Sankhyā : The Indian Journal of Statistics
1978, Volume 40, Series A, Pt. 3, pp. 253–277.

COMPARISON OF TESTS WITH SAME BAHADUR-EFFICIENCY

By TAPAS K. CHANDRA
Indian Statistical Institute
and
J. K. GHOSH
Indian Statistical Institute and University of Missouri

SUMMARY. Methods are developed to compare tests with same Bahadur efficiency. Measures of deficiency are introduced in the manner of Hodges and Lehmann. Since two tests which are equally efficient by Bahadur's criterion are usually equally efficient by Cochran's criterion also, the problem of measuring deficiency has also been approached from Cochran's point of view. It is shown that approaches based on Bahadur's and Cochran's ideas lead to the same measure of deficiency if one uses limits in probability in the definition of Bahadur slopes; the equivalence breaks down if one uses a.s. limits instead.

0. Introduction

In their paper Hodges and Lehmann (1970) studied the problem of discrimination between two statistical procedures which are according to some criterion, equally "efficient"; deficiency is essentially a quantitative measure of this discrimination. In the same spirit, we have discussed here the problem of discrimination between two test procedures which have equal Bahadur-efficiency.

It is suggested by Bahadur (1967 and 1971) that in many cases alternative test procedures might be compared on the basis of the associated limiting "attained levels". Following his suggestion we have introduced the notion of Bahadur-deficiency for two test procedures which are equally efficient from Bahadur's view-point. It appears that this approach of discrimination involves some difficulties; for example, the quantities involved are, in general, unlikely to be constants almost surely.

On the other hand, Cochran (1952) measured the efficiency of a test procedure by the rate of convergence (to zero) of its size, when the power is held fixed against a specified alternative. It is well known that Cochran's approach to efficiency usually leads precisely to the same conclusions as Bahadur's approach does. Motivated by this fact we have introduced in Section 2 the notion of Cochran-deficiency (to be referred to as BCD for reasons explained in the next paragraph) and have shown by means of an example that at the level of deficiency, the above equivalence between Cochran's and Bahadur's viewpoints is no longer true. A necessary and sufficient condition for the existence of Cochran-deficiency is proved. In most cases this condition does not hold and so Cochran-deficiency will rarely exist. When appropriate asymptotic expansions of the significance levels are available, an "approximate" Cochran-deficiency is calculated as a compensation. Conditions under which the said expansions are valid are also investigated.

As Bahadur-deficiency will, in general, be random, one may like to go to the considerations of taking some sort of average of Bahadur-deficiency. But since the computations involved in such considerations appear to be quite difficult, we proceed along a somewhat different route in Section 4, leading to a new interpretation of Cochran deficiency more in line with Bahadur's approach. In view of this interpretation we shall refer to Cochran deficiency as Bahadur-Cochran-deficiency (BCD).

1. Notations and preliminaries

Let $(X, \mathcal{B})$ be a measurable space; let $\{P_\theta;\ \theta \in \Theta\}$ be a family of probability distributions on X. Let $s = (x_1, x_2, \ldots)$ be an infinite sequence of independent observations on x. Let $T \equiv \{T_n(s) : n \geqslant 1\}$ be a *real-valued* statistic such that, for each n, $T_n(s)$ *depends on s only through* $(x_1, \ldots, x_n)$. In the next paragraph, a brief synopsis of Cochran's efficiency is given; for details consult Cochran (1952) and Bahadur (1967 and 1971).

Let Θ_0 be a proper subset of Θ. We are interested in testing $H_0 : \theta \in \Theta_0$ against $H_1 : \theta \in \Theta - \Theta_0$. For this purpose, we consider a test procedure which is based on a test statistic T and which regards *large values of $T_n(s)$ to be significant*; i.e., the critical region W_n of the test procedure is of the form

$$W_n = \{s : T_n(s) \geqslant k_n\}. \qquad \ldots \quad (1.1)$$

Fix a θ in $\Theta - \Theta_0$ and a β such that $0 < \beta < 1$. Choose $\{k_n : n \geqslant 1\}$ such that

$$P_\theta(W_n) \to \beta \qquad \ldots \quad (1.2)$$

as $n \to \infty$. Note that k_n will depend on β as well as on θ. Let $\alpha_n(\beta) \equiv \alpha_n(\theta, \beta)$ be the resulting size of the test, i.e., let

$$\alpha_n(\theta, \beta) = \sup\{P_{\theta_0}(W_n) : \theta_0 \in \Theta_0\}. \qquad \ldots \quad (1.3)$$

Cochran argued that the rate at which $\alpha_n(\theta, \beta)$ converges (to zero) is an indication of asymptotic efficiency of T against θ. Equivalently, one may proceed in the following way which is more suitable for our purpose : for each δ, $0 < \delta < 1$, let $M(\delta) \equiv M(\theta, \beta, \delta)$ be the *least integer* $m \geqslant 1$ such that $\alpha_n(\beta) < \delta$ for all $n \geqslant m$; and let $M(\delta) = \infty$ if no such m exists. Henceforth, we shall *assume that* $\alpha_n(\beta) \to 0$ *as* $n \to \infty$, which ensures that $M(\delta)$ is finite for all δ and that $\alpha_n(\beta) < 0$ for all sufficiently large n, which ensures that $M(\delta) \to \infty$ as $\delta \to 0$. The Cochran-efficiency of the test procedure, when it exists, is equal to the limit of $[2 \log (1/\delta)/M(\delta)]$ as $\delta \to 0$.

Consider now two testing procedures based on the statistics $T_1(s) = \{T_{1n}(s) : n \geqslant 1\}$ and $T_2(s) = \{T_{2n}(s) : n \geqslant 1\}$ for the problem of testing $H_0 = \theta \in \Theta_0$ against $H_1 : \theta \in \Theta - \Theta_0$. Fix a $\theta \in \Theta - \Theta_0$ and also a β such that $0 < \beta < 1$. We want to discriminate between these two procedures when θ obtains. Define $M_i(\delta)$, $i = 1, 2$, in the usual way. Clearly, the limit of $[M_2(\delta)/M_1(\delta)]$ as $\delta \to 0$ gives the Cochran-efficiency of T_1 relative to T_2 when θ obtains. When this efficiency is 1, $[M_1(\delta) - M_2(\delta)]/M_1(\delta) \to 0$. In typical cases, however, $[M_1(\delta) - M_2(\delta)]$ remains bounded as $\delta \to 0$, and so for the purpose of a more subtle distinction, one may use the limit of $[M_1(\delta) - M_2(\delta)]$ as $\delta \to 0$, whenever this limit exists.

Definition 1.1 : The lower (upper) Bahadur-Cochran-deficiency (BCD) at θ of the first testing procedure w.r.t. the second is

$$\underline{D}_C(\theta, \beta) = \liminf_{\delta \to 0} \, [M_1(\delta) - M_2(\delta)]$$

$$(\overline{D}_C(\theta, \beta) = \limsup_{\delta \to 0} \, [M_1(\delta) - M_2(\delta)]).$$

In case these two deficiencies are equal, we say that the BCD at θ exists and is equal to the common value.

Of course, $\underline{D}_C = \overline{D}_C = +\infty$ or $-\infty$ if $\lim \, [M_1(\delta)/M_2(\delta)]$ exists and $\neq 1$. The main use of deficiency is to discriminate tests for which $\lim \, [M_1(\delta)/M_2(\delta)]$ is 1. Note that although the relative Cochran-efficiency of two test procedures is usually free from β (see Proposition 11, Bahadur, 1967), their relative BCD need not be so.

Let $\Phi(x)$ stand for the distribution function of the standard normal distribution and $\phi(x)$ stand for its density function. For $0 < \beta < 1$, we define z_β by requiring that $\Phi(z_\beta) = 1 - \beta$. The following results will be needed in the sequel.

Lemma 1.1 : (See Chapter VII of Feller, 1968.) *If x is positive,*

$$1-\Phi(x) = \frac{\phi(x)}{x}\left(1-\frac{1}{x^2}+O(x^{-4})\right).$$

Lemma 1.2 : (See Chapter XV of Feller, 1966.) *Let $\{X_i\}$ be i.i.d. random variables with the common distribution $F(x)$ and with $E(X_1) = 0$, $E(X_1^2) = \sigma^2$. Let $F_n(x)$ be the normalized n-fold convolution of $F(x)$. If $F(x)$ is not a lattice distribution and if $m_3 \equiv E(X_1^3)$ is finite then one has*

$$F_n(x) = \Phi(x)+\frac{m_3}{6\sigma^3\sqrt{n}}(1-x^2)\,\phi(x)+\gamma_n(x)(n^{-\frac{1}{2}})$$

where $\gamma_n(x)\to 0$ uniformly in x.

For the next result, let $\{Y_n\}$ be a sequence of i.i.d. random vectors with values in $R^m (m \geqslant 1)$. Let $f_1, \ldots, f_k$ be real-valued Borel measurable functions on R^m. In the below, j stands for a positive integer $\geqslant 2$.

Assume

(A_{1j}) : $E|f_i(Y_1)|^j < +\infty, \quad i = 1, 2, \ldots, k.$

Write

$$Z_n = (f_1(Y_n), \ldots, f_k(Y_n))$$

$$\mu = EZ_1, \quad V = \text{dispersion matrix of } Z_1$$

Assume

(A_2) : V is nonsingular.

Let H be a *real-valued* function defined on some neighborhood $\mathcal{N}$ of μ.

Assume

$(A_{3,j})$: H has bounded continuous derivatives on $\mathcal{N}$ of all orders up to and including j.

Let

$$l = (D_1H, \ldots, D_kH)(\mu)$$

where D_i denotes differentiation w.r.t. the i-th coordinate.

Assume

(A_4) : $l \neq 0$.

Define H arbitrarily (but measurably) on all of R^k. We are interested in the asymptotic expansion of the distribution function of the statistic

$$W_n = \sqrt{n}(H(\overline{Z})-H(\mu)),$$

where

$$\bar{Z} = \frac{1}{n} \sum_{i=1}^{n} Z_i.$$

Lemma 1.3 : (See Theorem 2, Bhattacharya and Ghosh, 1978.) *Assume* (A_{1j}), (A_{3j}) *(for some integer* $j \geqslant 2$), (A_2) *and* (A_4) *hold. If* $j \geqslant 3$ *we furthermore assume that the distribution function* Q *of* $(Z_1-\mu)$ *satisfies Cramér's condition, namely,*

$$\limsup_{\|t\| \to \infty} \| \int_{R^k} \exp\{i < t, x >\} Q(dx) \| < 1,$$

then there exist polynomials q_r, $1 \leqslant r \leqslant j-2$, *whose coefficients depend only on the cumulants of* $(Z_1-\mu)$ *(of order j and less) and the derivatives of* H *at* μ *(of order* $j-1$ *and less) such that*

$$\sup_{u \in R^1} \left| Prob(W_n \leqslant u) - \int_{-\infty}^{u/\sigma} \phi(v) \left[1 + \sum_{r=1}^{j-2} n^{-r/2} q_r(v) \right] dv \right|$$

$$= o(n^{-(j-2)/2}),$$

where σ^2 *is the variance of* $< l, Z_1-\mu >$; *where* $< , >$ *stands for the usual Euclidean inner-product and* $\| \cdot \|$ *for the usual Euclidean norm.*

In Bhattacharya and Ghosh (1978), the methods of computing the polynomials q_r are also given. For details, the reader is referred to this paper.

2. Cochran's approach

In this section, we shall discuss our problem from the standpoint of Cochran's theory of efficiency. It is proved that for the existence of a finite BCD, the size functions of the test-procedures must be related in a very special way. Under appropriate asymptotic expansions of these size functions, bounds for the upper and lower BCD's are obtained; also methods of evaluating these deficiencies are discussed. All proofs are deferred to the appendix.

2.1. *Existence of a finite BCD.* Recall that θ is a fixed element in $\Theta - \Theta_0$ and β is a fixed real number s.t. $0 < \beta < 1$. α_{in} is the size of the i-th test procedure, $i = 1, 2$.

Obviously, $\lim_{\delta \to 0} (M_1(\delta)-M_2(\delta))$, if it exists, will be either an integer or, one of the two infinities; also

$$\lim_{\delta \to 0} (M_1(\delta)-M_2(\delta)) \text{ exists and is equal to an integer}$$
$$d \equiv d(0;\beta), \quad \ldots \quad (2.1.1)$$

iff for all sufficiently small δ, $M_1(\delta)-M_2(\delta)$ is identically equal to d. Thus at least in one case, e.g., when there exists an integer d such that

$$\alpha_{2,n} = \alpha_{1,\, n+d} \text{ for all sufficiently large } n, \quad \ldots \quad (2.1.2)$$

(2.1.1) holds true and hence the BCD is d. Our main Theorem 2.1.1 of this section states that the converse implication is also true.

Theorem 2.1.1 : *Suppose that for each $i = 1, 2$, α_{in} is a decreasing function of n for all sufficiently large n. Then* (2.1.1) *and* (2.1.2) *are equivalent.*

Remark 2.1.1 : It should be noted that the main reason why the existence of a finite BCD imposes a strong condition like (2.1.2) on the functions α_{in} is the discrete nature of the quantities $M_i(\delta)$. Unfortunately, any attempt to make the sizes continuous by taking resort to mixtures, as done by Hodges and Lehmann (1970), does not seem to work here.

2.2. *Bounds for the upper and lower BCD; the notion of approximate BCD:* As we know from the analysis of the previous section that BCD will exist rarely, we now turn to the problem of finding bounds for the upper and lower BCD's — of course, under suitable assumptions on the significance levels of the two test procedures.

We assume that

Each of $\{\alpha_{1n}\}$ and $\{\alpha_{2n}\}$ is a decreasing function of n for all sufficiently large n. ... (2.2.1)

For each $i = 1, 2$, there exists a function $\{\alpha_{i,x}\}$ defined for all $x > 0$ s.t.

(i) $\alpha_{i,x} = \alpha_{i,n}$ if x is the integer n; ... (2.2.2)

(ii) α_{ix} is a decreasing function of x, for all sufficiently large x;

(iii) if we define for each $n \geq 1$, a real number $m(n)$ s.t. $\alpha_{2,n} = \alpha_{1,m(n)}$, then the limit of $(m(n)-n)$ exists (it may be infinite).

We let $d(\beta) \equiv d(0, \beta)$ stand for this limit; in the rest of this paper, $d(\beta)$ will have this meaning only (unless otherwise is stated). Note that $d(\beta)$ need not be an integer.

Definition 2.2.1 : The approximate BCD of the first testing procedure w.r.t. the second one is $d(0, \beta)$.

Note that the approximate BCD depends on the particular extensions $\{\alpha_{i,x}\}$, $i = 1, 2$, we are using. We shall, however, suppress this dependence.

The above assumptions (2.2.1)-(2.2.2) are valid under appropriate asymptotic expansions of $\{\alpha_{i,n}\}$. (See Lemma 2.3.1.)

Note that, under (2.2.1), $M_i(\delta)$ is simply the first integer $m \geqslant 1$ such that $\alpha_{im} < \delta$. Consequently, if we let, for each $\delta(0 < \delta \leqslant 1)$, $M_{11}(\delta)$ and $M_{12}(\delta)$ satisfy

$$\alpha_{1,M_{11}(\delta)} = \alpha_{2,M_2(\delta)-1}, \quad \alpha_{1,M_{12}(\delta)} = \alpha_{2,M_2(\delta)}, \qquad \dots \quad (2.2.3)$$

then

$$M_{11}(\delta) \leqslant M_1(\delta) \leqslant M_{12}(\delta)+1. \qquad \dots \quad (2.2.4)$$

Because of the assumption (2.2.3), one has

$$\underline{D}c(0, \beta) \geqslant d(0, \beta)-1, \quad \overline{D}c(0, \beta) \geqslant d(0, \beta)+1;$$

or equivalently,

$$-[-d(0, \beta)]-1 \leqslant \underline{D}c(0, \beta) \leqslant \overline{D}c(0, \beta) \leqslant [d(0, \beta)]+1, \qquad \dots \quad (2.2.5)$$

where $[t]$ stands for the greatest integer $\leqslant t$.

Remark 2.2.1 : It follows that the BCD is $+\infty$ or $-\infty$ according as the approximate BCD is $+\infty$ or $-\infty$.

Remark 2.2.2 : If BCD exists and is finite then it must be equal to d and so d will be an integer; d may, however, be an integer even if BCD does not exist. (See Examples 2.4.1 and 2.4.2.)

If d is non-integral (and finite), say, $d = m+t$ $(0 < t < 1$, m is an integer), then the bounds given by (2.2.5) are sharp enough to conclude that

$$\underline{D}_C(\theta, \beta) = m \quad \text{and} \quad \overline{D}_C(\theta, \beta) = m+1.$$

In this case, it is possible to give the following interpretation of $\underline{D}_C$ and $\overline{D}_C$. For each $\delta (0 < \delta \leqslant 1)$, let $M_{13}(\delta)$ be the *smallest integer* k such that $\alpha_{2, M_2(\delta)} \geqslant \alpha_{1, k}$. Assumption (2.2.2) will then imply that

$$\lim_{\delta \to 0} (M_{13}(\delta) - M_2(\delta)) = m+1 = \overline{D}_C.$$

Similarly for $\underline{D}_C$.

Remark 2.2.3 : If d is an integer, the bounds given by (2,2.5) cannot immediately be used to find out the values of $\underline{D}_C$ and $\overline{D}_C$. However, when $m(n)-n$ always lies strictly on one side of d, one of these bounds can be improved upon as described in the next paragraph.

An improvement of the upper bound of $M_1(\delta)$ given by (2.2.4) is $M_1(\delta) \leqslant [M_{12}(\delta)]+1$; also from the assumption (2.2.2), $[M_{12}(\delta)] = M_2(\delta)+[d]$ or $M_2(\delta)+[d]-1$ for all sufficiently small δ.

Suppose now that d is an integer and that $m(n)-n-d < 0$ for all sufficiently large n. Then $[M_{12}(\delta)] \leqslant M_2(\delta)+d-1$ and hence $\overline{D}_C \leqslant d$. Clearly the BCD cannot exist. Thus one has from (2.2.5)

$$\underline{D}_C = d-1, \quad \overline{D}_C = d.$$

Similarly when d is an integer and $m(n)-n-d > 0$ for all sufficiently large n, one has $\underline{D}_C = d$, $\overline{D}_C = d+1$.

When asymptotic expansions of the size functions $\{\alpha_{in}\}$, $i = 1, 2$, are available, it may be possible to determine the exact rate of convergence (to zero) of $m(n)-n-d$ and hence to verify whether for all sufficiently large n, $m(n)-n-d$ is positive or negative. (See Section 2.4(A) for details.)

2.3. *Determination of the approximate BCD.* We assume throughout this section that the significance levels $\{\alpha_{in}\}$, $i = 1, 2$, of the test procedures admit of the following asymptotic expansions :

$$\log \alpha_{in}(\theta, \beta) = -na_i(\theta, \beta) + \sqrt{n}\, b_i(\theta, \beta) + c_i(\theta, \beta)\log n + d_i(\theta, \beta) + o(1) \quad (i = 1, 2) \qquad \ldots \quad (2.3.1)$$

where $a_i(\theta, \beta) > 0$, $i = 1, 2$.

In typical cases, $a_i(\theta, \beta)$ will be free from β; this will be the case if Bahadur-slopes of T_1 and T_2 exist; for a precise result, see Theorem 2 of Raghavachari (1970). Note that $M_1(\delta)/M_2(\delta) \rightarrow 1$ iff $a_1(\theta, \beta) = a_2(\theta, \beta) = a(\theta, \beta)$, say. Henceforth we shall assume that this is the case. For convenience, we shall suppress the dependence on θ, β of the quantities a_i, b_i, c_i, d_i.

The following lemma connects the two sets of assumptions made in the present and previous sections.

Lemma 2.3.1 : *Assume that* (2.3.1) *holds. Then* (2.2.1) *and* (2.2.2) *are valid. In fact,* (2.2.2) *holds in the following strong sense: There exist extensions* $\{\alpha_{ix}\}$, $i = 1, 2$, *which satisfy* (2.3.1) *for non-integral values* (>1) *of* x *as well.*

We shall work with those extensions $\{\alpha_{ix}\}$ which satisfy (2.3.1) for all real x in $(1, \infty)$. The next theorem gives the possible values of approximate BCD.

Theorem 2.3.1 : *Let* $\{\alpha_{i,n}\}$, $i = 1, 2$, *satisfy* (2.3.1). *Then one has*

(a) *if* $b_1 = b_2$ *and* $c_1 = c_2$, *then* $d = (d_1 - d_2)/a$;

(b) *if* $b_1 \neq b_2$, *then* d *is* $+\infty$ *or* $-\infty$, *according as* $b_1 > b_2$ *or* $b_1 < b_2$;

(c) *if* $b_1 = b_2$ *and* $c_1 \neq c_2$, *then* d *is* $+\infty$ *or* $-\infty$, *according as* $c_1 > c_2$ *or* $c_1 < c_2$.

Remark 2.3.1 : From the definition of approximate deficiency, it is apparent that the value of d will depend on the particular extensions $\{\alpha_{ix}\}$ of the sizes $\{\alpha_{in}\}$. However, from the above theorem it is clear that this dependence is slight and that the value of d will not depend on the extensions $\{\alpha_{ix}\}$ we are using, so long as they satisfy (2.3.1) for all real $x \in [1, \infty)$.

Remark 2.3.2 : In general, the value of the approximate BCD will depend both on θ and on β. However, in the examples discussed here, it will be free from β. In these examples, the situation (a) occurs and $d_1(\theta, \beta) - d_2(\theta, \beta)$ is free from both θ and β; as observed earlier, $a(\theta, \beta)$ will usually be free from β.

2.4. *Example:* In this section, we discuss two examples. As these examples will clarify different points of the next section 2.5, we prefer to go through the details. For simplicity, we determine the constants $k_n(\theta, \beta)$ (cf. (1.2)) such that

$$P_\theta(T_n > k_n(\theta, \beta)) \equiv \beta.$$

(A) In many examples, the size functions $\{\alpha_{in} : n \in I_+\}$, $i = 1, 2$ admit of extensions $\{\alpha_{ix}: x \in [1, \infty)\}$, $i = 1, 2$ which are decreasing functions of x and moreover, for large values of x, the following asymptotic expansions of these extensions are valid :

$$\log \alpha_{ix}(\theta, \beta) = -xa(\theta)+\sqrt{x}\, b(\theta, \beta)-\tfrac{1}{2}\log x+d_i(\theta, \beta)$$
$$+e(\theta, \beta)x^{-\frac{1}{2}}+o(x^{-\frac{1}{2}}) \quad (i = 1, 2) \qquad \dots \quad (2.4.1)$$

where $a(\theta) > 0$, $b(\theta, \beta) = 0$ iff $\beta = \frac{1}{2}$ and finally $d_1(\theta, \beta)-d_2(\theta, \beta)$ is always *nonzero*. Moreover, when $\beta = \frac{1}{2}$

$$\log \alpha_{ix}(\theta, \tfrac{1}{2}) = -xa(\theta)-\tfrac{1}{2}\log x+d_i(\theta, \tfrac{1}{2})+e(\theta, \tfrac{1}{2})x^{-\frac{1}{2}}$$
$$+f(\theta)x^{-1}+o(x^{-1}) \quad (i = 1, 2) \qquad \dots \quad (2.4.2)$$

From Theorem 2.3.1, the approximate BCD is $d(\theta, \beta) = (d_1(\theta, \beta)-d_2(\theta, \beta))/a(\theta)$. We want to compute $\underline{D}_C(\theta, \beta)$ and $\overline{D}_C(\theta, \beta)$, making use of the analysis made in Remark 2.2.3. For this let us first observe the following result about rate of convergence (to zero) of $t_n \equiv m(n)-n-d(\theta,\beta)$.

Lemma 2.4.1 : *Assume that the size functions* $\{\alpha_{i,n}\}$, $i = 1, 2$ *satisfy* (2.4.1) *and* (2.4.2). *Then* $\sqrt{n}\, t_n \to (d(\theta, \beta)\cdot b(\theta, \beta))/2a(\theta)$. *If* $\beta = \frac{1}{2}$, $nt_n \to d(\theta, \frac{1}{2})/2a(\theta)$.

We assume below, without loss of generality, that $d(\theta, \beta)$ is positive for all β. Three cases may arise.

Case I : Let β be such that $b(\theta, \beta) > 0$. In this case, $\sqrt{n}\, t_n$ converges to some positive number. Consequently, $m(n)-n-d(\theta, \beta)$ is positive for all sufficiently large n. It then follows from Remark 2.2.3 that the BCD does not exist and $\overline{D}_C(\theta, \beta) = d(\theta, \beta)+1$, while $\underline{D}_C(\theta, \beta) = d(\theta, \beta)$.

Case II : Let β be such that $b(\theta, \beta) < 0$. In this case, $m(n)-n-d(\theta, \beta)$ is negative for all sufficiently large n. So the BCD does not exist and $\overline{D}_C(\theta, \beta) = d(\theta, \beta)$ while $\underline{D}_C(\theta, \beta) = d(\theta, \beta)-1$.

Case III : Let β be such that $b(\theta, \beta) = 0$, i.e., let $\beta = \frac{1}{2}$. In this case nl_n converges to some positive number. So the conclusions of Case I are valid.

(B) The deficiency of one test procedure relative to another was defined in Section 1 for the *same* testing problem. In the following two examples, the two test procedures under comparison are for two *different* testing situations. More specifically, we want to compare one-sided test and two-sided test in what is apparently an one-sided testing problem. There are two good reasons for doing this. Let H_1 and H_2 correspond the two-sided and one-sided alternatives respectively ($H_2 \subset H_1$). Suppose now in a given problem the natural alternative is H_1 but there is some information (not entirely reliable) that the real alternative is H_2. In this case under the usual assumptions the likelihood ratio using H_2 is as Bahadur-efficient as the one using H_2 for all θ in H_1. So if Bahadur-efficiency were the only criterion, one should certainly ignore the information that H_2 is the real alternative. Our examples show the choice is not so clear if one also considers the deficiency. The second reason for considering these examples is a mathematical one; they are non-trivial and illustrate the various technical aspects of computing deficiency.

Example 1 : (*The Normal Distribution*) : Let H be the real line $(-\infty, +\infty)$, $H_0 = \{0\}$. For $\theta \epsilon H$, let P_θ stand for the normal distribution with mean θ and variance 1. Fix a θ s.t. $\theta > 0$.

For the testing problem H_0 : the population mean is zero against the alternative that it is non-zero, the critical region of the most powerful unbiased invariant test is given by $\{|\sqrt{n}\bar{X}_n| > k_{1,n}\}$ where $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$ and $k_{1,n}$ is such that $\beta = 1-\Phi(k_{1n}-\sqrt{n}\theta)+\Phi(-k_{1n}-\sqrt{n}\theta)$. Its power at θ is β and its size is $\alpha_{1n}(\theta, \beta) = 2(1-\Phi(k_{1n}))$. Bahadur's (as well as Cochran's) slope of $T_1 = \{|\sqrt{n}\bar{X}_n| : n \geqslant 1\}$ at θ is θ^2.

For the testing problem H_0 : the population mean is zero against the alternative that it is positive, the critical region of the most powerful test is given by $\{\sqrt{n}\,\bar{X}_n > k_{2,n}\}$ where k_{2n} is such that $\beta = 1-\Phi(k_{2n}-\sqrt{n}\theta)$, i.e., $k_{2n} = \sqrt{n}\theta + z_\beta$. Its power at θ is β and its size is $\alpha_{2n}(\theta, \beta) = 1-\Phi(k_{2n})$. Bahadur's (as well as Cochran's) efficiency at θ of $T_2 = \{\sqrt{n}\bar{X}_n : n \geqslant 1\}$ is θ^2.

Thus the two test procedures are equally efficient. At the level of deficiency, however, their performances are different. In this example, the situation described in (A) above holds. In fact, for non-integral values of x, one defines,

$$\alpha_{1x} = 2(1-\Phi(k_{1x})), \quad \alpha_{2x} = 1-\Phi(k_{2x}) \qquad \ldots \quad (2.4.3)$$

where k_{1x} is determined from

$$\beta = 1-\Phi(k_{1x}-\sqrt{x}\theta)+\Phi(-k_{1x}-\sqrt{x}\theta) \qquad \ldots \quad (2.4.4)$$

and $k_{2x} = \sqrt{x}\theta+z_\beta$. We claim that (2.4.1) and (2.4.2) hold with $a(\theta) = \theta^2/2$, $b(\theta,\beta) = -z_\beta\theta$, $d_1(\theta,\beta) = -(z_\beta^2+\log(2\pi\theta^2))/2+\log 2$, $d_2(\theta,\beta) = -(z_\beta^2+\log(2\pi\theta^2))/2$, $e(\theta,\beta) = -z_\beta\theta^{-1}$ and $f(\theta) = \theta^{-2}$. From (A), we can therefore conclude that BCD does not exist and that the approximate BCD $= 2\log 2/\theta^2$. Also, $\overline{D}_C(\theta,\beta) = 2\log 2/\theta^2+1$ if $\beta \geqslant \frac{1}{2}$, $= 2\log 2/\theta^2$ if $\beta < \frac{1}{2}$ while $\underline{D}_C(\theta,\beta) = 2\log 2/\theta^2$ if $\beta \geqslant \frac{1}{2}$, $= 2\log 2/\theta^2-1$ if $\beta < \frac{1}{2}$. Thus the approximate BCD does not depend on β. The upper and lower BCD do depend on β, but in a very weak sense.

We now proceed to the proof of our claim. That $\{\alpha_{2x}\}$ satisfies (2.4.1) and (2.4.2) is easy — one has to use Lemma 1.1. To show the same for $\{\alpha_{1x}\}$, it is sufficient to prove that as $x \to \infty$

$$\exp(x\theta^2)\,(\log\alpha_{1x}-\log\alpha_{2x}-\log 2) \to 0. \qquad \ldots \quad (2.4.5)$$

To this end, first observe that k_{2x} serves as a good asymptotic approximation to k_{1x}; for (2.4.4) implies that $(k_{1x}-\sqrt{x}\theta) \to z_\beta$, i.e., $(k_{1x}-k_{2x}) \to 0$. In fact, we have

$$\lim_{x\to\infty} \exp(2x\theta^2)(k_{1x}-k_{2x}) = 0; \qquad \ldots \quad (2.4.6)$$

for one has, from the definition of k_{2x} and (2.4.4),

$$1-\Phi(k_{1x}+\sqrt{x}\theta) = (k_{1x}-k_{2x})\Phi(\xi(x))$$

for some $\xi(x)$ satisfying $z_\beta < \xi(x) < k_{1x}-\sqrt{x}\theta$. By what has been proved, $\xi(x) \to z_\beta$ as $x \to \infty$. Thus

$$\exp(2x\theta^2)(k_{1x}-k_{2x}) \leqslant \frac{\exp(2x\theta^2)}{\Phi(\xi(x))}\cdot(1-\Phi(k_{2x}+\sqrt{x}\theta))$$

$$= O(x^{-\frac{1}{2}}), \quad \text{by Lemma 1.1.}$$

This completes the proof of (2.4.6). Now the l.h.s. of (2.4.5) is

$$\exp(x\theta^2)(k_{2x}-k_{1x})\Phi(\xi_1(x))/(1-\Phi(\xi_1(x)))$$

for some $\xi_1(x)$ satisfying $k_{2x} < \xi_1(x) < k_{1x}$; in particular, $\xi_1(x) = \sqrt{x}\theta(1+o(1))$. So (2.4.6) and Lemma 1.1 now complete the proof of (2.4.5).

Example 2 : (*The Student's Distribution*) : Here $\Theta = \{(\mu, \sigma) : -\infty < \mu < \infty,\ 0 < \sigma < \infty\}$, $\Theta_0 = \{0\} \times (0, \infty)$. For $\theta = (\mu, \sigma)$ in Θ, let P_θ stand for normal distribution with mean μ and variance σ^2. Fix a $\theta_1 = (\mu_1, \sigma_1)$ s.t. $\mu_1 > 0$, $\sigma_1 > 0$; put $\mu = \mu_1\sigma_1^{-1}$, $\theta = (\mu, 1)$ and $\theta_0 = (0, 1)$. Put $T_{2n} = \sqrt{n}\bar{X}_n/s_n$ and $T_{1n} = |T_{2n}|$ where $ns_n^2 = \sum_1^n (X_i - \bar{X}_n)^2$. Note that under θ_0, $\sqrt{n/(n-1)}T_{2n}$ is distributed as a Student's t-variable with $(n-1)$ d.f.

For the testing problem H_0 : the population mean is zero against the alternative that it is non-zero (the s.d. being unknown), the "best" test is based on the critical region $\{T_{1n} > k_{1n}\}$ where k_{1n} is such that $\beta = P_{\theta_1}(T_{1n} > k_{1n})$. Its power at θ_1 then is β and its size is $\alpha_{1n}(\theta_1, \beta) = 2P_{\theta_0}(T_{2n} > k_{1n})$. Bahadur's (as well as Cochran's) slope at θ_1 of T_1 is $\log(1+\mu^2)$.

For the testing problem H_0 : the population mean is zero against the alternative that it is positive (the s.d. being unknown), the "best" test is based on the critical region $\{T_{2n} > k_{2n}\}$ where k_{2n} is such that $\beta = P(T_{2n} > k_{2n})$. Its power at θ_0 is β and its size is $\alpha_{2n}(\theta_0, \beta) = P\ (T_{2n} > k_{2n})$. Bahadur's (as well as Cochran's) slope at θ_1 of T_2 is $\log(1+\mu^2)$.

Here too the two test procedures are equally efficient, though their deficiency is not zero. We shall show that $\{\alpha_{in} : n \in I_+\}$, $i = 1, 2$, satisfy (2.4.1) and (2.4.2) with $a(\theta_1) = \frac{1}{2}\log(1+\mu^2)$, $b(\theta_1, \beta) = -z_\beta \cdot \mu(1+\frac{1}{2}\mu^2)^{\frac{1}{2}}(1+\mu^2)^{-1}$ and $d_1(\theta_1, \beta) - d_2(\theta_1, \beta) = \log 2$. In this example, we cannot define the extensions $\{\alpha_{ix} : x \in (1, \infty)\}$, $i = 1, 2$, as we did in Example 1. One has to take linear averages of $\log \alpha_{in}$ and $\log \alpha_{i(n+1)}$; for details see the proof of Lemma 2.3.1.

From (A) we therefore conclude that BCD does not exist and that the approximate BCD is $d(\theta_1, \beta) = 2\log 2/\log(1+\mu^2)$. Also, $\overline{D}_C(\theta_1, \beta) = d(\theta_1, \beta)+1$ if $\beta \geqslant \frac{1}{2}$ $= d(\theta_1, \beta)$ if $\beta < \frac{1}{2}$ while $\underline{D}_C(\theta_1, \beta) = d(\theta_1, \beta)$ if $\beta \geqslant \frac{1}{2}$; $= d(\theta_1, \beta)-1$ if $\beta < \frac{1}{2}$. Thus in this example too, the approximate BCD is free from β. The upper and lower BCD do depend on β, but in a very weak sense.

The proofs of the different facts mentioned above run essentially along the same line as that of Example 1; however, some of the steps have to be justified in different ways. Lemma 1.1, e.g., is to be replaced by the following one, the proof of which depends on integration by parts.

Lemma 2.4.2 : *Let* $n \geqslant 5$ *and* $\alpha > 0$. *Put*

$$I_n(\alpha) = \int_\alpha^\infty (1+x^2)^{-n/2}\,dx$$

and

$$\gamma_n(\alpha) = (n-2)^{-1}\alpha^{-1}(1+\alpha^2)^{-(n-2)/2}.$$

Then one has

$$\gamma_n(\alpha)[1-(n-4)^{-1}(\alpha^{-2}+1)] \leqslant I_n(\alpha) \leqslant \gamma_n(\alpha)[1-(n-4)^{-1}(\alpha^{-2}+1)$$
$$-3(n-4)^{-1}(n-6)^{-1}(\alpha^{-2}+1)].$$

Using this lemma, one can verify

Lemma 2.4.3 : *Let* p_n *stand for* $P_{\theta_0}(T_{2n} > a\sqrt{n}+b+cn^{-\frac{1}{2}}+dn^{-1})$, $a > 0$. *Then*

$$\log p_n = -\frac{n}{2}\log(1+a^2)-\sqrt{n}\,ab(1+a^2)^{-1}-\tfrac{1}{2}\log n$$
$$+K_1(a,b,c)+K_2(a,b,c,d)n^{-\frac{1}{2}}+o(n^{-\frac{1}{2}}),$$

where K_1 *and* K_2 *are constants (free from* n*) which depend on* a, b, c, *etc., as indicated above.*

We also need to use the following result, which is a direct consequence of Lemma 1.3.

Lemma 2.4.4 : *Let* $W_n = T_{2n}-\sqrt{n}\mu$. *There exist two polynomials* P_1 *and* P_2 *whose coefficients are free from* n *such that*

$$\sup_{u\in R^1}\sqrt{n}\left|P_{\theta_1}(W_n > u\tilde{\sigma})-\int_u^\infty \phi(t)dt-\left(\frac{P_1(u)}{\sqrt{n}}+\frac{P_2(u)}{n}\right)\phi(u)\right| \to 0$$

where

$$\tilde{\sigma} = (1+\tfrac{1}{2}\mu^2)^{\frac{1}{2}}.$$

Define $G_n(u) = 1-\Phi(u)+(P_1(u)n^{-\frac{1}{2}}+P_2(u)n^{-1})\phi(u)$; determine the constants (free from n) d, d_1, d_2 such that $G_n(d+d_1n^{-\frac{1}{2}}+d_2n^{-1}) = \beta+o(n^{-1})$; one may verify that $d = z_\beta$ (we do not need the exact values of d_1 and d_2). Put $k'_{2n} = \sqrt{n}\mu+(d+d_1n^{-\frac{1}{2}}+d_2n^{-1})\tilde{\sigma}$. Then $k_{2n}-k'_{2n} = o(n^{-1})$—for a justification imitate the proof of Lemma 2.5.1. What we have achieved so far is simply an approximation k'_{2n} of k_{2n} which guarantees that

$$\log P_{\theta_0}(T_{2n} > k_{2n}) - \log P_{\theta_0}(T_{2n} > k'_{2n}) = o(n^{-1});$$

to verify this, one uses Lemma 2.4.2. An application of Lemma 2.4.3 then shows that $\{\alpha_{2n}\}$ satisfies (2.4.1) with the said values of $a(\theta)$ and $b(\theta, \beta)$.

Consider now the case of $\{\alpha_{1n}\}$. Here

$$\beta = P_{\theta_1}(W_n > k_{1n} - \sqrt{n}\mu) + P_{\theta_1}(W_n < -k_{1n} - \sqrt{n}\mu).$$

Because of Lemma 2.4.4, $P_{\theta_1}(W_n < -k_{1n} - \sqrt{n}\mu)$ is $o(n^{-1})$. (In fact it can be shown that it is $o(n^{-j})$ for each positive integer j.) This implies that $(k_{1n} - k'_{2n}) = o(n^{-1})$; and so

$$\log P_{\theta_0}(T_{2n} > k_{1n}) - \log P_{\theta_0}(T_{2n} > k'_{2n}) = o(n^{-1}).$$

Thus $\log \alpha_{1n} = \log 2 + \log \alpha_{2n} + o(n^{-1})$. This completes the proof of the fact that $\{\alpha_{1n}\}$ satisfies (2.4.1) and that $d_1(\theta, \beta) - d_2(\theta, \beta) = \log 2$.

The proof of the fact that $\{\alpha_{1n}\}$ and $\{\alpha_{2n}\}$ satisfy (2.4.2) should now be clear.

Remark 2.4.2 : Suppose we replace (1.2) by $\lim P_\theta(W_{in}) = \beta$ and put $R_{in}(\theta, \beta) = P_\theta(W_{in}) - \beta$, $(i = 1, 2)$; then in the above examples it can be shown that (a) if $R_{in}(\theta, \beta) = o(n^{-1})$ then the value of the approximate BCD remains unchanged; (b) if $R_{in}(\theta, \beta) = o(n^{-1})$ and $R_{in}(\theta, \frac{1}{2}) = o(n^{-3/2})$, BCD does not exist and the values of the upper and lower BCD remain unchanged.

2.5. *On the validity of* (2.3.1). Here we shall find conditions under which the asymptotic expansion of the form (2.3.1) of the significance level of a test procedure is valid. We motivate ourselves by considering a test procedure in which the critical region consists of large values of the sum of some sequence of i.i.d. random variables. We have the following general result in this direction; it is comparable with the results of Bahadur and Ranga Rao (1960).

Let $\{Y_n : n \geqslant 1\}$ be a sequence of i.i.d. r.v.s. with the m.g.f. $M(t)$. Put $T_n = n^{-\frac{1}{2}}\sum_{1}^{n} Y_i$. Let μ be a constant ($\neq 0$) and $\{q_n\}$ be a bounded sequence of real numbers. Define $p_n = \text{Prob}(T_n > \sqrt{n}\mu + q_n)$. Assume that the distribution of Y_1 and μ satisfy the following conditions : the distribution of Y_1 is non-lattice; if T stands for $\{t : M(t)$ is finite$\}$, then T is a nondegenerate interval; there exists a positive τ in the interior of T such that $\exp(-\mu\tau)M(\tau) = \inf\{\exp(-\mu t)M(t) : t \text{ in } T\} = \rho$ say.

Proposition 2.5.1 : Assume the above set-up. Then

$$\log p_n = n \log \rho - \sqrt{n}\alpha q_n - \frac{1}{2}\log n - (a + q_n^2/2) + o(1) \quad \ldots \quad (2.5.1)$$

where a, α, ρ are constants (free from n); $a > 0$, $\alpha > 0$, $0 < \rho < 1$.

Remark 2.5.1 : If we assume that the distribution of Y_1 satisfies Cramér's condition, we can get an asymptotic expansion of $\log p_n$ similar to the one given in Theorem 2 of Bahadur and Ranga Rao (1960).

Consider now the set-up of Section 1. Our main interest is to find an asymptotic expansion of $P_\theta(T_n > k_n)$ where k_n is to be determined from the condition (1.2).

Assume that the distributions of $\{T_n\}$ under P_θ and $\{P_{\theta_0} : \theta_0 \epsilon \Theta_0\}$ satisfy the following conditions :

There exist constants (free from n) $\mu \equiv \mu(\theta)$ and $\tilde{\sigma} \equiv \tilde{\sigma}(\theta) > 0$ and a polynomial q_θ such that if we let $F_n(x) = P_\theta(T_n - \sqrt{n}\mu \leqslant \tilde{\sigma}x)$, then

$$F_n(x) = \phi(x) + n^{-\frac{1}{2}} q_\theta(x)\phi(x) + o(n^{-\frac{1}{2}}), \text{ uniformly in } x. \quad \ldots \quad (2.5.2)$$

Whenever μ is a real number and $\{q_n\}$ is a bounded sequence of real numbers, (2.5.1) holds good with

$$p_n = \sup \{P^{\theta_0}(T_n > \sqrt{n}\mu + q_n) : \theta_0 \text{ in } \Theta_0\}. \quad \ldots \quad (2.5.3)$$

Lemma 2.5.1 : *Assume* (2.5.2). *Then there exists a constant* $d \equiv d(\theta, \beta)$ *such that* $k_n = \sqrt{n}\mu + z_\beta \tilde{\sigma} + n^{-\frac{1}{2}} d\tilde{\sigma} + o(n^{-\frac{1}{2}})$.

Theorem 2.5.1 : *Assume* (2.5.2) *and* (2.5.3). *Then*

$$\log \alpha_n = n \log \rho - \sqrt{n}\, \alpha z_\beta \tilde{\sigma} - \tfrac{1}{2} \log n$$
$$+ (a - \alpha\tilde{\sigma}d - \tfrac{1}{2} z_\beta \tilde{\sigma}^2) + o(1).$$

Remark 2.5.2 : It is well known that $\{n^{-\frac{1}{2}} \sum_1^n Y_i : n \geqslant 1\}$ satisfies (2.5.2) where $\{Y_i : i \geqslant 1\}$ is a sequence of i.i.d. r.v.s. with a finite third moment. The main result of Bhattacharya and Ghosh (1978) indicates that (2.5.2) is satisfied for a large collection of statistics.

3. Bahadur's approach

In this section, we shall consider two possible ways measuring deficiency from the standpoint of Bahadur's theory of efficiency. It is shown by means of an example that Bahadur-deficiency in the strong sense need not exist even though Bahadur-Cochran-deficiency exists. A new interpretation of the latter is suggested in Section 3.2.

3.1. *Bahadur's approach.* Assume the set-up of Section 1. For each real t, let $F_{in}(t) = \sup\{P_{\theta_0}(T_{1n} > t) : \theta_0 \epsilon \Theta_0\}$ and define $L_{in}(s) = F_{in}(T_{in}(s))$. For each $\delta\,(0 < \delta \leqslant 1)$ and for each s, let $N_i(\delta, s) \equiv N_i$ be the least integer $m \geqslant 1$ such that $L_{in}(s) < \delta$ for all $n \geqslant m$; and let $N_i\,(\delta, s) = \infty$ if no such m exists.

Definition 3.1.1 : The random lower (upper) Bahadur-deficiency at θ of the first testing procedure w.r.t. the second is

$$\underline{D}_B(\theta; \beta) = (\text{a.s. } P_\theta) \lim\ \inf\ (N_1(\delta, s) - N_2(\delta, s))$$

$$(\overline{D}_B(\theta; \beta) = (\text{a.s. } P_\theta) \lim\ \sup\ (N_1(\delta, s) - N_2(\delta, s)).$$

In case the above two deficiencies are equal, we say that the Bahadur-deficiency exists and is equal to the common value.

As in the case of Cochran-deficiency, the main use of studying these random deficiencies is to discriminate tests with the same Bahadur-efficiency, i.e., tests for which the (a.s. P_θ) limit of $(N_2(\delta, s)/N_1(\delta, s))$ is 1.

In this approach, the main source of difficulty is that the quantities $\sup\{L_{in}\,(s) : n \geqslant m\}$, $m \geqslant 1$ are difficult to expand — any possible expansion would seem to depend on the particular sample sequence considered.

Example 3 : (*The Uniform Distribution*) : Let θ be such that $0 < \theta < 1$; let $f_1(x)$ and $f_2(x)$ be respectively the densities of the uniform distributions over $[0, \theta]$ and $[0, 1]$. Consider the problem of testing $H_0 : f = f_2$ vs. $H_1 : f = f_1$ on the basis of the following two statistics, $T_{1n} = x_{(n)}$, $T_{2n} = y_{(n)}$ where $x_{(n)} = \max(x_2, x_4, \ldots, x_{2n})$ and $y_{(n)} = \max\,(x_1, x_3, \ldots, x_{2n-1})$, $(n \geqslant 2)$. [Here we deviate slightly from our basic convention that T_{in} must be a function of the first n observations; but we hope this will not lead to any confusion.] We choose the constants $k_{in}(\theta, \beta)$ such that $P_\theta(T_{in} > (\theta,\ \beta)) \equiv \beta$, $i = 1, 2$ (cf. (1.2)). Then $\alpha_{1n} = \beta\theta^{n-1}$ while $\alpha_{2n} = \beta\theta^n$ so that the BCD exists and is 1 for all β.

We are going to show that the (a.s. or stochastic) limit of $(N_1(\delta, s) - N_2\,(\delta, s))$, if it exists, cannot be degenerate. Note that $L_{1n}(s) = (x_{(n)})^{n-1}$ and $L_{2n}(s) = (y_{(n)})^n$. Clearly, $P_\theta(N_1(s, \delta) = m) = P_\theta(N_2(s, \delta) = m+1)$, for all $m \geqslant 2$. The lemma below gives the exact distribution of $N_2(s, \delta)$ under P_θ.

Lemma 3.1.1 : *Let* $p(\delta) \equiv p(\theta, \delta)$ *be the integer such that*

$$\frac{\log \delta}{\log \theta} \leqslant p(\delta) < \frac{\log \delta}{\log \theta}+1. \qquad \ldots \quad (3.1.1)$$

Then the distribution function of $N_2(\delta, s)$ *has the following expression*

$$P_\theta(N_2(s, \delta) \leqslant m) = \begin{cases} (\delta/\theta^{p-1}) \exp\left[\log \delta \left(\sum_{j=m+1}^{p-1} j^{-1} \right)\right] & \text{if} \quad m \leqslant p-2 \\ \partial/\theta^{p-1} & \text{if} \quad m = p-1 \\ 1 & \text{if} \quad m = p \end{cases}$$

$(p \equiv p(\delta))$

$$\ldots \quad (3.1.2)$$

The proof is straightforward.

The next lemma studies the weak convergence under P_θ of $p(\delta)-N_2(s, \delta)$ as $\delta \to 0$.

Lemma 3.1.2 : *For each* c, $0 \leqslant c \leqslant 1$, *let* X_c *be a r.v. such that* $P_\theta(X_c = 0) = 1-\theta^c$ *and* $P_\theta(X_c = i) = (1-\theta) \cdot \theta^{c+i-1}$, $i \geqslant 1$. *Let* $e(\delta)$ *be the excess over* $(p(\delta)-1)$ *of* $\log \delta/\log \theta$, $(0 < e(\delta) \leqslant 1)$. *Then*

(a) *if* $e(\delta_n) \to c$ *and* $\delta_n \to 0$, *then* $p(\delta_n)-N_2(\delta_n; s)$ *converges weakly to* X_c *under* P_θ.

(b) *if* $p(\delta_n)-N_2(\delta_n; s)$ *converges weakly to* X *under* P_θ *and* $\delta_n \to 0$, *then* $\{e(\delta_n)\}$ *is a convergent sequence; moreover,* X *can be taken to be* X_c *where* c *is the limit of* $e(\delta_n)$.

Proof: (a) By definition of $e(\delta)$, $\delta = \theta^{(p-1)+e(\delta)}$ so that $\delta_n/\theta^{p(\delta_n)-1} = \theta^{e(\delta_n)} \to \theta^c$. (3.1.1) implies that

$$\theta^{-1} \cdot \delta^{(p(\delta)-1)-1} < 1 \leqslant \theta^{-1} \delta^{(p(\delta)-1)-1}.$$

Thus for each $k \geqslant 1$, $\delta_n^{(p(\delta_n)-k)^{-1}} \to \theta$. It then follows from Lemma 3.1.1 that $P_\theta(p(\delta_n)-N_2(\delta_n, s) = m) \to P_\theta(X_c = m)$ $m \geqslant 0$, which completes the proof of (a).

(b) Because of (a), every convergent sub-sequence of $\{e(\delta_n)\}$, which is a bounded sequence, converges to the same real number.

It follows from the above lemma that the (a.s. P_θ or the stochastic) limit of $(N_1(\delta, s)-N_2(\delta, s))$ cannot be degenerate; to see this, one need only to note that $N_1(\delta, s)$ and $N_2(\delta, s)$ are independent and then use Theorem 3.2, Chapter VIII of Feller (1968). So in this case we cannot hope to get a single numerical value of deficiency from Bahadur's point of view.

3.2. *Another interpretation of BCD.* In his papers (1967) and (1971) Bahadur discussed asymptotic stochastic efficiencies in terms of almost sure convergence only. However, it is possible to define asymptotic stochastic efficiencies in terms of convergence in probability. Although exact slopes are easier to interpret in the almost sure convergence case, the convergence in probability definition is not only easier to use but perhaps more basic and stable. For these reasons, we now look at the convergence-in-probability case.

The following result is due to Raghavachari (1970) (see his Theorem 2). The set-up is the same as that given in Section 1.

Lemma 3.2.1 : *For all* $\beta (0 < \beta < 1)$,

$$\lim_{n \to \infty} n^{-1} \log \alpha_n(\theta, \beta) = -c(\theta), \quad 0 < c(\theta) < \infty$$

iff

$$n^{-1} \log L_n(s) \xrightarrow{P_\theta} -c(\theta).$$

This fact leads to the following definition.

Definition 3.2.1 : Fix a $\theta \notin \Theta_0$, an ε with $0 < \varepsilon < 1$ and a δ with $0 < \delta < 1$. Then $V(\varepsilon, \delta) \equiv V(\theta, \varepsilon, \delta)$ is the smallest integer $m \geqslant 1$ such that whenever $n \geqslant m$, $P_\theta(L_n(s) < \delta) > 1 - \varepsilon$, and let $V(\varepsilon, \delta) = +\infty$ if no such m exists.

The next lemma gives the asymptotic behavior of $V(\varepsilon, \delta)$ as $\delta \to 0$.

Lemma 3.2.2 : *Assume that*

$$n^{-1} \log L_n(s) \xrightarrow{P_\theta} -c(\theta) \quad 0 < c(\theta) < \infty.$$

Then for each ε, $0 < \varepsilon < 1$, *we have*

$$\frac{\log \delta}{V(\varepsilon, \delta)} \to -c(\theta)$$

as $\delta \to 0$.

Proof: Fix θ and ε. We abbreviate $V(\varepsilon, \delta)$ and $c(\theta)$ as $V(\delta)$ and c respectively.

Since c is finite, it follows that $V(\delta)$ tends to ∞ as δ tends to 0. Clearly then the proof of the lemma will be complete when we have established the following two inequalities.

$$\liminf_{\delta \to 0} \left[-\frac{\log \delta}{V(\delta)-1} \right] \geqslant c \qquad \dots \quad (3.2.1)$$

$$\limsup_{\delta \to 0} \left[-\frac{\log \delta}{V(\delta)} \right] \leqslant c. \qquad \dots \quad (3.2.2)$$

Suppose (3.2.1) is false. Then there exists η and a sequence $\{\delta_n\}$ s.t. $0 < \eta < c$, $\delta_n \to 0$ as $n \to \infty$ and

$$-\frac{\log \delta_n}{V(\delta_n)} < c-\eta.$$

From the definition of $V(\delta)$, we have

$$P_\theta \left\{ L_{V(\delta_n)-1}(s) < \delta_n \right\} \leqslant 1-\varepsilon$$

i.e.,

$$P_\theta \left\{ -\frac{\log \delta_n}{V(\delta_n)-1} \geqslant -\frac{\log L_{V(\delta_n)-1}(s)}{V(\delta_n)-1} \right\} \geqslant \varepsilon.$$

So certainly

$$P_\theta \left\{ c-\eta \geqslant -\frac{\log L_{V(\delta_n)-1}(s)}{V(\delta_n)-1} \right\} \geqslant \varepsilon.$$

Letting $n \to \infty$, we get a contradiction.

The proof of (3.2.2) is similar.

Lemma 3.2.2 suggests the following measure of deficiency; consider the set-up of Section 1 and define $V_1(\varepsilon, \delta)$ and $V_2(\varepsilon, \delta)$ similarly using $L_{1n}(s)$, $L_{2n}(s)$ for $L_n(s)$.

Definition 3.2.2: Fix a $\theta \notin \Theta_0$, an ε with $0 < \varepsilon < 1$. Then the lower (upper) deficiency of the first testing procedure w.r.t. the second at θ is

$$\liminf_{\delta \to 0} \left(V_1(\varepsilon, \delta) - V_2(\varepsilon, \delta)\right)$$

$$\left(\limsup_{\delta \to 0} \left(V_1(\varepsilon, \delta) - V_2(\varepsilon, \delta)\right)\right).$$

Let $F_{in}(t)$ be a strictly decreasing continuous function of t, $i = 1, 2$. For each $\theta \notin \Theta_0$, we make the same assumption about $P_\theta\{T_{in} > t\}$.

For each δ $(0 < \delta < 1)$, let $t_{in}(\delta) = F_{in}^{-1}(\delta)$. Consider the sequence of tests $\phi_{in}(\delta)$:

$$\text{Reject } H_0 \text{ iff } T_{in} > t_{in}(\delta).$$

Then the error of first kind for this test is δ. We denote its power by $\overline{\beta_{in}}$.

Fix $\theta \notin \Theta_0$. For each $0 < \beta < 1$, define the test $\psi_{in}(\beta)$: Reject H_0 iff $T_{in} > c_{in}(\beta)$ where $c_{in}(\beta)$ is such that $P_\theta\{T_{in} > c_{in}(\beta)\} = \beta$. Let its error of first kind be denoted by α_{in}.

Using the tests $\psi_{in}(\beta)$ define $M_i(\beta, \delta) \equiv M_i(\theta, \beta, \delta)$ as in Section 1. Then

$$M_i(\beta, \delta) = V_i(1-\beta, \delta). \qquad \dots \quad (3.2.1)$$

To see this, note that if $n \geqslant V_i(1-\beta, \delta)$ then by definition of V_i, the tests ϕ_{in} have error of first kind $= \delta$ and power (at θ) $> \beta$. Hence for $n \geqslant V_i(1-\beta, \delta)$ the tests ψ_{in} which have power $= \beta$, must have error of first kind $\alpha_{in} < \delta$. By definition of $M_i(\beta, \delta)$ this means $M_i(\beta, \delta) \leqslant V_i(1-\beta, \delta)$. Similarly the reverse inequality can be proved.

Thus BCD, upper, lower or approximate, agrees with the corresponding notion as defined here.

Appendix

Proofs of the results given in section 2

Proof of Theorem 2.1.1 : Let $\{\alpha_{in}\}$ be decreasing function of n if $n \geqslant m$ and let (2.1.1) hold. Then there exists $\delta_1 > 0$ s.t. $M_1(\delta) = d + M_2(\delta)$ if $\delta < \delta_1$. Assume that (2.1.2) does not hold, i.e., that $\alpha_{2,n_i} \neq \alpha_{1,n_i+d}$ where $n_1 < n_2 < n_3 < \dots$. Choose and fix a n_i s.t. $n_i > m$, $\alpha_{1,n_i+d} < \delta_1$ and $\alpha_{2,n_i} < \delta_1$. We may assume without any loss of generality that $\alpha_{2,n_i} < \alpha_{1,n_i+d}$. Then if $\delta = \alpha_{1,n_i+d}$, $M_1(\delta) > n_i + d$ and $M_1(\delta) \leqslant n_i$. Contradiction.

Proof of Lemma 2.3.1 : Assume (2.3.1). (2.2.1) is then immediate. For (2.2.2) we define $\{\alpha_{ix}\}$ for non-integral x as follows : let $n < x < n+1$; then $x = n\lambda + (n+1)\mu$ for some λ and μ such that $0 < \lambda < 1$, $\lambda + \mu = 1$. Define α_{ix} by

$$\log \alpha_{ix} = \lambda \log \alpha_{in} + \mu \log \alpha_{i(n+1)} \quad i = 1, 2.$$

It is obvious that the conditions (i) and (ii) of (2.2.2) hold. The proof of the fact the condition (iii) of (2.2.2) hold is included in that of Theorem 2.3.1.

Since $n = x-\mu$ and $n+1 = x+\lambda$, one has

$$\lambda \sqrt{n} + \mu \sqrt{n+1} = \sqrt{x} + o(x^{-1}).$$

and

$$\lambda \log n + \mu \log (n+1) = \log x + o(x^{-1}).$$

(2.3.1.) then shows that the last part of the lemma holds.

Proof of Theorem 2.3.1 : First note that since $\log \alpha_{i,n}/n \to -a$ for each $i = 1, 2$, $m(n)/n \to 1$.

(a) From the definition of $m(n)$, and the last part of Lemma 2.3.1 one has

i.e., $$(m(n)-n)\left[a-\frac{b_1}{\sqrt{n}+\sqrt{m}}\right] - c_1 \log \frac{m(n)}{n} \to (d_1-d_2)$$

$$m(n)-n \to (d_1-d_2)/a.$$

(b) Here we have,

$$\frac{m(n)-n}{\sqrt{n}}\left[a-\frac{b_1}{\sqrt{n}+\sqrt{m}}\right]-(b_1-b_2)+c_2 \cdot \frac{\log m(n)}{\sqrt{n}} - c_1 \cdot \frac{\log n}{\sqrt{n}} \to 0.$$

So $(m(n)-n)/\sqrt{n} \to (b_1-b_2)/a$. From this, (b) is immediate.

(c) The proof of (c) is similar to that of (b).

Proof of Lemma 2.4.1 : Similar to that of Theorem 2.3.1.

Proof of Proposition 2.5.1 : We shall follow Bahadur and Ranga Rao (1970). Let H_n be the distribution function of the standardized n-fold convolution, and v the s.d., of the conjugate distribution of $Y_1-\mu$. Put $\alpha = v\tau$. Proceeding exactly in the same way of Lemma 2 of Bahadur and Ranga Rao (1970), we have $p_n = \rho^n I_n$

where

$$I_n = \int_{e_n}^{\infty} e^{-\sqrt{n}\alpha x} \, dH_n(x).$$

Because of Lemma 1.2,

$$H_n(x) = \Phi(x) + \frac{K(1-x^2)}{\sqrt{n}}\Phi(x) + \gamma_n(x) n^{-\frac{1}{2}}$$

where $\gamma_n(x) \to 0$ uniformly in x and K is a constant.

(I) The contribution of $\phi(x)$ to I_n is

$$\int_{q_n}^{\infty} e^{-\sqrt{n}\,\alpha x}\, d\phi(x) = \frac{\exp(-\sqrt{n}\ \alpha q_n - q_n^2/2)}{\sqrt{2\pi}\sqrt{n}\alpha}\left\{1 - \frac{q_n}{\sqrt{n}\,\alpha} + \frac{q_n^2-1}{n\alpha^2} + o(n^{-1})\right\}$$

(use Lemma 1.1).

(II) The contribution of $K(1-x^2)\phi(x)/\sqrt{n}$ to I_n is

$$\frac{K}{\sqrt{\ }}\int_{q_n}^{\infty} e^{-\sqrt{n}\,\alpha x}(x^3-3x)\,\phi(x)\,dx$$

$$= \frac{K\exp(\frac{1}{2}n\alpha^2)}{\sqrt{n}}\int_{q_n+\sqrt{n}\,\alpha}^{\infty}\{(y-\sqrt{n}\,\alpha)^3 - 3(y-\sqrt{n}\,\alpha)\}\,\phi(y)\,dy$$

$$= \frac{\exp(-\sqrt{n}\alpha q_n - q_n^2(2))}{\sqrt{2\pi}\sqrt{n}\,\alpha}\cdot o(1).$$

(III) Let $\varepsilon > 0$. As $\sup_x |\gamma_n(x)| < \varepsilon/2$ for all sufficiently large n,

$$n^{-\frac{1}{2}}\left|\int_{q_n}^{\infty}\exp(-\sqrt{\ }\ \alpha x)\,d\gamma_n(x)\right| \leqslant \alpha\int_{q_n}^{\infty}\exp(-\sqrt{n}\,\alpha x)\,|\gamma_n(x)-\gamma_n(q_n)|\,dx$$

$$\leqslant \varepsilon n^{-\frac{1}{2}}\exp(-\sqrt{n}\,\alpha q_n).$$

Thus the contribution of $n^{-\frac{1}{2}}\gamma_n(x)$ to I_n is

$$n^{-\frac{1}{2}}\exp(-\sqrt{n}\,\alpha q_n - q_n^2/2)\cdot o(1).$$

From (I), (II) and (III), (2.5.1) follows.

Proof of Lemma 2.5.1 : Because of the uniformity condition involved in (2.5.2),

$$F_n(z_\beta + dn^{-\frac{1}{2}}) = \Phi(z_\beta) + n^{-\frac{1}{2}}(d + q_\theta(z_\beta))\phi(z_\beta) + o(n^{-\frac{1}{2}}).$$

Thus if we let $d = -q_\theta(z_\beta)$, then

$$F_n(z_\beta + dn^{-\frac{1}{2}}) = (1-\beta) + o(n^{-\frac{1}{2}}).$$

Put

$$G_n(x) = \Phi(x) + n^{-\frac{1}{2}}q_\theta(x)\phi(x),$$

$$\varepsilon_{1n} = \sup_x |F_n(x) - G_n(x)|$$

and

$$\varepsilon_{2n} = F_n(z_\beta + dn^{-\frac{1}{2}}) - (1-\beta).$$

Choose $\eta_n \to 0 (\eta_n \neq 0)$ such that $n^{\frac{1}{2}} \varepsilon_{in} = o(\eta_n)$ for each $i = 1, 2$. Then

$$\frac{\sqrt{n}}{|\eta_n|} \{F_n(z_\beta + (d+\eta_n)n^{-\frac{1}{2}}) - (1-\beta)\}$$

$$= \frac{\sqrt{n}}{|\eta_n|} \{G_n(z_\beta + (d+\eta_n)n^{-\frac{1}{2}}) - G_n(z_\beta + dn^{-\frac{1}{2}})\} + o(1)$$

$$= \frac{\sqrt{n}}{|\eta_n|} \cdot \frac{n}{\sqrt{n}} \cdot G_n'(\xi_n) + o(1)$$

for some ξ_n lying between $z_\beta + \frac{d}{\sqrt{n}}$ and $z_\beta + \frac{d}{\sqrt{n}} + \frac{\eta_n}{\sqrt{n}}$; in particular, $\xi_n \to z_\beta$. As $\{G_n'(u)\}$ is bounded away from zero in a neighborhood of z_β, one gets

$$\frac{\sqrt{n}}{|\eta_n|} \{F_n(z_\beta + (d+|\eta_n|)n^{-\frac{1}{2}} - F_n\left(\frac{k_n - \sqrt{n}\mu}{\bar{\sigma}}\right)\} > 0$$

for all sufficiently large n.

This implies that $(k_n - \sqrt{n}\,\mu)/\bar{\sigma} < z_\beta + dn^{-\frac{1}{2}} + |\eta_n| n^{-\frac{1}{2}}$ for all sufficiently large n. Similarly, $(k_n - \sqrt{n}\,\mu)/\bar{\sigma} > z_\beta + dn^{-\frac{1}{2}} - |\eta_n| n^{-\frac{1}{2}}$ for all sufficiently large n. As $\eta_n \to 0$, the proof of Lemma 2.5.1 is complete.

Proof of Theorem 2.5.1 : Let

$$k_n^* = \sqrt{n}\,\mu + z_\beta \bar{\sigma} + d\bar{\sigma} n^{-\frac{1}{2}}$$

and

$$\alpha_n^* = \sup \{P_{\theta_0}(T_n < k_n^*) : \theta_0 \in \Theta_0\}.$$

Because of (2.5.3),

$$\log \alpha_n^* = n \log \rho - \sqrt{n}\,\alpha z_\beta \bar{\sigma} - \frac{1}{2} \log n + \left(a - \alpha\bar{\sigma}d - \frac{1}{2} z_\beta^2 \bar{\sigma}^2\right) + o(1).$$

It suffices to show that

$$\log \alpha_n - \log \alpha_n^* = o(1). \qquad \ldots \quad (A.1)$$

Put $q_n = (k_n - \sqrt{n}\mu)/\bar{\sigma}$ and $q_n^* = (k_n^* - \sqrt{n}\mu)/\bar{\sigma}$. Then both the sequences $\{q_n\}$ and $\{q_n^*\}$ are bounded. By (2.5.3), one gets

$$\log \alpha_n - \log \alpha_n^*$$

$$= \sqrt{n}\,\alpha(q_n^* - q_n) + \frac{1}{2}(q_n^{*2} - q_n^*) + o(1)$$

$$= \sqrt{n}\,\alpha\frac{(k_n^* - k_n)}{\hat{\sigma}} + \frac{1}{2}\cdot(q_n + q_n^*)\,\frac{(k_n^* - k_n)}{\hat{\sigma}} + o(1)$$

$$= o(1) \text{ by Lemma 2.5.1.}$$

This completes the proof of (A.1).

Acknowledgement. The authors wish to thank Professors R. R. Bahadur and K. K. Roy for their interesting comments and invaluable help.

References

Bahadur, R. R. (1967): Rates of convergence of estimates and test statistics. *Ann. Math. Stat.* 38, 303-304.

——— (1971): Some limit theorems in statistics. *SIAM*, Philadelphia.

Bahadur, R. R. and Rao, R. Ranga (1960): On deviations of the sample mean. *Ann. Math. Stat.*, 31, 1015-1027.

Bhattacharya, R. N. and Ghosh, J. K. (1978): On the validity of the formal Edgeworth expansion. *Ann. Stat.*, 6, 434-451.

Cochran W. G. (1952): The χ^2-goodness of fit test. *Ann. Math. Stat.*, 23, 493-507.

Feller, W. (1966): *An Introduction to Probability Theory and its Applications*, Vol. I, John Wiley and Sons Inc.

——— (1968): *An Introduction to Probability Theory and its Applications*, Vol. II, John Wiley and Sons Inc.

Gnedenko, B. V. and Kolmogorov, A. N. (1954): Limit theorems for sums of independent random variables. English translation by K. L. Chung. Addison-Wesley, Reading, Mass.

Hodges, J. L., Jr. and Lehmann, E. L. (1970): Deficiency. *Ann. Math. Stat.*, 41, 783-801.

Killen, T. J., Hettmansperger, T. P. and Sievers, G. L. (1972): An elementary theorem on the probability of large deviations. *Ann. Math. Stat.*, 43, 181-192.

Raghavachari, M. (1970): On a theorem on the rate of convergence of test statistics. *Ann. Math. Stat.*, 41, 1695-1699.

Paper received : August, 1977.
Revised : May, 1978.