One-Sided Confidence About Functionals Over Tangent Cones

Helmut Rieder
University of Bayreuth, Germany

1. March 2002

Abstract

In the setup of i.i.d. observations and a real valued differentiable functional T, locally asymptotic upper bounds are derived for the power of one-sided tests (simple, versus large values of T) and for the confidence probability of lower confidence limits (for the value of T), in the case that the tangent set is only a convex cone. The bounds, and the tests and estimators which achieve the bounds, are based on the projection of the influence curve of the functional on the closed convex cone, as opposed to its closed linear span. The higher efficiency comes along with some weaker, only one-sided, regularity and stability.

Key Words and Phrases: semiparametric models; linear tangent spaces; convex tangent cones; projection; influence curves; differentiable functionals; asymptotically linear estimators; one-sided tests; lower confidence bounds; concentration bound; asymptotic median unbiasedness.
AMS/MSC-2000 classification: 62F35.

1 Introduction

Given a model $\mathcal{P}$ of probability measures on some sample space, let some one dimensional aspect be defined by some statistical functional $T\colon \mathcal{P} \to \mathbb{R}$. We consider the simplest case of n stochastically independent observations $x_1, \ldots, x_n$ with identical distribution any $P \in \mathcal{P}$, and the task is to make confidence statements on the unknown value $T(P)$ by means of tests and estimators.

In the usual testing problems concerning the value of T, the power of level α tests cannot exceed certain asymptotic upper bounds. Likewise, the accuracy of estimators of $T(P)$ is limited by some asymptotic upper bounds for one- and two-sided confidence probabilities. These bounds form a classical subject of non- and semiparametric theory; confer, for example, Bickel et al. (1993), Pfanzagl and Wefelmeyer (1982), Rieder (1994), and van der Vaart (1998).

Having fixed any $P \in \mathcal{P}$, either for the purpose of testing local alternatives or, in estimation, to be able to exclude artificial phenomena of superefficiency, local variations of P within $\mathcal{P}$ must be taken into account[1]. These variations are formulated as differentiable paths $(P_{g,s})_{s>0}$ in $\mathcal{P}$, in direction of certain tangents $g \in L_2(P)$ at P, such that, in the Hilbert space of square root densities,

$$\sqrt{dP_{g,s}} = \bigl(1 + \tfrac{1}{2}sg\bigr)\sqrt{dP} + \mathrm{o}(s) \qquad \text{as } s \downarrow 0 \tag{1.1}$$

The functions g necessarily have expectation $\mathrm{E}\,g = \langle g|1\rangle = 0$ under P; in other words, $g \perp$ the constants in $L_2(P)$. Given any $g \in L_2(P)$, $\langle g|1\rangle = 0$, a corresponding path (in the set of all probabilities) is

$$dP_{g,s} = \Bigl(\tfrac{1}{2}sg + \sqrt{1 - \tfrac{1}{4}s^2\,\|g\|^2}\,\Bigr)^2\,dP \tag{1.2}$$

or

$$dP_{g,s} = (1 + sg)\,dP \qquad \text{if } g \in L_\infty(P) \tag{1.3}$$

The set $\mathcal{G}$ of all tangents at P on one hand reflects the richness of the model $\mathcal{P}$. On the other hand, $\mathcal{G}$ is restricted by the differentiability requirement on the functional: There exist some function $\kappa \in L_2(P)$, such that for every $g \in \mathcal{G}$ and any path (1.1) in $\mathcal{P}$,

$$T(P_{g,s}) = T(P) + s\langle\kappa|g\rangle + \mathrm{o}(s) \qquad \text{as } s \downarrow 0 \tag{1.4}$$

The function κ, a so-called influence curve of T at P, may not be unique. But the orthogonal projection $\bar\kappa$ of κ on the closed linear span $c\ell\,\mathrm{lin}\,\mathcal{G}$ of $\mathcal{G}$ in $L_2(P)$ is unique—the canonical gradient, or efficient influence curve.

By definition, the tangent set $\mathcal{G}$ of $\mathcal{P}$ at P is a cone in $L_2(P) \cap \{1\}^\perp$ with vertex at 0, such that $\gamma g \in \mathcal{G}$ for $g \in \mathcal{G}$ and $\gamma \in [0,\infty)$. For example, the classical nonparametric alternative hypotheses of positive asymmetry and positive dependence naturally lead to cones.

Furthermore, there is a general argument why arbitrary tangent sets should be considered in theory. In testing, the null hypothesis usually is canonical and simple, but the alternative may be chosen freely, more complex, according to the particular case at hand. In estimation, as noted by one referee, tangent cones arise if the paramater value is a boundary point of the parameter set. Moreover, also for other parameter values, the previous argument may be enforced from a robustness viewpoint. In the setup of Rieder (1994; Chapter 4), any parametric model distribution may be enlarged to infinitesimal neighborhoods consisting of the local alternatives generated by, for example, a tangent cone (leading us to consider the smallest cone containing the neighborhood cone and the linear span of the parametric tangent).

In most of the literature on asymptotic bounds so far, the tangent set is assumed a linear space $\mathcal{G} = \bar{\mathcal{G}}$, such that $c\ell\,\mathrm{lin}\,\bar{\mathcal{G}}$ is just the closure $c\ell\,\bar{\mathcal{G}}$ of $\bar{\mathcal{G}}$. Then the said bounds are determined by the canonical gradient $\bar\kappa$, acting as a least favorable (limiting) tangent, and its norm $\|\bar\kappa\|$.

If the tangent set is not a linear space but a possibly nonconvex cone $\mathcal{G} = \tilde{\mathcal{G}}$, the situation is not quite clear[2]. In our paper, we shall settle on cones $\tilde{\mathcal{G}}$ that in addition are convex, such that $\gamma_1 g_1 + \gamma_2 g_2 \in \tilde{\mathcal{G}}$ for $g_i \in \tilde{\mathcal{G}}$ and $\gamma_i \in [0,\infty)$.

[1] implicitly, already, in the classical scores function—a derivative, of log densities.

[2] As for nonconvex cones, we refer to the footnote summary in van der Vaart (1998; p 367).

Even in this case, of a convex tangent cone $\tilde{\mathcal{G}}$, the results in literature seem somewhat contradictory: On one hand, the convolution representation and asymptotic minimax risk under symmetric subconvex loss given by van der Vaart (1998; Theorems 25.20 and 25.21) are still expressed by the canonical gradient $\bar{\kappa}$ (the orthogonal projection of κ on $c\ell \operatorname{lin} \tilde{\mathcal{G}}$). On the other hand, Pfanzagl and Wefelmeyer (1982; Theorem 9.2.2) state a two-sided concentration bound in terms of the (smaller) projection $\tilde{\kappa}$ of κ on a closed convex tangent cone $\tilde{\mathcal{G}} = c\ell\, \tilde{\mathcal{G}}$. Their proof, however, makes use of $-\tilde{\mathcal{G}} \subset \tilde{\mathcal{G}}$, so their cone must in fact be a (closed) linear space. Also Janssen (1999), in the context of testing, considers convex tangent cones $\tilde{\mathcal{G}}$ and argues by the projection $\tilde{\kappa}$ of κ on $c\ell\, \tilde{\mathcal{G}}$. But, throughout his paper, he treats $\tilde{\kappa}$ as if it were $\bar{\kappa}$, as he nowhere accounts for the nonorthogonality of the residual $\kappa - \tilde{\kappa}$ on $\tilde{\mathcal{G}}$ in the case that $\bar{\kappa} \notin c\ell\, \tilde{\mathcal{G}}$.

Thus, either by result or by implicit assumption, the asymptotic power and concentration bounds obtained so far for convex tangent cones agree with those for their linear spans.

The present investigation, in the case of convex tangent cones $\tilde{\mathcal{G}}$, derives locally asymptotic upper bounds for the power of one-sided tests (of a simple hypothesis against large values of T), as well as for the confidence probabilities of lower confidence limits for $T(P)$. These asymptotic bounds are given truly in terms of the projection $\tilde{\kappa}$ of the influence curve κ of the functional T on the closed convex cone $c\ell\, \tilde{\mathcal{G}}$ (Theorems 2.1 and 3.1). Since $\bar{\kappa} \in c\ell \operatorname{lin} \tilde{\mathcal{G}} \setminus c\ell\, \tilde{\mathcal{G}}$ in general, that is, $\tilde{\kappa} \neq \bar{\kappa}$ or, equivalently, $\|\tilde{\kappa}\| < \|\bar{\kappa}\|$, the upper bounds are larger than those based on $\bar{\kappa}$.

For the higher efficiency, however, a considerable price has to be paid, which constists in a weaker and merely one-sided regularity and stability: In the case of testing, the asymptotic size rises to 100% over an only slightly enlarged, and therefore over the larger one-sided, null hypothesis (Proposition 2.4). In the case of estimation, the asymptotic bias may become plus infinity under local alternatives (Proposition 3.5). As a consequence, and as the (positive parts of) efficient estimators are asymptotically unique (Proposition 3.4, Remark 3.6), the bound stated by Pfanzagl and Wefelmeyer (1982; Theorem 9.2.2) cannot possibly be attained under the condition of asymptotic median unbiasedness. The merely one-sided regularity, and one-sided asymmetric testing pseudo-loss function, are also responsible for the difference to van der Vaart's (1998) results.

The investigation originated from the attempt by Rieder (2000) to subject robust statistics to the semiparametric approach by treating neighborhoods as nuisance parameters, which leads to (the subtraction of the) nonlinear projection on balls (from the classical scores). Except for one-sided robust testing, however, the influence curves thus obtained may differ from the optimally robust ones of Rieder (1994; Chapter 5). Thus, contrary to what one would hopefully expect, the projection recipe does not always give the optimal procedures.

Therefore, the present extension from linear spaces to convex cones requires subtle modifictions of the proofs in the classical case. Once derived, the new results ask for a careful interpretation of the assumed regularity (Subsection 3.3) and the implied stability (Subsection 3.5), and a comparison for convex cones and their linear spans becomes due (Subsection 3.4).

For reasons of comparability, throughout this paper the cases of a linear tangent space $\bar{\mathcal{G}}$ and a convex tangent cone $\tilde{\mathcal{G}}$, respectively, are stated together. The $L_2(P)$-closure $c\ell\,\bar{\mathcal{G}}$ of a linear tangent space $\bar{\mathcal{G}}$ is again a linear space, the $L_2(P)$-closure $c\ell\,\tilde{\mathcal{G}}$ of a convex cone $\tilde{\mathcal{G}}$ again a convex cone. The canonical gradient, which is the projection of κ on $c\ell\,\bar{\mathcal{G}}$ and $c\ell\,\mathrm{lin}\,\tilde{\mathcal{G}}$, respectively, is denoted by $\bar{\kappa}$, the projection of κ on $c\ell\,\tilde{\mathcal{G}}$ is denoted by $\tilde{\kappa}$.

Convenient characterizations of the projections are supplied in the appendix; the criteria (4.1) and (4.2) for $\bar{\kappa}$ and $\tilde{\kappa}$ will be used without explicit reference. Throughout the paper, the influence curve κ, the tangent space $\bar{\mathcal{G}}$ and convex tangent cone $\tilde{\mathcal{G}}$ at P are assumed of such a kind that

$$\bar{\kappa} \neq 0\,, \qquad \tilde{\kappa} \neq 0 \tag{1.5}$$

As noted, the interesting case occurs if $\bar{\kappa} \neq \tilde{\kappa}$.

One-sided inference about non-smooth functionals of a density has been studied by Donoho (1988), by entirely different techniques and in an even more nonparametric setting. Nevertheless, we encounter a somehow similar impossibility of sensible upper confidence limits: The estimators that provide the best lower confidence limits, subject to some local asymptotic median nonnegativity, necessarily achieve overshoot probability 100% under local alternatives. This distinguishes convex tangent cones from linear tangent spaces, where the efficient estimator is unique and asymptotically median unbiased.

Notation $\langle .\,|.\,\rangle$ has already been used to denote the inner product in $L_2(P)$. $\mathbf{I}$ stands for the indicator function. Limits $\liminf_n$, $\limsup_n$, and $\lim_n$ are meant for $n \to \infty$. Asy. is our abbreviation of asymptotic/asymptotically.

2 One-Sided Tests

2.1 Definition of Hypotheses

For the fixed probability $P \in \mathcal{P}$ and tangent set $\mathcal{G}$, simple and one-sided composite asy. hypotheses about the sequence of laws Q_n of the i.i.d. observations at sample size $n = 1, 2, \ldots$ are defined by

$$J^0 : \;\; Q_n = P \quad \text{eventually} \tag{2.1}$$

$$J : \;\; \lim_n \sqrt{n}\,\big(T(Q_n) - T(P)\big) = 0 \tag{2.2}$$

$$H : \;\; \limsup_n \sqrt{n}\,\big(T(Q_n) - T(P)\big) \le 0 \tag{2.3}$$

$$K : \;\; \liminf_n \sqrt{n}\,\big(T(Q_n) - T(P)\big) \ge c \tag{2.4}$$

where $c \in (0, \infty)$ is some fixed constant. The measures Q_n in (2.2)–(2.4) may not be arbitrary elements of model $\mathcal{P}$ but are assumed to approach P along any path $(P_{g,s})_{s>0}$ in $\mathcal{P}$ such that, for some $g \in \mathcal{G}$ and $t \in (0, \infty)$, eventually,

$$Q_n = P_{n,t,g} = P_{g,t/\sqrt{n}} \tag{2.5}$$

In particular, every such sequence (Q_n^n) is contiguous to (P^n). Also, the expansion (1.4) of the functional is in force such that, for every $g \in \mathcal{G}$ and $t \in (0,\infty)$,

$$\sqrt{n}\,\bigl(T(P_{n,t,g}) - T(P)\bigr) = t\,\langle\kappa|g\rangle + \mathrm{o}(n^0) \tag{2.6}$$

Therefore, the asy. hypotheses J, H and K concern $(g,t) \in \mathcal{G} \times (0,\infty)$ and may be expressed by

$$J^0 : g = 0\,, \qquad J : \langle\kappa|g\rangle = 0\,, \qquad H : \langle\kappa|g\rangle \le 0\,, \qquad K : t\langle\kappa|g\rangle \ge c \tag{2.7}$$

Depending on whether the tangent set $\mathcal{G}$ is a convex cone $\tilde{\mathcal{G}}$ or a linear space $\bar{\mathcal{G}}$, the hypotheses J, H, and K will be denoted by $\tilde{J}$, $\tilde{H}$, and $\tilde{K}$, respectively by $\bar{J}$, $\bar{H}$, and $\bar{K}$; obviously, $\tilde{J}^0 = \bar{J}^0 = J^0$.

Overparametrization $P_{n,t,g} = P_{n,t/\gamma,\gamma g}$ with $t,\gamma > 0$, for $g \in \mathcal{G}$ (a cone), is allowed in (2.5) but, in view of (2.6) and (2.7), consistent with the functional and the hypotheses J^0, J, H, K. Distinction of three (actually, five) null hypotheses J^0, J, and H is essential to Theorem 2.1 and Proposition 2.4.

2.2 Asymptotic Power Bounds for Cones and Spaces

Let us fix some level $\alpha \in (0,1)$, and denote by u_α the upper α-point of the standard normal distribution function Φ, such that $\Phi(-u_\alpha) = \alpha$. We shall employ asy. tests, that is, sequences of tests τ_n at sample size n. Power and size of the tests τ_n are going to be evaluated under the n-fold product measures Q_n^n asy., as $n \to \infty$. An asy. test $(\hat{\tau}_n)$ is said to achieve an upper bound $\inf_K \limsup_n \int \tau_n\,dQ_n^n \le \beta$ with $\limsup_n$ replaced by $\liminf_n$, if itself fulfills the side conditions on the test sequences (τ_n) under consideration and $\inf_K \liminf_n \int \hat{\tau}_n\,dQ_n^n = \beta$ holds.

Theorem 2.1 *Let (τ_n) be an asy. test that maintains asy. level α under J^0,*

$$\limsup_n \int \tau_n\,dP^n \le \alpha \tag{2.8}$$

(a) *Then, in the case of a convex tangent cone $\tilde{\mathcal{G}}$,*

$$\inf_{\tilde{K}} \limsup_n \int \tau_n\,dQ_n^n \le \Phi\Bigl(-u_\alpha + \frac{c}{\|\tilde{\kappa}\|}\Bigr) \tag{2.9}$$

The upper bound (2.9), *with* $\limsup_n$ *replaced by* $\liminf_n$, *is achieved by the asy. test*

$$\tilde{\tau}_n = \mathbf{I}\bigl(\sqrt{n}\,\mathrm{ave}_{i=1}^n\,\tilde{\kappa}(x_i) > \|\tilde{\kappa}\|\,u_\alpha\bigr) \tag{2.10}$$

(b) *In the case of a linear tangent space $\bar{\mathcal{G}}$,*

$$\inf_{\bar{K}} \limsup_n \int \tau_n\,dQ_n^n \le \Phi\Bigl(-u_\alpha + \frac{c}{\|\bar{\kappa}\|}\Bigr) \tag{2.11}$$

The upper bound (2.11), *with* $\limsup_n$ *replaced by* $\liminf_n$, *is achieved by the asy. test*

$$\bar{\tau}_n = \mathbf{I}\big(\sqrt{n}\,\mathrm{ave}_{i=1}^n\,\bar{\kappa}(x_i) > \|\bar{\kappa}\|\,u_\alpha\big) \tag{2.12}$$

Moreover,

$$\sup_{\bar{H}} \limsup_n \int \bar{\tau}_n\,dQ_n^n \le \alpha \tag{2.13}$$

PROOF

(a) Given any $g \in \tilde{\mathcal{G}}$ such that $\langle\kappa|g\rangle > 0$, put $t_g = c/\langle\kappa|g\rangle$ and test J^0 vs. the simple subhypothesis $(P^n_{n,t_g,g})$ of $\tilde{K}$. Path differentiabilty (1.1) ensures the following well-known asy. expansion of loglikelihoods under P^n,

$$\log\frac{dP^n_{n,t_g,g}}{dP^n} = t_g\;\sqrt{n}\,\mathrm{ave}_1^n\,g(x_i) - \tfrac{1}{2}t_g^2\,\|g\|^2 + \mathrm{o}_{P^n}(n^0) \tag{2.14}$$

Thus Corollary 3.4.2 of Rieder[3](1994) is in force and bounds the asy. power under $P^n_{n,t_g,g}$ subject to (2.8) from above by $\Phi(-u_\alpha + t_g\,\|g\|\,)$. Now let $g \in \tilde{\mathcal{G}}$ approach $\tilde{\kappa}$ in $L_2(P)$. Then $t_g\,\|g\|$ tends to $c\|\tilde{\kappa}\|/\langle\kappa|\tilde{\kappa}\rangle = c/\|\tilde{\kappa}\|$ where we have used that $\langle\kappa|\tilde{\kappa}\rangle = \|\tilde{\kappa}\|^2$, and bound (2.9) is obtained as the limit

$$\lim_{g\to\tilde{\kappa}} \Phi(-u_\alpha + t_g\,\|g\|\,) = \Phi\Big(-u_\alpha + \frac{c}{\|\tilde{\kappa}\|}\Big) \tag{2.15}$$

Towards achieving bound (2.9) by the tests $\tilde{\tau}_n$, the sums $\sqrt{n}\,\mathrm{ave}_1^n\,\tilde{\kappa}(x_i)$ are, for every $(g,t) \in \tilde{\mathcal{G}} \times (0,\infty)$ asy. normal under $P^n_{n,t,g}$,

$$\big(\sqrt{n}\,\mathrm{ave}_1^n\,\tilde{\kappa}(x_i)\big)(P^n_{n,t,g}) \xrightarrow{\;\mathrm{w}\;} \mathcal{N}\big(t\langle\tilde{\kappa}|g\rangle, \|\tilde{\kappa}\|^2\,\big) \tag{2.16}$$

by (2.14) and a LeCam lemma, confer HR (1994; Corollary 2.2.6), and so

$$\lim_n \int \tilde{\tau}_n\,dP^n_{n,t,g} = \Phi\Big(-u_\alpha + \frac{t\langle\tilde{\kappa}|g\rangle}{\|\tilde{\kappa}\|}\Big) \tag{2.17}$$

Under $J^0 : g = 0$, this limit equals α. If $(P^n_{n,t,g}) \in \tilde{K}$ then, by (2.7), as $t > 0$, and since $\langle\tilde{\kappa}|g\rangle \ge \langle\kappa|g\rangle \;\; \forall g \in \tilde{\mathcal{G}}$, also $t\langle\tilde{\kappa}|g\rangle \ge t\,\langle\kappa|g\rangle \ge c$. Hence

$$\inf_{\tilde{K}} \lim_n \int \tilde{\tau}_n\,dP^n_{n,t,g} \ge \Phi\Big(-u_\alpha + \frac{c}{\|\tilde{\kappa}\|}\Big) \tag{2.18}$$

(b) With $\bar{\kappa}$ and $\bar{K}$ in the place of $\tilde{\kappa}$ and $\tilde{K}$, the proof of bound (2.11) is the same as in case (a). The limit corresponding to (2.17) for the tests $\bar{\tau}_n$ is

$$\lim_n \int \bar{\tau}_n\,dP^n_{n,t,g} = \Phi\Big(-u_\alpha + \frac{t\langle\bar{\kappa}|g\rangle}{\|\bar{\kappa}\|}\Big) = \Phi\Big(-u_\alpha + \frac{t\langle\kappa|g\rangle}{\|\bar{\kappa}\|}\Big) \tag{2.19}$$

since $\kappa - \bar{\kappa} \perp \bar{\mathcal{G}}$. If $(P^n_{n,t,g}) \in \bar{H}$, then $t\langle\kappa|g\rangle \le 0$ by (2.7) and $t > 0$. Therefore

$$\lim_n \int \bar{\tau}_n\,dP^n_{n,t,g} = \Phi\Big(-u_\alpha + \frac{t\langle\kappa|g\rangle}{\|\bar{\kappa}\|}\Big) \le \Phi(-u_\alpha + 0) = \alpha \tag{2.20}$$

is obtained from (2.19), and proves (2.13). ////

[3] HR, subsequently

Remark 2.2 Although Theorem 2.1 (a), for convex tangent cones, is straightforward to prove, it seems to have been omitted in literature so far. In its proof, $\tilde{\kappa}$ acts as a limiting least favorable tangent, as does $\bar{\kappa}$ in the proof of Theorem 2.1 (b). The latter result, for linear tangent spaces, compares with Pfanzagl and Wefelmeyer (1982; chapter 8), van der Vaart (1998; Theorem 25.44, Lemma 25.45), as well as Beran (1983; Theorem 1) and HR (1994; Theorem 4.3.8) who, in robust statistics, encounter linear tangent spaces $\bar{\mathcal{G}}$ with maximal closure $c\ell\,\bar{\mathcal{G}} = L_2(P) \cap \{1\}^\perp$ (so that $\bar{\kappa} = \kappa - \mathrm{E}\,\kappa$ there). ////

2.3 Comparison of Cones and Their Linear Spans

Let us consider P a member of two models $\tilde{\mathcal{P}} \subset \bar{\mathcal{P}}$ whose tangent sets at P are a convex cone $\tilde{\mathcal{G}}$, respectively the linear span of $\tilde{\mathcal{G}}$,

$$\bar{\mathcal{G}} = \operatorname{lin} \tilde{\mathcal{G}} \tag{2.21}$$

Power Comparison In this situation, we have $\tilde{J} \subset \bar{J}$, $\tilde{H} \subset \bar{H}$, $\tilde{K} \subset \bar{K}$, and J^0 should be easier to test vs. $\tilde{K}$ than vs. $\bar{K}$. In fact,

$$\Phi\Bigl(-u_\alpha + \frac{c}{\|\tilde{\kappa}\|}\Bigr) > \Phi\Bigl(-u_\alpha + \frac{c}{\|\bar{\kappa}\|}\Bigr) \tag{2.22}$$

because

$$\|\tilde{\kappa}\| < \|\bar{\kappa}\| \tag{2.23}$$

unless $\bar{\kappa} \in c\ell\,\tilde{\mathcal{G}}$, in which case $\tilde{\kappa} = \bar{\kappa}$ and the two power bounds coincide.

This is a consequence of $\|\tilde{\kappa}\|^2 = \langle \kappa|\tilde{\kappa}\rangle = \langle \bar{\kappa}|\tilde{\kappa}\rangle$ and the Cauchy–Schwarz inequality: $\langle \bar{\kappa}|\tilde{\kappa}\rangle \le \|\bar{\kappa}\|\,\|\tilde{\kappa}\|$, where equality holds iff $\tilde{\kappa}$ is some positive multiple of $\bar{\kappa}$, in which case $\bar{\kappa} \in c\ell\,\tilde{\mathcal{G}}$ and $\tilde{\kappa} = \bar{\kappa}$.

Sample Size Comparison Allowing for different sample sizes $\tilde{n}$ and $\bar{n}$, respectively, such that $\tilde{n}/n \to \tilde{\gamma}$ and $\bar{n}/n \to \bar{\gamma}$ for some $\tilde{\gamma}, \bar{\gamma} \in (0,\infty)$, the asy. power bounds (2.9) and (2.11) are the same iff

$$\bar{\gamma} : \tilde{\gamma} = \|\bar{\kappa}\|^2 : \|\tilde{\kappa}\|^2 \tag{2.24}$$

Thus, observations at the higher rate $\|\bar{\kappa}\|^2 / \|\tilde{\kappa}\|^2$ are needed by $(\bar{\tau}_{\bar{n}})$ to achieve, subject to level α on J^0, the same power vs. $\bar{K}$ as $(\tilde{\tau}_{\tilde{n}})$ vs. $\tilde{K}$.

Example 2.3 Consider the standard normal $P = \mathcal{N}(0,1)$ and $\kappa(x) = x$ the identity on the real line; κ is the influence curve at P of the expectation functional as well as of the one-sample normal scores rank functional,

$$E(Q) = \int_{-\infty}^{\infty} x\, Q(dx) \tag{2.25}$$

$$R(Q) = 2\int_0^{\infty} \Phi^{-1}\bigl(\tfrac{1}{2} + \tfrac{1}{2}[Q(x) - Q(-x)]\bigr)\, Q(dx) - 2\,\varphi(0) \tag{2.26}$$

where $\varphi = \dot{\Phi}$ denotes the standard normal density, and $Q(x) = Q\bigl((-\infty, x]\bigr)$.

As tangents at P, consider the sign-function $g_1(x) = \mathrm{sign}(x)$ and the function $g_2(x) = \mu\,\mathrm{sign}(x)\,\mathbf{I}(|x| \le a)$ with $\mu, a \in (0,\infty)$. Then $\|g_1\| = 1 = \|\kappa\|$, and $\mu = \mu_a$ may be determined by $\mu_a^{-2} = 2\Phi(a) - 1$ such that also $\|g_2\| = 1$. Then the coefficients $b_i = \langle \kappa | g_i \rangle$ and $c = \langle g_1 | g_2 \rangle$ are given by

$$b_1 = 2\varphi(0)\,, \qquad b_2 = 2\mu\big[\varphi(0) - \varphi(a)\big]\,, \qquad c = 2\mu\big[\Phi(a) - \tfrac{1}{2}\big] \tag{2.27}$$

As tangent sets at P, employ the (closed) convex cone $\tilde{\mathcal{G}} = c\ell\,\tilde{\mathcal{G}}$ and (closed) linear space $\bar{\mathcal{G}} = c\ell\,\bar{\mathcal{G}} = \mathrm{lin}\,\tilde{\mathcal{G}}$ spanned by the tangents g_1 and g_2,

$$\tilde{\mathcal{G}} = \big\{\, \gamma_1 g_1 + \gamma_2 g_2 \;\big|\; \gamma_i \ge 0 \,\big\}\,, \qquad \bar{\mathcal{G}} = \big\{\, \gamma_1 g_1 + \gamma_2 g_2 \;\big|\; \gamma_i \in \mathbb{R} \,\big\} \tag{2.28}$$

Via (1.3), the cone $\tilde{\mathcal{G}}$ defines a set of positively asymmetric alternatives to P.

Unconstrained minimization of $\|\kappa - \gamma_1 g_1 - \gamma_2 g_2\|$ being equivalent to the orthogonality relations $\gamma_1 + \gamma_2 c = b_1$ and $\gamma_1 c + \gamma_2 = b_2$, the canonical gradient is

$$\bar\kappa = \bar\gamma_1 g_1 + \bar\gamma_2 g_2 \qquad \text{where} \quad \bar\gamma_1 = \frac{b_1 - b_2 c}{1 - c^2}\,, \quad \bar\gamma_2 = \frac{b_2 - b_1 c}{1 - c^2} \tag{2.29}$$

In the appendix we show that $\bar\gamma_1 > 0 > \bar\gamma_2$; hence $\bar\kappa \in \bar{\mathcal{G}} \setminus \tilde{\mathcal{G}}$.

The constrained minimization of $\|\kappa - \gamma_1 g_1 - \gamma_2 g_2\|$ subject to $\gamma_i \ge 0$ is a convex and well-posed problem; HR (1994; Theorem B.2.3, Definition B.2.9). Thus there exist multipliers $\beta_i \ge 0$ such that the solutions $\tilde\gamma_i \ge 0$ minimize the following Lagrangian over $\gamma_i \in \mathbb{R}$,

$$\begin{aligned} &\|\kappa - \gamma_1 g_1 - \gamma_2 g_2\|^2 - 2\beta_1\gamma_1 - 2\beta_2\gamma_2 - \mathrm{const} \\ &\quad = \big[\gamma_1 - (b_1 + \beta_1)\big]^2 + \big[\gamma_2 - (b_2 + \beta_2)\big]^2 + 2c\,\gamma_1\gamma_2 \end{aligned} \tag{2.30}$$

Moreover, $\beta_i \tilde\gamma_i = 0$. Since $\tilde\kappa \ne \bar\kappa$, not both β_0 and β_1 can vanish.

In case $\beta_1 > 0$ we obtain that $\tilde\gamma_1 = 0$ and $\tilde\gamma_2 = b_2 + \beta_2$, where $\beta_2 = 0$ because $\beta_2\tilde\gamma_2 = 0$ and $b_2 \ge 0$. Hence $\tilde\gamma_2 = b_2$ and $\|\kappa - b_2 g_2\|^2 = 1 - b_2^2$. Likewise, if $\beta_2 > 0$ we obtain that $\tilde\gamma_2 = 0$ and $\tilde\gamma_1 = b_1 + \beta_1$, where $\beta_1 = 0$ because $\beta_1\tilde\gamma_1 = 0$, hence $\tilde\gamma_1 = b_1$ and $\|\kappa - b_1 g_1\|^2 = 1 - b_1^2$. Since $b_2 < b_1$, we have thus proved that $\tilde\kappa = b_1 g_1$ always.

Numerical values for $a = 1$ are

$$\begin{gathered} \mu = 1.210\,,\;\; b_1 = 0.798\,,\;\; b_2 = 0.380\,,\;\; c = 0.826 \\ \bar\gamma_1 = 1.525\,,\;\; \bar\gamma_2 = -0.880\,,\;\; \|\bar\kappa\|^2 = 0.882\,,\;\; \|\tilde\kappa\|^2 = 0.637 \\ \|\bar\kappa\|^2 : \|\tilde\kappa\|^2 = 1.386\,,\;\; \|\tilde\kappa\|^2 : \|\bar\kappa\|^2 = .721 \end{gathered} \tag{2.31}$$

The value $.721$, to the third digit, turns out to be the minimum of $\|\tilde\kappa\|^2 / \|\bar\kappa\|^2$ with respect to $a \in (0,\infty)$. ////

2.4 Level Breakdown of $(\tilde\tau_n)$

In the setup (2.21): $\bar{\mathcal{G}} = \mathrm{lin}\,\tilde{\mathcal{G}}$, in view of (2.13), the tests $\bar\tau_n$ automatically maintain asy. level α on the left-sided extension $\bar H$ of $\bar J$ and J^0, where $\bar H \supset \tilde H$.

On the contrary, the analogue to (2.13) for extensions $\tilde{H} \supset \tilde{J}$ of J^0 and the tests $\tilde{\tau}_n$ can in general not be achieved.

Note that

$$\bar{\kappa} \neq \tilde{\kappa} \iff \exists\, g \in \tilde{\mathcal{G}} : \langle \kappa | g \rangle < \langle \tilde{\kappa} | g \rangle \tag{2.32}$$

Proposition 2.4 *Assume the convex cone $\tilde{\mathcal{G}}$ contains a tangent g_0 such that*

$$\langle \kappa | g_0 \rangle \leq 0 < \langle \tilde{\kappa} | g_0 \rangle \tag{2.33}$$

Then

$$\sup\nolimits_{\tilde{J}} \limsup\nolimits_n \int \tilde{\tau}_n \, dQ_n^n = 1 \tag{2.34}$$

Proof If $\langle \kappa | g_0 \rangle = 0$, then $(P^n_{n,t,g_0}) \in \tilde{J}\ \forall t \in (0, \infty)$. In view of (2.17), therefore, the tests $\tilde{\tau}_n$ have asy. size at least

$$\sup\nolimits_{t>0} \lim\nolimits_n \int \tilde{\tau}_n \, dP^n_{n,t,g_0} = \sup_{t>0} \Phi\Bigl(-u_\alpha + \frac{t \langle \tilde{\kappa} | g_0 \rangle}{\|\tilde{\kappa}\|}\Bigr) = 1 \tag{2.35}$$

because $\lim_{t\to\infty} t\langle \tilde{\kappa} | g_0 \rangle = \infty$ due to $\langle \tilde{\kappa} | g_0 \rangle > 0$.

In case $\langle \kappa | g_0 \rangle < 0 < \langle \tilde{\kappa} | g_0 \rangle$, a suitable convex combination g_{01} of g_0 and $\tilde{\kappa}$, since $0 < \langle \kappa | \tilde{\kappa} \rangle = \|\tilde{\kappa}\|^2$, will satisfy $\langle \kappa | g_{01} \rangle = 0 < \langle \tilde{\kappa} | g_{01} \rangle$. ////

Example 2.5 In Example 2.3, although $\bar{\kappa} \neq \tilde{\kappa}$, condition (2.33) is not fulfilled, because $b_1, b_2 > 0$, and so $\langle \kappa | g \rangle \leq 0$ can hold for $g \in \tilde{\mathcal{G}}$ only if $g = 0$.

However, in the setup of Example 2.3, have tangent g_2 be replaced by the function

$$g_3(x) = -g_3(-x) = \begin{cases} \delta & \text{if } 0 < x \leq a \\ -\eta & \text{if } a < x \end{cases} \tag{2.36}$$

with $a, \delta, \eta \in (0, \infty)$. In the appendix we show that, given any $a \in (0, \infty)$, the constants $\eta = \eta_a$ and $\delta = \delta_a$ may be determined by $\delta_a = \sigma_a \eta_a$ and

$$\eta_a^{-2} = 2\bigl(\sigma_a^2 \bigl[\Phi(a) - \tfrac{1}{2}\bigr] + \bigl[1 - \Phi(a)\bigr]\bigr), \qquad \sigma_a = a\,\frac{1 - \Phi(a)}{\varphi(0) - \varphi(a)} \tag{2.37}$$

Then $\|g_3\| = 1$ and

$$\langle \kappa | g_3 \rangle < 0 < \langle g_1 | g_3 \rangle \tag{2.38}$$

By the method of Lagrange multipliers, in the appendix, we prove that

$$\tilde{\kappa} = \langle \kappa | g_1 \rangle \, g_1 \tag{2.39}$$

where $\langle \kappa | g_1 \rangle = 2\varphi(0) > 0$, and so $\langle \kappa | g_3 \rangle < 0 < \langle \kappa | g_1 \rangle \langle g_1 | g_3 \rangle = \langle \tilde{\kappa} | g_3 \rangle$, which implies (2.33) for $g_0 = g_3$. ////

Making use of the following uniqueness result (Proposition 2.7), we conclude that testing the slightly bigger null hypothesis $\tilde{J} \supset J^0$, or the even bigger one-sided extension $\tilde{H}$ of $\tilde{J}$, vs. $\tilde{K}$, is inevitably bound to larger error probabilities than those given in Theorem 2.1(a) for testing J^0 vs. $\tilde{K}$. This is contrary to the extension of J^0 to $\bar{J}$ and $\bar{H}$, vs. $\bar{K}$, which goes for free in Theorem 2.1(b).

Remark 2.6 The minimum asy. power $\Phi\big(-u_\alpha + c/\|\bar\kappa\|\big)$ achieved by the asy. test $(\bar\tau_n)$ under $\bar K$ stays the same under $\tilde K \subset \bar K$, that is, does not increase,

$$\inf_{\tilde K} \lim_n \int \bar\tau_n \, dQ_n^n = \Phi\Big(-u_\alpha + \frac{c}{\|\bar\kappa\|}\Big) \tag{2.40}$$

Indeed, pick any $g \in \tilde{\mathcal G}$ such that $\langle\kappa|g\rangle > 0$; for example, $g = \tilde\kappa$ itself. Then choose $t \in (0,\infty)$ such that $t\langle\kappa|g\rangle = c$, and apply (2.7) and (2.19).

Whether $\Phi\big(-u_\alpha + c/\|\bar\kappa\|\big)$ is the largest minimum asy. power that can be achieved vs. $\tilde K$, subject to asy. level α under $\tilde H$, respectively only under $\tilde J$, is unknown. In particular, we do not know if there exists some function $\eta \in L_2(P)$ of smaller norm $\|\eta\| < \|\bar\kappa\|$ and such that, for each $g \in \tilde{\mathcal G}$,

$$\left.\begin{array}{r} \tilde J : \langle\kappa|g\rangle = 0 \\ \tilde H : \langle\kappa|g\rangle \le 0 \end{array}\right\} \implies \langle\eta|g\rangle \le 0\,, \qquad \langle\kappa|g\rangle > 0 \implies \langle\eta|g\rangle \ge \langle\kappa|g\rangle \tag{2.41}$$

In connection with asy. median unbiased, two-sided confidence limits for cones, the corresponding function η cannot exist; confer Subsection 3.4, where instead of (2.41) the simpler condition (3.63) occurs. ////

2.5 Uniqueness of Most Powerful Tests

In the setup of Theorem 2.1, the optimal tests $\tilde\tau_n$ and $\bar\tau_n$ defined by (2.10) and (2.12), respectively, are unique up to terms $\mathrm{o}_{P^n}(n^0)$ tending stochastically to zero under P^n.

Proposition 2.7 *Suppose that an asy. test* (τ_n) *satisfies* (2.8), *and achieves the asy. power bound* (2.9) *in case* (a), *respectively bound* (2.11) *in case* (b). *Then necessarily*

$$\tau_n = \begin{cases} \tilde\tau_n + \mathrm{o}_{P^n}(n^0) & \textit{in case (a), respectively} \\ \bar\tau_n + \mathrm{o}_{P^n}(n^0) & \textit{in case (b).} \end{cases} \tag{2.42}$$

Conversely, form (2.42) *implies that the asy. test* (τ_n) *satisfies* (2.8) *and achieves bound* (2.9), *respectively satisfies* (2.13) *and achieves bound* (2.11).

PROOF In regard of the proof to Theorem 2.1, the proposition is a straightforward consequence of the uniqueness result in HR (1994; Corollary 3.4.2 b with $\sigma > 0$)—under the provision however that $\tilde\kappa \in \tilde{\mathcal G}$ and $\bar\kappa \in \bar{\mathcal G}$, respectively. Since, in general, the tangent set $\mathcal G$ (convex cone or linear space) needs not be closed, we have to incorporate an approximation in $L_2(P)$ of $\tilde\kappa$ and $\bar\kappa$ by elements of $\tilde{\mathcal G}$ and $\bar{\mathcal G}$, respectively. This is the reason for the following proof.

Thus, given any $t \in (0,\infty)$ and $h \in L_2(P)$, $h \ne 0$, $\mathrm{E}\, h = 0$, we shall show that (2.8) and

$$\liminf_{(s,g)\to(t,h)} \liminf_n \int \tau_n \, dP_{n,s,g}^n \ge \Phi\big(-u_\alpha + t\|h\|\big) \tag{2.43}$$

imply that

$$\tau_n = \mathbf{I}\big(\sqrt{n}\,\mathrm{ave}_1^n\, h(x_i) > \|h\|\, u_\alpha\big) + \mathrm{o}_{P^n}(n^0) \tag{2.44}$$

In proving this, it is no restriction to set $s = t = 1$, and then delete s and t from notation; in particular, we write $P_{n,s,g} = P_{n,g}$.

Given any $\delta \in (0,1)$, $\delta < \|h\|$, choose g so close to h that

$$\|g-h\|^2 < \delta^3\,,\ \big|\|g\|^2 - \|h\|^2\big| < 2\delta \quad\text{and}\quad |\beta_g - \beta_h| < \delta\,,\ |\ell_g - \ell_h| < \delta \tag{2.45}$$

for the norm based quantities $\beta_g = \Phi\bigl(-u_\alpha + \|g\|\bigr)$ and $\ell_g = \|g\|u_\alpha - \frac{1}{2}\|g\|^2$, and such that, making use of (2.43), moreover

$$\liminf_n \int \tau_n\, dP^n_{n,g} \ge \beta_h - \delta \tag{2.46}$$

The proof employs the following Neyman–Pearson tests $\tau^*_{n,g}$ for P^n vs. $P^n_{n,g}$,

$$\tau^*_{n,g} = \mathbf{I}(L_{n,g} > \ell_g)\,, \qquad L_{n,g} = \log dP^n_{n,g}/dP^n \tag{2.47}$$

As the loglikelihoods $L_{n,g}$ are asy. $\mathcal{N}\bigl(-\frac{1}{2}\|g\|^2, \|g\|^2\bigr)$ under P^n,

$$\alpha_n = \int \tau^*_{n,g}\, dP^n \longrightarrow \alpha\,, \qquad \beta_n = \int \tau^*_{n,g}\, dP^n_{n,g} \longrightarrow \beta_g \tag{2.48}$$

By (2.8), (2.45), and (2.46), some $n_0 = n_0(\delta)$ exists such that for all $n \ge n_0$,

$$\int \tau_n\, dP^n \le \alpha_n + 3\delta\,, \qquad \int \tau_n\, dP^n_{n,g} \ge \beta_n - 3\delta \tag{2.49}$$

Then Lemma 4.1 tells us that, for all such $n \ge n_0$ and for every $\varepsilon \in (0,1)$,

$$|\nu_{n,g}|\bigl\{|\tau_n - \tau^*_{n,g}| > \varepsilon\bigr\} \le 3\,(1 + c_g)\,\frac{\delta}{\varepsilon} \tag{2.50}$$

where

$$\nu_{n,g} = P^n_{n,g} - c_g P^n\,, \qquad c_g = e^{\ell_g} \tag{2.51}$$

Fix $\varepsilon \in (0,1)$ and set $A_{n,g} = \bigl\{|\tau_n - \tau^*_{n,g}| > \varepsilon\bigr\}$. Fix any $\rho \in (0,1)$. Then the probability $P^n\bigl(A_{n,g} \cap \{L_{n,g} > \ell_g + \rho\}\bigr)$ is bounded above by

$$\frac{1}{e^{(\ell_g+\rho)} - e^{\ell_g}} \int_{A_{n,g}\cap\{L_{n,g}>\ell_g+\rho\}} \bigl(e^{L_{n,g}} - e^{\ell_g}\bigr)\, dP^n \le \frac{|\nu_{n,g}|(A_{n,g})}{c_g\,(e^\rho - 1)} \tag{2.52}$$

Likewise, $P^n\bigl(A_{n,g} \cap \{L_{n,g} < \ell_g - \rho\}\bigr)$ is bounded above by

$$\frac{1}{e^{\ell_g} - e^{(\ell_g-\rho)}} \int_{A_{n,g}\cap\{L_{n,g}<\ell_g-\rho\}} \bigl(e^{\ell_g} - e^{L_{n,g}}\bigr)\, dP^n \le \frac{|\nu_{n,g}|(A_{n,g})}{c_g\,(1 - e^{-\rho})} \tag{2.53}$$

Put $\eta_\rho = (e^\rho + 1)/(e^\rho - 1)$ and use $|\ell_g - \ell_h| < \delta$, hence $c_g > e^{-\delta}c_h$, to conclude that

$$\begin{aligned} P^n(A_{n,g}) &\le 3\,\eta_\rho\,(1 + c_g^{-1})\,\frac{\delta}{\varepsilon} + P^n\bigl\{|L_{n,g} - \ell_g| \le \rho\bigr\} \\ &\le 3\,\eta_\rho\,(1 + e^\delta c_h^{-1})\,\frac{\delta}{\varepsilon} + P^n\bigl\{|L_{n,g} - \ell_g| \le \rho\bigr\} \end{aligned} \tag{2.54}$$

Asy. normality of $L_{n,g}$ under P^n, and (2.45) ensuring $\|g\| \ge \|h\| - \delta$, imply

$$\lim_n P^n\bigl\{|L_{n,g} - \ell_g| \le \rho\bigr\} \le 2\rho\,\frac{\varphi(0)}{\|g\|} \le 2\rho\,\frac{\varphi(0)}{\|h\| - \delta} \tag{2.55}$$

It follows that, for all $\delta \in (0,1)$, $\delta < \|h\|$, and for all $\rho \in (0,1)$,

$$\limsup_{g\to h} \limsup_n P^n(A_{n,g}) \le 3\,\eta_\rho\,(1 + e^\delta c_h^{-1})\,\frac{\delta}{\varepsilon} + 2\rho\,\frac{\varphi(0)}{\|h\| - \delta} \tag{2.56}$$

Hence

$$\lim_{g\to h} \limsup_n P^n\bigl\{|\tau_n - \tau^*_{n,g}| > \varepsilon\bigr\} = 0 \tag{2.57}$$

if we first let δ and then ρ approach 0 in (2.56).

Furthermore, comparing the Neyman–Pearson tests $\tau^*_{n,g}$ and $\tau^*_{n,h}$, we get

$$\begin{aligned} P^n\bigl\{|\tau^*_{n,g} - \tau^*_{n,h}| > \varepsilon\bigr\} \le P^n\bigl\{L_{n,g} > \ell_g\,,\ L_{n,h} < \ell_h - 4\delta\bigr\} \\ + P^n\bigl\{L_{n,g} \le \ell_g\,,\ L_{n,h} > \ell_h + 4\delta\bigr\} \\ + P^n\bigl\{|L_{n,h} - \ell_h| \le 4\delta\bigr\} \end{aligned} \tag{2.58}$$

The 3rd summand on the RHS, by the asy. normality of $L_{n,h}$ under P^n, satisfies

$$\lim_n P^n\bigl\{|L_{n,h} - \ell_h| \le 4\delta\bigr\} \le 8\delta\,\frac{\varphi(0)}{\|h\|} \tag{2.59}$$

The first two summands on the RHS in (2.58), since $|\ell_g - \ell_h| < \delta$, are bounded by $P^n\bigl\{|L_{n,g} - L_{n,h}| > 3\delta\bigr\}$. Invoke the loglikelihood expansion (2.14) and make use of $\bigl|\|g\|^2 - \|h\|^2\bigr| < 2\delta$ in order to bound $P^n\bigl\{|L_{n,g} - L_{n,h}| > 3\delta\bigr\}$ by

$$P^n\Bigl\{\bigl|\sqrt{n}\,\mathrm{ave}_1^n\,(g-h)(x_i)\bigr| > 2\delta - \mathrm{o}_{P^n}(n^0)\Bigr\} \tag{2.60}$$

which, in turn, is bounded by some $\mathrm{o}(n^0)$ plus

$$P^n\bigl\{\bigl|\sqrt{n}\,\mathrm{ave}_1^n\,(g-h)(x_i)\bigr| > \delta\bigr\} \le \frac{\|g-h\|^2}{\delta^2} \le \frac{\delta^3}{\delta^2} = \delta \tag{2.61}$$

This implies

$$\limsup_n P^n\bigl\{|\tau^*_{n,g} - \tau^*_{n,h}| > \varepsilon\bigr\} \le 8\delta\,\frac{\varphi(0)}{\|h\|} + \delta \tag{2.62}$$

hence

$$\lim_{g\to h} \limsup_n P^n\bigl\{|\tau^*_{n,g} - \tau^*_{n,h}| > \varepsilon\bigr\} = 0 \tag{2.63}$$

Observe that $\limsup_n P^n\bigl\{|\tau_n - \tau^*_{n,h}| > 2\varepsilon\bigr\}$ does not depend on g, therefore, may be bounded by

$$\begin{aligned} \lim_{g\to h} \limsup_n P^n\bigl\{|\tau_n - \tau^*_{n,g}| > \varepsilon\bigr\} + P^n\bigl\{|\tau^*_{n,g} - \tau^*_{n,h}| > \varepsilon\bigr\} \\ \le \lim_{g\to h} \limsup_n P^n\bigl\{|\tau_n - \tau^*_{n,g}| > \varepsilon\bigr\} \\ + \lim_{g\to h} \limsup_n P^n\bigl\{|\tau^*_{n,g} - \tau^*_{n,h}| > \varepsilon\bigr\} \end{aligned} \tag{2.64}$$

The upper bound equals zero by (2.57) and (2.63); thus,

$$\limsup_n P^n\bigl\{|\tau_n - \tau^*_{n,h}| > 2\varepsilon\bigr\} = 0 \tag{2.65}$$

It remains to prove that

$$\tau^*_{n,h} = \tau^\star_n + \mathrm{o}_{P^n}(n^0) \qquad \text{for} \quad \tau^\star_n = \mathbf{I}\bigl(\sqrt{n}\,\mathrm{ave}_1^n\, h(x_i) > \|h\|\, u_\alpha\bigr) \tag{2.66}$$

But $\sqrt{n}\,\mathrm{ave}_1^n\, h(x_i)$ is asy. normal $\mathcal{N}\bigl(0, \|h\|^2\bigr)$ and $\mathcal{N}\bigl(\|h\|^2, \|h\|^2\bigr)$ under P^n, respectively $P^n_{n,h}$, so that

$$\lim_n \int \tau^\star_n \, dP^n = \alpha\,, \qquad \lim_n \int \tau^\star_n \, dP^n_{n,h} = \Phi\bigl(-u_\alpha + \|h\|\,\bigr) \tag{2.67}$$

Thus, the uniqueness result of HR (1994; Corollary 3.4.2[4]) applies to $(\tau^\star_n)$, such that $\tau^\star_n = \tau^*_{n,h} + \mathrm{o}_{P^n}(n^0)$. Altogether, (2.65) and (2.66) imply (2.44).

The converse, that (2.42) entails optimality, is obvious, as all sequences (Q^n_n) in $H \cup K$ are contiguous to (P^n). ////

2.6 Invariant Tangent Cones and Spaces

Rank Functionals For the symmetry problem on the real line, one-sample rank functionals R_ϱ are given by

$$R_\varrho(Q) = 2\int_0^\infty \varrho\bigl(Q(x) - Q(-x)\bigr)\, Q(dx) - \int_0^1 \varrho\, d\lambda_0 \tag{2.68}$$

where $Q(x) = Q\bigl((-\infty, x]\bigr)$, λ_0 denotes Lebesgue measure on $(0,1)$, and ϱ is some (scores) function in $L_1(\lambda_0)$. Then $R_\varrho(Q)$ is defined for every $Q \in \mathcal{M}_c$, the set of all probabilities with continuous distribution functions. Let $\mathcal{M}_{cs}$ denote the subset of all symmetric $P \in \mathcal{M}_c$ (that is, $P(-x) = 1 - P(x)\ \ \forall\, x > 0$). Then $R_\varrho(P) = 0$ for all $P \in \mathcal{M}_{cs}$. A certain kind of asymmetry is defined through nonzero values of the functional. If ϱ is nonnegative increasing, then $R_\varrho(Q) \ge 0$ for all positively asymmetric $Q \in \mathcal{M}_c$ (that is, $Q(-x) \le 1 - Q(x)\ \ \forall\, x > 0$); more generally, $R_\varrho(Q'') \ge R_\varrho(Q')$ if $Q', Q'' \in \mathcal{M}_c$, $Q''(x) \le Q'(x)\ \ \forall\, x \in \mathbb{R}$.

Signed Linear Rank Statistics Linear rank statistics R_n are of the form

$$R_n = \mathrm{ave}_{i=1}^n\, \mathrm{sign}(x_i)\, \varrho_n(r^+_{n,i}) \tag{2.69}$$

where $r^+_{n,i}$ denote the absolute ranks (rank $|x_i|$ among $|x_1|, \dots, |x_n|$), and $\varrho_n(i)$ are some numbers (scores). The weak condition used by Hájek and Sidák (1967; V.1.7) to prove asy. normality of R_n under P^n (in fact, asy. linearity at P) is

$$\varrho_n([1 + ns]) \longrightarrow \varrho(s) \qquad \text{in } L_2(\lambda_0) \tag{2.70}$$

[4]Note that $\sigma > 0$ must be assumed in part (b).

Given any $\varrho \in L_2(\lambda_0)$, this condition is satisfied by the array $\varrho_n(i) = \mathrm{E}\,\varrho(u_{n(i)})$ (based on the order statistics $u_{n(i)}$ of an i.i.d. sample $u_1, \ldots, u_n \sim \lambda_0$), by the array $\varrho_n(i) = n\int_{I_n} \varrho\, d\lambda_0$ with $I_n = (\frac{i-1}{n}, \frac{i}{n})$, and the array $\varrho_n(i) = \varrho(\frac{i}{n+1})$ (under a mild extra condition on ϱ). Then, for every $P \in \mathcal{M}_{cs}$, the sequence of rank statistics (R_n) is asy. linear at P with influence curve κ_P,

$$R_n = \mathrm{ave}_{i=1}^n\, \kappa_P(x_i) + \mathrm{o}_{P^n}(1/\sqrt{n}\,) \tag{2.71}$$

where

$$\kappa_P(x) = \mathrm{sign}(x)\, \varrho\big(2\,P(|x|) - 1\big) \tag{2.72}$$

An alternative approach imposes bounds on the growth of the derivative(s) of the scores function ϱ; confer Hájek and Sidák (1967; VI.5.1). These Chernoff–Savage conditions have successively been weakened and ensure the asy. normality of $\sqrt{n}\,\big(R_n - R_\varrho(Q_n)\big)$, even under noncontiguous alternatives (Q_n^n), with R_ϱ as centering functional. Combining both sets of conditions, differentiablity of R_ϱ at $P \in \mathcal{M}_{cs}$ may be proved as in HR (1981 a; Proposition 4.1). Thus, at every $P \in \mathcal{M}_{cs}$, the functional R_ϱ is differentiable in the sense of (1.4) with influence curve the same κ_P given by (2.72).

Invariant Tangent Sets and Hypotheses Rank statistics R_n are not only distribution free under the null hypothesis $\mathcal{M}_{cs}$ but also under suitably defined alternatives. Let a family of sets $\mathcal{G}_P$, one for each $P \in \mathcal{M}_{cs}$, be generated by some set $\mathcal{G}_0 \subset L_2(\lambda_0)$ such that

$$\mathcal{G}_P = \big\{\, g_{P,q} \;\big|\; q \in \mathcal{G}_0 \,\big\}\,, \qquad g_{P,q}(x) = \mathrm{sign}(x)\, q\big(2\,P(|x|) - 1\big) \tag{2.73}$$

These sets $\mathcal{G}_P$, which obviously consist of odd functions, are invariant in the sense that the composition $\mathcal{G}_P \circ P^{-1} = \{\, g \circ P^{-1} \mid g \in \mathcal{G}_P\}$ with the pseudo-inverse $P^{-1}(s) = \inf\{x \in \mathbb{R} \mid P(x) \ge s\}$ is the same for all $P \in \mathcal{M}_{cs}$,

$$\mathcal{G}_P \circ P^{-1} = \big\{\, g_{0,q} \;\big|\; q \in \mathcal{G}_0 \,\big\}\,, \qquad g_{0,q}(s) = \mathrm{sign}(s - \tfrac{1}{2})\, q\big(|2s - 1|\big) \tag{2.74}$$

As $\int g_{P,q}\, dP = 0$ and $\int g_{P,q}^2\, dP = \int q^2\, d\lambda_0$, the sets $\mathcal{G}_P$ may actually serve as tangent sets at $P \in \mathcal{M}_{cs}$. Moreover, the properties of $\mathcal{G}_0$ to be closed, convex, a cone, a linear subspace of $L_2(\lambda_0)$, respectively, are each inherited to the sets $\mathcal{G}_P$ in $L_2(P)$ for every $P \in \mathcal{M}_{cs}$.

Remark 2.8 Conversely, given any set $\mathcal{G}_{P_0}$ of odd tangents at some $P_0 \in \mathcal{M}_{cs}$, define

$$\mathcal{G}_0 = \big\{\, q_g \;\big|\; g \in \mathcal{G}_{P_0} \,\big\}\,, \qquad q_g(s) = g\big(P_0^{-1}(\tfrac{1+s}{2})\big) \tag{2.75}$$

Then this set $\mathcal{G}_0$, via (2.73), reproduces the given tangent set $\mathcal{G}_{P_0}$ at P_0 and generates the following tangent sets $\mathcal{G}_P$ at other measures $P \in \mathcal{M}_{cs}$,

$$\mathcal{G}_P = \big\{\, g \circ P_0^{-1} \circ P \;\big|\; g \in \mathcal{G}_{P_0} \,\big\} \tag{2.76}$$

where $g \circ P_0^{-1}\big(P(x)\big) = \mathrm{sign}(x)\, g \circ P_0^{-1}\big(P(|x|)\big)$ a.e. $P(dx)$. Note that $P_0^{-1} \circ P$ is odd and strictly increasing a.e. P. For such tranformations applied to each x_i, the vector of signs and absolute ranks is (maximal) invariant.

Positive shifts, for example, of some $P_0 \in \mathcal{M}_{cs}$ which has finite Fisher information of location and a Lebesgue density p_0, lead to the tangent cone generated by the function $-(\dot{p}_0/p_0) = g_{P_0,q_0}$, where $q_0(s) = -(\dot{p}_0/p_0) \circ P_0^{-1}(\frac{1+s}{2})$, and then $g_{P_0,q_0} \circ P_0^{-1}\bigl(P(x)\bigr) = -\operatorname{sign}(x)\,(\dot{p}_0/p_0) \circ P_0^{-1}\bigl(P(|x|)\bigr)$ a.e. $P(dx)$. ////

Now suppose that $\mathcal{G}_0$ is (a) a convex cone, or (b) a linear space, in $L_2(\lambda_0)$. For each $P \in \mathcal{M}_{cs}$, let the hypotheses J_P^0, J_P, H_P, and K_P about the rank functional R_ϱ over the tangent set $\mathcal{G}_P$ be defined by (2.1)–(2.4). These hypotheses are invariant as they read

$$J_P^0 : q = 0\,, \quad J_P : \langle \varrho | q \rangle_0 = 0\,, \quad H_P : \langle \varrho | q \rangle_0 \le 0\,, \quad K_P : t\langle \varrho | q \rangle_0 \ge c \tag{2.77}$$

with reference to the tangent set $\mathcal{G}_P$ given by (2.73) at $P \in \mathcal{M}_{cs}$. In view of (2.7), representation (2.77) is a consequence of the following equality of scalar products and norms in $L_2(P)$ and $L_2(\lambda_0)$, respectively, for the tangents of form (2.73),

$$\langle \kappa_P | g \rangle_P = \langle \varrho | q \rangle_0\,, \qquad \| \kappa_P - g \|_P^2 = \| \varrho - q \|_0^2 \tag{2.78}$$

Invariant Optimality of Rank Tests As another consequence of (2.78) we observe that the approximation of κ_P by $g \in \mathcal{G}_P$ is equivalent to the approximation of ϱ by $q \in \mathcal{G}_0$. Therefore, the projection $\hat{\kappa}_P$ of κ_P on $c\ell\,\mathcal{G}_P$ in $L_2(P)$ is given in terms of the projection $\hat{\varrho}$ of ϱ on $c\ell\,\mathcal{G}_0$ in $L_2(\lambda_0)$,

$$\hat{\kappa}_P(x) = \operatorname{sign}(x)\,\hat{\varrho}\bigl(2\,P(|x|) - 1\bigr) \tag{2.79}$$

Then Theorem 2.1 is in force and yields the optimal asy. level α test sequence $(\hat{\tau}_{n,P})$ for J_P^0 vs. K_P,

$$\hat{\tau}_{n,P} = \mathbf{I}\bigl(\sqrt{n}\,\operatorname{ave}_{i=1}^n \hat{\kappa}_P(x_i) > \|\hat{\kappa}_P\|_P\, u_\alpha\bigr) \tag{2.80}$$

Now invoke any array of scores $\hat{\varrho}_n(i)$ that, via (2.70), are connected to $\hat{\varrho}$. Employ the corresponding rank statistics $\hat{R}_n$ to define the rank tests

$$\hat{\tau}_n = \mathbf{I}\bigl(\sqrt{n}\,\hat{R}_n > \|\hat{\varrho}\|_0\, u_\alpha\bigr) \tag{2.81}$$

independently of $P \in \mathcal{M}_{cs}$. Then, by (2.71), (2.72) for $\hat{R}_n$ and $\hat{\kappa}_P$, $\hat{\varrho}$, and by asy. normality,

$$\hat{\tau}_n = \hat{\tau}_{n,P} + \mathrm{o}_{P^n}(n^0) \tag{2.82}$$

for every $P \in \mathcal{M}_{cs}$. Thus, the sequence $(\hat{\tau}_n)$ of rank tests (2.81) is optimal for J_P^0—if $\mathcal{G}_0 = \operatorname{lin}\mathcal{G}_0$ even for H_P—against K_P, according to Theorem 2.1.

This optimality, in the two cases (a) $\mathcal{G}_0$ a convex cone, (b) $\mathcal{G}_0$ a linear space, holds true for every $P \in \mathcal{M}_{cs}$.

3 Confidence Limits

Let P be any element of $\mathcal{P}$, with tangent set $\mathcal{G} \subset L_2(P) \cap \{1\}^\perp$, and some constant $c \in (0,\infty)$. Similarly to the testing whether $T(Q) \ge T(P) + c/\sqrt{n}$, we now consider lower confidence limits $S_n - c/\sqrt{n}$ for the value $T(P)$; for example, the minimum amount of cash to be kept on a business account. Here and subsequently, the estimator sequence (S_n) may be any sequence of estimates S_n at sample size n. It is desirable that S_n underestimate $T(P) + c/\sqrt{n}$ with highest possible probability, under the i.i.d. observations $x_1, \ldots, x_n \sim P$. This aim, however, is not well-defined, as shown by arbitrary estimates $S_n \le T(P)$. Therefore, a side condition that also $S_n \ge T(P)$ with sufficiently high probability must be imposed. In addition, to cut out $S_n \equiv T(P)$, a local variation of P must be employed.

3.1 Confidence Bounds For Lower and Upper Limits

The following result requires some one-sided, respectively two-sided, asymptotic median unbiasedness under the local perturbations $P_{n,t,g}$ of P of kind (2.5), and is of the type intended by Pfanzagl and Wefelmeyer (1982; Theorem 9.2.2).

Qualitatively speaking, Theorem 3.1(a) bounds any 'limit distribution function' of $\sqrt{n}\,\bigl(S_n - T(Q)\bigr)$ under $Q = P$, subject to upper bound 1/2 at the origin under all $Q = P_{n,t,\hat\kappa}$, on the positve half-line by that of $\mathcal{N}\bigl(0, \|\hat\kappa\|^2\bigr)$ from above. In addition, Theorem 3.1(b) bounds such 'limit distribution functions' under $Q = P$, subject to the lower bound 1/2 at the origin under all $Q = P_{n,t,-\hat\kappa}$, at the same time on the negative half-line by that of $\mathcal{N}\bigl(0, \|\hat\kappa\|^2\bigr)$ from below; where $\hat\kappa = \tilde\kappa, \bar\kappa$, respectively. For best estimator accuracy, the limit distribution function should be maximal on $(0,\infty)$, and minimal on $(-\infty, 0)$. In general, 'limit distribution functions' need not exist nor need they be normal.

An estimator sequence $(\hat S_n)$ is said to attain a confidence upper bound

$$\limsup\nolimits_n P^n\bigl\{-t' < \sqrt{n}\,\bigl(S_n - T(P)\bigr) < t''\bigr\} \le \beta(t', t'') \tag{3.1}$$

uniformly in t', t'', and with $\limsup_n$ replaced by $\liminf_n$, if $(\hat S_n)$ itself satisfies the side conditions on the estimator sequences (S_n) and in fact achieves

$$\liminf\nolimits_n \inf\nolimits_{t',t''} \Bigl(P^n\bigl\{-t' < \sqrt{n}\,\bigl(\hat S_n - T(P)\bigr) < t''\bigr\} - \beta(t', t'')\Bigr) \ge 0 \tag{3.2}$$

As for asy. linear estimators, the reader is referred to the beginning of Subsection 3.3, where this kind of estimators are introduced in more generality.

Theorem 3.1 *Let (S_n) be any estimator sequence.*

(a) *Suppose $\mathcal{G} = \tilde{\mathcal{G}}$, a convex cone. Assume there exists some sequence of tangents $g_m \in \tilde{\mathcal{G}}$ such that $g_m \to \tilde\kappa$ in $L_2(P)$ and, for every convergent sequence $t_n \to t$ in $(0,\infty)$,*

$$\liminf\nolimits_m \liminf\nolimits_n P^n_{n,t_n,g_m}\bigl\{S_n \ge T(P_{n,t_n,g_m})\bigr\} \ge \tfrac{1}{2} \tag{3.3}$$

Then, for every $t \in (0,\infty)$ and every convergent sequence $t_n \to t$ in $(0,\infty)$,

$$\limsup\nolimits_n P^n\bigl\{\sqrt{n}\,\bigl(S_n - T(P)\bigr) < t_n\bigr\} \le \Phi\Bigl(\frac{t}{\|\tilde\kappa\|}\Bigr) \tag{3.4}$$

The upper bound (3.4) *is attained by the asy. linear estimator* $(\tilde{S}_n)$,

$$\sqrt{n}\,\bigl(\tilde{S}_n - T(P)\bigr) = \sqrt{n}\,\mathrm{ave}_{i=1}^{n}\,\tilde{\kappa}(x_i) + \mathrm{o}_{P^n}(n^0) \tag{3.5}$$

which achieves (3.4) *with* $t_n = t$, *uniformly in* $-\infty \le t \le \infty$, *and with* $\limsup_n$ *replaced by* $\liminf_n$.

(b) *Suppose* $\mathcal{G} = \bar{\mathcal{G}}$, *a linear space. Assume there exist two sequences of tangents* $g_m', g_m'' \in \bar{\mathcal{G}}$ *such that* $g_m' \to \bar{\kappa}$, $g_m'' \to -\bar{\kappa}$ *in* $L_2(P)$ *and, for every convergent sequence* $t_n \to t$ *in* $(0,\infty)$,

$$\liminf_m \liminf_n P^n_{n,t_n,g_m'}\bigl\{S_n \ge T(P_{n,t_n,g_m'})\bigr\} \ge \tfrac{1}{2} \tag{3.6}$$

$$\liminf_m \liminf_n P^n_{n,t_n,g_m''}\bigl\{S_n \le T(P_{n,t_n,g_m''})\bigr\} \ge \tfrac{1}{2} \tag{3.7}$$

Then, for every $t', t'' \in (0,\infty)$ *and all sequences* $t_n' \to t'$, $t_n'' \to t''$ *in* $(0,\infty)$,

$$\limsup_n P^n\bigl\{-t_n' < \sqrt{n}\,\bigl(S_n - T(P)\bigr) < t_n''\bigr\} \le \Phi\Bigl(\frac{t''}{\|\bar{\kappa}\|}\Bigr) - \Phi\Bigl(-\frac{t'}{\|\bar{\kappa}\|}\Bigr) \tag{3.8}$$

The upper bound (3.8) *is attained by the asy. linear estimator* $(\bar{S}_n)$,

$$\sqrt{n}\,\bigl(\bar{S}_n - T(P)\bigr) = \sqrt{n}\,\mathrm{ave}_{i=1}^{n}\,\bar{\kappa}(x_i) + \mathrm{o}_{P^n}(n^0) \tag{3.9}$$

which achieves (3.8) *with* $t_n' = t'$, $t_n'' = t''$, *uniformly in* $-\infty \le -t' < t'' \le \infty$, *and with* $\limsup_n$ *replaced by* $\liminf_n$.

Remark 3.2 [asymptotic median nonnegative, nonpositive]

Conditions (3.3), (3.6), and (3.7), respectively, mean that—in the iterated limit—the median of $\sqrt{n}\,\bigl(S_n - T(P_{n,t_n,g})\bigr)$ under $P^n_{n,t_n,g}$ for n large, $g \approx \tilde{\kappa}$, and $g \approx \bar{\kappa}$, respectively, becomes ≥ 0, and ≤ 0 for $g \approx -\bar{\kappa}$.

Of course, if $\tilde{\kappa} \in \tilde{\mathcal{G}}$, respectively $\bar{\kappa} \in \bar{\mathcal{G}}$, conditions (3.3) and (3.6), (3.7) are needed only for $g_m = \tilde{\kappa}$, respectively for $g_m' = \bar{\kappa}$ and $g_m'' = -\bar{\kappa}$.

Conditions (3.3), (3.6) and (3.7), respectively, are ensured by asymptotic median nonnegativity and nonpositivity, respectively, for every fixed tangent in the corresponding tangent set $\mathcal{G}$, in the sense of (3.48) and (3.49) below. ////

PROOF We start the derivation of the bounds simultaneously in both cases:

Fix any $g \in \mathcal{G}$ such that $\langle\kappa|g\rangle \ne 0$, any sequence $t_n \to t$ in $(0,\infty)$, and put $P_n = P_{n,t_n,g}$. Expansion (2.6), by (1.4), holds uniformly on t-compacts, so

$$\sqrt{n}\,\bigl(T(P_n) - T(P)\bigr) = t\,\langle\kappa|g\rangle + \mathrm{o}(n^0) = s_n\langle\kappa|g\rangle \tag{3.10}$$

for some suitable other sequence $s_n = s_{n,t_n,g} \to t$. Thus, we obtain

$$\sqrt{n}\,\bigl(S_n - T(P_n)\bigr) = R_n - s_n\langle\kappa|g\rangle \qquad \text{for} \qquad R_n = \sqrt{n}\,\bigl(S_n - T(P)\bigr) \tag{3.11}$$

Also the loglikelihood expansion (2.14) for fixed g, due to (1.1), holds uniformly on t-compacts. Therefore, and by mutual contiguity of (P_n^n) and (P^n),

$$\log\frac{dP^n}{dP_n^n} = -\log\frac{dP_n^n}{dP^n} + \mathrm{o}_n' = -t\,\sqrt{n}\,\mathrm{ave}_{i=1}^{n}\,g(x_i) + \tfrac{1}{2}t^2\,\|g\|^2 + \mathrm{o}_n'' \tag{3.12}$$

where o_n', o_n'' each are some $\mathrm{o}_{P^n}(n^0)$. By HR (1994; Proposition 2.2.12 and Corollary 3.4.2 a), the asy. power of any test sequence (τ_n) under (P^n), subject to asy. level α under (P_n^n), is bounded by $\Phi\bigl(-u_\alpha + t\|g\|\bigr)$.

Applying this bound to the sequence of tests

$$\tau_n = \mathbf{I}(R_n < s_n\langle\kappa|g\rangle) = \tau_{n,t_n,g} \tag{3.13}$$

and their asy. level

$$\alpha_g = \limsup_n P_n^n\bigl\{R_n < s_n\langle\kappa|g\rangle\bigr\} \tag{3.14}$$

we obtain

$$\Phi\bigl(-u_{\alpha_g} + t\|g\|\bigr) \ge \limsup_n P^n\bigl\{R_n < s_n\langle\kappa|g\rangle\bigr\} \tag{3.15}$$

(a) Observe that, by condition (3.3), as $g = g_m \in \tilde{\mathcal{G}}$ tends to $\tilde\kappa$ in $L_2(P)$,

$$\limsup \alpha_g \le \tfrac{1}{2}\,, \quad \text{hence} \quad \liminf u_{\alpha_g} \ge 0 \tag{3.16}$$

Therefore, given $\delta \in (0,1)$, one can choose $g = g_m \in \tilde{\mathcal{G}}$ so close to $\tilde\kappa$ that

$$-u_{\alpha_g} + t\,\|g\| \le t\,\|\tilde\kappa\| + \delta \qquad \text{and} \qquad s_n\langle\kappa|g\rangle \ge (t_n - \delta)\|\tilde\kappa\|^2 \tag{3.17}$$

eventually. Then (3.15) implies that

$$\limsup_n P^n\bigl\{R_n < (t_n - \delta)\|\tilde\kappa\|^2\bigr\} \le \Phi\bigl(t\,\|\tilde\kappa\| + \delta\bigr) \tag{3.18}$$

hence

$$\limsup_n P^n\bigl\{R_n < t_n\|\tilde\kappa\|^2\bigr\} \le \Phi\bigl(t\,\|\tilde\kappa\| + \delta\,\|\tilde\kappa\| + \delta\bigr) \tag{3.19}$$

where assumption (3.3) has been used for the shifted sequence $t_n+\delta$. Once more using (3.3) for the rescaled sequence $t_n/\|\tilde\kappa\|^2$, bound (3.4) follows from (3.19), if we let $\delta \to 0$.

(b) Starting from assumption (3.6), the proof (a) establishes the bound

$$\limsup_{n\to\infty} P^n\bigl\{R_n < t_n''\,\|\bar\kappa\|^2\bigr\} \le \Phi\bigl(t''\,\|\bar\kappa\|\bigr) \tag{3.20}$$

for every $t'' \in (0,\infty)$ and every convergent sequence $t_n'' \to t''$ in $(0,\infty)$.

In addition, given $g \in \bar{\mathcal{G}}$ and another sequence $t_n' \to t'$ in $(0,\infty)$, abbreviate $P_{n,t_n',g}$ by Q_n and choose $r_n = r_{n,t_n',g} \to t'$ to satisfy (3.10) for (t_n').

Then, like (3.15) has been obtained for the tests (3.13), we conclude that

$$\Phi\bigl(-u_{\beta_g} + t'\,\|g\|\bigr) \ge \limsup_n P^n\bigl\{R_n > r_n\langle\kappa|g\rangle\bigr\} \tag{3.21}$$

using the tests

$$\upsilon_n = \mathbf{I}(R_n > r_n\langle\kappa|g\rangle) \tag{3.22}$$

and their asy. level

$$\beta_g = \limsup_n Q_n^n\bigl\{R_n > r_n\langle\kappa|g\rangle\bigr\} \tag{3.23}$$

By condition (3.7), as $g = g_m'' \in \bar{\mathcal{G}}$ tends to $-\bar\kappa$ in $L_2(P)$,

$$\limsup \beta_g \le \tfrac{1}{2}\,, \quad \text{hence} \quad \liminf u_{\beta_g} \ge 0 \tag{3.24}$$

Therefore, given $\delta \in (0,1)$, we may choose $g = g''_m \in \bar{\mathcal{G}}$ so close to $-\bar{\kappa}$ that

$$-u_{\beta_g} + t'\,\|g\| \le t'\,\|\bar{\kappa}\| + \delta \qquad \text{and} \qquad r_n\langle\kappa|g\rangle \le -(t'_n - \delta)\|\bar{\kappa}\|^2 \tag{3.25}$$

eventually. Then (3.21) implies that, for each $\delta \in (0,1)$,

$$\limsup_n P^n\bigl\{R_n > -(t'_n - \delta)\|\bar{\kappa}\|^2\bigr\} \le \Phi\bigl(t'\,\|\bar{\kappa}\| + \delta\bigr) \tag{3.26}$$

hence

$$\limsup_n P^n\bigl\{R_n > -t'_n\,\|\bar{\kappa}\|^2\bigr\} \le \Phi\bigl(t'\,\|\bar{\kappa}\|\,\bigr) \tag{3.27}$$

that is,

$$\liminf_n P^n\bigl\{R_n \le -t'_n\,\|\bar{\kappa}\|^2\bigr\} \ge \Phi\bigl(-t'\,\|\bar{\kappa}\|\,\bigr) \tag{3.28}$$

As

$$\limsup_n P^n\bigl\{-t'_n\,\|\bar{\kappa}\|^2 < R_n < t''_n\,\|\bar{\kappa}\|^2\bigr\} \le \\ \limsup_n P^n\bigl\{R_n < t''_n\,\|\bar{\kappa}\|^2\bigr\} - \liminf_n P^n\bigl\{R_n \le -t'_n\,\|\bar{\kappa}\|^2\bigr\}$$

bound (3.8) follows from (3.20) and (3.28).

We shall check attainment of the bounds simultaneously in both cases: The asy. linearity (3.5) and (3.9) entail asy. normality under P^n,

$$\Bigl(\sqrt{n}\,\bigl(\hat{S}_n - T(P)\bigr)\Bigr)\,(P^n) \xrightarrow{\mathrm{w}} \mathcal{N}\bigl(0, \|\hat{\kappa}\|^2\bigr) \tag{3.29}$$

for $\hat{S}_n = \tilde{S}_n$ with $\hat{\kappa} = \tilde{\kappa}$, respectively for $\hat{S}_n = \bar{S}_n$ with $\hat{\kappa} = \bar{\kappa}$. It follows that

$$\lim_n P^n\bigl\{-t' < \sqrt{n}\,\bigl(\hat{S}_n - T(P)\bigr) < t''\bigr\} = \Phi\Bigl(\frac{t''}{\|\hat{\kappa}\|}\Bigr) - \Phi\Bigl(-\frac{t'}{\|\hat{\kappa}\|}\Bigr) \tag{3.30}$$

uniformly in $-\infty \le -t' < t'' \le \infty$, in both cases.

Verification of the regularity condition (3.3) for $(\tilde{S}_n)$, and of conditions (3.6) and (3.7) for $(\bar{S}_n)$, is postponed to Subsection 3.3.2. ////

Remark 3.3 In Theorem 3.1(a), the *upper* bound $\Phi\bigl(t/\|\tilde{\kappa}\|\bigr)$ on $(0,\infty)$ given by (3.4), for the 'limit distribution function' of $\sqrt{n}\,\bigl(S_n - T(Q)\bigr)$ under $Q = P$, does not extend to a *lower* bound on $(-\infty, 0)$, whereas the bound $\Phi\bigl(t/\|\bar{\kappa}\|\bigr)$ does in Theorem 3.1(b). For example, given any $a \in (0,\infty)$, consider the following modification $(\breve{S}_n)$ of $(\tilde{S}_n)$,

$$\breve{S}_n = \tilde{S}_n \vee \bigl(T(P) - a/\sqrt{n}\,\bigr) \tag{3.31}$$

Then, if $g \in \tilde{\mathcal{G}}$ is such that $\langle\kappa|g\rangle \ge 0$, and $t_n \to t$ in $(0,\infty)$, it holds that, eventually, $T(P_{n,t_n,g}) \ge T(P) - a/\sqrt{n}$. Using the asymptotic median nonnegativity (3.48) of $(\tilde{S}_n)$ to be proved in Subsection 3.3.2, we obtain that, eventually,

$$P^n_{n,t_n,g}\bigl\{\breve{S}_n \ge T(P_{n,t_n,g})\bigr\} = P^n_{n,t_n,g}\bigl\{\tilde{S}_n \ge T(P_{n,t_n,g})\bigr\} \ge \tfrac{1}{2} + \mathrm{o}(n^0) \tag{3.32}$$

Under P, however, since $\sqrt{n}\,\bigl(\breve{S}_n - T(P)\bigr) = (-a) \vee \sqrt{n}\,\bigl(\tilde{S}_n - T(P)\bigr)$, we have

$$\begin{aligned} P^n\bigl\{\sqrt{n}\,\bigl(\breve{S}_n - T(P)\bigr) \le t\bigr\} &= 0 && \text{if } t < -a \\ &\longrightarrow \Phi\Bigl(\frac{t}{\|\tilde{\kappa}\|}\Bigr) && \text{if } t \ge -a \end{aligned} \tag{3.33}$$

The choice $a = 0$ is possible if the asymptotic median ≥ 0 condition (3.48) is required, instead of for $\langle\kappa|g\rangle \ge 0$, only for $\langle\kappa|g\rangle > 0$, which suffices for (3.3).////

3.2 Uniqueness of Efficient Estimators

In the setup of Theorem 3.1(b), the optimal estimates $\bar{S}_n$ defined by (3.9) are unique, up to terms tending stochastically to zero under (P^n). In the setup of Theorem 3.1(a), on the contrary, only the positive part $\big(\tilde{S}_n - T(P)\big)_+$ of the optimal estimates (3.5) centered at $T(P)$ will be asymptotically unique; confer Remark 3.3 for an example.

Proposition 3.4 *Let $(\breve{S}_n)$ and $(\hat{S}_n)$ be two estimator sequences.*

(a) *In the case of a convex tangent cone $\tilde{\mathcal{G}}$, suppose $(\breve{S}_n)$ satisfies condition* (3.3) *and achieves the confidence bound* (3.4), *with* $\limsup_n$ *replaced by* $\liminf_n$. *Then necessarily*

$$\begin{aligned} \sqrt{n}\,\big(\breve{S}_n - T(P)\big)_+ + \breve{\mathrm{o}}_{P^n}(n^0) &= \big(\sqrt{n}\,\mathrm{ave}_{i=1}^n\,\tilde{\kappa}(x_i)\big)_+ \\ &= \sqrt{n}\,\big(\tilde{S}_n - T(P)\big)_+ + \tilde{\mathrm{o}}_{P^n}(n^0) \end{aligned} \tag{3.34}$$

Conversely, form (3.34) *of* $(\breve{S}_n)$ *implies* (3.48) *and achievement of bound* (3.4), *uniformly in* $-\infty \le t_n = t \le \infty$, *and with* $\limsup_n$ *replaced by* $\liminf_n$.

(b) *In the case of a linear tangent space $\bar{\mathcal{G}}$, assume $(\hat{S}_n)$ satisfies conditions* (3.6) *and* (3.7), *and achieves the confidence bound* (3.8), *with* $\limsup_n$ *replaced by* $\liminf_n$. *Then necessarily*

$$\begin{aligned} \sqrt{n}\,\big(\hat{S}_n - T(P)\big) + \hat{\mathrm{o}}_{P^n}(n^0) &= \sqrt{n}\,\mathrm{ave}_{i=1}^n\,\bar{\kappa}(x_i) \\ &= \sqrt{n}\,\big(\bar{S}_n - T(P)\big) + \bar{\mathrm{o}}_{P^n}(n^0) \end{aligned} \tag{3.35}$$

Conversely, if $(\hat{S}_n)$ is of form (3.35), *then it satisfies* (3.48), (3.49), *and achieves bound* (3.8), *with $t_n' = t'$, $t_n'' = t''$, uniformly in $-\infty \le -t' < t'' \le \infty$, and with* $\limsup_n$ *replaced by* $\liminf_n$.

PROOF The proof draws on the proofs to Proposition 2.7 and Theorem 3.1:

In case (a), let $(\breve{S}_n)$ satisfy (3.3) and achieve bound (3.4) such that, for every constant sequence $t_n = t \in (0,\infty)$,

$$\limsup_m \limsup_n P^n_{n,t,g_m}\big\{\breve{S}_n < T(P_{n,t,g_m})\big\} \le \tfrac{1}{2} \tag{3.3}$$

and

$$\liminf_n P^n\big\{\breve{R}_n < t\|\tilde{\kappa}\|^2\big\} \ge \Phi(t\|\tilde{\kappa}\|) \tag{3.36}$$

where $\breve{R}_n = \sqrt{n}\,\big(\breve{S}_n - T(P)\big)$. Fix some $t_n = t \in (0,\infty)$ and any $\delta_a \in (0,1)$. Choose $\delta \in (0,t)$ small enough and then $g = g_m \in \tilde{\mathcal{G}}$ so close to $\tilde{\kappa}$ that

$$\alpha_g < \tfrac{1}{2} + \delta_a\,, \qquad \Phi\big((t-\delta)\|\tilde{\kappa}\|\big) + \delta_a \ge \Phi(t\|g\|) \ge \Phi(t\|\tilde{\kappa}\|) - \delta_a \tag{3.37}$$

and such that (3.17) is fulfilled, too. Recall (3.10), (3.11), and (3.14). Then

$$\begin{aligned} \liminf_n P^n\big\{\breve{R}_n < s_n\langle\kappa|g\rangle\big\} &\ge \liminf_n P^n\big\{\breve{R}_n < (t-\delta)\|\tilde{\kappa}\|^2\big\} \\ &\ge \Phi\big((t-\delta)\|\tilde{\kappa}\|\big) \ge \Phi(t\|g\|) - \delta_a \end{aligned} \tag{3.38}$$

while

$$\limsup_n P^n_{n,t,g}\big\{\breve{R}_n < s_n\langle\kappa|g\rangle\big\} = \alpha_g < \tfrac{1}{2} + \delta_a \tag{3.39}$$

Therefore, the tests $\tau'_{n,g} = 1 - \tau_{n,t,g} = \mathbf{I}(\breve{R}_n \ge s_n\langle\kappa|g\rangle)$ given by (3.13) satisfy

$$\limsup_n \int \tau'_{n,g}\, dP^n \le \alpha'_g + \delta_a \le \alpha' + 2\delta_a \tag{3.40}$$

$$\liminf_n \int \tau'_{n,g}\, dP^n_{n,t,g} \ge \tfrac{1}{2} - \delta_a \tag{3.41}$$

where $\alpha' = \Phi(-t\|\tilde\kappa\|)$, $\alpha'_g = \Phi(-t\|g\|)$, and so $u_{\alpha'} = t\|\tilde\kappa\|$, $u_{\alpha'_g} = t\|g\|$. Replacing α and β_g in (2.48) by α'_g and $\beta'_g = \Phi\bigl(-u_{\alpha'_g} + t\|g\|\bigr) = 1/2 = \beta'$, respectively, (2.49) is satisfied by the tests $\tau'_{n,g}$ and leeway δ_a, in the place of τ_n and δ there. Via (2.57) and (2.63), we reach (2.65). Taking already the asymptotic equivalence (2.66) into account, where $\|\tilde\kappa\|\, u_{\alpha'} = t\|\tilde\kappa\|^2$, and the fact that the present tests are all nonrandomized, we thus obtain

$$\lim_{g\to\tilde\kappa} \limsup_n P^n(\tau'_{n,g} \ne \tau^\star_n) = 0\,, \qquad \tau^\star_n = \mathbf{I}\bigl(\sqrt{n}\,\mathrm{ave}_1^n\,\tilde\kappa(x_i) \ge t\|\tilde\kappa\|^2\bigr) \tag{3.42}$$

The tests $\tau'_{n,g} = \mathbf{I}(\breve{R}_n \ge s_n\langle\kappa|g\rangle)$ may be compared with $\tau'_n = \mathbf{I}(\breve{R}_n \ge t\|\tilde\kappa\|^2)$. In (3.36), $P^n\bigl(\breve{R}_n < t\|\tilde\kappa\|^2\bigr)$ must actually converge to $\Phi(t\|\tilde\kappa\|)$, and $s_n \to t$. Therefore, employing the modulus ω_Φ of uniform continuity of Φ, we obtain

$$\limsup_n P^n(\tau'_{n,g} \ne \tau'_n) \le \omega_\Phi(t|\lambda_g|)\,, \qquad \lambda_g = \frac{\langle\kappa|g\rangle}{\|\tilde\kappa\|} - \|\tilde\kappa\| \tag{3.43}$$

As $\lim_{g\to\tilde\kappa}\lambda_g = 0$, it follows that

$$\lim_{g\to\tilde\kappa} \limsup_n P^n(\tau'_{n,g} \ne \tau'_n) = 0 \tag{3.44}$$

Using the triangle inequality, we deduce from (3.42) and (3.44) that

$$\lim_n P^n\bigl\{\mathbf{I}(\breve{R}_n \ge t) \ne \mathbf{I}\bigl(\sqrt{n}\,\mathrm{ave}_1^n\,\tilde\kappa(x_i) \ge t\bigr)\bigr\} = 0 \tag{3.45}$$

for every $t \in (0,\infty)$. Because $\sqrt{n}\,\mathrm{ave}_1^n\,\tilde\kappa(x_i)$ is tight under (P^n), the difference between the positive parts of $\breve{R}_n$ and $\sqrt{n}\,\mathrm{ave}_1^n\,\tilde\kappa(x_i)$ must converge to zero in P^n-probability; confer HR (1981 b), fact (3.12)–(3.13). Thus (3.34) is proved.

In case (b), we may now continue the same way as part (b) of the proof to Theorem 3.1 proceeds after part (a). From (2.48) and (2.49) onwards, plug the tests $\upsilon'_{n,g} = 1 - \upsilon_{n,t',g} = \mathbf{I}(\hat{R}_n \le r_n\langle\kappa|g\rangle)$ given by (3.22) in the proof of Proposition 2.7. Letting g tend to $-\bar\kappa$, one similarly obtains that

$$\lim_n P^n\bigl\{\mathbf{I}(\hat{R}_n \le -t) \ne \mathbf{I}\bigl(\sqrt{n}\,\mathrm{ave}_1^n\,\bar\kappa(x_i) \le -t\bigr)\bigr\} = 0 \tag{3.46}$$

for every $t \in (0,\infty)$. This implies that also the difference between the negative parts of $\hat{R}_n$ and $\sqrt{n}\,\mathrm{ave}_1^n\,\bar\kappa(x_i)$ must go to zero in P^n-probability, hence (3.35).

As for the converse, which is obvious in case (b), observe in case (a) that, for some stochastic term $\mathrm{o}_n = \mathrm{o}_{P^n}(n^0)$, and for every $t \in (0,\infty)$,

$$\begin{aligned} P^n(\breve{R}_n < t) &= P^n(\breve{R}_n^+ < t) = P^n(\tilde{R}_n^+ < t + \mathrm{o}_n) \\ &= P^n(\tilde{R}_n^+ < t) + \mathrm{o}(n^0) = P^n(\tilde{R}_n < t) + \mathrm{o}(n^0) \end{aligned} \tag{3.47}$$

where the third equality is true because the limit $\Phi(t/\|\tilde{\kappa}\|)$ is continuous in t. Thus, $(\breve{S}_n)$ inherits the optimality from $(\tilde{S}_n)$.

Verification of the regularity conditions is postponed to Subsection 3.3.2. ////

3.3 Regularity of Efficient Estimators

The asymptotic upper bounds (3.4) and (3.8) for the confidence probabilities derived in Theorem 3.1 seem to involve only P. The model $\mathcal{P}$ and its tangent set $\mathcal{G}$ at P, however, enter through the regularity condition. As indicated above, the bounds are not meaningful without such regularity conditions.

3.3.1 Modified Regularity, Asymptotic Linearity and Normality

Asymptotic Median Bias In Theorem 3.1, the regularity conditions (3.3), (3.6) and (3.7), respectively, are certainly fulfilled if asymptotic median nonnegativity, respectively nonpositivity, holds for every fixed tangent in the respective tangent set $\mathcal{G}$, in the sense that

$$\liminf_n P^n_{n,t_n,g}\bigl\{S_n \ge T(P_{n,t_n,g})\bigr\} \ge \tfrac{1}{2} \tag{3.48}$$

$$\liminf_n P^n_{n,t_n,g}\bigl\{S_n \le T(P_{n,t_n,g})\bigr\} \ge \tfrac{1}{2} \tag{3.49}$$

respectively, for every $g \in \mathcal{G}$ and every convergent sequence $t_n \to t \in (0,\infty)$. The notion implicitly depends on P, the model $\mathcal{P}$, and its tangent set $\mathcal{G}$ at P.

Asymptotic median unbiasedness, that is, (3.48) and (3.49), for every $g \in \mathcal{G}$, is the regularity assumption of Pfanzagl and Wefelmeyer (1982; Theorem 9.2.2).

Asymptotic Linear Estimators An estimator sequence (S_n) is asymptotically linear at P if there exists some function $\eta \in L_2(P) \cap \{\mathrm{const}\}^{\perp}$, the (unique) influence curve of (S_n) at P, such that

$$\sqrt{n}\,\bigl(S_n - T(P)\bigr) = \sqrt{n}\,\mathrm{ave}_{i=1}^{n}\,\eta(x_i) + \mathrm{o}_{P^n}(n^0) \tag{3.50}$$

For example, the estimator sequences $(\bar{S}_n)$ and $(\tilde{S}_n)$, in view of (3.9) and (3.5), are asymptotically linear at P with influence curves $\bar{\kappa}$ and $\tilde{\kappa}$, respectively.

The construction of such estimators, given a family of prescribed influence curves η_P, one for each (unknown) element P of model $\mathcal{P}$, under very general conditions, is a topic in Bickel et al. (1993; Appendix A.10), van der Vaart (1998; Sections 25.8–10), and HR (1994; Chapters 1 and 6).

Asymptotic Normality The expansions (3.50), (3.10), and (3.12), of the estimator, the functional, and loglikelihoods, respectively, imply the following asymptotic normality extending (3.29),

$$\Bigl(\sqrt{n}\,\bigl(S_n - T(P_{n,t_n,g})\bigr)\Bigr)\bigl(P^n_{n,t_n,g}\bigr) \xrightarrow{\;\mathrm{w}\;} \mathcal{N}\bigl(t\,\langle \eta - \kappa | g\rangle, \|\eta\|^2\bigr) \tag{3.51}$$

for all convergent $t_n \to t \in (0,\infty)$, every $g \in \mathcal{G}$, and so, for each $c \in [0,\infty)$,

$$\lim_n P^n_{n,t_n,g}\bigl\{\sqrt{n}\,\bigl(S_n - T(P_{n,t_n,g})\bigr) < c\bigr\} = \Phi\Bigl(\frac{c - t\langle \eta - \kappa|g\rangle}{\|\eta\|}\Bigr) \tag{3.52}$$

$$\lim_n P^n_{n,t_n,g}\bigl\{\sqrt{n}\,\bigl(S_n - T(P_{n,t_n,g})\bigr) > -c\bigr\} = \Phi\Bigl(\frac{c + t\langle \eta - \kappa|g\rangle}{\|\eta\|}\Bigr) \tag{3.53}$$

where of course $< c$ may also be replaced by $\le c$. These convergences in particular apply to $(\bar S_n)$ and $(\tilde S_n)$, with $\eta = \bar\kappa$, respectively $\eta = \tilde\kappa$.

Asymptotic Confidence Probabilities Based on $(\breve S_n)$: Besides $(\tilde S_n)$, we consider any optimal estimator sequences $(\breve S_n)$ as described by (3.34). Then, by the asymptotic normality (3.51) of $(\tilde S_n)$, and contiguity, we conclude that, for all convergent $t_n \to t$ in $(0,\infty)$, every tangent $g \in \tilde{\mathcal{G}}$, and each $c \in \mathbb{R}$,

$$\begin{aligned}
&P^n_{n,t_n,g}\bigl\{\sqrt{n}\,\bigl(\breve S_n - T(P_{n,t_n,g})\bigr) < c\bigr\} &(3.54)\\
&\quad= P^n_{n,t_n,g}\bigl\{\sqrt{n}\,\bigl(\breve S_n - T(P)\bigr) < c + t\,\langle\kappa|g\rangle + \mathrm{o}(n^0)\bigr\}\\
&\quad= P^n_{n,t_n,g}\bigl\{\sqrt{n}\,\bigl(\breve S_n - T(P)\bigr)_+ < c + t\,\langle\kappa|g\rangle + \mathrm{o}(n^0)\bigr\}\\
&\quad= P^n_{n,t_n,g}\bigl\{\sqrt{n}\,\bigl(\tilde S_n - T(P)\bigr)_+ < c + t\,\langle\kappa|g\rangle + \mathrm{o}_{P^n}(n^0)\bigr\}\\
&\quad= P^n_{n,t_n,g}\bigl\{\sqrt{n}\,\bigl(\tilde S_n - T(P)\bigr)_+ < c + t\,\langle\kappa|g\rangle\bigr\} + \mathrm{o}(n^0)\\
&\quad= P^n_{n,t_n,g}\bigl\{\sqrt{n}\,\bigl(\tilde S_n - T(P)\bigr) < c + t\,\langle\kappa|g\rangle\bigr\} + \mathrm{o}(n^0)\\
&\quad= P^n_{n,t_n,g}\bigl\{\sqrt{n}\,\bigl(\tilde S_n - T(P_{n,t_n,g})\bigr) < c\bigr\} + \mathrm{o}(n^0) &(3.55)
\end{aligned}$$

provided that

$$t\,\langle\kappa|g\rangle > -c \tag{3.56}$$

In (3.54)–(3.55), we may replace $<$ by $\le$, hence by any inequality sign.

Thus (3.52) and (3.53), using $-c$ instead of c, extend from $(\tilde S_n)$ to $(\breve S_n)$.

3.3.2 One-Sided Regularity of $(\tilde S_n)$, $(\breve S_n)$

Let $c = 0$, as in the regularity assumptions of Theorem 3.1.

Regularity of $(\bar S_n)$: For $\eta = \bar\kappa$, since $\langle\bar\kappa - \kappa|g\rangle = 0\ \ \forall g \in \bar{\mathcal{G}}$, the two limits in (3.52) and (3.53) are always $1/2$. Hence $(\bar S_n)$ is asymptotically median unbiased, for each $g \in \bar{\mathcal{G}}$. In particular, $(\bar S_n)$ satisfies conditions (3.6) and (3.7).

Regularity of $(\tilde S_n)$, $(\breve S_n)$: For $\eta = \tilde\kappa$, since $t > 0$ and $\langle\tilde\kappa - \kappa|g\rangle \ge 0\ \ \forall g \in \tilde{\mathcal{G}}$, the limit in (3.52) is always $\le 1/2$ (it is $= 1/2$, e.g. for[5] $g = 0, \tilde\kappa$).

Thus, $(\tilde S_n)$ satisfies the asymptotic median nonnegativity condition (3.48), for every $g \in \tilde{\mathcal{G}}$, hence, in particular, $(\tilde S_n)$ fulfills condition (3.3).

If $(\breve S_n)$ satisfying (3.34) is another optimal estimator sequence, (3.54)–(3.56) apply with $c = 0$, hence asymptotic median nonnegativity (3.48) of $(\tilde S_n)$ is

[5] In Subsections 3.3 and 3.5, the choice $g = \tilde\kappa$ stands under the provision that $\tilde\kappa \in \tilde{\mathcal{G}}$.

inherited to $(\breve{S}_n)$, for every $g \in \tilde{\mathcal{G}}$ such that $\langle \kappa | g \rangle > 0$. This suffices to fulfill condition (3.3), because $\langle \kappa | \tilde{\kappa} \rangle = \|\tilde{\kappa}\|^2 > 0$, and so eventually $\langle \kappa | g_m \rangle > 0$ for any tangents $g_m \in \tilde{\mathcal{G}}$ approaching $\tilde{\kappa}$.

3.3.3 Positive Median Bias of $(\tilde{S}_n)$, $(\breve{S}_n)$

For $c = 0$ and $\eta = \tilde{\kappa}$, the limit in (3.53) ($= 1/2$ for $g = 0, \tilde{\kappa}$) in general falls below $1/2$. We shall prove this for any estimator sequence $(\breve{S}_n)$ which is optimal in the sense of Theorem 3.1(a).

Consequently, all these estimators violate asymptotic median nonpositivity (3.49). The result corresponds to the level breakdown we have encountered in Subsection 2.4.

Proposition 3.5 *Let $\tilde{\mathcal{G}}$ be a convex tangent cone such that*

$$\bar{\kappa} \neq \tilde{\kappa} \tag{3.57}$$

Then there is some tangent $g_0 \in \tilde{\mathcal{G}}$ such that $\langle \kappa | g_0 \rangle > 0$ and

$$\inf_{t>0} \lim_n P^n_{n,t,g_0}\bigl\{\breve{S}_n \le T(P_{n,t,g_0})\bigr\} = 0 \tag{3.58}$$

for all estimator sequences $(\breve{S}_n)$ of kind (3.34).

Proof If $\bar{\kappa} \neq \tilde{\kappa}$ there is some tangent $g_0 \in \tilde{\mathcal{G}}$ such that

$$\langle \kappa | g_0 \rangle < \langle \tilde{\kappa} | g_0 \rangle \tag{3.59}$$

Then, for $c = 0$ and $\eta = \tilde{\kappa}$, (3.53) implies

$$\inf_{t>0} \lim_n P^n_{n,t,g_0}\bigl\{\tilde{S}_n \le T(P_{n,t,g_0})\bigr\} = \inf_{t>0} \Phi\Bigl(-\frac{t\langle \tilde{\kappa} - \kappa | g \rangle}{\|\tilde{\kappa}\|}\Bigr) = 0 \tag{3.60}$$

But g_0 may always be chosen such that, in addition to (3.59),

$$0 < \langle \kappa | g_0 \rangle < \langle \tilde{\kappa} | g_0 \rangle \tag{3.61}$$

If necessary, pass to a suitable convex combination g_{02} of g_0 satisfying (3.59) and $\tilde{\kappa}$, in order to achieve (3.61).

Then the arguments (3.54)–(3.56) go through, with $c = 0$, and with $\le$ in the place of $<$. Thus the positive asymptotic median bias (3.58) carries over from $(\tilde{S}_n)$ to all estimator sequences $(\breve{S}_n)$ satisfying (3.34). ////

Remark 3.6 The result implies that bound (3.4) cannot possibly be achieved if, in addition to (3.3), asymptotic median nonpositivity (3.49) is imposed (for all $g \in \tilde{\mathcal{G}}$, or only all $g \in \tilde{\mathcal{G}}$ such that $\langle \kappa | g \rangle > 0$). In particular, asymptotic median unbiasedness cannot be afforded if bound (3.4) is to be attained.

As a consequence, Theorem 9.2.2 of Pfanzagl and Wefelmeyer (1982) for (closed) convex tangent cones $\tilde{\mathcal{G}}$ is ailing in two respects:

First, since $-g \notin \tilde{\mathcal{G}}$ in general and $-\tilde{\kappa} \notin \tilde{\mathcal{G}}$ in particular, their bound of form (3.8) for two-sided confidence limits, with $\tilde{\kappa}$ in the place of $\bar{\kappa}$, is not available over cones, but only bound (3.4) for lower confidence limits.

Second, their regularity condition is too strict: Contrary to what they believe (Section 9.1, p 154), even the one-sided bound (3.4), let alone the asserted two-sided extension (3.8), cannot possibly be achieved by any estimator sequence that is asymptotically median unbiased. ////

3.4 Comparison of Cones and Spaces

Variance and Sample Size Recall the setup of Subsection 2.3: $P \in \tilde{\mathcal{P}} \subset \bar{\mathcal{P}}$, with tangent set a convex cone $\tilde{\mathcal{G}}$, respectively the linear span (2.21): $\bar{\mathcal{G}} = \operatorname{lin} \tilde{\mathcal{G}}$.

Then, in view of the asymptotic normality (3.29), the previous comparison of $\|\bar{\kappa}\|^2$ and $\|\tilde{\kappa}\|^2$ now concerns the variances $\|\tilde{\kappa}\|^2/n$ and $\|\bar{\kappa}\|^2/n$ of the approximate normal distributions of $\tilde{S}_n - T(P)$ and $\bar{S}_n - T(P)$, respectively.

Thus, the value $T(P)$, in terms of variance or width of confidence intervals, can be estimated under P^n more accurately in model $\mathcal{P}$ with tangent set $\tilde{\mathcal{G}}$ than it is possible in the larger model $\bar{\mathcal{P}}$ with tangent set $\bar{\mathcal{G}}$. Observations at the higher rate $\bar{n}/\tilde{n} \to \|\bar{\kappa}\|^2/\|\tilde{\kappa}\|^2$ are needed under P to estimate $T(P)$ with the same asymptotic accuracy by $\bar{S}_{\bar{n}}$ as by $\tilde{S}_{\tilde{n}}$. Again Example 2.3 applies.

Lower Confidence Limits for Spaces The preceding comparison does not explicitly take the different sets of regularity assumptions into account: In the case of $(\tilde{S}_n)$, it is condition (3.3), and conditions (3.6), (3.7) in the case of $(\bar{S}_n)$.

However, in the case of a linear tangent space $\bar{\mathcal{G}}$, suppose we dispense of condition (3.7) and, keeping (3.6), wish to maximize the asymptotic confidence probability merely of the sequence of lower confidence limits $S_n - c/\sqrt{n}$ of $T(P)$, under (P^n). In particular, the statistical task seems to be made easier.

Nevertheless, the previous upper bound $\Phi\bigl(c/\|\bar{\kappa}\|\bigr)$ established under Theorem 3.1(b), with $t'_n = t' = \infty$, does not increase, and $(\bar{S}_n)$ remains an optimal estimator sequence. This is true, simply because $\bar{\mathcal{G}}$ also is a convex tangent cone, so Theorem 3.1(a) especially holds for $\bar{\mathcal{G}}$.

Thus, under condition (3.6), asymptotic median nonpositivity (3.49) for $\bar{\mathcal{G}}$, as well as the maximization, subject to (3.7), of the asymptotic confidence probability under (P^n) of the sequence of upper confidence limits $S_n + c/\sqrt{n}$ for $T(P)$, come free with $(\bar{S}_n)$, which achieves the corresponding upper bound, which is $\Phi\bigl(c/\|\bar{\kappa}\|\bigr)$ again.

Two-Sided Confidence Limits for Cones In the case of a convex tangent cone $\tilde{\mathcal{G}}$, suppose we want to maximize the asymptotic confidence probability under (P^n) of the sequence of lower confidence limits $S_n - c/\sqrt{n}$ of $T(P)$, as in Theorem 3.1(a), but insist on asymptotic median unbiasedness, that is, (3.48) and (3.49) for every $g \in \tilde{\mathcal{G}}$. As (3.48) and (3.49) imply (3.3), the statistical task is made more difficult, and one expects the upper bound $\Phi\bigl(c/\|\tilde{\kappa}\|\bigr)$ to decrease. According to Proposition 3.5, it must strictly decrease if $\bar{\kappa} \neq \tilde{\kappa}$.

We clarify the amount of decrease, at least in the class of estimator sequences (S_n) which are asymptotically linear at P. For such an estimator with influence curve η at P, the lower/upper confidence limits $S_n \mp c/\sqrt{n}$ satisfy

$$\begin{aligned} \lim_n P^n\bigl\{\sqrt{n}\,\bigl(S_n - T(P)\bigr) < c\bigr\} &= \Phi\Bigl(\frac{c}{\|\eta\|}\Bigr) \\ &= \lim_n P^n\bigl\{\sqrt{n}\,\bigl(S_n - T(P)\bigr) > -c\bigr\} \end{aligned} \tag{3.62}$$

Under local alternatives, in view of the limits (3.52) and (3.53) for each $g \in \tilde{\mathcal{G}}$, (S_n) is asymptotically median unbiased iff $\langle\eta|g\rangle = \langle\kappa|g\rangle$ $\forall g \in \tilde{\mathcal{G}}$, which holds if and only if

$$\langle\eta|g\rangle = \langle\kappa|g\rangle \qquad \forall g \in c\ell\,\mathrm{lin}\,\tilde{\mathcal{G}} \tag{3.63}$$

Introducing the projections $\bar{\kappa}$ of κ, and $\bar{\eta}$ of η, on $c\ell\,\mathrm{lin}\,\tilde{\mathcal{G}}$, $\bar{\eta}$ must equal $\bar{\kappa}$. But, subject to $\bar{\eta} = \bar{\kappa}$, the asymptotic confidence probability $\Phi\bigl(c/\|\eta\|\bigr)$ is maximized iff $\|\eta\|$ is minimized, which is the case iff $\eta = \bar{\kappa}$.

Therefore, in the class of estimator sequences which are asymptotically linear at P, the unique solution is the estimator sequence $(\bar{S}_n)$ with influence curve $\bar{\kappa}$. And the achievable upper bound decreases from $\Phi\bigl(c/\|\tilde{\kappa}\|\bigr)$ to $\Phi\bigl(c/\|\bar{\kappa}\|\bigr)$.

So the answer to the corresponding (open) question raised for testing in Remark 2.6 turns out negative in the estimation context.

In addition, in view of (3.62), the upper confidence limits $\bar{S}_n + c/\sqrt{n}$ of $T(P)$ supplied by $(\bar{S}_n)$ have the same asymptotic confidence probability $\Phi\bigl(c/\|\bar{\kappa}\|\bigr)$ under P^n as the lower confidence limits $\bar{S}_n - c/\sqrt{n}$. And the two-sided bounds $\bar{S}_n \mp c/\sqrt{n}$, in view of (3.64) below, maintain their asymptotic confidence probability for T even under local perturbations $P_{n,t,g}$ of P, $g \in \tilde{\mathcal{G}}$.

3.5 Local Behaviour of Efficient Confidence Limits

3.5.1 Confidence Probabilities Under Perturbations

Given $c \in (0,\infty)$, we study the two sequences of lower/upper limits $\bar{S}_n \mp c/\sqrt{n}$ and $\tilde{S}_n \mp c/\sqrt{n}$ under local perturbations $P_{n,t,g}$ of P.

Stability of Confidence Limits Based on $(\bar{S}_n)$: For $\eta = \bar{\kappa}$, the two limits in (3.52) and (3.53), since $\langle\bar{\kappa} - \kappa|g\rangle = 0$ $\forall g \in \bar{\mathcal{G}}$, are always the same,

$$\begin{aligned} \lim_n P^n_{n,t_n,g}\bigl\{\sqrt{n}\,\bigl(\bar{S}_n - T(P_{n,t_n,g})\bigr) < c\bigr\} &= \Phi\Bigl(\frac{c}{\|\bar{\kappa}\|}\Bigr) \\ &= \lim_n P^n_{n,t_n,g}\bigl\{\sqrt{n}\,\bigl(\bar{S}_n - T(P_{n,t_n,g})\bigr) > -c\bigr\} \end{aligned} \tag{3.64}$$

for every convergent sequence $t_n \to t$ in $(0,\infty)$, every $g \in \bar{\mathcal{G}}$, which reveals some stability of the lower/upper limits based on $(\bar{S}_n)$.

Instability of Confidence Limits Based on $(\tilde{S}_n)$, $(\breve{S}_n)$: For $\eta = \tilde{\kappa}$, the limits in (3.52) and (3.53) are, respectively,

$$\lim_n P^n_{n,t_n,g}\bigl\{\sqrt{n}\,\bigl(\tilde{S}_n - T(P_{n,t_n,g})\bigr) < c\,\bigr\} = \Phi\Bigl(\frac{c - t\langle\tilde{\kappa} - \kappa|g\rangle}{\|\tilde{\kappa}\|}\Bigr) \tag{3.65}$$

$$\lim_n P^n_{n,t_n,g}\bigl\{\sqrt{n}\,\bigl(\tilde{S}_n - T(P_{n,t_n,g})\bigr) > -c\,\bigr\} = \Phi\Bigl(\frac{c + t\langle\tilde{\kappa} - \kappa|g\rangle}{\|\tilde{\kappa}\|}\Bigr) \tag{3.66}$$

Under-Coverage by Lower Confidence Limits The limit (3.65) is always $\le \Phi\bigl(c/\|\tilde{\kappa}\|\bigr)$, since $t > 0$ and $\langle\tilde{\kappa} - \kappa|g\rangle \ge 0$ $\forall g \in \tilde{\mathcal{G}}$; the upper bound is achieved, e.g. for $g = 0, \tilde{\kappa}$. In general, e.g. for g_0 taken from (3.61), the limit in (3.65), with $\le c$ in the place of $< c$, may become arbitrarily close to 0 as

$$\inf_{t>0} \lim_n P^n_{n,t,g_0}\bigl\{\sqrt{n}\,\bigl(\tilde{S}_n - T(P_{n,t,g_0})\bigr) \le c\,\bigr\} = 0 \tag{3.67}$$

In view of (3.54)–(3.56), the limit statement (3.65) for $t\langle\kappa|g\rangle > -c$ extends to $(\breve{S}_n)$, hence also (3.67) extends to $(\breve{S}_n)$. Obviously, (3.67) generalizes (3.58).

Over-Coverage by Upper Confidence Limits The limit in (3.66) is always $\ge \Phi\bigl(c/\|\tilde{\kappa}\|\bigr)$; and $= \Phi\bigl(c/\|\tilde{\kappa}\|\bigr)$ e.g. for $g = 0, \tilde{\kappa}$. In general, e.g. for g_0 taken from (3.61), the limit in (3.66) may become arbitrarily close to 1,

$$\sup_{t>0} \lim_n P^n_{n,t,g_0}\bigl\{\sqrt{n}\,\bigl(\tilde{S}_n - T(P_{n,t,g_0})\bigr) > -c\,\bigr\} = 1 \tag{3.68}$$

In view of (3.54)–(3.56), with $-c$ in the place of c, the limit statement (3.66) extends from $(\tilde{S}_n)$ to $(\breve{S}_n)$ of form (3.34), provided that $t\langle\kappa|g_0\rangle > c$, and hence also (3.68) extends to $(\breve{S}_n)$.

The degenerate limits (3.67) and (3.68) indicate an instability of the lower and upper confidence limits based on $(\tilde{S}_n)$, which is not accounted for by the criterion maximized in Theorem 3.1(a) merely under (P^n), nor by the (only one-sided) asymptotic median nonnegativity conditions (3.3) or (3.48).

3.5.2 $(\bar{S}_n)$, $(\tilde{S}_n)$, $(\breve{S}_n)$ in the Light of the Convolution Theorem

Superefficiency The convolution theorem by van der Vaart (1998; Theorem 25.20) states the lower bound $\|\bar{\kappa}\|^2$ for the asymptotic variance, which is attained by $(\bar{S}_n)$ in (3.29), but which seems to contradict the smaller asymptotic variance $\|\tilde{\kappa}\|^2$ of $(\tilde{S}_n)$ in (3.29), in case (3.57): $\bar{\kappa} \ne \tilde{\kappa}$.

Hájek–Regularity This convolution result concerns the asymptotic variance of estimator sequences (S_n) which are Hájek–regular. (S_n) is called Hájek–regular at P, for the functional T, along the tangent set $\mathcal{G}$, if there is some (limit) distribution M such that, for every $g \in \tilde{\mathcal{G}}$ and $t_n \to t$ in $(0,\infty)$,

$$\Bigl(\sqrt{n}\,\bigl(S_n - T(P_{n,t_n,g})\bigr)\Bigr)\bigl(P^n_{n,t_n,g}\bigr) \xrightarrow{\;\mathrm{w}\;} M \tag{3.69}$$

If (S_n) is Hájek–regular with limit M, then $M(0,\infty) \geq 1/2$ implies asymptotic median nonnegativity (3.48), $M(-\infty,0) \geq 1/2$ implies asymptotic median nonpositivity (3.49), and $M(0,\infty) = 1/2$, $M(\{0\}) = 0$, implies that (S_n) is asymptotically median unbiased.

Hájek–Nonregularity Contrary to $(\bar{S}_n)$, whose limit distribution in (3.51) is always $\mathcal{N}\big(0, \|\bar{\kappa}\|^2\big)$, hence is Hájek–regular, the limit distribution of $(\tilde{S}_n)$ in (3.51) clearly does depend on the particular $(t,g) \in (0,\infty)\times\tilde{\mathcal{G}}$. Therefore, the estimator sequence $(\tilde{S}_n)$ is not Hájek–regular. As (3.65), (3.67) with $c \in (0,\infty)$ also hold for $(\check{S}_n)$ and $g = 0$, respectively for the tangent g_0 taken from (3.61), neither estimator sequence $(\check{S}_n)$ which is optimal in the sense of Theorem 3.1(a) can be Hájek–regular, unless $\tilde{\kappa} = \bar{\kappa}$.

4 Appendix

4.1 Projection—Generalities

Let $\mathcal{H}$ be a Hilbert space—for example, $\mathcal{H} = L_2(P)$—and fix some $\kappa \in \mathcal{H}$.

If $\bar{\mathcal{G}}$ is a closed linear subspace of $\mathcal{H}$, the orthogonal projection $\bar{\kappa} \in \bar{\mathcal{G}}$ of κ on $\bar{\mathcal{G}}$, and unique element of $\bar{\mathcal{G}}$ closest to κ in norm $\|\,.\,\|$, is characterized by

$$\langle \kappa - \bar{\kappa} | g \rangle = 0 \quad \forall g \in \bar{\mathcal{G}} \tag{4.1}$$

If $\tilde{\mathcal{G}}$ is a closed convex cone in $\mathcal{H}$, the projection $\tilde{\kappa} \in \tilde{\mathcal{G}}$ of κ on $\tilde{\mathcal{G}}$, that is, the unique element of $\tilde{\mathcal{G}}$ closest to κ in norm $\|\,.\,\|$, is characterized by

$$\langle \kappa | \tilde{\kappa} \rangle = \|\tilde{\kappa}\|^2 , \qquad \langle \kappa | g \rangle \leq \langle \tilde{\kappa} | g \rangle \quad \forall g \in \tilde{\mathcal{G}} \tag{4.2}$$

If $\hat{\mathcal{G}}$ is an arbitrary nonempty closed convex subset of $\mathcal{H}$, the unique minimum norm element $\hat{g}$ of $\hat{\mathcal{G}}$ is characterized by

$$\|\hat{g}\|^2 \leq \langle g | \hat{g} \rangle \quad \forall g \in \hat{\mathcal{G}} \tag{4.3}$$

These facts are well-known; see, for example, Proposition 4.2.1 in Pfanzagl and Wefelmeyer (1982). (4.3) may be proved by differentiation at $s = 0$ of the function $\|(1-s)\hat{g} + sg\|^2$, which is convex in $0 \leq s \leq 1$, for any $g \in \hat{\mathcal{G}}$. Passing to $\kappa - \tilde{\mathcal{G}}$ and using the structure of cones, (4.2) may be derived from (4.3). Using $-\bar{\mathcal{G}} = \bar{\mathcal{G}}$ for the linear space $\bar{\mathcal{G}}$, (4.1) follows from (4.2).

4.2 Projection—Examples

ad **Example 2.3:** Recall (2.27) and (2.29). Then

$$\bar{\gamma}_1 > 0 \iff b_1 > b_2\, c \iff \varphi(0) - \varphi(a) < \varphi(0) \tag{4.4}$$

Introduce the function $r(a) = \big[\varphi(0) - \varphi(a)\big] \big/ \big[2\Phi(a) - 1\big]$. Then

$$\bar{\gamma}_2 < 0 \iff b_2 < b_1 c \iff r(a) < \varphi(0) \tag{4.5}$$

However, $\varphi(0) = \lim_{a\to\infty} r(a)$ and $\lim_{a\downarrow 0} r(a) = 0$ (de l'Hospital). Moreover,

$$\dot r(a) > 0 \iff \varphi(0) - \varphi(a) < a\big[\Phi(a) - \tfrac{1}{2}\big] \tag{4.6}$$

But

$$\varphi(0) - \varphi(a) = \int_0^a x\varphi(x)\,dx < a\int_0^a \varphi(x)\,dx = a\big[\Phi(a) - \tfrac{1}{2}\big] \tag{4.7}$$

Also $b_2 < b_1$, since $b_2 < b_1 c$ and $c < 1$ (Cauchy–Schwarz).

ad **Example 2.5:** Recall $b_1 = \langle\kappa|g_1\rangle = 2\varphi(0)$ from (2.27), g_3 from (2.36), and put $b_3 = \langle\kappa|g_3\rangle$, $c = \langle g_1|g_3\rangle$. Set $\sigma = \delta/\eta$. Then

$$1 = \|g_3\|^2 = 2\eta^2\Big(\sigma^2\big[\Phi(a) - \tfrac{1}{2}\big] + \big[1 - \Phi(a)\big]\Big) \tag{4.8}$$

As $\frac{1}{2}b_3 = \delta\big[\varphi(0) - \varphi(a)\big] - \eta\varphi(a)$, we have

$$b_3 < 0 \iff \sigma\big[\varphi(0) - \varphi(a)\big] < \varphi(a) \tag{4.9}$$

And as $\frac{1}{2}c = \delta\big[\Phi(a) - \frac{1}{2}\big] - \eta\big[1 - \Phi(a)\big]$, we have

$$c > 0 \iff \sigma\big[\Phi(a) - \tfrac{1}{2}\big] > \big[1 - \Phi(a)\big] \tag{4.10}$$

But

$$a\big[1 - \Phi(a)\big] < \int_a^\infty x\varphi(x)\,dx = \varphi(a) \tag{4.11}$$

and (4.7),

$$\frac{\varphi(0) - \varphi(a)}{\Phi(a) - \Phi(0)} < a < \frac{\varphi(a) - \varphi(\infty)}{\Phi(\infty) - \Phi(a)} \tag{4.12}$$

imply

$$\frac{\varphi(a)}{\varphi(0) - \varphi(a)} > \sigma > \frac{1 - \Phi(a)}{\Phi(a) - \Phi(0)} \tag{4.13}$$

for $\sigma = \sigma_a = a\big[1 - \Phi(a)\big]/\big[\varphi(0) - \varphi(a)\big]$. Then (4.8) defines us $\eta = \eta_a$.

As $b_3 < 0 < c, b_1$, the coefficients of the projection $\bar\kappa$ on $\bar{\mathcal G} = (c\ell)\operatorname{lin}\{g_1, g_2\}$ satisfy $\bar\gamma_1 > 0 > \bar\gamma_3$; confer (2.29). Therefore, $\bar\kappa \neq$ the projection $\tilde\kappa$ on the (closed) convex cone $\tilde{\mathcal G}$ generated by g_1 and g_3, and so $\|\tilde\kappa\| < \|\bar\kappa\|$.

In minimizing the Lagrangian corresponding to (2.30), we can again rule out that both multipliers vanish. If $\beta_1 > 0$ then $\tilde\gamma_1 = 0$ and $\tilde\gamma_3 = b_3 + \beta_3 \geq 0$. As $b_3 < 0$, necessarily $\beta_3 > 0$, hence $\tilde\gamma_3 = 0$ as $\beta_3\tilde\gamma_3 = 0$, which leads to an approximation error of $\|\kappa - 0\|^2 = 1$. This is worse than the error obtained under the assumption that $\beta_3 > 0$. For in this case, $\tilde\gamma_3 = 0$ and $\tilde\gamma_1 = b_1 + \beta_1$ where $\beta_1 = 0$ due to $\beta_1\tilde\gamma_1 = 0$ and $b_1 > 0$. Hence $\tilde\gamma_1 = b_1$, and the error amounts to $\|\kappa - b_1 g_1\|^2 = 1 - b_1^2 < 1$. Altogether, this proves that $\tilde\kappa = b_1 g_1$.

4.3 Approximate Uniqueness

Given two probabilites P and Q on some sample space, let τ^* be a Neyman–Pearson test for P vs. Q, with critical value $c \in [0,\infty]$,

$$\tau^* = \mathbf{I}(dQ > c\,dP) \qquad \text{on } \{dQ \neq c\,dP\} \tag{4.14}$$

and possibly nonconstant randomization on $\{dQ = c\,dP\}$. By $|\nu_c|$ we denote the total variation measure of $d\nu_c = dQ - c\,dP$.

Lemma 4.1 *Consider any test* τ *for* P *vs.* Q *such that, for some* $\delta \in (0,1)$,

$$\int \tau\,dP \leq \int \tau^*\,dP + \delta\,, \qquad \int \tau\,dQ \geq \int \tau^*\,dQ - \delta \tag{4.15}$$

Then

$$|\nu_c|\bigl\{|\tau - \tau^*| > \varepsilon\bigr\} \leq (1+c)\tfrac{\delta}{\varepsilon} \qquad \forall\,\varepsilon > 0 \tag{4.16}$$

Proof Choose any dominating positive measure μ, and densities p, q such that $dP = p\,d\mu$ and $dQ = q\,d\mu$. Then $d\nu_c = (q - cp)\,d\mu$ and, by Rudin (1974; Theorem 6.13), $d|\nu_c| = |q - cp|\,d\mu$. Since $\int (\tau^* - \tau)\,d\nu_c \leq (1+c)\,\delta$ by (4.15), and $(\tau^* - \tau)(q - cp) \geq 0$ a.e. μ, we conclude that

$$\int |\tau^* - \tau|\,d|\nu_c| = \int (\tau^* - \tau)\,d\nu_c \leq (1+c)\,\delta \tag{4.17}$$

Via the Chebyshev–Markov inequality, (4.16) follows. ////

Acknowledgement I thank P. Ruckdeschel for the numerical computations.

References

[1] Bickel, P.J., Klaassen, C.A.J., Ritov, Y., and Wellner, J.A. (1993): *Efficient and Adaptive Estimation for Semiparametric Models*. Springer, New York.

[2] Donoho, D.L. (1988): One-sided inference about functionals of a density. *Ann. Statist.* **16** 1390–1420.

[3] Hájek, J. and Šidák, Z. (1967): *Theory of Rank Tests*. Academic Press, New York.

[4] Janssen, A. (1999): Testing nonparametric statistical functionals with applications to rank tests. *J. Statist. Plann. Inf.* **81** 71–93.

[5] Pfanzagl, J. and Wefelmeyer, W. (1982): *Contributions to a General Asymptotic Statistical Theory*. Lecture Notes in Statistics #13. Springer, Berlin.

[6] Rieder, H. (1981 a): Robustness of one- and two-sample rank tests against gross errors. *Ann. Statist.* **9** 245–265.

[7] Rieder, H. (1981 b): On local asymptotic minimaxity and admissibility in robust estimation. *Ann. Statist.* **9** 266–277.

[8] Rieder, H. (1994): *Robust Asymptotic Statistics*. Springer, New York.

[9] Rieder, H. (2000): Neighborhoods as nuisance parameters? Robustness vs. semiparametrics. Submitted for publication.

[10] Rudin, W. (1974): *Real and Complex Analysis* (2^{nd} ed.). McGraw-Hill, New York.

[11] van der Vaart, A.W. (1998): *Asymptotic Statistics*. CUP, Cambridge.

Department of Mathematics
University of Bayreuth, NW II
D-95440 Bayreuth, Germany
e-mail: helmut.rieder@uni-bayreuth.de