

## STATISTICAL EXPANSIONS AND LOCALLY UNIFORM FRÉCHET DIFFERENTIABILITY

T. BEDNARSKI, B. R. CLARKE and W. KOLKIEWICZ

(Received 23 June 1989; revised 9 November 1989)

Communicated by Timothy C. Brown

### Abstract

Estimators which have locally uniform expansions are shown in this paper to be asymptotically equivalent to  $M$ -estimators. The  $M$ -functionals corresponding to these  $M$ -estimators are seen to be locally uniformly Fréchet differentiable. Other conditions for  $M$ -functionals to be locally uniformly Fréchet differentiable are given. An example of a commonly used estimator which is robust against outliers is given to illustrate that the locally uniform expansion need not be valid.

1980 *Mathematics subject classification* (*Amer. Math. Soc.*): primary 62 E 20; secondary 62 G 35.

*Keywords and phrases*: infinitesimal neighbourhoods, Fréchet differentiability, strong expansions,  $M$ -functionals.

### 1. Introduction

Robust inference for a statistical model denoted by  $\{F_\theta : \theta \in \Theta\}$  where  $\Theta$  is an open subset of  $R^k$ , is often formalized in the so called “infinitesimal set-up”. Here the true distribution  $F$  describing the underlying process is in a small vicinity of the model  $\{F_\theta\}$ ; the vicinity being described by the union of small neighbourhoods of all possible  $F_\theta$ , where  $\theta \in \Theta$ . For any given neighbourhood of a particular  $F_\theta$  it is assumed here that the size of the neighbourhood tends to zero at the rate  $1/\sqrt{n}$ , where  $n$  is the sample size. The infinitesimal approach constitutes a convenient and relatively universal tool for construction and verification of important asymptotic properties of estimates and tests.

Infinitesimal models are considered by Jaeckel (1971), Beran (1977), Huber-Carol (1970) and Rieder (1978), and a justification can be found in Bickel (1981) for their use. Bednarski (1985) illustrates the link to the influence curve of Hampel (1974), whence the infinitesimal approach can be considered to have broad application. Further rigor is given to the influence curve in proofs of Fréchet differentiability for classes of  $M$ -estimators in Clarke (1983, 1986), while authors Rieder (1980) and Bickel (1981) give infinitesimal approaches for  $M$ -estimators.

Let  $\mathcal{F}_{\theta, n}$  denote the shrinking neighbourhood about  $F_{\theta}$ , and let  $T_n$  be an estimator of  $\theta$ . Studying properties of  $T_n$ , one often assumes (Bickel (1981), Rieder (1980), Bednarski (1985)) that, for some function  $\psi$ ,

$$(1.1) \quad \sqrt{n}(T_n - \theta) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi(X_i, \theta) + o_{F_{\theta}^{\otimes n}}(1).$$

Here  $X_1, \dots, X_n$  is a sample from  $F_{\theta}$ . Asymptotic properties of  $T_n$  are derivable from expansion (1.1) and estimators having this property are called regular. The expansions hold automatically for product measures from  $\mathcal{F}_{\theta, n}^{\otimes n}$  which form sequences contiguous to  $\{F_{\theta}^{\otimes n}\}$ . Bickel and Rieder devote essential parts of their quoted papers to the analysis of regular estimators under contiguous departures from the basic model. Both authors however discuss briefly the question concerning the validity of (1.1) for more general departures (arbitrary sequences from  $\mathcal{F}_{\theta, n}^{\otimes n}$ ). In fact, the estimators which satisfy the stronger property form an important subclass of regular estimators, for which a complete characterization of asymptotic distributions under infinitesimal departures can be given. If in addition we require the expansion to be locally uniform, we are led to the following condition for  $T_n$ :

$$(1.2) \quad \sqrt{n}(T_n - \tau_n) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi(X_i, \tau_n) + o_{Q_n}(1)$$

for every sequence  $Q_n \in \mathcal{F}_{\tau_n, n}^{\otimes n}$ , where  $\tau_n$  converges to  $\theta$  so that  $\sqrt{n}(\tau_n - \theta)$  stays bounded.

Let us say that  $T_n$  satisfying (1.2) has a strong expansion at  $F_{\theta}$  with respect to the family of neighbourhoods. In Section 2 the family of infinitesimal neighbourhoods will be given by a metric and then we shall say that the strong expansion is with respect to the metric.

An important class of estimators for which we have general methods for constructing these asymptotic expansions is the class of  $M$ -estimators. The expansions are frequently obtained by differentiability of  $M$ -functionals, for example by making use of the Fréchet-differentiability of  $M$ -estimators as is presented concisely in Clarke (1983, 1986). Other basic references concern-

ing differentiability are Kallianpur (1963), Reeds (1976), Fernholz (1983), and Gill (1989).

The objective of this paper is to characterize relationships between differentiability and strong expansions. We show in Section 2 that if an estimator has a strong expansion (at  $F_\theta$  with respect to a metric  $d$ ), with a function  $\psi$ , then under additional mild conditions the  $M$ -functional related to  $\psi$  is locally uniformly Fréchet differentiable, meaning that

$$(1.3) \quad \left| T[G] - \tau - M^{-1} \int \psi(x, \tau) dG \right| = o[d(F_\theta, F_\tau) + d(F_\tau, G)]$$

for some matrix  $M$ . Use is made of the fixed point theorem.

For an  $M$ -estimator the strong expansion (1.2) is implied by (1.3) above. The converse statement need not be true for estimators more generally. In Section 3, conditions for local uniform Fréchet differentiability are derived. These conditions are implied by the assumptions of Clarke (1983, 1986), whence the “usual”  $M$ -estimators have strong expansions. Also examples are presented which are meant to help understand limitations of the main theorem.

### 2. Strong expansions and $M$ -estimators

The objective of this section is to show that any estimator having a strong expansion in a parametric model must be in an asymptotic linear relation with an  $M$ -estimator. Moreover the functional induced by this  $M$ -estimator will be locally uniformly Fréchet differentiable.

The principal result is given in Theorem 2.1, which makes use of Conditions C. Some preliminary notation will be adopted in what follows. Let  $\{F_\theta\}_{\theta \in \Theta}$  be a family of probability measures defined on the Borel  $\sigma$ -field of Euclidean space  $R^m$ , where  $\Theta$  is an open subset of  $R^k$ . Let  $d(\cdot, \cdot)$  denote a metric on the space of measures on  $R^m$ , which is denoted by  $G$  say. Let the parameter  $\theta$  be fixed. Consider the sequences of parameters and neighbourhoods of probabilities defined as follows:

$$J_\theta = \{ \{ \tau_n \} : \exists b > 0 \text{ such that } \sqrt{n} | \tau_n - \theta | \leq b \text{ for all } \tau_n \in \{ \tau_n \} \};$$
$$\mathcal{F}_{\tau, n}(\eta) = \{ Q \in G : \sqrt{n} d(Q, F_\tau) < \eta \}.$$

Here  $|\cdot|$  is the usual Euclidean norm. The functional  $T$  defined on  $G$  and the shrinking neighbourhoods are assumed to satisfy the following conditions.

*Conditions C.*

C1. There is a neighbourhood  $U_\theta$  of  $\theta$  and  $0 < k_1 < k_2$  such that

$$k_1 | \theta - \tau | \leq d(F_\theta, F_\tau) \leq k_2 | \theta - \tau | \quad \text{for all } \tau \in U_\theta.$$

C2. For every  $\{\tau_n\} \in J_\theta$ ,  $\eta > 0$  and  $G_n \in \mathcal{F}_{\tau_n, n}(\eta)$ , we have that

$$\sqrt{n} \left| T[F_n] - \tau_n - \int \psi(x, \tau_n) dF_n \right|$$

converges to zero in  $G_n^{\otimes n}$  probability, where  $F_n$  is the empirical distribution function based on the sample of size  $n$  and  $\psi(\cdot, \tau)$  is a measurable function.

C3. The function  $\psi$  from C2 is bounded.

C4. For every  $\{\tau_n\} \in J_\theta$ , every  $\eta > 0$  and  $\{G_n\}$  such that  $G_n \in \mathcal{F}_{\tau_n, n}(\eta)$ ,  $n \geq 1$ ,

$$\sqrt{n}(T[F_n] - \theta)$$

is relatively compact in  $G_n^{\otimes n}$ .

C5. With  $K_G(\tau)$  defined equal to  $\int \psi(x, \tau) dG(x)$ , we suppose  $K_{F_\theta}(\tau)$  is differentiable at  $\tau = \theta$  and has nonsingular derivative  $M$  at  $\tau = \theta$ .

C6. For every sequence of  $G_n \in \mathcal{F}_{\tau_n, n}(\eta)$  there is a neighbourhood  $U_\theta$  of  $\theta$  such that for every  $n$ ,  $K_{G_n}(\tau)$  are continuous functions on  $U_\theta$ .

These conditions are associated with the usual estimating functionals involved in robustness studies and which satisfy the strong expansion as is stated in C2. Condition C1 is satisfied by regular parametric models under suitable choices of metric  $d$ . The condition C4 denotes the square root consistency of  $T$  under infinitesimal departures from the parametric model.

It can be remarked that the conditions imply the usual Fisher consistency requirement that  $K_{F_\theta}(\theta) = 0$ , since by condition C2 and C4 the  $\int \psi(x, \theta) dF_n$  converges in  $F_\theta^{\otimes n}$  probability to zero, which is also  $K_{F_\theta}(\theta)$  by the law of large numbers.

In the following discussion  $T^*[G]$  is denoted to be the  $M$ -estimator defined as the solution of  $\int \psi(x, \tau) dG(x) = 0$  which is closest to  $\theta$ .

**THEOREM 2.1.** *Under conditions C1 to C6,  $T^*$  is locally uniformly Fréchet differentiable, that is,*

$$|T^*[G] - \tau - M^{-1}K_G(\tau)| = o(d[G, F_\tau] + d[F_\tau, F_\theta]).$$

The proof of Theorem 2.1 will be preceded by several lemmas, given below.

**LEMMA 2.1.** *For every  $\eta > 0$  and every sequence  $\{G_n\}$  such that  $G_n \in \mathcal{F}_{\theta, n}(\eta)$ , we have that*

$$\sqrt{n}|K_{G_n}(\theta) - K_{F_\theta}(\theta)|$$

*is bounded.*

PROOF. By Condition C2,

$$\sqrt{n} \left| T[F_n] - \theta - \int \psi(x, \theta) dF_n \right| \xrightarrow{\text{in } G_n^{\otimes n}} 0.$$

We obtain, on rewriting this quantity,

$$\left| \sqrt{n}(T[F_n] - \theta) - \frac{1}{\sqrt{n}} \sum_{i=1}^n [\psi(X_i, \theta) - K_{G_n}(\theta)] - \sqrt{n}K_{G_n}(\theta) \right|.$$

Condition C4 gives us that the first term is bounded. For the second term it suffices to consider the univariate case of  $\psi$ . There are two possibilities. Either  $\text{Var}_{G_n}[\psi(X, \theta)]$  is bounded away from zero or there are subsequences  $G_n$  for which  $\text{Var}_{G_n}[\psi(X, \theta)]$  tends to zero. In the former case an application of the Lindenberg-Feller theorem implies convergence to a normal law in which case the second term is bounded. In the latter case it is clear that the second term tends to zero in probability. Thus the result follows since  $K_{F_\theta}(\theta) = 0$ . A multivariate result is easily obtained by using a method analogous to the proof of the central limit theorem by Cramér and Wold (1936).

LEMMA 2.2. For every sequence  $\{G_n\}$  such that  $G_n \in \mathcal{F}_{\theta, n}(\eta)$  and every  $\{\tau_n^1\}, \{\tau_n^2\} \in J_\theta$ , we have

$$\sqrt{n} |(K_{G_n}(\tau_n^1) - K_{G_n}(\tau_n^2)) - (K_{F_\theta}(\tau_n^1) - K_{F_\theta}(\tau_n^2))| \xrightarrow{n \rightarrow \infty} 0.$$

PROOF. By assumption C2 it follows that

$$\sqrt{n} \left| \tau_n^1 - \tau_n^2 + \int \psi(x, \tau_n^1) dF_n - \int \psi(x, \tau_n^2) dF_n \right|$$

converges to zero in  $G_n^{\otimes n}$ . The above expression can be rewritten as

$$\left| \sqrt{n}(\tau_n^1 - \tau_n^2) + \frac{1}{\sqrt{n}} \sum_{i=1}^n [\psi(X_i, \tau_n^1) - \psi(X_i, \tau_n^2) - (K_{G_n}(\tau_n^1) - K_{G_n}(\tau_n^2))] + \sqrt{n}(K_{G_n}(\tau_n^1) - K_{G_n}(\tau_n^2)) \right|.$$

The second term involving the summation has independent identically distributed summands with expectation zero. Hence, using similar reasoning as in Lemma 2.1, we see that the normed summation tends either to zero in probability or to a limiting normal distribution. If it tends to a limiting normal distribution then it can be noted that if  $Y_n \xrightarrow{d} N(0, \sigma)$  and for some

sequence of numbers  $C_n$  it is true that  $|Y_n + C_n| \rightarrow 0$  in  $G^{\otimes n}$ , then  $\sigma = 0$  and  $C_n \rightarrow 0$ . Consequently it follows in either case that

$$|\sqrt{n}(\tau_n^1 - \tau_n^2) + \sqrt{n}(K_{G_n}(\tau_n^1) - K_{G_n}(\tau_n^2))| \rightarrow 0.$$

Making the same argument but replacing  $G_n$  by  $F_\theta$  gives by difference the result of the lemma.

LEMMA 2.3. For every  $\eta > 0$  and  $G_n \in \mathcal{F}_{\theta, n}(\eta)$ ,  $\sqrt{n}|T^*[G_n] - \theta|$  is bounded.

PROOF. By Lemma 2.1, it follows that there is an  $\alpha > 0$  such that

$$\sqrt{n}|K_{G_n}(\theta) - K_{F_\theta}(\theta)| < \alpha.$$

Let  $\varepsilon > 0$  be given. For every  $\{\tau_n\}$  such that  $\sqrt{n}|\tau_n - \theta| < \eta$ ,

$$\sqrt{n}|K_{G_n}(\tau_n) - K_{G_n}(\theta) - (K_{F_\theta}(\tau_n) - K_{F_\theta}(\theta))| < \frac{\varepsilon}{2}$$

holds for  $n \geq n_0$ . Consequently it follows that

$$\sqrt{n}|K_{G_n}(\tau_n) - K_{F_\theta}(\tau_n)| < \alpha + \frac{\varepsilon}{2}.$$

By differentiability of  $K_{F_\theta}(\tau)$  at  $\tau = \theta$ , we have

$$K_{F_\theta}(\tau) = M(\tau - \theta) + o(1/\sqrt{n})$$

when  $\sqrt{n}|\tau - \theta| < \eta$ . So

$$\sqrt{n}|K_{G_n}(\tau) - M(\tau - \theta)| < \alpha + \varepsilon \quad \text{if } \sqrt{n}|\tau - \theta| < \eta.$$

This is rewritten as

$$\sqrt{n}|K_{G_n}(\tau + \theta) - M\tau| < \alpha + \varepsilon \quad \text{if } \sqrt{n}|\tau| < \eta,$$

whence it follows that

$$\sqrt{n}|M^{-1}K_{G_n}(\theta + \tau) - \tau| \leq |M^{-1}|(\alpha + \varepsilon)$$

for  $n$  large enough and for  $\sqrt{n}|\tau - \theta| \leq |M^{-1}|(\alpha + \varepsilon)$  (since  $\eta$  can be chosen). Hence for  $n$  large enough,  $-M^{-1}K_{G_n}(\theta + \tau) + \tau$  transforms the ball  $|\tau| \leq (|M^{-1}|(\alpha + \varepsilon))/\sqrt{n}$  into the same ball and moreover the transformation is continuous. By the fixed point theorem there is a  $\tau^*$  such that  $|\tau^*| \leq |M^{-1}|(\alpha + \varepsilon)/\sqrt{n}$  and  $-M^{-1}K_{G_n}(\theta + \tau^*) + \tau^* = \tau^*$ , which gives us that  $-M^{-1}K_{G_n}(\theta + \tau^*) = 0$  implies  $K_{G_n}(\theta + \tau^*) = 0$ . Thus  $T^*[G_n] = \theta + \tau^*$  and the result is proved.

**PROOF OF THE THEOREM.** Consider the sequences of  $G_n$  and  $\tau_n$  so that  $C_n = d(G_n, F_{\tau_n}) + d(F_{\tau_n}, F_\theta)$  satisfy  $0 < \lim \sqrt{n}C_n < \infty$ . Since  $\sqrt{n}|T^*[G_n] - \theta|$  is bounded, by Lemma 2.3, it follows from Lemma 2.2 and C5 that

$$\begin{aligned} K_{G_n}(T^*[G_n]) - K_{G_n}(\tau_n) &= K_{F_\theta}(T^*[G_n]) - K_{F_\theta}(\tau_n) + o(1/\sqrt{n}) \\ &= K_{F_\theta}(T^*[G_n]) - K_{F_\theta}(\theta) + K_{F_\theta}(\theta) - K_{F_\theta}(\tau_n) + o(1/\sqrt{n}) \\ &= M[T^*[G_n] - \theta] + M[\theta - \tau_n] + o(1/\sqrt{n}) \\ &= M[T^*[G_n] - \tau_n] + o(1/\sqrt{n}). \end{aligned}$$

Therefore  $\sqrt{n}|T^*[G_n] - \tau_n + M^{-1}K_{G_n}(\tau_n)|$  must converge to zero. Thus the result follows.

### 3. Other conditions for locally uniform Fréchet differentiability and examples

By considering the parametric family  $\{F_\theta\}_{\theta \in \Theta}$ , where as before

$$k_1|\theta - \tau| < d(F_\tau, F_\theta) < k_2|\tau - \theta|$$

for some  $0 < k_1 < k_2$  and all  $\tau$  in a neighbourhood of  $\theta$ , a set of conditions on  $K_G(\tau) = \int \psi(x, \tau)dG$ , giving locally uniform Fréchet differentiability, can be specified.

*Conditions B.*

**B<sub>1</sub>.**  $K_{F_\theta}(\theta) = 0$ .  $K_{F_\theta}(\tau)$  is differentiable at  $\tau = \theta$  and the derivative  $M$  is nonsingular.

**B<sub>2</sub>.** For every sequence of probability measures  $\{G_n\}$  such that  $G_n \in \mathcal{F}_{\theta,n}(\eta)$ , for some  $\eta > 0$ , there is a neighbourhood  $U_\theta$  of  $\theta$  for which  $K_{G_n}(\tau)$  is continuous for every  $n$ .

**B<sub>3</sub>.** For every sequence  $\{G_n\}$ ,  $G_n \in \mathcal{F}_{\theta,n}(\eta)$  we have  $\sqrt{n}K_{G_n}(\theta)$  is bounded.

**B<sub>4</sub>.** For every sequence  $\{G_n\}$ ,  $G_n \in \mathcal{F}_{\theta,n}(\eta)$  and  $\{\tau_n^1\}, \{\tau_n^2\} \in J_\theta$  we have

$$\sqrt{n} |(K_{G_n}(\tau_n^1) - K_{G_n}(\tau_n^2)) - (K_{F_\theta}(\tau_n^1) - K_{F_\theta}(\tau_n^2))| \xrightarrow{n \rightarrow \infty} 0.$$

These conditions hold true under conditions presented in Clarke (1986). Condition **B<sub>1</sub>** corresponds to  $A'_0, A'_3$  and a general assumption of Section 4 of that paper, while **B<sub>2</sub>** follows from  $A'_1$ . Condition **B<sub>3</sub>** follows from assumption (4.1) in Clarke (1986). Finally one can show that **B<sub>4</sub>** is obtained from assumption  $A'_4$  with the aid of upper semicontinuity of the generalized Jacobian.

If  $T^*[G]$  denotes the  $M$ -functional corresponding to  $\psi$  it can be inferred as in Section 3 that  $\sqrt{n}|T^*[G_n] - \theta|$  is bounded for sequences  $G_n \in \mathcal{F}_{\theta, n}(\eta)$ . Arguing as in the proof of Theorem 2.1, we obtain the following result.

**THEOREM 3.1.** *Assume Conditions B hold. Then the  $M$ -functional  $T^*$  is locally uniformly Fréchet differentiable. That is,*

$$|T^*[G] - \tau - M^{-1}K_G(\tau)| = o[d(G, F_\tau) + d(F_\tau, F_\theta)].$$

**COROLLARY 3.1.** *Under assumptions B, using the Kolmogorov metric, by the inequality of Dvoretzky, Kiefer and Wolfowitz (cf. Serfling (1980)), we have*

$$\sqrt{n}(T^*[F_n] - \tau_n) = \frac{1}{\sqrt{n}}M^{-1} \sum_{i=1}^n \psi(X_i, \theta) + o_{G_n^{\otimes n}}(1),$$

where  $X_1, \dots, X_n$  is a sample from  $G_n \in \mathcal{F}_{\theta, n}(\eta)$ .

There are many examples of  $M$ -estimators which satisfy Conditions B. They can be found in Clarke (1986), where in particular Huber's minimax  $M$ -estimator for location is analysed. Therefore these  $M$ -estimators have the strong expansion.

Even though the main result of the paper is meant to indicate that there is a close relationship between strong expansions and Fréchet differentiability, there are examples which illustrate that the strong expansion must be supported by various non-stochastic assumptions, to imply Fréchet differentiability.

It appears that the functional  $T$  given in assumption C2 need not be Fréchet differentiable. Suppose that  $T$  is an  $M$ -functional which satisfies conditions C1 to C6 for the Kolmogorov distance. Such functionals exist as noted above. Moreover, suppose  $F_\theta$  is a distribution on the real line which is absolutely continuous with respect to the Lebesgue measure.

Let  $\{D_n\}$  denote a sequence of absolutely continuous distributions such that the sequence  $\sqrt{n}(d_k(D_n, F_\theta))$  is bounded. Then a new functional  $T' = T + I_{\{D_n\}}$ , where  $I_{\{D_n\}}$  denotes the indication function, satisfies C1 to C6. Since  $T'$  is not continuous at  $F_\theta$  it cannot be differentiable there.

It is perhaps important to realize that some well known estimators which are robust in the sense that they are resistant to large outliers need not at the same time have a strong expansion. The median is well known not to be Fréchet differentiable (cf. Fernholz, (1983)). It does not seem however so well known that the expansion admitted by the empirical median does not hold uniformly under shrinking neighbourhoods. Since the functional induced by the median does not satisfy Condition C6 our main theorem does not imply

that there is no strong expansion. We shall validate this fact below, showing then that assumptions of our main theorem can perhaps be weakened.

Here we define the median to be  $F_n^{-1}(\frac{1}{2})$  where  $F^{-1}(\alpha) = \inf\{t : F(t) \geq \alpha\}$ . The influence curve for the median at the normal distribution  $\Phi$  is given by a discontinuous function

$$\psi(x) = \{2\phi(\Phi^{-1}(\frac{1}{2}))\}^{-1} \text{sign}(x - \Phi^{-1}(\frac{1}{2})),$$

where  $\phi(x)$  is the standard normal density. The expansion

$$\sqrt{n}F_n^{-1}(\frac{1}{2}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \psi(X_i) + o_{\Phi^{\otimes n}}(1)$$

is true at the normal distribution, whence the asymptotic normal distribution of  $N(0, 1/\{4\phi(0)^2\})$  is achieved. The same cannot be said in infinitesimal departures of  $\Phi^{\otimes n}$ . The following result follows in a straightforward way from the Malmquist (1950) representation of the distribution of uniform order statistics by exponential random variables.

**LEMMA 3.1.** *Let  $X_1, \dots, X_n$  be independent identically distributed variables from*

$$G_n = (1 - \varepsilon/\sqrt{n})\Phi + (\varepsilon/\sqrt{n})\delta_0,$$

where  $\delta_0$  places atomic mass one at the origin and  $0 < \varepsilon < 1$ . Then asymptotically  $\sqrt{n}F_n^{-1}(\frac{1}{2})$  has a distribution that attributes positive mass at the origin, while  $(1/\sqrt{n}) \sum_{i=1}^n \psi(X_i)$  tends in distribution to a normal variable.

### Acknowledgement

We thank the referee for comments which helped improve the final version of the paper.

### References

- T. Bednarski (1985), 'On minimum bias and variance estimation for parametric models with shrinking contamination', *Probab. Math. Statist.* **6**, 121–129.
- R. Beran (1977), 'Minimum Hellinger distance estimates for parametric models', *Ann. Statist.* **5**, 445–463.
- P. J. Bickel (1981), 'Quelques aspects de la statistique robuste', *Lecture Notes in Mathematics* pp. 1–72, (Ecole d'Été de Probabilités de Saint-Fleur IX, 1971).

- B. R. Clarke (1983), 'Uniqueness and Fréchet differentiability of functional solutions to maximum likelihood type equations', *Ann. Statist.* **11**, 1196–1205.
- B. R. Clarke (1986), 'Nonsmooth analysis and Fréchet differentiability of  $M$ -functionals', *Probab. Th. Rel. Fields* **73**, 197–209.
- H. Cramér and H. Wold (1936), 'Some theorems on distribution functions', *J. London Math. Soc.* **11**, 290–295.
- L. T. Fernholz (1983), *Von Mises calculus for statistical functionals*, (Lecture Notes in Statistics 19, Springer Berlin, Heidelberg, New York).
- R. D. Gill (1989), 'Non- and semi-parametric maximum likelihood estimators and the von Mises method (Part 1)', *Scand. J. Statist.* **16**, 97–128.
- F. R. Hampel (1974), 'The influence curve and its role in robust estimation', *J. Amer. Statist. Assoc.* **62**, 1179–1186.
- C. Huber-Carol (1970), *Etude asymptotique de tests robustes*, (These a l'Ecole Polytechnique Federale de Zürich).
- L. A. Jaeckel (1971), 'Robust estimates of location: symmetric and asymmetric contamination', *Ann. Math. Stat.* **42**, 1020–1035.
- G. Kallianpur (1963), 'Von Mises functions and maximum likelihood estimation', *Sankhya Ser. A* **23**, 149–158.
- S. Malmquist (1950), 'On a property of order statistics from a rectangular distribution', *Skand. Akt.* **33**, 214–222.
- J. A. Reeds (1976), *On the definition of Von Mises functions*, (Ph.D. Thesis, Dept. of Statistics, Harvard Univ., Cambridge).
- H. Rieder (1978), 'A robust asymptotic testing model', *Ann. Statist.* **6**, 1080–1094.
- H. Rieder (1980), 'Estimators derived from robust tests', *Ann. Statist.* **8**, 106–115.
- R. J. Serfling (1980), *Approximation theorems of mathematical statistics*, (Wiley, New York).

Institute of Mathematics  
Polish Academy of Sciences  
Warsaw  
Poland

Institute of Mathematics  
Polish Academy of Sciences  
Warsaw  
Poland

Murdoch University  
Murdoch, WA 6150  
Australia