$(2.5) \qquad e^B_{\tilde{t},t}(\beta,\theta) = \tilde{c}(\theta)/c(\theta) \qquad \text{for all } \beta \in (0,1).$$

Hence, if $e^B_{\tilde{t},t} > 1$, the tail probability of $\{\underline{\tilde{t}}_n\}$ tends faster to zero than that of $\{\underline{t}_n\}$ for $n \to \infty$.

It remains to find simple ways to compute Bahadur slopes of sequences of test statistics $\{\underline{t}_n\}$. Suppose that the following two conditions are satisfied: for $n \to \infty$

$$(2.6) \qquad n^{-\frac{1}{2}}\underline{t}_n \to b(\theta) \text{ in } P_\theta\text{-probability } (\theta\in\Theta_1),$$

(2.7) $\quad - n^{-1} \log \{1 - G_n(n^{\frac{1}{2}}t)\} \to f(t) \quad (\forall t \in \mathbb{R}),$

where $b(\theta)$ is an arbitrary real number, G_n is defined in (2.1) and f is an arbitrary nonnegative function continuous at $t = b(\theta)$. It is not difficult to show that conditions (2.6) and (2.7) imply

(2.8) $\quad - n^{-1} \log \underline{L}_n \to f(b(\theta))$ in P_θ-probability,

i.e. the Bahadur slope of $\{\underline{t}_n\}$ is $2f(b(\theta))$.

Condition (2.7) shows that in computing Bahadur slopes we are concerned with probabilities of large deviations. In section 3 a survey of some results in this field will be given.

In the first papers on Bahadur efficiency attention was focussed on *approximate Bahadur efficiency*. To introduce this concept, suppose that G_n is a left continuous df converging weakly to a left continuous df G, for $n \to \infty$. Define the approximate tail probability of $\underline{t}_n$ by

$$\underline{L}_n^* := 1 - G(\underline{t}_n), \quad n = 1,2,...$$

(note that in many practical problems such approximate tail probabilities are actually computed). In addition suppose that for $n \to \infty$

$$- n^{-1} \log \underline{L}_n^* \to c^*(\theta) \text{ in } P_\theta\text{-probability } (\theta \in \Theta_1).$$

Then $2c^*(\theta)$ is called the *approximate Bahadur slope* of $\{\underline{t}_n\}$. The approximate Bahadur efficiency of $\{\underline{\tilde{t}}_n\}$ with respect to $\{\underline{t}_n\}$ is now defined as the ratio of the respective approximate Bahadur slopes, cf. (2.5).

It is again easily shown that (2.6) and

(2.9) $\quad - n^{-1} \log\{1 - G(n^{\frac{1}{2}}t)\} \to f(t) \quad (\forall t \in \mathbb{R}),$

for $n \to \infty$, imply that the approximate Bahadur slope of $\{\underline{t}_n\}$ is $2f(b(\theta))$. Since (2.9) is often trivially satisfied (e.g. if $G=\Phi$), approximate Bahadur slopes are easier to handle than exact slopes, explaining the original preference for approximate Bahadur efficiencies. In addition, under conditions slightly stronger than those of theorem 1.1 ((1.5) must also hold for fixed $\theta \in \Theta_1$), the approximate Bahadur efficiency $\tilde{c}^*(\theta)/c^*(\theta)$ of $\{\tilde{t}_n\}$ with respect to $\{t_n\}$ exists for all $\theta \in \Theta_1$ and tends to the Pitman efficiency as $\theta \downarrow \theta_0$. This nice property gave further support to the use of approximate Bahadur efficiencies.

However, it is well-known that monotone transformations of $\{t_n\}$ and $\{\tilde{t}_n\}$ may lead to entirely different approximate Bahadur efficiencies (cf. BAHADUR [4] and GROENEBOOM *et al.* [12]), although of course the corresponding tests remain invariant under such transformations. Hence the whole concept of approximate Bahadur efficiency is of little value and we will not consider it any further.

Incidentally we remark that under appropriate conditions exact Bahadur efficiencies also tend to the Pitman efficiency for alternatives tending to the hypothesis, cf. example 1.1. Unfortunately, the conditions are more severe than in the case of approximate Bahadur efficiencies.

Consider a sequence of test statistics $\{t_n\}$ for testing H: $\theta \in \Theta_0$. This sequence is said to be *efficient in the sense of Bahadur* if the Bahadur slope of $\{t_n\}$ is a maximum among all sequences of test statistics. If we could find an upper bound for the Bahadur slope of sequences of test statistics for testing H, the sequence $\{t_n\}$ is obviously efficient in the sense of Bahadur if its slope is equal to this upper bound. This approach will be further explored in the remainder of this section.

Let the family of distributions $\{P_\theta: \theta \in \Theta\}$ of the observations be dominated by a σ-finite measure μ and let $p_\theta = dP_\theta/d\mu$ be the density of P_θ with respect to μ. Define

$$(2.10) \qquad K(\theta,\theta') := \int p_\theta \log(p_\theta/p_{\theta'})d\mu,$$

the *Kullback-Leibler information number* of P_θ with respect to $P_{\theta'}$. Note that $0 \le K \le \infty$ (in view of Jensen's inequality), $K = 0$ iff $P_\theta = P_{\theta'}$ and $K = \infty$ if P_θ is not absolutely continuous with respect to $P_{\theta'}$. Furthermore let

$$(2.11) \qquad K(\theta,\Theta_0) := \inf\{K(\theta,\theta'): \theta' \in \Theta_0\}.$$

First consider the case of a simple hypothesis H_0: $\theta = \theta_0$. Let $\alpha_n^+(\beta,\theta)$ denote the minimal size of the most powerful (MP) test of H_0 against a simple alternative θ based on n observations with power $\geq \beta$ at θ. C. Stein has shown (cf. BAHADUR [5]) that without any condition whatsoever

$$\lim_{n\to\infty} - n^{-1} \log \alpha_n^+(\beta,\theta) = K(\theta,\theta_0) \qquad \text{for all } \beta \in (0,1).$$

Hence the Bahadur slope of the sequence of MP tests is $2K(\theta,\theta_0)$.

Now let Θ_0 be an arbitrary composite hypothesis. Since the power at θ of any test of H can never exceed the power of an MP test of H_0: $\theta = \theta_0$ $(\theta_0 \in \Theta_0)$ against the simple alternative θ, it immediately follows from the preceding result that the Bahadur slope $2c(\theta)$ of a sequence of test statistics, if it exists, satisfies the inequality

$$(2.12) \qquad 2c(\theta) \leq 2K(\theta,\Theta_0)$$

for all $\theta \in \Theta_1$. If the Bahadur slope of $\{\underline{t}_n\}$ does not exist, we can replace (2.12) by

$$P_\theta(\limsup_{n\to\infty} - n^{-1} \log \underline{L}_n \leq K(\theta,\Theta_0')) = 1,$$

which has been proved by RAGHAVACHARI [28].

A general method for testing a composite hypothesis is furnished by likelihood ratio tests. It is well-known that for many testing problems likelihood ratio tests have satisfactory power properties, especially for large sample sizes. BAHADUR [3], [4] has shown that under certain regularity conditions the Bahadur slope of the sequence of likelihood ratio test statistics exists and attains the upper bound $2K(\theta,\Theta_0)$ in (2.12) for all $\theta \in \Theta_1$. Hence, under these conditions, likelihood ratio tests are efficient in the sense of Bahadur.

BROWN [9] has obtained slightly stronger results of the same type for appropriately modified likelihood ratio tests. He has also shown that these tests are efficient in the sense of Hodges-Lehmann under very general regularity conditions. More detailed results have been obtained by HOEFFDING [18] for likelihood ratio tests and chi-squared tests in multinomial families of distributions.

A quite general framework for obtaining efficient tests in the sense of Bahadur is discussed in BAHADUR-RAGHAVACHARI [6]. In this paper it is also investigated what sort of conditioning is helpful in making conditional tests.

3. PROBABILITIES OF LARGE DEVIATIONS

In this section some limit theorems on probabilities of large deviations are reviewed.

Let $\{\underline{t}_n\}$ be a sequence of real-valued statistics defined on a probability space $(\Omega,\mathcal{T},\mathcal{P})$ and let $\{t_n\}$ be a sequence of real numbers. If

$$I = \lim_{n\to\infty} n^{-1} \log \mathcal{P}(\underline{t}_n \geq t_n)$$

exists and $-\infty < I < 0$, the events $\{\underline{t}_n \geq t_n\}$ with n large will be called *large deviations*. Thus probabilities of large deviations deal with the extreme (right-hand) tail of the distribution of $\underline{t}_n$. In view of (2.7) such extreme tail probabilities are needed to determine Bahadur slopes of tests based on the statistics $\{\underline{t}_n\}$.

In example 1.1 asymptotic expressions for probabilities of large deviations were derived in two particular cases. Here we describe subsequently two general approaches to obtain asymptotic expressions for probabilities of large deviations: the moment generating function approach and the empirical distribution function approach. A third method based on probability densities will not be discussed here; we refer to KILLEEN *et al.* [21] and SIEVERS [31].

3.1. *The moment generating function approach*

The best known result on the relationship between probabilities of large deviations and moment generating functions is the following theorem of CHERNOFF [10].

<u>THEOREM 3.1</u>. (Chernoff). *Let* $\underline{x}_1,\underline{x}_2,\underline{x}_3,\ldots$ *be real-valued i.i.d. random variables. Then, for each* $t \in \mathbb{R}$,

$$(3.1)\qquad \lim_{n\to\infty} n^{-1} \log P(n^{-1} \sum_{i=1}^{n} \underline{x}_i \ge t) = \inf_{\tau\ge 0} \{\log(Ee^{\tau\underline{x}_1}) - \tau t\}.$$

Note that this theorem holds without any conditions at all. If $E\exp(\tau\underline{x}_1) = \infty$ for all $\tau > 0$, the infimum in the right-hand member is obviously assumed for $\tau = 0$ and is equal to zero. However, the generality of the theorem is counterbalanced by the weakness of the statement (3.1) which gives rise to approximations of the probability of the event $\{n^{-1}\sum_1^n \underline{x}_i \ge t\}$ up to factors $\exp(o(n))$ only.

Proofs of theorem 3.1 have been given by CHERNOFF [10], BAHADUR & RANGA RAO [7], BAHADUR [5] and HAMMERSLEY [15]. These proofs are all based on somewhat different ideas. We present a sketch of a proof in the spirit of Chernoff's original proof.

SKETCH OF PROOF OF THEOREM 3.1.

By Markov's inequality

$$P(\sum_{i=1}^{n} \underline{x}_i \ge nt) = P(\exp(\tau \sum_{i=1}^{n} \underline{x}_i) \ge \exp(n\tau t)) \le$$
$$\le \exp(-n\tau t)\cdot(E\exp(\tau\underline{x}_1))^n$$

for $\tau > 0$. The inequality is trivially satisfied for $\tau = 0$. Hence

$$\limsup_{n\to\infty} n^{-1} \log P(\sum_{i=1}^{n} \underline{x}_i \ge nt) \le \inf_{\tau\ge 0} \{\log(Ee^{\tau\underline{x}_1}) - \tau t\}.$$

We still have to show that conversely

$$(3.2)\qquad \liminf_{n\to\infty} n^{-1} \log P(\sum_{i=1}^{n} \underline{x}_i \ge nt) \ge \inf_{\tau\ge 0} \{\log(Ee^{\tau\underline{x}_1}) - \tau t\}.$$

Suppose the distribution of $\underline{x}_1$ has finite support: $P(\underline{x}_1 = a_j) = p_j > 0$ $(j=1,\ldots,k)$, $p_1 + \ldots + p_k = 1$. Moreover, let $E\underline{x}_1 \le t < \max\{a_j : 1 \le j \le k\}$. Define

$$\alpha(\tau) = \log(Ee^{\tau\underline{x}_1}) - \tau t \qquad (\tau\ge 0).$$

A little algebra shows that there exists a unique $\tau_1 \ge 0$ satisfying

If $t < E\underline{x}_1$, it is easily seen that both members of (3.2) are equal to zero. The particular cases $t = \max a_j$ and $t > \max a_j$ can also easily be dealt with directly.

Thus relation (3.2) is proved for distributions with finite support. The general case is proved by approximating the distribution of $\underline{x}_1$ by distributions with finite support. □

EXAMPLE 3.1. Let $\underline{x}_1, \underline{x}_2, \underline{x}_3, \ldots$ be i.i.d. random variables with normal $N(\theta,1)$ distributions and consider the testing problem of example 1.1. Let $\underline{x}_{m(n)}$ denote the sample median of a sample of size n. Now Chernoff's theorem implies

$$\begin{aligned}\lim_{n\to\infty} n^{-1} \log P_0(\underline{x}_{m(n)} \geq t) &= \lim_{n\to\infty} n^{-1} \log P_0\Big(\sum_{i=1}^{n} 1_{[t,\infty)}(\underline{x}_i) \geq \tfrac{1}{2}n\Big)\\ &= \inf_{\tau\geq 0}(\log\{\Phi(t)+e^{\tau}(1-\Phi(t))\}-\tfrac{1}{2}\tau)\\ &= \tfrac{1}{2}\log\,\{4\Phi(t)(1-\Phi(t))\}.\end{aligned}$$

It is also easily exhibited that $\underline{x}_{m(n)} \to \theta$ in P_θ-probability for $n \to \infty$. Hence (2.6) and (2.7) are satisfied and (2.8) yields that the Bahadur slope of the test based on the sample median is equal to

$$-\log\,\{4\Phi(\theta)(1-\Phi(\theta))\}.$$

It follows (cf. example 1.1) that the Bahadur efficiency of the test based on the sample median with respect to the Gauss test is equal to

$$-\theta^{-2}\log\,\{4\Phi(\theta)(1-\Phi(\theta))\}.$$

For $\theta \downarrow 0$ the efficiency tends to $2/\pi$ (= Pitman efficiency) and for $\theta \to \infty$ it tends to $\frac{1}{2}$.

Refinements of Chernoff's theorem under additional regularity conditions on the underlying distributions have been given by BAHADUR & RANGO RAO [7] and PETROV [24]. In these papers only large deviations of sample means are treated.

Large deviation theorems for more general statistics than the sample mean have been obtained by SIEVERS [30], PLACHKY [26] and PLACHKY & STEINEBACH [27], who also used the moment generating function technique. We mention the most recent - and strongest - result.

THEOREM 3.2. (Plachky & Steinebach). *Let* $\{\underline{t}_n\}$ *be a sequence of real-valued random variables and let* $m_n(\tau) = E\exp(\tau\underline{t}_n)$ *and* $\psi_n(\tau) = \log m_n(\tau)$. *Assume real numbers* T_0 *and* T_1 *exist*, $0 \le T_0 < T_1$, *such that the following three conditions are satisfied:*

(i) $m_n(\tau) < \infty$ *for all* $\tau \in [0,T_1)$,

(ii) $\lim_{n\to\infty} \psi_n(\tau)/n = c_0(\tau) \in \mathbb{R}$ *exists for all* $\tau \in (T_0,T_1)$,

(iii) $c_0'(\tau)$ *exists, is continuous from the right and strictly monotone for all* $\tau \in (T_0,T_1)$.

Let $\{a_n\}$ *be a sequence of real numbers satisfying* $\lim_{n\to\infty} a_n = a \in \{c_0'(\tau): \tau \in (T_0,T_1)\}$. *Then*

$$\lim_{n\to\infty} n^{-1} \log P(\underline{t}_n \ge na_n) = \inf_{\tau>0} \{c_0(\tau) - \tau a\}.$$

3.2. *The empirical distribution function approach*

Let $\underline{x}_1,\underline{x}_2,\underline{x}_3,\ldots$ be i.i.d. random variables with df F. Let $\hat{\underline{F}}_n$ denote the empirical df of $\underline{x}_1,\ldots,\underline{x}_n$, i.e. $\hat{\underline{F}}_n(x)$ is the fraction of the n random variables $\le x$ ($x\in\mathbb{R}$). Almost all test statistics which are in common use are in fact functions of empirical dfs. Examples of such statistics are

(i) the sample mean $n^{-1}\sum_{i=1}^n \underline{x}_i = \int_{\mathbb{R}} x\,d\hat{\underline{F}}_n(x)$,

(ii) the α-trimmed mean

$$(n-2n_\alpha)^{-1} \sum_{i=n_\alpha+1}^{n-n_\alpha} \underline{x}_{i:n} = \int_{\mathbb{R}} xJ_n(\hat{\underline{F}}_n(x))\,d\hat{\underline{F}}_n(x),$$

where $0 < \alpha < \frac{1}{2}$, n_α is the largest integer $\le n\alpha$, $\underline{x}_{i:n}$ is the i-th order statistic of $\underline{x}_1,\ldots,\underline{x}_n$ and J_n is defined by $J_n = n(n-2n_\alpha)^{-1} \times$ $\times\, 1_{[(n_\alpha+1)/n,(n-n_\alpha)/n]}$,

(iii) signed rank statistics of the form

$$(3.6)\qquad \int_{(0,\infty)} J_n(\hat{\underline{H}}_n(x))\,d\hat{\underline{F}}_n(x),$$

where $\hat{\underline{H}}_n(x) = \hat{\underline{F}}_n(x) - \hat{\underline{F}}_n(-x-0)$ for $x \geq 0$ represents the empirical df of $|\underline{x}_1|,\ldots,|\underline{x}_n|$ and J_n is any score function defined on the interval $[0,1]$. In particular, $J_n(u) \equiv 1$ yields the sign test statistic, $J_n(u) \equiv u$ yields the Wilcoxon signed rank test statistic and $J_n(u) \equiv \Phi^{-1}(\frac{1}{2}+\frac{1}{2}nu/(n+1))$ leads to the test statistic of van der Waerden's test.

In the sequel statistics of type (iii) will be studied. To this end we introduce some further notation. The space of one-dimensional dfs endowed with the topology of uniform convergence induced by the supremum metric is denoted by $\mathcal{D}$. For $F,G \in \mathcal{D}$ let f and g represent the densities of F and G, respectively, with respect to a σ-finite measure ν. Then the Kullback-Leibler information number $K(G,F)$ of G with respect to F is defined by

$$K(G,F) = \int_{\mathbb{R}} g \log(g/f)d\nu$$

(cf. (2.10)). Furthermore for $A \subset \mathcal{D}$ let $K(A,F) = \inf\{K(G,F)\colon G \in A\}$. With this notation we have the following theorem.

THEOREM 3.3. *Let* T *be a real-valued continuous function on* $\mathcal{D}$ *and let* Ω_τ *be defined by*

$$\Omega_\tau = \{G \in \mathcal{D}\colon T(G) \geq \tau\} \qquad \textit{for each } \tau \in \mathbb{R}.$$

If the mapping $\tau \to K(\Omega_\tau,F)$ *is continuous from the right at* $\tau = t$, *then for any sequence* $\{u_n\}$ *of real numbers tending to zero*

$$\text{(3.7)} \qquad \lim_{n\to\infty} n^{-1} \log P(T(\hat{\underline{F}}_n)\geq t+u_n) = -K(\Omega_t,F).$$

This theorem has been proved under several additional conditions by HOADLEY [17] and also under these additional conditions, but by simpler means, by STONE [32]. A proof of the present theorem 3.3 will be given in another paper where it will also be shown that theorem 3.3 continues to hold for multi-dimensional dfs F.

In the particular case that $u_n = 0$ for all n (3.7) can also be formulated in the more transparent form

$$(3.8) \qquad \lim_{n\to\infty} n^{-1} \log \mathcal{P}(\hat{\underline{F}}_n \in \Omega) = -K(\Omega,F),$$

where $\Omega = \Omega_t$. The relation (3.8) for general Ω has a long history in the statistical literature. We mention only a few papers. In 1957 SANOV [29] proved (3.8) for the first time under very general conditions on Ω and F. However, his proof contains some obscure points and the conditions are complicated. BOROVKOV [8] proved (3.8) under a very neat set of conditions. He required F to be continuous and Ω to be an open set of dfs in $\mathcal{D}$ satisfying $K(\bar{\Omega},F) = K(\Omega,F)$, where $\bar{\Omega}$ is the closure of Ω in $\mathcal{D}$. A more precise analysis of (3.8) has been given by HOEFFDING [19] in the particular case that the $\underline{x}_i$'s assume only finitely many values.

The proofs of theorem 3.3 and of the results of Hoadley, Stone and Sanov are all based on the same idea which we briefly explain. For each $n \in \mathbb{N}$ a finite partition $\mathcal{R}_n = \{B_{n,1},\dots,B_{n,m(n)}\}$ of $\mathbb{R}$ is introduced. Let $\underline{k}_j$ be the number of random variables among $\underline{x}_1,\dots,\underline{x}_n$ taking a value in $B_{n,j}$ $(j=1,\dots,m(n))$. Then $(\underline{k}_1,\dots,\underline{k}_{m(n)})$ has a multinomial distribution with parameters n and $p_{n,1},\dots,p_{n,m(n)}$, where $p_{n,j} = P(\underline{x}_1 \in B_{n,j})$. Define $K_n(G,F)$ for $G \in \mathcal{D}$ by

$$K_n(G,F) = \sum_{j=1}^{m(n)} P_G(B_{n,j}) \log \{P_G(B_{n,j})/p_{n,j}\},$$

where $P_G(B)$ is the probability that a random variable with df G takes a value in B. Put $K_n(\Omega,F) = \inf\{K_n(G,F): G \in \Omega\}$. Since $\hat{\underline{F}}_n \in \Omega$ implies $K_n(\hat{\underline{F}}_n,F) \geq K_n(\Omega,F)$, one has

$$\mathcal{P}(\hat{\underline{F}}_n \in \Omega) \leq \mathcal{P}(K_n(\hat{\underline{F}}_n,F) \geq K_n(\Omega,F)) \leq$$

$$\leq \sum \{\mathcal{P}(\underline{k}_1=k_1,\dots,\underline{k}_{m(n)}=k_{m(n)}): \sum_{j=1}^{m(n)} (k_j/n) \log(k_j/np_{n,j}) \geq K_n(\Omega,F)\}.$$

Conversely,

$$(3.9) \qquad \mathcal{P}(\hat{\underline{F}}_n \in \Omega) \geq \mathcal{P}(\underline{k}_1=k_1,\dots,\underline{k}_{m(n)}=k_{m(n)})$$

for each sequence $(k_1,\dots,k_{m(n)})$ of integers satisfying

$$\{G \in \mathcal{D}\colon P_G(B_{n,j}) = k_j/n,\ j = 1,\dots,m(n)\} \subset \Omega.$$

Under suitable conditions on the partitions $\mathcal{R}_n$ and the set Ω one has $K_n(\Omega,F) \to K(\Omega,F)$ for $n \to \infty$. Moreover, under additional conditions, the integers $k_1,\dots,k_{m(n)}$ in (3.9) can be chosen such that $\Sigma(k_j/n)\log(k_j/np_{j,n})$ approximates $K(\Omega,F)$ for $n \to \infty$. The proofs are completed by an application of Stirling's formula to the multinomial coefficients. We omit the rather technical details.

Borovkov's approach is quite different. He relates the distribution of $\hat{\underline{F}}_n$ to a Poisson process and uses Fourier analysis of random walks to obtain (3.8). For continuous F theorem 3.3 is a consequence of Borovkov's result.

As previously noted sample means can be expressed as a function of empirical dfs. Although this function is not continuous, theorem 3.3 together with a truncation argument implies that

$$\lim_{n\to\infty} n^{-1} \log \mathcal{P}(n^{-1} \sum_{i=1}^{n} \underline{x}_i \geq t) = - K(\Omega_t^*,F),$$

where $\Omega_t^* = \{G \in \mathcal{D}\colon \int x dG(x) \geq t\}$. Since by a simple argument (cf. HOEFFDING [19])

$$- K(\Omega_t^*,F) = \inf_{\tau \geq 0} \{\log(Ee^{\tau \underline{x}_1}) - \tau t\},$$

Chernoff's theorem is again obtained. This illustrates the power of the empirical df technique.

To determine probabilities of large deviations for signed rank statistics theorem 3.3 is very useful. HO [16] proved the following theorem under an additional condition on the score functions.

<u>THEOREM 3.4</u> (Ho). *Let* F *be a continuous df symmetric about zero* (i.e. $F(x) = 1-F(-x)$). *Let* $\underline{t}_n$ *denote the signed rank statistic* (3.6), $n = 1,2,\dots$, *where the score functions* J_n *satisfy the condition*

$$(3.10) \qquad \lim_{n\to\infty} \sum_{i=1}^{n} \int_{(i-1)/n}^{i/n} |J_n(\tfrac{i}{n}) - J(u)|\,du = 0$$

for some limiting score function J. *If*

$$\tfrac{1}{2}\int_0^1 J(u)du < t < \int_0^1 J(u)1_{(0,\infty)}(J(u))du$$

and $\{u_n\}$ *is a sequence of real numbers tending to zero, then*

$$\lim_{n\to\infty} n^{-1} \log P(\underline{t}_n \geq t+u_n) = - I(t),$$

where

$$I(t) = \lambda t - \tfrac{1}{2}\lambda\int_0^1 J(u)du - \int_0^1 \log \cosh(\tfrac{1}{2}\lambda J(u))du$$

and $\lambda > 0$ *is the unique root of the equation*

$$\int_0^1 J(u)\ \{1 + \exp(-\lambda J(u))\}^{-1}du = t.$$

This theorem has been proved by Ho along the lines of the proof of a similar theorem for two-sample problems in WOODWORTH [33], but it can more easily be derived from theorem 3.3.

If the score functions J_n satisfy the conditions of theorem 3.4, if F is a continuous df (not necessarily symmetric) and $H(x) = F(x) - F(-x)$, $x \geq 0$, then

$$\int_{(0,\infty)} J_n(\underline{\hat{H}}_n(x))\ d\underline{\hat{F}}_n(x) \to \int_{(0,\infty)} J(H(x))dF(x) \tag{3.11}$$

in probability, for $n \to \infty$. This result together with theorem 3.4 will be applied in section 4 to determine Bahadur efficiencies of signed rank statistics.

Let $\Theta_0 = \{G \in \mathcal{D}$: G is continuous and symmetric about zero$\}$. The next theorem due to HO [16] yields the signed rank tests with best Bahadur slope with respect to a simple alternative.

<u>THEOREM 3.5</u> (Ho). *Let* $G \notin \Theta_0$ *be a df with density* g *with respect to Lebesgue measure. Define*

$$J_G(u) = \log \{g(H^{-1}(u))/g(-H^{-1}(u))\}, \qquad u \in (0,1),$$

where $H(x) = G(x) - G(-x)$, $x \geq 0$, and $H^{-1}(u) = \inf\{x: H(x) \geq u\}$. *Consider the sequence of signed rank statistics* $\{\underline{t}_n\}$ *with* $\underline{t}_n$ *defined by* (3.6). *If the score functions* J_n *satisfy condition* (3.10) *with* $J = J_G$, *then the sequence* $\{\underline{t}_n\}$ *attains the best possible Bahadur slope for testing* Θ_0 *against the simple alternative* G. *Moreover, this best Bahadur slope is equal to*

$$(3.12) \qquad 2K(G,\Theta_0) = 2\int_{\mathbb{R}} g(x)\log\frac{2g(x)}{g(x)+g(-x)}\,dx.$$

<u>SKETCH OF PROOF</u>. By (2.12) the Bahadur slopes of tests of Θ_0 against the simple alternative G are bounded above by $2K(G,\Theta_0)$. To prove the equality (3.12) let $F \in \Theta_0$ be arbitrary and let f_ν, g_ν and h_ν be densities of F, G and H with respect to a σ-finite measure ν dominating both F and G. The df $\overline{G} \in \Theta_0$ is defined by its density $\overline{g}_\nu(x) = \frac{1}{2}\{g_\nu(x) + g_\nu(-x)\}$ w.r.t.ν. Then

$$K(G,F) - K(G,\overline{G}) = \int_{\mathbb{R}} g_\nu \log(g_\nu/f_\nu)d\nu - \int_{\mathbb{R}} g_\nu \log(g_\nu/\overline{g}_\nu)d\nu$$

$$= \int_{\mathbb{R}} g_\nu \log(\overline{g}_\nu/f_\nu)d\nu$$

$$= \int_{(0,\infty)} h_\nu \log(\tfrac{1}{2}h_\nu/f_\nu)d\nu \geq 0$$

since $f_\nu(x) = f_\nu(-x)$ a.e.(ν) and Kullback-Leibler information numbers are nonnegative. Hence $K(G,\overline{G}) = K(G,\Theta_0)$ and $\overline{G}$ is the "least favorable" df for testing Θ_0 against G. Since G has a density g w.r.t. Lebesgue measure, (3.12) follows.

To complete the proof it must be shown that the maximal slope $2K(G,\Theta_0)$ is actually attained by the sequence $\{\underline{t}_n\}$. This follows by an application of theorem 3.4 in combination with (3.11). □

The property that the best possible slope is attained by signed rank statistics is called *asymptotic sufficiency* in the Bahadur sense of the vector of signs and ranks. The concept of asymptotic sufficiency in the Pitman sense is treated in HÁJEK & ŠIDÁK [14] Ch.7. HÁJEK [13] introduced asymptotic sufficiency in the Bahadur sense and obtained a result similar to theorem 3.5 for linear rank statistics in the two-sample problem. An analogous result for rank tests for independence has been proved by GROENEBOOM *et al.* [11].

4. APPLICATIONS: ONE-SAMPLE TESTS FOR LOCATION

Let $\underline{x}_1, \underline{x}_2, \underline{x}_3, \ldots$ be i.i.d. random variables with df $F(x-\theta)$, where θ is an unknown location parameter and F is an unknown continuous df symmetric about 0. The hypothesis H: $\theta = 0$ is to be tested against $\theta > 0$. Since F is unknown, the hypothesis set of dfs consists of all continuous dfs symmetric about 0. Thus the *hypothesis of symmetry* is to be tested against a *shift in location* to the right. Note that a simple alternative is determined by a pair (F,θ), $\theta > 0$.

We consider three well-known rank tests for this testing problem: the *sign test*, *Wilcoxon's signed rank test* and *van der Waerden's signed rank test*. The properly scaled test statistics of these tests - based on n observations - are

$$\underline{t}_n^{(1)} = n^{-\frac{1}{2}} \sum_{i=1}^{n} 1_{(0,\infty)}(\underline{x}_i),$$

$$\underline{t}_n^{(2)} = n^{-\frac{3}{2}} \sum_{i=1}^{n} \underline{r}_i^{+}\, 1_{(0,\infty)}(\underline{x}_i) \quad \text{and}$$

$$\underline{t}_n^{(3)} = n^{-\frac{1}{2}} \sum_{i=1}^{n} \Phi^{-1}(\tfrac{1}{2}+\tfrac{1}{2}\underline{r}_i^{+}/(n+1))\, 1_{(0,\infty)}(\underline{x}_i),$$

respectively, where $\underline{r}_i^{+}$ is the rank of $|\underline{x}_i|$ in the sample of absolute values $|\underline{x}_1|, \ldots, |\underline{x}_n|$.

If the df F is known to be normal, the Student t-test is the best test for the given testing problem. In this case the hypothesis reduces to a much smaller set of dfs. If the variance of F is also known, the hypothesis is simple and the Gauss test is UMP.

For normally distributed observations with a given variance van der Waerden's test is asymptotically most powerful (against a shift in location). This is a very strong property, since in this case the hypothesis is restricted to a simple hypothesis but van der Waerden's test is distributionfree and therefore a valid test of the entire hypothesis H. For a detailed exposition of the theory involved we refer to HÁJEK & ŠIDÁK [14]. Similarly, Wilcoxon's test and the sign test are asymptotically most powerful if the observations have logistic or double exponential distributions, respectively (recall that the logistic distribution has df $F(x) = (1+\exp(-x))^{-1}$ and the double exponential distribution density $f(x) = \frac{1}{2}\exp(-|x|)$).

These properties are reflected in their Pitman efficiencies. These efficiencies are complicated expressions depending heavily on F, but the following table is informative (cf. HÁJEK & ŠIDÁK [14] Ch.7).

TABLE 4.1. Pitman efficiencies of some distribution free tests

	normal F	logistic F	double exponential F
sign test w.r.t. Wilcoxon's test	$\frac{2}{3}$	$\frac{3}{4}$	$\frac{4}{3}$
sign test w.r.t. v.d. Waerden's test	$\frac{2}{\pi}$	$\frac{\pi}{4}$	$\frac{\pi}{2}$
Wilcoxon's test w.r.t. v.d. Waerden's test	$\frac{3}{\pi}$	$\frac{\pi}{3}$	$\frac{3\pi}{8}$

Moreover, for normal distributions the Pitman efficiencies of the sign test, Wilcoxon's test and van der Waerden's test with respect to the Student t-test (or the Gauss test) are $2/\pi$, $3/\pi$ and 1, respectively.

Roughly speaking, these results indicate that van der Waerden's test is satisfactory for thin-tailed distributions, that the sign test is adequate for heavy-tailed distributions and that Wilcoxon's test is a good choice in intermediate cases.

However, the preceding points of view take into account power at near alternatives only. Since the selection of a test is often motivated by these results, it is of some interest to compare asymptotic powers of these tests at fixed alternatives by way of their Bahadur efficiencies. In KLOTZ [23] these efficiencies were obtained for the first time. With this purpose in mind we proceed to determine the Bahadur slopes of the tests involved. In the sequel we consider a fixed alternative (F,θ) such that $\theta > 0$ and $F(\theta) < 1$.

Let $L_n^{(1)}(t) = P_0(n^{-\frac{1}{2}}\underline{t}_n^{(1)} \geq t)$ denote the tail probability of the sign test. By Chernoff's theorem (theorem 3.1) for $t \in (0,1)$

$$f_1(t) = -\lim_{n\to\infty} n^{-1} \log L_n^{(1)}(t) = -\inf_{\tau\geq 0} \{\log(\tfrac{1}{2}(1+e^{\tau})) - \tau t\} =$$

$$= t\log(2t) + (1-t)\log(2(1-t)).$$

Since f_1 is continuous on the interval $(0,1)$ and $n^{-\frac{1}{2}}\underline{t}_n^{(1)} \to F(\theta)$ in P_θ-probability for $n \to \infty$ by the weak law of large numbers and the symmetry of F, (2.6) and (2.7) are satisfied. Hence, in view of (2.8), the Bahadur slope of the sign test is equal to

$$2c_1(\theta) = 2F(\theta)\log(2F(\theta)) + 2(1-F(\theta))\log\{2(1-F(\theta))\}.$$

Next we turn to Wilcoxon's signed rank test. The test statistic $\underline{t}_n^{(2)}$ can be written in the form

$$\underline{t}_n^{(2)} = n^{-\frac{1}{2}} \sum_{i=1}^{n} J_n(\underline{r}_i^+/n)1_{(0,\infty)}(\underline{x}_i),$$

where $J_n(u) \equiv u$. Hence theorem 3.4 is applicable (with $J(u)\equiv u$) and we find for $t \in (\frac{1}{4},\frac{1}{2})$

$$f_2(t) = -\lim_{n\to\infty} n^{-1} \log P_0(n^{-\frac{1}{2}}\underline{t}_n^{(2)} \geq t) =$$

$$= \lambda(t-\tfrac{1}{4}) - \int_0^1 \log\cosh(\tfrac{1}{2}\lambda u)du,$$

where $\lambda > 0$ is the unique root of the integral equation

$$\int_0^1 u(1+e^{-\lambda u})^{-1} du = t.$$

This equation cannot be solved explicitly, but it may be used to simplify $f_2(t)$ to

$$f_2(t) = \lambda(2t-\tfrac{1}{2}) - \log\cosh(\tfrac{1}{2}\lambda).$$

Thus (2.7) is satisfied and it remains to verify (2.6). Putting

$$b_2(\theta) = \int_0^\infty \{F(x-\theta) - F(-x-\theta)\}dF(x-\theta),$$

application of (3.11) with $F(x)$ replaced by $F(x-\theta)$ yields that for $n \to \infty$

$$n^{-\frac{1}{2}} \underline{t}_n^{(2)} \to b_2(\theta) \text{ in } P_\theta\text{-probability.}$$

It follows that the Bahadur slope of Wilcoxon's test is

$$2c_2(\theta) = 2f_2(b_2(\theta)).$$

Although $b_2(\theta)$ can be evaluated in particular cases, explicit expressions for $f_2(b_2(\theta))$ are not available and its computation must be accomplished by numerical methods. We mention a few results for $b_2(\theta)$:

$$\begin{aligned} b_2(\theta) &= \tfrac{1}{2}\Phi(2^{\frac{1}{2}}\theta) && \text{for } F = \Phi \\ &\tfrac{1}{2} - ((\theta-\tfrac{1}{2})e^{2\theta}+\tfrac{1}{2})(e^{2\theta}-1)^{-2} && \text{for logistic } F \\ &\tfrac{1}{2} - \tfrac{1}{4}(\theta+1)e^{-2\theta} && \text{for double exponential } F. \end{aligned}$$

Now let us consider van der Waerden's test. Proceeding as in the previous case we find that the Bahadur slope of van der Waerden's test is equal to

$$2c_3(\theta) = 2f_3(b_3(\theta)),$$

where

$$f_3(t) = \lambda t - \lambda(2\pi)^{-\frac{1}{2}} - \int_0^1 \log\cosh(\tfrac{1}{2}\lambda\ \Phi^{-1}(\tfrac{1}{2}+\tfrac{1}{2}u))du,$$

$\lambda > 0$ is the unique root of the equation

$$\int_0^1 \Phi^{-1}(\tfrac{1}{2}+\tfrac{1}{2}u)\{1 + \exp(-\lambda\Phi^{-1}(\tfrac{1}{2}+\tfrac{1}{2}u))\}^{-1}du = t$$

and

$$b_3(\theta) = \int_0^\infty \Phi^{-1}(\tfrac{1}{2}+\tfrac{1}{2}F(x-\theta)-\tfrac{1}{2}F(-x-\theta))dF(x-\theta).$$

Simple expressions for $b_3(\theta)$ do not exist in this case.

In example 1.1 we have already seen that for normally distributed observations the Bahadur slope of the Student test is equal to $\log(1+\theta^2)$. However, the Student test has an unfair advantage over the distributionfree tests since its hypothesis set of distributions - in the normal case - is so much smaller. A better standard of comparison is furnished by the MP test of H against simple shift alternatives. In section 3 it was demonstrated (theorem 3.5) that for a simple alternative (F,θ), where F has a density f, the Bahadur slope of the MP test is equal to

$$2K(\theta,\Theta_0) = 2\int_{\mathbb{R}} f(x-\theta)\log\frac{2f(x-\theta)}{f(x-\theta)+f(x+\theta)}\,dx.$$

A little algebra shows that

$$K(\theta,\Theta_0) =$$

$$= \log 2 - \int_{\mathbb{R}} \phi(x-\theta)\log(1+e^{-2\theta x})dx \qquad \text{for } F=\Phi$$

$$\log 2 - \log(1+e^{2\theta}) + 2e^{\theta}(e^{2\theta}-1)^{-1}\{\theta e^{\theta}-\arctan(\tfrac{1}{2}(e^{\theta}-e^{-\theta}))\} \qquad \text{for logistic } F$$

$$\log 2 - \log(1+e^{-2\theta}) - e^{-\theta}\{\arctan e^{\theta}-\arctan e^{-\theta}\} \qquad \text{for double exponential } F.$$

In the figures 4.1, 4.2 and 4.3 the Bahadur efficiencies of the three rank tests with respect to the MP tests of H against simple alternatives (F,θ) are sketched for standard normal, logistic and double exponential F, respectively. The amount by which these efficiencies fall short of one describes the loss due to not precisely knowing the alternative. In figure 4.1 the Bahadur efficiency of the Student test with respect to the MP tests of H has also been shown, to facilitate a comparison of the first three tests to the Student test in the normal case.

Note that the efficiency of van der Waerden's test in figure 4.1 is very close to one; it is slightly smaller than the efficiency of Wilcoxon's test for θ greater than some constant between 1 and 1.1.

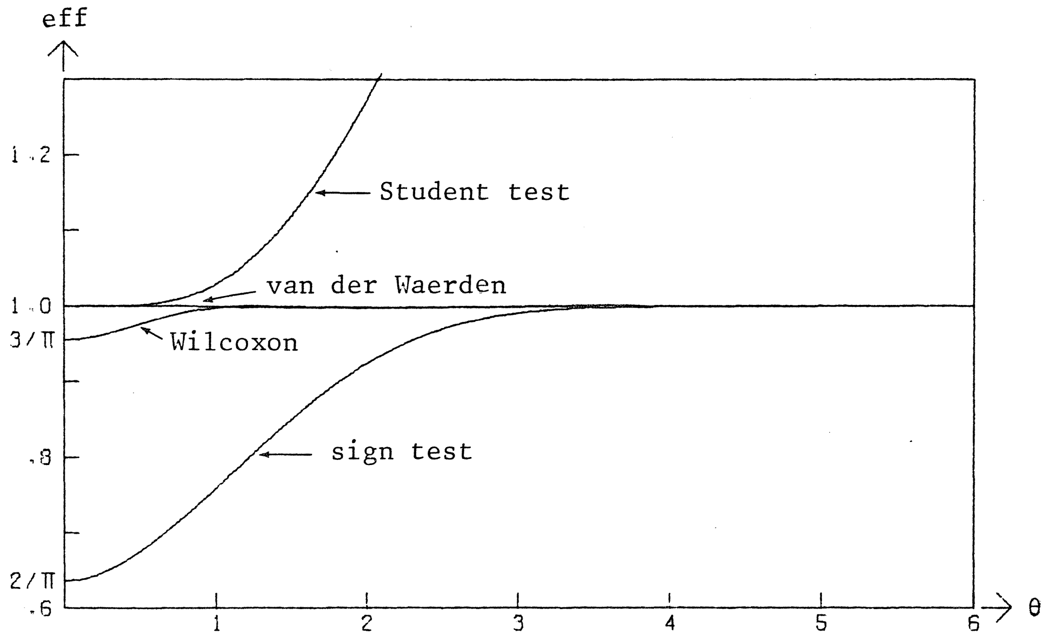

FIG 4.1 BAHADUR EFFICIENCIES FOR NORMAL ALTERNATIVES

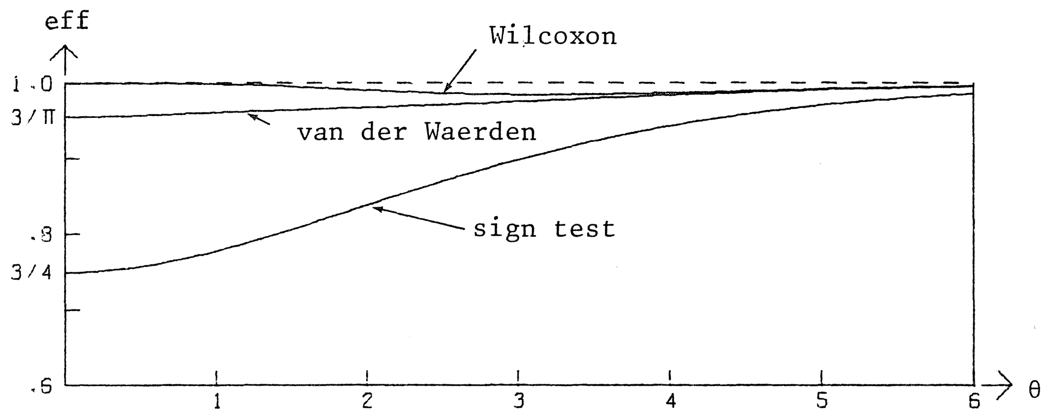

FIG 4.2 BAHADUR EFFICIENCIES FOR LOGISTIC ALTERNATIVES

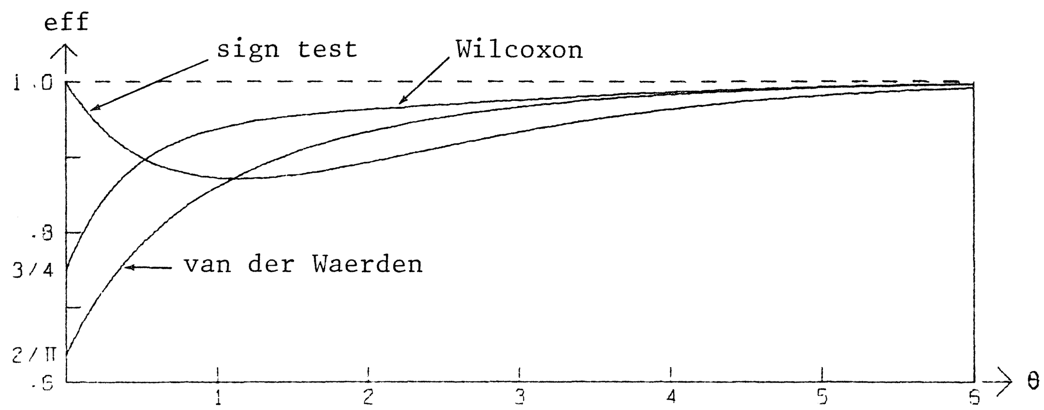

FIG 4.3 BAHADUR EFFICIENCIES FOR DOUBLE EXPONENTIAL ALTERNATIVES

Inspection of the figures shows that for fixed alternatives near the hypothesis the Bahadur efficiencies with respect to MP tests of H agree with the Pitman efficiencies. For large θ the efficiencies of the rank tests tend to one. This is explained by the fact that $c_i(\theta) \to \log 2$ for $\theta \to \infty$ $(i=1,2,3)$ for all symmetric F and $K(\theta,\theta_0) \to \log 2$ for $\theta \to \infty$ for all symmetric F with a density. The rather technical proofs are omitted. Yet it is remarkable that for alternatives far from the hypothesis the rank tests have such similar asymptotic power properties for all underlying dfs. From an overall point of view the Wilcoxon signed rank test does surprisingly well for the three distributions considered. A few small-sample results are reported by KLOTZ [22] for normal shift alternatives.

To conclude we mention some further results from the literature. Bahadur efficiencies of two-sample tests against shift alternatives have been studied in WOODWORTH [33]. It turns out that in the case of equal sample sizes the two-sample tests of van der Waerden and Wilcoxon and the median test have efficiencies identical to the corresponding one-sample tests. However, for unequal sample sizes the results are quite different. In the same paper the tests of Kendall and Spearman for independence are also compared. In HWANG & KLOTZ [20] two-sample rank tests against normal scale alternatives are considered. In BAHADUR [5] Kolmogorov-Smirnov tests and some related tests are studied. This monograph also contains an exhaustive list of references (up to 1970).

Acknowledgements

The authors are very grateful to R. van der Horst, who carried out the computational work resulting in figures 4.1 to 4.3. We also wish to thank the referee for his valuable comments.

REFERENCES

[1] ALBERS, W. (1975), *Efficiency and deficiency considerations in the symmetry problem*, Statistica Neerlandica 29, 81-92.

[2] BAHADUR, R.R. (1960), *Stochastic comparison of tests*, Ann. Math. Statist. 31, 276-295.

[3] BAHADUR, R.R. (1967), *An optimal property of the likelihood ratio statistic*, Proc. Fifth Berkeley Symp. Math. Stat. Prob. 1, 13-26.

[4] BAHADUR, R.R. (1967), *Rates of convergence of estimates and test statistics*, Ann. Math. Statist 38, 303-324.

[5] BAHADUR, R.R. (1971), *Some limit theorems in statistics*, SIAM, Philadelphia.

[6] BAHADUR, R.R. and M. RAGHAVACHARI (1972), *Some asymptotic properties of likelihood ratios on general sample spaces*, Proc. Sixth Berkeley Symp. Math. Stat. Prob. 1, 129-152.

[7] BAHADUR, R.R. and R. RANGA RAO (1960), *On deviations of the sample mean*, Ann. Math. Statist. 31, 1015-1027.

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

[9] BROWN, L.D. (1971), *Non-local asymptotic optimality of appropriate likelihood ratio tests*, Ann. Math. Statist. 42, 1206-1240.

[10] CHERNOFF, H. (1952), *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations*, Ann. Math. Statist. 23, 491-507.

[11] GROENEBOOM, P., Y. LEPAGE and F.H. RUYMGAART (1976), *Rank tests for independence with best strong exact Bahadur slope*. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 36, 119-127.

[12] GROENEBOOM, P., R. POTHARST, R. HELMERS and J. OOSTERHOFF (1976), *Efficiency begrippen in de statistiek*, MC Syllabus 30, Math. Centrum, Amsterdam.