Limit Behavior of the Empirical Influence Function of the Median

Christophe Croux *

Université Libre de Bruxelles

Abstract: The empirical influence function $\mathrm{EIF}(x, T_n; X)$ measures the influence of an observation x on the estimator T_n at a sample X of size n. In this note we show that the empirical influence function of the median is not a consistent estimator of the corresponding influence function. This observation leads to a reconsideration of the most B-robustness property of the median. We will prove and show by simulations that the median is less robust to single outliers than commonly believed.

Key words: Breakdown point, Influence function, Location estimator, Robustness, Sensitivity.

AMS subject classifications: 62F35, 62G35

*ECARE and Institut de Statistique, Université Libre de Bruxelles, CP 139, Av. F.D. Roosevelt 50, B-1050 Bruxelles, Belgium

1 Introduction

In the robustness literature, various measures of robustness of an estimator T_n have been proposed (Hampel 1974, Huber 1981, Hampel et al 1986): influence function, change-of-variance curve, breakdown point, maxbias curve,.... In this note we will focus on the influence function of an estimator. This requires the existence of a functional version T of the estimator such that $T(F_n) = T_n$, where F_n denotes the empirical distribution function of the sample $X = \{x_1, \ldots, x_n\}$. The influence function of an estimator T at a fixed distribution F is then defined as

$$\mathrm{IF}(x, T; F) = \lim_{\varepsilon \downarrow 0} \frac{T((1-\varepsilon)F + \varepsilon\Delta_x) - T(F)}{\varepsilon}, \tag{1.1}$$

where Δ_x is a distribution which puts all its mass at the point x. A finite-sample version of the influence function is obtained by suppressing the limit in (1.1) and replacing ε by $1/(n+1)$ and F by F_n. This yields the *empirical influence function* (sometimes also called *sensitivity function*, Hampel et al 1986, page 93)

$$\mathrm{EIF}(x, T_n; X) = (n+1)\{T_{n+1}(x_1, \ldots, x_n, x) - T_n(x_1, \ldots, x_n)\}. \tag{1.2}$$

Since the EIF depends on the observed sample $X = \{x_1, \ldots, x_n\}$, where each observation x_i follows the distribution F, we need to consider it as a random variable. [1] One can see $\mathrm{EIF}(x, T_n; X)$ as the (standardized) influence of an observation x on the estimator T_n at the sample X. A high value of $\mathrm{EIF}(x, T_n; X)$ means that a potential outlier at x will have a large influence on the estimator, indicating non-robustness. This interpretation carries over to the functional based $\mathrm{IF}(x, T; F)$ since for most estimators

$$\lim_{n\to\infty} \mathrm{EIF}(x, T_n; X) = \mathrm{IF}(x, T; F) \qquad a.s. \tag{1.3}$$

However, if the above equality does not hold, one should be more careful with relating the form of the influence function to the finite-sample behavior of the estimator, even for large values of n .

In Section 2 we proof that (1.3) holds for α-trimmed means (with $\alpha < 0.5$) and for smooth M-estimators. However, (1.3) does *not* hold for the median. We show in Section 3 that the median has no longer the smallest gross-error sensitivity when we use a "sample-based" instead of a "functional-based" definition of the gross-error sensitivity. In Section

[1] Sometimes X is taken to be a stylized sample (Andrews et al. 1972, page 96) from F, eliminating the random character of $\mathrm{EIF}(x, T_n; X)$. Unfortunately, stylized samples never occur in practice.

4 a simulation study shows that the resistance of the sample median to single outlying observations at finite samples corresponds well with our theoretical findings. Some discussion is presented in the last Section.

2 Convergence of the Empirical Influence Function

The univariate sample median is defined as

$$\mathrm{med}_n(X) = \frac{x_{(\lfloor n/2 \rfloor + 1)} + x_{(\lfloor (n+1)/2 \rfloor)}}{2},$$

where $x_{(1)} \leq \ldots \leq x_{(n)}$ are the ordered observations from X. The corresponding functional is defined for every distribution G and equals

$$\mathrm{med}(G) = \frac{\inf\{t|G(t) > 1/2\} + \sup\{t|G(t) < 1/2\}}{2}.$$

It is known (Hampel et al 1986, page 109) and easy to verify that the influence function of the median at a distribution F with positive density f at its median is given by

$$\mathrm{IF}(x, \mathrm{med}; F) = \begin{cases} -1/(2f(\mathrm{med}(F))) & \text{for } x < \mathrm{med}(F) \\ 0 & \text{for } x = \mathrm{med}(F) \\ 1/(2f(\mathrm{med}(F))) & \text{for } x > \mathrm{med}(F). \end{cases} \tag{2.1}$$

To avoid too much technical detail, we assume throughout the paper:

(F) The distribution F has a continuous, strictly positive density f, is unimodal and symmetric around its median $\mathrm{med}(F)$.

The following proposition shows that $\mathrm{EIF}(x, \mathrm{med}_n; X)$ is not a consistent estimator for $\mathrm{IF}(x, \mathrm{med}; F)$. The empirical influence function evaluated in x converges in distribution to an exponential distribution with mean $\mathrm{IF}(x, \mathrm{med}; F)$ for $x > 0$. For $x < 0$ we obtain a negative exponential and for $x = 0$ a double exponential limit distribution. All proofs can be found in the Appendix.

Proposition 1. *The empirical influence function of the median EIF($x, med_n; X$), where each observation from $X = \{x_1, \ldots, x_n\}$ is distributed according to F, converges in law to a distribution with density*

$$\begin{cases} 2f(med(F)) \exp\left(2uf(med(F))\right) I(u \leq 0) & \text{if } x < med(F) \\ f(med(F)) \exp\left(-2|u|f(med(F))\right) & \text{if } x = med(F) \\ 2f(med(F)) \exp\left(-2uf(med(F))\right) I(u \geq 0) & \text{if } x > med(F). \end{cases} \tag{2.2}$$

We will compare this limit behavior of the empirical influence function with that of trimmed means and M-estimators. The α-trimmed mean T_n^α (with $0 < \alpha < 1/2$) is defined as

$$T_n^\alpha(X) = \frac{1}{n - 2\lfloor \alpha n \rfloor} \sum_{i=\lfloor \alpha n \rfloor + 1}^{n - \lfloor \alpha n \rfloor} x_{(i)},$$

where $x_{(1)} \leq \ldots \leq x_{(n)}$ are the ordered observations. The influence function of this estimator is given by (see Huber 1981, page 58)

$$\begin{aligned} \mathrm{IF}(x, T^\alpha; F) &= (F^{-1}(\alpha) - \mathrm{med}(F))/(1 - 2\alpha) && \text{for } x < F^{-1}(\alpha) \\ &= (x - \mathrm{med}(F))/(1 - 2\alpha) && \text{for } F^{-1}(\alpha) \leq x \leq F^{-1}(1-\alpha) \\ &= (F^{-1}(1-\alpha) - \mathrm{med}(F))/(1 - 2\alpha) && \text{for } F^{-1}(1-\alpha) < x. \end{aligned} \tag{2.3}$$

The following proposition shows that the EIF of the α-trimmed mean converges uniformly and almost surely to the corresponding influence function.

Proposition 2. *For any distribution F satisfying condition (F), and for any $0 < \alpha < 1/2$ we have*

$$\lim_{n \to \infty} \sup_x |EIF(x, T_n^\alpha; X) - IF(x, T^\alpha; F)| = 0 \quad a.s.$$

Other popular robust estimators of location are M-estimators $T_{n,\psi}$, which are defined as the solution of the following equation in t:

$$\sum_{i=1}^n \psi(x_i - t) = 0. \tag{2.4}$$

We suppose that

(Ψ) The function ψ is odd, non-decreasing, bounded, continuous, and almost everywhere differentiable with $\psi'(0) > 0$.

These conditions exclude the function $\psi(u) = \mathrm{sign}(u)$, which corresponds to the median estimator. Our main example will be the Huber M-estimator of location defined by $\psi_b(u) = \max(-b, \min(u, b))$ where b is a positive tuning parameter. The influence function of an M-estimator T_ψ at the model distribution F is given by (Hampel et al 1986, page 103)

$$\mathrm{IF}(x, T_\psi; F) = \frac{\psi(x - \mathrm{med}(F))}{E_F[\psi'(X - \mathrm{med}(F))]}. \tag{2.5}$$

Comparing (2.3) and (2.5) shows that the influence function of the Huber M-estimator equals the IF of the trimmed mean for $b = F^{-1}(1-\alpha) - \mathrm{med}(F)$. Just like for trimmed means, but opposed to the median, we obtain almost sure convergence for the EIF of (smooth) M-estimators.

Proposition 3. *If the distribution F satisfies (F) and the M-estimator T_ψ satisfies (Ψ) then*

$$\lim_{n\to\infty} \sup_x |EIF(x, T_{n,\psi}; X) - IF(x, T_\psi; F)| = 0 \quad a.s.$$

3 Reconsidering the Gross-error Sensitivity

The definition of the gross-error sensitivity of T at the distribution F uses the functional version of the estimator and is given by

$$\gamma^*(T, F) = \sup_x |IF(x, T; F)|. \tag{3.1}$$

It is often interpreted as the maximum (standardized) influence that one single outlier can have on the estimator when the "good" data come from a distribution F. However, this interpretation corresponds merely with the sample-based quantity

$$\gamma(T_n, X) = \sup_x |\mathrm{EIF}(x, T_n; X)|, \tag{3.2}$$

where $X \sim F$, which is a stochastic variable. For α-trimmed means and M-estimators Propositions 2 and 3 yield that $\gamma(T_n^\alpha, X)$ and $\gamma(T_{n,\psi}, X)$ converge almost surely to $\gamma^*(T^\alpha, F)$ and $\gamma^*(T_\psi, F)$. For the median however, the following proposition shows that the variable $\gamma(\mathrm{med}_n, X)$ lacks consistency.

Proposition 4. *For every distribution satisfying (F), we have*

$$\mathit{limsup}_n\, P_F(\gamma(\mathit{med}_n, X) \le u) = (1 - \exp(-2uf(\mathit{med}(F))))^2 \quad \mathit{for}\ \ u \ge 0. \tag{3.3}$$

A sample-based definition of gross-error sensitivity could be given by

$$\operatorname*{limsup}_n \mathrm{Loss}(\gamma(T_n, X)) \tag{3.4}$$

where the loss function measures the deviation of the random variable $\gamma(T_n, X)$ from the target zero. In Table 1 we compare the sample-based gross-error sensitivity (3.4) of the median with the 10%-, 25%- and 45%-trimmed mean. We used the mean, median, expected squared value, and 95%-quantile of the distribution of $\gamma(T_n, X)$ to summarize the loss of robustness. The distribution F was taken to be the standard normal. We see that the sample-based measures of gross-error sensitivity for 25%- and 45%-trimmed means are smaller than those for the median, and this for all loss functions considered here. Even the 10%-trimmed mean

Table 1: Sample-based gross-error sensitivity $\limsup_n \text{Loss}(\gamma(T_n, X))$, where X follows the standard normal, for the median and several trimmed means .

Estimator		Loss function $L(Y)$			
		$E\|Y\|$	$\text{med}\|Y\|$	EY^2	$F_Y^{-1}(0.95)$
	median	1.880	1.539	5.498	4.607
T^α	$\alpha = 0.10$	1.602	1.602	2.566	1.602
	$\alpha = 0.25$	1.349	1.349	1.820	1.349
	$\alpha = 0.45$	1.257	1.257	1.579	1.257
	$\alpha \uparrow 0.50$	1.253	1.253	1.571	1.253

performs quite well w.r.t. the median. These observations lead us to reconsider the most B-robustness property of the median. This property (Hampel et al 1986, page 133; Rousseeuw 1982) says that the median minimizes $\gamma^*(T, F)$ within the class of translation equivariant location estimators. Working with a sample-based measure of gross-error sensitivity reveals however that this doesn't imply that the median is also most robust with respect to single outliers. If we focus on the mean absolute deviation as loss function in (3.4), we obtain a new measure of robustness:

$$\tilde{\gamma}(T, F) = \limsup_n E[|\gamma(T_n, X)|]. \tag{3.5}$$

As we can see from the first column of Table 1, there are many other estimators having a lower $\tilde{\gamma}(T, F)$ than the median.

A strange discontinuity is observed from the last line of Table 1 and can be derived from equations (2.1) and (2.3).

$$\lim_{\alpha \uparrow 0.5} \tilde{\gamma}(T^\alpha, F) = 1/(2f(\text{med}(F))) = \gamma^*(\text{med}, F) < \tilde{\gamma}(\text{med}, F).$$

We come close to the lowest possible value for $\tilde{\gamma}(T^\alpha, F)$ when the trimming proportion tends to 50%. For the limiting estimator (the median) however, $\tilde{\gamma}(T^{0.5}, F)$ lifts up. One should not forget that although $\tilde{\gamma}$ is a sample-based measure, it is still an asymptotic measure. At finite samples the median behaves essentially like a trimmed mean with α almost equal to 1/2. The question which arises now is whether the first or the last row of Table 1 describes better the distribution of $\gamma(\text{med}_n, X)$ for finite values of n. In the next Section it is shown by means of

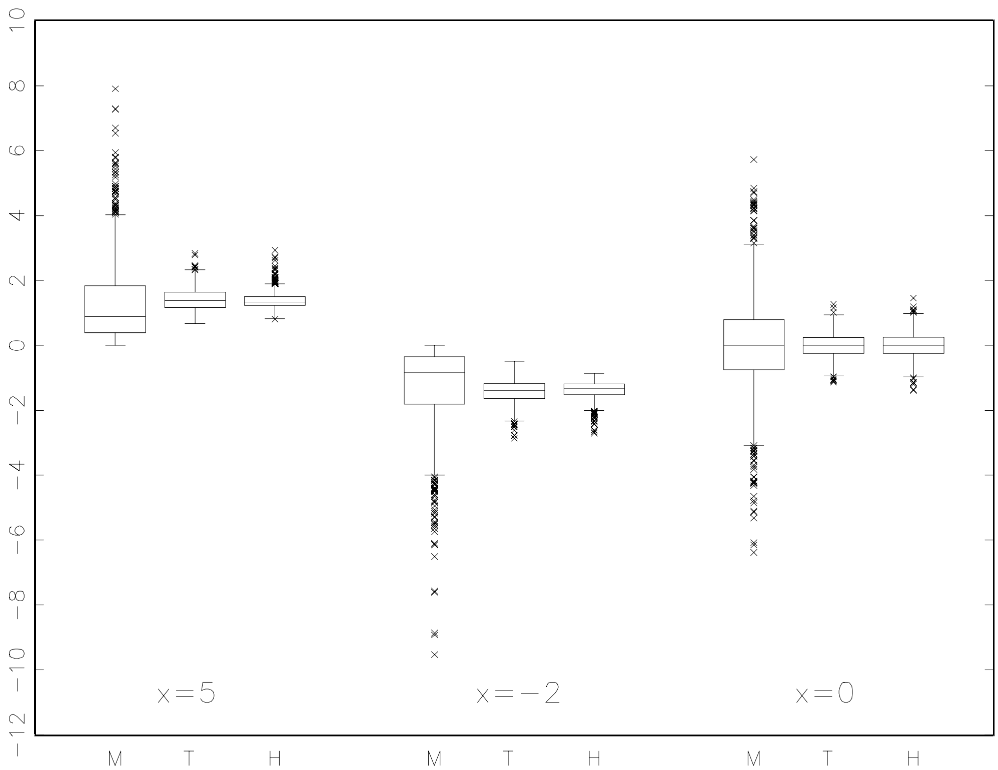

Figure 1: Boxplots of 1000 numbers generated from EIF$(x, T_n; X)$, where X follows a standard normal distribution and $n = 30$. We considered the median (M), the 25%-trimmed mean (T), and the corresponding Huber M-estimator (H) for $x = 5, -2,$ and 0.

a simulation study that the asymptotic results obtained for the median estimator correspond well with its finite-sample behavior. The rather large value of $\tilde{\gamma}(\text{med}, F)$ is therefore not a pure theoretical phenomenon, it is also observed in practice.

4 Simulations

We performed a simulation study to confirm our theoretical findings. First we generated 1000 samples X of size $n = 30$ from a standard normal distribution Φ and computed $\text{EIF}(x, T_n; X)$ for each of these samples. Boxplots of these 1000 numbers are pictured in Figure 1 for the median, the 25%-trimmed mean, and the Huber M-estimator with $b = \Phi^{-1}(0.75)$. We repeated this for three values of x. The exponential form of the limit distributions of $\text{EIF}(x, \text{med}_n; X)$ appears clearly. The distributions of the EIF for the trimmed mean and the Huber M-estimator are much less dispersed and have only a slightly different median.

A similar simulation was done to investigate the distribution of $\gamma(T_n, X)$ at finite samples.

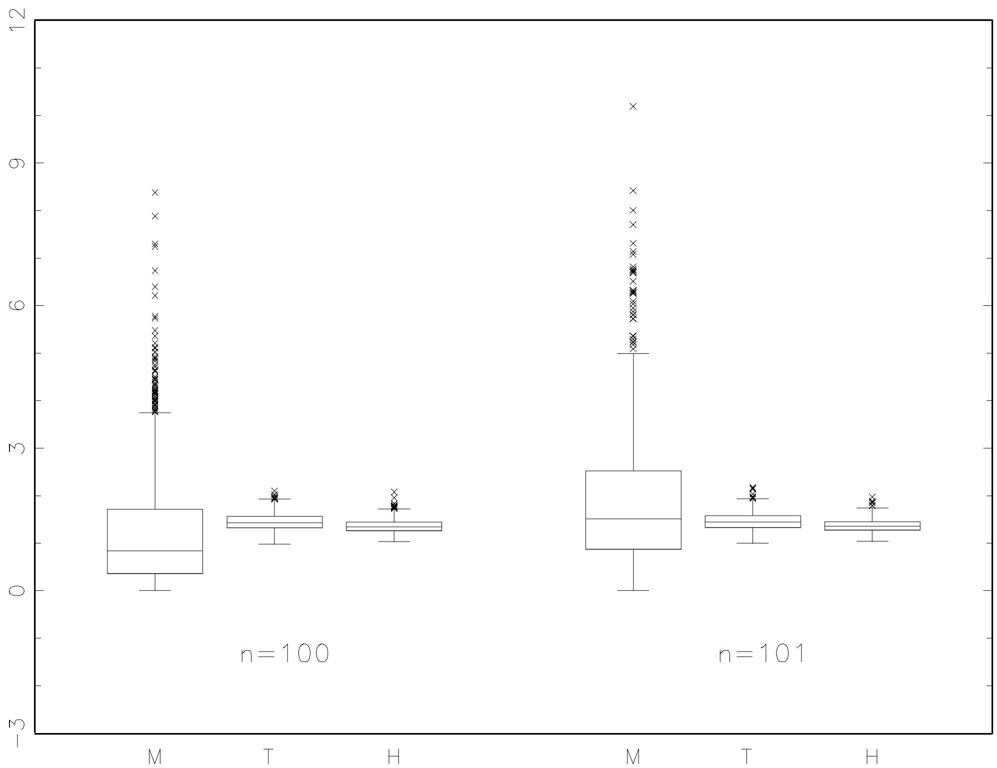

Figure 2: Boxplots of 1000 numbers generated from $\gamma(T_n, X)$, where X follows a standard normal distribution and $n = 100$ and 101. The considered estimators are the median (M), the 25%-trimmed mean (T), and the corresponding Huber M-estimator (H).

For the odd sample size $n = 101$ we see from Figure 2 that boxplots for the trimmed mean and the Huber M-estimator are less dispersed and have smaller medians than the boxplot for the sample median. In this case, we may conclude that the sample median is less resistant to single outliers than the other two estimators. Notice the difference between the odd and the even sample size $n = 100$ (This is another odd property of the median, cfr. Cabrera et al 1994).

Finally we looked at the distribution of $\sup_{X'} |T_{n+m}(X') - T_n(X)|$, where X' is obtained from X by adding a small number m of arbitrary observations. We simulated this distribution for 2%, respectively 5% of contamination. The results are shown in Figure 3. (We only pictured the results for the 25%-trimmed median. The Huber M-estimator behaves similarly.) The difference between n odd and n even becomes much smaller. For 5 outliers added to 100 good observations, we see again that the distribution of $\sup_{X'} |T_{n+m}(X') - T_n(X)|$ for the trimmed mean has a lower median and is less dispersed than for the sample median. Our results for single outliers seem thus to remain valid for small amounts of outliers. These

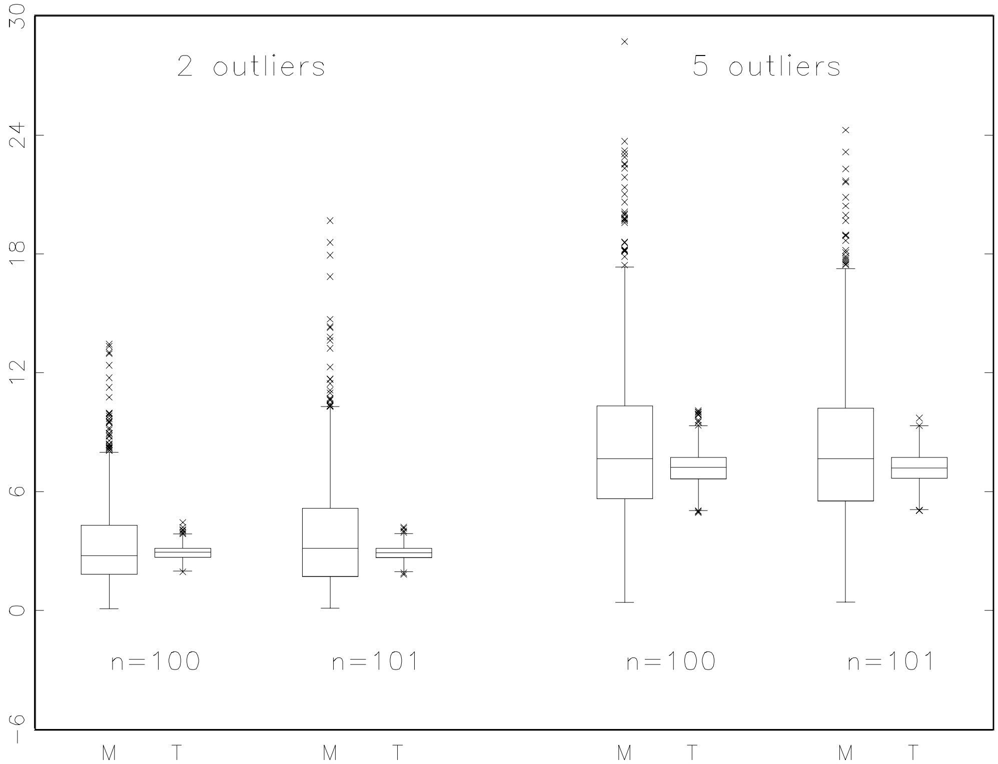

Figure 3: Boxplots of 1000 numbers generated from $\sup_{X'} |T_{n+m}(X') - T_n(X)|$*, where X follows a standard normal distribution and X' is obtained from X by adding m arbitrary observations. We considered* $m = 2$ *and* 5*,* $n = 100$ *and* 101*, and as estimators the median (M) and the 25%-trimmed mean (T).*

findings are not in contradiction with the famous minimax bias property of the median (Huber 1981, page 75). The latter is an optimality property for the functional version of the median. This simulation experiment dealt with the (finite-)sample median. Moreover, we see from Figure 1 that the median of the distibution of the bias of the sample median under a *specified* contaminating distribution is still consistently smaller than for the other considered estimators.

5 Conclusions

This paper shows that the empirical influence functions of trimmed means and M-estimators of location converge almost surely to the theoretical influence function, while the EIF of the median is not a consistent estimator for the corresponding IF. This result resembles the lack of consistency of the jackknife variance estimator for the median. Another example of a non-convergent EIF is given by Christmann et al (1994).

By introducing a sample-based measure of gross-error sensitivity, we showed that the median is no longer most robust with respect to single outliers, although it minimizes the functional-based gross-error sensitivity. Simulations have shown that this is a real phenomenon, and not just a theoretical artefact. Heavily trimmed means appear to be more robust to single outliers, while they achieve at the same time a higher efficiency at normal models. Other arguments in favor of using heavily trimmed means are given by Croux (1996) and Oosterhoff (1994).

The influence function needs to be quite carefully interpreted . If we have one outlier out of $n+1$ observations, the percentage of contamination tends to zero at the rate n^{-1}. Suppose now that we put a total of $\lfloor \varepsilon n \rfloor + 1$ identical observations at position x (one can argue whether this is a realistic situation), implying that the number of contaminants is linear in the number of observations. The following equality holds (at least for consistent estimators) :

$$\lim_{\varepsilon \downarrow 0} \lim_{n \to \infty} \frac{T((1 - \frac{\lfloor \varepsilon n \rfloor + 1}{n+1})F_n + \frac{\lfloor \varepsilon n \rfloor + 1}{n+1}\Delta_x) - T(F_n)}{\frac{\lfloor \varepsilon n \rfloor + 1}{n+1}} = \mathrm{IF}(x, T; F) \quad a.s. \tag{5.1}$$

It follows from Proposition 1 that it is not always allowed to change the order of the two limits in (5.1). Equation (5.1) shows that the influence function $\mathrm{IF}(x, T; F)$ needs to be considered as an approximation for the standardized influence of a proportion of ε % of outliers at the same position x, with ε very small and n huge, but the total number of outliers still fairly large and certainly different from 1. Indeed, for the sample mean the percentage of contamination ε should tend at a slower rate than $1/n$ to zero to obtain coherency between the influence function and its empirical counterpart. For smoother estimators however, we can still take $\varepsilon \approx 1/n$.

6 Appendix

Proof of Proposition 1: Denote $h = \lfloor n/2 \rfloor + 1$. First suppose that n is odd. We have $\mathrm{med}_n(X) = x_{(h)}$. Adding the observation x gives rise to a shift

$$\mathrm{med}_{n+1}(X \cup \{x\}) - \mathrm{med}_n(X) = \begin{cases} (x_{(h+1)} - x_{(h)})/2 & \text{for } x \geq x_{(h+1)} \\ (x - x_{(h)})/2 & \text{for } x_{(h+1)} > x > x_{(h-1)} \\ (x_{(h-1)} - x_{(h)})/2 & \text{for } x_{(h-1)} \geq x \end{cases} \tag{6.1}$$

The theory of spacings provides the following fact (David 1981 page 257):

(S) Consider the random variables $U_1 = f(\mathrm{med}(F))n(X_{(h+1)} - X_{(h)})$, $U_2 = 2f(\mathrm{med}(F))\sqrt{n}(X_{(h)} - \mathrm{med}(F))$, and $U_3 = f(\mathrm{med}(F))n(X_{(h)} - X_{(h-1)})$, where $h/n \to 0.5$. The vector (U_1, U_2, U_3) converges in law to (Z_1, Z_2, Z_3), where Z_1 and Z_3 have an exponential distribution with mean 1 and Z_2 has a standard normal distribution. Furthermore, Z_1, Z_2 and Z_3 are independent.

Since both $X_{(h+1)}$ and $X_{(h-1)}$ converge almost surely to $\mathrm{med}(F)$ (Jurečková and Sen 1996, page 94), the result follows from (S) for $x \neq 0$. For $x = 0$ we combine (S) with the fact that both $P(x \geq X_{(h+1)})$ and $P(x \leq X_{(h-1)})$ tend to 1/2.

The proof for n even is analogous. There we have

$$\mathrm{med}_{n+1}(X \cup \{x\}) - \mathrm{med}_n(X) = \begin{cases} (x_{(h)} - x_{(h-1)})/2 & \text{for } x \geq x_{(h)} \\ (2x - x_{(h)} - x_{(h-1)})/2 & \text{for } x_{(h)} > x > x_{(h-1)} \\ (x_{(h-1)} - x_{(h)})/2 & \text{for } x_{(h-1)} \geq x. \end{cases} \tag{6.2}$$

□

Proof of Proposition 2: Denote $r = \lfloor \alpha n \rfloor$ and take $i_0(x)$ such that $X_{(i_0(x))} \leq x \leq X_{(i_0(x)+1)}$. Suppose $r = \lfloor \alpha(n+1) \rfloor$ (The case $r + 1 = \lfloor \alpha(n+1) \rfloor$ can be proved in the same way.) We can verify that

$$\mathrm{EIF}(x, T_n^\alpha; X) = \begin{cases} \frac{n+1}{n+1-2r} X_{(r)} - \frac{n+1}{(n+1-2r)(n-2r)} \sum_{i=r+1}^{n-r} X_{(i)} & i_0(x) < r \\ \frac{n+1}{n+1-2r} x - \frac{n+1}{(n+1-2r)(n-2r)} \sum_{i=r+1}^{n-r} X_{(i)} & r \leq i_0(x) \leq n-r \\ \frac{n+1}{n+1-2r} X_{(n+1-r)} - \frac{n+1}{(n+1-2r)(n-2r)} \sum_{i=r+1}^{n-r} X_{(i)} & n-r < i_0(x) \end{cases} \tag{6.3}$$

Since $X_{(r)}$ converges almost surely to $F^{-1}(\alpha)$, $X_{(n-r)}$ a.s. to $F^{-1}(1-\alpha)$, and

$$\frac{n+1}{n+1-2r} \frac{\sum_{i=r+1}^{n-r} X_{(i)}}{n-2r} \to \frac{\int_\alpha^{1-\alpha} t\, dF(t)}{(1-2\alpha)^2} = \frac{\mathrm{med}(F)}{1-2\alpha} \quad a.s.$$

(cfr. Jurečková and Sen 1996, page 95), it follows from (6.3) and (2.3) that

$$\lim_{n\to\infty} \sup_x |\mathrm{EIF}(x, T_n^\alpha; X) - \mathrm{IF}(x, T^\alpha; F)| \leq \lim_{n\to\infty} \sup_{|x| \leq F^{-1}(1-\alpha)} \left| (\frac{n+1}{n+1-2r} - \frac{1}{1-2\alpha})\, x \right| = 0$$

almost surely. □

Proof of Proposition 3: Denote $t_n(x) = T_{n+1,\psi}(X \cup \{x\}), t_n = T_{n,\psi}(X)$ and $\lambda_n(t) = \sum_{i=1}^n \psi(X_i - t)$. First we show that

$$\sup_x |t_n(x) - t_n| \to 0 \quad a.s. \tag{6.4}$$

Indeed, if (6.4) does not hold, there exists an ε_0 and subsequences $\{n_k\}$ and $\{x_k\}$ such that $|t_{n_k}(x_k) - t_{n_k}| > \varepsilon_0$. By taking a further subsequence we may assume w.l.o.g. that $t_{n_k}(x_k) > t_{n_k} + \varepsilon_0$. Since the functions λ_n are non-increasing we have

$$\lambda_{n_k}(t_{n_k}(x_k)) \leq \lambda_{n_k}(t_{n_k} + \varepsilon_0). \tag{6.5}$$

Using

$$\lambda_n(t_n(x)) = \sum_{i=1}^{n} \psi(X_i - t_n(x)) + \psi(x - t_n(x)) - \psi(x - t_n(x)) = -\psi(x - t_n(x)) \tag{6.6}$$

and the boundedness of ψ we conclude

$$\limsup_k \frac{\lambda_{n_k}(t_{n_k}(x_k))}{n_k} = 0. \tag{6.7}$$

On the other hand, it follows from (6.5), the law of large numbers and the fact that t_n is a consistent estimator of med(F) (cfr. Huber, page 48) that

$$\begin{aligned} \limsup_k \frac{\lambda_{n_k}(t_{n_k}(x_k))}{n_k} &\leq \limsup_k \frac{\sum_{i=1}^{n_k} \psi(X_i - t_{n_k} - \varepsilon_0)}{n_k} \\ &= E_F[\psi(X - \mathrm{med}(F) - \varepsilon_0)] < E_F[\psi(X - \mathrm{med}(F))] = 0. \end{aligned} \tag{6.8}$$

(Here, we used the condition that ψ is strictly increasing at the origin.) Since (6.8) contradicts (6.7), (6.4) should be valid.

A Taylor development of λ_n around t_n yields

$$\lambda_n(t_n(x)) = \lambda_n(t_n) + \lambda_n'(t_n)(t_n(x) - t_n) + o(t_n(x) - t_n).$$

Since $\lambda_n(t_n) = 0$, we have

$$\mathrm{EIF}(x, T_{n,\psi}, X) = (n+1)(t_n(x) - t_n) = \frac{\lambda_n(t_n(x))}{(\lambda_n'(t_n) + o(1))/(n+1)}.$$

Due to (6.4), (6.6), the law of large numbers applied to $\lambda_n'(t_n)/(n+1) = -\sum_{i=1}^{n} \psi'(X_i - t_n)/(n+1)$, and the fact that t_n tends to med(F) almost surely, we conclude

$$\lim_n \sup_x |\mathrm{EIF}(x, T_{n,\psi}; X) - \frac{\psi(x - \mathrm{med}(F))}{E_F[\psi'(X - \mathrm{med}(F))]}| = 0.$$

□

Proof of Proposition 4: Take $u \geq 0$. From (6.1) it follows that, for n odd,

$$P_F(\sup_x |\mathrm{EIF}(x, \mathrm{med}_n; X)| \leq u) = P_F\left((n+1)\max(X_{(h+1)} - X_{(h)}, X_{(h)} - X_{(h-1)}) \leq 2u\right).$$

Using (S) of Proposition 2 we obtain

$$\lim_{n\to\infty, n \text{ odd}} P_F(\sup_x |\text{EIF}(x, \text{med}_n; X)| \le u) = P(\max(Y_1, Y_2) \le 2f(\text{med}(F))u), \qquad (6.9)$$

where Y_1 and Y_2 are two independent exponential variables with mean 1. On the other hand, (6.2) yields

$$\lim_{n\to\infty, n \text{ even}} P_F(\sup_x |\text{EIF}(x, \text{med}_n; X)| \le u) = P(Y_1 \le 2f(\text{med}(F))u). \qquad (6.10)$$

Equations (6.9) and (6.10) yield the stated result (3.3). □

References

Andrews, D.F., P.J. Bickel, P.J., Hampel, F.R., Huber, P.J., Tukey, J.W., and Rogers, W.H. (1972), *Robust Estimates of Location: Survey and Advances,* Princeton University Press, Princeton NJ.

Cabrera, J., Maguluri, G., and Singh, K. (1994), "An Odd Property of the Median," *Statistics and Probability Letters,* 19, 349-354.

Christmann, A., Gather, U., and Scholz, G. (1994), "Some Properties of the Length of the Shortest Half," *Statistica Neerlandica,* 48, 209-213.

Croux, C. (1996), "Maximum Deviation Curves for Location Estimators," *Statistics,* 28, 285-305.

David, H.A. (1981), *Order Statistics,* New York: John Wiley.

Hampel, F.R. (1974), "The Influence Curve and Its Role in Robust Estimation", *Journal of the American Statistical Association,* 69, 383–393.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., and Stahel, W.A. (1986), *Robust Statistics: the Approach based on Influence Functions,* New York: John Wiley.

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

Jurečková, J. and Sen, P.K. (1996), *Robust Statistical Procedures, Asymptotics and Interrelations,* New York: John Wiley.

Oosterhoff, J. (1994), "Trimmed Mean or Sample Median?" *Statistics and Probability Letters,* 20, 401-409.

Rousseeuw, P.J. (1982), "Most Robust M-Estimators in the Infinitesimal Sense," *Zeitschrift für Wahrscheinlichkeitstheorie,* 61, 541-551.