On Robustness of Median Estimator in Bernoulli Logistic Regression

Tomáš Hobza, Leandro Pardo

Abstract: The paper deals with generalized logistic regression models which include the classical model with binary responses governed by the Bernoulli law depending on the logistic regression function. The median estimator of the logistic regression parameters employing smoothed data in the discrete case, introduced in Hobza et al (2005), is considered. Sensitivity of this estimator to contaminations of the logistic regression data is studied by simulations and compared with the sensitivity of some robust estimators previously introduced to logistic regression. The median estimator is demonstrated to be more robust for higher levels of contamination.

MSC 2000: 62F10, 62F35
Key words: Logistic regression, Median estimator, Morgenthaler estimator, Bianco and Yohai estimator, Robustness

1 Introduction and basic concepts

In this paper we are interested in estimation of the parameter $\boldsymbol{\beta}_0 \in \mathbb{R}^d$ in the statistical models with independent real valued observations $Y_1, ..., Y_n$ of the form

$$Y_i \sim F_{\pi\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_0\right)}(y), \qquad 1 \leq i \leq n \tag{1}$$

where $\boldsymbol{x}_i \in \mathbb{R}^d$ are vectors of explanatory variables (regressors), $\boldsymbol{\beta}_0 \in \mathbb{R}^d$ is a vector of true parameters and $\boldsymbol{x}^T\boldsymbol{\beta} = \sum_{j=1}^d x_j\beta_j$ denotes the scalar product of $\boldsymbol{x} = (x_1, ..., x_d)^T \in \mathbb{R}^d$ and $\boldsymbol{\beta} = (\beta_1, ..., \beta_d)^T \in \mathbb{R}^d$. Further,

$$\pi(t) = \frac{e^t}{1+e^t} \qquad \text{for every } t \in \mathbb{R} \tag{2}$$

is the logistic regression function and $\mathcal{F} = \{F_\pi : \pi \in (0,1)\}$ is an arbitrary family of distribution functions on $\mathbb{R}$. The models given by (1), (2) are called *general logistic regression models.* In these models $\pi = \pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}_0\right)$ represents a nonlinear logistic regression and the distribution function F_π specifies the random response to this regression, see e.g. Andersen (1990), Agresti (2002), Pardo et al (2006) and others cited there.

Important special logistic regression models are obtained if for all $\pi \in (0,1)$ the random response functions $F_\pi(y)$ are either a right-continuous distribution functions with jumps $p_\pi(k) = F_\pi(k) - F_\pi(k-0)$, $k = 0, 1, \ldots$ summing up to 1,

This work was supported by the grants DGES PB2003-892, AV CR 107 5403 and MSMT 1M0572

or continuous piecewise differentiable distribution functions with densities $f_\pi(y) = dF_\pi(y)/dy$, $y \in \mathbb{R}$. In the first case we speak about *discrete models* and in the second case about *continuous models.*

Of particular attention is the discrete Bernoulli model obtained for the Bernoulli distribution functions

$$F_\pi(y) = (1-\pi)\, I(0 \le y < 1) + I(y \ge 1), \qquad \pi \in (0,1) \tag{3}$$

with jumps $1-\pi$ and π at $y = 0$ and $y = 1$. For these functions the problem reduces to the classical logistic regression with binary observations $Y_1, ..., Y_n$ taking on value 1 with probabilities $\pi\left(\boldsymbol{x}_1^T\boldsymbol{\beta}_0\right), \dots, \pi\left(\boldsymbol{x}_n^T\boldsymbol{\beta}_0\right)$ and value 0 with the complementary probabilities $1-\pi\left(\boldsymbol{x}_1^T\boldsymbol{\beta}_0\right), \dots, 1-\pi\left(\boldsymbol{x}_n^T\boldsymbol{\beta}_0\right)$.

In Hobza et al (2005) we proposed the median estimator of parameter $\boldsymbol{\beta}_0$ which can be formally defined as follows.

Definition 1.1. *The median estimator* $\widehat{\beta}_n$ of the true parameter $\boldsymbol{\beta}_0$ in the general logistic regression model given by (1) and (2) is defined by the formula

$$\widehat{\boldsymbol{\beta}}_n = \arg\min_{\boldsymbol{\beta}} \sum_{i=1}^n \left|Y_i - m\left(\pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}\right)\right)\right| \tag{4}$$

where $m(\pi)$ is for every $\pi \in (0,1)$ the median

$$m(\pi) = F_\pi^{-1}(1/2) = \inf\{y \in \mathbb{R} : F_\pi(y) \ge 1/2\}. \tag{5}$$

Note that this estimator is in fact member of the class of so-called *least absolute deviation estimators* (or briefly L_1-estimators) defined by

$$\widetilde{\boldsymbol{\beta}}_n = \arg\min_{\boldsymbol{\beta}} \sum_{i=1}^n \left|Y_i - \mu\left(u\left(\boldsymbol{x}_i^T\boldsymbol{\beta}\right)\right)\right| \tag{6}$$

where $\mu : \Theta \to \mathbb{R}$ and $u : \mathbb{R} \to \Theta$ are given functions. From the extensive literature dealing with these estimators one can mention Richardson and Bhattacharyya (1987), Yohai (1987), Pollard (1991), Morgenthaler (1992), Arcones (2001), Liese and Vajda (2003, 2004) and others cited in these papers.

The median estimator was not previously considered in the logistic regression because for the most important Bernoulli model (as well as for all other discrete logistic regression models) the median function $m(\pi) = F_\pi^{-1}(1/2)$ is not sensitive to small variations of the parameter $\pi \in (0,1)$. For example in the Bernoulli model we have the piecewise constant $m(\pi) = 0$ if $\pi \le 1/2$ and $m(\pi) = 1$ if $\pi > 1/2$. Therefore the classical median estimator (4) cannot be consistent in these models. The above mentioned sensitivity means the strict monotonicity of the median function $m(\pi)$ on its domain $(0,1)$. The originality of our approach consists in replacing the models with discrete responses by their standard modifications defined as follows.

Definition 1.2. The *standard modification* of a discrete logistic regression model (1) is the continuous logistic regression model with the observations

$$\widetilde{Y}_i = Y_i + W_i, \qquad 1 \le i \le n, \tag{7}$$

where W_i are an independent noise random variables uniformly distributed on the interval $(0,1)$ and independently added to the discrete observations Y_i of the original model (1).

The transformation introduced in the previous definition is statistically sufficient since the original observations Y_i can be recovered from $\widetilde{Y}_i$ as the integer parts $[\widetilde{Y}_i]$. At the same time the median functions $\widetilde{m}\,(\pi) = \widetilde{F}_\pi^{-1}\,(1/2)$ of the transformed observations (7) are already one-one on the interval $(0,1)$. For example the

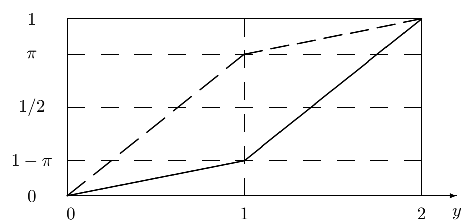

Figure 1: $F_\pi\,(y)$ full line, $F_{1-\pi}\,(y)$ dashed line.

standardly modified Bernoulli model with the discrete probabilities $1-\pi$ and π is the continuous model with the response distribution function

$$F_\pi\,(y) = (1-\pi)\;\; y\; I\,(0 < y \le 1) + [1 - \pi + \pi\,(y-1)]\, I\,(1 < y \le 2) \tag{8}$$

given in Figure 1. Consequently the median function has the form

$$m\,(\pi) = 1 + \frac{\pi - 1/2}{1/2 + |\pi - 1/2|}, \qquad \pi \in (0,1) \tag{9}$$

and is strictly increasing on $(0,1)$.

In the above mentioned paper of Hobza et al (2005) we proved under some regularity assumptions the consistency and asymptotic normality for the median estimator (4) in the continuous logistic regression model or in a standardly modified discrete regression model under consideration.

It is known (cf. e. g. Hampel et al (1986), Yohai (1987), Jurečková and Sen (1996), Zwanzig (1997)) that the median estimator of parameters of linear and nonlinear regression is robust with respect to contamination of observations from the assumed statistical models. This naturally leads to the hope that the median estimator for the general logistic regression is robust too. The aim of the present paper is to demonstrate the robustness of the median estimator and to compare it with the robustness of some well known estimators tailor-made for robust estimation in logistic regression. For this comparison were selected the L_1-estimators of Morgenthaler (1992) and the M-estimators of Bianco and Yohai (1996). These estimators, and also the MLE, were compared with the median estimator by performances on simulated noncontaminated and contaminated logistic regression observations.

In the simulation experiment we study particularly the Bernoulli discrete logistic regression models of the small dimension $d = 2$. The Bernoulli models are typical and the small dimensions are simpler and sufficient to provide an insight into the general properties of estimators. Furthermore, the considered logistic regression models are the same as used in Bianco and Yohai (1996) for demonstration of robustness of their estimator.

2 Robustness of the median estimator

From what was said at the end of the previous section it follows that one can expect more robustness, namely better resistance to the gross errors in observations $Y_1, \ldots, Y_n$, from our median estimator than e.g. from the classical MLE of the logistic regression parameters. Similar robust alternatives to the classical MLE's seem so far been considered only for the logistic regression models with Bernoulli responses. In this section we compare the median estimator with the L_1-estimators of Morgenthaler (1992) and the M-estimators of Bianco and Yohai (1996) which are the two most recent estimators considered for robust estimation in logistic regression with Bernoulli responses known to us. We start with description of the mentioned estimators.

Morgenthaler (1992) started with the weighted L_1-estimator

$$\boldsymbol{\beta}_n^{(0)} = \arg\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \frac{\left|Y_i - \pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}\right)\right|}{\sqrt{\pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}\right)\left(1-\pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}\right)\right)}},$$

more precisely with the solutions $\boldsymbol{\beta}_n^{(0)}$ of the system of equations $U_n^{(0)}(\beta) = 0$, for

$$U_n^{(0)}(\boldsymbol{\beta}) = D^T \left\{diag\left(\sigma_1(\boldsymbol{\beta}), ..., \sigma_n(\boldsymbol{\beta})\right)\right\}^{-1/2} \left(sgn\,\Delta_1(\boldsymbol{\beta}), ..., sgn\,\Delta_n(\boldsymbol{\beta})\right)^T$$

where $D = \left(D_{ij} = \partial\pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}\right)/\partial\beta_j\right)_{i,j=1}^T$, $\sigma_i^2(\boldsymbol{\beta}) = \pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}\right)\left(1-\pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}\right)\right)$ and sgn denotes the sign. Since the resulting estimator $\boldsymbol{\beta}_n^{(0)}$ was inconsistent, he proposed a slight modification $\boldsymbol{\beta}_n^{(1)}$ which solves the equation $U_n^{(1)}(\boldsymbol{\beta}) = 0$ for the centered version $U_n^{(1)}(\boldsymbol{\beta}) = U_n^{(0)}(\boldsymbol{\beta}) - E_\beta U_n^{(0)}(\boldsymbol{\beta})$. One can find an explicit formula for $U_n^{(1)}$, namely

$$U_n^{(1)}(\boldsymbol{\beta}) = \sum_{i=1}^{n} \sqrt{\pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}\right)\left(1-\pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}\right)\right)} \left(Y_i - \pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}\right)\right)\boldsymbol{x}_i. \tag{10}$$

An alternative robust estimator $\boldsymbol{\beta}_n^{(2)}$ for the logistic regression was proposed by Bianco and Yohai (1996) who also assumed the Bernoulli responses Y_i given in (1) for the $F_\pi(y)$ of (3). They started with the MLE

$$\boldsymbol{\beta}_n = \arg\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} D_i(\boldsymbol{\beta}) \tag{11}$$

where

$$D_i(\boldsymbol{\beta}) = -Y_i \ln \mu_i(\boldsymbol{\beta}) - (1 - Y_i)\ln(1 - \mu_i(\boldsymbol{\beta})) \tag{12}$$

and $\mu_i(\boldsymbol{\beta}) = E_{\boldsymbol{\beta}} Y_i = \pi\left(\boldsymbol{x}_i^T \boldsymbol{\beta}\right)$. Bianco and Yohai proved nice asymptotic properties of $\boldsymbol{\beta}_n$ like consistency and asymptotic normality with the variances at the Cramér-Rao lower bound.

However, this estimator is too sensitive to the gross errors (outliers) among the data $(\boldsymbol{x}_1, Y_1), ..., (\boldsymbol{x}_n, Y_n)$ which are the pairs $(\boldsymbol{x}_i, Y_i)$ where Y_i are not generated by the Bernoulli model $Be\left(\pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}_0\right)\right)$. Indeed, typical outliers are $Y_i = 0$ when the regressors $\boldsymbol{x}_i$ are leading to $\pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}_0\right) \approx 1$ or $Y_i = 1$ when $\pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}_0\right) \approx 0$. A simple source of outliers taking place with a probability $0 < \varepsilon < 1/2$ is the transmission of the true observations $Y_i \sim Be\left(\pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}_0\right)\right)$ through a binary symmetric channel BSC(ε) with independent inputs Y_i, additive (mod 2) independent noise $W_i \sim Be(\varepsilon)$ and independent outputs $Z_i = Y_i + W_i$ (mod 2) presented in Figure 2. Then the actual data $(\boldsymbol{x}_1, Z_1), ..., (\boldsymbol{x}_n, Z_n)$ contain responses Z_i generated by the

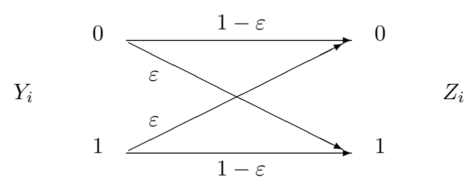

Figure 2: Binary Symmetric channel BSC(ε).

stochastic mixture

$$(1-\varepsilon)\, Be\left(\pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}_0\right)\right) + \varepsilon Be\left(1 - \pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}_0\right)\right) \tag{13}$$

of the Bernoulli models with parameters $\pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}_0\right)$ and $1 - \pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}_0\right)$.

To restrict the influence of the outliers $(\boldsymbol{x}_i, 0)$ with probabilities $\pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}_0\right) = 1 - \delta_i$ close to 1 and the outliers $(\boldsymbol{x}_i, 1)$ with probabilities $\pi\left(\boldsymbol{x}_i^T\boldsymbol{\beta}_0\right) = \delta_i$ close to 0, both of them leading to large $D_i(\boldsymbol{\beta}_0) = -\ln\delta_i$, Bianco and Yohai proposed the modified estimator

$$\boldsymbol{\beta}_n^{(2)} = \arg\min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \left[\rho\left(D_i(\boldsymbol{\beta})\right) + G\left(\mu_i(\boldsymbol{\beta})\right) + G\left(1 - \mu_i(\boldsymbol{\beta})\right)\right] \tag{14}$$

where

$$\rho(y) = \left(y - \frac{y^2}{2c}\right) I(0 \le y \le c) + \frac{c}{2} I(y > c) \tag{15}$$

is a hard-limiter defined on the real line and specified by a limiting constant $c > \ln 2$, $D_i(\beta)$ is defined by (12) and terms involving the function

$$G(\pi) = \int_0^\pi \rho'(-\ln t)\, dt \quad \text{for } \pi \in (0,1) \tag{16}$$

represent a bias correction. Under some regularity assumptions about the regressors $\boldsymbol{x}_1, ..., \boldsymbol{x}_n$ these authors proved that $\boldsymbol{\beta}_n^{(2)}$ consistently estimates $\boldsymbol{\beta}_0$.

In the simulation experiment we compare our median estimator $\widehat{\boldsymbol{\beta}}_n$ with the Morgenthaler estimator $\boldsymbol{\beta}_n^{(1)}$ and the Bianco-Yohai estimator $\boldsymbol{\beta}_n^{(2)}$ in the same logistic regression model as used in Bianco and Yohai (1996) to demonstrate the robustness of their estimator $\boldsymbol{\beta}_n^{(2)}$. The estimates $\boldsymbol{\beta}_n^{(1)}$ are evaluated as solutions of the equation $U_n^{(1)}(\boldsymbol{\beta}) = 0$ for $U_n^{(1)}(\boldsymbol{\beta})$ defined by (10) and the estimates $\boldsymbol{\beta}_n^{(2)}$ are evaluated by the minimization specified in (14). The constant c used in (15) is $-\ln 0.03 \approx \ln 33.3$, the same as used in the simulations of Bianco and Yohai. In addition to the estimates $\widehat{\boldsymbol{\beta}}_n$, $\boldsymbol{\beta}_n^{(1)}$ and $\boldsymbol{\beta}_n^{(2)}$ we evaluate also the MLE's $\boldsymbol{\beta}_n$. All four these estimates are evaluated from the same simulated data, namely the independent realizations

$$Y_i \sim (1-\varepsilon)\, Be\left(\pi\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_0\right)\right) + \varepsilon Be\left(1 - \pi\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_0\right)\right), \quad 1 \le i \le n \quad \text{(cf. (13))}$$

for a fixed $\boldsymbol{\beta}_0 = (\beta_{01}, \beta_{02})$ and $\boldsymbol{x}_i = (1, \xi_i)$ where ξ_i are random mutually independent $N(0,1)-$distributed regressors. The same four specifications will be used as in Bianco and Yohai (1996), defined by the expectations

$$E\pi\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_0\right) \in \{0.2, 0.3, 0.4, 0.5\}. \tag{17}$$

These expectations coincide with the probabilities $\Pr(Y_i = 1)$ and are one-one related to the parameters $\boldsymbol{\beta}_0 = (\beta_{01}, \beta_{02})$. Values of the corresponding parameters are given under the Tables 1-4 below.

In Tables 1-4 one can find for $\varepsilon \in \{0, 0.05, 0.1, 0.15, 0.2\}$ and $n \in \{400, 800, 1600\}$ the mean absolute errors

$$MAE(n) = \frac{1}{2000} \sum_{l=1}^{1000} \left(|\beta_{n1}(l) - \beta_{01}| + |\beta_{n2}(l) - \beta_{02}|\right)$$

for 1000 simulated realizations of $(Y_1, ..., Y_n)$ and the corresponding 1000 values $\boldsymbol{\beta}_n(l) = (\beta_{n1}(l), \beta_{n2}(l))$ of the MLE $\boldsymbol{\beta}_n = (\beta_{n1}, \beta_{n2})$, and the same errors also for the estimators $\boldsymbol{\beta}_n^{(1)}$, $\boldsymbol{\beta}_n^{(2)}$ and $\widehat{\boldsymbol{\beta}}_n$ denoted briefly as Morg, B&Y and Median. The values of all four these estimates were computed by using the subroutines for minimization and solving equations from the standard IMSL numerical package.

Using Tables 1-4 one compare performances of these four estimators measured by the corresponding mean absolute errors *MAE(n)*. We see from the first rows that if there is not contamination ($\varepsilon = 0$) then the best estimator is MLE. For light

ε	$(\widehat{\beta}_1, \widehat{\beta}_2)$	$MAE(400)$	NEF	$MAE(800)$	NEF	$MAE(1600)$	NEF
0	MLE	0.246	0	0.176	0	0.126	0
	Morg	0.270	0	0.193	0	0.135	0
	B&Y	0.340	0	0.232	0	0.161	0
	Median	1.128	7	0.678	0	0.378	0
0.05	MLE	1.011	0	1.037	0	1.048	0
	Morg	0.731	16	0.760	3	0.768	0
	B&Y	0.525	0	0.520	0	0.528	0
	Median	1.118	0	0.634	0	0.514	0
0.1	MLE	1.368	0	1.414	0	1.441	0
	Morg	3.518	1826	4.836	2568	5.114	3391
	B&Y	0.796	0	0.888	0	0.942	0
	Median	1.070	1	0.789	0	0.792	0
0.15	MLE	1.814	0	1.819	0	1.819	0
	Morg	-	-	-	-	-	-
	B&Y	1.590	0	1.608	0	1.609	0
	Median	1.317	2	1.336	0	1.369	0
0.2	MLE	2.029	0	2.036	0	2.036	0
	Morg	-	-	-	-	-	-
	B&Y	1.940	0	1.951	0	1.953	0
	Median	1.643	0	1.716	0	1.737	0

Table 1: Mean absolute errors *MAE(n)* of the four estimators in the model of Bianco and Yohai with Pr(Y=1)=0.2 and the true parameters $(\beta_{01}, \beta_{02}) = (-2.82, 2.82)$. Column NEF presents the numbers of simulated observation vectors $(Y_1, ..., Y_n)$ for which the evaluation of the corresponding estimates failed. If NEF exceeds 10000, neither *MAE(n)* nor NEF is presented.

and medium contaminations $(0 < \varepsilon < 0.1)$ the best is the estimator $\boldsymbol{\beta}_n^{(2)}$ of Bianco and Yohai. For heavier contamination $(\varepsilon \geq 0.1)$ the best is the median estimator $\widehat{\boldsymbol{\beta}}_n$.

The Morgenthaler's $\boldsymbol{\beta}_n^{(1)}$ is outperformed by B&Y and Median in each of the present contamination model. Moreover, it faces evaluation problems when the equations (10) are solved using the corresponding IMSL subroutines. This is indicated by the NEF numbers increasing with the contamination ε to unacceptable levels for $\varepsilon > 0.05$. Note that NEF is the count of the simulated realizations of $(Y_1, ..., Y_n)$ for which either the estimate cannot be evaluated or it is evaluated but its absolute error exceeds 100.

As a conclusion maybe said that the comparisons confirm the expected fact that for noncontaminated data the MLE dominates all three remaining estimators. On the other hand, they also confirm that even for lightly contaminated data this relation is reversed. Another conclusion clearly demonstrated is that for heavier contaminations and larger sample sizes the median estimator is more robust than any of the remaining three estimators.

ε	$(\widehat{\beta}_1, \widehat{\beta}_2)$	$MAE(400)$	NEF	$MAE(800)$	NEF	$MAE(1600)$	NEF
0	MLE	0.263	0	0.194	0	0.136	0
	Morg	0.295	0	0.211	0	0.149	0
	B&Y	0.365	0	0.249	0	0.173	0
	Median	1.105	7	0.671	0	0.404	0
0.05	MLE	1.102	0	1.137	0	1.145	0
	Morg	0.753	17	0.791	2	0.792	0
	B&Y	0.553	0	0.543	0	0.537	0
	Median	1.056	1	0.629	0	0.526	0
0.1	MLE	1.488	0	1.535	0	1.562	0
	Morg	2.470	1622	2.526	2086	3.153	3024
	B&Y	0.804	0	0.898	0	0.948	0
	Median	0.959	5	0.835	0	0.809	0
0.15	MLE	1.943	0	1.948	0	1.948	0
	Morg	-	-	-	-	-	-
	B&Y	1.637	0	1.654	0	1.657	0
	Median	1.395	1	1.364	0	1.402	0
0.2	MLE	2.161	0	2.168	0	2.168	0
	Morg	-	-	-	-	-	-
	B&Y	2.031	0	2.043	0	2.046	0
	Median	1.810	0	1.769	0	1.799	0

Table 2: The same as in Table 1 for $\Pr(Y = 1) = 0.3$ and $(\beta_{01}, \beta_{02}) = (-2.16, 3.71)$.

ε	$(\widehat{\beta}_1, \widehat{\beta}_2)$	$MAE(400)$	NEF	$MAE(800)$	NEF	$MAE(1600)$	NEF
0	MLE	0.264	0	0.187	0	0.137	0
	Morg	0.288	0	0.205	0	0.149	0
	B&Y	0.346	0	0.244	0	0.171	0
	Median	1.169	5	0.607	0	0.396	0
0.05	MLE	1.018	0	1.049	0	1.056	0
	Morg	0.686	19	0.718	1	0.718	0
	B&Y	0.510	0	0.492	0	0.484	0
	Median	1.028	4	0.614	0	0.489	0
0.1	MLE	1.378	0	1.413	0	1.439	0
	Morg	2.359	1643	2.347	2009	2.378	2845
	B&Y	0.742	0	0.814	0	0.871	0
	Median	0.888	2	0.736	0	0.729	0
0.15	MLE	1.786	0	1.793	0	1.793	0
	Morg	-	-	-	-	-	-
	B&Y	1.481	0	1.507	0	1.511	0
	Median	1.275	0	1.237	0	1.281	0
0.2	MLE	1.985	0	1.993	0	1.993	0
	Morg	-	-	-	-	-	-
	B&Y	1.856	0	1.869	0	1.870	0
	Median	1.563	0	1.622	0	1.638	0

Table 3: The same as in Table 1 for $\Pr(Y = 1) = 0.4$ and $(\beta_{01}, \beta_{02}) = (-1.16, 4.20)$.

ε	$(\widehat{\beta}_1, \widehat{\beta}_2)$	$MAE(400)$	NEF	$MAE(800)$	NEF	$MAE(1600)$	NEF
0	MLE	0.262	0	0.182	0	0.128	0
	Morg	0.283	0	0.196	0	0.139	0
	B&Y	0.342	0	0.229	0	0.159	0
	Median	0.963	4	0.559	0	0.341	0
0.05	MLE	0.886	0	0.893	0	0.887	0
	Morg	0.626	14	0.626	0	0.614	0
	B&Y	0.477	0	0.445	0	0.427	0
	Median	0.882	6	0.537	1	0.439	0
0.1	MLE	1.173	0	1.188	0	1.199	0
	Morg	1.684	1762	1.695	2004	1.579	2686
	B&Y	0.664	0	0.709	0	0.736	0
	Median	0.887	3	0.673	0	0.644	0
0.15	MLE	1.502	0	1.492	0	1.484	0
	Morg	-	-	-	-	-	-
	B&Y	1.263	0	1.262	0	1.258	0
	Median	1.129	0	1.070	0	1.084	0
0.2	MLE	1.662	0	1.654	0	1.646	0
	Morg	-	-	-	-	-	-
	B&Y	1.558	0	1.552	0	1.545	0
	Median	1.337	4	1.357	0	1.368	0

Table 4: The same as in Table 1 for $\Pr(Y = 1) = 0.5$ and $(\beta_{01}, \beta_{02}) = (0, 4.36)$.

References

[1] A. Agresti (2002). *Categorical Data Analysis* (Second Edition). Wiley, New York.

[2] E. B. Andersen (1990). *The Statistical Analysis of Categorical data.* Springer-Verlag, Berlin.

[3] M. A. Arcones (2001), Asymptotic distribution of regression M-estimators. *Journal of Statistical Plannning and Inference,* **97,** 235-261.

[4] A. M. Bianco and V. J. Yohai (1996), Robust estimation in the logistic regression model. *Robust Statistics, Data Analysis, and Computer Intensive Methods* (Schloss Thurnau, 1994), 17-34. Lecture Notes in Statistics, 109, Springer, New York.

[5] F. R. Hampel, P. J. Rousseeuw, E. M. Ronchetti, W. A. Stahel (1986), *Robust Statistics: The Approach Based on Influence Functions.* Wiley, New York.

[6] T. Hobza, L. Pardo and I. Vajda (2005): *Median Estimators in Generalized Logistic Regressio*n. Research Report No. 2153. Institute of Information Theory, Prague.

[7] J. Jurečková and P. K. Sen (1996), *Robust Statistical Procedures.* Wiley, New York.

[8] F. Liese and I. Vajda (2003), A general asymptotic theory of M-estimators, I. *Mathematical Methods of Statistics,* **12,** 454-477.

[9] F. Liese and I. Vajda (2004), A general asymptotic theory of M-estimators, II. *Mathematical Methods of Statistics,* **13,** 82-95.

[10] S. Morgenthaler (1992), Least-absolute-derivations fits for generalized linear models. *Biometrika,* **79**, 747-754.

[11] J. A. Pardo, L. Pardo and M. C. Pardo (2006), Testing in logistic regression models based on ϕ-divergences measures. *Journal of Statistical Planning and Inference*, **136,** 982-1006.

[12] D. Pollard (1991), Asymptotic for least absolute derivation regression estimators. *Econometric Theory* **7,** 186-199.

[13] G. D. Richardson and B. B. Bhattacharyya (1987), Consistent L_1-estimates in nonlinear regression for a noncompact parameter space. *Sankhya,* **49,** ser. A, 377-387.

[14] V. J. Yohai (1987), High breakdown point high efficiency robust estimates for regression. *Annals Statistics,* **15,** 692–656.

[15] S. Zwanzig (1997), *On L_1-norm Estimators in Nonlinear Regression and in Nonlinear Error-in-Variables Models.* IMS Lecture Notes 31, 101-118, Hayward.

Tomáš Hobza: Institute of Information Theory and Automation , Pod vodárenskou věží 4, Praha, 182 08, Czech Republic, *hobza@km1.fjfi.cvut.cz*

Leandro Pardo: Department of Statistics and O.R., Complutense University of Madrid, 28040, Madrid, Spain, *lpardo@mat.ucm.es*