ASYMPTOTICALLY MOST ACCURATE CONFIDENCE INTERVALS IN THE SEMIPARAMETRIC SYMMETRIC LOCATION MODEL

CHRIS A.J. KLAASSEN
University of Amsterdam

One- and two-sided confidence intervals are considered for the location parameter in the semiparametric symmetric location model. Asymptotic bounds are proved and confidence intervals are constructed that attain these bounds locally asymptotically uniformly. Global uniformity is studied as well.

1. Introduction

Let $\mathcal{G}$ be the class of distribution functions G with densities g w.r.t. Lebesgue measure that are symmetric about 0 and that have finite Fisher information $I(G)$ for location. This means that every $G \in \mathcal{G}$ has a density g satisfying

$$g(-x) = g(x), \quad x \in \mathbb{R}, \tag{1.1}$$

and being absolutely continuous with derivative g' such that

$$I(G) = \int (g'/g)^2 g < \infty \tag{1.2}$$

holds. The semiparametric symmetric location model

$$\mathcal{P} = \{P_{\theta,G} : \theta \in \mathbb{R}, G \in \mathcal{G}\} \tag{1.3}$$

consists of all distributions $P_{\theta,G}$ with density $g(x-\theta)$, $x \in \mathbb{R}$, with respect to Lebesgue measure.

Based on i.i.d. random variables $X_1, \dots, X_n$ with distribution $P_{\theta,G}$ estimation of the location parameter θ is possible by estimator sequences $(T_n)_{n\in\mathbb{N}} = \big(t_n(X_1, \dots, X_n)\big)_{n\in\mathbb{N}}$ satisfying

$$\sqrt{n}(T_n - \theta) \xrightarrow{\mathcal{D}} \mathcal{N}\big(0, I^{-1}(G)\big), \quad \theta \in \mathbb{R}, G \in \mathcal{G}, \tag{1.4}$$

and even

$$\sqrt{n}\left(T_n - \theta + \frac{1}{n}\sum_{i=1}^{n} I^{-1}(G)\frac{g'}{g}(X_i - \theta)\right) \xrightarrow{P_{\theta,G}} 0, \tag{1.5}$$

AMS subject classifications: Primary 62G15; Secondary 62G05, 62G20.
Keywords and phrases: Adaptive estimation, local uniformity, global uniformity, asymptotic bounds.

as $n \to \infty$. An estimator sequence $(\tilde{T}_n)$ is called regular at $P_{\theta,G}$ within $\mathcal{P}_1 = \{P_{\tilde{\theta},G} : \tilde{\theta} \in \mathbb{R}\}$ if there exists a law L_θ such that for every sequence (θ_n) with $\big(\sqrt{n}(\theta_n - \theta)\big)$ bounded

$$\sqrt{n}(\tilde{T}_n - \theta_n) \xrightarrow{\mathcal{D}} L_\theta \tag{1.6}$$

holds under $P_{\theta_n,G}$. By the Hájek–Le Cam convolution theorem such estimator sequences satisfy

$$\sqrt{n}\left(\tilde{T}_n - \theta + \frac{1}{n}\sum_{i=1}^{n} I^{-1}(G)\frac{g'}{g}(X_i - \theta)\right) \xrightarrow{\mathcal{D}} M_\theta \tag{1.7}$$

under $P_{\theta,G}$ with $L_\theta = \mathcal{N}\big(0, I^{-1}(G)\big) * M_\theta$. The asymptotic linearity (1.5) implies (1.4) and even implies regularity of (T_n) within $\mathcal{P}_1$ at $P_{\theta,G}$ with $L_\theta = \mathcal{N}\big(0, I^{-1}(G)\big)$. This follows from local asymptotic normality within $\mathcal{P}_1$. Therefore, estimator sequences (T_n) with the asymptotic linearity property (1.5) are called asymptotically efficient; see, e.g., Section 2.3 of Bickel, Klaassen, Ritov, and Wellner (1993), henceforth referred to as BKRW (1993). Such estimator sequences are also asymptotically optimal within the class of all (not just regular) estimator sequences in that they attain the lower bound in the local asymptotic minimax theorem.

As above, fix θ and G. Furthermore, let V be a random variable and define the random location parameter

$$\vartheta = \theta + \frac{\sigma}{\sqrt{n}}V. \tag{1.8}$$

Given $\vartheta = \tilde{\theta}$ the common distribution of $X_1, \ldots, X_n$ is $P_{\tilde{\theta},G}$. The limit behavior of the sequence $\big(\sqrt{n}(T_n - \vartheta)\big)$ is governed by the local asymptotic behavior of (T_n) around $P_{\theta,G}$. Under weak conditions on the distribution of V the finite sample spread inequality (Klaassen, 1989a) implies that as $n \to \infty$ and subsequently $\sigma \to \infty$, all possibly defective limit points of the sequence of distributions of $\sqrt{n}(T_n - \vartheta)$ are at least as spread out as a normal distribution with variance $I^{-1}(G)$; see Klaassen (1989b). One distribution is at least as spread out as another if any two quantiles of the first distribution are at least as far apart as the corresponding quantiles of the latter; cf. Bickel and Lehmann (1975). Asymptotically linear estimators satisfying (1.5) are efficient also in this local asymptotic spread framework. Le Cam's one step estimators based on a discretized preliminary estimator and on knowledge of the underlying density g, are asymptotically linear in the sense of (1.5) for every $G \in \mathcal{G}$. For strongly unimodal g, for which g'/g is nonincreasing, it is particularly easy to see that the maximum likelihood estimator is efficient.

Stein (1956) has noted that it should be possible to construct estimators T_n satisfying (1.5) for all $G \in \mathcal{G}$ simultaneously. van Eeden (1970) was the first to construct estimators T_n satisfying (1.5) for all strongly unimodal g, $G \in \mathcal{G}$, simultaneously; she called her estimator efficiency-robust and noted the analogy to the uniformly asymptotically efficient tests of Hájek (1962). In fact, she constructed the first semiparametrically efficient estimator, namely for the semiparametric model

$$\{P_{\theta,G} : \theta \in \mathbb{R}, G \in \mathcal{G}, G \text{ has a strongly unimodal density}\}. \tag{1.9}$$

Stone (1975) and Beran (1974, 1978) constructed such estimators for the complete semiparametric model $\mathcal{P}$. These estimators were called adaptive, since they adapted, so to say, to the underlying density g. This model and these semiparametrically efficient estimators stimulated the development of semiparametric estimation theory, which includes regression models, Cox' proportional hazards model, transformation models, and many more; see BKRW (1993). In fact, van Eeden (1970) also constructed an adaptive, efficiency-robust, semiparametrically efficient estimator of the shift parameter in the two-sample location model for strongly unimodal distributions.

Given these fully efficient estimators of the location parameter θ within the semiparametric symmetric location model $\mathcal{P}$, a natural next step is the construction of asymptotically optimal confidence intervals for θ. In Section 2 we will derive asymptotic bounds to the performance of one-sided $(1-\alpha)$-confidence intervals and of two-sided asymptotically unbiased $(1-\alpha)$-confidence intervals. Constructions of efficient confidence intervals, which attain these asymptotic lower bounds, will be given in Section 4. It is well known that there is a close relationship between confidence intervals and testing hypotheses. Choi, Hall, and Schick (1996) have developed a general theory for semiparametric hypothesis testing and they have derived the asymptotic performance of asymptotically uniformly most powerful (unbiased) tests. In their Section 6 they note that these results can be translated into optimality results for confidence intervals and in their Example 8.1(a) they treat the symmetric location case. Strictly speaking, our results can be derived from theirs. However, we have chosen to present explicit results and self-contained, quite straightforward proofs in order to set the proper stage for our discussion of uniformity. Indeed, in Sections 5 and 6 we will discuss local and global uniformity issues for our confidence intervals. In our approach we keep the confidence level $1-\alpha$ fixed. Asymptotically efficient fixed width confidence intervals for model (1.3) have been given by Martinsek (1991) and Chang (1992).

2. Parametric asymptotic bounds

Fix $G \in \mathcal{G}$. Within the parametric symmetric location model

$$\mathcal{P}_1 = \{P_{\theta,G} : \theta \in \mathbb{R}\} \tag{2.1}$$

we consider random intervals I_n that are based on $X_1, \dots, X_n$ and that have asymptotic coverage probability at least $1 - \alpha$ for a small positive value α at $\theta_0 \in \mathbb{R}$, i.e.

$$\liminf_{n\to\infty} P_{\theta_0,G}(\theta_0 \in I_n) \geq 1 - \alpha. \tag{2.2}$$

First we will study one-sided intervals of type $I_n = (-\infty, A_n]$ that have locally asymptotically uniform coverage probability at least $1 - \alpha$ at $\theta_0 \in \mathbb{R}$ in the sense that for all sequences $(\theta_n)_{n\in\mathbb{N}}$ with $\left(\sqrt{n}(\theta_n - \theta_0)\right)_{n\in\mathbb{N}}$ bounded,

$$\liminf_{n\to\infty} P_{\theta_n,G}(\theta_n \in I_n) \geq 1 - \alpha \tag{2.3}$$

holds.

Theorem 2.1. *Fix $G \in \mathcal{G}$, $\theta_0 \in \mathbb{R}$. If the sequence of one-sided confidence intervals $I_n = (-\infty, A_n]$ has locally asymptotically uniform coverage probability at least $1 - \alpha$ at θ_0 in the sense of* (2.3), *then for every $v > 0$,*

$$\liminf_{n\to\infty} P_{\theta_0,G}\left(\theta_0 + \frac{v}{\sqrt{nI(G)}} \in I_n\right) \geq \Phi(\Phi^{-1}(1-\alpha) - v) \tag{2.4}$$

holds.

Proof. By the standard relationship between confidence sets and testing theory we reject the hypothesis that the location parameter θ equals $\theta_n = \theta_0 + v/\sqrt{nI(G)}$ iff $\theta_n \notin I_n$. In view of (2.3) this yields a sequence of size α_n tests with $\limsup_{n\to\infty} \alpha_n \leq \alpha$. We compare this sequence to the sequence of most powerful level α_n tests for the null hypothesis $\theta = \theta_n$ against the simple alternative $\theta = \theta_0$, which rejects if, for some nonnegative constant $c_{0,n}$,

$$\prod_{i=1}^{n} g(X_i - \theta_0) - c_{0,n} \prod_{i=1}^{n} g(X_i - \theta_n) \geq 0 \tag{2.5}$$

holds with possibly randomization at equality. Now, under $P_{\tilde{\theta}_n,G}$ with $\tilde{\theta}_n = \theta_0 + w/\sqrt{nI(G)}$ the left hand side of (2.5) behaves asymptotically as

$$\begin{aligned}\prod_{i=1}^{n} g(X_i - \tilde{\theta}_n)\Bigg\{&\exp\left[-\frac{w}{\sqrt{nI(G)}} \sum_{i=1}^{n} -\frac{g'}{g}(X_i - \tilde{\theta}_n) - \frac{1}{2}w^2\right] \\ &- c_{0,n} \exp\left[\frac{v-w}{\sqrt{nI(G)}} \sum_{i=1}^{n} -\frac{g'}{g}(X_i - \tilde{\theta}_n) - \frac{1}{2}(v-w)^2\right]\Bigg\}\end{aligned} \tag{2.6}$$

in the sense that the ratio of both expressions tends to 1 in probability, by the local asymptotic normality property of regular parametric families; see

Example 2.1.2 and Proposition 2.1.2. of BKRW (1993). The second factor behaves asymptotically, with $U \sim \mathcal{N}(0,1)$, as

$$\exp[-wU - \tfrac{1}{2}w^2] - c_{0,n}\exp[(v-w)U - \tfrac{1}{2}(v-w)^2]. \tag{2.7}$$

From (2.5), (2.6), and (2.7) we obtain

$$\begin{aligned}
\alpha &\geq \limsup_{n\to\infty} P_{\theta_n,G}\Big(\prod_{i=1}^n g(X_i-\theta_0) - c_{0,n}\prod_{i=1}^n g(X_i-\theta_n) > 0\Big)\\
&= P(\exp[-vU - \tfrac{1}{2}v^2] > c_0)\\
&= P\Big(U < -\frac{\log c_0}{v} - \frac{1}{2}v\Big)
\end{aligned}$$

and hence

$$-\log c_0 \leq v\Phi^{-1}(\alpha) + \tfrac{1}{2}v^2$$

with $c_0 = \liminf_{n\to\infty} c_{0,n}$. This implies

$$\begin{aligned}
\limsup_{n\to\infty} P_{\theta_0,G}\Big(\prod_{i=1}^n g(X_i-\theta_0) - c_{0,n}\prod_{i=1}^n g(X_i-\theta_n) \geq 0\Big)&\\
&= P(1 - c_0\exp[vU - \tfrac{1}{2}v^2] \geq 0)\\
&\leq P(0 \geq -v\Phi^{-1}(\alpha) - v^2 + vU)\\
&= P(U \leq \Phi^{-1}(\alpha) + v)
\end{aligned}$$

and hence (2.4). □

Note that inequality (2.4) of Theorem 2.1 may be interpreted also as an inequality for the overshoot $A_n - \theta_0$. Furthermore, with

$$\hat{\theta}_n = A_n + \frac{1}{\sqrt{nI(G)}}\Phi^{-1}(\alpha) \tag{2.8}$$

it may be rewritten as

$$\limsup_{n\to\infty} P_{\theta_0,G}\big(\sqrt{nI(G)}(\hat{\theta}_n - \theta_0) \leq y\big) \leq \Phi(y), \quad y > \Phi^{-1}(\alpha), \tag{2.9}$$

which states that $\hat{\theta}_n$ may be viewed as an estimator of θ and that the right tail of $\hat{\theta}_n$ is asymptotically under $P_{\theta_0,G}$ stochastically larger than a standard normal distribution, provided $\hat{\theta}_n$ is standardized to $\sqrt{nI(G)}(\hat{\theta}_n - \theta_0)$. Note also that (2.8) suggests a way to construct confidence intervals from estimators, as we will exploit in Section 4.

We will call a sequence (I_n) of confidence intervals with asymptotic coverage probability at least $1-\alpha$ locally asymptotically unbiased at θ_0 if for every sequence (θ_n) with $\big(\sqrt{n}(\theta_n-\theta_0)\big)$ bounded

$$\limsup_{n\to\infty} P_{\theta_n,G}(\theta_0\in I_n)\le 1-\alpha,\quad G\in\mathcal{G}, \tag{2.10}$$

holds. For two-sided confidence intervals $[B_n,C_n]$ the following analogue of Theorem 2.1 holds.

Theorem 2.2. *Fix $0<\alpha<1$, $G\in\mathcal{G}$, and $\theta_0\in\mathbb{R}$, and write $I_n=[B_n,C_n]$. If for a sequence of such two-sided intervals* (2.2) *and* (2.10) *hold, then for all $v\in\mathbb{R}$ the inequality*

$$\begin{aligned}\liminf_{n\to\infty} & P_{\theta_0+v/\sqrt{nI(G)},G}(\theta_0\in I_n)\\ &\ge \Phi\left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)+v\right)+\Phi\left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)-v\right)-1\end{aligned} \tag{2.11}$$

holds.

Proof. Fix $\theta_0\in\mathbb{R}$, $G\in\mathcal{G}$, $0<\alpha<1$, and choose $v>0$ without loss of generality. By the standard relation between confidence sets and testing theory we reject the hypothesis that the location parameter θ equals θ_0 iff $\theta_0\notin I_n$. In view of (2.2) and (2.10) this yields a sequence of size α_n-tests with $\alpha_n\to\alpha$, asymptotically. Let u be arbitrary and define $\theta_n=\theta_0-u/\sqrt{nI(G)}$, $\theta_n'=\theta_0+v/\sqrt{nI(G)}$, and

$$\beta_n=P_{\theta_n,G}(\theta_0\notin I_n). \tag{2.12}$$

By (2.10) we have

$$\liminf_{n\to\infty}\beta_n\ge\alpha. \tag{2.13}$$

We will compare this sequence of tests to the sequence of most powerful level α_n-tests φ_n for the null hypothesis $\theta=\theta_0$ against the alternative $\theta=\theta_n'$ under the side condition

$$E_{\theta_n,G}\varphi_n(X_1,\dots,X_n)=\beta_n. \tag{2.14}$$

By the Neyman–Pearson lemma the test φ_n rejects iff

$$\prod_{i=1}^n g(X_i-\theta_n')+k_{0,n}(u)\prod_{i=1}^n g(X_i-\theta_0)+k_{1,n}(u)\prod_{i=1}^n g(X_i-\theta_n)\ge 0 \tag{2.15}$$

(with possibly randomization at equality), where $k_{0,n}(u)$ and $k_{1,n}(u)$ are determined by the level α_n and by (2.14). Fix $w\in\mathbb{R}$. Since g has finite Fisher information for location by assumption, the corresponding location family of distributions is regular and hence it has the local asymptotic normality property. Consequently, under $P_{\tilde{\theta}_n,G}$ with $\tilde{\theta}_n = \theta_0 + w/\sqrt{nI(G)}$, the left hand side of (2.15) divided by $\prod_{i=1}^n g(X_i - \tilde{\theta}_n)$ behaves asymptotically as

$$\begin{aligned}(2.16) \quad \exp[-(v-w)Z - \tfrac{1}{2}(v-w)^2] + k_0(u)\exp[wZ - \tfrac{1}{2}w^2] \\ + k_1(u)\exp[(u+w)Z - \tfrac{1}{2}(u+w)^2]\end{aligned}$$

with Z a standard normal random variable, where for each subsequence (n') of (n) there exists a further subsequence (n'') such that $k_0(u)$ and $k_1(u)$ exist as limits of $\big(k_{0,n''}(u)\big)$ and $\big(k_{1,n''}(u)\big)$, respectively. Since the subsequent argument is valid for each subsequence (n'), we may assume without loss of generality $(n'') = (n)$, $\beta_n \to \beta \geq \alpha$, and also $\gamma_n \to \gamma$, where

$$(2.17) \qquad \gamma_n = E_{\theta_n',g}\varphi_n(X_1,\ldots,X_n)$$

denotes the power at θ_n' of the most powerful test.

Taking $w = 0, -u$, and v, respectively, we obtain from (2.16), $\alpha_n \to \alpha$, (2.14), $\beta_n \to \beta$, (2.17), and $\gamma_n \to \gamma$ the system of equations

$$(2.18) \qquad \begin{aligned} \alpha &= P(\psi_u(Z+v) \geq 0), \\ \beta &= P(\psi_u(Z+v+u) \geq 0), \\ \gamma &= P(\psi_u(Z) \geq 0), \\ \psi_u(y) &= \exp[-vy + \tfrac{1}{2}v^2] + k_0(u) \\ &\qquad + k_1(u)\exp[uy - \tfrac{1}{2}u^2 - uv], \quad y \in \mathbb{R}. \end{aligned}$$

Since v is positive, $\psi_u(\cdot)$ is either decreasing or decreasing-increasing depending on the sign of $uk_1(u)$. In any case, $\psi_u(y) \geq 0$ is equivalent to

$$(2.19) \qquad y \leq b_u \text{ or } y \geq c_u$$

for some $b_u \leq c_u \in (-\infty, \infty]$. It follows that the first two equations from (2.18) become

$$(2.20) \qquad \begin{aligned} \alpha &= \Phi(b_u - v) + \Phi(-c_u + v), \\ \beta &= \Phi(b_u - v - u) + \Phi(-c_u + v + u). \end{aligned}$$

In view of $\alpha < 1$ the first equality in (2.20) shows that the smallest value among the arguments $b_u - v$ and $-c_u + v$ is negative and largest in absolute value. This implies that the function $w \mapsto \Phi(b_u - v - w) + \Phi(-c_u + v + w)$ is increasing near $w = 0$ in case of $b_u - v < -c_u + v$ and decreasing in case of $b_u - v > -c_u + v$. Consequently, (2.13) and (2.20) show that for small positive values of u the inequality $b_u - v \leq -c_u + v$ has to hold and for negative values of u close to 0 the inequality $b_u - v \geq -c_u + v$ has to hold. It follows that as $u \to 0$ any limit point (b_0, c_0) satisfies

$$b_0 - v = -c_0 + v, \tag{2.21}$$

which by the first equality in (2.20) has to be equal to $\Phi^{-1}(\alpha/2)$. Combining this with the third equation of (2.18) and with (2.19), we arrive with $\gamma = \gamma_u$ at

$$\begin{aligned}\lim_{u\to 0} \gamma_u &= \lim_{u\to 0} \big(\Phi(b_u) + \Phi(-c_u)\big) \\ &= 2 - \Phi\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) + v\right) - \Phi\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) - v\right).\end{aligned} \tag{2.22}$$

Since the tests ϕ_n are most powerful, $1 - \gamma_u$ is a lower bound to the left hand side of (2.11) for all u. Together with (2.22) this proves (2.11) and the theorem. □

Of course, there is a close relationship between Theorem 2.2 and the unbiased test in the normal location model for the simple hypothesis $\theta = 0$ against the alternative $\theta \neq 0$. If we would have strengthened (2.10) to

$$\lim_{n\to\infty} \frac{\partial}{\partial \theta_0} P_{\theta_0, G}(\theta_0 \in I_n) = 0, \tag{2.23}$$

this relationship would have been still more apparent. With our much weaker condition (2.10) the extra limiting procedure $u \downarrow 0$ has been needed in the proof.

As we have seen, these bounds in the theory of confidence intervals are based on the Neyman–Pearson Lemma. Therefore, they are easier to prove than the bounds in estimation theory as discussed in the introduction.

3. Semiparametric asymptotic bounds

Our focus is on the semiparametric symmetric location model

$$\mathcal{P} = \{P_{\theta,G} : \theta \in \mathbb{R}, G \in \mathcal{G}\}$$

from (1.3), in which we consider random intervals I_n that are based on $X_1, \ldots, X_n$. Of course, the bounds from Theorems 2.1 and 2.2 are still valid for every $\theta_0 \in \mathbb{R}$ and every $G \in \mathcal{G}$. Like in estimation theory, they are sharp in the sense that there exist sequences $(I_n)_{n\in\mathbb{N}}$ of confidence intervals for the location parameter θ that attain equality for all $G \in \mathcal{G}$ simultaneously. Such sequences of confidence intervals will be constructed in the next Section. These sequences might be called adaptive. Needless to say, that in the generic semiparametric model this phenomenon of adaptiveness does not occur, typically.

4. Asymptotically Most Accurate Confidence Intervals

4.1. G known, one-sided

Let G be known and let $\hat{\theta}_{G,n}$ be an efficient estimator of the location parameter θ based on i.i.d. random variables $X_1, \ldots, X_n$ with distribution function $G(\cdot - \theta)$ and Fisher information $I(G)$. This means that for every $\theta_0 \in \mathbb{R}$ and every sequence $(\theta_n)_{n\in\mathbb{N}}$ with $\left(\sqrt{n}(\theta_n - \theta_0)\right)_{n\in\mathbb{N}}$ bounded, under $P_{\theta_n,G}$,

$$\sqrt{nI(G)}\left(\hat{\theta}_{G,n} - \theta_n\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad \text{as } n \to \infty; \tag{4.1}$$

see, e.g., Section 2.3 of BKRW (1993).

With

$$A_{G,n} = \hat{\theta}_{G,n} - \frac{1}{\sqrt{nI(G)}}\Phi^{-1}(\alpha), \tag{4.2}$$

as suggested by (2.8), the one-sided confidence interval $I_n = (-\infty, A_{G,n}]$ is asymptotically most accurate at level $1-\alpha$ in the sense that it attains equality in both (2.3) and (2.4) of Theorem 2.1. Indeed, by (4.1) we have

$$\begin{aligned} \lim_{n\to\infty} P_{\theta_n,G}(\theta_n \in I_n) &= \lim_{n\to\infty} P_{\theta_n,G}\big(\sqrt{nI(G)}(\hat{\theta}_{G,n} - \theta_n) \geq \Phi^{-1}(\alpha)\big) \\ &= 1 - \Phi\big(\Phi^{-1}(\alpha)\big) = 1 - \alpha \end{aligned} \tag{4.3}$$

and

$$\begin{aligned} &\lim_{n\to\infty} P_{\theta_0,G}\left(\theta_0 + \frac{v}{\sqrt{nI(G)}} \in I_n\right) \\ &\qquad = \lim_{n\to\infty} P_{\theta_0,G}(\sqrt{nI(G)}(\hat{\theta}_{G,n} - \theta_0) \geq \Phi^{-1}(\alpha) + v) \\ &\qquad = \Phi(\Phi^{-1}(1-\alpha) - v). \end{aligned} \tag{4.4}$$

4.2. G known, two-sided

Similarly, we may construct asymptotically optimal two-sided confidence intervals. With

$$B_{G,n} = \hat{\theta}_{G,n} - \frac{1}{\sqrt{nI(G)}}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \tag{4.5}$$

and

$$C_{G,n} = \hat{\theta}_{G,n} - \frac{1}{\sqrt{nI(G)}}\Phi^{-1}\left(\frac{\alpha}{2}\right) \tag{4.6}$$

the two-sided confidence interval $I_n = [B_{G,n}, C_{G,n}]$ is asymptotically most accurate at level $1-\alpha$ in the sense that it satisfies (2.10) and attains equality in (2.2) and (2.11) of Theorem 2.2. Indeed, by (4.1) we have, as above,

(4.7)
$$\begin{aligned}
&\lim_{n\to\infty} P_{\theta_0,G}(\theta_0 \in I_n)\\
&\quad = \lim_{n\to\infty} P_{\theta_0,G}\left(\Phi^{-1}\left(\frac{\alpha}{2}\right) \le \sqrt{nI(G)}(\hat\theta_{G,n} - \theta_0) \le \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right)\\
&\quad = \Phi\left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)\right) - \Phi\left(\Phi^{-1}\left(\frac{\alpha}{2}\right)\right) = 1-\alpha,
\end{aligned}$$

(4.8)
$$\begin{aligned}
&\limsup_{n\to\infty} P_{\theta_n,G}(\theta_0 \in I_n)\\
&\quad = \limsup_{n\to\infty}\Big\{ P_{\theta_n,G}\left(\sqrt{nI(G)}(\hat\theta_{G,n}-\theta_n) \le \Phi^{-1}\left(1-\frac{\alpha}{2}\right) + \sqrt{nI(G)}(\theta_0-\theta_n)\right)\\
&\qquad\qquad - P_{\theta_n,G}\left(\sqrt{nI(G)}(\hat\theta_{G,n}-\theta_n) \le \Phi^{-1}\left(\frac{\alpha}{2}\right) + \sqrt{nI(G)}(\theta_0-\theta_n)\right)\Big\}\\
&\quad \le \sup_{x\in\mathbb{R}}\left(\Phi\left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)+x\right) - \Phi\left(\Phi^{-1}\left(\frac{\alpha}{2}\right)+x\right)\right) = 1-\alpha,
\end{aligned}$$

and

$$\begin{aligned}
(4.9)\quad &\lim_{n\to\infty} P_{\theta_0,G}\left(\theta_0 + \frac{v}{\sqrt{nI(G)}} \in I_n\right)\\
&\quad = \lim_{n\to\infty} P_{\theta_0,G}\left(\Phi^{-1}\left(\frac{\alpha}{2}\right) + v \le \sqrt{nI(G)}(\hat\theta_{G,n} - \theta_0) \le \Phi^{-1}\left(1-\frac{\alpha}{2}\right) + v\right)\\
&\quad = \Phi\left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right)+v\right) - \Phi\left(\Phi^{-1}\left(\frac{\alpha}{2}\right)+v\right),
\end{aligned}$$

for $v > 0$.

4.3. *G* unknown, one-sided

If G is unknown, we have the semiparametric model $\mathcal{P}$ from (1.3) and estimation of the location parameter θ with the same asymptotic performance as for the case G known, is still possible as mentioned in the Introduction; see also Example 7.8.1 of BKRW (1993). Consequently, there exists a sequence $(\hat\theta_n)_{n\in\mathbb{N}}$ of estimators of θ such that for every $G \in \mathcal{G}$, for every $\theta_0 \in \mathbb{R}$, and every sequence $(\theta_n)_{n\in\mathbb{N}}$ with $\left(\sqrt{n}(\theta_n - \theta_0)\right)_{n\in\mathbb{N}}$ bounded, under $P_{\theta_n,G}$ the convergence

$$(4.10)\qquad \sqrt{nI(G)}(\hat\theta_n - \theta_n) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad \text{as } n \to \infty,$$

holds. To construct confidence intervals of the location parameter along the lines of (4.2), (4.5), and (4.6) for the semiparametric model $\mathcal{P}$, we may use this semiparametrically efficient estimator sequence, but we also have to estimate the unknown Fisher information $I(G)$ consistently. However, estimation of the Fisher information hardly causes additional difficulties. In fact, all constructions of semiparametrically efficient estimators of θ implicitly generate or explicitly use estimators of $I(G)$; see, e.g., Lemma 3.3 of van Eeden (1970) and Theorem 7.8.1 and Example 7.8.1, continued, of BKRW (1993). Essentially, existence of semiparametrically efficient estimators of θ is equivalent to existence of $\sqrt{n}$-unbiased, consistent estimators of the efficient influence function $-I^{-1}(G)g'/g(\cdot)$; see Klaassen (1987). Let $(\hat{I}_n)_{n\in\mathbb{N}}$ be a locally uniformly consistent sequence of estimators of $I(G)$, i.e., for every $G \in \mathcal{G}$, for every $\theta_0 \in \mathbb{R}$, and every sequence $(\theta_n)_{n\in\mathbb{N}}$ with $\left(\sqrt{n}(\theta_n - \theta_0)\right)_{n\in\mathbb{N}}$ bounded,

$$\hat{I}_n \xrightarrow{P_{\theta_n,G}} I(G), \quad n \to \infty, \tag{4.11}$$

holds. The existence of such an estimator sequence is guaranteed by e.g. Proposition 7.8.1 of BKRW (1993) and the contiguity of $(P_{\theta_n,G})_{n\in\mathbb{N}}$ and $(P_{\theta_0,G})_{n\in\mathbb{N}}$, which is implied by local asymptotic normality, which in turn is a consequence of the finiteness of $I(G)$. From (4.10) and (4.11) it follows that for every $G \in \mathcal{G}$, for every $\theta_0 \in \mathbb{R}$, and every sequence $(\theta_n)_{n\in\mathbb{N}}$ with $\left(\sqrt{n}(\theta_n - \theta_0)\right)_{n\in\mathbb{N}}$ bounded, under $P_{\theta_n,G}$,

$$\sqrt{n\hat{I}_n}(\hat{\theta}_n - \theta_n) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad \text{as } n \to \infty. \tag{4.12}$$

Using (4.10)–(4.12) we define

$$A_n = \hat{\theta}_n - \frac{1}{\sqrt{n\hat{I}_n}}\Phi^{-1}(\alpha) \tag{4.13}$$

and we note that the one-sided confidence interval $I_n = (-\infty, A_n]$ is asymptotically most accurate at level $1 - \alpha$ in the sense that it attains equality in both (2.3) and (2.4) of Theorem 2.1 for all $G \in \mathcal{G}$ simultaneously. Indeed, by (4.12) we have

$$\begin{aligned}\lim_{n\to\infty} P_{\theta_n,G}(\theta_n \in I_n) &= \lim_{n\to\infty} P_{\theta_n,G}\left(\sqrt{n\hat{I}_n}(\hat{\theta}_n - \theta_n) \geq \Phi^{-1}(\alpha)\right) \\ &= 1 - \Phi\left(\Phi^{-1}(\alpha)\right) = 1 - \alpha\end{aligned} \tag{4.14}$$

and, by (4.12) and (4.11), we obtain

$$\begin{aligned}(4.15)\quad \lim_{n\to\infty} P_{\theta_0,G}\left(\theta_0 + \frac{v}{\sqrt{nI(G)}} \in I_n\right) &= \lim_{n\to\infty} P_{\theta_0,G}\left(\sqrt{n\hat{I}_n}(\hat{\theta}_n - \theta_0) \geq \Phi^{-1}(\alpha) + v\sqrt{\frac{\hat{I}_n}{I(G)}}\right) \\ &= \Phi(\Phi^{-1}(1-\alpha) - v).\end{aligned}$$

4.4. G unknown, two-sided

Similarly, we may construct semiparametric, asymptotically optimal two-sided confidence intervals. With

$$(4.16)\qquad B_n = \hat{\theta}_n - \frac{1}{\sqrt{n\hat{I}_n}}\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

and

$$(4.17)\qquad C_n = \hat{\theta}_n - \frac{1}{\sqrt{n\hat{I}_n}}\Phi^{-1}\left(\frac{\alpha}{2}\right)$$

the two-sided confidence interval $I_n = [B_n, C_n]$ is asymptotically most accurate at level $1-\alpha$ in the sense that it satisfies (2.10) and attains equality in (2.2) and (2.11) of Theorem 2.2 for all $G \in \mathcal{G}$ simultaneously. Indeed, by (4.12) and (4.11) we have, as above,

(4.18)

$$\begin{aligned}&\lim_{n\to\infty} P_{\theta_0,G}(\theta_0 \in I_n) \\ &= \lim_{n\to\infty} P_{\theta_0,G}\left(\Phi^{-1}\left(\frac{\alpha}{2}\right) \leq \sqrt{n\hat{I}_n}(\hat{\theta}_{G,n} - \theta_0) \leq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right) \\ &= \Phi\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right) - \Phi\left(\Phi^{-1}\left(\frac{\alpha}{2}\right)\right) = 1 - \alpha,\end{aligned}$$

(4.19)

$$\begin{aligned}&\limsup_{n\to\infty} P_{\theta_n,G}(\theta_0 \in I_n) \\ &= \limsup_{n\to\infty}\Big\{P_{\theta_n,G}\left(\sqrt{n\hat{I}_n}(\hat{\theta}_{G,n} - \theta_n) \leq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) + \sqrt{n\hat{I}_n}(\theta_0 - \theta_n)\right) \\ &\qquad - P_{\theta_n,G}\left(\sqrt{n\hat{I}_n}(\hat{\theta}_{G,n} - \theta_n) \leq \Phi^{-1}\left(\frac{\alpha}{2}\right) + \sqrt{n\hat{I}_n}(\theta_0 - \theta_n)\right)\Big\} \\ &\leq \sup_{x\in\mathbb{R}}\left(\Phi\left(\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) + x\right) - \Phi\left(\Phi^{-1}\left(\frac{\alpha}{2}\right) + x\right)\right) = 1 - \alpha,\end{aligned}$$

and

$$(4.20)\quad \lim_{n\to\infty} P_{\theta_0,G}\Big(\theta_0 + \frac{v}{\sqrt{nI(G)}} \in I_n\Big)$$
$$= \lim_{n\to\infty} P_{\theta_0,G}\Big(\Phi^{-1}\Big(\frac{\alpha}{2}\Big) + v\sqrt{\frac{\hat{I}_n}{I(G)}} \le \sqrt{n\hat{I}_n}(\hat{\theta}_n - \theta_0) \le \Phi^{-1}\Big(1-\frac{\alpha}{2}\Big) + v\sqrt{\frac{\hat{I}_n}{I(G)}}\Big)$$
$$= \Phi\Big(\Phi^{-1}\Big(1-\frac{\alpha}{2}\Big) + v\Big) - \Phi\Big(\Phi^{-1}\Big(\frac{\alpha}{2}\Big) + v\Big),$$

for $v > 0$.

Semiparametrically efficient estimators are asymptoticallly linear, which implies (4.10) and (4.12) locally uniformly in G. This local uniformity is inherited by our semiparametric confidence intervals as we will discuss in the next section.

5. Local uniformity

Fisher information for location $G \mapsto I(G)$ is lower semicontinuous; see, e.g., Definition 4.1, Theorem 4.2, and the paragraph below the proof of Theorem 4.2 in Huber (1981). In fact, for any G with finite Fisher information $I(G)$ and with density g, and for any $\epsilon > 0$ there exists a density g_ϵ with distribution function G_ϵ and with

$$(5.1)\qquad \int \big(g_\epsilon^{1/2} - g^{1/2}\big)^2 < \epsilon, \quad I(G_\epsilon) > 1/\epsilon.$$

This phenomenon has been used in Klaassen (1979, 1980) to prove that adaptive estimators of location cannot converge uniformly; more precisely he showed that for any $n \in \mathbb{N}$, for any $\theta_0 \in \mathbb{R}$, and for any translation equivariant estimator $\hat{\theta}_n$ of the location parameter θ

$$(5.2)\qquad \sup_{\epsilon>0}\sup_{x\in\mathbb{R}} \big|P_{\theta_0,G_\epsilon}\big(\sqrt{nI(G_\epsilon)}(\hat{\theta}_n - \theta_0) \le x\big) - \Phi(x)\big| = \frac{1}{2}$$

holds.

For semiparametric confidence intervals satisfying (2.2) the strict lower semicontinuity property (5.1) has dramatic consequences as well. In view of

$$(5.3)\qquad \sup_{\epsilon>0} P_{\theta_0+v/\sqrt{nI(G_\epsilon)},G_\epsilon}(\theta_0 \in I_n) \ge \lim_{\epsilon\downarrow 0} P_{\theta_0+v/\sqrt{nI(G_\epsilon)},G_\epsilon}(\theta_0 \in I_n) = P_{\theta_0,G}(\theta_0 \in I_n)$$

these intervals satisfy

$$(5.4)\qquad \liminf_{n\to\infty}\sup_{\epsilon>0} P_{\theta_0+v/\sqrt{nI(G_\epsilon)},G_\epsilon}(\theta_0 \in I_n) \ge \liminf_{n\to\infty} P_{\theta_0,G}(\theta_0 \in I_n) \ge 1-\alpha,$$

and consequently the convergences (4.15) and (4.20) are as far from uniform as possible.

Semiparametric estimation theory is based on a study of the local asymptotic properties of estimators within parametric submodels that are regular at a value of the parameter. A k-dimensional parametric model

$$\mathcal{P}_H = \{P_\eta : \eta \in H\}, \quad H \subset \mathbb{R}^k, \tag{5.5}$$

of probability measures dominated by a measure μ, is regular at η_0 if η_0 is an interior point of H and if the map $\eta \mapsto \sqrt{dP_\eta/d\mu}$ from H to $\mathcal{L}_2(\mu)$ is continuously Fréchet (or equivalently Hadamard) differentiable of rank k at η_0. Such parametric models have the local asymptotic normality property, which implies the contiguity

$$(P^n_{\eta_n})_{n\in\mathbb{N}} \lhd\rhd (P^n_{\eta_0})_{n\in\mathbb{N}}, \quad \big(\sqrt{n}(\eta_n - \eta_0)\big)_{n\in\mathbb{N}} \text{ bounded.} \tag{5.6}$$

Estimator sequences $(\hat\theta_n)_{n\in\mathbb{N}}$ within a general semiparametric model $\mathcal{P} = \{P_{\theta,G} : \theta \in \mathbb{R}, G \in \mathcal{G}\}$ are called semiparametrically locally asymptotically efficient at P_{θ_0,G_0}, or efficient at P_{θ_0,G_0} for short, if there exists a number I_0 such that for every regular parametric submodel

$$\{P_{\theta,G_\eta} : \theta \in \mathbb{R}, \eta \in H\}, \quad 0 \in H \subset \mathbb{R}^k \text{ open,} \tag{5.7}$$

of $\mathcal{P}$

$$\sqrt{nI_0}(\hat\theta_n - \theta_n) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \text{ under } P_{\theta_n,G_{\eta_n}}, \tag{5.8}$$

holds for all sequences (θ_n) and (η_n) with both $\big(\sqrt{n}(\theta_n - \theta_0)\big)$ and $\big(\sqrt{n}\eta_n\big)$ bounded. In our semiparametric symmetric location model $\mathcal{P}$ from (1.3) efficient estimator sequences exist and they satisfy (5.8) with $I_0 = I(G_0)$; see the Introduction for references. Efficiency is proved by showing that the estimator sequence is asymptotically linear in the sense of (1.5) under P_{θ_0,G_0} and by invoking Le Cam's third lemma based on the local asymptotic normality of regular parametric submodels.

For our one-sided semiparametric confidence intervals property (5.8) of efficient estimators yields the following local asymptotic result.

Theorem 5.1. *Let $(\hat\theta_n)_{n\in\mathbb{N}}$ be an efficient sequence of estimators of the location parameter θ in the symmetric location model $\mathcal{P}$ from (1.3). Let $\hat I_n$ be a consistent estimator of the Fisher information for location $I(G)$ satisfying (4.11) with $\theta_n = \theta_0$. Fix $\alpha \in (0,1)$ and define*

$$I_n = \left(-\infty, \hat\theta_n - \frac{1}{\sqrt{n\hat I_n}}\Phi^{-1}(\alpha)\right]. \tag{5.9}$$

For every $\theta_0 \in \mathbb{R}$ and $G_0 \in \mathcal{G}$ and for every regular parametric submodel of $\mathcal{P}$ of type (5.7)

$$\lim_{c\to\infty} \limsup_{n\to\infty} \sup_{\sqrt{n}|\eta|\le c} P_{\theta_0,G_\eta}\left(\theta_0 + \frac{v}{\sqrt{nI(G_\eta)}} \in I_n\right) = \Phi(\Phi^{-1}(1-\alpha) - v), \quad v > 0, \tag{5.10}$$

holds.

Proof. Fix $c > 0$ and let $(\eta_n)_{n\in\mathbb{N}}$, $\sqrt{n}|\eta_n| < c$, be such that

$$\lim_{n\to\infty} P_{\theta_0,G_{\eta_n}}\left(\theta_0 + \frac{v}{\sqrt{nI(G_{\eta_n})}} \in I_n\right) = \limsup_{n\to\infty} \sup_{\sqrt{n}|\eta|\le c} P_{\theta_0,G_\eta}\left(\theta_0 + \frac{v}{\sqrt{nI(G_\eta)}} \in I_n\right). \tag{5.11}$$

The contiguity (5.6) with $\eta_0 = 0$ and the continuity on the chosen regular parametric submodel of the map $G \mapsto I(G)$ imply

$$\hat{I}_n \xrightarrow{P_{\theta_0,G_{\eta_n}}} I(G_0), \quad \frac{\hat{I}_n}{I(G_{\eta_n})} \xrightarrow{P_{\theta_0,G_{\eta_n}}} 1. \tag{5.12}$$

Together with (5.8) with $I_0 = I(G_0)$, this shows that the left hand side of (5.11) equals

$$\lim_{n\to\infty} P_{\theta_0,G_{\eta_n}}\left(\sqrt{n\hat{I}_n}(\hat{\theta}_n - \theta_0) \ge \Phi^{-1}(\alpha) + \sqrt{\frac{\hat{I}_n}{I(G_{\eta_n})}}v\right) = \Phi(\Phi^{-1}(1-\alpha) - v) \quad v > 0. \tag{5.13}$$

Taking $\lim_{c\to\infty}$ we obtain (5.10). □

Similarly to semiparametric estimation theory, this result justifies the terminology to call $(I_n)_{n\in\mathbb{N}}$ defined in (5.9) semiparametrically locally asymptotically efficient or efficient for short. We might also call the sequence $(I_n)_{n\in\mathbb{N}}$ locally asymptotically most accurate; cf. Sections 3.5 and 5.5 of Lehmann (1959) and the title of the present paper. In the last section we will show that this sequence of confidence intervals is globally asymptotically efficient—or most accurate for that matter—within appropriately restricted submodels. Finally, we note that also the semiparametric two-sided confidence interval from (4.16) and (4.17) is locally asymptotically optimal.

6. Global uniformity under strong unimodality

As explained in Section 5 the discontinuity of the Fisher information causes non-uniformity. Therefore, uniformity can hold only on subsets of $\mathcal{P}$ on which the Fisher information for location is continuous. In Bickel and Klaassen (1986) an adaptive, semiparametrically efficient estimator $\hat{\theta}_n$ of θ has been constructed with the following strengthening of (4.10) and (5.8) with $I_0 = I(G_0)$. For every $\theta_0 \in \mathbb{R}$, and every sequence $(\theta_n)_{n\in\mathbb{N}}$ with $\left(\sqrt{n}(\theta_n - \theta_0)\right)_{n\in\mathbb{N}}$ bounded, for every $G_0 \in \mathcal{G}$ and every sequence $(G_n)_{n\in\mathbb{N}}$ in $\mathcal{G}$ satisfying

$$G_n \xrightarrow{w} G_0, \quad I(G_n) \to I(G_0), \tag{6.1}$$

the convergence

$$\sqrt{nI(G_n)}(\hat{\theta}_n - \theta_n) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad \text{as } n \to \infty, \tag{6.2}$$

holds under P_{θ_n,G_n}. Moreover, their estimator $\hat{I}_n$ of the Fisher information is consistent along such sequences, i.e.,

$$\hat{I}_n \xrightarrow{P_{\theta_n,G_n}} I(G_0). \tag{6.3}$$

For the proof of (6.2) we refer to Theorem 1.1 of Bickel and Klaassen (1986). By (2.5), (2.9), and (2.12) of ibid. with $c_n = \sigma_n$, $n\sigma_n^6 \to \infty$, and $\nu = \nu_n^* = \mathcal{O}_p(1/\sqrt{n})$ in their notation, we see that (6.3) holds.

The one-sided confidence interval from (5.9) based on these estimators is globally asymptotically most accurate or globally asymptotically efficient in the following sense.

Theorem 6.1. *Let $\mathcal{G}_0 \subset \mathcal{G}$ be such that $I\colon \mathcal{G}_0 \to (0,\infty)$, $G \mapsto I(G)$, is continuous under the topology of weak convergence. The confidence interval I_n is as defined in* (5.9) *with $\hat{\theta}_n$ and $\hat{I}_n$ satisfying* (6.1) *through* (6.3)*. If $\mathcal{G}_0$ is compact, then*

$$\limsup_{n\to\infty} \sup_{G\in\mathcal{G}_0} P_{\theta_0,G}\left(\theta_0 + \frac{v}{\sqrt{nI(G)}} \in I_n\right) = \Phi(\Phi^{-1}(1-\alpha) - v), \quad v > 0, \tag{6.4}$$

holds.

Proof. Let (G_n), $G_n \in \mathcal{G}_0$, be such that the left hand side of (6.4) equals

$$\lim_{n\to\infty} P_{\theta_0,G_n}\left(\theta_0 + \frac{v}{\sqrt{nI(G_n)}} \in I_n\right). \tag{6.5}$$

Because of the compactness we may assume without loss of generality that there exists a $G_0 \in \mathcal{G}_0$ such that the sequence of distribution functions G_n converges weakly to the distribution function G_0. By (6.2) and (6.3) the theorem is proved. □

A distribution function G on $\mathbb{R}$ has a strongly unimodal density g if its convolution with any unimodal distribution is unimodal again. Ibragimov (1956) has shown that g is strongly unimodal iff it is log-concave. Consequently, a strongly unimodal density g is absolutely continuous with derivative g' such that the score function for location $-g'/g$ is nondecreasing. As mentioned already in the Introduction, Van Eeden (1970) has studied her adaptive estimator of location for the class of symmetric densities with monotone score function, in other words for the class of symmetric, strongly unimodal densities. This motivates us to study global uniformity of our confidence intervals over this class of densities.

We have the following continuity result on substantial subclasses of the class of all strongly unimodal densities.

Theorem 6.2. *Denote the variance of a random variable with distribution function G by σ_G^2, the density of G by g and its score function for location by $-g'/g$. Let $\mathcal{G}_{\epsilon,\kappa}$ be the set of distribution functions with symmetric, strongly unimodal densities g on $\mathbb{R}$ satisfying*

$$\sigma_G^2\left(\frac{g'}{g}\left(G^{-1}(u)\right)\right)^2 \leq \kappa\big(u(1-u)\big)^{\epsilon-1}, \quad 0 < u < 1. \tag{6.6}$$

For every $\epsilon \in (0, \frac{1}{2}]$ and $\kappa \geq 1$ the set $\mathcal{G}_{\epsilon,\kappa}$ is nonempty and it contains the normal distribution functions with mean zero. Furthermore, the Fisher information $G \mapsto I(G)$ is continuous as a map from $\mathcal{G}_{\epsilon,\kappa}$ with the topology of weak convergence, to $(0, \infty)$.

Proof. In view of

$$\begin{aligned}(6.7)\quad \sup_x x^2\Big(\Phi(x)\big(1-\Phi(x)\big)\Big)^{1-\epsilon} & \\ &\leq \sup_{x>0} x^{1+\epsilon}\varphi^{1-\epsilon}(x) \\ &= (2\pi)^{(\epsilon-1)/2}\exp\Big[\frac{1}{2}(1+\epsilon)\Big(\log\Big(\frac{1+\epsilon}{1-\epsilon}\Big)-1\Big)\Big] \\ &\leq 3^{3/4}(2\pi e^3)^{-1/4} < 1\end{aligned}$$

the first statement has been proved. For measurable functions f, f_n, $n \in \mathbb{N}$, we will say that (f_n) converges weakly to f if $(f_n(x))$ converges to $f(x)$ for all continuity points x of f. We will denote this convergence by $f_n \xrightarrow{w} f$.

Fix ϵ and κ. Let (G_n) be a sequence in $\mathcal{G}_{\epsilon,\kappa}$ with (g_n) the sequence of corresponding densities and $(\psi_n) = (-g_n'/g_n(G_n^{-1}))$ the sequence of corresponding score functions. Let (G_n) converge weakly to $G \in \mathcal{G}_{\epsilon,\kappa}$ with density g and score function $\psi = -g'/g(G^{-1})$. By, e.g., (6.1) of Klaassen (1989b) this implies $G_n^{-1} \xrightarrow{w} G^{-1}$. The symmetry and unimodality of g imply

$$G^{-1}(u) = \int_{1/2}^{u} \frac{1}{g\big(G^{-1}(s)\big)}\, ds, \quad 0 < u < 1, \tag{6.8}$$

with the integrand monotone both for $u < \frac{1}{2}$ and $u > \frac{1}{2}$. The same holds mutatis mutandis for $G_n^{-1}(\cdot)$. Consequently, Lemma 6.1 below, applied to $(0, \frac{1}{2})$ and $(\frac{1}{2}, 1)$ respectively, yields

$$g_n\big(G_n^{-1}(\cdot)\big) \xrightarrow{w} g\big(G^{-1}(\cdot)\big). \tag{6.9}$$

The strong unimodality of g implies that

$$g(G^{-1}(s)) = \int_s^1 \psi(t)dt \tag{6.10}$$

has a nondecreasing integrand $\psi(\cdot) = -g'/g\big(G^{-1}(\cdot)\big)$. The same holds for $g_n(G_n^{-1}(\cdot))$. Consequently, Lemma 6.1 can be applied again to yield

$$\psi_n \xrightarrow{w} \psi. \tag{6.11}$$

By (6.6), $\limsup_{n\to\infty} \sigma_{G_n}^2 = \limsup_{n\to\infty} E_{G_n}X^2 \geq E_G X^2$, and dominated convergence we obtain

$$I(G_n) = \int \psi_n^2 to \int \psi^2 = I(G) \tag{6.12}$$

and hence the theorem. □

In this proof we have used the following lemma repeatedly.

Lemma 6.1. *Let (a,b) be an interval and let f, f_n, $n \in \mathbb{N}$, be measurable nondecreasing functions on (a,b). If for all $s, t \in (a,b)$ the convergence*

$$\int_s^t f_n(u)\, du \to \int_s^t f(u)\, du, \quad \text{as } n \to \infty, \tag{6.13}$$

holds, then

$$f_n(s) \to f(s), \quad \text{as } n \to \infty, \tag{6.14}$$

for all $s \in (a,b)$ at which f is continuous.

Proof. Let f be continuous at s and assume that there exists a positive ϵ with

$$\liminf_{n\to\infty} f_n(s) \leq f(s) - \epsilon. \tag{6.15}$$

By the continuity there exists a positive δ with

$$f(s-\delta) \geq f(s) - \tfrac{1}{2}\epsilon. \tag{6.16}$$

Choose the subsequence (n') of (n) in such a way that

$$\lim_{n'\to\infty} f_{n'}(s) = \liminf_{n\to\infty} f_n(s) \tag{6.17}$$

is valid. From these relations, the monotonicity, (6.13) and Fatou's Lemma we get the contradiction

$$\begin{aligned} \delta(f(s)-\epsilon) &= \int_{s-\delta}^{s} (f(s)-\epsilon)\,du \geq \int_{s-\delta}^{s} \limsup_{n'\to\infty} f_{n'}(u)\,du \\ &\geq \limsup_{n'\to\infty} \int_{s-\delta}^{s} f_{n'}(u)\,du = \int_{s-\delta}^{s} f(u)\,du \\ &\geq \delta(f(s) - \tfrac{1}{2}\epsilon). \end{aligned} \tag{6.18}$$

□

Applying Theorem 6.2 to any compact subset of $\mathcal{G}_{\epsilon,\kappa}$ for any $\epsilon \in (0, \frac{1}{2}]$ and any $\kappa \geq 1$ we obtain a global uniformity result for strongly unimodal densities. Needless to say that this global uniformity carries over to adaptive estimation.

Acknowledgements. I would like to thank editor Marc Moore, two referees, and Ge Hong for pointing out some typos and for suggesting some clarifications in the text.

REFERENCES

Beran, R. (1974). Asymptotically efficient adaptive rank estimates in location models. *Ann. Statist.* 2, 63–74.

Beran, R. (1978). An efficient and robust adaptive estimator of location. *Ann. Statist.* 6, 292–313.

Bickel, P.J. and Klaassen, C.A.J. (1986). Empirical Bayes estimation in functional and structural models, and uniformly adaptive estimation of location. *Adv. in Appl. Math.* 7, 55–69.

Bickel, P.J., Klaassen, C.A.J., Ritov, Y. and Wellner, J.A. (1993). *Efficient and Adaptive Estimation in Semiparametric Models.* Johns Hopkins Univ. Press, Baltimore; (1998). Springer, New York. Reprint.

Bickel, P.J., Lehmann, E.L. (1975). Descriptive statistics for nonparametric models II. Location. *Ann. Statist.* 3, 1045–1069.

Chang, Y.C. (1992). Adaptive sequential confidence intervals for a location parameter. *Sequential Anal.* 11, 257–272.

Choi, S., Hall, W.J. and Schick, A. (1996). Asymptotically uniformly most powerful tests in parametric and semiparametric models. *Ann. Statist.* 24, 841–861.

Hájek, J. (1962). Asymptotically most powerful rank-order tests. *Ann. Math. Statist.* 33, 1124–1147.

Huber, P. (1981). *Robust Statistics.* Wiley, New York.

Ibragimov, I.A. (1956). On the composition of unimodal distributions. *Theory Probab. Appl.* 1, 255–260.

Klaassen, C.A.J. (1979). Nonuniformity of the convergence of location estimators. *Proceedings of the Second Prague Symposium on Asymptotic Statistics* (Mandl, P. and Hušková eds.), pp. 251–258. North Holland, Amsterdam.

Klaassen, C.A.J. (1980). *Statistical Performance of Location Estimators.* Proefschrift, Leiden; (1981). Mathematical Centre Tracts, vol. 133. Mathematisch Centrum, Amsterdam.

Klaassen, C.A.J. (1987). Consistent estimation of the influence function of locally asymptotically linear estimators. *Ann. Statist.* 15, 1548–1562.

Klaassen, C.A.J. (1989a). Estimators and spread. *Ann. Statist.* 17, 859–867.

Klaassen, C.A.J. (1989b). The asymptotic spread of estimators. *J. Statist. Plann. Inference* 23, 267–285.

Lehmann, E.L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

Martinsek, A.T. (1991). Asymptotically fully efficient fixed-width confidence intervals for a location parameter. *Amer. J. Math. Management Sci.* 11, 95–112.

Stein, C. (1956). Efficient nonparametric testing and estimation. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* Vol. 1, pp. 187–195. Univ. California Press, Berkeley.

Stone, C.J. (1975). Adaptive maximum liklehood estimators of a location parameter. *Ann. Statist.* 3, 267–284.

van Eeden, C. (1970). Efficiency-robust estimation of location. *Ann. Math. Statist.* 41, 172–181.

Chris A.J. Klaassen
Korteweg-de Vries Institute for Mathematics
University of Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
The Netherlands
chrisk@science.uva.nl