Transnational Journal of Mathematical Analysis and Applications
Vol. 8, Issue 1, 2020, Pages 37-74
ISSN 2347-9086
Published Online on March 06, 2021
© 2020 Jyoti Academic Press
http://jyotiacademicpress.org

SECOND ORDER ASYMPTOTIC OPTIMALITY IN TESTING PROBLEM WITH ONE DIMENTIONAL PARAMETER

V. E. BENING

Faculty of Computational Mathematics and Cybernetics
Moscow State University
Vorobyovy Gory
Moscow 119899
Russia
e-mail: bening@yandex.ru

Abstract

In this paper, a brief survey of the asymptotic theory of hypotheses testing is presented and some of the author's recent results are given. The survey is not intended to be complete; it contains mainly results related to the author's interests (for detailed proofs, see [51]). A detailed review of this field can be found in Pfanzagl [45], Pfanzagl and Wefelmeyer [46, 47], Chibisov [23], and Götze and Milbrodt [25].

We consider the asymptotic approach with the probabilities of errors of first and second kind being bounded away from zero and therefore we study the power of tests against local alternatives. Special attention is paid to asymptotically efficient tests for testing a simple hypothesis concerning a univariate parameter. We shall consider only "regular" families for which local alternatives approach the hypothesis at a rate of $n^{-1/2}$.

2020 Mathematics Subject Classification: Primary 60F05; Secondary 62E20, 91B30, 91B70.
Keywords and phrases: hypotheses testing, asymptotic expansions, deficiency, efficiency, power loss, test, level, errors of first and second kind.
This work is supported by the Moscow Center for Fundamental and Applied Mathematics.
Received December 15, 2020; Revised February 17, 2021

1. First-Order Asymptotic Theory

Let $\{\mathbf{P}_\theta, \theta \in \Theta \subset \mathbf{R}^1\}$ be a family of probability measures on a measurable space $(\mathcal{X}, \mathcal{A})$ having densities $p(x, \theta)$ with respect to a σ-finite measure ν. Assuming without loss of generality that $\Theta \subset \mathbf{R}^1$ contains an interval $[0, \varepsilon], \varepsilon > 0$. Suppose we have independent and identically distributed $\mathcal{X}$-valued observations $(X_1, \cdots, X_n)$ distributed according to $\{\mathbf{P}_\theta, \theta \in \Theta \subset \mathbf{R}^1\}$. Our problem is to test the hypothesis

$$\mathbf{H}_0 : \theta = 0 \text{ against } \mathbf{H}_1 : \theta > 0.$$

Note that alternative hypothesis is composite. This is a most general one-sided hypothesis that we need to consider. The test $\mathbf{H}_0 : \theta = \theta_0, \theta_0$ specified against $\mathbf{H}_1 : \theta > \theta_0$, can be reduced to the above case by considering the family $\{\mathbf{P}_{\theta_0+\theta}, \theta \in \Theta \subset \mathbf{R}^1\}$. By a test (or critical function) for the sample size n we mean a measurable map

$$\Psi_n : \mathcal{X}^n \to [0, 1].$$

If Ψ_n assumes the values 0 and 1 only, the set

$$\{x \in \mathcal{X}^n : \Psi_n(x) = 1\}$$

will be called critical region.

We denote $\mathbf{P}_{n,0}$ and $\mathbf{P}_{n,\theta}$ the joint distributions of $(X_1, \cdots, X_n)$ under $\mathbf{H}_0$ and $\mathbf{H}_1$, respectively. The respective expectations will be denoted by $\mathbf{E}_{n,0}$ and $\mathbf{E}_{n,\theta}$ (with subscript n dropped when applied to a function of single X_i).

A test Ψ_n is of level $\alpha \in (0, 1)$ if

$$\mathbf{E}_{n,0}\Psi_n = \alpha.$$

The power of a test Ψ_n is defined by

$$\beta_n(\theta) = \mathbf{E}_{n,\theta}\Psi_n.$$

It is well-known that for a fixed test size $\alpha \in (0, 1)$ and a fixed alternative θ this definition is not so useful, since the power of every reasonable test will tend to 1 (consistency of a test), i.e.,

$$\operatorname*{Lim}_{n\to\infty} \beta_n(\theta) = 1$$

for every $\theta > 0$. Indeed (see also Hettmansperger [32], Subsection 1.3), for every $\theta > 0$ according to the Neyman-Pearson fundamental Lemma, the most powerful test for $\mathbf{H}_0$ rejects $\mathbf{H}_0$ if

$$\Lambda_n(\theta) = \sum_{i=1}^{n}(l_\theta(X_i) - l_0(X_i)) > c_{n,\theta},$$

where $l_\theta(x) = \log p(x, \theta)$ and $c_{n,\theta}$ is defined by

$$\mathbf{P}_{n,0}\{\Lambda_n(\theta) > c_{n,\theta}\} = \alpha.$$

(We tacitly assume continuity of the corresponding distribution).

By the Central Limit Theorem

$$\mathcal{L}\left(\frac{\Lambda_n(\theta) - n\mu_0}{\sigma_0\sqrt{n}}\middle|\mathbf{H}_0\right) \to \mathcal{N}(0, 1),$$

where

$$\mu_0 = \mathbf{E}_0(l_\theta(X_1) - l_0(X_1)), \quad \sigma_0^2 = \mathbf{D}_0(l_\theta(X_1) - l_0(X_1)),$$

with $\mathbf{D}_0$ stands for variance under $\mathbf{H}_0$.

Hence

$$c_{n,\theta} = u_\alpha \sigma_0 \sqrt{n} + n\mu_0 + \cdots, \tag{1.1}$$

where $u_\alpha = \Phi^{-1}(1-\alpha)$ denotes the upper α-point of the standard normal distribution and $\Phi(x)$ stands for the standard normal distribution function.

Now we are in a position to prove that

$$\beta_n(\theta) = \mathbf{P}_{n,\theta}\{\Lambda_n(\theta) > c_{n,\theta}\} \to 1.$$

Let μ_θ and σ_θ^2 be the expectation and the variance of $l_\theta(X_1) - l_0(X_1)$ under $\mathbf{H}_1$, respectively. Applying once again the Central Limit Theorem to $\Lambda_n(\theta)$, we obtain

$$\mathcal{L}\left(\frac{\Lambda_n(\theta) - n\mu_\theta}{\sigma_\theta\sqrt{n}} \middle| \mathbf{H}_1\right) \to \mathcal{N}(0, 1).$$

Therefore, in view of (1.1)

$$\beta_n(\theta) = \mathbf{P}_{n,\theta}\{\Lambda_n(\theta) > c_{n,\theta}\} = 1 - \Phi\left(\frac{c_{n,\theta} - n\mu_\theta}{\sigma_\theta\sqrt{n}}\right) + o(1)$$

$$= \Phi\left(\frac{\sqrt{n}(\mu_\theta - \mu_0) - u_\alpha\sigma_0}{\sigma_\theta}\right) + o(1). \tag{1.2}$$

Application of Jensen's inequality to μ_0 and μ_θ yields

$$\mu_0 = \mathbf{E}_0 \log\frac{p(X_1, \theta)}{p(X_1, 0)} < 0, \quad \mu_\theta = \mathbf{E}_\theta \log\frac{p(X_1, \theta)}{p(X_1, 0)} > 0.$$

It follows that

$$\sqrt{n}(\mu_\theta - \mu_0) \to +\infty,$$

and due to (1.2)

$$\beta_n(\theta) = \mathbf{P}_{n,\theta}\{\Lambda_n(\theta) > c_{n,\theta}\} \to 1.$$

Any reasonable test should be consistent. If a test is not consistent for a reasonable set of alternatives, it should be rejected as defective.

This result is not sufficiently informative for the comparison of tests performance because such an evaluation would require knowledge of the rate of convergence of their powers to 1. This, however, is a complicated matter and we will not consider this problem here (see, e.g., Chernoff [19, 20], Bahadur [7, 8, 9], Groeneboom and Oosterhoff [26, 28], Groeneboom [27], Kallenberg [34, 35] and Nikitin [38]).

Usually, the following Pitman's approach (see Pitman [48] and Noether [39]) is used: the test size $\alpha \in (0, 1)$ remains fixed but instead of a fixed alternative $\theta > 0$ we consider so called local or contiguous alternatives $\{\theta_n\}$ for which $\theta_n \to 0$ as $n \to \infty$ at such rate that the power tends to a limit which lies strictly between α and 1. Under natural regularity conditions (see, e.g., (1.2)), it can be easily shown that the class of these sequences is the class of sequences θ_n for which

$$\operatorname*{Lim}_{n\to\infty} \sqrt{n}\theta_n = t,$$

for some constant t with $0 < t < \infty$. In justification of this approach, we might argue that large sample sizes would be relevant in practice only if the alternatives of interest were close to the null hypothesis and thus hard to distinguish with only a small sample. Relation (1.2) shows that we are concerned with θ_n such that

$$\mu_\theta - \mu_0 = O(n^{-1/2}).$$

So for any $0 < t \le C$, $C > 0$ we will consider testing

$$\mathbf{H}_0 : \theta = 0 \quad \text{against} \quad \mathbf{H}_{n1} : \theta = \tau t, \quad 0 < t \le C. \tag{1.3}$$

Throughout the paper, we use the abbreviation

$$\tau = n^{-1/2},$$

and we denote $\mathbf{P}_{n,0}$ and $\mathbf{P}_{n,t}$ the distributions of $(X_1, \cdots, X_n)$ under $\mathbf{H}_0$ and $\mathbf{H}_{n1}$, respectively. Obviously, they have densities

$$p_{n,0}(\mathbf{x}_n) = \prod_1^n p(x_i, 0) \quad \text{and} \quad p_{n,t}(\mathbf{x}_n) = \prod_1^n p(x_i, \tau t), \tag{1.4}$$

with respect to the corresponding product measure, $\mathbf{x}_n = (x_1, \cdots, x_n)$. The respective expectations will be denoted by $\mathbf{E}_{n,0}$ and $\mathbf{E}_{n,t}$ (with subscript n omitted when applied to a function of single X_i). Denote by

$$\beta_n(t) = \mathbf{E}_{n,t}\Psi_n, \tag{1.5}$$

the power of a test Ψ_n for $\mathbf{H}_0$ against the local alternative $\mathbf{H}_{n,t} : \theta = \tau t$. Considered as a function of t, this sequence converges for every reasonable test to a monotone continuous function assuming its values in $(0, 1)$ (see (1.14)). Using the argument from Pfanzagl [45, p.43], we note that this reparametrization is not only a matter of technical convenience. In meaningful applications of tests, we know which alternatives we wish to discriminate from the hypothesis with high probability, say, γ, and we choose the sample size accordingly. Hence it is reasonable to compare the power of different tests for alternatives with maximal rejection probability γ, irrespective of the sample size. Exactly this is achieved if we compare the power functions $\beta_n(t)$ for fixed t.

Assume that all measures $\mathbf{P}_{n,t}$ are mutually absolutely continuous. Consider the log-likelihood ratio

$$\Lambda_n(t) = \log \frac{d\mathbf{P}_{n,t}}{d\mathbf{P}_{n,0}} = \log \frac{p_{n,t}}{p_{n,0}}.$$

Then by (1.4)

$$\Lambda_n(t) = \sum_{i=1}^{n} [l_{\tau t}(X_i) - l_0(X_i)]. \tag{1.6}$$

By the Taylor series expansion,

$$l_{\tau t}(X_i) - l_0(X_i) = \tau t l^{(1)}(X_i) + \frac{1}{2}(\tau t)^2 l^{(2)}(X_i) + \cdots. \tag{1.7}$$

Here and in what follows the k-th derivative of a function with respect to θ will be denoted by the superscript k. For a function of θ at $\theta = 0$, the argument θ will be often suppressed, e.g.,

$$l^{(2)}(x) = \frac{\partial^2}{\partial \theta^2} l_\theta(x) \bigg|_{\theta=0}.$$

Denote

$$L_n^{(1)} = \tau \sum_{i=1}^{n} l^{(1)}(X_i), \quad L_n^{(2)} = \tau \sum_{i=1}^{n} [l^{(2)}(X_i) - \mathbf{E}_0 l^{(2)}(X_1)], \cdots \tag{1.8}$$

The sums are centered by the corresponding $\mathbf{E}_0$-expectations; the first sum contains no centering because

$$\mathbf{E}_0 l^{(1)}(X_1) = 0.$$

Further, denote by I the Fisher information

$$I = \mathbf{E}_0 (l^{(1)}(X_1))^2.$$

It is well-known that

$$\mathbf{E}_0 l^{(2)}(X_1) = -I.$$

With this notation, putting (1.7) into (1.6) yields

$$\Lambda_n(t) = t L_n^{(1)} - \frac{1}{2} t^2 I + \frac{1}{2} \tau t^2 L_n^{(2)} + \cdots. \tag{1.9}$$

The first two terms in the right-hand side of (1.9) express the local asymptotic normality (LAN) of the family of distributions. In the case of one-parameter family, a simple and sufficient condition for LAN was obtained in Hájek [29].

The omitted terms in (1.9) include the nonrandom term

$$\frac{1}{6}\tau t^3 \mathbf{E}_0 l^{(3)}(X_1)$$

and the terms of higher order than τ.

The Neyman-Pearson test, i.e., the most powerful size-α test for $\mathbf{H}_0$ against $\mathbf{H}_{n,t} : \theta = \tau t$ rejects $\mathbf{H}_0$ when

$$\Lambda_n(t) > c_{n,t}$$

with $c_{n,t}$ defined by (assuming continuity of the corresponding distribution)

$$\mathbf{P}_{n,0}\{\Lambda_n(t) > c_{n,t}\} = \alpha.$$

Using (1.9) and the Central Limit Theorem, we obtain

$$\mathcal{L}(\Lambda_n(t)|\mathbf{H}_0) \to \mathcal{N}\left(-\frac{1}{2}t^2 I,\, t^2 I\right). \tag{1.10}$$

Hence

$$c_{n,t} \to c_t = t\sqrt{I}u_\alpha - 1/2t^2 I, \tag{1.11}$$

$$\Phi(u_\alpha) = 1 - \alpha.$$

The power of this most powerful test is

$$\beta_n^*(t) = \mathbf{P}_{n,t}(\Lambda_n(t) > c_{n,t}). \tag{1.12}$$

It is well-known from the LAN theory (see also (1.9)) that

$$\mathcal{L}\left(\Lambda_n(t)|\mathbf{H}_{n,t}\right) \to \mathcal{N}\left(\frac{1}{2}t^2 I,\, t^2 I\right). \tag{1.13}$$

Thus (1.11)-(1.13) yield

$$\beta_n^*(t) \to \beta^*(t) = \Phi(t\sqrt{I} - u_\alpha). \tag{1.14}$$

These results have been obtained by Wald [50].

Note that $\beta_n^*(t)$, known as the *envelope power function* (i.e., the supremum over all size-α tests of the power at τt), is not the power function of a single test. The envelope power function renders a standard for evaluating the power function of any particular test. For each $t > 0$, it is the power of the most powerful test against $\mathbf{H}_{n,t}$ based on $\Lambda_n(t)$. Thus it provides an upper bound for the power of any test for $\mathbf{H}_0$ against $\mathbf{H}_1 : t > 0$.

It is well-known that there are many (first order) asymptotically efficient tests, i.e., tests whose power function $\beta_n(t)$ converges to the same limit as $\beta_n^*(t)$. So are, for example, tests based on $L_n^{(1)}$, on $\Lambda_n(t_0)$ with an arbitrary $t_0 > 0$, on the maximum likelihood estimator $\hat{\theta}_n$, on a certain linear combination of order statistics; on a certan U-statistics; for θ location parameter there are asymptotically efficient rank tests (see Section 4). Hence, there is an abundance of tests fulfilling

$$\beta_n(t) \to \beta^*(t), \quad t > 0, \tag{1.15}$$

i.e., of tests which are most powerful for $\mathbf{H}_0$ against $\mathbf{H}_{n,t}$ up to an error $o(1)$ for every $t > 0$. They can be compared with each other by higher order terms of their power. We even have the result (see Pfanzagl [41], p.31, Theorem 6) that if (1.15) is satisfied for one $t > 0$ then (1.15) holds for all $t > 0$ (efficiency up to $o(1)$ for one $t > 0$ implies efficiency up to $o(1)$ for all $t > 0$).

Before proceeding to the higher-order theory, we will derive some simple formulas to be used in the sequel.

Denote by $p_{0,t}(x)$ and $p_{1,t}(x)$ the limiting densities of $\Lambda_n(t)$ under $\mathbf{H}_0$ and $\mathbf{H}_{n,t}$ respectively, which correspond to the normal distributions in (1.10) and (1.13). Note that they are related to each other by

$$e^x p_{0,t}(x) = p_{1,t}(x),$$

which follows from the properties of the log-likelihood ratio or can be verified directly. We will need expressions for $p_{0,t}(c_t)$ and $p_{1,t}(c_t)$. Putting (1.11) into the explicit expressions for normal densities (1.10), (1.13) yields

$$p_{0,t}(c_t) = \frac{1}{t\sqrt{I}}\varphi(u_\alpha), \quad p_{1,t}(c_t) = \frac{1}{t\sqrt{I}}\varphi(u_\alpha - t\sqrt{I}). \tag{1.16}$$

Next, suppose instead of $c_{n,t}$ we use another critical value $\overline{c}_{n,t}$, say, which also converges to c_t (see (1.10), (1.11)). Then the test with the critical region

$$\Lambda_n(t) > \overline{c}_{n,t}$$

has size $\overline{\alpha}_n$ and power $\overline{\beta}_n^*(t)$ converging to α and $\beta^*(t)$, respectively. Let us now have two such sequences $\overline{c}_{n,t}$ and $\widetilde{c}_{n,t}$ converging to c_t with

$$\delta_n = \overline{c}_{n,t} - \widetilde{c}_{n,t} \to 0,$$

and we need expressions for the differences of the corresponding sizes and powers up to $o(\delta_n)$. Assuming certain regularity, so that the distribution functions of $\Lambda_n(t)$ under $\mathbf{P}_{n,0}$ and $\mathbf{P}_{n,t}$ have Edgeworth expansions, it is easy to see that these differences are entirely determined by the leading terms of these expansions, because the next terms will contribute at most $O(\tau\delta_n) = o(\delta_n)$. The leading terms are the normal distributions we have just discussed. Thus, it is readily seen that

$$\widetilde{\alpha}_n - \overline{\alpha}_n = \delta_n p_{0,t}(c_t) + o(\delta_n) = \frac{\delta_n}{t\sqrt{I}}\varphi(u_\alpha) + o(\delta_n), \tag{1.17}$$

$$\widetilde{\beta}_n^*(t) - \overline{\beta}_n^*(t) = \delta_n p_{1,t}(c_t) + o(\delta_n) = \frac{\delta_n}{t\sqrt{I}}\varphi(u_\alpha - t\sqrt{I}) + o(\delta_n). \tag{1.18}$$

2. Second Order Efficiency

Typically, an asymptotically efficient test statistic (suitably normalized) has the score function $L_n^{(1)}$ as its leading term, so that it has the form

$$T_n = L_n^{(1)} + \tau Q_n + \cdots, \tag{2.1}$$

with Q_n bounded in probability. For example (see (1.9)), $\Lambda_n(t_0)$ is equivalent to

$$T_n = L_n^{(1)} + \frac{1}{2}\tau t_0 L_n^{(2)}.$$

For rank statistics (R-statistics) and linear combinations of order statistics (L-statistics) Q_n can be written as a quadratic functional of the empirical process (centered and normalized empirical distribution function) (see Bening [11, 12]).

In 70-ies for the power functions $\beta_n(t)$ of various asymptotically efficient tests an expansion in τ to terms of order τ^2 was obtained. The purpose was to study the deficiencies (see Section 3) of the corresponding tests, which we will briefly discuss later on. Writing down such expansions in an explicit form required very involved calculations. For "parametric" test statistics first a "stochastic expansion" of the form (2.1), but containing also the τ^2 term was derived. It was used to obtain the

Edgeworth expansions for the distributions of T_n under $\mathbf{H}_0$ and $\mathbf{H}_{n,t}$. For rank statistics a different technique based on a certain conditioning was used by Albers et al. [1] and Bickel and Van Zwet [17]. The Edgeworth expansion under $\mathbf{H}_0$ was used to obtain an expansion in τ for the critical value a_n defined by

$$\mathbf{P}_{n,0}\{T_n > a_n\} = \alpha.$$

Then the Edgeworth expansion for

$$\beta_n(t) = \mathbf{P}_{n,t}\{T_n > a_n\}$$

was derived by the substitution of the expansion for a_n into the Edgeworth expansion under $\mathbf{H}_{n,t}$. The Edgeworth expansions for $\beta_n^*(t)$ with error terms $o(\tau)$ and $o(\tau^2)$ have been obtained independently by Chibisov ([21], p.40, Theorem 9.1; [22], Section 9) and Pfanzagl ([40], Section 4; [41], pp.223 and 225).

Though the Edgeworth expansions for the distributions of various asymptotically efficient test statistics and of $\Lambda_n(t)$ differ by terms of order τ, it was observed that their powers $\beta_n(t)$ differ from each other and from $\beta_n^*(t)$ by $o(\tau)$ (and typically by $O(\tau^2)$), so that “first-order efficiency implies second-order efficiency”(see Pfanzagl [44]), the latter meaning that the power agrees with $\beta_n^*(t)$ up to terms of order τ. The approach of comparing the expansions for $\beta_n^*(t)$ and $\beta_n(t)$ described above gave no insight into the nature of this phenomenon. A simple and intuitively clear proof of this general property was given by Bickel et al. [18]. We outline here that proof adapted to the present setup.

The idea was, first, to treat directly the difference

$$\beta_n^*(t) - \beta_n(t)$$

and, secondly, to adjust the test statistic to the log-likelihood ratio (rather than to adjust test statistics and the log-likelihood ratio to $L_n^{(1)}$), so that the difference

$$\Delta_{n,t} \equiv S_{n,t} - \Lambda_n(t) \tag{2.2}$$

is small. For example, (2.1) as a test statistic is equivalent to

$$S_{n,t} = tT_n - \frac{1}{2}t^2 I$$

(note that this transformation does not influence the test function, and hence, the power) and then (see (1.9))

$$\Delta_{n,t} = -\tau\left(\frac{1}{2}\tau t^2 L_n^{(2)} - tQ_n\right) + \cdots.$$

(We state this expression to show that $\Delta_{n,t}$ is of order τ and do not need its particular form). Throughout the rest of this section we mostly suppress the subscript and argument t. Let c_n and b_n be the corresponding critical values defined by

$$\mathbf{P}_{n,0}\{\Lambda_n > c_n\} = \mathbf{P}_{n,0}\{S_n > b_n\} = \alpha. \tag{2.3}$$

Then the corresponding powers are

$$\beta_n^*(t) = \mathbf{P}_{n,t}\{\Lambda_n > c_n\}, \quad \beta_n(t) = \mathbf{P}_{n,t}\{S_n > b_n\}.$$

Their difference is

$$\beta_n^*(t) - \beta_n(t) = \int_{\{\Lambda_n > c_n\}} d\mathbf{P}_{n,t} - \int_{\{S_n > b_n\}} d\mathbf{P}_{n,t}$$

$$= \int_{A_+} d\mathbf{P}_{n,t} - \int_{A_-} d\mathbf{P}_{n,t}, \tag{2.4}$$

where

$$A_+ = \{\Lambda_n > c_n,\, S_n \le b_n\}, \quad A_- = \{\Lambda_n \le c_n,\, S_n > b_n\}. \tag{2.5}$$

Since

$$d\mathbf{P}_{n,t} = e^{\Lambda_n} d\mathbf{P}_{n,0},$$

and by (2.3)

$$\int_{\{\Lambda_n > c_n\}} d\mathbf{P}_{n,0} - \int_{\{S_n > b_n\}} d\mathbf{P}_{n,0} = 0,$$

we can rewrite (2.4) as

$$\beta_n^*(t) - \beta_n(t) = \left(\int_{A_+} - \int_{A_-} \right) (e^{\Lambda_n} - e^{c_n}) d\mathbf{P}_{n,0}. \tag{2.6}$$

Using (2.2) rewrite (2.5) as

$$A_+ = \{c_n < \Lambda_n \le b_n - \Delta_n\}, \quad A_- = \{b_n - \Delta_n < \Lambda_n \le c_n\}. \tag{2.7}$$

Since Δ_n is of order τ, so is the difference of distribution functions of Λ_n and S_n, hence so is $c_n - b_n$. Thus Λ_n in (2.6) varies in the layer (2.7) having of order τ. Moreover, the integrand in (2.6) vanishes on one side of this layer, namely, on the surface $\Lambda_n = c_n$, so that it remains $O(\tau)$ in the domain of integration. Its integration over the thin layer results in

$$\beta_n^*(t) - \beta_n(t) = o(\tau). \tag{2.8}$$

An argument of this type was used in Bickel et al. [18] to obtain (2.8) under very general conditions, in particular, on the magnitude of Δ_n. When Δ_n is of order τ, it is seen from the above argument that the difference in (2.8) is likely to be $O(\tau^2)$.

Thus asymptotically efficient tests can be compared to each other by considering the higher order terms of their power and one may ask whether there exists an asymptotically (up to $o(\tau^2)$) most powerful test and, if not, whether one can find a sufficiently small asymptotically complete class. This problem was solved by Pfanzagl [42] and Pfanzagl and Wefelmeyer [43] who showed that, under suitable regularity conditions, the family of tests based on

$$\{\Lambda_n(s),\ s \geq 0\}$$

forms an asymptotically complete class. This means that for any sequence of size-α tests having powers $\beta_n(t)$, there exists a sequence $s_n \geq 0$ such that

$$\beta_n(t) \leq \beta^*_{n, s_n}(t) + o(\tau^2),$$

for all $t > 0$. Here $\beta^*_{n, s_n}(t)$ stands for the power of the size-α test based on $\Lambda_n(s_n)$. It was shown in Chibisov [23] (p.1069, Example 2.2) that

$$\beta^*_n(t) - \beta^*_{n,s}(t) = \tau^2 \frac{D_{t,s}}{2t\sqrt{I}} \varphi(t\sqrt{I} - u_\alpha),$$

where

$$D_{t,s} = \frac{1}{4} t^2 (t-s)^2 \left(\mathbf{Var}_0 l^{(2)}(X_1) - I^{-1}\mathbf{Cov}_0^2(l^{(1)}(X_1),\ l^{(2)}(X_1))\right).$$

These formulas show that the log-likelihood ratio tests for different s do not dominate each other and their powers differ by terms of order τ^2, unless $D_{t,s}$ vanishes. It does so when $\{\mathbf{P}_\theta,\ \theta \in \Theta \subset \mathbf{R}^1\}$ is an exponential family because then a uniformly most powerful test exists and $\beta^*_{n,s}(t)$ does not depend on $s \geq 0$.

3. Power Loss

The difference

$$\beta_n^*(t) - \beta_n(t)$$

is closely related to the *deficiency* of the corresponding test, which is the number of additional observations needed for this test to achieve the same power as the most powerful test. This notion was introduced by Hodges and Lehmann [33]. Deficiencies of various tests were extensively studied in 70-ies by Albers et al. [1] (for rank tests), by Chibisov [23], Pfanzagl [45] (for "parametric" tests), Bender [10], Albers [2, 3], Klaassen and Van Zwet [36] and Bening and Chibisov [16] (for the test theory with nuisance parameters) and others.

When the limit

$$r(t) := \operatorname*{Lim}_{n\to\infty} n(\beta_n^*(t) - \beta_n(t)) \tag{3.1}$$

exists, the asymptotic deficiency is finite and can be directly expressed through this limit. We will not state this relationship here. Rather, we will directly deal with the quantity (3.1), which we will refer to as the *power loss*. This quantity was actually the object of the studies on deficiency. As we pointed out, its derivation was very involved.

An elaboration of the argument given in the previous section leads to the following formula for the power loss. Suppose that

$$\Delta_{n,t} = S_{n,t} - \Lambda_n(t)$$

as in (2.2) is of order τ in a somewhat stronger sense then it was meant before. Namely, assume that

$$(\sqrt{n}\Delta_{n,t},\ \Lambda_n(t))$$

converges in distribution under $\mathbf{P}_{n,0}$ to a certain bivariate random variable. Denoting

$$\prod_{n,t} = \sqrt{n}\Delta_{n,t},$$

we write it as

$$(\prod_{n,t} \Lambda_n(t)) \overset{\mathbf{P}_{n,0}}{\to} (\prod, \Lambda). \tag{3.2}$$

In all regular cases Λ is a normal random variable (see (1.10)). Denote its distribution function and density by $\Phi_{0,t}(x)$ and $p_{0,t}(x)$. Let c_t be the limiting critical value defined by

$$\Phi_{0,t}(c_t) = 1 - \alpha.$$

Then

$$r(t) = \operatorname*{Lim}_{n\to\infty} n(\beta_n^*(t) - \beta_n(t)) = \frac{1}{2} e^{c_t} p_{0,t}(c_t)\mathbf{Var}[\prod \mid \Lambda = c_t]. \tag{3.3}$$

Note that (see (1.10)-(1.13), (1.16))

$$e^{c_t} p_{0,t}(c_t) = p_{1,t}(c_t),$$

where $p_{1,t}(x)$ is the limiting density of $\Lambda_n(t)$ under $\mathbf{P}_{n,t}$ and

$$c_t = t\sqrt{I}u_\alpha - \frac{1}{2}t^2 I, \; p_{0,t}(x) = \frac{1}{t\sqrt{I}}\varphi\left(\frac{2x + t^2 I}{2t\sqrt{I}}\right), \tag{3.4}$$

$$p_{1,t}(x) = \frac{1}{t\sqrt{I}}\varphi\left(\frac{2x - t^2 I}{2t\sqrt{I}}\right), \; \Phi_{0,t}(x) = \Phi\left(\frac{2x + t^2 I}{2t\sqrt{I}}\right). \tag{3.5}$$

Combined with (3.3) and (1.16) these relations imply

$$r(t) = \frac{1}{2t\sqrt{I}}\varphi(u_\alpha - t\sqrt{I})\mathbf{Var}[\prod \mid \Lambda = c_t]. \tag{3.6}$$

Example 3.1. Let $(X_1, \cdots, X_n)$ be independent identically distributed (i.i.d.) observations with distribution function $F(x, \theta)$ and density $p(x, \theta), \theta$ ranging over an open set $\Theta \subset \mathbf{R}^1$ containing 0. Let the hypothesis

$$\mathbf{H}_0 : \theta = 0$$

be tested against a sequences of local alternatives

$$\mathbf{H}_{n1} : \theta = \tau t,\ 0 < t \le C,\ \tau = n^{-1/2}.$$

Consider an asymptotically efficient test based on

$$T_n = \tau \sum_{i=1}^{n} l^{(1)}(X_i) = L_n^{(1)}, \tag{3.7}$$

$$l^{(1)}(x) = \frac{\partial}{\partial \theta} \log p(x, \theta)\Big|_{\theta=0}.$$

Writing out the Taylor expansion of $\Lambda_n(t)$ as described in Section 1 (see (1.8) for the notation and (1.9)) we have

$$\Lambda_n(t) = tL_n^{(1)} - \frac{1}{2}t^2 I + \frac{1}{2}\tau t^2\left(L_n^{(2)} + \frac{1}{3}tm_3\right) + \cdots, \tag{3.8}$$

$$m_k = \mathbf{E}_0 l^{(k)}(X_1), \quad k = 1, 2, \cdots.$$

We introduce

$$S_{n,t} = tT_n - \frac{1}{2}t^2 I.$$

Assume that the joint distribution of

$$\left(L_n^{(1)}, L_n^{(2)}\right)$$

converges under $\mathbf{H}_0$ to a normal one. Denote by

$$\left(L^{(1)}, L^{(2)}\right)$$

a random vector in $\mathbf{R}^2$ having this limiting distribution.

Then (3.2) holds with

$$\Delta_{n,t} = -\frac{1}{2}\tau t^2\left(L_n^{(2)} + \frac{1}{3}tm_3\right) + \cdots, \tag{3.9}$$

$$\Lambda = tL^{(1)} - \frac{1}{2}t^2 I, \quad \Pi = -\frac{1}{2}t^2\left(L^{(2)} + \frac{1}{3}tm_3\right).$$

Then we have (see (3.4), (3.6))

$$r(t) = \frac{1}{2t\sqrt{I}}\varphi(u_\alpha - t\sqrt{I})\mathbf{Var}[\Pi | L^{(1)} = \sqrt{I}u_\alpha]$$

$$= \frac{t^3}{8\sqrt{I}}\varphi(u_\alpha - t\sqrt{I})\left(\mathbf{Var}_0 l^{(2)}(X_1) - I^{-1}\mathbf{E}_0^2 l^{(1)}(X_1) l^{(2)}(X_1)\right). \tag{3.10}$$

In the above argument we assumed that the tests have exactly size α, but the formula (3.3) remains valid when the sizes converge to α and equal each other up to $o(\tau^2)$, i.e.,

$$\mathbf{P}_{n,0}\{\Lambda_n(t) > c_{n,t}\} - \mathbf{P}_{n,0}\{T_n > a_n\} = o(\tau^2).$$

The formula (3.3) demonstrates, in particular, that the power loss (hence the deficiency) is determined by the terms of order τ of the asymptotically efficient test statistic.

The formula (3.3) was proved by Chibisov [24] for statistics admitting a stochastic expansion in terms of sums of independent identically distributed random variables (which is typical for “parametric” problems) subject to certain conditions. Bening [11, 12, 13] proved the formula (3.3) for rank statistics, linear combinations of order statistics and U-statistics (see Section 4).

4. Tests Based on L-, R- and U-Statistics

In this section, a heuristic derivation of explicit formulas for $\prod$, Λ and $r(t)$ (see (3.1), (3.2) and (3.6)) related to L-, R- and U-tests is given. Its justification under suitable regularity conditions along with some details of derivation is given in Bening [11, 12, 13, 14].

Let $(X_1, \cdots, X_n)$ be independent identically distributed random observations with distribution function $F(x, \theta)$ and density $p(x, \theta)$, θ ranging over an open set $\Theta \subset \mathbf{R}^1$ containing 0. Let the hypothesis

$$\mathbf{H}_0 : \theta = 0 \tag{4.1}$$

be tested against a sequences of local alternatives

$$\mathbf{H}_{n1} : \theta = \tau t,\ 0 < t \leq C,\ C > 0, \tag{4.2}$$

where $\tau = n^{-1/2}$. We will write $F(x)$ for the hypothesized distribution function $F(x, 0)$.

4.1. L-test

Consider an asymptotically efficient L-test based on

$$T_{n1} = \tau \sum_{i=1}^{n} b_{in} X_{i:n}, \tag{4.3}$$

where $(X_{1:n}, \cdots, X_{n:n})$ are the order statistics of $(X_1, \cdots, X_n)$

$$l(x, \theta) = \log p(x, \theta),\ l^{(i)}(x) = \frac{\partial^i}{\partial \theta^i} \log p(x, \theta)\bigg|_{\theta=0},\ i = 1, 2, \ldots,$$

$$b_{in} = n \int_{(i-1)/n}^{i/n} J_1(s) ds,\ J_k(s) = (l^{(k)}(x))'\big|_{x=F^{-1}(s)},\ k \in N = \{1, 2, \cdots\}, \tag{4.4}$$

$$F^{-1}(s) = \inf\{x : F(x) \geq s\},$$

and a prime denoting differentiation with respect to x.

Given $\alpha \in (0, 1)$, we need to find

$$r_1(t) = \operatorname*{Lim}_{n \to \infty} n(\beta_n^*(t) - \beta_{n1}(t)),$$

where $\beta_{n1}(t)$ and $\beta_n^*(t)$ are the powers of the size $\alpha \in (0, 1)$ tests based on T_{n1} and on

$$\Lambda_n(t) = \sum_{i=1}^{n} \log \frac{p(X_i, \tau t)}{p(X_i, 0)}, \tag{4.5}$$

respectively.

Let $\Gamma_n(s)$ denote the empirical distribution function of $F(X_1), \cdots, F(X_n)$, i.e.,

$$\Gamma_n(s) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{I}_{[0, s)}(F(X_i)),$$

and let

$$B_n(s) = \sqrt{n}(\Gamma_n(s) - s),\ s \in [0, 1].$$

Put

$$L_n^{(k)} = -\int_0^1 J_k(s) B_n(s) dF^{-1}(s),\ a_k = \mathbf{E}_0 l^{(k)}(X_1),\ k \in N. \tag{4.6}$$

Then, as is readily seen,

$$\tau \sum_{i=1}^{n} l^{(k)}(X_i) = L_n^{(k)} + \sqrt{n} a_k,\ k \in N. \tag{4.7}$$

In a regular case, when the identity

$$\int p(x, \theta) dx \equiv 1, \quad \theta \in \Theta$$

can be differentiated under the integral sign, we have

$$a_1 = 0, \quad a_2 = -I,$$

where

$$I = \mathbf{E}_0\left(l^{(1)}(X_1)\right)^2.$$

Using the Taylor's formula in (4.5), one has

$$\Lambda_n(t) \approx tL_n^{(1)} - \frac{t^2}{2} I + \tau\left(\frac{t^2}{2} L_n^{(2)} + \frac{t^3}{6} a_3\right). \tag{4.8}$$

Let $B(s)$ stand for a Brownian bridge on $[0, 1]$, i.e., $B(t)$ is centered Gaussian process with covariances

$$\mathbf{E} B(s)B(u) = \min(s, u) - su, \ s, u \in [0, 1],$$

and having continuous paths and $L^{(k)}$ be defined by (4.6) with $B_n(s)$ substituted by $B(s)$. Then since

$$B_n(s) \xrightarrow{D} B(s)$$

so (see (4.8))

$$\Lambda_n(t) \xrightarrow{D} \Lambda = tL^{(1)} - \frac{1}{2} t^2 I. \tag{4.9}$$

Thus, as above, (3.4) and (3.5) are valid.

Now we will express statistic T_{n1} in terms of $\Gamma_n(s)$ (cf. Helmers [30, 31]). Put

$$\overline{J}_1(s) = sA - \int_0^s J_1(s)ds, \ A = \int_0^1 J_1(s)ds, \tag{4.10}$$

$$M_{n1} = A\tau\sum_{i=1}^{n} X_i.$$

Then (cf. (4.3) and (4.4))

$$T_{n1} = -\sqrt{n}\sum_{i=1}^{n} X_{(i)} \int_{\frac{i-1}{n}}^{\frac{i}{n}} d\overline{J}_1(s) + M_{n1}$$

$$= \sqrt{n}\sum_{i=1}^{n-1} \overline{J}_1(i/n)\left(X_{(i+1)} - X_{(i)}\right) + M_{n1}$$

$$= \sqrt{n}\int_0^1 \overline{J}_1(\Gamma_n(s))dF^{-1}(s) + M_{n1}.$$

Writing the argument of $\overline{J}_1$ as $\Gamma_n(s) = s + \tau B_n(s)$ and expanding $\overline{J}_1(\Gamma_n(s))$ by the Taylor formula, one obtains

$$T_{n1} \approx L_n^{(1)} + \frac{1}{2}\tau K_{n1} + M_n, \tag{4.11}$$

$$K_{n1} = -\int_0^1 J_1'(s)B_n^2(s)dF^{-1}(s), \tag{4.12}$$

$$M_n = \sqrt{n}\int_0^1 F^{-1}(s)J_1(s)ds.$$

Therefore, one has from (4.8) and (4.11)

$$\Delta_{n1}(t) \equiv S_{n1,t} - \Lambda_n(t) \approx -\tau\left(\frac{t}{2}K_{n1} - \frac{t^2}{2}L_n^{(2)}\right), \tag{4.13}$$

with

$$S_{n1,t} = tT_{n1} + b_{n,t},\ b_{n,t} = -\frac{t^2}{2}I + \tau\frac{t^3}{6}a_3 - tM_n.$$

Define K_1 by (4.12) with $B_n(\cdot)$ substituted by $B(\cdot)$. It is seen from (4.8) and (4.13) that

$$\left(\sqrt{n}\Delta_{n2}(t),\ \Lambda_n(t)\right)\xrightarrow{D}(\Pi_1,\ \Lambda),$$

with

$$\Pi_1 = -\frac{t}{2} K_1 + \frac{t^2}{2} L^{(2)}$$

and Λ as in (4.9). Then by (3.6)

$$r_1(t) = \lim_{n\to\infty} n(\beta_n^*(t) - \beta_{n1}(t))$$

$$= \frac{t}{8\sqrt{I}} \varphi\left(u_\alpha - t\sqrt{I}\right)\mathbf{Var}\left[K_1 - tL^{(2)} | L^{(1)} = u_\alpha\sqrt{I}\right]. \tag{4.14}$$

The latter conditional variance has the form:

$$\mathbf{Var}\left[K_1 - tL^{(2)} | L^{(1)} = u_\alpha\sqrt{I}\right] = v_0 + v_1 t + v_2 t^2, \tag{4.15}$$

where

$$v_0 = 2\left(I_2^2 - 2I_1 I_2 + I_{112}\right) + 4(1 - u_\alpha^2)\left(I_1^2 - I_{111}\right),$$

$$v_1 = 4u_\alpha(I_0 I_1 - I_{011}),$$

$$v_2 = I_{001} - I_0^2,$$

$$I_i = \int_0^1 J_{2-i}^{(i)}(s)\mu^{i+1}(s)dF^{-1}(s),\ i = 0, 1;$$

$$I_2 = \int_0^1 J_1'(s)s(1-s)dF^{-1}(s),$$

$$I_{ijl} = \int_0^1\int_0^1 J_{2-i}^{(i)}(s)J_{2-j}^{(j)}(t)\mu^{i-l+1}(s)\mu^{j-l+1}(t)K^l(s, t)dF^{-1}(s)dF^{-1}(t),$$

$$i, j = 0, 1;\ l = 1, 2;$$

$$K(s, t) = \min(s, t) - st, \tag{4.16}$$

$$\mu(s) = \frac{1}{\sqrt{I}}\int_0^s l^{(1)}(F^{-1}(t))dt.$$

4.2. R-test

Assume now that θ is a location parameter, $p(x, \theta) = p(x - \theta)$, and density $p(x)$ is symmetric, $p(-x) = p(x)$. Consider an asymptotically efficient rank test (R-test) for testing $\mathbf{H}_0$ against $\mathbf{H}_{n1}$ based on

$$T_{n2} = \tau \sum_{i=1}^{n} a(R_i^+) \operatorname{sgn}(X_i), \tag{4.17}$$

where $(R_1^+, \cdots, R_n^+)$ is the vector of ranks of $(|X_1|, \cdots, |X_n|)$ and

$$a(i) = I\left(\frac{i}{n+1}\right) \text{ with } I(s) = l^{(1)}\left(F^{-1}\left(\frac{1+s}{2}\right)\right). \tag{4.18}$$

Denote the powers of the size $\alpha \in (0, 1)$ tests based on T_{n2} and $\Lambda_n(t)$ by $\beta_{n2}(t)$ and $\beta_n^*(t)$, respectively. As above, (3.4), (3.5), (4.8) and (4.9) are valid.

Let us approximate T_{n2} by a functional of $B_n(\cdot)$. Let $\overline{F}(x) = 2F(x) - 1$ be the distribution function of $|X_1|$ and let $F_n(x)$, $\overline{F}_n(x)$ and $\Gamma_n(x)$ be the empirical distribution functions of $X_1, \cdots, X_n$, of $|X_1|, \cdots, |X_n|$ and $F(X_1), \cdots, F(X_n)$, respectively. One has

$$R_i^+ = n\overline{F}_n(|X_i|) = n\left(\overline{F}(|X_i|) + \left(\overline{F}_n(|X_i|) - \overline{F}(|X_i|)\right)\right).$$

The Taylor series expansion of $a(R_i^+) = I(R_i^+/(n+1))$ in (4.17) yields

$$T_{n2} \approx \tau \sum_{i=1}^{n} I\left(\overline{F}(|X_i|)\right) \operatorname{sgn}(X_i)$$

$$+ \tau \sum_{i=1}^{n} I'\left(\overline{F}(|X_i|)\right)\left(\overline{F}_n(|X_i|) - \overline{F}(|X_i|)\right) \operatorname{sgn}(X_i). \tag{4.19}$$

Notice that, by definition (4.18),

$$I\left(\overline{F}(|X_i|)\right) \operatorname{sgn}(X_i) = l^{(1)}(X_i).$$

Hence the first term in the right hand side of (4.19) is $L_n^{(1)}$ (cf. (4.7)). In the second term, we use formulas

$$\overline{F}_n(x) = F_n(x) - F_n(-x),\ x \geq 0,$$

$$\Gamma_n(s) = F_n(F^{-1}(s)),\ \Gamma_n(1-s) = F_n(-F^{-1}(s)),$$

and integration by part, which yields

$$T_{n2} \approx L_n^{(1)} + \tau \frac{1}{2} K_{n2}, \tag{4.20}$$

where

$$K_{n2} = \int_0^1 J_0'(s)(B_n(1-s) - B_n(s))dB_n(s), \tag{4.21}$$

$$J_0'(s) = -\frac{J_1(s)}{p(F^{-1}(s))}.$$

Therefore (4.8) and (4.20) imply

$$\Delta_{n2}(t) \equiv S_{n2,t} - \Lambda_n(t) \approx -\tau\left(\frac{t}{2} K_{n2} - \frac{t^2}{2} L_n^{(2)}\right), \tag{4.22}$$

with

$$S_{n2,t} = tT_{n2} - \frac{1}{2} t^2 I.$$

Define K_2 by (4.21) with $B_n(\cdot)$ substituted by $B(\cdot)$. Then (4.8) and (4.22) imply

$$\left(\sqrt{n}\Delta_{n2}(t),\ \Lambda_n(t)\right) \xrightarrow{D} (\Pi_2,\ \Lambda),$$

with

$$\Pi_2 = -\frac{t}{2} K_2 + \frac{t^2}{2} L^{(2)}$$

and Λ as in (4.9).

Using (3.6), we have

$$r_2(t) = \lim_{n\to\infty} n(\beta_n^*(t) - \beta_{n2}(t))$$

$$= \frac{t}{8\sqrt{I}}\varphi\left(u_\alpha - t\sqrt{I}\right)\mathbf{Var}\left[K_2 - tL^{(2)}\middle| L^{(1)} = u_\alpha\sqrt{I}\right]. \tag{4.23}$$

The latter conditional variance is given by

$$\mathbf{Var}\left[K_2 - tL^{(2)}\middle| L^{(1)} = u_\alpha\sqrt{I}\right] = w_0 + w_1 t + w_2 t^2, \tag{4.24}$$

where

$$w_0 = 4\int_0^1 (I'(s))^2 s(1-s)ds$$

$$+ \frac{4}{I}(u_\alpha^2 - 1)\int_0^1\int_0^1 I'(s)I'(t)I(s)I(t)K(s, t)dsdt,$$

$$w_1 = -\frac{2u_\alpha}{\sqrt{I}}\int_0^1\int_0^1 I'(s)I(s)g(t)K(s, t)dsdt,$$

$$w_2 = \frac{1}{4}\int_0^1\int_0^1 g(s)g(t)K(s, t)dsdt,$$

$$g(s) = \frac{l^{(3)}\left(F^{-1}((1+s)/2)\right)}{p\left(F^{-1}((1+s)/2)\right)},$$

with $K(\cdot, \cdot)$ and $I(\cdot)$ defined by (4.16) and (4.18). The expression for $r_2(t)$ given by (4.23), (4.24) agrees with formula (6.3) in Albers et al. [1].

4.3. U-test

Finally, let us now consider an asymptotically efficient U-test for testing $\mathbf{H}_0$ versus $\mathbf{H}_{n1}$ (cf. (4.1) and (4.2)) based on

$$T_{n3} = \sum_{1\le i<j\le n} h(X_i, X_j), \tag{4.25}$$

where $\Psi(x, y)$ is measurable and symmetric in its two arguments, i.e., $\Psi(x, y) = \Psi(y, x)$, real function,

$$\mathbf{E}_0[\Psi(X_1, X_2)|X_1] = 0 \quad \text{a.s.} \tag{4.26}$$

and

$$h(x, y) = l^{(1)}(x) + l^{(1)}(y) + \Psi(x, y). \tag{4.27}$$

Denote the powers of the size $\alpha \in (0, 1)$ tests based on T_{n3} and $\Lambda_n(t)$ by $\beta_{n3}(t)$ and $\beta_n^*(t)$, respectively. As above, (3.4), (3.5), (4.8) and (4.9) are valid.

Let

$$\Delta_{n3}(t) \equiv S_{n3,t} - \Lambda_n(t), \tag{4.28}$$

where

$$S_{n3,t} = \frac{\tau t}{n-1} T_{n3} - \frac{t^2 I}{2} + \tau \frac{t^3}{6} a_3.$$

Combining (4.25)-(4.28) and (4.8), we can write

$$\Lambda_{n3}(t) \approx -\tau\left(\frac{t}{n-1} \sum_{1 \le i < j \le n} \Psi(X_i, X_j) - \frac{t^2}{2} L_n^{(2)}\right). \tag{4.29}$$

For the first member of this representation, being the U-statistic with degenerate kernal (see (4.27)) and Theorem 4.3.1 from Koroljuk and Borovskikh [37] can be applied. It follows that (see also (4.9))

$$\sqrt{n}\Delta_{n3}(t) \xrightarrow{D} \prod_3, \tag{4.30}$$

with

$$\prod_3 = -\frac{t}{2} K_3 + \frac{t^2}{2} L^{(2)}, \quad \Lambda = tL^{(1)} - \frac{1}{2} t^2 I,$$

$$K_3 = \sum_{j=1}^{\infty} \lambda_j \left(\xi_j^2 - 1\right), \tag{4.31}$$

$\{\xi_j\}$ are independent identically distributed normal random variables,

$$\xi_j = \int_0^1 e_j\left(F^{-1}(s)\right) dB(s),$$

$$L^{(2)} = -\int_0^1 J_2(s) B(s) dF^{-1}(s),$$

$$L^{(1)} = -\int_0^1 J_1(s) B(s) dF^{-1}(s), \quad J_k(s) = (l^{(k)}(x))'|_{x=F^{-1}(s)}, \; k \in N,$$

$B(s)$ stands for a Brownian bridge and $\{\lambda_j\}$, $\{e_j\}$ are the eigenvalues and eigenfunctions of the linear operator $S : L_2[0, 1] \to L_2[0, 1]$

$$S : f \to \mathbf{E}_0 \Psi(X_1, x) f(X_1). \tag{4.32}$$

As above, using (3.6), we have

$$\begin{aligned} r_3(t) &= \operatorname*{Lim}_{n\to\infty} n(\beta_n^*(t) - \beta_{n3}(t)) \\ &= \frac{t}{8\sqrt{I}} \varphi\left(u_\alpha - t\sqrt{I}\right) \mathbf{D}\left[K_3 - tL^{(2)} | L^{(1)} = u_\alpha \sqrt{I}\right]. \end{aligned} \tag{4.33}$$

The latter conditional variance can be calculated

$$\begin{aligned} \mathbf{Var}\left[K_3 - tL^{(2)} | L^{(1)} = u_\alpha \sqrt{I}\right] &= 4(1 - u_\alpha^2) I_{(1)}^2 \\ &+ t4u_\alpha I_{(1)}\left(I_{(2)} - \sqrt{I}\right) + t^2\left(I_{(3)} - I_{(2)}^2\right), \end{aligned} \tag{4.34}$$

where

$$I_{(1)} = I^{-1} \mathbf{E}_0 \Psi(X_1, X_2) l^{(1)}(X_1) l^{(1)}(X_2),$$

$$I_{(2)} = I^{-1/2} \mathbf{E}_0 l^{(2)}(X_1) l^{(1)}(X_1),$$

$$I_{(3)} = \mathbf{E}_0 \left(l^{(2)}(X_1)\right)^2 - I^2.$$

5. Combined L-Tests

In this section, we consider combined L-tests in the one-sample problem and investigate the difference (power loss) between the power of a given combined asymptotically efficient L-test and that of the most powerful test.

Consider a testing problem in which the total sample has been divided into $m \geq 2$ subsamples. Suppose that for each of these subsamples a separate asymptotically efficient L-statistics can be obtained and the best combination of these statistics is then compared to the ordinary undivided asymptotically efficient L-statistic. One easily sees that, under natural conditions, splitting causes no first order efficiency loss. Hence, it becomes interesting to derive second order results and it would be nice to obtain more precise results on effect of splitting. This subject and related ones have received considerable attention in the literature. For earlier works, we refer to Van Zwet and Oosterhoff [49], Albers and Akritas [4]. One- and two-sample combined rank tests have been considered in Albers [5, 6]. Note again that the used method was essentially based on asymptotic expansions. Although the basic ideas underlying these papers are simple, the proofs are highly technical matter. The method we used for proving our results carries over LeCam's approach to higher order asymptotics and based on the likelihood ratio properties.

Let $(X_1, \cdots, X_n)$ be independent identically distributed random observations with distribution function $F(x, \theta)$ and density $p(x, \theta)$, θ ranging over an open set $\Theta \subset \mathbf{R}^1$ containing 0. Let the hypothesis

$$\mathbf{H}_0 : \theta = 0$$

be tested against a sequences of local alternatives

$$\mathbf{H}_{n1} : \theta = \tau t.\ 0 < t \leq C, \quad C > 0,$$

where $\tau = n^{-1/2}$. We will write $F(x)$ for the hypothesized distribution function $F(x, 0)$.

Consider an asymptotically efficient L-test based on

$$T_{n1} = \tau \sum_{i=1}^{n} b_{in} X_{i:n}, \tag{5.1}$$

where $(X_{1:n}, \cdots, X_{n:n})$ are the order statistics of $(X_1, \cdots, X_n)$

$$l(x, \theta) = \log p(x, \theta),\ l^{(i)}(x) = \frac{\partial^i}{\partial \theta^i} \log p(x, \theta)\Big|_{\theta=0},\ i = 1, 2, \ldots$$

$$b_{in} = n \int_{(i-1)/n}^{i/n} J_1(s) ds,\ J_k(s) = (l^{(k)}(x))'\big|_{x=F^{-1}(s)},\ k \in N, \tag{5.2}$$

$$F^{-1}(s) = \inf\{x : F(x) \geq s\},$$

and a prime denoting differentiation with respect to x.

Given $\alpha \in (0, 1)$, denote $\beta_{n1}(t)$ and $\beta_n^*(t)$ the powers of the size α tests based on T_{n1} and on

$$\Lambda_n(t) = \sum_{i=1}^{n} \log \frac{p(X_i, \tau t)}{p(X_i, 0)}$$

respectively.

Now suppose that our sample has been split into $m \geq 2$ subsamples of sizes n_l, $l = 1, \cdots, m$,

$$(X_1, \cdots, X_{n1}), \cdots, (X_{n-n_m+1}, \cdots, X_n),\ n_1 + \cdots + n_m = n,$$

and

$$\frac{n_l}{n} \to \gamma_l > 0,\ l = 1, \cdots, m;\ \sum_{l=1}^{m} \gamma_l = 1, \text{ as } n \to \infty.$$

For each of these samples we obtain the L-statistics

$$T_{n1}^{(l)},\ l = 1, \cdots, m$$

as in (5.1) and (5.2).

Consider the combined L-statistic

$$\overline{T}_{n1} = \sum_{l=1}^{m} T_{n1}^{(l)}$$

and the level-$\alpha \in (0, 1)$ test based on $\overline{T}_{n1}$. Let $\overline{\beta}_{n1}(t)$ be the power of this combined L-test. The purpose of the present section is to find the following limits:

$$\overline{r}_1^{(m)}(t) = \operatorname*{Lim}_{n\to\infty} n(\beta_{n1}(t) - \overline{\beta}_{n1}(t)), \tag{5.3}$$

$$r_1^{(m)}(t) = \lim_{n\to\infty} n(\beta_n^*(t) - \overline{\beta}_{n1}(t)). \tag{5.4}$$

By the arguments similar to those used for obtaining (4.14) and (4.15) (see Bening [15] for detail), we get

$$r_1^{(m)}(t) = \frac{t}{8\sqrt{I}}\varphi(u_\alpha - t\sqrt{I})\mathbf{Var}\left[\overline{K}_1 - tL^{(2)} | L^{(1)} = u_\alpha\sqrt{I}\right],$$

$$\overline{r}_1^{(m)}(t) = \frac{t}{8\sqrt{I}}\varphi(u_\alpha - t\sqrt{I})\mathbf{Var}\left[\overline{K}_1 - tL^{(2)} | L^{(1)} = u_\alpha\sqrt{I}\right]$$

$$- \frac{t}{8\sqrt{I}}\varphi(u_\alpha - t\sqrt{I})\mathbf{Var}\left[K_1 - tL^{(2)} | L^{(1)} = u_\alpha\sqrt{I}\right],$$

where

$$L^{(k)} = -\int_0^1 J_k(s)B(s)dF^{-1}(s), \ k = 1, 2;$$

$$K_1 = -\int_0^1 J_1'(s)B^2(s)dF^{-1}(s),$$

$$\overline{K}_1 = -\sum_{l=1}^{m}\int_0^1 J_1'(s)B_l^2(s)dF^{-1}(s),$$

and

$$B(s) = \sum_{l=1}^{m} \sqrt{\gamma_l} B_l(s)$$

is the Brownian bridge on $[0, 1]$, $B_1(s), \cdots, B_r(s)$ are mutually independent Brownian bridges defined on $[0, 1]$ and $u_\alpha = \Phi^{-1}(1 - \alpha)$ denotes the upper α-point of the standard normal distribution.

The latter conditional variances can be found:

(i) $$\mathbf{Var}\left[K_1 - tL^{(2)} | L^{(1)} = u_\alpha \sqrt{I}\right] = v_0 + v_1 t + v_2 t^2,$$

where

$$v_0 = 2\left(I_2^2 - 2I_2 I_1 + I_{112}\right) + 4\left(1 - u_\alpha^2\right)\left(I_1^2 - I_{111}\right),$$

$$v_1 = 4u_\alpha (I_0 I_1 - I_{011}),$$

$$v_2 = I_{001} - I_0^2,$$

$$I_i = \int_0^1 J_{2-i}^{(i)}(s) \mu^{i+1}(s) dF^{-1}(s),\ i = 0, 1;$$

$$I_2 = \int_0^1 J_1'(s) s(1 - s) dF^{-1}(s),$$

$$I_{ijl} = \int_0^1 \int_0^1 J_{2-i}^{(i)}(s) J_{2-j}^{(j)}(u) \mu^{i-l+1}(s) \mu^{j-l+1}(u) K^l(s, u) dF^{-1}(s) dF^{-1}(u),$$

$$i,\ j = 0, 1;\ l = 1, 2;$$

$$K(s, u) = \min(s, u) - su,$$

$$\mu(s) = \frac{1}{\sqrt{I}} \int_0^s l^{(1)}(F^{-1}(u)) du.$$

(ii)
$$\mathbf{Var}\left[\overline{K}_1 - tL^{(2)} | L^{(1)} = u_\alpha \sqrt{I}\right]$$

$$= \mathbf{Var}\left[K_1 - tL^{(2)} | L^{(1)} = u_\alpha \sqrt{I}\right] + v_{0m} + v_{1m}t + v_{2m}t^2,$$

where

$$v_{0m} = 2(m-1)I_{112},$$

$$v_{1m} = 4u_\alpha \left(\sum_{l=1}^{m} \sqrt{\gamma_l} - 1 \right) (I_0 I_1 - I_{011}),$$

$$v_{2m} = (m-1)I_{001} - \left(\left(\sum_{l=1}^{m} \sqrt{\gamma_i} \right)^2 - 1 \right) I_0^2.$$

(iii)
$$\overline{r}_1^{(m)}(t) = \frac{t}{8\sqrt{I}} \varphi(u_\alpha - t\sqrt{I})(v_{0m} + tv_{1m} + t^2 v_{2m}),$$

$$r_1^{(m)}(t) = \frac{t}{8\sqrt{I}} \varphi(u_\alpha - t\sqrt{I})(v_0 + v_{0m} + t(v_1 + v_{1m}) + t^2(v_2 + v_{2m})).$$

References

[1] W. Albers, P. J. Bickel and W. R. Van Zwet, Asymptotic expansion for the power of distribution free tests in the one-sample problem, Annals of Statistics 4(1) (1976), 108-156; Correction 6(5) (1978), 1170-1171.

DOI: https://doi.org/10.1214/aos/1176343350

[2] W. Albers, Some Asymptotic Results for Rank Tests with Nuisance Parameter, In: Proceedings of the 3rd Prague Symposium on Asymptotical Statistics, P. Mandl and M. Hušková (Editors), Elsevier Science Publishers, Amsterdam-New-York-Oxford (1984), 159-164.

[3] W. Albers, Rates of Convergence and Expansions for Rank Tests: A Review and Some Remarks on the Nuisance Parameter Case, In: Proceedings of the 7th Conference on Probability Theory, (M. Iosifescu, Editor), Brasov, Romania (1982), 23-32, Bucurecsi, 1984.

[4] W. Albers and M. G. Akritas, Combined rank tests for the two-sample problem with randomly censored data, Journal of the American Statistical Association 82(398) (1987), 648-655.

DOI: https://doi.org/10.1080/01621459.1987.10478479

[5] W. Albers, Deficiencies of Combined Rank Tests, Medical Informatics and Statistics, Report 19, University of Limburg, 1989.

[6] W. Albers, Second order analysis of two-stage rank tests for the one-sample problem, Annals of Statistics 19(2) (1991), 1042-1052.

DOI: https://doi.org/10.1214/aos/1176348136

[7] R. R. Bahadur, Stochastic comparison of tests, Annals of Mathematical Statistics 31(2) (1960), 276-295.

DOI: https://doi.org/10.1214/aoms/1177705894

[8] R. R. Bahadur, On the asymptotic efficiency of tests and estimates, Sankhya 22(3-4) (1960), 229-252.

[9] R. R. Bahadur, Some Limit Theorems in Statistics, SIAM, Philadelphia, 1971.

[10] K.-L. Bender, Vergleich der Einhullenden Gutefunktionen Dritter Ordnung fur Verschiedene Klassen von Testverfahren, Inaugural-Dissertation, Köln, 1980.

[11] V. E. Bening, A formula for deficiency: One sample L- and R-tests I, Mathematical Methods of Statistics 4(2) (1995), 167-188.

[12] V. E. Bening, A formula for deficiency: One sample L- and R-tests II, Mathematical Methods of Statistics 4(3) (1995), 274-293.

[13] V. E. Bening, Asymptotic Analysis of the Distributions of Some Asymptotically Efficient Test Statistics, Doct. Sci. Thesis, Moscow State University, Moscow, 1997 (in Russian).

[14] V. E. Bening, Some properties of asymptotically efficient U-statistics, Teor. Veroyatn. Primen. 42 (1997), 382-384 (in Russian).

[15] V. E. Bening, Deficiencies of combined L-tests, Journal of Mathematical Sciences 92(4) (1998), 4003-4009.

DOI: https://doi.org/10.1007/BF02432333

[16] V. E. Bening and D. M. Chibisov, Higher order asymptotic optimality in testing problem with nuisance parameters I, Mathematical Methods of Statistics 8(2) (1999), 142-165.

[17] P. J. Bickel and W. R. Van Zwet, Asymptotic expansions for the power of distribution free tests in the two-sample problem, Annals of Statistics 6(5) (1978), 937-1004.

DOI: https://doi.org/10.1214/aos/1176344305

[18] P. J. Bickel, D. M. Chibisov and W. R. Van Zwet, On efficiency of first and second order, International Statistical Review 49(2) (1981), 169-175.

DOI: https://doi.org/10.2307/1403070

[19] H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on sum of observations, Annals of Mathematical Statistics 23(4) (1952), 493-507.

DOI: https://doi.org/10.1214/aoms/1177729330

[20] H. Chernoff, Large-sample theory: Parametric case, Annals of Mathematical Statistics 27(1) (1956), 1-22.

DOI: https://doi.org/10.1214/aoms/1177728347

[21] D. M. Chibisov, Asymptotic Expansions for Neyman's $C(\alpha)$ Tests, In: Proceedings of the Second Japan-USSR Symposium on Probability Theory, Lecture Notes in Mathematics 330 (1973), 16-45.

DOI: https://doi.org/10.1007/BFb0061479

[22] D. M. Chibisov, Asymptotic Expansions for Some Asymptotically Optimal Tests, In: Proceedings of the Prague Symposium on Asymptotic Statistics, 2, (J. Hájek, Editor), 37-68, Charles University, Prague, 1974.

[23] D. M. Chibisov, Asymptotic Expansions and Deficiencies of Tests, In: Proceedings of the International Congress of Mathematicians, Warszawa 2 (1983), 1063-1079.

[24] D. M. Chibisov, Calculation of the deficiency of asymptotically efficient tests, Theory of Probability & Its Applications 30(2) (1986), 289-310.

DOI: https://doi.org/10.1137/1130037

[25] F. Götze and H. Milbrodt, The work of Johan Pfanzagl, Mathematical Methods of Statistics 8(2) (1999), 121-141.

[26] P. Groeneboom and J. Oosterhoff, Bahadur efficiency and probabilities of large deviations, Statistica Neerlandica 31(1) (1977), 1-24.

DOI: https://doi.org/10.1002/stan.1977.31.1.1

[27] P. Groeneboom, Large Deviations and Asymptotic Efficiencies, Mathematical Centre Tracts 118, Amsterdam: Mathematisch Centrum, 1979.

[28] P. Groeneboom and J. Oosterhoff, Bahadur efficiency and small-sample efficiency, International Statistical Review 49(2) (1981), 127-141.

DOI: https://doi.org/10.2307/1403067

[29] J. Hájek, Local Asymptotic Minimax and Admissibility in Estimation, In: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability 1, University California Press (1972), 175-194.

[30] R. Helmers, Edgeworth expansions for linear combinations of order statistics with smooth weight functions, Annals of Statistics 8(6) (1980), 1361-1374.

DOI: https://doi.org/10.1214/aos/1176345207

[31] R. Helmers, Edgeworth Expansions for Linear Combinations of Order Statistics, Mathematisch Centrum Amsterdam, 1984.

[32] T. P. Hettmansperger, Statistical Inference Based on Ranks, Wiley, New York, 1984.

[33] J. L. Hodges Jr. and E. L. Lehmann, Deficiency, Annals of Mathematical Statistics 41(3) (1970), 783-801.

DOI: https://doi.org/10.1214/aoms/1177696959

[34] W. C. M. Kallenberg, Asymptotic Optimality of Likelihood Ratio Tests in Exponential Families, Report 77, Department of Mathematics, Vrije Universiteit, Amsterdam, 1978.

[35] W. C. M. Kallenberg, Bahadur deficiency likelihood ratio tests in exponential families, Journal of Multivariate Analysis 11(4) (1981), 506-531.

DOI: https://doi.org/10.1016/0047-259X(81)90093-2

[36] C. A. J. Klaassen and W. R. Van Zwet, On Estimating a Parameter and Its Score Function, In: Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer, (L. M. LeCam and R. A. Olsen Editors), 2 (1985), 827-839, Wadsworth, Inc.

[37] V. S. Koroljuk and Yu. V. Borovskich, Theory of U-Statistics, Naukova Dumka, Kiev, 1989 (in Russian).

[38] Ya. Yu. Nikitin, Hodges-Lehmann and Chernoff efficiencies of linear rank statistics, Journal of Statistical Planning and Inference 29(3) (1991), 309-323.

DOI: https://doi.org/10.1016/0378-3758(91)90006-Z

[39] G. E. Noether, On a theorem of Pitman, Annals of Mathematical Statistics 26(1) (1955), 64-68.

DOI: https://doi.org/10.1214/aoms/1177728593

[40] J. Pfanzagl, Asymptotic expansions related to minimum contrast estimators, Annals of Statistics 1(6) (1973), 993-1026; Corrections 2(6) (1974), 1357-1358.

DOI: https://doi.org/10.1214/aos/1176342554

[41] J. Pfanzagl, Asymptotically Optimum Estimation and Test Procedures, In: Proceedings of the Prague Symposium on Asymptotic Statistics (J. Hájek, Editor) 1, pp. 201-272, Charles University, Prague, 1974.

[42] J. Pfanzagl, On Asymptotically Complete Classes, In: Statistical Inference and Related Topics, (M. L. Puri, Editor), 2, pp. 1-43, Academic Press, New York, 1975.

[43] J. Pfanzagl and W. Wefelmeyer, An asymptotically complete class of tests, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 45(1) (1978), 49-72.

DOI: https://doi.org/10.1007/BF00635963

[44] J. Pfanzagl, First Order Efficiency Implies Second Order Efficiency, In: Contributions to Statistics, Jaroslav Hájek Memorial Volume (J. Jurecková, Editor), pp. 167-196, Academia, Prague, 1979.

[45] J. Pfanzagl, Asymptotic Expansions in Parametric Statistical Theory, In: Developments in Statistics, 3, pp. 1-97, (P. R. Krishnaiah, Editor), Academic Press, New York, 1980.

DOI: https://doi.org/10.1016/B978-0-12-426603-2.50007-5

[46] J. Pfanzagl and W. Wefelmeyer, Contributions to a General Asymptotic Statistical Theory, Lecture Notes in Statistics 13, Springer-Verlag, New York, 1982.

[47] J. Pfanzagl and W. Wefelmeyer, Asymptotic expansions for general statistical models, Lecture Notes in Statistics 31, New-York, Springer-Verlag, 1985.

[48] E. J. G. Pitman, Lecture Notes on Nonparametric Statistical Inference, Lectures given for the University of North Carolina, Institute of Statistics, 1948.

[49] W. R. Van Zwet and J. Oosterhoff, On the combinations of independent test statistics, Annals of Mathematical Statistics 38(3) (1967), 659-680.

DOI: https://doi.org/10.1214/aoms/1177698861

[50] A. Wald, Asymptotically most powerful tests of statistical hypothesis, Annals of Mathematical Statistics 12(1) (1941), 1-19.

DOI: https://doi.org/10.1214/aoms/1177731783

[51] V. E. Bening, Asymptotic Theory of Testing Statistical Hypotheses: Efficient Statistics, Optimality, Power Loss and Deficiency, Berlin: Walter de Gruyter (2011), 277 pp., ISBN 978-3-11-093599-8.

■