KYBERNETIKA — VOLUME 22 (1986), NUMBER 1

EFFICIENCY AND ROBUSTNESS CONTROL VIA DISTORTED MAXIMUM LIKELIHOOD ESTIMATION

IGOR VAJDA

Families of M. L. E.'s with likelihood functions distorted by a parameter $\alpha \geq 0$ are introduced so that $\alpha = 0$ yields the classical non-distorted M. L. E.. The M. L. E. is known to be efficient but not robust while the distorted estimators are shown to be robust but not efficient. For quite general types of distortions and statistical families, the distorted estimates as well as the corresponding influence curves and asymptotic variances are shown to be continuous at $\alpha = 0$. Thus the parameter α controls the efficiency and robustness of estimators under consideration so that one can easily review the set of attainable compromises and select the most appropriate one. General location and scale families are analyzed from this point of view in more detail.

1. INTRODUCTION

$\mathbb{N}$ is the set of all natural and $\mathbb{R}$ the set of all real numbers. $(\mathcal{X}, \mathcal{A})$ is a measurable sample space, $\mathcal{P}$ the class of all probability distributions on $(\mathcal{X}, \mathcal{A})$, $\delta_x \in \mathcal{P}$ the distribution with all probability concentrated at $x \in \mathcal{X}$, and $\mathcal{P}_e$ the subclass of all empirical distributions defined by the mapping

$$x \mapsto P_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i} \quad \text{for all} \quad x = (x_1, \ldots, x_n) \in \mathcal{X}^n , \quad n \in \mathbb{N} .$$

Θ is a locally compact Hausdorff topological space with countable base and Borel sigma-algebra $\mathcal{B}$, $\mathcal{P}_\Theta = \{P_\theta: \theta \in \Theta\} \ll \lambda$ (sigma-finite), $p_\theta = \mathrm{d}P_\theta/\mathrm{d}\lambda$ on $\Theta \times \mathcal{X}$, and $D(\theta, P)$, with $D(\theta, x)$ written instead of $D(\theta, P_n)$ for $x \mapsto P_n$, is a mapping $\Theta \times \mathcal{P}_0 \mapsto [-\infty, \infty]$, where $\mathcal{P}_\Theta \cup \mathcal{P}_e \subset \mathcal{P}_0 \subset \mathcal{P}$. A mapping $T: \mathcal{P}_0 \mapsto \Theta$ is said a D-estimator if[1])

$$(1.1) \qquad T(P) \in \operatorname{argmin}_\Theta D(\theta, P) \subset \Theta \quad \text{for every} \quad P \in \mathcal{P}_0$$

and if $T_n(x) = T(P_n)$ for $x \mapsto P_n$ is $(\mathcal{A}^n, \mathcal{B})$-measurable for every $n \in \mathbb{N}$.

[1]) The symbol $\operatorname{argmin}_\Theta D(\theta, P)$ denotes the non-empty subset of Θ at which the function $D(\theta, P)$ of variable $\theta \in \Theta$ attains its minimum value, provided the minimum exists. Otherwise this symbol denotes the empty subset. Analogically for argmax.

Various functions D have been proposed in [13] but the concept of D-estimator was introduced in a too narrow sense there. Namely, (1.1) was replaced by $T(P) = \tau(\operatorname{argmin}_\Theta D(\theta, P))$ for every $P \in \mathscr{P}_0$, where $\tau: \exp \Theta \mapsto \Theta$ was a fixed rule of choice. In the meantime we found that quite complicated topologies have to be considered on $\exp \Theta$ if the continuity of τ required in [13] has to be ensured on sufficiently general parameter spaces Θ and that the whole theory is considerably simpler if the concept of D-estimator is based on (1.1).

The next result which follows from Theorem 1.1 of [14] is stated here for references later. There and in the sequel, $\Theta \cup \{\theta^*\}$ denotes the one-point compactification of Θ (cf. Chap. 5 of Kelley [7]).

Lemma 1.1. Let $D(\theta, P)$ be continuous on Θ and continuously extendable to θ^* with

$$D(\theta, P) < D(\theta^*, P) \quad \text{for some} \quad \theta = \theta(P) \in \Theta \quad \text{and every} \quad P \in \mathscr{P}_0 . \tag{1.2}$$

Then the D-estimator T exists.

In the present paper we are interested in D-estimators which can be interpreted as distorted maximum likelihood estimators (M. L. E.'s) with a distortion parameter $\alpha \geqq 0$. These estimators are denoted by T^α and defined as follows. Let f_α, $\alpha \geqq 0$, be a class of increasing, continuous, extended real-valued functions defined on $[0, \infty]$ and twice continuously differentiable on $(0, \infty)$ with $(f_\alpha(u), f'_\alpha(u), f''_\alpha(u)) \to (f_0(u), f'_0(u), f''_0(u))$ as $\alpha \to 0$ for all $u \in (0, \infty)$ where $f_0(u) = \ln u$. Then T^α is the D-estimator defined by $D(\theta, P) = -l_\alpha(\theta, P)$ where

$$l_\alpha(\theta, P) = \mathsf{E}_P f_\alpha(p_\theta) \quad \text{for every} \quad \theta \in \Theta\,, \quad P \in \mathscr{P}_0\,, \tag{1.3}$$

and where the expectations are supposed to be well-defined. Obviously, T^0 is the well-known M. L. E. Note that, if restricted to the subdomains $\mathscr{P}_e \subset \mathscr{P}_0$, all distorted M. L. E.'s T^α of the present paper become minimum contrast estimators of Pfanzagl [10].

Theorem 1.1. Let $p_\theta(x)$ be continuous on Θ with

$$\mathsf{E}_P p_\theta > 0 \quad \text{for some} \quad \theta = \theta(P) \in \Theta \quad \text{and every} \quad P \in \mathscr{P}_0 \tag{1.4}$$

and with

$$\lim_{\theta \to \theta^*} p_\theta(x) = 0 \quad \text{for every} \quad x \in \mathscr{X} . \tag{1.5}$$

Further, let for every $\tilde{\theta} \in \Theta \cup \{\theta^*\}$ there exists an open neighborhood $B(\tilde{\theta}) \subset \Theta$ such that

$$-\infty \leqq \mathsf{E}_P f_\alpha({}_{B(\tilde{\theta})}p) \leqq \mathsf{E}_P f_\alpha(p_{B(\tilde{\theta})}) < \infty \quad \text{for every} \quad P \in \mathscr{P}_0\,, \tag{1.6}$$

where

$${}_{B(\tilde{\theta})}p(x) = \inf_{B(\tilde{\theta})} p_\theta(x)\,, \quad p_{B(\tilde{\theta})}(x) = \sup_{B(\tilde{\theta})} p_\theta(x)\,, \tag{1.7}$$

and where the left-hand equality in (1.6) takes place iff $\tilde{\theta} = \theta^*$ and $f_\alpha(0) = -\infty$. Then T^α exists.

Proof. By (1.6) and the Lebesgue bounded convergence theorem, the continuity of p_θ on Θ implies that of $l_\alpha(\theta, P)$ for every $P \in \mathscr{P}_0$. If $f_\alpha(0) > -\infty$ then the same argument together with (1.5) implies that $l_\alpha(\theta, P)$ continuously extends to θ^* and $l_\alpha(\theta^*, P) = f_\alpha(0)$. If $f_\alpha(0) = -\infty$ then the assumption $\limsup_{j\to\infty} l_\alpha(\theta_j, P) > -\infty$ for some $\theta_j \to \theta^*$ together with the Fatou lemma and (1.5) leads to the contradiction $f_\alpha(0) > -\infty$. Therefore the continuous extendability of $l_\alpha(\theta, P)$ to θ^* holds for $f_\alpha(0) = -\infty$ as well. Finally, the monotony of f_α together with (1.4) implies that for every $P \in \mathscr{P}_0$ there exists $\theta \in \Theta$ such that $l_\alpha(\theta, P) > f_\alpha(0)$. Therefore all assumptions of Lemma 1.1 hold for $D(\theta, P) = -l_\alpha(\theta, P)$ and the desired result follows from Lemma 1.1. □

Next we present conditions under which the estimates $T^\alpha(P)$ tend to the M.L. E.'s $T^0(P)$ for $P \in \mathscr{P}_0$.

Theorem 1.2. Let the assumptions of Theorem 1.1 hold for all $\alpha \geqq 0$ and let for some $P \in \mathscr{P}_0$ there exists a compact $B(P) \subset \Theta$ such that

$$\max_{\Theta} l_\alpha(\theta, P) = \max_{B(P)} l_\alpha(\theta, P) \quad \text{for all} \quad \alpha \geqq 0\,, \tag{1.8}$$

$$\lim_{\alpha\to 0+} \sup_{B(P)} |l_\alpha(\theta, P) - l_0(\theta, P)| = 0\,, \tag{1.9}$$

and

$$\{T^0(P)\} = \operatorname{argmax}_{B(P)} l_0(\theta, P)\,. \tag{1.10}$$

Then

$$\lim_{\alpha\to 0+} T^\alpha(P) = T^0(P)\,. \tag{1.11}$$

Proof. Let $\{\alpha_j: j \in \mathbb{N}\}$ be arbitrary fixed sequence tending to zero and $\theta_j = T^{\alpha_j}(P) \in B(P)$. By Theorem 5 of Chapter 5 of [7] there exists in $B(P)$ at least one limit point θ_0 of $\{\theta_j: j \in \mathbb{N}\}$. If we prove $\theta_0 = T^0(P)$ then (1.11) will be proved. Suppose for simplicity $\theta_j \to \theta_0$ as $j \to \infty$ (there exists exactly one limit point). By the definition of $T^{\alpha_j}(P)$, $l_{\alpha_j}(\theta_j, P) \geqq l_{\alpha_j}(\theta, P)$ and, by (1.9), $l_{\alpha_j}(\theta, P) \to l_0(\theta, P)$ for all $\theta \in B(P)$. Therefore

$$\liminf_{j\to\infty} l_{\alpha_j}(\theta_j, P) \geqq l_0(\theta, P) \quad \text{for all} \quad \theta \in B(P)\,. \tag{1.12}$$

Take now into account the inequality

$$|l_{\alpha_j}(\theta_j, P) - l_0(\theta_0, P)| \leqq |l_{\alpha_j}(\theta_j, P) - l_0(\theta_j, P)| + |l_0(\theta_j, P) - l_0(\theta_0, P)|\,.$$

By the proof of Theorem 1.1, $l_0(\theta, P)$ is continuous so that the second right-hand term tends to zero as $j \to \infty$. The first right-hand term tends to zero by (1.9). Therefore $l_{\alpha_j}(\theta_j, P) \to l_0(\theta_0, P)$. This and (1.12) imply $l_0(\theta_0, P) \geqq l_0(\theta, P)$ for all $\theta \in B(P)$ which implies $\theta_0 \in \operatorname{argmax}_{B(P)} l_0(\theta, P)$. The rest follows from (1.10). □

T^α is said strongly consistent for $P \in \mathscr{P}_0$ if $T_n^\alpha(\boldsymbol{x}) \to T^\alpha(P)$ as $n \to \infty$ P^∞ – a.s. The next result extends the theorem of Le Cam [8] on strong consistency of M. L. E.'s.

Theorem 2.1. If the conditions of Theorem 1.1 hold and $\{T^\alpha(P)\} = \operatorname{argmax}_\Theta l_\alpha(\theta, P)$ for some $P \in \mathscr{P}_0$ then T^α is strongly consistent for P.

Proof. (I) Let $p_B(x) = \sup_B p_\theta(x)$ for every $B \subset \Theta$ (cf. (1.7)) and let for every $B \subset \Theta$

$$z(x, B) = \inf_B \left[f_\alpha(p_{T^\alpha(P)}(x)) - f_\alpha(p_\theta(x))\right] =$$
$$= f_\alpha(p_{T^\alpha(P)}(x)) - \sup_B f_\alpha(p_\theta(x)) = f_\alpha(p_{T^\alpha(P)}(x)) - f_\alpha(p_B(x)) .$$

Since Θ has a countable base, for every $\theta \in \Theta \cup \theta^*$ there exist open neighborhoods $\Theta \supset B_1(\theta) \supset B_2(\theta) \supset \ldots$ with the intersection $\{\theta\}$. The monotony of the neighborhoods implies

$$z(x, B_j(\theta)) \leqq z(x, B_{j+1}(\theta)) \quad \text{for every} \quad j \in N$$

and the continuity of $p_\theta(x)$ implies

$$\lim_{j\to\infty} z(x, B_j(\theta)) = f_\alpha(p_{T^\alpha(P)}(x)) - f_\alpha(p_\theta(x)) = z(x, \theta)$$

provided $p_{\theta^*}(x) = 0$ on $\mathscr{X}$. Moreover the assumption $\{T^\alpha(P)\} = \operatorname{argmax}_\Theta l_\alpha(\theta, P)$ implies $\mathsf{E}_P\, z(x, \theta) = l_\alpha(T^\alpha(P), P) - l_\alpha(\theta, P) > 0$ for all θ under consideration different from $T^\alpha(P)$. Therefore, by the Lebesgue monotone convergence theorem,

$$\lim_{j\to\infty} \mathsf{E}_P\, z(x, B_j(\theta)) = \mathsf{E}_P\, z(x, \theta) > 0 \quad \text{for every} \quad \theta \in \Theta \cup \{\theta^*\}, \quad \theta \neq T^\alpha(P). \tag{2.1}$$

Consequently, for every θ considered in (2.1), there exists an open neighborhood $B(\theta) \subset \Theta$ such that $\mathsf{E}_P\, z(x, B(\theta)) > 0$.

(II) Let $B \subset \Theta$ be an arbitrary open neighborhood of $T^\alpha(P)$. Since $\Theta \cup \{\theta^*\}$ is compact, the Heine-Borel theorem implies that there exists a finite set $H = \{\theta^*, \theta_1, \ldots, \theta_k\} \subset \Theta \cup \{\theta^*\}$ disjoint with B and open neighborhoods $B(\theta) \subset \Theta$ of points $\theta \in H$ such that

$$\bigcup_{\theta \in H} B(\theta) \supset \Theta - B \quad \text{and} \quad \mathsf{E}_P\, z(x, B(\theta)) > 0 \quad \text{for all} \quad \theta \in H . \tag{2.2}$$

(cf. part (I) of the proof). Obviously, $T_n^\alpha(\boldsymbol{x}) \in B$ if $\boldsymbol{x} \in A_n$ where

$$A_n = \{\boldsymbol{x} \in \mathscr{X}^n : \sum_{i=1}^n f_\alpha(p_{T(P)}(x_i)) - \sup_{\Theta - B} \sum_{i=1}^n f_\alpha(p_\theta(x_i)) > 0\} .$$

(III) To prove the desired statement take first into account that the inclusion in (2.2) implies

$$\max_{\theta \in H} \sup_{B(\theta)} \sum_{i=1}^n f_\alpha(p_\theta(x_i)) \geqq \sup_{\Theta - B} \sum_{i=1}^n f_\alpha(p_\theta(x_i)) ,$$

for every $x \in \mathscr{X}^n$. Therefore

$$A_n \supset \tilde{A}_n = \{x \in \mathscr{X}^n : \max_{\theta \in H} (\sum_{i=1}^n f_\alpha(p_{T(P)}(x_i)) - \sup_{B(\theta)} \sum_{i=1}^n f_\alpha(p_\theta(x_i))) > 0\} .$$

Since for every $x \in \mathscr{X}^n$

$$\sum_{i=1}^n \sup_{B(\theta)} f_\alpha(p_\theta(x_i)) \geqq \sup_{B(\theta)} \sum_{i=1}^n f_\alpha(p_\theta(x_i)) ,$$

it holds

$$\tilde{A}_n \supset \bar{A}_n = \{x \in \mathscr{X}^n : \max_{\theta \in H} \sum_{i=1}^n z(x_i, B(\theta)) > 0\} .$$

Further, the right-hand inequality in (2.2) and the strong law of large numbers imply that there exists a measurable set $A \subset \mathscr{X}^\infty$ with $P^\infty(A) = 1$ such that for every infinite sequence $x^\infty = (x_1, x_2, \ldots) \in A$ there exists n_{x^∞} such that $(x_1, \ldots, x_n) \in \bar{A}_n$ for all $n > n_{x^\infty}$. This however implies that $T_n^\alpha(x)$ tends P^∞ – a.s. to $T^\alpha(P)$ for the above established inclusions $\bar{A}_n \subset \tilde{A}_n \subset A_n$ together with part (II) of the proof imply that the event $x \in \bar{A}_n$ results in that $T_n^\alpha(x)$ belongs to the neighborhood B of $T^\alpha(P)$. □

In the rest of this section we consider the following

Regularity assumptions. Let $\Theta \subset \mathbb{R}^m$ for $m \in \mathbb{N}$, let us consider the Euclidean $\mathbb{R}^m$-topology on Θ, let $\operatorname{int} \Theta \neq \emptyset$, let for every $\alpha \geqq 0$ and $P \in \mathscr{P}_0 \subset \mathscr{P}$ there exist derivatives

$$l'_\alpha(\theta, P) = \frac{d}{d\theta} l_\alpha(\theta, P) = \mathsf{E}_P[f'_\alpha(p_\theta) p'_\theta] \tag{2.4}$$

$$l''_\alpha(\theta, P) = \left(\frac{d}{d\theta}\right)^T l_\alpha(\theta, P) = \mathsf{E}_P[f''_\alpha(p_\theta) p'_\theta p'^T_\theta + f'_\alpha(p_\theta) p''_\theta] \tag{2.5}$$

with components continuous on $\operatorname{int} \Theta$, where $(d/d\theta)^T = (\partial/\partial\theta_1, \ldots, \partial/\partial\theta_m)$ (throughout the paper the superscript ^{T} denotes the matrix transposition) and

$$p'_\theta = \frac{d}{d\theta} p_\theta , \quad p''_\theta = \left(\frac{d}{d\theta}\right)^T p'_\theta , \quad \mathsf{E}_\lambda p'_\theta = \mathsf{E}_\lambda p''_\theta = 0 , \tag{2.6}$$

and let $\mathscr{P}_0$ be a convex subclass of $\mathscr{P}$.

We shall say that $\Omega_P: \mathscr{X} \mapsto \Theta$ is an influence curve of an estimator $T: \mathscr{P}_0 \mapsto \Theta$ at $P \in \mathscr{P}_0$ if for every $x \in \mathscr{X}$

$$\Omega_P(x) = \lim_{\varepsilon \to 0+} \frac{T(P_{\varepsilon,x}) - T(P)}{\varepsilon} , \quad \text{where } P_{\varepsilon,x} = (1 - \varepsilon) P + \varepsilon \delta_x \text{ for } \varepsilon \in (0, 1). \tag{2.7}$$

Theorem 2.2. Let the assumptions of Theorem 2.1 hold for some $P \in \mathscr{P}_0$ and let $T^\alpha(P) \in \operatorname{int} \Theta$, $\det [l''_\alpha(T^\alpha(P), P)] \neq 0$, and

$$\lim_{\varepsilon \to 0+} T^\alpha(P_{\varepsilon,x}) = T^\alpha(P) \quad \text{for all} \quad x \in \mathscr{X} . \tag{2.8}$$

Then the influence curve of T^α at P exists and is given by

$$(2.9) \qquad \Omega_P^\alpha(x) = -l''_\alpha(T^\alpha(P), P)^{-1}\,\psi_P(x)\,, \quad \psi_P(x) = f'_\alpha(p_{T^\alpha(P)}(x))\,p'_{T^\alpha(P)}(x)\,.$$

Proof. Let $x \in \mathscr{X}$ be arbitrary fixed. Since $\delta_x \in \mathscr{P}_e \subset \mathscr{P}_0$ and since $\mathscr{P}_0$ is convex, $P_{\varepsilon,x} \in \mathscr{P}_0$ and $T^\alpha(P_{\varepsilon,x})$ exists for all $\varepsilon \in [0, 1]$. Further, since for all sufficiently small $\varepsilon > 0$ it holds

$$\begin{aligned} &l'_\alpha(T^\alpha(P), P) = 0\,, \\ (2.10) \qquad &l'_\alpha(T^\alpha(P_{\varepsilon,x}), P_{\varepsilon,x}) = 0\,, \\ &l'_\alpha(T^\alpha(P), P_{\varepsilon,x}) = (1-\varepsilon)\,l'_\alpha(T^\alpha(P), P) + \varepsilon l'_\alpha(T^\alpha(P), 1_{\{x\}}) = \\ &\qquad = \varepsilon\psi_P(x) \quad (\text{cf. } (2.9), (2.4) \text{ and } (2.10))\,, \end{aligned}$$

it also holds

$$l'_\alpha(T^\alpha(P), P_{\varepsilon,x}) - l'_\alpha(T^\alpha(P_{\varepsilon,x}), P_{\varepsilon,x}) = \varepsilon\psi_P(x)\,.$$

By the mean value theorem applied coordinatewise to the function $L(u) = l'_\alpha(uT^\alpha(P) + (1-u)\,T^\alpha(P_{\varepsilon,x}), P_{\varepsilon,x})$, $u \in [0, 1]$, each coordinate of the left-hand difference is equal to the corresponding coordinate of

$$l''_\alpha(\theta_\varepsilon, P_{\varepsilon,x})\,(T^\alpha(P) - T^\alpha(P_{\varepsilon,x}))$$

where θ_ε, possibly depending on the coordinate, tends to $T(P)$ as $\varepsilon \to 0$ (cf. the assumption (2.8)). Therefore

$$(2.11) \quad l''_\alpha(\theta_\varepsilon, P_{\varepsilon,x})\,\frac{T^\alpha(P_{\varepsilon,x}) - T^\alpha(P)}{\varepsilon} = -\psi_P(x) \quad \text{for all sufficiently small } \varepsilon > 0\,.$$

On the other hand, by (2.5),

$$l''_\alpha(\theta_\varepsilon, P_{\varepsilon,x}) = (1-\varepsilon)\,l''_\alpha(\theta_\varepsilon, P) + \varepsilon l''_\alpha(\theta_\varepsilon, \delta_x)\,, \quad \text{where} \quad P, \delta_x \in \mathscr{P}_0\,.$$

Since we assume that $l''_\alpha(\theta, P)$ is continuous on $\operatorname{int}\Theta$ for every $\tilde{P} \in \mathscr{P}_0$ and since $\theta_\varepsilon \in \operatorname{int}\Theta$ for all sufficiently small ε, the last identity implies

$$\lim_{\varepsilon\to 0+} l''_\alpha(\theta_\varepsilon, P_{\varepsilon,x}) = l''_\alpha(T(P), P)\,.$$

This together with (2.11) implies that the limit $\Omega_P^\alpha(x)$ defined by (2.7) with T replaced by T^α exists and satisfies (2.9). □

Theorem 2.3. Let the assumptions of Theorem 2.2 hold, let all components of the $m \times m$ matrix $\mathsf{E}_P(\Omega_P^\alpha \Omega_P^{\alpha\mathrm{T}})$ be finite and let for every $\tilde{\theta} \in \operatorname{int}\Theta$ there exists an open neighborhood $B(\tilde{\theta}) \subset \Theta$ such that

$$(2.14) \qquad \mathsf{E}_P \sup_{B(\tilde{\theta})} \varphi_\theta^2 < \infty$$

for all components φ_θ of the matrix $f''_\alpha(p_\theta)\,p'_\theta p'^{\mathrm{T}}_\theta + f'_\alpha(p_\theta)\,p''_\theta$. Then $\sqrt{(n)}\,(T_n^\alpha(x) - T^\alpha(P))$ converges P^∞-weakly to $N(0, \mathsf{E}_P(\Omega_P^\alpha \Omega_P^{\alpha\mathrm{T}}))$ as $n \to \infty$.

Proof. (I) We first prove that if $\theta_n \to \theta_0 \in \operatorname{int} \Theta$ as $n \to \infty$ and $x \to P_n$, then

$$\lim_{n\to\infty} l''_\alpha(\theta_n, P_n) = l''_\alpha(\theta_0, P) \quad \text{in} \quad P^\infty\text{-probability}. \tag{2.15}$$

The assumed continuity of $\varphi_\theta(x)$ on Θ for every $x \in \mathscr{X}$ (cf. (2.14), (2.5)) together with (2.14) and with the Lebesgue bounded convergence theorem imply $\mathsf{E}_P\varphi_{\theta_n} \to \to \mathsf{E}_P\varphi_{\theta_0}$ as $n \to \infty$. The same argument implies that there exists $c > 0$ such that

$$\mathsf{E}_P(\varphi_{\theta_n} - \mathsf{E}_P\varphi_{\theta_n})^2 < c \quad \text{for all} \quad n \in \mathbb{N}.$$

This assumption and the Chebyshev inequality imply (cf. the proof of Theorem 2 in Sec. VIII. 3 of Rényi [12])

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (\varphi_{\theta_n}(x_i) - \mathsf{E}_P\varphi_{\theta_n}) = 0 \quad \text{in} \quad P^\infty\text{-probability}.$$

Since

$$\frac{1}{n} \sum_{i=1}^{n} \varphi_{\theta_n}(x_i) \quad \text{and} \quad \mathsf{E}_P\varphi_{\theta_0} = \lim_{n\to\infty} \mathsf{E}_P\varphi_{\theta_n}$$

are nothing but the components of the matrices $l''_\alpha(\theta_n, P_n)$ and $l''_\alpha(\theta_0, P_n)$ respectively (cf. (2.14), (2.5)), the statement (2.15) holds.

(II) By Theorem 2.1, $T_n^\alpha = T_n^\alpha(x) \to T^\alpha(P)$ as $n \to \infty$ P^∞-a.s. Hence for $x \mapsto P_n$ and all sufficiently large $n \in \mathbb{N}$

$$l'_\alpha(T_n^\alpha, P_n) = 0$$

$$l'_\alpha(T^\alpha(P), P_n) = \frac{1}{n} \sum_{i=1}^{n} \psi_P(x_i) \quad \text{(cf. (2.4), (2.9))}.$$

Therefore

$$l'_\alpha(T^\alpha(P), P_n) - l'_\alpha(T_n^\alpha, P_n) = \frac{1}{n} \sum_{i=1}^{n} \psi_P(x_i). \tag{2.16}$$

Using coordinatewise the mean value theorem analogically as in the proof of Theorem 2.2, we obtain

$$l'_\alpha(T^\alpha(P), P_n) - l'_\alpha(T_n^\alpha, P_n) = l''_\alpha(\theta_n, P_n)\,(T^\alpha(P) - T_n^\alpha) \tag{2.17}$$

where $\theta_n \to T^\alpha(P)$ as $n \to \infty$ P^∞-a.s. Combining (2.16) and (2.17) we get the identity

$$\sqrt{(n)}\,(T_n^\alpha - T^\alpha(P)) = M_n n^{-1/2} \sum_{i=1}^{n} \Omega_P^\alpha(x_i)$$

where $M_n = l''_\alpha(\theta_n, P_n)^{-1}\, l''_\alpha(T^\alpha(P), P)$ for all $\mathscr{X}^n \ni x \mapsto P_n$ such that $\det(l''_\alpha(\theta_n, P_n)) \neq 0$. By (2.15), $l''_\alpha(\theta_n, P_n) \to l''_\alpha(T^\alpha(P), P)$ in P^∞-probability, where $\det(l''_\alpha(T^\alpha(P), P)) \neq 0$ (cf. Theorem 2.2). Hence, M_n tends to the unit $m \times m$ matrix as $n \to \infty$ in P^∞-probability. The assertion of the corollary now follows from the multidimensional central limit theorem (cf. e.g. Anděl [1], p. 185), from the multidimensional version of the Cramér-Slutskij theorem (cf. Fuller [4], pp. 140–145) and from the identities

$$\mathsf{E}_P \Omega_P^\alpha = -l_\alpha''(T^\alpha(P), P)^{-1} \mathsf{E}_P \psi_P = \qquad \text{(cf. (2.9))}$$
$$= -l_\alpha''(T^\alpha(P), P)^{-1} l_\alpha'(T^\alpha(P), P) = 0 \qquad \text{(cf. (2.9), (2.4), 2.10))}.$$

Efficiency and robustness control. Let us suppose that the assumptions of Theorems 2.2 and 2.3 hold for $\alpha \in [0, \alpha_0]$, $\alpha_0 > 0$, and let for some $P \in \mathscr{P}_0$

$$\lim_{\alpha \to 0} \mathsf{E}_P(\Omega_P^\alpha \Omega_P^{\alpha T}) = \mathsf{E}_P(\Omega_P^0 \Omega_P^{0T}).$$

It is typical that the influence curves Ω_P^α are bounded on $\mathscr{X}$ for $\alpha > 0$ and unbounded for $\alpha = 0$. In this situation, choosing T^α with suitably small $\alpha > 0$, one estimates the unknown parameter in a robust manner with efficiency arbitrarily close to the efficiency of M. L. E. T^0. Generally, one can always control the asymptotic variance $\mathsf{E}_P(\Omega_P^\alpha \Omega_P^{\alpha T})$ and the gross-error sensitivity $\sup_{\mathscr{X}} \Omega_P^\alpha(x)$ (cf. Hampel [5]) by the parameter $\alpha \geqq 0$. We shall analyze this possibility in more detail in Sections 5 and 6, where two concrete families of distorted M. L. E.'s T^α, $\alpha \geqq 0$, are considered.

3. DISTORTED M. L. E.'s OF STRUCTURAL PARAMETERS

$\mathscr{P}_\Theta$ is said structural with a parent $P_o \in \mathscr{P}$ (in symbols $\mathscr{P}_\Theta = P_{o/\Theta}$) if (a) Θ is a group with a neutral element $o \in \Theta$ and with the property $\theta_n \to \theta_0$ for some $\theta_n \in \Theta$, $\theta_0 \in \Theta \cup \{\theta^*\}$ iff $\theta\theta_n \to \theta\theta_0$ for all $\theta \in \Theta$, (b) Θ is homomorphic with a group $[\Theta]$ of one-to-one $\mathscr{A}$-measurable mappings $[\theta]: \mathscr{X} \mapsto \mathscr{X}$, (c) $P_\theta = P_o[\theta]^{-1}$ for all $\theta \in \Theta$, where $[\theta]^{-1}: \mathscr{X} \mapsto \mathscr{X}$ is inverse to $[\theta]$.

Throughout this section we consider a structural family $\mathscr{P}_\Theta$ and the corresponding distorted likelihood function (1.3) satisfying the assumptions of Theorem 1.1. We denote $\theta B = \{\theta\tilde\theta: \tilde\theta \in B\}$ for all $\theta \in \Theta$, $B \subset \Theta$, and say that T^α is equivariant if, for every fixed $\theta \in \Theta$ and $P \in \mathscr{P}_0$, it holds $P[\theta]^{-1} \in \mathscr{P}_0$ and

$$\operatorname{argmax}_\Theta l_\alpha(\tilde\theta, P[\theta]^{-1}) = \theta \operatorname{argmax}_\Theta l_\alpha(\tilde\theta, P). \tag{3.1}$$

Theorem 3.1. If there exist bounded continuous mappings $A: \Theta \to (0, \infty)$, $B: \Theta \to \mathbb{R}$ such that

$$l_\alpha(\tilde\theta, P[\theta]^{-1}) = A(\theta)\, l_\alpha(\theta^{-1}\tilde\theta, P) + B(\theta) \quad \text{for every} \quad \theta, \tilde\theta \in \Theta, \quad P \in \mathscr{P}_0, \tag{3.2}$$

and either $\mathscr{P}_0 = \mathscr{P}$ or $l_\alpha(\theta^*, P) = -\infty$ for all $P \in \mathscr{P}_0$ then T^α is equivariant.

Proof. Let $\theta \in \Theta$, $P \in \mathscr{P}_0$ be arbitrary fixed. Generally $P \in \mathscr{P}_0$ implies $P[\theta]^{-1} \in \mathscr{P}_0$, for if $\mathscr{P}_0 \neq \mathscr{P}$ then $l_\alpha(\tilde\theta, P[\theta]^{-1})$ satisfies the assumptions of Theorem 1.1 iff $l_\alpha(\tilde\theta, P)$ does so. Therefore $P[\theta]^{-1} \in \mathscr{P}_0$. Further if (1.14) holds then $l_\alpha(\tilde\theta_0, P[\theta]^{-1}) = \max_\Theta l_\alpha(\tilde\theta, P[\theta]^{-1})$ iff $l_\alpha(\bar\theta, P) = \max_\Theta l_\alpha(\tilde\theta, P)$ for $\bar\theta = \theta^{-1}\tilde\theta_0$. In other words, $\tilde\theta_0 \in \operatorname{argmax}_\Theta l_\alpha(\tilde\theta, P[\theta]^{-1})$ iff $\tilde\theta_0 = \theta\bar\theta$ for $\bar\theta \in \operatorname{argmax}_\Theta l_\alpha(\tilde\theta, P)$, i.e. (3.1) holds. □

Denote $[\theta] . \boldsymbol{x} = ([\theta](x_1), \ldots, [\theta](x_n))$ for every $\boldsymbol{x} \in \mathscr{X}^n$. It holds $\boldsymbol{x} \to P_n$ iff

$[\theta] . x \to P_n[\theta]^{-1}$. Hence, in the case $\{T_n^\alpha(x)\} = \text{argmax}_\Theta\, l_\alpha(\tilde{\theta}, P)$, (3.1) takes on the well-known form

(3.3) $$T_n^\alpha([\theta] . x) = \theta\, T_n^\alpha(x) \quad \text{for all} \quad \theta \in \Theta .$$

Theorem 3.2. Let for a structural family $\mathscr{P}_\Theta$ the conditions of Theorem 1.1 hold, let T^α be equivariant and let $\{T^\alpha(P)\} = \text{argmax}_\Theta\, l_\alpha(\theta, P)$ for some $P \in \mathscr{P}_0$. Then $T_n^\alpha(x)\, T^\alpha(P)^{-1} \to \theta$ as $n \to \infty$ $(P[\theta]^{-1})^\infty$-a.s. for all $\theta \in \Theta$.

Proof. If $\{T^\alpha(P)\} = \text{argmax}_\Theta\, l_\alpha(\tilde{\theta}, P)$ then, by (3.1), $\{T^\alpha(P[\theta]^{-1})\} = \{\theta\, T^\alpha(P)\} = \text{argmax}_\Theta\, l_\alpha(\tilde{\theta}, P[\theta]^{-1})$ for every fixed $\theta \in \Theta$. This implies that the assumptions of Theorem 2.1 hold for all $P = P[\theta]^{-1}$, $\theta \in \Theta$. Hence, by Theorem 2.1, $T_n^\alpha(x) \to T^\alpha(P[\theta]^{-1}) = \theta\, T^\alpha(P)$ as $n \to \infty$ $(P[\theta]^{-1})^\infty$-a.s. But, by the assumption (a) in the definition of structural family, this implies the desired statement. □

Theorem 3.3. Let the assumptions of Theorem 3.2 hold, let $\tilde{\theta}\theta = A(\tilde{\theta})\,\theta + a(\tilde{\theta})$ for all $\theta, \tilde{\theta} \in \Theta$ and for some $m \times m$ or $m \times 1$ matrix functions $A(\tilde{\theta})$ or $a(\tilde{\theta})$ with real-valued components, and let for all $x \in \mathscr{X}$ and all sufficiently small $\varepsilon > 0$

(3.4) $$\{T^\alpha(P_{\varepsilon,x})\} = \text{argmax}_\Theta\, l_\alpha(\theta, P_{\varepsilon,x}) \quad (\text{cf. } (2.7)) .$$

Then the influence curve Ω_θ^α of T^α at $P[\theta]^{-1}$ exists if Ω_P^α exists and

(3.5) $$\Omega_\theta^\alpha(x) = A(\theta)\, \Omega_P^\alpha([\theta]^{-1}(x)) \quad \text{for every} \quad x \in \mathscr{X} , \quad \theta \in \Theta .$$

If, moreover, $\tilde{\theta}\theta = B(\theta)\,\tilde{\theta} + b(\theta)$ for all θ, $\tilde{\theta} \in \Theta$ and for some matrix functions $B(\theta)$, $b(\theta)$ with real-valued components, then the influence curve $\tilde{\Omega}_\theta^\alpha$ of the 'right-modified version' $\tilde{T}^\alpha = T^\alpha T^\alpha(P)^{-1}$ of T^α (cf. Theorem 3.2) at $P[\theta]^{-1}$ exists iff the influence curve Ω_P^α of T^α at P exists and

(3.6) $$\tilde{\Omega}_0^\alpha(x) = C(\theta)\, \Omega_P^\alpha([\theta]^{-1}(x)) \quad \text{for every} \quad x \in \mathscr{X} , \quad \theta \in \Theta ,$$

where

(3.7) $$C(\theta) = A(\theta)\, B(T^\alpha(P)^{-1}) .$$

Proof. Put $\tilde{P}_\theta = P[\theta]^{-1}$. It holds

(3.8) $$(\tilde{P}_\theta)_{\varepsilon,x} = (P[\theta]^{-1})_{\varepsilon,x} = P_{\varepsilon,[\theta]^{-1}(x)}[\theta]^{-1}$$

for all $\varepsilon \in (0,1)$ and $x \in \mathscr{X}$, $\theta \in \Theta$. Therefore, by (3.1), (3.4) and by the linear representation of the associative group multiplication,

$$T^\alpha((\tilde{P}_\theta)_{\varepsilon,x}) - T^\alpha(\tilde{P}_\theta) = \theta\, T^\alpha(P_{\varepsilon,[\theta]^{-1}(x)}) - \theta\, T^\alpha(P) = $$
$$= A(\theta)\,(T^\alpha(P_{\varepsilon,[\theta]^{-1}(x)}) - T^\alpha(P)) .$$

Therefore, by the definition of Ω_θ^α and Ω_P^α (cf. (2.7)), Ω_θ^α exists if Ω_P^α does so and the identity (3.5) holds. Replacing T^α by $\tilde{T}^\alpha = T^\alpha T^\alpha(P)^{-1}$ we get from (3.8)

$$\tilde{T}^\alpha((\tilde{P}_\theta)_{\varepsilon,x}) - \tilde{T}^\alpha(\tilde{P}_\theta) = C(\theta)\,(T^\alpha(P_{\varepsilon,[\theta]^{-1}(x)}) - T^\alpha(P))$$

with $C(\theta)$ given by (3.7). This identity and the preceding result imply the second assertion of Theorem 3.3. □

Theorem 3.4. Let the assumptions of Theorems 3.3 and 2.3 hold. Then $\sqrt{(n)}\,(T_n^\alpha(x) - T^\alpha P[\theta]^{-1}) \to N(0, A(\theta)\,\mathsf{E}_P(\Omega_P^\alpha \Omega_P^{\alpha\mathrm{T}})\,A(\theta)^\mathrm{T})$ for all $\theta \in \operatorname{int}\Theta$ and $\sqrt{(n)}\,(\tilde{T}_n^\alpha(x) - \theta) \to N(0, C(\theta)\,\mathsf{E}_P(\Omega_P^\alpha \Omega_P^{\alpha\mathrm{T}})\,C(\theta)^\mathrm{T})$ for all $\theta \in \operatorname{int}\Theta$, where Ω_P^α and $C(\theta)$ are as in Theorem 3.3 and both convergences are $(P[\theta]^{-1})^\infty$ – weak as $n \to \infty$.

Proof. If the conditions of Theorem 2.3 hold for P then, by (3.5), they hold for all $P[\theta]^{-1}$ with $\theta \in \operatorname{int}\Theta$. Therefore, by Theorem 2.3,

$$\sqrt{(n)}\,(T_n^\alpha(x) - T^\alpha(P[\theta]^{-1})) \to N(0, \mathsf{E}_{P[\theta]^{-1}}(\Omega_\theta^\alpha \Omega_\theta^{\alpha\mathrm{T}}))\quad (P[\theta]^{-1})^\infty\text{-weakly.}$$

But by (3.5) it holds

$$\mathsf{E}_{P[\theta]^{-1}}(\Omega_\theta^\alpha \Omega_\theta^{\alpha\mathrm{T}}) = A(\theta)\,\mathsf{E}_{P[\theta]^{-1}}[\Omega_P^\alpha([\theta]^{-1}(x))\,\Omega_P^\alpha([\theta]^{-1}(x))^\mathrm{T}]\,A(\theta)^\mathrm{T} = A(\theta)\,\mathsf{E}_P[\Omega_P^\alpha(x)\,\Omega_P^\alpha(x)^\mathrm{T}]\,A(\theta)^\mathrm{T}$$

so that the first convergence is proved. At the beginning of the proof of Theorem 3.2 it was proved that $T^\alpha(P[\theta]^{-1}) = \theta\,T^\alpha(P)$ for all $\theta \in \Theta$. Therefore $\tilde{T}^\alpha(P[\theta]^{-1}) = T^\alpha(P[\theta]^{-1})\,T^\alpha(P)^{-1} = \theta$ for all $\theta \in \Theta$ and, consequently,

$$\sqrt{(n)}\,(\tilde{T}_n^\alpha(x) - \theta) = \sqrt{(n)}\,[T_n^\alpha(x)\,T^\alpha(P)^{-1} - T^\alpha(P[\theta]^{-1})\,T^\alpha(P)^{-1}] = B(T^\alpha(P)^{-1})\,\sqrt{(n)}\,(T_n^\alpha(x) - T^\alpha(P[\theta]^{-1}))\,.$$

Thus the second assertion of Theorem 3.4 follows from the first one. □

4. IMPLICATIONS FOR M. L. E.'s

In the present section we summarize some implications of the general theory of Sections 1–3 for the M. L. E.'s T^0. Remind that T^0 is defined in the present paper by the condition $T^0(P) \in \operatorname{argmax}_\Theta l_0(\theta, P)$ for all $P \in \mathscr{P}_0 \supset \mathscr{P}_\Theta \cup \mathscr{P}_e$, where $\mathscr{P}_\Theta$ is a family of theoretical distributions dominated by a sigma-finite measure λ with densities $p_\theta(x)$ on $\Theta \times \mathscr{X}$, $\mathscr{P}_e$ is a family of empirical distributions and

(4.1) $$l_0(\theta, P) = \mathsf{E}_P f_0(p_\theta) = \mathsf{E}_P \ln p_\theta \quad \text{for all} \quad \theta \in \Theta\,, \quad P \in \mathscr{P}_0\,.$$

We shall consider the following conditions concerning $\mathscr{P}_\Theta$ and $\mathscr{P}_0$:

A 1: $p_\theta(x)$ is positive and bounded on $\Theta \times \mathscr{X}$,

A 2: $p_\theta(x)$ is continuous on $\Theta \times \mathscr{X}$ with $\lim_{\theta \to \theta^*} p_\theta(x) = 0$ for every $x \in \mathscr{X}$ (cf. (1.5)),

A 3: for every $\tilde{\theta} \in \Theta$ there exists an open neighborhood $B(\tilde{\theta}) \subset \Theta$ such that $\mathsf{E}_P \ln (\inf_{B(\tilde{\theta})} p_\Theta) > -\infty$,

A 4: $p_\theta = p_{\tilde{\theta}}$ λ-a.s. for no different $\theta, \tilde{\theta} \in \Theta$,

A 5: $p_\theta(x)$ satisfies the regularity assumptions of Section 2 for $\alpha = 0$ and $-l_0''(\theta, P) = \mathsf{E}_P[(p_\theta'/p_\theta)\,(p_\theta'/p_\theta)^\mathrm{T} - p_\theta''/p_\theta] > 0$ on $\operatorname{int}\Theta$ for every $P \in \mathscr{P}_0$.

A 6: The squares of components of the matrix $(p_\theta'/p_\theta)\,(p_\theta'/p_\theta)^\mathrm{T} - p_\theta''/p_\theta$ are uniformly P-integrable in an open neighborhood of each $\tilde{\theta} \in \operatorname{int}\Theta$.

Assertions similar to those that follow have been widely established in the literature since R. A. Fisher [3] introduced the concept of M. L. E. We avoid the tedious task to list all relevant references here.

Theorem 4.1. If A 1 – A 3 hold then all expectations in (4.1) are well-defined and the T^0 exists.

Proof. The first assertion follows from A 3. Since under A 1 – A 3 all assumptions of Theorem 1.1 hold, the second assertion follows from Theorem 1.1. □

Theorem 4.2. If A 1 – A 4 hold then, for every fixed $\theta \in \Theta$,

$$\{T^0(P_\theta)\} = \{\theta\} = \operatorname{argmax}_\Theta l_0(\tilde{\theta}, P_\theta) . \tag{4.2}$$

If (4.2) holds for P_θ replaced by any $P \in \mathscr{P}_0$ then

$$\lim_{n\to\infty} T_n^0(x) = \theta \quad P^\infty\text{-a.s.} \tag{4.3}$$

Proof. (4.2) follows from the fact that $l_0(\tilde{\theta}, P_\theta) \leqq l_0(\theta, P_\theta)$ for every $\theta, \tilde{\theta}$ with the equality iff $P_\theta = P_{\tilde{\theta}}$ (cf. Theorem 5 of Perez [9]), where, by A 4, the equality takes place iff $\theta = \tilde{\theta}$. By (4.2) and by what has been said in the proof of Theorem 4.1, the assumptions of Theorem 2.1 hold for P under consideration with $T^0(P) = \theta$. Therefore (4.3) follows from Theorem 2.1. □

Theorem 4.3. If A 1 – A 5 hold and if for some $P \in \mathscr{P}_0$ there is a unique root $\theta = \theta(P)$ of the equation $l_0'(\theta, P) = \mathsf{E}_P(p_\theta'/p_\theta) = 0$ on int Θ, then

$$\{T^0(P)\} = \{\theta\} = \operatorname{argmax}_\Theta l_0(\tilde{\theta}, P) \tag{4.4}$$

and the influence curve of T^0 at P is given by

$$\Omega_P^0(x) = -l_0''(\theta, P)^{-1} \frac{p_\theta'(x)}{p_\theta(x)} \quad \text{for all} \quad x \in \mathscr{X} . \tag{4.5}$$

For all $P = P_\theta \in \mathscr{P}_0$, $\theta \in \operatorname{int} \Theta$, the root $\theta(P)$ is unique and equal θ and $l_0''(\theta, P) = -I(\theta)$, where $I(\theta)$ is the Fisher information of $\mathscr{P}_\Theta$ at θ.

Proof. (4.4) follows from the fact that, under A 5, $l_0(\theta, P)$ is strictly concave on Θ. Since $\delta_x \in \mathscr{P}_e \subset \mathscr{P}_0$ for every $x \in \mathscr{X}$, $l_0(\theta, \delta_x) = \ln p_\theta(x)$ is strictly concave on Θ for every $x \in \mathscr{X}$ as well. Therefore, $l_0(\theta, P_{\varepsilon,x}) = (1 - \varepsilon)\, l_0(\theta, P) + \varepsilon \ln p_\theta(x)$ is a system of strictly concave functions tending to the strictly concave function $l_0(\theta, P)$ as $\varepsilon \to 0$. It follows from here that (2.8) holds and, moreover,

$$\{T(P_{\varepsilon,x})\} = \operatorname{argmax}_\Theta l_0(\theta, P_{\varepsilon,x}) . \tag{4.6}$$

Since the remaining conditions of Theorem 2.2 with $\alpha = 0$ under A 1 – A 5 hold as well (cf. A 5 and the proof of Theorem 4.2), (4.5) follows from (2.9) with $\alpha = 0$ and from (4.4). The last assertion follows from (4.2) and from the fact that $-l_0''(\theta, P_\theta) = \mathsf{E}_{P_\theta}[(p_\theta'/p_\theta)(p_\theta'/p_\theta)^{\mathrm{T}}] - \mathsf{E}_{P_\theta}(p_\theta''/p_\theta) = I(\theta) - \mathsf{E}_\lambda p_\theta''$, where $\mathsf{E}_\lambda p_\theta'' = 0$ by (2.6). □

Theorem 4.4. If the assumptions of Theorem 4.3 and A 6 hold then $\sqrt{(n)}$.

$\cdot(T_n^0(x) - \theta) \to N(0, \mathsf{E}_P(\Omega_P^0 \Omega_P^{0\mathrm{T}}))$ P^∞-weakly as $n \to \infty$. If $P = P_\theta \in \mathscr{P}_0$ for $\theta \in \operatorname{int} \Theta$, then $\mathsf{E}_P(\Omega_P^0 \Omega_P^{0\mathrm{T}}) = I(\theta)^{-1}$, where $I(\theta)$ is the Fisher information of $\mathscr{P}_\Theta$ at θ.

Proof. By A 5, $-l_0''(\theta, P)$ is positive definite so that $-l_0''(\theta, P)^{-1}$ has all components bounded. By A 6, $\mathsf{E}_P[(p_\theta'/p_\theta)(p_\theta'/p_\theta)^\mathrm{T}]$ has all components bounded too. Hence, by (4.5), all components of the matrix $\mathsf{E}_P(\Omega_P^0, \Omega_P^{0\mathrm{T}})$ are finite. By A 6, (2.14) holds as well. Consequently, the first assertion of Theorem 4.4 follows from Theorem 2.3 and the second one from Theorem 4.3. □

In the rest of this section we consider Euclidean sample spaces $\mathscr{X} = \mathbb{R}^m$, $m \in \mathbb{N}$, with the norm $\|x\|^2 = xx^\mathrm{T}$, $x \in \mathbb{R}^m$, Euclidean parameter spaces $\Theta = \mathbb{R}^m \times [\delta, \delta^{-1}]$, $\delta \in (0, 1)$, and Lebesgue dominating measures λ on $\mathscr{X}$. We introduce a general location and scale structure as follows. Let $\theta^\mathrm{T} = (\mu, \sigma) \in \mathbb{R}^m \times [\delta, \delta^{-1}]$, where $\mu = (\mu_1, \ldots, \mu_m) \in \mathbb{R}^m$ (unless necessary, we do not distinguish between (μ, σ) and $(\mu, \sigma)^\mathrm{T}$). If we put

$$\theta\tilde{\theta} = (\mu, \sigma)(\tilde{\mu}, \tilde{\sigma}) = (\mu + \sigma\tilde{\mu}, \sigma\tilde{\sigma}) \quad \text{and} \quad [\theta](x) = \mu + \sigma x\,, \tag{4.7}$$

then (a), (b), (c) in Section 3 hold ($\theta = (\mu, \sigma) \to \theta^*$ iff $\|\mu\| \to \infty$). Let $\mathscr{P}_\Theta$ be a dominated structural family with a parent density p_o. It holds

$$p_\theta(x) = \frac{1}{\sigma^m}\, p_o\left(\frac{x - \mu}{\sigma}\right) \quad \text{for every} \quad x \in \mathscr{X}\,, \quad \theta = (\mu, \sigma) \in \Theta\,. \tag{4.8}$$

Denote by M^0 and S^0 the μ- and σ-components of the M. L. E. T^0 defined by this family $\mathscr{P}_\Theta$.

We shall consider the following conditions concerning p_o and $\mathscr{P}_0$:

B 1: $p_o(x)$ is a continuous, bounded, decreasing function of $\|x\|$ for $x \in \mathscr{X}$,

B 2: $l_0(\theta, P) = \mathsf{E}_{P[\theta]} \ln p_o > -\infty$ for all $\theta \in \Theta$, $P \in \mathscr{P}_0$.

Theorem 4.5. If B 1, B 2 hold, then A 1 – A 4 hold too, T^0 exists, $\{T^0(P_\theta)\} = \{\theta\} = \operatorname{argmax}_\Theta l_0(\tilde{\theta}, P_\theta)$ for every fixed $\theta \in \Theta$ and $T_n^0(x) \to \theta$ P_θ^∞-a.s. for all $\theta \in \Theta$.

Proof. The fact that A 1, A 2, A 4 follow from B 1 is clear from (4.8) and from the equivalence between $\theta = (\mu, \sigma) \to \theta^*$ and $\|\mu\| \to \infty$. If we prove that A 3 holds too, then the desired assertions will follow from Theorem 4.1 and 4.2. By (4.8), A 3 holds if, for each fixed $(\bar{\mu}, \bar{\sigma}) \in \Theta$ and every $\varepsilon < \delta$,

$$\mathsf{E}_P \ln\left(\max_{\|\mu\|, |\sigma| \leqq \varepsilon} p_o\left(\frac{x - \bar{\mu} + \mu}{\bar{\sigma} + \sigma}\right)\right) > -\infty \tag{4.9}$$

(take into account the inequalities $\delta < \bar{\sigma} < 1/\delta$, $\ln(\bar{\sigma} + \sigma)^{-m} \geqq -m \ln(\varepsilon + 1/\delta) > -m \ln(\delta + 1/\delta)$). Let $\mathbf{1}(x) = x/\|x\|$ for all $x \neq 0$ and $\mathbf{1}(x) = (1, 0, \ldots, 0) \in \mathscr{X}$ for $x = 0$. Since

$$\max_{\|\mu\|, |\sigma| \leqq \varepsilon} \left\|\frac{x - \bar{\mu} + \mu}{\bar{\sigma} + \sigma}\right\| = \left\|\frac{x - \bar{\mu} + \varepsilon\,\mathbf{1}(x - \bar{\mu})}{\bar{\sigma} - \varepsilon}\right\|,$$

the monotony of p_o assumed in B 1 implies

$$\mathsf{E}_P \ln\left(\max_{\|\mu\|,|\sigma|\geqq\varepsilon} p_o\left(\frac{x-\bar{\mu}+\mu}{\bar{\sigma}+\sigma}\right)\right) = \mathsf{E}_P \ln p_o\left(\frac{x-\bar{\mu}+\varepsilon\,\mathbf{1}(x-\bar{\mu})}{\bar{\sigma}-\varepsilon}\right) =$$

$$= m\ln(\bar{\sigma}-\varepsilon) + \mathsf{E}_{P[\bar{\mu},\bar{\sigma}-\varepsilon]} \ln p_o(x+\bar{\varepsilon}\,\mathbf{1}(x)) \quad \text{where} \quad \bar{\varepsilon} = \frac{\varepsilon}{\sigma-\varepsilon} > 0\,.$$

Since $\|x+\bar{\varepsilon}\,\mathbf{1}(x)\| < \|2x\|$ for $\|x\| > \bar{\varepsilon}$ and since $p_o(x+\bar{\varepsilon}\,\mathbf{1}(x))$, $p_o(2x)$ are bounded from below as well as from above by positive constants for $\|x\| < \bar{\varepsilon}$, the monotony of p_o assumed in B 1 now implies that (4.9) holds provided

$$\mathsf{E}_{P[\bar{\mu},\bar{\sigma}-\varepsilon]} \ln p_o(2x) > -\infty\,.$$

But

$$\mathsf{E}_{P[\bar{\mu},\bar{\sigma}-\varepsilon]} \ln p_o(2x) = \mathsf{E}_{P[\bar{\mu},(\bar{\sigma}-\varepsilon)/2]} \ln p_o(x)$$

where the right-hand expectation is bounded from below by B 2, so that (4.9) holds. □

Theorem 4.6. If B 1, B 2 hold, then T^0 is equivariant.

Proof. Let $\theta = (\mu, \sigma)$, $\tilde{\theta} = (\tilde{\mu}, \tilde{\sigma}) \in \Theta$ and $P \in \mathscr{P}_0$ be arbitrary fixed. By (4.1) and (4.8),

$$l_0(\tilde{\theta}, P[\theta]^{-1}) = \mathsf{E}_{P[\theta]^{-1}} \ln p_\theta = \mathsf{E}_{P[\theta]^{-1}} \ln\left(\frac{1}{\tilde{\sigma}}\, p_o([\tilde{\theta}]^{-1}(x))\right) =$$

$$= \mathsf{E}_{P[\theta]^{-1}} \ln\left(\frac{1}{\tilde{\sigma}}\, p_o([\tilde{\theta}]^{-1}[\theta][\theta]^{-1}(x))\right) = \mathsf{E}_P \ln\left(\frac{1}{\tilde{\sigma}}\, p_o([\theta^{-1}\tilde{\theta}]^{-1}(x))\right) =$$

$$= \mathsf{E}_P \ln\left(\frac{1}{\tilde{\sigma}}\, p_o\left(\left[\frac{\tilde{\mu}-\mu}{\sigma}, \frac{\tilde{\sigma}}{\sigma}\right]^{-1}(x)\right)\right) = \mathsf{E}_P \ln\left(\frac{\sigma}{\tilde{\sigma}}\, p_o\left(\left[\frac{\tilde{\mu}-\mu}{\sigma}, \frac{\tilde{\sigma}}{\sigma}\right]^{-1}(x)\right)\right) -$$

$$- \ln\sigma = \mathsf{E}_P \ln p_{\theta^{-1}\tilde{\theta}} - \ln\sigma = l_0(\theta^{-1}\tilde{\theta}, P) - \ln\sigma\,.$$

Since, moreover, $l_0(\theta^*, P) = -\infty$ for all $P \in \mathscr{P}_0$ (cf. Theorem 4.5 and the proofs of Theorems 4.1 and 1.1), all assumptions of Theorem 3.1 hold and the equivariance of T^0 follows from Theorem 3.1. □

Theorem 4.7. Let B 1, B 2 and A 5 hold and let there is a unique root $\theta = (\mu(P), \sigma(P))$ of the equation $\mathsf{E}_P(p'_\theta/p_\theta) = 0$ on int Θ. If $P \ll \lambda$ and $p = \mathrm{d}P/\mathrm{d}\lambda$ satisfies B 1 then $\mu(P) = 0$,

$$\text{(4.10)}\qquad \tilde{T}_n^0(x) = \left(M_n^0(x), \frac{S_n^0(x)}{\sigma(P)}\right) \to \theta \quad \text{as} \quad n\to\infty \quad (P[\theta]^{-1})^\infty\text{-a.s.}$$

for all $\theta = (\mu, \sigma) \in \Theta$, and the influence curve of $\tilde{T}^0 = (M^0, S^0/\sigma(P))^{\mathrm{T}}$ at $P[\theta]^{-1}$ is given by

$$\text{(4.11)}\qquad \tilde{\Omega}_\theta^0(x) = \sigma C\Omega_P^0\left(\frac{x-\mu}{\sigma}\right) \quad \text{for every} \quad x \in \mathscr{X}\,, \quad \theta \in \Theta\,,$$

where Ω_P^0 is given by (4.5) and

$$C = \begin{bmatrix} I_m & 0 \\ 0 & 1/\sigma(P) \end{bmatrix} \quad (I_m \text{ is the } m \times m \text{ unit matrix}). \tag{4.12}$$

Proof. By Theorem 4.4, (4.4) holds. By (4.4), $T^0(P) = (M^0(P), S^0(P)) = (\mu(P), \sigma(P))$. By Theorem 1 of Anderson [2] (cf. also Example 2 of Pfanzagl [11]), if p satisfies B 1 then, for every fixed $\sigma > 0$, $\mathsf{E}_P \ln p_{\mu,\sigma}$ is maximized by $\mu = 0$. This implies $\mu(P) = 0$. Since by (4.7) it holds $(M^0(P), S^0(P))^{-1} = (0, \sigma(P))^{-1} = (0, \sigma(P)^{-1})$, since the conditions of Theorem 1.1 hold (cf. Theorem 4.5 and the proof of Theorem 4.1) and since T^0 is equivariant (cf. Theorem 4.6), (4.10) follows from Theorem 3.2. (4.11) follows from Theorems 4.5, 4.3 and 3.3 and from the fact that $A(\tilde{\mu}, \tilde{\sigma}) = \tilde{\sigma} I_{m+1}$, $a(\tilde{\mu}, \tilde{\sigma}) = (\tilde{\mu}, 0)^{\mathrm{T}} \in \mathbb{R}^{m+1}$, and

$$B(\mu, \sigma) = \begin{bmatrix} I_m & \mu^{\mathrm{T}} \\ 0 & \sigma \end{bmatrix}, \quad b(\mu, \sigma) = 0 \tag{4.13}$$

(cf. Theorem 3.3; by (4.6), the assumption (3.4) of Theorem 3.3 holds). □

Theorem 4.8. If all conditions of Theorem 4.7 as well as A 6 hold, then

$$\sqrt{(n)}\,(\tilde{T}_n^0(x) - \theta) \to N(0, \mathsf{E}_{P[\theta]^{-1}}(\tilde{\Omega}_\theta^0 \tilde{\Omega}_\theta^{0\mathrm{T}})) \quad (P[\theta]^{-1})^\infty\text{-weakly as } n \to \infty \tag{4.14}$$

for the estimator $\tilde{T}^0$ and influence curves $\tilde{\Omega}_\theta^0$ defined in Theorem 4.7 and for all $\theta \in \operatorname{int} \Theta$.

Proof. By the preceding proof, the assumptions of Theorem 3.3 hold. By A 6 and Theorem 4.3, the assumptions of Theorem 2.3 hold for $\alpha = 0$ too. Thus the assumptions of Theorem 3.4 hold and the desired assertion follows from Theorem 3.4. □

Example 4.1. Let $p_o(x) = (2\pi)^{-m/2} \exp(-\|x\|^2/2)$ for every $x \in \mathscr{X}$. This p_o obviously satisfies B 1. Moreover, by (4.8), it holds for all $\theta \in \operatorname{int} \Theta$

$$\frac{p'_\theta(x)}{p_\theta(x)} = \frac{\mathrm{d}}{\mathrm{d}\theta} \ln p_\theta(x) = \frac{1}{\sigma} \begin{bmatrix} \left(\dfrac{x-\mu}{\sigma}\right)^{\mathrm{T}} \\ \left\|\dfrac{x-\mu}{\sigma}\right\|^2 - m \end{bmatrix} \tag{4.15}$$

$$\frac{p''_\theta(x)}{p_\theta(x)} - \left(\frac{p'_\theta}{p_\theta}\right)\left(\frac{p'_\theta}{p_\theta}\right)^{\mathrm{T}} = \left(\frac{\mathrm{d}}{\mathrm{d}\theta}\right)^{\mathrm{T}}\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\right) \ln p_\theta(x) = \tag{4.16}$$

$$= -\frac{1}{\sigma^2} \begin{bmatrix} I_m & 2\left(\dfrac{x-\mu}{\sigma}\right)^{\mathrm{T}} \\ 2\left(\dfrac{x-\mu}{\sigma}\right) & 3\left\|\dfrac{x-\mu}{\sigma}\right\|^2 - m \end{bmatrix}.$$

Let $P \in \mathscr{P}$, $P \ll \lambda$, be a fixed distribution with a density $p = \mathrm{d}P/\mathrm{d}\lambda$ satisfying the condition B 1 with $\mathsf{E}_P\|x\|^2 < \infty$. By (4.1), (4.15), (4.16),

$$(4.17)\quad l_0(\theta, P) = -\frac{1}{2}\left[m \ln \sigma^2 + \frac{\mathsf{E}_P\|x\|^2 + \|\mu\|^2}{\sigma^2}\right] \quad \text{for all} \quad \theta = (\mu, \sigma)^{\mathrm{T}} \in \Theta\,,$$

and

$$(4.18)\qquad l_0'(\theta, P) = \mathsf{E}_P \frac{\mathrm{d}}{\mathrm{d}\theta} \ln p_\theta(x) = \frac{1}{\sigma}\begin{bmatrix} -\left(\dfrac{\mu}{\sigma}\right)^{\mathrm{T}} \\ \dfrac{\mathsf{E}_P\|x\|^2 + \|\mu\|^2}{\sigma^2} - m \end{bmatrix},$$

$$(4.19)\quad l_0''(\theta, P) = \mathsf{E}_P \left(\frac{\mathrm{d}}{\mathrm{d}\theta}\right)^{\mathrm{T}} \left(\frac{\mathrm{d}}{\mathrm{d}\theta}\right) \ln p_\theta(x) = -\frac{1}{\sigma^2}\left[\begin{array}{c|c} I_m & -2\left(\dfrac{\mu}{\sigma}\right)^{\mathrm{T}} \\ \hline -2\dfrac{\mu}{\sigma} & 3\dfrac{\mathsf{E}_P\|x\|^2 + \|\mu\|^2}{\sigma^2} - m \end{array}\right]$$

for all $\theta = (\mu, \sigma)^{\mathrm{T}} \in \operatorname{int} \Theta$. We shall prove that B 1, B 2, A 5, A 6 (consequently also A 1 – A 4, cf. Theorem 4.5) hold for $\mathscr{P}_0 = \{\tilde{P} \in \mathscr{P}: \mathsf{E}_{\tilde{P}}\|x\|^2 < \infty\} \supset \mathscr{P}_\Theta \cup \mathscr{P}_{\mathrm{e}}$. Indeed, for every $\tilde{P} \in \mathscr{P}_0$

$$(4.20)\quad l_0(\theta, \tilde{P}) = -\frac{1}{2}\left[m \ln \sigma^2 + \frac{\mathsf{E}_{\tilde{P}}\|x\|^2 - 2\mu(\mathsf{E}_{\tilde{P}}x)^{\mathrm{T}} + \|\mu\|^2}{\sigma^2}\right] \quad (\text{cf. } (4.17))\,,$$

and, analogically, (4.18) remain true for arbitrary $\tilde{P} \in \mathscr{P}_0$ provided all dividents $\mathsf{E}_P\|x\|^2 + \|\mu\|^2$ are replaced by the more general divident of (4.20). B 1 is clear. B 2 and A 5 follow from (4.18) – (4.19). A 6 follows from (4.16) and (4.19).

By (4.17), if for P under consideration

$$(4.21)\qquad \sigma(P) = \left[\frac{1}{m}\,\mathsf{E}_P\|x\|^2\right]^{1/2} \in (\delta, \delta^{-1})\,,$$

then $\theta(P) = (0, \sigma(P))^{\mathrm{T}} \in \operatorname{int} \Theta$ is the only point of Θ which maximizes $l_0(\theta, P)$ on Θ (it is at the same time the unique solution of the equation $l_0'(\theta, P) = 0$ on $\operatorname{int} \Theta$, cf. (4.18)). Hereafter we suppose that $\delta > 0$ is selected small enough so that (4.21) holds for P (if $P = P_o$ then $\sigma(P) = 1$ so that (4.21) holds for every $\delta \in (0, 1)$).

By Theorem 4.5, the M. L. E. $T^0 = (M^0, S^0): \mathscr{P}_0 \to \Theta = \mathbb{R}^m \times [\delta, \delta^{-1}]$ exists. By Theorem 4.6, (M^0, S^0) is equivariant. Since $\{(M^0(\tilde{P}), S^0(\tilde{P}))\} = \{(\mathsf{E}_{\tilde{P}}x, (\mathsf{E}_{\tilde{P}} \,.\, \|x\|^2/m)^{1/2})\} = \operatorname{argmax}_\Theta l_0(\theta, \tilde{P})$ for all $\tilde{P} \in \mathscr{P}_0$ (cf. (4.20) for the standard normal p_o) and since $\boldsymbol{x} \to P_n \in \mathscr{P}_{\mathrm{e}} \subset \mathscr{P}_0$ for all $\boldsymbol{x} \in \mathscr{X}^n$, $n \in \mathbb{N}$, (M^0, S^0) as well as $(M^0, S^0/\sigma(P))$ are equivariant in the sense specified in (3.3) for all $\boldsymbol{x} \in \mathscr{X}^n$, $n \in \mathbb{N}$. By Theorem 4.7

$$\lim_{n \to \infty} (M_n^0(\boldsymbol{x}), S_n^0(\boldsymbol{x})/\sigma(P)) = \theta \quad (P[\theta]^{-1})^\infty\text{-a.s.} \quad \text{for all} \quad \theta \in \Theta$$

and, moreover,

$$(4.22)\qquad \tilde{\Omega}_\theta^0(x) = \begin{bmatrix} (x-\mu)^{\mathrm{T}} \\ \dfrac{\sigma}{2\sigma(P)}\left(\dfrac{1}{m\,\sigma(P)^2}\left\|\dfrac{x-\mu}{\sigma}\right\|^2 - 1\right)\end{bmatrix} \quad \text{for every} \quad x \in \mathscr{X}$$

is the influence curve of $(M^0, S^0/\sigma(P))^{\mathrm{T}}$ at $P[\theta]^{-1}$ for every $\theta = (\mu, \sigma) \in \operatorname{int}\Theta$. The formula (4.22) follows from (4.5), (4.11), (4.12) and from the fact that, by (4.15), (4.19),

$$\frac{p'_{0,\sigma(P)}(x)}{p_{0,\sigma(P)}(x)} = \frac{1}{\sigma(P)}\begin{bmatrix}\left(\dfrac{x}{\sigma(P)}\right)^{\mathrm{T}} \\ \dfrac{\|x\|^2}{\sigma(P)^2} - m\end{bmatrix}$$

$$-\, l''_0((0, \sigma(P)), P) = \frac{1}{\sigma(P)^2}\left[\begin{array}{c|c} I_m & \\ \hline 0 & \dfrac{3\mathsf{E}_P\|x\|^2}{\sigma(P)^2} - m\end{array}\right] = \frac{1}{\sigma(P)^2}\left[\begin{array}{c|c} I_m & 0 \\ \hline 0 & 2m\end{array}\right] \quad (\text{cf. } (4.21))\,.$$

In the particular case of location, the influence curves $\tilde{\Omega}_\mu^0(x) = \Omega^0(x) = x - \mu$ of the sample mean M^0 (defined by $M^0(\tilde{P}) = \mathsf{E}_{\tilde{P}}x$ for every $\tilde{P} \in \mathscr{P}_0$) at $P[\mu]^{-1}$, $\mu \in \mathbb{R}$, are well known (cf. e.g. Hampel [5]). For the general right-modified least square estimators $(M^0, S^0/\sigma(P))$ the curves (4.22) seem to be new. Note that e.g. for the standard doubly exponential density $p(x) = \exp(-|x|)/2$ it holds $\sigma(P) = 2$ while the standard Cauchy distribution is outside $\mathscr{P}_0$. Note also that, by Theorem 4.8, $\sqrt{(n)}\,(M_n^0(x), S_n^0(x)/\sigma(P)) - (\mu, \sigma))$ is asymptotically normal with the asymptotic mean zero and the asymptotic variance $\mathsf{E}_{P[\theta]^{-1}}(\tilde{\Omega}_\theta^0 \tilde{\Omega}_\theta^{0\mathrm{T}})$ (cf. (4.22)) for every $\theta = (\mu, \sigma) \in \operatorname{int}\Theta$. If the Fisher information $\tilde{I}(\theta)$ of the family $\{P[\theta]^{-1} : \theta \in \Theta\}$ exists and is positive on int Θ then, using the idea employed in the end of Section 2 of Huber [6], one can establish the inequality $\mathsf{E}_{P[\theta]^{-1}}(\tilde{\Omega}_\theta^0 \tilde{\Omega}_\theta^{0\mathrm{T}}) \geqq \tilde{I}(\theta)^{-1}$.

5. TYPE 1 DISTORTED M. L. E.'s

In the present section we consider one concrete class of estimators T^α, $\alpha \in (0, 1)$, defined by the functions

$$(5.1)\quad f_\alpha(u) = \frac{u^\alpha - 1}{\alpha}\,, \quad f'_\alpha(u) = u^{\alpha-1}\,, \quad f''_\alpha(u) = (\alpha - 1)\,u^{\alpha-2}\,, \quad u \in [0, \infty]\,,$$
$$\alpha \in [0, 1]\,.$$

These functions are satisfying the assumptions of Section 1 and, for $\alpha \in (0, 1)$, $f_\alpha(u)$ are uniformly bounded from below by $-1/\alpha$ so that the left-hand inequality in (1.6) is strict for all $\tilde{\theta} \in \Theta$ and all families $\mathscr{P}_\Theta$. This fact considerably simplifies the theory for T^α, $\alpha \in (0, 1)$, against the theory for T^0 presented in Section 4. Throughout this section we preserve the notation introduced in Section 4 and the preceding sections and we consider $\mathscr{P}_0 = \mathscr{P}$.

Theorem 5.1. If A 1, A 2 hold then all T^α, $\alpha \in (0, 1)$, exist.

Proof. All assumptions of Theorem 1.1 follow from A 1, A 2 and the desired assertion follows from Theorem 1.1. □

Lemma 5.1. If A 1, A 2 hold and, moreover,

A*3: $p_\theta'(x)$, $p_\theta''(x)$ defined in (2.6) are continuous and $p_\theta^\alpha(x)\, p_\theta'(x)/p_\theta(x)$, $p_\theta^\alpha(x) \,.\, p_\theta''(x)/p_\theta(x)$ bounded on int $\Theta \times \mathscr{X}$ for all $\alpha \in (0, 1]$

then the regularity assumptions of Section 2 as well as (2.14) hold for all $P \in \mathscr{P}_0 = \mathscr{P}$ and $\alpha \in (0, 1)$.

Proof. By (5.1), A*3 implies the continuity and boundedness of the integrands in (2.4), (2.5). Hence. by the Lebesgue dominated convergence theorem and by the mean value theorem applied componentwise to the integrands, (2.4)–(2.6) hold i.e., in particular,

$$l_\alpha'(\theta, P) = \mathsf{E}_P p_\theta^\alpha \frac{p_\theta'}{p_\theta} \quad \text{for every} \quad \theta \in \text{int}\, \Theta\,, \quad P \in \mathscr{P}\,, \tag{5.2}$$

$$-l_\alpha''(\theta, P) = \mathsf{E}_P p^\alpha \left((1-\alpha)\left(\frac{p_\theta'}{p_\theta}\right)\left(\frac{p_\theta'}{p_\theta}\right)^{\mathrm{T}} - \frac{p_\theta''}{p_\theta}\right) \quad \text{for every } \theta \in \text{int}\, \Theta\,, \quad P \in \mathscr{P}\,. \tag{5.3}$$

(2.14) follows from the fact that, for every $P \in \mathscr{P}$, the integrands in (5.3) are assumed bounded on (int Θ) $\times\, \mathscr{X}$ so that the squares of components of these integrands are bounded on (int Θ) $\times\, \mathscr{X}$ too. □

Lemma 5.2. If for some $P \in \mathscr{P}$

A*4: there is a unique solution $\theta = \theta^\alpha(P) \in \text{int}\, \Theta$ of the equation $l_\alpha'(\theta, P) = 0$ on int Θ and $l_\alpha(\theta^\alpha(P), P) > l_\alpha(\theta, P)$ for all $\theta \in \Theta - \text{int}\, \Theta$, $\alpha \in (0, 1)$,

then $\{T^\alpha(P)\} = \{\theta^\alpha(P)\} = \text{argmax}_\Theta\, l_\alpha(\theta, P)$ as well as (2.8) hold for all $\alpha \in (0, 1)$.

Proof. Let $\alpha \in (0, 1)$ and $P \in \mathscr{P}$ be arbitrary fixed. By Theorem 5.1 there exists $T^\alpha(P) \in \Theta$ maximizing $l_\alpha(\theta, P)$ on Θ. By the inequality assumed in A*4, $T^\alpha(P) \in \text{int}\, \Theta$. Therefore $T^\alpha(P)$ is a stationary point of $l_\alpha(\theta, P)$, i.e. $l_\alpha'(T^\alpha(P), P) = 0$. By assumptions, $T^\alpha(P) = \theta^\alpha(P)$ is unique and the first assertion holds. As to the assertion (2.8), take at first into account that $l_\alpha(\theta, P_{\varepsilon,x}) = (1 - \varepsilon)\, l_\alpha(\theta, P) + \varepsilon(p_\theta(x)^\alpha - 1)/\alpha$ tends for arbitrary fixed $x \in \mathscr{X}$ to $l_\alpha(\theta, P)$ uniformly on Θ as $\varepsilon \to 0$. Since Θ is locally compact with countable base, it is also sigma-compact i.e. there exist compact subsets $B_1 \subset \subset B_2 \subset \ldots \subset \Theta$ with a union equal Θ. The assumption $T^\alpha(P_{\varepsilon_j,x}) \notin B_j$ for all $j \in \mathbb{N}$ and some $\varepsilon_j \to 0$ as $j \to \infty$ contradicts A 1, A 2. Therefore, for all sufficiently small $\varepsilon > 0$ there exists a compact subset $B \subset \Theta$, $T^\alpha(P) \in B$, such that $T^\alpha(P_{\varepsilon,x}) \in B$. The uniform convergence of $l_\alpha(\theta, P_{\varepsilon,x})$ to $l_\alpha(\theta, P)$ together with the assumption that $l_\alpha(\theta, P)$ is maximized at exactly one point $T^\alpha(P) \in \Theta$ imply (2.8). □

Theorem 5.2. If A 1, A 2, A*3, A*4 hold and, moreover,

$$-l_\alpha''(T^\alpha(P), P) > 0 \quad \text{for all} \quad \alpha \in (0, 1) \tag{5.4}$$

then, for all $\alpha \in (0, 1)$,

(i) $T_n^\alpha(x) \to T^\alpha(P) \quad P^\infty$-a.s.,

(ii) the influence curve of T^α at P is bounded and given on $\mathscr{X}$ by

$$\Omega_P^\alpha(x) = -l''_\alpha(\theta, P)^{-1}\, p_\theta^\alpha(x)\, \frac{p'_\theta(x)}{p_\theta(x)} \quad \text{where} \quad \theta = T^\alpha(P)\,, \tag{5.5}$$

(iii) $\sqrt{(n)}\,(T_n^\alpha(x) - T^\alpha(P)) \to N(0, \mathsf{E}_P(\Omega_P^\alpha \Omega_P^{\alpha\mathrm{T}})) \quad P^\infty$-weakly as $n \to \infty$.

Proof. (i) By Lemma 5.2 and the proof of Theorem 5.1, all assumptions of Theorem 2.1 hold so that (i) holds too.

(ii) By (5.4), Lemmas 5.1 and 5.2, and by what has been said in part (i) of this proof, all assumptions of Theorem 2.2 hold. Thus (5.5) follows from (2.9) and (5.1). The boundedness of Ω_P^α follows from A*3.

(iii) By what has been said in part (ii) of this proof and by Lemma 5.1, all assumptions of Theorem 2.3 hold. This Theorem implies the assertion (iii). □

In the rest of this section we consider the general location and scale structural model introduced in Section 4. We assume that the family $\mathscr{P}_\Theta$ defined by a parent density p_o satisfies B 1, A*3 (by what has been said in the roof of Theorem 4.5, A 1, A 2 follow from B 1). We also consider a fixed distribution $P \in \mathscr{P}$ satisfying A*4 with a density $p = \mathrm{d}P/\mathrm{d}\lambda$ satisfying B 1 as well.

Lemma 5.3. It holds $\theta^\alpha(P) = (\mu^\alpha(P), \sigma^\alpha(P)) = (0, \sigma^\alpha(P))$ for all $\alpha \in (0, 1)$.

Proof. This statement follows from the same Theorem 1 of Anderson [2] as the analogical statement of Theorem 4.7. □

Let us suppose in addition to what has been supposed above that

$$-l''_\alpha((0, \sigma^\alpha(P)), P) > 0 \quad \text{for all} \quad \alpha \in (0, 1)\,. \tag{5.6}$$

Theorem 5.3. It holds for all $\alpha \in (0, 1)$ that (i) $T^\alpha = (M^\alpha, S^\alpha)\colon \mathscr{P} \to \Theta = \mathbb{R} \times [\delta, \delta^{-1}]$ exists and is equivariant, (ii) $\tilde{T}^\alpha = T^\alpha T^\alpha(P)^{-1} = (M^\alpha, S^\alpha/\sigma^\alpha(P))$, (iii) $(M_n^\alpha(x), S_n^\alpha(x)/\sigma^\alpha(P)) \to \theta$ as $n \to \infty$ $(P[\theta]^{-1})^\infty$-a.s. for all $\theta = (\mu, \sigma) \in \Theta$, (iv) the influence curve of $(M^\alpha, S^\alpha/\sigma^\alpha(P))$ at $P[\theta]^{-1}$ is bounded and given by

$$\tilde{\Omega}_\theta^\alpha(x) = \sigma \begin{bmatrix} I_m & 0 \\ 0 & 1/\sigma^\alpha(P) \end{bmatrix} (-\,l''_\alpha((0, \sigma^\alpha(P)), P)^{-1}) \times \tag{5.7}$$

$$\times\, p_{0,\sigma^\alpha(P)}\left(\frac{x-\mu}{\sigma}\right) \frac{p'_{0,\sigma^\alpha(P)}\left(\dfrac{x-\mu}{\sigma}\right)}{p_{0,\sigma^\alpha(P)}\left(\dfrac{x-\mu}{\sigma}\right)}$$

for all $\theta = (\mu, \sigma) \in \operatorname{int}\Theta$ (cf. (5.3)),

(v) $$\sqrt{(n)}\,((M_n^\alpha(x), S_n^\alpha(x)/\sigma^\alpha(P)) - (\mu, \sigma)) \to N(0, \mathsf{E}_{P[\theta]^{-1}}(\tilde{\Omega}_\theta^\alpha \tilde{\Omega}_\theta^{\alpha\mathrm{T}}))$$

$$(P[\theta]^{-1})^\infty\text{-weakly} \quad \text{as} \quad n \to \infty \quad \text{for all} \quad \theta = (\mu, \sigma) \in \operatorname{int}\Theta\,.$$

Proof. (i) The existence follows from Theorem 5.1. The equivariance follows from Theorem 3.1 and from the following identities (cf. (3.2))

$$\alpha\, l_\alpha(\tilde\theta, P[\theta]^{-1}) = \mathsf{E}_{P[\theta]^{-1}}(p_{\tilde\theta}^\alpha - 1) = \mathsf{E}_{P[\theta]^{-1}}\left(\frac{1}{\tilde\sigma^\alpha}\, p_o\left(\frac{x-\tilde\mu}{\tilde\sigma}\right)^\alpha - 1\right) =$$

$$= \mathsf{E}_P\left(\frac{1}{\tilde\sigma^\alpha}\, p_o\left(\frac{\mu+\sigma x-\tilde\mu}{\tilde\sigma}\right)^\alpha - 1\right) = \mathsf{E}_P\left(\frac{1}{\sigma^\alpha}\, p_o\left([\theta^{-1}\tilde\theta]^{-1}(x)\right)^\alpha - 1\right) =$$

$$= \frac{1}{\sigma^\alpha}\left(\mathsf{E}_P p^\alpha_{\theta^{-1}\tilde\theta^{-1}}\right) + \frac{1}{\sigma^\alpha} - 1 = \frac{\alpha}{\sigma^\alpha}\, l_\alpha(\theta^{-1}\tilde\theta, P) + \frac{1}{\sigma^\alpha} - 1\,.$$

(ii) follows from Lemma 5.3.

(iii) follows from Theorem 3.2 (cf. (i) of the present Theorem, Lemma 5.2, and part (i) of the proof of Theorem 5.2).

(iv) If (3.4) holds then all assumptions of Theorem 3.3 holds and the desired assertion follows from Theorems 3.3 and 5.2 (cf. (5.6) and (5.4) and (4.11), (4.12)). Thus we shall prove (3.4). By A*3 and (5.6), $-l''_\alpha(\theta, P) > 0$ in an open neighborhood $B \subset \Theta$ of $T^\alpha(P) = (0, \sigma^\alpha(P))$. By A*3 and (5.3). $l''_\alpha(0, P_{\varepsilon,x}) = (1-\varepsilon)\, l''_\alpha(\theta, P) + \varepsilon\, l''_\alpha(0, \delta_x) = (1-\varepsilon)\, l''_\alpha(\theta, P) + \varepsilon[p''_\theta(x)/p_\theta(x) - (1-\alpha)(p'_\theta(x)/p_\theta(x))(p'_\theta(x)/p_\theta(x))^T]$ tends to the locally concave $l''_\alpha(\theta, P)$ uniformly on a compact covering of B as $\varepsilon \to 0$. Therefore there exists an open neighbourhood $\tilde B$ of $T^\alpha(P)$ such that $-l''_\alpha(\theta, P_{\varepsilon,x}) > 0$ for all $\theta \in \tilde B$ and all sufficiently small $\varepsilon > 0$. Since, by Lemma 5.2, $T^\alpha(P_{\varepsilon,x}) \in \tilde B$ for all sufficiently small $\varepsilon > 0$ (cf. (2.8)) and since $l_\alpha(\theta, P_{\varepsilon,x})$ is concave on $\tilde B$ for all sufficiently small $\varepsilon > 0$, the points of maxima of $l_\alpha(\theta, P_{\varepsilon,x})$ on $\tilde B$ are unique for all sufficiently small $\varepsilon > 0$ which proves (3.4).

(v) Since we proved in the part (iii) that all assumptions of Theorem 3.3 hold, (v) follows from the second assertion of Theorem 3.4. □

Example 5.1. Let us consider the same standard normal p_o as in Example 4.1. By (4.15), (4.16), the assumptions B 1, A*3 hold for the corresponding normal family $\mathscr{P}_\Theta$. A*4 holds for P with densities p of normal, doubly-exponential, uniform or Cauchy type. For simplicity of calculations let us consider $p = p_o$. For this P it holds

$$(5.3)\quad \mathsf{E}_P p^\alpha_{0,\sigma}\xi = c(\alpha, \sigma)^m\, \mathsf{E}_{P_{0,\tilde\sigma}}\xi \quad \text{for } c(\alpha, \sigma) = \frac{\sigma^{1-\alpha}}{(2\pi)^{\alpha/2}(\alpha+\sigma^2)^{1/2}},\quad \tilde\sigma^2 = \frac{\sigma^2}{\alpha+\sigma^2},$$

and for any vector-valued measurable $\xi(x)$ defined on $\mathscr{X}$ and any $\sigma > 0$, $\alpha \in (0.\ 1)$. By (5.2), (4.15) and (5.8),

$$l'_\alpha((0, \sigma), P) = \frac{c(\alpha, \sigma)^m}{\sigma}\begin{bmatrix} 0 \\ m\left(\dfrac{1}{\alpha+\sigma^2} - 1\right)\end{bmatrix}.$$

Therefore A*4 holds for the standard normal P with

$$(5.9)\qquad \sigma^\alpha(P) = (1-\alpha)^{1/2} \quad \text{for all} \quad \alpha \in (0, 1-\delta^2)\,.$$

Analogically by (5.3), (4.16) and (5.8),

$$(5.10) \qquad -l''_\alpha((0, \sigma^\alpha(P)), P) = c(\alpha, (1-\alpha)^{1/2})^m \left[\begin{array}{c|c} I_m & 0 \\ \hline 0 & 2m \end{array}\right]$$

$$\text{for all} \quad \alpha \in (1, 1-\delta^2)$$

so that P satisfies (5.6).

Therefore the distorted M. L. E.'s $(M^\alpha, S^\alpha/(1-\alpha)^{1/2})$, $\alpha \in (0, 1-\delta^2)$, are euqi-variant strongly consistent estimators of parameters $(\mu, \sigma) \in \mathbb{R}^m \times [\delta, \delta^{-1}]$.

By (5.7), (4.15) and (5.10), the influence curves of these estimators at $P_{\mu,\sigma}$ are bounded and given by

$$(5.11) \qquad \tilde{\Omega}^\alpha_\theta(x) = \frac{\exp\left(\frac{\alpha}{2(1-\alpha)}\left\|\frac{x-\mu}{\sigma}\right\|^2\right)}{(1-\alpha)^{m/2+1}} \begin{bmatrix} (x-\mu)^{\mathrm{T}} \\ \frac{\sigma}{2}\left(\frac{1}{m(1-\alpha)}\left\|\frac{x-\mu}{\sigma}\right\|^2 - 1\right) \end{bmatrix}$$

$$\text{for all} \quad x \in \mathcal{X}, \theta = (\mu, \sigma) \in \operatorname{int} \Theta$$

and for all $\alpha \in (0, 1)$. As said in Section 4, $\sigma(P) = 1$ in (4.22) provided P is the standard normal distribution. This together with (4.22) imply that (5.10) for $\alpha = 0$ yields the influence curve of the M. L. E. $(M^0(\tilde{P}), S^0(\tilde{P})) = (\mathsf{E}_{\tilde{P}}x, \mathsf{E}_{\tilde{P}}\|x\|^2/m)$ at $P_{\mu,\sigma}$ for all $(\mu, \sigma) \in \operatorname{int} \Theta$. Since all $(M^\alpha, S^\alpha/(1-\alpha)^{1/2})$ are asymptotically normal with the asymptotic mean zero and the asymptotic variance $\mathsf{E}_{P_\theta}(\tilde{\Omega}^\alpha_\theta \tilde{\Omega}^{\alpha\mathrm{T}}_\theta)$ where, by (5.10), $\tilde{\Omega}^\alpha_\theta(x)$ are uniformly bounded by the integrable $\tilde{\Omega}^0_0(x)$ outside the circle $\|x\| = r$ of a large radius r as $\alpha \to 0$. Thus $\mathsf{E}_{P_\theta}(\tilde{\Omega}^\alpha_\theta \tilde{\Omega}^{\alpha\mathrm{T}}_\theta)$ tends to the variance $\mathsf{E}_{P_\theta}(\tilde{\Omega}^0_\theta \tilde{\Omega}^{0\mathrm{T}}_\theta)$ of the M. L. E. which is equal to the Cramér-Rao lower bound $I(\theta)^{-1}$. Therefore, by a proper choice of $\alpha \in (0, 1)$, one can control the efficiency and robustness of the estimates of location and scale $(\mu, \sigma) \in \mathbb{R}^m \times [\delta, \delta^{-1}]$ as claimed at the end of Section 2.

6. TYPE 2 DISTORTED M. L. E.'s

An alternative class of distorted M. L. E.'s T^α, $\alpha \in (0, 1)$, is obtained when, instead of (5.1), one considers the functions

$$(6.1) \qquad f_\alpha(u) = \ln(\alpha + u), \quad f'_\alpha(u) = \frac{1}{\alpha + u}, \quad f''_\alpha(u) = \frac{1}{(\alpha + u)^2},$$

$$u \in [0, \infty], \quad \alpha \in [0, \infty).$$

Here again the functions $f_\alpha(u)$ as well as their derivatives are bounded from below so that T^α can be considered on domains $\mathscr{P}_0 = \mathscr{P}$ and their theory is even simpler than that presented in Section 5. In particular, the assumptions of Theorem 1.1 follow from A 1, A 2 so that the following theorem holds.

Theorem 6.1. If A 1, A 2 hold then all T^α, $\alpha \in (0, \infty)$, exist.

Analogically, Lemmas 5.1 and 5.2 remain true with

$$(6.2) \qquad l'_\alpha(\theta, P) = \mathsf{E}_P \frac{p'_\theta}{\alpha + p_\theta}, \quad -l''_\alpha(\theta, P) = \mathsf{E}_P \left[\frac{p'_\theta p'^{\mathrm{T}}_\theta}{(\alpha + p_\theta)^2} - \frac{p''_\theta}{\alpha + p_\theta} \right]$$

for $\theta \in \operatorname{int} \Theta$, $P \in \mathscr{P}$ (the arguments used in the proof transfer to the present situation without any modifications). Consequently, Theorem 5.1 remains true with $p^\alpha_\theta(x) \, . \, p'_\theta(x)/p_\theta(x)$ in (5.5) replaced by $p'_\theta(x)/(\alpha + p_\theta(x))$.

Lemma 5.3 still follows from the same argument as used in Section 5. However, T^α are only location and not location and scale equivariant (they are equivariant only if we consider the structural subfamily with $(\mu, \sigma) \in \mathbb{R} \times \{1\}$). For location families with the same parent densities as considered in Section 5 one can easily reformulate Theorem 5.2 and analyze the standard normal example. The same control of efficiency and robustness as with the type 1 distorted M. L. E.'s is possible. Details are omitted here.

(Received December 3, 1984.)

REFERENCES

[1] J. Anděl: Mathematical Statistics (in Czech). SNTL — ALFA, Praha—Bratislava 1978.

[2] T. W. Anderson: The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. Proc. Amer. Math. Soc. *6* (1955), 170—176.

[3] R. A. Fisher: On the mathematical foundations of theoretical statistics. Reprinted in: Contributions to Mathematical Statistics (by R. A. Fisher). J. Wiley, New York 1950.

[4] W. Fuller: Introduction to the Statistical Time Series. J. Wiley, New York 1976.

[5] F. R. Hampel: The influence curve and its role in robust estimation. J. Amer. Statist. Assoc. *69* (1974), 383—393.

[6] P. I. Huber: Robust estimation of a location parameter. Ann. Math. Statist. *35* (1964), 73—101.

[7] J. L. Kelley: General Topology. Van Nostrand, Princeton 1957.

[8] L. Le Cam: On the asymptotic theory of estimation and testing hypotheses. In: Proc. 3rd Berkeley Symp. Math. Statist. Prob. Vol. *1* (1956), 129—156.

[9] A. Perez: Notions generalisées d'incertitude, d'entropie et d'information du point de vue de la théorie des martingales. In: Trans. 1st Prague Conf. on Inform. Theory, etc. Publ. House Czechosl. Acad. Sci., Prague 1957.

[10] J. Pfanzagl: On the measurability and consistency of minimum contrast estimators. Metrika *14* (1969), 249—272.

[11] J. Pfanzagl: The second order optimality of tests and estimators for minimum contrast functionals. Probab. and Math. Statist. *2* (1981), 55—70.

[12] A. Rényi: Theory of probability (in Czech). Academia, Prague 1972.

[13] I. Vajda: Motivation, existence and equivariance of *D*-estimators. Kybernetika *20* (1984), 189—208.

[14] I. Vajda: Robust estimation in discrete and continuous families by means of a minimum chi-square method. Problems Control Inform. Theory *15* (1986), No. 2.

Ing. Igor Vajda, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Prague 8. Czechoslovakia.