Faculty of Mathematical Sciences

University of Twente

University for Technical and Social Sciences

P.O. Box 217
7500 AE Enschede
The Netherlands
Phone: +31-53-4893400
Fax: +31-53-4893114
Email: memo@math.utwente.nl

Memorandum No. 1467

Vanishing shortcoming and asymptotic relative efficiency

T. Inglot[1], W.C.M. Kallenberg and T. Ledwina[2]

October 1998

ISSN 0169-2690

[1]Wrocław University of Technology
[2]Polish Academy of Sciences

Vanishing shortcoming and asymptotic relative efficiency

Tadeusz Inglot*
Wrocław University of Technology

Wilbert C. M. Kallenberg
University of Twente

Teresa Ledwina
Polish Academy of Sciences

Abstract
The shortcoming of a test is the difference between the maximal attainable power and the power of the test under consideration. Vanishing shortcoming, when the number of observations tends to infinity, is therefore an optimality property of a test. Other familiar optimality criteria are based on the asymptotic relative efficiency of the test. The relations between these optimality criteria are investigated. It turns out that vanishing shortcoming is seemingly slightly stronger than first order efficiency, but in regular cases there is equivalence. The results are in particular applied on tests for goodness-of-fit.

Keywords and phrases: Shortcoming, Pitman efficiency, Bahadur efficiency, intermediate or Kallenberg efficiency, Cramér-von-Mises test, Anderson-Darling test.

1991 Mathematics Subject Classification: 62 F 05, 62 G 10, 62 G 20

* Supported by grant KBN 2 P03A 054 11,

1 Introduction

Comparison of tests is in principle based on the power of the tests. Exact powers of tests are hard to compute and, if they can be computed, it is in general not easy to compare them for large sets of alternatives simultaneously. Therefore, often an asymptotic approach is applied to simplify things. The simplifications by the asymptotics should be such that computation can be done and comparison can be made. Moreover, the conclusions based on the asymptotics should be in line with the finite sample behavior, which means that the asymptotics should provide good approximations.

A direct way of comparison of two tests is to consider the difference in power of the two tests. In particular, the difference between the most powerful test and a given test is of interest. It is called the *shortcoming* of that test. This concept is used to express *optimality* of a test: if the shortcoming of a test tends to 0 when the number of observations n tends to infinity, the test is asymptotically optimal.

Another, indirect, way of comparison of tests is based on the number of observations $N(\alpha, \beta, \theta)$ needed to get power β at the alternative θ when the level of the test equals α. If we have two tests with corresponding numbers $N_1(\alpha, \beta, \theta)$ and $N_2(\alpha, \beta, \theta)$, the ratio $N_2(\alpha, \beta, \theta)/N_1(\alpha, \beta, \theta)$ is called the *relative efficiency* of test 1 w.r.t test 2. If the relative efficiency equals r, test 2 needs r times as much observations to perform equally well as test 1 and hence test 1 is called r times as efficient as test 2. To investigate *optimality* we consider $N_*(\alpha, \beta, \theta)/N(\alpha, \beta, \theta)$, where $N_*(\alpha, \beta, \theta)$ corresponds to the most powerful test.

Again, an asymptotic approach is welcome to simplify the calculation and evaluation, as $N_i(\alpha, \beta, \theta)$ depends on three parameters. When sending n to infinity, two principles are (a) to "decrease the significance probability as n increases", i.e. to send α to 0, or (b) to "move the alternative hypothesis steadily closer to the null hypothesis", i.e. to send θ to the null hypothesis H_0. Both principles are attractive: with more observations it seems reasonable to have a stronger requirement on the level and on the other hand, for alternatives far away from H_0 there is no need for statistical methods, since those alternatives are obviously different from H_0. In *Pitman's* asymptotic efficiency concept, method (b) is used, while one deals with fixed levels, thus ignoring principle (a). In *Bahadur's* asymptotic efficiency concept, method (a) is actually used, while fixed alternatives are under consideration, thereby ignoring principle (b). *Intermediate or Kallenberg* efficiency, as defined in Kallenberg (1983a), applies *both* attractive principles simultaneously.

Optimality of a test can be expressed by first order efficiency, which means that $N_*(\alpha, \beta, \theta)/N(\alpha, \beta, \theta)$ converges to 1, where the limit is taken according to the efficiency concept involved.

It is the purpose of this paper to investigate the relations between the asymptotic optimality concepts based on vanishing shortcoming on the one hand and the three asymptotic efficiency concepts on the other hand. It turns out that asymptotic vanishing shortcoming is in the Pitman case equivalent to first order efficiency, while in the Bahadur and intermediate case vanishing shortcoming seems slightly stronger than first order efficiency. However, in regular cases first order efficient tests do also have asymptotic vanishing shortcoming. Here is a parallel with the phenomenon of "first order efficiency implies second order efficiency" (cf. Bickel, Chibisov and van Zwet (1981) and Kallenberg (1983a)).

The main results on the relationship between vanishing shortcoming and first order efficiency are *very general*: there is a very general set-up, very mild conditions on the most powerful tests and no condition at all on the (form of the) first order efficient tests. Moreover, the results hold simultaneously for all three efficiency concepts.

The paper is organized as follows. Section 2 contains notation, definitions and basic assumptions. The main results describing in great generality the relations between vanishing shortcoming and asymptotic relative efficiency are given in Section 3. These results are based on an asymptotic expression for $N_*(\alpha, \beta, \theta)$ and on an investigation of the number of extra observations needed to get a gain in asymptotic power for the most powerful test. These theorems (Theorem 3.2, 3.3, 3.2′ and 3.3′) and their extensions to general (regular) tests (Theorem 4.1 and 5.11) may be of independent interest. Some examples in Section 3 show the great generality of the results. Apart from an asymptotic expression for $N(\alpha, \beta, \theta)$, when the test is based on a (regular) test statistic, it is shown in Section 4 that as a rule vanishing shortcoming and first order efficiency are equivalent. Section 5 is devoted to further elaboration of the examples of Section 3. Applications are made to some first order efficient tests in these cases with special attention to tests for testing goodness-of-fit. In particular, some useful formulas for their asymptotic relative efficiencies are derived, showing that both for the Cramér-von-Mises test and for the Anderson-Darling test first order efficiency is only attained in one direction. In all other directions the asymptotic relative efficiency is less than one and often much lower. Moreover, these applications give a nice illustration of the phenomenon that equality of asymptotic optimal shift implies also equality of scale terms.

2 Notation, definitions and basic assumptions

Let $\mathcal{S}$ be a space of points s, $\mathcal{A}$ a σ-field of subsets of $\mathcal{S}$ and for each point θ in a set Θ let P_θ be a probability measure on $\mathcal{A}$. The random element S with values in $\mathcal{S}$ is distributed according to P_θ. In typical cases $S = (X_1, X_2, \dots)$ is a sequence of i.i.d. r.v.'s, but as yet $(\mathcal{S}, \mathcal{A}, P_\theta)$ is a quite general probability space. Note that θ is an abstract "parameter" and hence the results apply as well to parametric, nonparametric or semiparametric testing problems.

Suppose the hypothesis $H_0 : \theta \in \Theta_0$ has to be tested against $H_1 : \theta \in \Theta_1 \subset$

$\Theta - \Theta_0$. Let $\{\psi_{n,\alpha} : n \in \mathbb{N}, 0 < \alpha < 1\}$ be a family of (randomized) level-α tests of H_0, i.e., for each $n \in \mathbb{N}$ and $0 < \alpha < 1$ the function $\psi_{n;\alpha}$ is a measurable function on $\mathcal{S}$ with values in $[0, 1]$ satisfying

$$\sup\{E_{\theta_0}\psi_{n;\alpha}(S) : \theta_0 \in \Theta_0\} \leq \alpha.$$

Under very weak conditions a most powerful (MP) test of H_0 against the simple alternative $H_1^* : \theta = \theta_1$ exists (cf. Lehmann (1986) p. 576). In general, such a MP test depends on the particular alternative. The existence of such MP tests will be assumed in the sequel. They will be denoted by $\{\psi^*_{n;\alpha}\}$, suppressing their dependence on the particular alternative θ_1. In case of ambiguity we write $\psi^*_{n;\alpha}(S;\theta)$.

The power $E_\theta\psi_{n;\alpha}(S)$ of the level-α test $\psi_{n;\alpha}$ is denoted by $\beta_n(\alpha,\theta)$. When taking the supremum over all level-α tests, we get the envelope power function, denoted by $\beta^*_n(\alpha,\theta)$. In formula

$$\beta^*_n(\alpha,\theta) = \sup_\psi \beta_n(\alpha,\theta),$$

where $\psi = \{\psi_{n;\alpha}\}$ runs through all families of level-α tests of H_0. Obviously, $\beta^*_n(\alpha,\theta)$ is the power of $\psi^*_{n;\alpha}$.

The *shortcoming* of $\psi_{n;\alpha}$ is defined by

$$R_n(\alpha,\theta) = \beta^*_n(\alpha,\theta) - \beta_n(\alpha,\theta). \tag{2.1}$$

Asymptotic relative efficiency is defined in terms of $N(\alpha,\beta,\theta)$, which is the smallest number of observations N such that the level-α test $\psi_{m;\alpha}$ has power at least β at the alternative θ for all $m \geq N$. In formula

$$N(\alpha,\beta,\theta) = \inf\{N : \beta_m(\alpha,\theta) \geq \beta \text{ for all } m \geq N\}.$$

In case of the MP test of H_0 against the simple alternative θ we write $N_*(\alpha,\beta,\theta)$.

For given sequences $\{\alpha_n\}$ and $\{\theta_n\}$ with $0 < \alpha_n < 1$ and $\theta_n \in \Theta_1$ we want to relate the shortcoming with $N(\alpha_n,\beta,\theta_n) - N_*(\alpha_n,\beta,\theta_n)$, where $0 < \beta < 1$. If $\alpha_n \to \alpha \in (0,1)$ and θ_n converges to Θ_0, we speak of Pitman efficiency. If $\alpha_n \to 0$ and θ_n is fixed, we deal with Bahadur efficiency. The case $\alpha_n \to 0$ and θ_n converging to Θ_0 is called intermediate or Kallenberg efficiency. All throughout the paper $\{\theta_n\}$ and $\{\alpha_n\}$ are given sequences with $0 < \alpha_n < 1, \lim_{n\to\infty}\alpha_n = \bar{\alpha} \in [0,1)$ and $\theta_n \in \Theta_1$.

As n is the number of available observations, $\psi^*_{n;\alpha}$ can be considered for $m \geq n$ also as a test based on m observations for testing H_0, simply obtained by ignoring the $m-n$ extra observations. Since $\psi^*_{m;\alpha}$ is the MP test based on m observations, we have for all $0<\alpha<1$ and $\theta \in \Theta_1$

$$\beta^*_m(\alpha,\theta) = E_\theta \psi^*_{m;\alpha}(S;\theta) \geq \beta^*_n(\alpha,\theta) = E_\theta \psi^*_{n;\alpha}(S;\theta) \text{ if } m \geq n. \tag{2.2}$$

We assume that the MP test of H_0 against the simple alternative θ for each $n \in \mathbb{N}$ and $0<\alpha<1$ is based on a real valued test statistic T^*_n (which as a rule will depend on θ), rejecting for large values of T^*_n. More precisely, the level-α MP test of H_0 against θ is given by

(A1)

$$\psi^*_{n;\alpha}(s) = \begin{cases} 1 & \text{if } T^*_n(s) > c_n \\ \delta_n & \text{if } T^*_n(s) = c_n \\ 0 & \text{if } T^*_n(s) < c_n, \end{cases} \tag{2.3}$$

where $c_n = c_n(\alpha) = \inf\{c : \sup\{P_{\theta_0}(T^*_n(S) > c) : \theta_0 \in \Theta_0\} \leq \alpha\}$ and $\delta_n = \delta_n(\alpha) = \sup\{\delta \in [0,1] : \sup\{P_{\theta_0}(T^*_n(S) > c_n) + \delta P_{\theta_0}(T^*_n(S) = c_n) : \theta_0 \in \Theta_0\} \leq \alpha\}$. Then we have for all $c < c_n$

$$\sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T^*_n(S) > c_n) \leq \alpha < \sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T^*_n(S) > c). \tag{2.4}$$

The next assumption concerns the behavior of the test statistics T^*_n, given in (2.3)

(A2) There exist a continuous distribution function G_* on $\mathbb{R}$, strictly increasing on its support, which is $\mathbb{R}$ or $[a,\infty)$ with $a \in \mathbb{R}$, and a function $\mu_* : \Theta_1 \to (0,\infty)$ such that for every sequence $N = N(n)$ of natural numbers satisfying

$$\begin{aligned} &\sqrt{N}\mu_*(\theta_n) - G_*^{-1}(1-\alpha_n) = O(1) \text{ and} \\ &\lim_{n\to\infty} \inf \sqrt{N}\mu_*(\theta_n) > 0 \text{ in case } \bar{\alpha} > 0 \end{aligned} \tag{2.5}$$

it holds that (with T^*_N the MP test of H_0 against θ_n)

$$\lim_{n\to\infty} P_{\theta_n}(T^*_N - \sqrt{N}\mu_*(\theta_n) \leq x) = G_*(x) \text{ for every } x \in \mathbb{R}, \tag{2.6}$$

and, as $n \to \infty$,

$$\begin{aligned} &\sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T^*_N > t_n) = 1 - G_*(t_n + o(1)) \\ &\text{for all } t_n = \sqrt{N}\mu_*(\theta_n) + c \end{aligned} \tag{2.7}$$

with $c \in \mathbb{R}$, not depending on n. Moreover, $\limsup_{n\to\infty} \mu_*(\theta_n) < \infty$ and $\lim_{n\to\infty} \mu_*(\theta_n) = 0$ in case $\bar{\alpha} > 0$.

In typical cases T_N^* is a (standardized) sum of i.i.d. r.v.'s and G_* is the standard normal distribution function.

Condition (A2) is appropriate for the more local situations as Pitman efficiency and, partly, intermediate efficiency. In the nonlocal case the following condition is suitable.

(A2$'$) Let $\bar{\alpha} = 0$. There exist a continuous distribution function G_* on $\mathbb{R}$, strictly increasing on its support, which is $\mathbb{R}$ or $[a, \infty)$ with $a \in \mathbb{R}$, functions $\mu_{n*} : \Theta_1 \to (0, \infty)$ and functions r_{1n*} and r_{2n*}, defined on an open interval containing the limiting points of $\{\mu_{n*}(\theta_n)\}$ and satisfying for $i = 1, 2$

$$0 < b_1 \leq \frac{r_{in*}(x)}{x} \leq b_2 < \infty \tag{2.8}$$

for all x and some constants b_1, b_2, such that for every sequence $N = N(n)$ of natural numbers satisfying

$$\sqrt{N}\, r_{1n*}(\mu_{n*}(\theta_n)) - |2 \log \alpha_n|^{1/2} = O(1) \tag{2.9}$$

it holds that (with T_N^* the MP test of H_0 against θ_n)

$$\lim_{n\to\infty} P_{\theta_n} \left(T_N^* - \sqrt{N} \mu_{n*}(\theta_n) \leq x \right) = G_*(x) \text{ for every } x \in \mathbb{R}, \tag{2.10}$$

and, as $n \to \infty$,

$$\begin{aligned} -N^{-1} \log \left\{ \sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T_N^* > t_n) \right\} &= \tfrac{1}{2} r_{1n*}^2(\mu_{n*}(\theta_n)) + \\ d_n N^{-1/2} r_{2n*}(\mu_{n*}(\theta_n)) &+ o(N^{-1/2} \mu_{n*}(\theta_n)) \end{aligned} \tag{2.11}$$

for all $t_n = \sqrt{N} \mu_{n*}(\theta_n) + d_n$

with $\{d_n\} \subset \mathbb{R}$ being any bounded sequence. Moreover,

$$\limsup_{n\to\infty} \mu_{n*}(\theta_n) < \infty.$$

Remark 2.1 As a rule r_{2n*} is the derivative of $\frac{1}{2} r_{1n*}^2$. In typical cases

$$r_{in*}(\mu_{n*}(\theta_n))/\mu_{n*}(\theta_n) \to 1 \text{ if } \theta_n \to \Theta_0,$$

see also Example 3.6. This gives the connection between (A2) and (A2$'$) if G_* is the standard normal distribution function, cf. also (3.27).

Note that also in (A2$'$) G_* is often the standard normal distribution function, but in the Bahadur case the variance as a rule is not equal to 1. Checking of (A2) and (A2$'$) is exemplified in Examples 3.4 – 3.6. □

3 Main results

The idea behind the relationship between vanishing shortcoming and optimality in the sense of asymptotic relative efficiency is as follows. The asymptotic shift of the MP test statistic T_n^* for testing H_0 against the simple alternative θ_n is equal to $\sqrt{n}\mu_*(\theta_n)$. To obtain with k additional observations asymptotically a gain in power at θ_n requires $\liminf_{n\to\infty}(\sqrt{n+k} - \sqrt{n})\mu_*(\theta_n) > 0$. Therefore, vanishing shortcoming corresponds to an additional number k of observations satisfying $(\sqrt{n+k} - \sqrt{n})\mu_*(\theta_n) = o(1)$, or, equivalently, an additional number of observations of order $o(\sqrt{n}/\mu_*(\theta_n))$.

In the following theorem the latter result is indeed established under (A1) and (A2).

Theorem 3.1 *Assume* (A1) *and* (A2). *The following statements are equivalent*

(i) $\lim_{n\to\infty} R_m(\alpha_n, \theta_n) = 0$ *for each sequence* $m = m(n)$*, provided that in case* $\bar{\alpha} > 0$, $\liminf\limits_{n\to\infty} \beta_m^*(\alpha_n, \theta_n) > \bar{\alpha}$,

(ii) $N(\alpha_n, \beta, \theta_n) - N_*(\alpha_n, \beta, \theta_n) = o\left(\sqrt{N_*(\alpha_n, \beta, \theta_n)}/\mu_*(\theta_n)\right)$ *for each* $\beta \in (\bar{\alpha}, 1)$.

The related version of this theorem under (A1) and (A2′) is as follows.

Theorem 3.1′ *Assume* (A1) *and* (A2′). *The following statements are equivalent*

(i) $\lim_{n\to\infty} R_m(\alpha_n, \theta_n) = 0$ *for each sequence* $m = m(n)$.

(ii) $N(\alpha_n, \beta, \theta_n) - N_*(\alpha_n, \beta, \theta_n) = o\left(\sqrt{N_*(\alpha_n, \beta, \theta_n)}/\mu_{n*}(\theta_n)\right)$ *for each* $\beta \in (0, 1)$.

Before proving Theorem 3.1 and Theorem 3.1′ we give some comments. First order optimality in the sense of asymptotic relative efficiency means

$$N_*(\alpha_n, \beta, \theta_n)/N(\alpha_n, \beta, \theta_n) \to 1. \tag{3.1}$$

In the Pitman case, typically $\mu_*(\theta_n)$ is of exact order $\{N_*(\alpha_n, \beta, \theta_n)\}^{-1/2}$ and vanishing shortcoming corresponds to first order optimality in the sense of (3.1).

For Bahadur efficiency $\theta_n = \theta$ is fixed and vanishing shortcoming seems to be a stronger property than first order optimality in the sense of asymptotic relative efficiency, since $N(\alpha_n, \beta, \theta)$ should not only have the same first order term, but also the same $\sqrt{N}$-term as $N_*(\alpha_n, \beta, \theta)$. So, it seems that first order optimality in the sense of Bahadur efficiency, i.e. (3.1) with $\theta_n = \theta$, is not sufficient to guarantee vanishing shortcoming as (3.1) does not automatically imply (ii) in case $\theta_n = \theta$. However, in regular cases it turns out, by a simple argument, that nevertheless most tests which are first order efficient in the sense of Bahadur do also have vanishing shortcoming. The same argument applies to intermediate or Kallenberg efficiency. In Section 4 more details are given.

The proof of Theorem 3.1 and 3.1′ consists of a few steps. First we establish a formula for $N_*(\alpha_n, \beta, \theta_n)$.

Theorem 3.2 *Assume* (A1) *and* (A2). *For all* $\beta \in (\bar{\alpha}, 1)$ *we have*

$$N_*(\alpha_n, \beta, \theta_n) = \left\{ \frac{G_*^{-1}(1-\alpha_n) - G_*^{-1}(1-\beta) + o(1)}{\mu_*(\theta_n)} \right\}^2 . \tag{3.2}$$

Proof. For a (sufficiently small) positive ϵ set

$$\sqrt{N} = \left\{ G_*^{-1}(1-\alpha_n) - G_*^{-1}(1-\beta) - \epsilon \right\} / \mu_*(\theta_n). \tag{3.3}$$

Consider the test which rejects H_0 if

$$T_N^* > \sqrt{N}\mu_*(\theta_n) + G_*^{-1}(1-\beta) + \epsilon/2.$$

Denote the level of this test by $\alpha_n^{(1)}$. Then, by (2.7) and (3.3), we get for sufficiently large n

$$\begin{aligned} \alpha_n^{(1)} &= \sup_{\theta_0 \in \Theta_0} P_{\theta_0} \left(T_N^* > \sqrt{N}\mu_*(\theta_n) + G_*^{-1}(1-\beta) + \epsilon/2 \right) \\ &= 1 - G_* \left(G_*^{-1}(1-\alpha_n) - \epsilon/2 + o(1) \right) > \alpha_n \end{aligned}$$

and, by (2.6),

$$\begin{aligned} &\lim_{n\to\infty} \beta_N^*(\alpha_n^{(1)}, \theta_n) = 1 - \lim_{n\to\infty} P_{\theta_n} \left(T_N^* - \sqrt{N}\mu_*(\theta_n) \le G_*^{-1}(1-\beta) + \epsilon/2 \right) \\ &= 1 - G_*(G_*^{-1}(1-\beta) + \epsilon/2) < \beta. \end{aligned}$$

Hence, for sufficiently large n,

$$N_*(\alpha_n, \beta, \theta_n) \ge N_*(\alpha_n^{(1)}, \beta, \theta_n) > N. \tag{3.4}$$

By sending $\epsilon \to 0$ in (3.4), cf. also (3.3), we arrive at

$$\liminf_{n\to\infty} \sqrt{N^*(\alpha_n, \beta, \theta_n)}\mu_*(\theta_n) - G_*^{-1}(1-\alpha_n) \ge -G_*^{-1}(1-\beta). \tag{3.5}$$

Similarly, define for $\epsilon > 0$

$$\sqrt{N} = \left\{ G_*^{-1}(1-\alpha_n) - G_*^{-1}(1-\beta) + \epsilon \right\} / \mu_*(\theta_n) \tag{3.6}$$

and consider the test which rejects H_0 if

$$T_N^* > \sqrt{N}\mu_*(\theta_n) + G_*^{-1}(1-\beta) - \epsilon/2.$$

By (2.7) we get for sufficiently large n

$$\alpha_n^{(2)} = \sup_{\theta_0\in\Theta_0} P_{\theta_0}\left(T_N^* > \sqrt{N}\mu_*(\theta_n) + G_*^{-1}(1-\beta) - \epsilon/2\right) < \alpha_n$$

and, by (2.6),

$$\lim_{n\to\infty} \beta_N^*(\alpha_n^{(2)}, \theta_n) = 1 - G_*\left(G_*^{-1}(1-\beta) - \epsilon/2\right) > \beta. \tag{3.7}$$

In view of (2.2) and (3.7) we therefore have

$$\beta_m^*(\alpha_n^{(2)}, \theta_n) > \beta \text{ for all } m \geq N$$

and hence

$$N_*(\alpha_n, \beta, \theta_n) \leq N_*(\alpha_n^{(2)}, \beta, \theta_n) \leq N \tag{3.8}$$

for sufficiently large n. Sending $\epsilon \to 0$ in (3.8), cf. also (3.6), we end up with

$$\limsup_{n\to\infty} \sqrt{N_*(\alpha_n, \beta, \theta_n)}\mu_*(\theta_n) - G_*^{-1}(1-\alpha_n) \leq -G_*^{-1}(1-\beta). \tag{3.9}$$

Combination of (3.5) and (3.9) completes the proof. □

Under (A1) and (A2′) Theorem 3.2 takes the following form.

Theorem 3.2′ *Assume* (A1) *and* (A2′). *For all* $\beta \in (0,1)$

$$N_*(\alpha_n, \beta, \theta_n) = \left\{ \frac{|2\log\alpha_n|^{1/2} - \{r_{2n*}(\mu_{n*}(\theta_n))/r_{1n*}(\mu_{n*}(\theta_n))\}G_*^{-1}(1-\beta) + o(1)}{r_{1n*}(\mu_{n*}(\theta_n))} \right\}^2.$$

Proof. Write r_{in*} for short instead of $r_{in*}(\mu_{n*}(\theta_n))$. Note that by (2.8) r_{in*} and μ_{n*} are of the same exact order, implying e.g. that r_{2n*}/r_{1n*} is bounded away from 0 and ∞. For a (sufficiently small) positive ϵ set

$$\sqrt{N} = \left\{|2\log\alpha_n|^{1/2} - (r_{2n*}/r_{1n*})G_*^{-1}(1-\beta) - \epsilon\right\}/r_{1n*}. \tag{3.10}$$

Consider the test which rejects H_0 if

$$T_N^* > \sqrt{N}\mu_{n*}(\theta_n) + G_*^{-1}(1-\beta) + (\epsilon/2)(r_{1n*}/r_{2n*}).$$

Denote the level of this test by $\alpha_n^{(1)}$. Then, by (2.11) and (3.10), we get for sufficiently large n

$$\log \alpha_n^{(1)} = \log \left\{ \sup_{\theta_0 \in \Theta_0} P_{\theta_0} \left(T_N^* > \sqrt{N}\mu_{n*}(\theta_n) + G_*^{-1}(1-\beta) + (\epsilon/2)(r_{1n*}/r_{2n*}) \right) \right\}$$

$$= -\tfrac{1}{2} N r_{1n*}^2 - \{G_*^{-1}(1-\beta) + (\epsilon/2)(r_{1n*}/r_{2n*})\} N^{1/2} r_{2n*} + o(N^{1/2} r_{1n*})$$

$$= -\tfrac{1}{2} \left\{ |2\log \alpha_n|^{1/2} - (r_{2n*}/r_{1n*}) G_*^{-1}(1-\beta) - \epsilon \right\}^2$$

$$-(r_{2n*}/r_{1n*}) \{G_*^{-1}(1-\beta) + (\epsilon/2)(r_{1n*}/r_{2n*})\} |2\log\alpha_n|^{1/2} + o(|\log \alpha_n|^{1/2})$$

$$= \log \alpha_n + (\epsilon/2)|2\log\alpha_n|^{1/2} + o(|\log\alpha_n|^{1/2}) > \log \alpha_n$$

and, by (2.10), for sufficiently large n

$$\beta_N^*(\alpha_n^{(1)}, \theta_n) = 1 - P_{\theta_n} \left(T_N^* - \sqrt{N}\mu_{n*}(\theta_n) \le G_*^{-1}(1-\beta) + (\epsilon/2)(r_{1n*}/r_{2n*}) \right)$$

$$\le 1 - P_{\theta_n}(T_N^* - \sqrt{N}\mu_{n*}(\theta_n) \le G_*^{-1}(1-\beta) + (\epsilon/2)(b_1/b_2))$$

$$= 1 - G_*(G_*^{-1}(1-\beta) + (\epsilon/2)(b_1/b_2)) + o(1) < \beta.$$

Hence, for sufficiently large n,

$$N_*(\alpha_n, \beta, \theta_n) \ge N_*(\alpha_n^{(1)}, \beta, \theta_n) > N. \tag{3.11}$$

By sending $\epsilon \to 0$ in (3.11), cf. also (3.10), we arrive at

$$\liminf_{n\to\infty} \left\{ \sqrt{N^*(\alpha_n, \beta, \theta_n)}\, r_{1n*} - |2\log\alpha_n|^{1/2} + (r_{2n*}/r_{1n*}) G_*^{-1}(1-\beta) \right\} \ge 0. \tag{3.12}$$

Similarly, define for $\epsilon > 0$

$$\sqrt{N} = \left\{ |2\log\alpha_n|^{1/2} - (r_{2n*}/r_{1n*}) G_*^{-1}(1-\beta) + \epsilon \right\} / r_{1n*} \tag{3.13}$$

and consider the test which rejects H_0 if

$$T_N^* > \sqrt{N}\mu_{n*}(\theta_n) + G_*^{-1}(1-\beta) - (\epsilon/2)(r_{1n*}/r_{2n*}).$$

By (2.11) we get for sufficiently large n

$$\log \alpha_n^{(2)} = \log \left\{ \sup_{\theta_0 \in \Theta_0} P_{\theta_0} \left(T_N^* > \sqrt{N}\mu_{n*}(\theta_n) + G_*^{-1}(1-\beta) - (\epsilon/2)(r_{1n*}/r_{2n*}) \right) \right\}$$

$$< \log \alpha_n$$

and, by (2.10), for sufficiently large n

$$\beta_N^*(\alpha_n^{(2)},\theta_n) = 1 - P_{\theta_n}\left(T_N^* - \sqrt{N}\mu_{n*}(\theta_n) \le G_*^{-1}(1-\beta) - (\epsilon/2)(r_{1n*}/r_{2n*})\right)$$
$$\ge 1 - P_{\theta_n}\left(T_N^* - \sqrt{N}\mu_{n*}(\theta_n) \le G_*^{-1}(1-\beta) - (\epsilon/2)(b_1/b_2)\right)$$
$$= 1 - G_*\left(G_*^{-1}(1-\beta) - (\epsilon/2)(b_1/b_2)\right) + o(1) > \beta. \tag{3.14}$$

In view of (2.2) and (3.14) we therefore have

$$\beta_m^*(\alpha_n^{(2)},\theta_n) > \beta \text{ for all } m \ge N$$

and hence

$$N_*(\alpha_n,\beta,\theta_n) \le N_*(\alpha_n^{(2)},\beta,\theta_n) \le N \tag{3.15}$$

for sufficiently large n. Sending $\epsilon \to 0$ in (3.15), cf. also (3.13), we end up with

$$\limsup_{n\to\infty}\left\{\sqrt{N_*(\alpha_n,\beta,\theta_n)}r_{1n*} - |2\log\alpha_n|^{1/2} + (r_{2n*}/r_{1n*})G_*^{-1}(1-\beta)\right\} \le 0. \tag{3.16}$$

Combination of (3.12) and (3.16) completes the proof. □

In view of Remark 2.1 we get in typical cases the same expression (up to $o(1)$) for $\sqrt{N_*(\alpha_n,\beta,\theta_n)}\mu_*(\theta_n)$ in Theorem 3.2 as for $\sqrt{N_*(\alpha_n,\beta,\theta_n)}r_{1n*}(\mu_{n*}(\theta_n))$ in Theorem 3.2′ if G_* is the standard normal distribution function, cf. also (3.27).

The next theorem specifies the argument mentioned in the first lines of this section: "to obtain with k additional observations asymptotically a gain in power at θ_n requires $\lim_{n\to\infty}\left(\sqrt{n+k}-\sqrt{n}\right)\mu_*(\theta_n) > 0$".

Theorem 3.3 *Assume* (A1) *and* (A2). *For each* $\beta \in (\bar{\alpha},1)$ *and for each sequence* $m = m(n)$ *we have*

$$\lim_{n\to\infty}\Bigg[\beta_m^*(\alpha_n,\theta_n) - \left\{1 - G_*\left(G_*^{-1}(1-\beta) - \left[\sqrt{m} - \sqrt{N_*(\alpha_n,\beta,\theta_n)}\right]\mu_*(\theta_n)\right)\right\}\Bigg] = 0. \tag{3.17}$$

Proof. Consider first a sequence $m = m(n)$ of the form

$$\sqrt{m} = \sqrt{N_*(\alpha_n,\beta,\theta_n)} + \{b_n/\mu_*(\theta_n)\}, \tag{3.18}$$

where $b_n \to b, b \in \mathbb{R}$ with $b > G_*^{-1}(1-\beta) - G_*^{-1}(1-\bar{\alpha})$ if $\bar{\alpha} > 0$. In view of (3.2) and (3.18), m is of the form (2.5). By (2.4) and (2.7) the critical value c_m in (2.3) of the level-α_n MP test of H_0 against θ_n based on m observations satisfies

$$c_m = G_*^{-1}(1-\alpha_n) + o(1).$$

(Note that $G_*^{-1}(1-\alpha_n) + o(1)$ is of the form $\sqrt{m}\mu_*(\theta_n) + O(1)$ as required in (2.7).)

Since

$$P_{\theta_n}(T_m^* > c_m) \leq \beta_m^*(\alpha_n, \theta_n) \leq P_{\theta_n}(T_m^* \geq c_m)$$

and G_* is continuous, it follows by (2.6) and (3.2) that

$$\lim_{n\to\infty} \beta_m^*(\alpha_n, \theta_n) = 1 - G_*(G_*^{-1}(1-\beta) - b) \tag{3.19}$$
$$= 1 - \lim_{n\to\infty} G_*\left(G_*^{-1}(1-\beta) - \left[\sqrt{m} - \sqrt{N_*(\alpha_n, \beta, \theta_n)}\right]\mu_*(\theta_n)\right).$$

So, (3.17) holds for sequences of the form (3.18).

If $\sqrt{m} = \sqrt{N_*(\alpha_n, \beta, \theta_n)} + \{b_n/\mu_*(\theta_n)\}$ with $b_n \to \infty$, then for any b we have $b_n > b$ for sufficiently large n and hence $\sqrt{m} > \sqrt{N_*(\alpha_n, \beta, \theta_n)} + \{b/\mu_*(\theta_n)\}$. Consequently, by (2.2) and (3.19),

$$\beta_m^*(\alpha_n, \theta_n) \geq 1 - G_*(G_*^{-1}(1-\beta) - b)$$

for every $b \in \mathbb{R}$ and n sufficiently large. Thus $\lim_{n\to\infty} \beta_m^*(\alpha_n, \theta_n) = 1$. On the other hand,

$$\lim_{n\to\infty} G_*\left(G_*^{-1}(1-\beta) - \left[\sqrt{m} - \sqrt{N_*(\alpha_n, \beta, \theta_n)}\right]\mu_*(\theta_n)\right)$$
$$= \lim_{n\to\infty} G_*\left(G_*^{-1}(1-\beta) - b_n\right) = 0.$$

So, (3.17) holds true also in this case. We can proceed similarly in case $b_n \to -\infty$ getting (3.17) again. The general case follows now by a subsequence argument.□

A counterpart of Theorem 3.3 when (A1) and (A2′) hold is as follows.

Theorem 3.3′ *Assume* (A1) *and* (A2′). *Then we have for each* $\beta \in (0,1)$ *and for each sequence* $m = m(n)$

$$\lim_{n\to\infty}\Bigg[\beta_m^*(\alpha_n, \theta_n) -$$
$$\left\{1 - G_*\left(G_*^{-1}(1-\beta) - \left[\sqrt{m} - \sqrt{N_*(\alpha_n, \beta, \theta_n)}\right]\frac{r_{1n*}^2(\mu_{n*}(\theta_n))}{r_{2n*}(\mu_{n*}(\theta_n))}\right)\right\}\Bigg] = 0. \tag{3.20}$$

Proof. Write r_{in*} for short instead of $r_{in*}(\mu_{n*}(\theta_n))$ and N^* instead of $N^*(\alpha_n, \beta, \theta_n)$. Consider first a sequence $m = m(n)$ of the form

$$\sqrt{m} = \sqrt{N_*} + \{b_n r_{2n*}/r_{1n*}^2\}, \tag{3.21}$$

where $b_n \to b, b \in \mathbb{R}$. In view of Theorem 3.2′ and (3.21), m is of the form (2.9). Let $\tilde{c}_m = \sqrt{m}\mu_{n*}(\theta_n) - b + G_*^{-1}(1-\beta) + \epsilon$ for some $\epsilon \in \mathbb{R}$. By (2.11) we get

$$\begin{aligned}
&\log\left\{\sup_{\theta_0\in\Theta_0} P_{\theta_0}(T_m^* > \tilde{c}_m)\right\} \\
&= -\tfrac{1}{2}\left\{\sqrt{N_*}r_{1n*} + b_n r_{2n*}/r_{1n*}\right\}^2 - \left\{-b + G_*^{-1}(1-\beta) + \epsilon\right\}\sqrt{N_*}r_{2n*} \\
&\quad + o(|\log\alpha_n|^{1/2}) \\
&= -\tfrac{1}{2}\left\{|2\log\alpha_n|^{1/2} - (r_{2n*}/r_{1n*})G_*^{-1}(1-\beta)\right\}^2 \\
&\quad - \left\{G_*^{-1}(1-\beta) + \epsilon\right\}(r_{2n*}/r_{1n*})|2\log\alpha_n|^{1/2} + o(|\log\alpha_n|^{1/2}) \\
&= \log\alpha_n - \epsilon(r_{2n*}/r_{1n*})|2\log\alpha_n|^{1/2} + o(|\log\alpha_n|^{1/2}).
\end{aligned}$$

Therefore, considering $\epsilon > 0$ and $\epsilon < 0$ in $\tilde{c}_m$, it follows that the critical value c_m, of the level-α_n test based on T_m^* satisfies

$$c_m = \sqrt{m}\mu_{n*}(\theta_n) - b + G_*^{-1}(1-\beta) + o(1).$$

Since

$$P_{\theta_n}(T_m^* > c_m) \le \beta_m^*(\alpha_n, \theta_n) \le P_{\theta_n}(T_m^* \ge c_m)$$

and G_* is continuous, it follows by (2.10) and Theorem 3.2′ that

$$\begin{aligned}
\lim_{n\to\infty}\beta_m^*(\alpha_n,\theta_n) &= 1 - G_*(G_*^{-1}(1-\beta) - b) \\
&= 1 - \lim_{n\to\infty} G_*\left(G_*^{-1}(1-\beta) - \left(\sqrt{m} - \sqrt{N_*}\right)(r_{1n*}^2/r_{2n*})\right).
\end{aligned} \tag{3.22}$$

So, (3.20) holds for sequences of the form (3.21).

If $\sqrt{m} = \sqrt{N_*} + \{b_n(r_{2n*}/r_{1n*}^2)\}$ with $b_n \to \infty$, then for any b we have $b_n > b$ for sufficiently large n and hence $\sqrt{m} > \sqrt{N_*} + b(r_{2n*}/r_{1n*}^2)$. Consequently, by (2.2) and (3.22),

$$\beta_m^*(\alpha_n, \theta_n) \ge 1 - G_*(G_*^{-1}(1-\beta) - b)$$

for every $b \in \mathbb{R}$ and n sufficiently large. Thus $\lim_{n\to\infty} \beta_m^*(\alpha_n, \theta_n) = 1$. On the other hand,

$$\lim_{n\to\infty} G_* \left(G_*^{-1}(1-\beta) - \left(\sqrt{m} - \sqrt{N_*}\right) (r_{1n*}^2 / r_{2n*}) \right)$$
$$= \lim_{n\to\infty} G_* \left(G_*^{-1}(1-\beta) - b_n \right) = 0.$$

So, (3.20) holds true also in this case. We can proceed similarly in case $b_n \to -\infty$ getting (3.20) again. The general case follows now by a subsequence argument.□

Proof of Theorem 3.1 (i) $\Longrightarrow$ (ii)
Let $\bar{\alpha} < \beta < 1$ and $\epsilon > 0$. For each m satisfying

$$\sqrt{m} \geq \sqrt{N_*(\alpha_n, \beta, \theta_n)} + \epsilon / \mu_*(\theta_n)$$

we have, in view of (3.17),

$$\liminf_{n\to\infty} \beta_m(\alpha_n, \theta_n) = \liminf_{n\to\infty} \{\beta_m^*(\alpha_n, \theta_n) - R_m(\alpha_n, \theta_n)\} > \beta.$$

Hence

$$N(\alpha_n, \beta, \theta_n) \leq \left[\sqrt{N_*(\alpha_n, \beta, \theta_n)} + \epsilon / \mu_*(\theta_n) \right]^2 + 1$$

for sufficiently large n. Since $\epsilon > 0$ is arbitrarily chosen, (ii) now follows.
(ii) $\Longrightarrow$ (i)
Let $m = m(n)$ be some sequence. By a subsequence argument we may for proving $R_m(\alpha_n, \theta_n) \to 0$ assume w.l.o.g. that $\beta_m(\alpha_n, \theta_n) \to \beta \in [0, 1)$. (The case $\beta = 1$ can be excluded, since $\beta_m(\alpha_n, \theta_n) \to 1$ automatically implies $R_m(\alpha_n, \theta_n) \to 0$.) If $\beta_m^*(\alpha_n, \theta_n) \to \beta$, we get $R_m(\alpha_n, \theta_n) \to 0$ and hence it suffices to prove that the statement $\beta_m^*(\alpha_n, \theta_n) \to \beta + \epsilon$ for some $\epsilon > 0$ and $\beta + \epsilon > \bar{\alpha}$ leads to a contradiction.

So, assume $\beta_m(\alpha_n, \theta_n) \to \beta$ and $\beta_m^*(\alpha_n, \theta_n) \to \beta + \epsilon$ with $0 \leq \beta < 1$ and $\bar{\alpha} < \beta + \epsilon \leq 1$. For sufficiently large n we have for all β^* and $\tilde{\beta}$ such that $\beta < \beta^* < \tilde{\beta} < \beta + \epsilon$ and $\beta^* > \bar{\alpha}$

$$m \leq N(\alpha_n, \beta^*, \theta_n) \text{ and } m \geq N_*(\alpha_n, \tilde{\beta}, \theta_n).$$

Hence, we get

$$N_*(\alpha_n, \tilde{\beta}, \theta_n) \leq N(\alpha_n, \beta^*, \theta_n)$$

and therefore, by (ii),

$$N_*(\alpha_n, \tilde{\beta}, \theta_n) - N_*(\alpha_n, \beta^*, \theta_n) \leq$$
$$N(\alpha_n, \beta^*, \theta_n) - N_*(\alpha_n, \beta^*, \theta_n) = o\left(\sqrt{N_*(\alpha_n, \beta^*, \theta_n)} / \mu_*(\theta_n) \right).$$

Application of (3.17) with $m = N_*(\alpha_n, \tilde{\beta}, \theta_n)$ and β replaced by β^* gives a contradiction. □

Proof of Theorem 3.1′ (i) $\Longrightarrow$ (ii)
Let $0 < \beta < 1$ and $\epsilon > 0$. For each m satisfying

$$\sqrt{m} \geq \sqrt{N_*(\alpha_n, \beta, \theta_n)} + \epsilon/\mu_{n*}(\theta_n)$$

we have, in view of Theorem 3.3′ and noting that μ_{n*}, $r_{1n*}(\mu_{n*}(\theta_n))$ and $r_{2n*}(\mu_{n*}(\theta_n))$ are all of the same exact order,

$$\liminf_{n\to\infty} \beta_m(\alpha_n, \theta_n) = \liminf_{n\to\infty} \{\beta_m^*(\alpha_n, \theta_n) - R_m(\alpha_n, \theta_n)\} > \beta.$$

Hence

$$N(\alpha_n, \beta, \theta_n) \leq \left[\sqrt{N_*(\alpha_n, \beta, \theta_n)} + \epsilon/\mu_{n*}(\theta_n)\right]^2 + 1$$

for sufficiently large n. Since $\epsilon > 0$ is arbitrarily chosen, (ii) now follows.
(ii) $\Longrightarrow$ (i)
Let $m = m(n)$ be some sequence. By a subsequence argument we may for proving $R_m(\alpha_n, \theta_n) \to 0$ assume w.l.o.g. that $\beta_m(\alpha_n, \theta_n) \to \beta \in [0, 1)$. (The case $\beta = 1$ can be excluded, since $\beta_m(\alpha_n, \theta_n) \to 1$ automatically implies $R_m(\alpha_n, \theta_n) \to 0$.) If $\beta_m^*(\alpha_n, \theta_n) \to \beta$, we get $R_m(\alpha_n, \theta_n) \to 0$ and hence it suffices to prove that the statement $\beta_m^*(\alpha_n, \theta_n) \to \beta + \epsilon$ for some $\epsilon > 0$, leads to a contradiction.

So, assume $\beta_m(\alpha_n, \theta_n) \to \beta$ and $\beta_m^*(\alpha_n, \theta_n) \to \beta + \epsilon$ with $0 \leq \beta < 1$ and $\beta + \epsilon \leq 1$. For sufficiently large n we have for all β^* and $\tilde{\beta}$ such that $\beta < \beta^* < \tilde{\beta} < \beta + \epsilon$

$$m \leq N(\alpha_n, \beta^*, \theta_n) \text{ and } m \geq N_*(\alpha_n, \tilde{\beta}, \theta_n).$$

Hence, we get

$$N_*(\alpha_n, \tilde{\beta}, \theta_n) \leq N(\alpha_n, \beta^*, \theta_n)$$

and therefore, by (ii),

$$N_*(\alpha_n, \tilde{\beta}, \theta_n) - N_*(\alpha_n, \beta^*, \theta_n) \leq$$
$$N(\alpha_n, \beta^*, \theta_n) - N_*(\alpha_n, \beta^*, \theta_n) = o\left(\sqrt{N_*(\alpha_n, \beta^*, \theta_n)}/\mu_{n*}(\theta_n)\right).$$

Application of Theorem 3.3′ with $m = N_*(\alpha_n, \tilde{\beta}, \theta_n)$ and β replaced by β^* gives a contradiction. □

As is seen from the definitions and basic assumptions in Section 2, there are no conditions or assumptions on the tests ψ. This means that Theorems 3.1 and 3.1′ are in fact properties of MP tests.

We end this section by presenting some testing problems, where assumptions (A1) and (A2) hold and hence Theorem 3.1 can be applied. In Example 3.6 we also show that (A1) and (A2′) hold and apply Theorem 3.1′ in the given context.

Example 3.4 *Gauss-test.* Let $X_1, \dots, X_n$ be i.i.d. r.v.'s each with a normal $N(\theta, 1)$-distribution. Consider the testing problem

$$H_0 : \theta = 0 \text{ against } H_1 : \theta > 0.$$

It is easily seen that (A1) and (A2) hold with $T_N^* = \overline{X}\sqrt{N}$, where $\overline{X} = N^{-1}\sum_{i=1}^N X_i, G_* = \Phi$, the standard normal distribution function and $\mu_*(\theta) = \theta$. Theorem 3.1 can be applied for all sequences $\{\theta_n\}$ and $\{\alpha_n\}$ with $\theta_n > 0$ and $\lim_{n\to\infty} \alpha_n = \bar{\alpha} \in [0, 1)$, provided of course that $\theta_n \to 0$ if $\bar{\alpha} > 0$. □

Example 3.5 *Curved exponential families.* Let $X_1, \dots, X_n$ be i.i.d. k-dimensional r.v.'s with density

$$p_\theta(x) = \exp\{\gamma_\theta' x - \psi(\gamma_\theta)\},$$

$\theta \in \Theta$, with respect to a σ-finite measure ν on $\mathbb{R}^k$. Here $\gamma_\theta' x$ denotes the inner product of γ_θ and x, while $\psi(\gamma_\theta) = \log \int \exp(\gamma_\theta' x) d\nu(x)$. Set $\Gamma = \{\gamma \in \mathbb{R}^k : \psi(\gamma) < \infty\}$, where $\psi(\gamma) = \log \int \exp(\gamma' x) d\nu(x)$. Assume int $\Gamma \neq \phi$, Θ is an open interval in $\mathbb{R}^1$, while γ_θ is a differentiable bijection from Θ on $\gamma(\Theta) \subset \Gamma$. Note that $p_\theta(x)$ is a curved exponential family in the terminology of Efron (1975).

Consider the testing problem $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$, where $\theta_0 \in \Theta$ is given. Additionally to the above assume that $\gamma_{\theta_0} \in \operatorname{int}\Gamma$ and that the covariance matrix of X_1 under p_{θ_0} is nonsingular.

Let $\{\alpha_n\}$ be a sequence of levels satisfying

$$\Phi^{-1}(1 - \alpha_n) = o(n^{1/4}) \text{ as } n \to \infty, \tag{3.23}$$

or, equivalently,

$$|\log \alpha_n| = o(n^{1/2}) \text{ as } n \to \infty. \tag{3.24}$$

Let $\{\theta_n\}$ be a sequence of alternatives satisfying

$$\lim_{n\to\infty} (\theta_n - \theta_0)^2 |\log \alpha_n|^{1/2} = 0. \tag{3.25}$$

Condition (A1) holds with

$$T_n^* = n^{-1/2} \sum_{i=1}^n (\gamma_{\theta_n} - \gamma_{\theta_0})'(X_i - E_{\theta_0} X)/\{\mathrm{var}_{\theta_0}(\gamma_{\theta_n} - \gamma_{\theta_0})' X\}^{1/2}.$$

To check (A2) set

$$\mu_*(\theta) = (\gamma_\theta - \gamma_{\theta_0})'(E_\theta X - E_{\theta_0} X)/\{\mathrm{var}_{\theta_0}(\gamma_\theta - \gamma_{\theta_0})' X\}^{1/2}.$$

We have

$$
\begin{aligned}
&(\gamma_\theta - \gamma_{\theta_0})'(E_\theta X - E_{\theta_0} X) \\
&= \int [(\Delta\gamma)'x - \Delta\psi]\,[\exp\{(\Delta\gamma)'x - \Delta\psi\} - 1]\, p_{\theta_0}(x) d\nu(x),
\end{aligned}
$$

where for short we write $\Delta\gamma = \gamma_\theta - \gamma_{\theta_0}$ and $\Delta\psi = \psi(\gamma_\theta) - \psi(\gamma_{\theta_0})$. Since $u\,\{\exp(u) - 1\} \geq 0, u \in \mathbb{R}^1$, we get $\mu_*(\theta) \geq 0$ for any θ. Due to the differentiability of the bijection γ_θ we have that $(\gamma_{\theta_n} - \gamma_{\theta_0})/\{\mathrm{var}_{\theta_0}(\gamma_{\theta_n} - \gamma_{\theta_0})'X\}^{1/2} \to \dot\gamma_{\theta_0}/s$, where $\cdot$ denotes the derivative w.r.t. θ, while $s = \{\mathrm{var}_{\theta_0} \dot\gamma'_{\theta_0} X\}^{1/2}$. Observe now that the above and the assumption $\gamma_{\theta_0} \in \operatorname{int}\Gamma$ imply that there exists $\delta > 0$ such that $\gamma_{\theta_0} + t(\gamma_{\theta_n} - \gamma_{\theta_0})/\{\mathrm{var}_{\theta_0}(\gamma_{\theta_n} - \gamma_{\theta_0})'X\}^{1/2} \in \operatorname{int}\Gamma$ and $\gamma_{\theta_n} + t(\gamma_{\theta_n} - \gamma_{\theta_0})/ \{\mathrm{var}_{\theta_0}(\gamma_{\theta_n} - \gamma_{\theta_0})'X\}^{1/2} \in \operatorname{int}\Gamma$ for all $|t| < \delta$ and n sufficiently large.

Since the mapping $\psi(\gamma)$ is continuous on $\operatorname{int}\Gamma$, the preceding implies that both $E_{\theta_n} \exp\{t(\gamma_{\theta_n} - \gamma_{\theta_0})'X\}$ and $E_{\theta_0} \exp\{t(\gamma_{\theta_0} - \gamma_{\theta_0})'X\}$ are uniformly bounded for t in a neighborhood of 0 and n sufficiently large. This ensures that also moments of $(\gamma_{\theta_n} - \gamma_{\theta_0})'X$ are uniformly bounded under p_{θ_0} and p_{θ_n}.

So, under the above assumptions, Liapounov's theorem easily yields

$$
\lim_{n\to\infty} P_{\theta_n}\left(T^*_M - \sqrt{M}\mu_*(\theta_n) \leq x\right) = \Phi(x),\ x \in \mathbb{R},
$$

for an arbitrary sequence $M = M(n)$ tending to infinity and $\theta_n \to \theta_0$. This, in particular, implies that (2.6) of (A2) holds with $G_* = \Phi$.

Exploiting again the continuity of $\psi(\gamma)$ on $\operatorname{int}\Gamma$, in a similar way as done in the proof of Theorem 5.8 in Inglot and Ledwina (1996), we can apply the Cramér-type large deviation result for triangular arrays obtained by Book (1976) (cf. Lemma 4.1 in Jurečková et al. (1988)). This yields

$$
P_{\theta_0}(T^*_N > t_n) = \exp\{-\tfrac{1}{2}t_n^2 + o(t_n)\} \tag{3.26}
$$

if $t_n \to \infty$ and $t_n = \sqrt{N}\mu_*(\theta_n) + c$ with $c \in \mathbb{R}$ and N of the form (2.5). (Note that for getting $o(t_n)$ in (3.26) we need $t_n = o(n^{1/4})$ and this is one of the reasons to require (3.23).) Combination of

$$
\Phi^{-1}(u) = \sqrt{-2\log(1-u)} + o(1),\ \text{as } u \to 1, \tag{3.27}
$$

with (3.26) gives (2.7) if $t_n \to \infty$. For bounded t_n (2.7) follows from the central limit theorem. Since (3.25) implies $\theta_n \to \theta_0$, we have always $\lim_{n\to\infty} \mu_*(\theta_n) = 0$. Therefore, (A2) holds for this testing problem and Theorem 3.1 can be applied for all sequences $\{\theta_n\}$ and $\{\alpha_n\}$ satisfying (3.23) (or (3.24)) and (3.25). □

Example 3.6 *Goodness-of-fit.* Let $X_1, \ldots, X_n$ be i.i.d. r.v.'s with values in $[0, 1]$. The null hypothesis states that the X_i have a uniform distribution on

$[0,1]$. Note that by an application of the integral transformation this is no loss of generality in the goodness-of-fit testing problem for a simple null hypothesis.

For simplicity of presentation we restrict attention to alternatives defined by densities (with respect to Lebesgue measure on $[0,1]$)

$$p_n(x) = 1 + \theta_n a(x),\ \theta_n > 0, \tag{3.28}$$

with $\theta_n \to 0 = \theta_0$ as $n \to \infty$ and where a is bounded, $\int_0^1 a(x)dx = 0$ and $\int_0^1 a^2(x)dx = 1$. The MP test of $H_0 : p(x) = 1$ against the simple alternative p_n, given by (3.28), satisfies (A1) with

$$T_n^* = (\sqrt{n}\sigma_{0n})^{-1} \sum_{i=1}^{n} \{\log p_n(X_i) - e_{0n}\},$$

where

$$e_{0n} = E_{\theta_0} \log p_n(X),\ \sigma_{0n}^2 = \mathrm{var}_{\theta_0} \log p_n(X).$$

To check (A2), define $G_* = \Phi$ and $\mu_*(\theta) = \theta$. Consider sequences of levels $\{\alpha_n\}$ and alternatives $\{\theta_n\}$ such that

$$\lim_{n\to\infty} \alpha_n = 0,\ \lim_{n\to\infty} \theta_n = 0 \text{ and } \lim_{n\to\infty} \theta_n \Phi^{-1}(1-\alpha_n) = 0. \tag{3.29}$$

According to (2.5) take

$$N = \{\Phi^{-1}(1-\alpha_n) + O(1)\}^2/\theta_n^2.$$

By Lemma 5.4 and Proposition 6.6 in Inglot and Ledwina (1996) we get (2.6). Indeed, due to (3.29) we get $\sqrt{N}(\theta_n - b^{(1)}(p_n)) \to 0$, where $b^{(1)}$ is given by (5.18) of Inglot and ledwina (1996). In view of Proposition 5.12 and Lemma 5.4 in Inglot and Ledwina (1996) we can apply Book's (1976) result as in the previous example. The assumption $\theta_n \Phi^{-1}(1-\alpha_n) \to 0$ (see (3.29)) yields $\sqrt{N}\theta_n = o(N^{1/4})$ and hence Book's result gives (cf. also (3.26))

$$P_{\theta_0}(T_N^* > t_n) = \exp\{-\tfrac{1}{2}\, t_n^2 + o(t_n)\} = 1 - \Phi(t_n + o(1))$$

for all $t_n = \sqrt{N}\mu_*(\theta_n) + c$ with $c \in \mathbb{R}$. Therefore, (A2) holds and Theorem 3.1 can be applied for all sequences $\{\theta_n\}$ and $\{\alpha_n\}$ satisfying (3.29).

Next it will be shown that (A2′) holds for all sequences of levels $\{\alpha_n\}$ and alternatives $\{\theta_n\}$ such that

$$\lim_{n\to\infty} \alpha_n = 0 \text{ and } \lim_{n\to\infty} \theta_n = \theta_1 \in \mathbb{R}.$$

We assume that $\inf\{1 + \theta_1 a(x) : x \in [0,1]\} > 0$. Note that θ_1 can be 0 (intermediate case) or unequal to 0 (Bahadur case). So, in the intermediate case this is no restriction at all and in the Bahadur case this is only slightly more than stating that the alternative is well-defined. To avoid writing every time "for sufficiently large n" we further assume w.l.o.g. that $\inf\{1+\theta_n a(x) : x \in [0,1]\} \geq \delta > 0$ for all n. Define

$$\mu_{n*}(\theta) = \sigma_{0n}^{-1}\{E_\theta \log p_n(X) - e_{on}\}.$$

We have $\mu_{n*}(\theta) \geq 0$ for all $\theta \geq 0$. Next set

$$G_*(x) = \Phi(x/\sigma_{\theta_1}) \text{ with}$$
$$\sigma_\theta^2 = \begin{cases} \dfrac{\mathrm{var}_\theta\{\log(1+\theta a(X))\}}{\mathrm{var}_{\theta_0}\{\log(1+\theta a(X))\}} & \text{if } \theta \neq \theta_0 = 0 \\ 1 & \text{if } \theta = \theta_0 = 0. \end{cases}$$

Using Liapounov's theorem (as in the proof of Proposition 6.6 in Inglot and Ledwina (1996)) we have for any sequence $N = N(n) \to \infty$ that

$$\lim_{n\to\infty} P_{\theta_n}(T_N^* - \sqrt{N}\mu_{n*}(\theta_n) \leq x) = G_*(x) \text{ for every } x \in \mathbb{R},$$

which proves (2.10) of (A2′).

Write

$$Y_{ni} = \sigma_{0n}^{-1}\{\log p_n(X_i) - e_{0n}\}$$

and denote by $P_{n\theta_0}$ the distribution of Y_{ni} when X_i has the uniform distribution. Define the exponential family associated with $P_{n\theta_0}$ by

$$dQ_{nu}(y) = \exp(uy - \psi_n(u))dP_{n\theta_0}(y),$$

when $\psi_n(u)$ is the normalizing factor, given by

$$\psi_n(u) = \log\{E_{\theta_0}\exp(uY_{ni})\}.$$

Standard exponential family theory yields

$$\psi_n'(u) = E_{nu}Y_{ni} \text{ and } \psi_n''(u) = \mathrm{var}_{nu}Y_{ni},$$

where E_{nu} and var_{nu} denote the expectation and variance under Q_{nu}. We write

$$\lambda_n(u) = \psi_n'(u).$$

It is immediately seen that $\lambda_n(0) = E_{\theta_0}Y_{ni} = 0$ and that λ_n is increasing. The so called Kullback-Leibler information number of Q_{nu} w.r.t. $P_{n\theta_0}$ is given by

$$E_{nu}\log\left(dQ_{nu}/dP_{n\theta_0}\right) = u\lambda_n(u) - \psi_n(u).$$

The functions r_{1n*} and r_{2n*} are defined by

$$\tfrac{1}{2}r_{1n*}^2(v) = \lambda_n^{-1}(v)v - \psi_n(\lambda_n^{-1}(v))$$

$$r_{2n*}(v) = \lambda_n^{-1}(v).$$

It is seen that $\frac{1}{2}r_{1n*}^2$ is the Kullback-Leibler information number of $Q_{n\lambda_n^{-1}(v)}$ w.r.t. $P_{n\theta_0}$. Further, it is easily checked that r_{2n*} is the derivative of $\frac{1}{2}r_{1n*}^2$. As $\lambda_n(0) = 0$ and $\lambda_n'(0) = \mathrm{var}_{\theta_0} Y_{ni} = 1$, we have

$$\lambda_n(z_n) = \lambda_n(0) + z_n\lambda_n'(0) + \tfrac{1}{2}z_n^2\lambda_n''(\xi_n)$$

for some $0 < \xi_n < z_n$. Since Y_{ni} is uniformly bounded, $\lambda_n''(\xi_n)$ converges to $\lim_{n\to\infty} E_{\theta_0} Y_{ni}^3$ as $z_n \to 0$ and hence is uniformly bounded too. Therefore, $v_n \to 0$ implies

$$\lim_{n\to\infty} r_{1n*}^2(v_n)/v_n^2 = 1 \text{ and } \lim_{n\to\infty} r_{2n*}(v_n)/v_n = 1.$$

Hence, (2.8) now easily follows.

Let N be of the form (2.9) and define u_n by

$$\lambda_n(u_n) = \mu_{n*}(\theta_n) + d_n/\sqrt{N}$$

with $\{d_n\} \subset \mathbb{R}$ a bounded sequence. Let $\overline{Y}_n = N^{-1}\sum_{i=1}^N Y_{ni}$ and when Y_{ni} has distribution Q_{nu}, denote the distribution of $\overline{Y}_n$ by $\overline{Q}_{nu}$. Then we have

$$d\overline{Q}_{nu}(y) = \exp\{N[uy - \psi_n(u)]\}\, d\overline{P}_{n\theta_0}(y),$$

where $\overline{P}_{n\theta_0} = \overline{Q}_{n0}$ is the distribution of $\overline{Y}_n$ when Y_{ni} has distribution $P_{n\theta_0}$. Hence, we get

$$\begin{aligned}
&P_{\theta_0}\left(T_N^* > \sqrt{N}\mu_{n*}(\theta_n) + d_n\right) = P_{\theta_0}(\overline{Y}_n > \lambda_n(u_n)) \\
&= \int_{(\lambda_n(u_n),\infty)} \exp\{-N[u_n y - \psi_n(u_n)]\}\, d\overline{Q}_{nu_n}(y) \\
&\le \int_{(\lambda_n(u_n),\infty)} \exp\{-N[u_n\lambda_n(u_n) - \psi_n(u_n)]\}\, d\overline{Q}_{nu_n}(y) \\
&= \exp\left\{-\tfrac{1}{2}Nr_{1n*}^2(\lambda_n(u_n))\right\} Q_{nu_n}(\overline{Y}_n > \lambda_n(u_n)) \\
&\le \exp\left\{-\tfrac{1}{2}Nr_{1n*}^2(\lambda_n(u_n))\right\}.
\end{aligned}$$

(With the notation $Q_{nu_n}(\overline{Y}_n > \lambda_n(u_n))$ we mean that $Y_{n1}, \ldots, Y_{nn}$ are i.i.d. r.v.'s with distribution Q_{nu_n}.) On the other hand, we have for any $C > 0$, writing A_j for $(\lambda_n(u_n) + (j-1)CN^{-1}, \lambda_n(u_n) + jCN^{-1})$,

$$\begin{aligned}
&P_{\theta_0}\left(T_N^* > \sqrt{N}\mu_{n*}(\theta_n) + d_n\right) \\
&\geq \sum_{j=1}^{C^{-1}N^{1/2}} \int_{A_j} \exp\left\{-N\left[u_n y - \psi_n(u_n)\right]\right\} d\overline{Q}_{nu_n}(y) \\
&\geq \sum_{j=1}^{C^{-1}N^{1/2}} \int_{A_j} \exp\left\{-N\left[u_n\lambda_n(u_n) + u_n jCN^{-1} - \psi_n(u_n)\right]\right\} d\overline{Q}_{nu_n}(y) \\
&= \sum_{j=1}^{C^{-1}N^{1/2}} \exp\left\{-\tfrac{1}{2}Nr_{1n*}^2(\lambda_n(u_n)) - u_n jC\right\} Q_{nu_n}(\overline{Y}_n \in A_j).
\end{aligned}$$

Noting that Y_{ni} is uniformly bounded and hence also its third central moment, it follows from the Berry-Esseen theorem that, for some $c > 0$,

$$\sup_{x\in\mathbb{R}}\left|Q_{nu_n}\left(\frac{\overline{Y}_n - \lambda_n(u_n)}{\{\psi_n''(u_n)\}^{1/2}}\sqrt{N} \leq x\right) - \Phi(x)\right| \leq cN^{-1/2}$$

for all n. Note that $\psi_n''(u_n)$ is bounded away from 0 and infinity. (This can be seen as follows. If $\theta_n \to 0$, then $\psi_n''(u_n) = \text{var}_{nu_n}Y_{ni} \to 1$. If $\theta_n \to \theta_1 \neq 0$, then $\psi_n''(u_n) \to \text{var}_u Y$, where

$$Y = \frac{\log(1+\theta_1 a(X)) - E_{\theta_0}\log(1+\theta_1 a(X))}{\{\text{var}_{\theta_0}\log(1+\theta_1 a(X))\}^{1/2}},$$

u is given by $\lambda(u) = E_{\theta_1}Y, \lambda = \psi', \psi(v) = \log E_{\theta_0}\exp(vY)$ and $\text{var}_u Y$ refers to the variance of Y under Q_u with

$$dQ_u(y) = \exp(uy - \psi(u))dR_{\theta_0}(y),$$

where R_{θ_0} denotes the distribution of Y with X uniformly distributed on $[0,1]$.)

Hence, by taking C large enough we get, for all $1 \leq j \leq C^{-1}N^{1/2}$,

$$\begin{aligned}
&\liminf_{n\to\infty} N^{1/2}Q_{nu_n}(\overline{Y}_n \in A_j) \\
&= \liminf_{n\to\infty} N^{1/2}Q_{nu_n}\left(\frac{(j-1)CN^{-1/2}}{\{\psi_n''(u_n)\}^{1/2}} < \frac{\overline{Y}_n - \lambda_n(u_n)}{\{\psi_n''(u_n)\}^{1/2}}\sqrt{N} < \frac{jCN^{-1/2}}{\{\psi_n''(u_n)\}^{1/2}}\right) \geq 1.
\end{aligned}$$

Therefore,

$$
\begin{aligned}
&P_{\theta_0}(T_N^* > \sqrt{N}\mu_{n*}(\theta_n) + d_n) \\
&\geq N^{-1/2}(1+o(1)) \exp\left\{-\tfrac{1}{2}N r_{1n*}^2(\lambda_n(u_n))\right\} \sum_{j=1}^{C^{-1}N^{1/2}} \exp(-u_n jC) \\
&= N^{-1/2}(1+o(1)) \exp\left\{-\tfrac{1}{2}N r_{1n*}^2(\lambda_n(u_n)) - u_n C\right\} \frac{1-\exp(-u_n N^{1/2})}{1-\exp(-u_n C)}.
\end{aligned}
$$

Since $v_n \to 0$ implies $r_{1n*}(v_n)/v_n \to 1$, it follows from (2.9) that $u_n N^{1/2} \to \infty$ and thus we get

$$
\begin{aligned}
&P_{\theta_0}(T_N^* > \sqrt{N}\mu_{n*}(\theta_n) + d_n) \\
&\geq (N u_n^2)^{-1/2} \exp\left\{-\tfrac{1}{2}N r_{1n*}^2(\lambda_n(u_n)) + O(1)\right\}.
\end{aligned}
$$

(The upper bound can be sharpened by the same method as used in the lower bound, cf. (the proof of) Lemma 3.2 in Kallenberg (1981b).)The upper and lower bound result in

$$
\begin{aligned}
-N^{-1}\log\left\{P_{\theta_0}(T_N^* > \sqrt{N}\mu_{n*}(\theta_n) + d_n)\right\} &= \tfrac{1}{2}r_{1n*}^2(\lambda_n(u_n)) + O(N^{-1}\log(N u_n^2)) \\
= \tfrac{1}{2}r_{1n*}^2(\mu_{n*}(\theta_n)) + d_n N^{-1/2} r_{2n*}(\mu_{n*}(\theta_n)) &+ d_n^2 N^{-1} r_{2n*}'(\xi_n) + O(N^{-1}\log(N u_n^2))
\end{aligned}
$$

with ξ_n between $\mu_{n*}(\theta_n)$ and $\mu_{n*}(\theta_n) + d_n N^{-1/2}$. Since $r_{2n*}'(v) = \{\psi_n''(\lambda_n^{-1}(v))\}^{-1}$, it easily follows that $r_{2n*}'(\xi_n)$ is bounded. Noting that $\sqrt{N}\mu_{n*}(\theta_n) \to \infty$ by (2.9) we get $d_n^2 N^{-1} r_{2n*}'(\xi_n) = o(N^{-1/2}\mu_{n*}(\theta_n))$. Moreover, $\lambda_n(u_n)/u_n \to 1$ if $u_n \to 0$ and hence $O(N^{-1}\log(N u_n^2)) = O(N^{-1/2}\mu_{n*}(\theta_n)\{N\mu_{n*}^2(\theta_n)\}^{-1/2}\log\{N\mu_{n*}^2(\theta_n)\}) = o(N^{-1/2}\mu_{n*}(\theta_n))$, since $N\mu_{n*}^2(\theta_n) \to \infty$. This completes the proof that (A2') holds true. Therefore, Theorem 3.1 can be applied in the Pitman case, using (A2), and in the whole intermediate and Bahadur case we can apply Theorem 3.1′, using (A2′). □

Remark 3.7 It is seen in Example 3.6 that the whole range of sequences $\{\alpha_n\}$ and $\{\theta_n\}$ under consideration in Pitman-, Bahadur- and intermediate effciency is covered. Moreover, the method applied in Example 3.6 can be generalized to other testing problems where MP tests of Θ_0 against a simple alternative are based on sums of i.i.d. r.v.'s, provided that the moment generating function of the involved r.v. exists on a sufficiently large interval. □

The examples will be investigated further in Section 5, where Theorem 3.1 will be applied on several first order efficient tests after the general discussion on these tests in Section 4.

4 First order efficient tests

In this section we present an expansion for $N(\alpha_n, \beta, \theta_n)$ (a counterpart of Theorem 3.2) when ψ is based on a test statistic. The result is applied to show that as a rule vanishing shortcoming is equivalent to first order efficiency.

The basic assumptions of this section are modifications of (A1) and (A2) and an extra condition to replace (2.2), which obviously holds for MP tests, but not automatically for other tests. Condition (A1) is replaced by (B1), which is obtained from (A1) by writing $\psi_{n;\alpha}$ and T_n instead of $\psi^*_{n;\alpha}$ and T^*_n, i.e.

(B1)

$$\psi_{n;\alpha}(s) = \begin{cases} 1 & \text{if } T_n(s) > c_n \\ \delta_n & \text{if } T_n(s) = c_n \\ 0 & \text{if } T_n(s) < c_n, \end{cases} \tag{4.1}$$

where $c_n = c_n(\alpha) = \inf\{c : \sup\{P_{\theta_0}(T_n(S) > c) : \theta_0 \in \Theta_0\} \leq \alpha\}$ and $\delta_n = \delta_n(\alpha) = \sup\{\delta \in [0,1] : \sup\{P_{\theta_0}(T_n(S) > c_n) + \delta P_{\theta_0}(T_n(S) = c_n) : \theta_0 \in \Theta_0\} \leq \alpha\}$. Then we have for all $c < c_n$

$$\sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T_n(S) > c_n) \leq \alpha < \sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T_n(S) > c).$$

Condition (A2) is replaced by the following one.

(B2) There exist continuous distribution functions G_1 and G_2 on $\mathbb{R}$, strictly increasing on its support, which is $\mathbb{R}$ or $[a, \infty)$ with $a \in \mathbb{R}$, and a function $\mu : \Theta_1 \to (0, \infty)$ such that for every sequence $N = N(n)$ of natural numbers satisfying

$$\begin{aligned} &\sqrt{N}\mu(\theta_n) - G_1^{-1}(1 - \alpha_n) = O(1) \text{ and} \\ &\liminf_{n\to\infty} \sqrt{N}\mu(\theta_n) > 0 \text{ in case } \bar{\alpha} > 0 \end{aligned} \tag{4.2}$$

it holds that

$$\lim_{n\to\infty} P_{\theta_n}(T_N - \sqrt{N}\mu(\theta_n) \leq x) = G_2(x) \text{ for every } x \in \mathbb{R}, \tag{4.3}$$

and, as $n \to \infty$,

$$\begin{aligned} &\sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T_N > t_n) = 1 - G_1(t_n + o(1)) \\ &\text{for all } t_n = \sqrt{N}\mu(\theta_n) + c \end{aligned} \tag{4.4}$$

with $c \in \mathbb{R}$, not depending on n. Moreover, $\limsup_{n\to\infty} \mu(\theta_n) < \infty$ and $\lim_{n\to\infty} \mu(\theta_n) = 0$ in case $\bar{\alpha} > 0$.

The extra condition replacing (2.2) is as follows.

(B3) For every sequence $N = N(n)$ of natural numbers satisfying

$$\lim_{n\to\infty} \left\{\sqrt{N}\mu(\theta_n) - G_1^{-1}(1-\alpha_n)\right\} = \infty \text{ we have } \lim_{n\to\infty} \beta_N(\alpha_n, \theta_n) = 1.$$

This condition is used to show that if the power tends to 1 for some sequence $m(n)$, it still goes to 1 if we have even more observations. Such a property is needed, due to the definition of $N(\alpha, \beta, \theta)$.

Inspection of the proof of Theorem 3.2 shows that the first part of the proof can be copied, replacing $T_N^*, N_*(\alpha_n, \beta, \theta_n), \beta_N^*$ by $T_N, N(\alpha_n, \beta, \theta_n), \beta_N$ and G_* by G_1 or G_2 in an obvious way. To prove the modified version of (3.9), suppose that there exist $\delta > 0$ and a sequence $n(k), k \geq 1$, of natural numbers with $n(k) \to \infty$, such that for all $k \geq 1$

$$\sqrt{N(\alpha_{n(k)}, \beta, \theta_{n(k)})}\, \mu(\theta_{n(k)}) - G_1^{-1}(1-\alpha_{n(k)}) > -G_2^{-1}(1-\beta) + \delta.$$

By definition of $N(\alpha, \beta, \theta)$ we have

$$\beta_{N(\alpha,\beta,\theta)-1}(\alpha, \theta) < \beta.$$

If $\sqrt{N(\alpha_{n(k)}, \beta, \theta_{n(k)}) - 1}\, \mu(\theta_{n(k)}) - G_1^{-1}(1-\alpha_{n(k)})$ contains a subsequence tending to ∞, then by (B3) we get a contradiction. If this sequence is bounded, we can proceed as in (3.6)–(3.7), leading again to a contradiction. Hence, we have the following result.

Theorem 4.1 *Assume* (B1), (B2) *and* (B3). *For all $\beta \in (\bar{\alpha}, 1)$, with G_2^{-1} $(1-\beta) < G_1^{-1}(1-\bar{\alpha})$ in case $\bar{\alpha} > 0$, we have*

$$N(\alpha_n, \beta, \theta_n) = \left\{ \frac{G_1^{-1}(1-\alpha_n) - G_2^{-1}(1-\beta) + o(1)}{\mu(\theta_n)} \right\}^2.$$

Suppose that (B2) holds for $\psi_{n;\alpha}$ with $G_1 = G_*$ (see Sections 5.3 and 5.4). Often this can be established by a monotone transformation of the test statistic, thus keeping the same test. For instance, if

$$\sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T_N > t_n) = 1 - G_1(t_n + o(1))$$

we might consider $\tilde{T}_N = G_*^{-1}(G_1(T_N))$. Then we have

$$\sup_{\theta_0 \in \Theta_0} P_{\theta_0}(\tilde{T}_N > t_n) = \sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T_N > G_1^{-1}(G_*(t_n)))$$

$$= 1 - G_1\left(G_1^{-1}(G_*(t_n)) + o(1)\right),$$

which often reduces to $1-G_*(t_n+o(1))$. Assume further that the asymptotic shift represented by μ and μ_* is the same: $\mu = \mu_*$. If we have first order efficiency, i.e. $N_*(\alpha_n, \beta, \theta_n)/N(\alpha_n, \beta, \theta_n) = 1 + o(1)$, this equality will often hold true. Having no difference in shift, suppose that there is a possible difference in scale. So, suppose that $G_2(x) = G_*(x/\sigma)$ for some $\sigma \in \mathbb{R}$.

Since $N(\alpha_n, \beta, \theta_n) \geq N_*(\alpha_n, \beta, \theta_n)$ for all $\beta \in (\bar{\alpha}, 1)$, by definition of N_*, we get in view of Theorem 3.2 and Theorem 4.1,

$$\left\{ \frac{G_*^{-1}(1-\alpha_n) - \sigma G_*^{-1}(1-\beta) + o(1)}{\mu_*(\theta_n)} \right\}^2 \geq \left\{ \frac{G_*^{-1}(1-\alpha_n) - G_*^{-1}(1-\beta) + o(1)}{\mu_*(\theta_n)} \right\}^2$$

for all $\beta \in (\bar{\alpha}, 1)$. If $G_*^{-1}(1-\beta)$ takes positive as well as negative values (which for instance is the case if G_* is the standard normal distribution function), then we get $\sigma = 1$. Therefore, the same shift implies automatically the same scale. Concrete examples of this phenomenon are presented in Sections 5.3 and 5.4, when discussing the Cramér-von-Mises test and the Anderson-Darling test. A similar argument as above is used in proving first order efficiency implies second order efficiency in Kallenberg (1983a).

As a conclusion we may state that in many situations tests which are first order efficient ($N_*(\alpha_n, \beta, \theta_n)/N(\alpha_n, \beta, \theta_n) = 1 + o(1)$) satisfy not only $\mu_* = \mu$ and $G_* = G_1$, but automatically also $G_1 = G_2$. We call such a situation a "regular case". It follows in that case that for each $\beta \in (\bar{\alpha}, 1)$

$$\begin{aligned} & N(\alpha_n, \beta, \theta_n) - N_*(\alpha_n, \beta, \theta_n) = o\left(G_*^{-1}(1-\alpha_n)/\mu_*^2(\theta_n)\right) \\ & = o\left(\sqrt{N_*(\alpha_n, \beta, \theta_n)}/\mu_*(\theta_n)\right). \end{aligned} \tag{4.5}$$

This, together with Theorem 3.1, yields

Corollary 4.2 *Assume* (A1) *and* (A2) *for* $\psi^*_{n;\alpha}$ *and* (B1), (B2) *and* (B3) *for* $\psi_{n;\alpha}$ *with* $\mu_* = \mu$, $G_* = G_1 = G_2$. *Then*

$$N(\alpha_n, \beta, \theta_n) - N_*(\alpha_n, \beta, \theta_n) = o\left(\sqrt{N_*(\alpha_n, \beta, \theta_n)}/\mu_*(\theta_n)\right)$$

for each $\beta \in (\bar{\alpha}, 1)$

and

$$\lim_{n\to\infty} R_m(\alpha_n, \theta_n) = 0 \quad \text{for each sequence } m = m(n), \text{ provided that } \liminf_{n\to\infty} \beta^*_m(\alpha_n, \theta_n) > \bar{\alpha} \text{ in case } \bar{\alpha} > 0.$$

From Corollary 4.2 and the discussion above it is seen that the seemingly stronger property of vanishing shortcoming is, in regular cases, in fact equivalent to first order efficiency.

The exception in Corollary 4.2 with respect to the shortcoming is not very serious. Let $\bar{\alpha} > 0$ and $\lim_{n\to\infty} \beta_m^*(\alpha_n, \theta_n) = \bar{\alpha}$. This means that with the $m(n)$ observations θ_n is too close to Θ_0 to detect. As a rule $\beta_m(\alpha_n, \theta_n)$ will also converge to $\bar{\alpha}$ in this case, at least if the test ψ is asymptotically unbiased. However, the conditions (B1), (B2) and (B3) give no information on this exceptional occasion. So, formally the restriction should be there.

5 Applications and extensions

It has been shown in the previous sections that shortcoming and first order efficiency are strongly related optimality concepts. The equivalence holds in quite generality as is seen from the very general structure of the testing problem, the different types of efficiency concepts which are involved, from local to non-local, and the rather weak conditions imposed on the test statistics. Here we consider some concrete examples and applications. Also some extensions to comparison of tests which are not efficient are discussed.

5.1 Student test

Consider the situation from Example 3.4.

In the Pitman situation with $\theta_n = cn^{-1/2}$ for some $c > 0$ and $\alpha_n = \bar{\alpha} \in (0, 1)$, we consider the one-sided Student's t-test. It is easily seen that (B1), (B2) and (B3) hold with $G_1 = G_2 = G_* = \Phi$ and $\mu(\theta) = \mu_*(\theta) = \theta$. By Corollary 4.2 we get the well-known result that the t-test is Pitman-efficient (cf. e.g. Serfling (1980), p. 320) and that its shortcoming tends to 0.

Next consider a sequence $\{\theta_n\}$ with $\theta_n \to 0$ and $n^{1/2}\theta_n \to \infty$ as $n \to \infty$. Let $\{\alpha_n\}$ be a sequence of levels satisfying $\lim_{n\to\infty} \alpha_n = 0$ and $\Phi^{-1}(1 - \alpha_n) = o(n^{1/2})$. Again conditions (B1), (B2) and (B3) hold (note that the power of the t-test increases with n). Hence, Corollary 4.2 gives that the shortcoming of the t-test tends to 0 for all these alternatives and, equivalently, that its efficiency equals $N_*(\alpha_n, \beta, \theta_n)/N(\alpha_n, \beta, \theta_n) = 1 + o\left(\{\theta_n N_*(\alpha_n, \beta, \theta_n)\}^{-1/2}\right) = 1 + o\left(\{\Phi^{-1}(1 - \alpha_n)\}^{-1/2}\right)$.

Finally, in the Bahadur case with $\theta_n = \theta > 0$, (B1) and (4.3) continue to hold with $G_2 = G_* = \Phi$ and $\mu(\theta) = \mu_*(\theta) = \theta$, but

$$N^{-1} \log P_0 \left(T_N > \theta\sqrt{N}\right) \to -\tfrac{1}{2} \log(1 + \theta^2)$$

and hence (4.4) does not hold with $G_1 = G_* = \Phi$. This corresponds to the fact that the t-test is not Bahadur-efficient.

5.2 Curved exponential families

Consider the situation from Example 3.5 and the locally most powerful test. Its test statistic is of the form $T_n = \dot{\gamma}'_{\theta_0}(\overline{X}_n - E_{\theta_0}X)\sqrt{n}/s$. It is seen from Lemma 3.7 in Kallenberg (1981a) with the number of observation n replaced by m that $R_m(\alpha_n, \theta_n) \to 0$ in all cases, provided that

$$m^{-1}|\log \alpha_n|^{3/2} \to 0 \text{ if } s(\theta_n - \theta_0)m^{1/2}\{-2\log \alpha_n\}^{-1/2} \to 1. \tag{5.1}$$

Writing

$$m^{-1}|\log \alpha_n|^{3/2} = (\theta_n - \theta_0)^{-2}m^{-1}|\log \alpha_n|(\theta_n - \theta_0)^2|\log \alpha_n|^{1/2}$$

it immediately follows that (5.1) is implied by (3.25). Hence, for all sequences of levels $\{\alpha_n\}$ and all sequences of alternatives $\{\theta_n\}$ satisfying (3.24) and (3.25) $R_m(\alpha_n, \theta_n) \to 0$ for each sequence $m = m(n)$.

Therefore, by Theorem 3.1, the locally most powerful test is first order efficient in the Pitman case and in the intermediate case, provided in the latter situation that (3.24) and (3.25) are satisfied. (Of course, the locally most powerful test is in general not first-order efficient in the sense of Bahadur, see Kallenberg (1981a), p. 673.)

5.3 Cramér-von-Mises test

Consider the situation from Example 3.6. The Cramér-von-Mises test statistic for this testing problem is defined by rejecting H_0 for large values of

$$\left\{n\int_0^1 (F_n(x) - x)^2 dx\right\}^{1/2}, \tag{5.2}$$

where F_n is the empirical distribution function. We will show that (within the class (3.28)) there is only one function $a(x), a(x) = C_1(x) = \sqrt{2}\cos(\pi x)$, under which the Cramér-von-Mises test is efficient (cf. Corollary 5.9). This result supplies analogous results stated in another framework by Neuhaus (1976) and Nikitin (1995). Moreover, the asymptotic relative efficiency under other alternatives than the one given by $C_1(x)$ will be calculated. The phenomenon that equality of asymptotic optimal shift implies also equality of scale terms, as discussed in general terms in Section 4, is illustrated clearly in the case. Similar results for the Anderson-Darling statistic are treated in Section 5.4.

To apply Theorem 4.1 we have still to check conditions (B1), (B2) and (B3). First we introduce some notation

$$A(t) = \int_0^t a(x)dx,\ \|A\| = \left\{\int_0^1 A^2(t)dt\right\}^{1/2},\ \mu(\theta) = \pi\|A\|\theta. \tag{5.3}$$

Instead of using (5.2) as test statistic we take the equivalent test statistic

$$T_n = \pi \left\{ n \int_0^1 (F_n(x) - x)^2 dx \right\}^{1/2}. \tag{5.4}$$

This simplifies notation a little bit and has no influence upon the results on $N(\alpha_n, \beta, \theta_n)$.

Condition (B1) holds. Lemmas 5.1 and 5.2, given below, yield (4.4) with $G_1 = \Phi$. The proof of Lemma 5.1 follows from inequalities in Inglot and Ledwina (1990). The constant π involved in (5.4) can be deduced from Nikitin (1995), e.g., while a related result on the tails of the limiting distribution of T_n can be found in Gregory (1980), cf. also Theorem 2.1 and Proposition 2.3 in Inglot et al. (1993).

Lemma 5.1 *If $x_n \to 0$ and $nx_n^2 \to \infty$ then for any $2 < \rho < 3$ it holds that*

$$\log P_{\theta_0}(T_n > x_n\sqrt{n}) = -\tfrac{1}{2} n x_n^2 + O(n x_n^\rho). \tag{5.5}$$

Lemma 5.2 *For any $\{\theta_n\}$ and $\{\alpha_n\}$ such that for $n \to \infty$*

$$\alpha_n \to 0 \text{ and } \theta_n^\gamma \Phi^{-1}(1 - \alpha_n) \to 0 \text{ for some } \gamma \in (0, 1) \tag{5.6}$$

(4.4) *holds with $G_1 = \Phi$.*

Proof. Using $\Phi^{-1}(1 - \epsilon) = (-2 \log \epsilon)^{1/2} + o(1)$ and writing $t_n = \sqrt{N}\mu(\theta_n) + c$ with $c \in \mathbb{R}$ and N satisfying $\sqrt{N}\mu(\theta_n) = \Phi^{-1}(1 - \alpha_n) + O(1)$, cf. (4.2), we derive from (5.5)

$$\Phi^{-1}(1 - P_{\theta_0}(T_N > t_n)) = \{t_n^2 + O(t_n^2 \theta_n^{\rho-2})\}^{1/2} + o(1)$$
$$= t_n(1 + O(\theta_n^{\rho-2})) + o(1) = t_n + o(1)$$

by taking $\rho - 2 = \gamma$, applying (5.6) and noting that $t_n = \Phi^{-1}(1 - \alpha_n) + O(1)$. The result now easily follows. □

The next lemma yields (4.3).

Lemma 5.3 *Let $\theta_n \to 0$ and let $M = M(n)$ be an arbitrary sequence satisfying*

$$M \to \infty \text{ and } \sqrt{M}\theta_n \to \infty.$$

Then

$$\lim_{n\to\infty} P_{\theta_n}\left(T_M - \sqrt{M}\mu(\theta_n) \le x\right) = \Phi\left(\frac{\|A\|}{\sigma\pi} x\right),$$

where

$$\sigma^2 = \int_0^1 \int_0^1 (s \wedge t - st) A(s) A(t) ds dt. \tag{5.7}$$

Hence, (4.3) *holds with $G_2(x) = \Phi(\|A\| x / (\sigma\pi))$.*

Proof. Set $A_n(t) = \int_0^t p_n(x)dx$. Then $A_n(t) = t + \theta_n A(t)$ and $\sup_t |A_n(t) - t| \leq \theta_n$. Let $\{\alpha_M^*\}$ and $\{B_M\}$ be sequences of the empirical processes and the Brownian bridges defined via the KMT construction. Then, under p_n, we have $\sqrt{M}(F_M - A_n) \stackrel{\mathcal{D}}{=} \alpha_M^*(A_n) = \{\alpha_M^*(A_n) - B_M(A_n)\} + \{B_M(A_n) - B_M\} + B_M$. Therefore, under p_n, $\sqrt{M}(F_M - A_n) \xrightarrow{\mathcal{D}} B$, where B is a Brownian bridge. Hence, for $T_M = \pi\|\sqrt{M}(F_M - \mathbb{I})\| = \pi\|\sqrt{M}(F_M - A_n) + \sqrt{M}\theta_n A\|$, with $\mathbb{I}(t) = t$, it follows that

$$\left(\pi^{-2}T_M^2 - M\theta_n^2\|A\|^2\right) / \left(2\sqrt{M}\theta_n\right) \xrightarrow{\mathcal{D}} \int_0^1 B(t)A(t)dt$$

and the rest of the proof can be easily deduced. □

Lemmas 5.1 and 5.3 can also be used to prove (B3).

Lemma 5.4 *Assume* (5.6). *Then condition* (B3) *holds true.*

Proof. Let $N = N(n)$ be a sequence of natural numbers satisfying

$$\lim_{n\to\infty} \left\{\sqrt{N}\mu(\theta_n) - \Phi^{-1}(1-\alpha_n)\right\} = \infty. \tag{5.8}$$

Consider the test which rejects the null hypothesis if

$$T_N > \Phi^{-1}(1-\alpha_n) + 1. \tag{5.9}$$

Note that $\Phi^{-1}(1-\alpha_n) \leq \sqrt{N}\mu(\theta_n)$ for sufficiently large n and hence $\left\{\Phi^{-1}(1-\alpha_n)+1\right\}/\sqrt{N} \to 0$. By Lemma 5.1 we get for the level $\tilde{\alpha}_n$ of the test given in (5.9)

$$\log\tilde{\alpha}_n = -\tfrac{1}{2}\left\{\Phi^{-1}(1-\alpha_n)+1\right\}^2 + O\left(N\left[\left(\Phi^{-1}(1-\alpha_n)+1\right)/\sqrt{N}\right]^\rho\right). \tag{5.10}$$

Writing

$$N\left[\left(\Phi^{-1}(1-\alpha_n)+1\right)/\sqrt{N}\right]^\rho = \left\{\Phi^{-1}(1-\alpha_n)+1\right\}\theta_n^{\rho-2}\left\{\Phi^{-1}(1-\alpha_n)+1\right\}$$
$$\left[\left(\Phi^{-1}(1-\alpha_n)+1\right)/\left(\sqrt{N}\theta_n\right)\right]^{\rho-2}$$

and taking $\rho = \gamma + 2$ it is seen from (5.6) and (5.8) that (5.10) reduces to

$$\log\tilde{\alpha}_n = -\tfrac{1}{2}\left\{\Phi^{-1}(1-\alpha_n)+1\right\}^2 + o(\Phi^{-1}(1-\alpha_n))$$

and hence

$$\log\tilde{\alpha}_n = \log\alpha_n - (-2\log\alpha_n)^{1/2} + o(|\log\alpha_n|^{1/2}) < \log\alpha_n$$

for sufficiently large n. Therefore $\beta_N(\tilde{\alpha}_n, \theta_n) \le \beta_N(\alpha_n, \theta_n)$ for sufficiently large n. Since $\Phi^{-1}(1-\alpha_n) + 1 - \sqrt{N}\mu(\theta_n) \to -\infty$, application of Lemma 5.3 now yields

$$\lim_{n\to\infty} \beta_N(\tilde{\alpha}_n, \theta_n) = \lim_{n\to\infty} P_{\theta_n}(T_N > \Phi^{-1}(1-\alpha_n) + 1) = 1 = \lim_{n\to\infty} \beta_N(\alpha_n, \theta_n).$$

This completes the proof of the lemma. □

Theorem 4.1 can now be applied and we get the following result.

Theorem 5.5 *Let $p_n(x) = 1 + \theta_n a(x)$ be as in* (3.28)*. Suppose that* (5.6) *holds. Then for each $\beta \in (0, 1)$*

$$N(\alpha_n, \beta, \theta_n) = \left\{ \Phi^{-1}(1-\alpha_n) - \frac{\pi\sigma}{\|A\|}\Phi^{-1}(1-\beta) + o(1) \right\}^2 / \{\pi\|A\|\theta_n\}^2, \quad (5.11)$$

where σ and A are given in (5.7) *and* (5.3)*.*

Application of Theorem 3.2, cf. Example 3.6, yields

Theorem 5.6 *Let $p_n(x) = 1 + \theta_n a(x)$ be as in* (3.28)*. Suppose that* (3.29) *holds. Then for each $\beta \in (0, 1)$*

$$N_*(\alpha_n, \beta, \theta_n) = \left\{ \Phi^{-1}(1-\alpha_n) - \Phi^{-1}(1-\beta) + o(1) \right\}^2 / \theta_n^2. \quad (5.12)$$

Theorems 5.5 and 5.6 imply the following corollary.

Corollary 5.7 *Let $p_n(x) = 1 + \theta_n a(x)$ be as in* (3.28)*. Suppose that* (5.6) *holds. Then for each $\beta \in (0, 1)$*

$$\lim_{n\to\infty} \frac{N_*(\alpha_n, \beta, \theta_n)}{N(\alpha_n, \beta, \theta_n)} = (\pi\|A\|)^2. \quad (5.13)$$

To study the magnitude of $\pi\|A\|$ the following lemma is convenient.

Lemma 5.8 *Let a be bounded, $\int_0^1 a(x)dx = 0$ and $\int_0^1 a^2(x)dx = 1$. Then $\|A\| \le \pi^{-1}$ and $\sigma \le \pi^{-2}$. Besides, in both inequalities equality holds if and only if $a(x) = C_1(x) = \sqrt{2}\cos(\pi x)$.*

Proof. Write

$$a(x) = \sum_{k=1}^{\infty} \boldsymbol{c}_k C_k(x), \text{ where } C_k(x) = \sqrt{2}\cos(\pi k x)$$

and the series converges in $L_2[0, 1]$. Observe that $\sum_{k=1}^{\infty} \boldsymbol{c}_k^2 = 1$. By continuity of the inner product in $L_2[0, 1]$ we have for each $x \in [0, 1]$

$$A(x) = \int_0^1 1_{[0,x]}(u)a(u)du = \sum_{k=1}^{\infty} \boldsymbol{c}_k \int_0^1 1_{[0,x]}(u)\boldsymbol{c}_k(u)du = \sum_{k=1}^{\infty} \frac{\boldsymbol{c}_k}{\pi k}\sqrt{2}\sin(\pi k x)$$

and the series on the right converges also in $L_2[0,1]$. Hence

$$\|A\|^2 = \sum_{k=1}^{\infty} \frac{\boldsymbol{c}_k^2}{\pi^2 k^2}\,. \tag{5.14}$$

As the integral is a continuous linear functional in $L_2[0,1]$ it easily follows that, writing $\overline{C}_k(x) = \sqrt{2}\sin(\pi k x)$,

$$\sigma^2 = \int_0^1 \int_0^1 (s \wedge t - st)A(s)A(t)dsdt \tag{5.15}$$

$$= \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \frac{\boldsymbol{c}_k \boldsymbol{c}_j}{\pi^2 kj} \int_0^1 \overline{C}_j(t) \left\{ \int_0^1 (s \wedge t - st)\overline{C}_k(s)ds \right\} dt$$

$$= \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} \frac{\boldsymbol{c}_k \boldsymbol{c}_j}{\pi^2 kj} \int_0^1 \frac{\overline{C}_j(t)\overline{C}_k(t)}{\pi^2 k^2} dt = \sum_{k=1}^{\infty} \frac{\boldsymbol{c}_k^2}{\pi^4 k^4}.$$

Hence the conclusion follows. □

Lemma 5.8 together with Theorems 5.5, 5.6 and 3.1 give a useful corollary. For short introduce the notation

$$ARE(T_n, T_n^*) = \lim_{n\to\infty} \frac{N_*(\alpha_n, \beta, \theta_n)}{N(\alpha_n, \beta, \theta_n)}$$

for the asymptotic relative efficiency of T_n with respect to T_n^*.

Corollary 5.9 *Under the assumptions of Theorem 5.5 we have the following situation.*

> *If $p_n(x) = 1 + \theta_n C_1(x)$, $C_1(x) = \sqrt{2}\cos(\pi x)$, then $ARE(T_n, T_n^*) = 1$ and moreover, $N_*(\alpha_n, \beta, \theta_n) - N(\alpha_n, \beta, \theta_n) = o\left(\sqrt{N_*(\alpha_n, \beta, \theta_n)}/\theta_n\right)$. Hence, for these alternatives $\lim_{n\to\infty} R_m(\alpha_n, \theta_n) = 0$ for each sequence $m = m(n)$. For any other alternative $p_n(x) = 1 + \theta_n a(x)$, satisfying our assumptions, $ARE(T_n, T_n^*) < 1$. In particular, for $p_n(x) = 1 + \theta_n C_k(x)$, $C_k(x) = \sqrt{2}\cos(k\pi x)$, $ARE(T_n, T_n^*) = k^{-2}$.*

Theorems 5.5 and 5.6 clearly illustrate the phenomenon discussed in Section 4. To get first order efficiency, the asymptotic shift $\mu(\theta_n) = \pi\|A\|\theta_n$ of T_n should be the same as that of T_n^*, which equals $\mu_*(\theta_n) = \theta_n$. Hence, $\pi\|A\|$ should be equal to 1. However, (5.11) and (5.12) show that also the terms corresponding to the scale of the limiting distributions G_2 and G_* are different: $\pi\sigma/\|A\|$ and

1, respectively. But, in case $\pi\|A\| = 1$, these scale terms should be the same as well due to the phenomenon discussed in Section 4. And, indeed this is the case: if $\pi\|A\|=1$, then $\sigma = \pi^{-2}$ and hence $\pi\sigma/\|A\| = 1$. This is shown in Lemma 5.8 analytically, but could also be derived from (5.11) and (5.12) directly. Let $\pi\|A\| = 1$. Since $N_*(\alpha_n, \beta, \theta_n) \leq N(\alpha_n, \beta, \theta_n)$ for all $\beta \in (0, 1)$ and $\Phi^{-1}(1-\beta)$ can be positive as well as negative, it follows that the coefficients of $\Phi^{-1}(1-\beta)$ in (5.11) and (5.12) should be the same. Together with $\pi\|A\| = 1$ this gives $\sigma = \pi^{-2}$.

The next section contains similar results for the Anderson-Darling test.

5.4 Anderson-Darling test

Consider the situation from Example 3.6. The Anderson-Darling test rejects for large values of

$$T_n = \left\{ 2n \int_0^1 \frac{(F_n(t) - t)^2}{t(1-t)} dt \right\}^{1/2}.$$

Set, cf. (5.3),

$$\mu(\theta_n) = \sqrt{2}\theta_n \|A_w\|, \; A_w(t) = A(t)/\sqrt{t(1-t)}.$$

Theorem 5.10

(a) *Suppose* $\theta_n \to 0$ *and* $M = M(n)$ *is an arbitrary sequence satisfying* $M \to \infty$ *and* $\sqrt{M}\theta_n \to \infty$. *Then*

$$\lim_{n\to\infty} P_{\theta_n}(T_M - \sqrt{M}\mu(\theta_n) \leq x) = \Phi(x\|A_w\|/(\sqrt{2}\sigma_w)),$$

where

$$\sigma_w^2 = \int_0^1 \int_0^1 \frac{(s \wedge t - st)A(s)A(t)}{s(1-s)t(1-t)} dsdt.$$

(b) *Let* $\{L_n(x)\}_{n\geq 0}$ *be the system of orthonormal Legendre polynomials on* $[0, 1]$ *and let* $\ell_n = \int_0^1 a(x)L_n(x)dx$. *We have*

$$\|A_w\|^2 = \sum_{k=1}^{\infty} \ell_k^2/\{k(k+1)\} \text{ and } \sigma_w^2 = \sum_{k=1}^{\infty} \ell_k^2/\{k^2(k+1)^2\}, \tag{5.16}$$

provided that a *is bounded,* $\int_0^1 a(x)dx = 0$ *and* $\int_0^1 a^2(x)dx = 1$.

(c) *If* $x_n \to 0$, $nx_n^2 \to \infty$, *then for any* $\rho \in (2, 3)$

$$\log P_{\theta_0}\left(T_n \geq x_n\sqrt{n}\right) = -\tfrac{1}{2}nx_n^2 + O(nx_n^\rho),$$

provided that $nx_n^\rho \to 0$.

(d) *Let* $p_n(x) = 1 + \theta_n a(x)$ *be as in* (3.28). *Suppose that* $\alpha_n \to 0$ *and* $\theta_n^\gamma \Phi^{-1}(1-\alpha_n) \to 0$ *for some* $\gamma \in (0, 1/2)$. *Then conditions* (B1), (B2) *and* (B3) *hold with* $G_1(x) = \Phi(x)$ *and* $G_2(x) = \Phi(x\|A_w\|/(\sqrt{2}\sigma_w))$. *Hence, for each* $\beta \in (0,1)$

$$N(\alpha_n, \beta, \theta_n) = \left\{\Phi^{-1}(1-\alpha_n) - \frac{\sqrt{2}\sigma_w}{\|A_w\|}\Phi^{-1}(1-\beta) + o(1)\right\}^2 / \{\sqrt{2}\|A_w\|\theta_n\}^2$$

and

$$ARE(T_n, T_n^*) = 2\|A_w\|^2. \tag{5.17}$$

(e) *Suppose that* $\alpha_n \to 0$ *and* $\theta_n^\gamma \Phi^{-1}(1-\alpha_n) \to 0$ *for some* $\gamma \in (0, 1/2)$. *If* $p_n(x) = 1 + \theta_n L_1(x)$, $L_1(x) = \sqrt{3}(2x-1)$, *then* $ARE(T_n, T_n^*) = 1$ *and, moreover,* $N_*(\alpha_n, \beta, \theta_n) - N(\alpha_n, \beta, \theta_n) = o\left(\sqrt{N_*(\alpha_n, \beta, \theta_n)}/\theta_n\right)$. *Hence, for these alternatives* $\lim_{n\to\infty} R_m(\alpha_n, \theta_n) = 0$ *for each sequence* $m = m(n)$. *For any other alternative* $p_n(x) = 1 + \theta_n a(x)$, *satisfying our assumptions,* $ARE(T_n, T_n^*) < 1$. *In particular, if* $p_n(x) = 1 + \theta_n L_k(x)$, L_k *the* k^{th} *Legendre polynomial,* $ARE(T_n, T_n^*) = 2/\{k(k+1)\}$.

Proof. Theorem 5.10 can be proved similarly as the corresponding results for the Cramér-von-Mises test. In particular, to get (5.16) it is useful to apply formulas of Sansone (1959), p. 175 to see that for $\{L_n(x)\}_{n\geq 0}$ one gets $(x^2 - x)L_n'(x) = n(n+1)\mathbb{L}_n(x)$, where $\mathbb{L}_n(x) = \int_0^x L_n(t)dt$, and to prove that $\{\mathbb{L}_n(x)\}_{n\geq 1}$, are orthogonal with the weight $\{x(1-x)\}^{-1}$. Expanding $a(x)$ in the $\{L_n(x)\}$ system, using continuity of the inner product and proceeding exactly as in the proof of Lemma 5.8, (5.16) follows.
To get (c) one can use results of Section 4 of Inglot and Ledwina (1993) or of Example 2.1 in Inglot et al. (1993). Note that γ should be smaller than 1/2 in (d) and (e), due to (c). □

The same remark on the effect that optimal asymptotic shift implies also equality of scale terms as after Corollary 5.9 applies here. If $\sqrt{2}\|A_w\| = 1$, σ_w should be 1 by this property. And indeed, this is the case.

5.5 Concluding remarks and extensions

In Sections 5.1–5.4 the equivalence of vanishing shortcoming and first-order efficiency is illustrated on a lot of examples. Numerous other examples and applications could be added, both in the Pitman case, the intermediate one and for fixed alternatives. The equivalence is partly due to the phenomenon that equality of asymptotic optimal shift implies also equality of scale terms. This is clearly illustrated in the applications in Sections 5.3 and 5.4, where for most directions there is no first order efficiency. In those cases the asymptotic shift *and* scale of T_n and T_n^* differ. As soon as the asymptotic shifts (first order efficiency) coincide, automatically also the asymptotic scales become the same, which in turn is equivalent to vanishing shortcoming.

The results of the Sections 5.3 and 5.4 can be extended to other statistics being bilinear functionals of the empirical process. For appropriate limit theorems see Inglot et al. (1993).

Results like (5.14), (5.15), (5.16) and (5.13), (5.17) are of independent interest. They provide a simple and intuitive way of comparison of quadratic tests with the best possible one or of two quadratic tests with each other. They supply Nikitin's (1995) results, where Bahadur's approach has been exploited. On the other hand, they can be nicely confronted with the two step approach proposed by Hájek and Sidák (1967) and applied in Neuhaus (1976) as well as Wieand's approach exploited by Gregory (1980). Moreover, the results coincide with those obtained in Section 7.7 of Inglot and Ledwina (1996), where a slightly different definition of asymptotic intermediate relative efficiency (cf. also Kallenberg (1983b)) has been applied.

Although the main theme of the paper concerns the relation between vanishing shortcoming and first order efficiency, Theorem 4.1 can also be applied to compare two tests with each other. This gives an easy way to calculate the asymptotic relative efficiency of T_n w.r.t. $\tilde{T}_n$, where T_n and $\tilde{T}_n$ are two test statistics.

Similarly, Theorem 3.3 can be generalized to other statistics than the MP.

Theorem 5.11 *Assume* (B1), (B2) *and* (B3). *Moreover, assume that for every sequence* $N = N(n)$ *of natural numbers satisfying* $\lim_{n\to\infty}\{\sqrt{N}\mu(\theta_n) - G_1^{-1}(1-\alpha_n)\} = -\infty$ *we have* $\lim_{n\to\infty}\beta_N(\alpha_n,\theta_n) = 0$. *Then, for each* $\beta \in (\bar{\alpha}, 1)$ *and for each sequence* $m = m(n)$ *we have*

$$\lim_{n\to\infty}\left[\beta_m(\alpha_n,\theta_n) - \left\{1 - G_2\left(G_2^{-1}(1-\beta) - \left[\sqrt{m} - \sqrt{N(\alpha_n,\beta,\theta_n)}\right]\mu(\theta_n)\right)\right\}\right] = 0.$$

The proof of Theorem 5.11 is obtained from the proof of Theorem 3.3 by obvious modifications and is therefore omitted.

As Theorem 4.1 can be used to get results on the asymptotic relative efficiency of T_n w.r.t. $\tilde{T}_n$, Theorem 5.11 in combination with Theorem 4.1 can be applied to obtain results on the "shortcoming of T_n w.r.t. $\tilde{T}_n$": $\beta_m(\alpha_n,\theta_n) - \tilde{\beta}_m(\alpha_n,\theta_n)$.

References

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

Book, S. A., (1976). The Cramér-Feller-Petrov large deviation theorem for triangular arrays, Technical Report, California State College.

Efron, B. (1975). Defining the curvature of a statistical problem. *Ann. Statist.* **3** 1189-1242.

Gregory, G. G. (1980). On efficiency and optimality of quadratic tests. *Ann. Statist.* **8** 116-131.

Hájek, J. and Sidák, Z. (1967). *Theory of Rank Tests.* Academic Press, New York.

Inglot, T., Kallenberg, W. C. M. and Ledwina, T. (1993). Asymptotic behavior of some bilinear functionals of the empirical process. *Math. Methods of Statist.* **2** 316-336.

Inglot, T. and Ledwina, T. (1990). On probabilities of excessive deviations for Kolmogorov-Smirnov, Cramér-von-Mises and chi-square statistics. *Ann. Statist.* **18** 1491-1495.

Inglot, T. and Ledwina, T. (1993). Moderately large deviations and expansions of large deviations for some functionals of weighted empirical processes. *Ann. Probab.* **21** 1691-1705.

Inglot, T. and Ledwina, T. (1996). Asymptotic optimality of data-driven Neyman's tests for uniformity. *Ann. Statist.* **24** 1982-2019.

Jurečková, J., Kallenberg, W. C. M. and Veraverbeke, N. (1988). Moderate and Cramér-type large deviation theorems for M-estimators. *Statist. Probab. Lett.* **6** 191-199.

Kallenberg, W. C. M. (1981a). The shortcoming of locally most powerful tests in curved exponential families. *Ann. Statist.* **9** 673-677.

Kallenberg, W. C. M. (1981b). Bahadur deficiency of likelihood ratio tests in exponential families. *J. Multivariate Anal.* **11** 506-531.

Kallenberg, W. C. M. (1983a). Asymptotic efficiency and deficiency of tests. In *Proceedings of the 44^{th} Session of the ISI.* Madrid, Invited Papers **2** 1173-1189.

Kallenberg, W. C. M. (1983b). Intermediate efficiency, theory and examples. *Ann. Statist.* **11** 170-182.

Lehmann, E. L. (1986). *Testing Statistical Hypotheses*, Second Edition. Wiley, New York.

Neuhaus, G. (1976). Asymptotic power properties of the Cramér-von-Mises test under contiguous alternatives. *J. Multivariate Anal.* **6** 95-110.

Nikitin, Ya. Yu. (1995). *Asymptotic Efficiency of Nonparametric Tests.* Cambridge Univ. Press.

Sansone, G. (1959). *Orthogonal Functions.* Interscience, New York.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* Wiley, New York.