ROBUST TESTS BASED ON MINIMUM DENSITY POWER DIVERGENCE ESTIMATORS AND SADDLEPOINT APPROXIMATIONS

AIDA TOMA

The nonrobustness of classical tests for parametric models is a well known problem and various robust alternatives have been proposed in literature. Usually, the robust tests are based on first order asymptotic theory and their accuracy in small samples is often an open question. In this paper we propose tests which have both robustness properties, as well as good accuracy in small samples. These tests are based on robust minimum density power divergence estimators and saddlepoint approximations.

AMS 2000 Subject Classification: 62F35, 62F03.

Key words: robust tests, saddlepoint approximations, divergences.

1. INTRODUCTION

In various statistical applications, approximations of the probability that a random variable θ_n exceed a given threshold value are important, because in most cases the exact distribution function of θ_n could be difficult or even impossible to be determined. Particularly, such approximations are useful for computing p-values for hypothesis testing. The normal approximation is one of the widely used, but often it doesn't ensures a good accuracy to the testing procedure for moderate to small samples. An alternative is to use saddlepoint approximations which provide a very good approximation with a small relative error to the tail, as well as in the center of the distribution. Saddlepoint approximations have been widely studied and applied due to their excellent performances. For details on theory and applications of saddlepoint approximations we refer for example to the books Field and Ronchetti [5] and Jensen [10], as well as to the papers Field [3], Field and Hampel [4], Gatto and Ronchetti [8], Gatto and Jammalamadaka [7], Almudevar et al. [1], Field et al. [6].

Beside the accuracy in moderate to small samples, the robustness is an important requirement in hypotheses testing. In this paper, we combine robust minimum density power divergence estimators and saddlepoint approximations as presented in Robinson et al. [12] and obtain robust test statistics for

MATH. REPORTS **12(62)**, *4* (2010), 383–392

hypotheses testing which are asymptotically χ^2-distributed with a relative error of order $\mathcal{O}(n^{-1})$. The minimum density power divergence estimators were introduced by Basu et al. [2] for robust and efficient estimation in general parametric models. Particularly, they are M-estimators and the associated ψ-functions can satisfy conditions such that the tests based on these estimators to be robust and accurate in small samples.

The paper is organized as follows. In Section 2 we present the class of minimum density power divergence estimators. In Section 3, using the minimum density power divergence estimators, we define saddlepoint test statistics and give approximations for the p-values of the corresponding tests. In Section 4, we prove robustness properties of the tests, by means of the influence function and the asymptotic breakdown point. In Section 5, we consider the normal location model and the exponential scale model to exemplify the proposed robust testing method. We show that conditions for robust testing and good accuracy are simultaneously satisfied. We also prove results regarding the asymptotic breakdown point of the test statistics.

2. MINIMUM DENSITY POWER DIVERGENCE ESTIMATORS

The class of minimum density power divergence estimators was introduced in Basu et al. [2] for robust and efficient estimation in general parametric models.

The family of density power divergences is defined by

$$d_\alpha(g,f) = \int \left\{ f^{1+\alpha}(z) - \left(1+\frac{1}{\alpha}\right) g(z) f^\alpha(z) + \frac{1}{\alpha} g^{1+\alpha}(z) \right\} \mathrm{d}z, \quad \alpha > 0,$$

for g and f density functions. When $\alpha \to 0$ this family reduces to the Kullback-Leibler divergence (see Basu et al. [2]) and $\alpha = 1$ leads to the square of the standard L_2 distance between g and f.

Let F_θ be a distribution with density f_θ relative to the Lebesgue measure, where the parameter θ is known to belong to a subset Θ of $\mathbb{R}^d$. Given a random sample $X_1, \ldots, X_n$ from F_{θ_0} and $\alpha > 0$, a minimum density power divergence estimator of the unknown true value of the parameter θ_0 is defined by

$$\widehat{\theta}_n(\alpha) := \arg\min_{\theta \in \Theta} \left\{ \int f_\theta^{1+\alpha}(z) \mathrm{d}z - \left(1+\frac{1}{\alpha}\right) \frac{1}{n} \sum_{i=1}^n f_\theta^\alpha(X_i) \right\}. \tag{1}$$

Formula (1) determines a class of M-estimators indexed by the parameter α that specifies the divergence. The choice of α represents an important aspect, because α controls the trade-off between robustness and asymptotic efficiency of the estimation procedure. It is known that estimators with small α have strong robustness properties with little loss in asymptotic efficiency relative to maximum likelihood under the model conditions.

We recall that a map T which sends an arbitrary distribution function into the parameter space is a statistical functional corresponding to an estimator $\widehat{\theta}_n$ of the parameter θ_0 whenever $T(F_n) = \widehat{\theta}_n$, F_n being the empirical distribution function associated to the sample on F_{θ_0}. The influence function of the functional T in F_0 measures the effect on T of adding a small mass at x and is defined as

$$\mathrm{IF}(x; T, F_{\theta_0}) = \lim_{\varepsilon \to 0} \frac{T(\widetilde{F}_{\varepsilon x}) - T(F_{\theta_0})}{\varepsilon},$$

where $\widetilde{F}_{\varepsilon x} = (1-\varepsilon)F_{\theta_0} + \varepsilon\delta_x$ and δ_x represents the Dirac distribution. When the influence function is bounded, the corresponding estimator is called robust. More details on this robustness measure are given for example in Hampel et al. [9].

For any given α, the minimum density power divergence functional at the distribution G with density g is defined by

$$T_\alpha(G) := \arg\min_{\theta \in \Theta} d_\alpha(g, f_\theta).$$

From M-estimation theory, the influence function of the density power divergence functional is

$$\mathrm{IF}(x; T_\alpha, F_{\theta_0}) = J^{-1}\{\dot{f}_{\theta_0}(x) f_{\theta_0}^{\alpha-1}(x) - \xi\}, \tag{2}$$

where $\dot{f}_\theta$ denotes the derivative with respect to θ of f_θ, $\xi := \int \dot{f}_{\theta_0}(z) f_{\theta_0}^{\alpha}(z)\mathrm{d}z$ and $J := \int \dot{f}_{\theta_0}(z)\dot{f}_{\theta_0}(z)^t f_{\theta_0}^{\alpha-1}(z)\mathrm{d}z$. When J is finite, $\mathrm{IF}(x; T_\alpha, F_{\theta_0})$ is bounded, and hence $\widehat{\theta}_n(\alpha)$ is robust, whenever $\dot{f}_{\theta_0}(x) f_{\theta_0}^{\alpha-1}(x)$ is bounded.

3. SADDLEPOINT TEST STATISTICS BASED ON MINIMUM DENSITY POWER DIVERGENCE ESTIMATORS

In order to test the hypothesis H_0: $\theta = \theta_0$ in $\mathbb{R}^d$ with respect to the alternatives H_1: $\theta \neq \theta_0$, we define test statistics based on minimum density power divergence estimators.

Let α be fixed. Notice that, a minimum density power divergence estimator of the parameter θ_0 defined by (1) is an M-estimator obtained as solution of the equation

$$\sum_{i=1}^{n} \psi_\alpha(X_i, \widehat{\theta}_n(\alpha)) = 0,$$

where

$$\psi_\alpha(x, \theta) = f_\theta^{\alpha-1}(x)\dot{f}_\theta(x) - \int f_\theta^{\alpha}(z)\dot{f}_\theta(z)\mathrm{d}z.$$

Assume that the cumulant generating function of the vector of scores $\psi_\alpha(X,\theta)$ defined by

$$K_{\psi_\alpha}(\lambda,\theta) := \log \mathrm{E}_{F_{\theta_0}}\{\mathrm{e}^{\lambda^t\psi_\alpha(X,\theta)}\}$$

exists.

The test statistic which we consider is $h_\alpha(\widehat{\theta}_n(\alpha))$, where

$$h_\alpha(\theta) := \sup_\lambda\{-K_{\psi_\alpha}(\lambda,\theta)\}.$$

The p-value of the test based on $h_\alpha(\widehat{\theta}_n(\alpha))$ is

$$p := P_{H_0}(h_\alpha(\widehat{\theta}_n(\alpha)) \geq h_\alpha(\theta_n(\alpha))), \tag{3}$$

where $\theta_n(\alpha)$ is the observed value of $\widehat{\theta}_n(\alpha)$. This p-value can be approximated as in Robinson et al. [12], as soon as the density of $\widehat{\theta}_n(\alpha)$ exists and admits the saddlepoint approximation

$$f_{\widehat{\theta}_n(\alpha)}(t) = (2\pi/n)^{-d/2}\mathrm{e}^{nK_{\psi_\alpha}(\lambda_\alpha(t),t)}|B_\alpha(t)|\,|\Sigma_\alpha(t)|^{-1/2}(1+\mathcal{O}(n^{-1})), \tag{4}$$

where $\lambda_\alpha(t)$ is the saddlepoint satisfying

$$K'_{\psi_\alpha}(\lambda,t) := \frac{\partial}{\partial\lambda}K_{\psi_\alpha}(\lambda,t) = 0,$$

$|\cdot|$ denotes the determinant,

$$B_\alpha(t) := \mathrm{e}^{-K_{\psi_\alpha}(\lambda_\alpha(t),t)}\mathrm{E}_{F_{\theta_0}}\left\{\frac{\partial}{\partial t}\psi_\alpha(X,t)\mathrm{e}^{\lambda^t\psi_\alpha(X,t)}\right\}$$

and

$$\Sigma_\alpha(t) := \mathrm{e}^{-K_{\psi_\alpha}(\lambda_\alpha(t),t)}\mathrm{E}_{F_{\theta_0}}\left\{\psi_\alpha(X,t)\psi_\alpha(X,t)^t\mathrm{e}^{\lambda^t\psi_\alpha(X,t)}\right\}.$$

Conditions ensuring the saddlepoint approximation (4) are given in Almudevar et al. [1] for density of a general M-estimator and apply here as well. Under these conditions, using the general result from Robinson et al. [12], the p-value (3) corresponding to the test that we propose is given by

$$p = \overline{Q}_d(n\widehat{u}_\alpha^2) + n^{-1}c_n\widehat{u}_\alpha^d\mathrm{e}^{-n\widehat{u}_\alpha^2/2}\left[\frac{G_\alpha(\widehat{u}_\alpha)-1}{\widehat{u}_\alpha^2}\right] + \overline{Q}_d(n\widehat{u}_\alpha^2)\mathcal{O}\left(n^{-1}\right), \tag{5}$$

where $\widehat{u}_\alpha = \sqrt{2h_\alpha(\theta_n(\alpha))}$, $c_n = \frac{n^{d/2}}{2^{d/2-1}\Gamma(d/2)}$, $Q_d = 1-\overline{Q}_d$ is the distribution function of a χ^2 variate with d degrees of freedom,

$$G_\alpha(u) = \int_{S_d}\delta_\alpha(u,s)\mathrm{d}s = 1 + u^2k(u)$$

for

$$\delta_\alpha(u,s) = \frac{\Gamma(d/2)|B_\alpha(y)|\,|\Sigma_\alpha(y)|^{-1/2}J_1(y)J_2(y)}{2\pi^{d/2}u^{d-1}},$$

where for any $y \in \mathbb{R}^d$, (r, s) are the polar coordinates corresponding to y, $r = \sqrt{y^t y}$ is the radial component and $s \in S_d$, the d-dimensional sphere of unit radius, $u = \sqrt{2h_\alpha(y)}$, $J_1(y) = r^{d-1}$, $J_2(y) = ru/(h'_\alpha(y)^t y)$ and $k(\widehat{u}_\alpha)$ is bounded and the order terms are uniform for $\widehat{u}_\alpha < \varepsilon$, for some $\varepsilon > 0$.

Moreover, p admits the following simpler approximation

$$p = \overline{Q}_d(n\widehat{u}_\alpha^2)(1 + \mathcal{O}((1 + n\widehat{u}_\alpha^2)/n)). \tag{6}$$

The accuracy of the test in small samples as it is assured by the approximations given by (5) and (6) can be accompanied by robustness properties. This will be discussed in the next sections.

4. ROBUSTNESS RESULTS

The use of a robust minimum density power divergence estimator leads to a test statistic $h_\alpha(\widehat{\theta}_n(\alpha))$ which is robust, too. This can be proved by computing the influence function, respectively the breakdown point for $h_\alpha(\widehat{\theta}_n(\alpha))$.

The statistical functional corresponding to the test statistic $h_\alpha(\widehat{\theta}_n(\alpha))$ is defined by $U_\alpha(G) := h_\alpha(T_\alpha(G))$, where T_α is the minimum density power divergence functional associated to $\widehat{\theta}_n(\alpha)$. The following proposition show that the influence function of the test statistic is bounded whenever the influence function of the minimum density power divergence estimator is bounded.

PROPOSITION 1. *The influence function of the test statistic* $h_\alpha(\widehat{\theta}_n(\alpha))$ *is*

$$\mathrm{IF}(x; U_\alpha, F_{\theta_0}) = h'_\alpha(\theta_0)^t J^{-1}\{\dot{f}_{\theta_0}(x) f_{\theta_0}^{\alpha-1}(x) - \xi\}, \tag{7}$$

where J *and* ξ *are those from* (2) *and*

$$h'_\alpha(\theta_0) = -\frac{\mathrm{E}_{F_{\theta_0}}\{\mathrm{e}^{\lambda_\alpha(\theta_0)^t \psi_\alpha(X,\theta_0)} \frac{\partial}{\partial\theta}\psi_\alpha(X,\theta_0)\lambda_\alpha(\theta_0)\}}{\mathrm{E}_{F_{\theta_0}}\{\mathrm{e}^{\lambda_\alpha(\theta_0)^t \psi_\alpha(X,\theta_0)}\}}.$$

Proof. The minimum density power divergence functional T_α is Fisher consistent, meaning that $T_\alpha(F_{\theta_0}) = \theta_0$ (see Basu et al. [2], p. 551). Using this, derivation yields

$$\mathrm{IF}(x; U_\alpha, F_{\theta_0}) = \frac{\partial}{\partial\varepsilon}[U_\alpha((1-\varepsilon)F_{\theta_0} + \varepsilon\delta_x)]_{\varepsilon=0} = h'_\alpha(\theta_0)^t\, \mathrm{IF}(x; T_\alpha, F_{\theta_0}). \tag{8}$$

Derivation with respect to θ gives

$$\begin{aligned} h'_\alpha(\theta_0) &= -\frac{\partial}{\partial\lambda}K_{\psi_\alpha}(\lambda_\alpha(\theta_0), \theta_0)\lambda'_\alpha(\theta_0) - \frac{\partial}{\partial\theta}K_{\psi_\alpha}(\lambda_\alpha(\theta_0), \theta_0) \\ &= -\frac{\partial}{\partial\theta}K_{\psi_\alpha}(\lambda_\alpha(\theta_0), \theta_0) = -\frac{\mathrm{E}_{F_{\theta_0}}\{\mathrm{e}^{\lambda_\alpha(\theta_0)^t \psi_\alpha(X,\theta_0)} \frac{\partial}{\partial\theta}\psi_\alpha(X,\theta_0)\lambda_\alpha(\theta_0)\}}{\mathrm{E}_{F_{\theta_0}}\{\mathrm{e}^{\lambda_\alpha(\theta_0)^t \psi_\alpha(X,\theta_0)}\}} \end{aligned}$$

using the definition of $\lambda_\alpha(\theta_0)$. Then (7) holds, by replacing $h'_\alpha(\theta_0)$ and (2) in (8). $\square$

The breakdown point also measures the resistance of an estimator or of a test statistic to small changes in the underlying distribution.

Maronna et al. [11] (p. 58) define the asymptotic contamination breakdown point of an estimator $\widehat{\theta}_n$ at F_{θ_0}, denoting it by $\varepsilon^*(\widehat{\theta}_n, F_{\theta_0})$, as the largest $\varepsilon^* \in (0,1)$ such that for $\varepsilon < \varepsilon^*$, $T((1-\varepsilon)F_{\theta_0} + \varepsilon G)$ as function of G remains bounded and also bounded away from the boundary $\partial\Theta$ of Θ. Here $T((1-\varepsilon)F_{\theta_0} + \varepsilon G)$ represents the asymptotic value of the estimator by the means of the convergence in probability, when the observations come from $(1-\varepsilon)F_{\theta_0} + \varepsilon G$. This definition can be considered for a test statistic, too.

The aforesaid robustness measure is appropriate to be applied for both minimum density power divergence estimators $\widehat{\theta}_n(\alpha)$, as well as for corresponding test statistics $h_\alpha(\widehat{\theta}_n(\alpha))$. This is due to the consistency of $\widehat{\theta}_n(\alpha)$ and to the continuity to h_α, allowing to use $T_\alpha(G)$ and $h_\alpha(T_\alpha(G))$ as asymptotic values by means of the convergence in probability. The consistency of $\widehat{\theta}_n(\alpha)$ for $T_\alpha(G)$ when the observations $X_1, \ldots, X_n$ are i.i.d. with distribution G is given in Basu et al. [2].

Remark 1. The asymptotic contamination breakdown point of the test statistic $h_\alpha(\widehat{\theta}_n(\alpha))$ at F_{θ_0} satisfies the inequality

$$\varepsilon^*(h_\alpha(\widehat{\theta}_n(\alpha)), F_{\theta_0}) \geq \varepsilon^*(\widehat{\theta}_n(\alpha), F_{\theta_0}). \tag{9}$$

Indeed, when $\varepsilon < \varepsilon^*(\widehat{\theta}_n(\alpha), F_{\theta_0})$, there exists a bounded and closed set $K \subset \Theta$, $K \cap \partial\Theta = \emptyset$ and $T_\alpha((1-\varepsilon)F_{\theta_0} + \varepsilon G) \in K$, for all G. By the continuity of h_α as function of θ, $h_\alpha(K)$ is bounded and closed in $[0,\infty)$ and

$$U_\alpha((1-\varepsilon)F_{\theta_0} + \varepsilon G) = h_\alpha(T_\alpha((1-\varepsilon)F_{\theta_0} + \varepsilon G)) \in h_\alpha(K),$$

for all G. This prove the inequality (9).

Remark 1 shows that the test statistic cope with at least as many outliers as the minimum density power divergence estimator.

5. ROBUST SADDLEPOINT TEST STATISTICS FOR SOME PARTICULAR MODELS

In this section we consider two particular parametric models, namely the normal location model and the exponential scale model, in order to provide examples regarding the robust testing method presented in previous sections.

For both models, the conditions assuring the approximation (5) of the test p-value, as well as conditions for robust testing are simultaneously satisfied. This confirms the qualities of the proposed testing method from good accuracy in small samples and robustness point of views.

Consider the univariate normal model $\mathcal{N}(m, \sigma^2)$ with σ known, m being the parameter of interest. For α fixed, the ψ-function associated to the minimum density power divergence estimator $\widehat{\theta}_n(\alpha)$ of the parameter $\theta_0 = m$ is

$$\psi_\alpha(x, m) = \frac{x - m}{\sigma^{\alpha+2}(\sqrt{2\pi})^\alpha}\left[\mathrm{e}^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}\right]^\alpha .$$

For any $\alpha > 0$, the conditions from Almudevar et al. [1] assuring the approximation (4) for the density of $\widehat{\theta}_n(\alpha)$ are verified, therefore the p-value of the test based $h_\alpha(\widehat{\theta}_n(\alpha))$ can be approximated in accordance with (5). On the other hand, for any $\alpha > 0$, the test statistic is robust on the basis of the Proposition 1 and to the fact that the influence function of $\widehat{\theta}_n(\alpha)$ is bounded. Moreover, for any $\alpha > 0$, the minimum density power divergence estimator has the asymptotic breakdown point 0.5, because it is in fact a redescending M-estimator (see Maronna et al. [11], p. 59, for discussions regarding the asymptotic breakdown point of redescending M-estimators). Then Remark 1 ensures that $\varepsilon^*(h_\alpha(\widehat{\theta}_n(\alpha)), F_{\theta_0}) \geq 0.5$.

Let us consider now the exponential scale model with density $f_\theta(x) = \frac{1}{\theta}\mathrm{e}^{-x/\theta}$ for $x \geq 0$. The exponential distribution is widely used to model random inter-arrival times and failure times, and it also arises in the context of time series spectral analysis. The parameter θ is the expected value of the random variable X and the sample mean is the maximum likelihood for θ. It is known that the sample mean and classical test statistics lack robustness and can be influenced by outliers. Therefore, robust alternatives need to be used.

For α fixed, the ψ-function of a minimum density power divergence estimator $\widehat{\theta}_n(\alpha)$ of the parameter θ is

$$\psi_\alpha(x, \theta) = \frac{x - \theta}{\theta^{\alpha+2}}\left(\mathrm{e}^{-\frac{x}{\theta}}\right)^\alpha + \frac{\alpha}{(\alpha+1)^2\theta^{\alpha+1}}, \quad x \geq 0.$$

For any $\alpha > 0$, the conditions from Almudevar et al. [1] assuring the approximation (4) for the density of $\widehat{\theta}_n(\alpha)$ are verified, therefore the p-value of the test based $h_\alpha(\widehat{\theta}_n(\alpha))$ can be approximated in accordance with (5). Moreover, for any $\alpha > 0$, the test statistic is robust on the basis of the Proposition 1 and to the fact that the influence function of $\widehat{\theta}_n(\alpha)$ is bounded. We also establish an inferior bound of the asymptotic contamination breakdown point of $h_\alpha(\widehat{\theta}_n(\alpha))$, using Remark 1 and the following proposition:

PROPOSITION 2. *The asymptotic contamination breakdown point of* $\widehat{\theta}_n(\alpha)$ *at* F_{θ_0} *satisfies the inequality*

$$\frac{\alpha}{(\alpha+1)^2} \leq \varepsilon^*(\widehat{\theta}_n(\alpha), F_{\theta_0}) \leq \frac{\left(\mathrm{e}^{-(1+\frac{1}{\alpha})}\right)^\alpha + \frac{\alpha^2}{(\alpha+1)^2}}{\left(\mathrm{e}^{-(1+\frac{1}{\alpha})}\right)^\alpha + \alpha}.$$

Proof. Put $\widetilde{F}_{\varepsilon G} = (1-\varepsilon)F_{\theta_0} + \varepsilon G$, $\varepsilon > 0$ for a given G. Then

$$(1-\varepsilon)\mathrm{E}_{F_{\theta_0}}\psi_\alpha(X, T_\alpha(\widetilde{F}_{\varepsilon G})) + \varepsilon \mathrm{E}_G \psi_\alpha(X, T_\alpha(\widetilde{F}_{\varepsilon G})) = 0. \tag{10}$$

We first prove that

$$\varepsilon^*(\widehat{\theta}_n(\alpha), F_{\theta_0}) \leq \frac{\left(\mathrm{e}^{-(1+\frac{1}{\alpha})}\right)^\alpha + \frac{\alpha^2}{(\alpha+1)^2}}{\left(\mathrm{e}^{-(1+\frac{1}{\alpha})}\right)^\alpha + \alpha}. \tag{11}$$

Let $\varepsilon < \varepsilon^*(\widehat{\theta}_n(\alpha), F_{\theta_0})$. Then, for some $C_1 > 0$ and $C_2 > 0$ we have $C_1 \leq |T_\alpha(\widetilde{F}_{\varepsilon G})| \leq C_2$, for all G. Take $G = \delta_{x_0}$, with $x_0 > 0$. Then (10) becomes

$$(1-\varepsilon)\mathrm{E}_{F_{\theta_0}}\psi_\alpha(X, T_\alpha(\widetilde{F}_{\varepsilon x_0})) + \varepsilon \psi_\alpha(x_0, T_\alpha(\widetilde{F}_{\varepsilon x_0})) = 0. \tag{12}$$

Taking into account that

$$\psi_\alpha(x, T_\alpha(\widetilde{F}_{\varepsilon x_0})) \leq \frac{1}{\alpha(T_\alpha(\widetilde{F}_{\varepsilon x_0}))^{\alpha+1}}\left(\mathrm{e}^{-(1+\frac{1}{\alpha})}\right)^\alpha + \frac{\alpha}{(\alpha+1)^2(T_\alpha(\widetilde{F}_{\varepsilon x_0}))^{\alpha+1}}$$

for all x (the right hand term is the maximum value of the function $x \to \psi_\alpha(x, T_\alpha(\widetilde{F}_{\varepsilon x_0}))$), from (12) we deduce

$$0 \leq (1-\varepsilon)\left[\frac{1}{\alpha}\left(\mathrm{e}^{-(1+\frac{1}{\alpha})}\right)^\alpha + \frac{\alpha}{(\alpha+1)^2}\right] +$$
$$+\varepsilon\left[\left(\frac{x_0}{T_\alpha(\widetilde{F}_{\varepsilon x_0})} - 1\right)\left(\mathrm{e}^{-\frac{x_0}{T_\alpha(\widetilde{F}_{\varepsilon x_0})}}\right)^\alpha + \frac{\alpha}{(\alpha+1)^2}\right].$$

Letting $x_0 \to 0$, since $T_\alpha(\widetilde{F}_{\varepsilon x_0})$ is bounded we have

$$\varepsilon \leq \frac{\left(\mathrm{e}^{-(1+\frac{1}{\alpha})}\right)^\alpha + \frac{\alpha^2}{(\alpha+1)^2}}{\left(\mathrm{e}^{-(1+\frac{1}{\alpha})}\right)^\alpha + \alpha}$$

and consequently (11) is satisfied.

We now prove the inequality $\frac{\alpha}{(\alpha+1)^2} \leq \varepsilon^*(\widehat{\theta}_n(\alpha), F_{\theta_0})$. Let $\varepsilon > \varepsilon^*(\widehat{\theta}_n(\alpha), F_{\theta_0})$. Then, there exists a sequence of distributions (G_n) such that $T_\alpha(\widetilde{F}_{\varepsilon G_n}) \to \infty$ or $T_\alpha(\widetilde{F}_{\varepsilon G_n}) \to 0$.

Suppose that $T_\alpha(\widetilde{F}_{\varepsilon G_n}) \to 0$. Since

$$\psi_\alpha(x, T_\alpha(\widetilde{F}_{\varepsilon G_n})) \geq \frac{1}{(T_\alpha(\widetilde{F}_{\varepsilon G_n}))^{\alpha+1}} \left[\frac{\alpha}{(\alpha+1)^2} - 1\right]$$

for all x, (10) implies

$$(13)\quad 0 \geq (1-\varepsilon)\mathrm{E}_{F_{\theta_0}} \left\{(T_\alpha(\widetilde{F}_{\varepsilon G_n}))^{\alpha+1}\psi_\alpha(X, T_\alpha(\widetilde{F}_{\varepsilon G_n}))\right\} + \varepsilon\left[\frac{\alpha}{(\alpha+1)^2} - 1\right].$$

Using the bounded convergence theorem, we have

$$\lim_{n\to\infty} \mathrm{E}_{F_{\theta_0}} \left\{(T_\alpha(\widetilde{F}_{\varepsilon G_n}))^{\alpha+1}\psi_\alpha(X, T_\alpha(\widetilde{F}_{\varepsilon G_n}))\right\} = \frac{\alpha}{(\alpha+1)^2}.$$

Then (13) yields

$$0 \geq (1-\varepsilon)\frac{\alpha}{(\alpha+1)^2} + \varepsilon\left[\frac{\alpha}{(\alpha+1)^2} - 1\right]$$

which implies that

$$(14)\qquad \varepsilon \geq \frac{\alpha}{(\alpha+1)^2}.$$

Suppose now that $T_\alpha(\widetilde{F}_{\varepsilon G_n}) \to \infty$. Since

$$\psi_\alpha(x, T_\alpha(\widetilde{F}_{\varepsilon G_n})) \leq \frac{1}{\alpha(T_\alpha(\widetilde{F}_{\varepsilon G_n}))^{\alpha+1}}\left(\mathrm{e}^{-(1+\frac{1}{\alpha})}\right)^\alpha + \frac{\alpha}{(\alpha+1)^2(T_\alpha(\widetilde{F}_{\varepsilon G_n}))^{\alpha+1}},$$

relation (10) implies

$$0 \leq (1-\varepsilon)\mathrm{E}_{F_{\theta_0}} \left\{(T_\alpha(\widetilde{F}_{\varepsilon G_n}))^{\alpha+1}\psi_\alpha(X, T_\alpha(\widetilde{F}_{\varepsilon G_n}))\right\} + $$
$$+\varepsilon\left[\frac{1}{\alpha}\left(\mathrm{e}^{-(1+\frac{1}{\alpha})}\right)^\alpha + \frac{\alpha}{(\alpha+1)^2}\right].$$

Using again the bounded convergence theorem, we obtain

$$0 \leq (1-\varepsilon)\left[-1 + \frac{\alpha}{(\alpha+1)^2}\right] + \varepsilon\left[\frac{1}{\alpha}(\mathrm{e}^{-(1+\frac{1}{\alpha})})^\alpha + \frac{\alpha}{(\alpha+1)^2}\right]$$

which yields

$$(15)\qquad \varepsilon \geq \frac{1 - \frac{\alpha}{(\alpha+1)^2}}{\frac{1}{\alpha}\left(\mathrm{e}^{-(1+\frac{1}{\alpha})}\right)^\alpha + 1}.$$

Since the bound in (14) is smaller than the bound in (15), we deduce that $\varepsilon^*(\widehat{\theta}_n(\alpha), F_{\theta_0}) \geq \frac{\alpha}{(\alpha+1)^2}$. □

Remark 2. The asymptotic breakdown point of the test statistic satisfies $\varepsilon^*(h_\alpha(\widehat{\theta}_n(\alpha), F_{\theta_0})) \geq \frac{\alpha}{(\alpha+1)^2}$. This is obtained by applying Remark 1 and

Proposition 2. Particularly, when using the density power divergence corresponding to $\alpha = 1$, $\varepsilon^*(h_\alpha(\widehat{\theta}_n(\alpha), F_{\theta_0})) \geq 0.25$.

REFERENCES

[1] A. Almudevar, C. Field and J. Robinson, *The density of multivariate M-estimates.* Ann. Statist. **28** (2000), 275–297.

[2] A. Basu, I. Harris, N. Hjort and M. Jones, *Robust and efficient estimation by minimizing a density power divergence.* Biometrika **85** (1998), 549–559.

[3] C.A. Field, *Small sample asymptotic expansion for multivariate M-estimates.* Ann. Statist. **10** (1982), 672–689.

[4] C.A. Field and F.R. Hampel, *Small-sample asymptotic distributions of M-estimators of location.* Biometrika **69** (1982), 29–46.

[5] C.A. Field and E. Ronchetti, *Small Sample Asymptotics.* IMS, Hayward, CA., 1990.

[6] C.A. Field, J. Robinson and E. Ronchetti, *Saddlepoint approximations for multivariate M-estimates with applications to bootstrap accuracy.* Ann. Inst. Statist. Math. **60** (2008), 205–224.

[7] R. Gatto and S.R. Jammalamadaka, *A conditional saddlepoint approximation for testing problems.* J. Amer. Statist. Assoc. **94** (1999), 533–541.

[8] R. Gatto and E. Ronchetti, *General saddlepoint approximations of marginal densities and tail probabilities.* J. Amer. Statist. Assoc. **91** (1996), 666–673.

[9] F. Hampel, E. Ronchetti, P.J. Rousseeuw and W. Stahel, *Robust Statistics: the Approach based on Influence Functions.* Wiley, New York, 1986.

[10] J.L. Jensen, *Saddlepoint Approximations.* Oxford University Press, Oxford, 1995.

[11] R.A. Maronna, R.D. Martin and V.J Yohai, *Robust Statistics – Theory and Methods.* Wiley, New York, 2006.

[12] J. Robinson, E. Ronchetti and G.A. Young, *Saddlepoint approximations and tests based on multivariate M-estimates.* Ann. Statist. **31** (2003), 1154–1169.

Received 31 March 2009

Academy of Economic Studies
Mathematics Department
Piaţa Romană no. 6
010374 Bucharest, Romania

and

"Gheorghe Mihoc–Caius Iacob" Institute
of Mathematical Statistics and Applied Mathematics
Calea 13 Septembrie no. 13
050711 Bucharest, Romania