Jana Jurečková

Asymptotic comparison of maximum likelihood and a rank estimate in simple linear regression model

Commentationes Mathematicae Universitatis Carolinae, Vol. 16 (1975), No. 1, 87--97

Persistent URL: `http://dml.cz/dmlcz/105607`

COMMENTATIONES MATHEMATICAE UNIVERSITATIS CAROLINAE

16,1 (1975)

ASYMPTOTIC COMPARISON OF MAXIMUM LIKELIHOOD AND A RANK ESTIMATE IN SIMPLE LINEAR REGRESSION MODEL

Jana JUREČKOVÁ, Praha

Abstract: In the simple linear regression model, two estimates of the regression parameter vector, a maximum likelihood and a robust rank one, are compared under the non-standard condition that the supposed distribution differs from the real one. The comparison is asymptotic for the number of observations increasing and is based on the asymptotic distribution of the difference of both estimates which is determined under some regularity conditions.

Key words: Simple linear regression model, maximum likelihood estimate, robust procedures, rank statistics, asymptotically normal distribution.

AMS: 62G05, 62F10, 62G35 Ref. Ž. 9.741

---

1. Introduction. For $N = 1,2,\dots,$ let $X_{N1},\dots,X_{NN}$ be independent observations such that X_{Ni} has the distribution function

$$(1.1) \qquad G(x_i - \sum_{j=1}^{p} \Delta_j c_{ji}) , \quad i = 1,\dots, N$$

where $\Delta^0 = (\Delta_1^0 \dots, \Delta_p^0)$ is an unknown parameter and c_{ji} , $j = 1,\dots, p$; $i = 1,\dots, N$ are given real numbers dependent on N ; the distribution function G is supposed to be unknown.

Let $\hat{\Delta}_1$ be the maximum likelihood estimate of Δ^0

computed under a false assumption that, instead of $G(x_i - \sum_{j=1}^{p} \Delta_j^0 c_{ji})$, the distribution function of X_{Ni} is equal to $F(x_i - \sum_{j=1}^{p} \Delta_j^0 c_{ji})$, $i = 1, \ldots, N$. It means that $\hat{\boldsymbol{\Delta}}_1$ is any solution of the following system of equations

$$(1.2) \quad M_j(\boldsymbol{X}_N, \boldsymbol{\Delta}) = \sum_{i=1}^{N} c_{ji}\, \psi(X_{Ni} - \sum_{\ell=1}^{p} \Delta_\ell c_{\ell i}) = 0 ,$$

$$j = 1, \ldots, p$$

where

$$(1.3) \quad \psi(x) = -\frac{f'(x)}{f(x)} , \quad x \in R^1 , \quad f'(x) = \frac{dF(x)}{dx} .$$

Let $\hat{\boldsymbol{\Delta}}_2$ be the rank estimate of $\boldsymbol{\Delta}^0$, suggested by the author in [3], being also determined under the assumption that F is the underlying distribution; i.e. $\hat{\boldsymbol{\Delta}}_2$ is any solution of the minimization problem

$$(1.4) \quad \sum_{j=1}^{p} |S_j(\boldsymbol{X}_N, \boldsymbol{\Delta})| = \min , \quad \text{where}$$

$$(1.5) \quad S_j(\boldsymbol{X}_N, \boldsymbol{\Delta}) = \sum_{i=1}^{N} c_{ji}\, \psi(F^{-1}(\frac{R_{Ni}^{\boldsymbol{\Delta}}}{N+1})) ,$$

$$j = 1, \ldots, p$$

with $R_{Ni}^{\boldsymbol{\Delta}}$ being the rank of $X_{Ni} - \sum_{j=1}^{p} \Delta_j c_{ji}$ among

$X_{N1} - \sum_{j=1}^{p} \Delta_j c_{j1}, \ldots, X_{NN} - \sum_{j=1}^{p} \Delta_j c_{jN}$, i.e.

$$(1.6) \qquad R^{\Delta}_{Ni} = \sum_{k=1}^{N} u(X_{Ni} - X_{Nk} - \sum_{j=1}^{n} \Delta_j (c_{ji} - c_{jk}))$$

where $u(x) = 1$ if $x \geq 0$ and $u(x) = 0$ if $x < 0$. Further, in (1.5),

$$F^{-1}(u) = \inf\{x: F(x) \geq u\} , \quad 0 < u < 1 .$$

It is known that, if $F \equiv G$, both estimators are asymptotically efficient for $N \to \infty$. Nevertheless, the estimators differ e.g. in the number of operations needed for their computation, in the speed in which they become asymptotically efficient and in their robustness with respect to the individual pairs F , G . It is the purpose of the present study to make an asymptotic comparison for $N \to \infty$ of the estimators with respect to their robustness. There exist the pairs F , G such that the asymptotic variance of $\hat{\Delta}_1$ is infinite while the asymptotic variance of $\hat{\Delta}_2$ is always finite. On the other hand, $\hat{\Delta}_1$ and $\hat{\Delta}_2$ may become asymptotically equivalent in probability even if $F \not\equiv G$.

It is shown that the sequence of differences $\{\hat{\Delta}_1 - \hat{\Delta}_2\}_{N=1}^{\infty}$ has for $N \to \infty$ an asymptotically normal distribution, generally non-degenerate. Some special cases in which this distribution becomes degenerate are indicated (it trivially happens if $F \equiv G$). More general consequences of the form of the asymptotic covariance matrix are still an open problem and are a subject of study.

2. Assumptions, the main result. We shall study the

asymptotic behavior of $\hat{\Delta}_1 - \hat{\Delta}_2$ under the following assumptions, some of which mean no loss of generality while other mean the real restriction:

1^o <u>Assumptions on</u> c_{ji}. Let $\mathbf{C}_N = [c_{ji}]$, $j = 1, \ldots, p$ and $i = 1, \ldots, N$ be a $p \times N$ design matrix with the rows $\mathbf{c}_{(j)}$ and columns $\mathbf{c}^{(i)}$ satisfying the conditions

(a) $\sum_{i=1}^{N} c_{ji} = 0$, $j = 1, \ldots, p$

(b) $c_{ji} = c'_{ji} + c''_{ji}$, $j = 1, \ldots, p$; $i = 1, \ldots, N$

(c) The vectors $\mathbf{c}'_{(j)} = (c'_{j1}, \ldots, c'_{jN})$, $j = 1, \ldots, p$ satisfy either

$$(2.1) \qquad (\mathbf{c}'_{(j)} - \bar{c}'_j)(\mathbf{c}'_{(j)} - \bar{c}'_j)^T = 0$$

for all but a finite number of N, or

$$(2.2) \qquad (\mathbf{c}'_{(j)} - \bar{c}'_j)(\mathbf{c}'_{(j)} - \bar{c}'_j)^T > 0$$

for all but a finite number of N; further

$$(2.3) \qquad (\mathbf{c}'_{(j)} - \bar{c}'_j)(\mathbf{c}'_{(j)} - \bar{c}'_j)^T \leq M \quad \text{for} \quad N = 1, 2, \ldots$$

and if (2.2) holds, then

$$(2.4) \qquad \lim_{N \to \infty} \left\{ \max_{1 \leq i \leq N} (c'_{ji} - \bar{c}'_j)^2 \left[\sum_{i=1}^{N} (c'_{ji} - \bar{c}'_j)^2 \right] \right\} = 0;$$

here $\bar{c}'_j = N^{-1} \sum_{i=1}^{N} c'_{ji}$ and $M > 0$ is a constant independent of N.

Analogous assumptions are to be satisfied for vectors $\mathbf{c}''_{(j)}$, $j = 1, \ldots, p$.

(d) It holds for all pairs $j, \ell = 1,\dots, p$ and $i, k = 1,\dots, N$; $N = 2,3,\dots$ that

$$(2.5)\qquad (c'_{ji} - c'_{jk})\,(c'_{\ell i} - c'_{\ell k}) \geq 0$$

$$(c'_{ji} - c'_{jk})\,(c''_{\ell i} - c''_{\ell k}) \leq 0$$

$$(c''_{ji} - c''_{jk})\,(c''_{\ell i} - c''_{\ell k}) \geq 0 \,.$$

(e) $\lim_{N\to\infty} \mathbb{C}_N \mathbb{C}_N^T = \boldsymbol{\Sigma}$ exists and is a positively definite matrix.

2^0 <u>Assumptions on distributions</u> F <u>and</u> G .

(a) Let f and g be the respective densities corresponding to F and G and suppose that both f and g are absolutely continuous and have finite Fisher's informations, i.e.

$$(2.7)\qquad I(f) = \int_{-\infty}^{\infty} \left[\frac{f'(x)}{f(x)}\right]^2 f(x)\,dx < \infty$$

$$I(g) = \int_{-\infty}^{\infty} \left[\frac{g'(x)}{g(x)}\right]^2 g(x)\,dx < \infty$$

and moreover, that f is unimodal, i.e. $(-\log f(x))$ is convex in x .

Let us denote

$$(2.8)\qquad \gamma = -\int_{-\infty}^{\infty} \psi\left[F^{-1}(G(x))\right] g'(x)\,dx$$

$$(2.9)\qquad \omega = -\int_{-\infty}^{\infty} \psi(x)\, g'(x)\,dx$$

and

$$(2.10) \qquad \rho^2 = \int_{-\infty}^{\infty} \psi^2(x)\ g(x)\ dx - \left(\int_{-\infty}^{\infty} \psi(x)\ g(x)\ dx\right)^2$$

where ψ is given by (1.3).
Suppose that $\rho^2 < \infty$.

Under the assumptions 1° and 2°, we have the following theorem.

Theorem 2.1. Under the assumptions 1° and 2°, for $N \longrightarrow \infty$,

$$(2.11) \qquad \mathcal{L}(\hat{\Delta}_1 - \hat{\Delta}_2) \longrightarrow$$

$$\mathcal{N}_p(\mathbb{0}, \int_{-\infty}^{\infty} [\omega^{-1} \frac{f'(x)}{f(x)} - \gamma^{-1} \frac{f'(F^{-1}(G(x)))}{f(F^{-1}(G(x)))}]^2 g(x)\,dx \,.\, \Sigma^{-1})$$

where $\mathcal{N}_p(\mathbf{a}, \mathbb{A})$ denotes the p-dimensional normal distribution with the expectation $\mathbf{a}$ and the covariance matrix $\mathbb{A}$.

3. Sketch of the proof of Theorem 2.1. Here we shall only sketch the main idea of the proof. A more detailed proof together with other results will appear in a more comprehensive study, being now prepared.

First of all, we may suppose without loss of generality that $c''_{ji} = 0$ for all i , j , so that c_{ji} equals to c'_{ji} . In view of (2.5) and the unimodality of f , we have

Lemma 3.1. If f is unimodal, and if $c''_{ji} = 0$ for $i = 1,\dots, N$ and $j = 1,\dots, p$ then $M_j(\mathcal{X}_N, \Delta)$ and $S_j(\mathcal{X}_N, \Delta)$, $j = 1,\dots, p$ are non-increasing in Δ_ℓ ,

$\ell = 1, \ldots, p$, with probability one.

Otherwise, if $c''_{ji} \neq 0$, both $M_j(\mathcal{X}_N, \Delta)$ and $S_j(\mathcal{X}_N, \Delta)$ could be written down as the differences of two functions, non-increasing in each component.

Asymptotic properties of $S_j(\mathcal{X}_N, \Delta)$ and of $\hat{\Delta}_2$ were investigated by Hájek-Šidák in [1] and by Jurečková in [3] in details. We shall thus restrict our attention to the asymptotic properties of $M_j(\mathcal{X}_N, \Delta)$. We have the lemma

<u>Lemma 3.2</u>. Put $\Delta^0 = 0$. For a fixed ℓ , $1 \leq \ell \leq p$, let $A_\ell = \{\Delta : \Delta_k = 0 \text{ for } k \neq \ell\}$. Then for any fixed $\Delta \in A_\ell$ and for $N \to \infty$ it holds that

$$(3.1) \quad \mathcal{L}\{M_j(\mathcal{X}_N, \Delta)\} \longrightarrow \mathcal{N}_1(-\Delta_\ell \, \omega \, \sigma_{j\ell} \, , \, \varrho^2 \sigma_{jj}) \, ,$$

$$j = 1, \ldots, p \, ;$$

$\sigma_{j\ell}$ and σ_{jj} are elements of Σ .

The proof of Lemma 3.2 is based on the concept of contiguity and on three Le Cam's lemmas (for the definition of contiguity and Le Cam's lemmas see Hájek-Šidák [1]).

The following theorem is substantial for treating the asymptotic behavior of $\hat{\Delta}_1$. It tells that $M_j(\mathcal{X}_N, \Delta)$ are uniformly asymptotically linear in Δ in the sense of the convergence in probability. Analogous theorems for $S_j(\mathcal{X}_N, \Delta)$ were proved by the author in [2] and [3].

<u>Theorem 3.1</u>. Under the assumptions 1^0 and 2^0,

(3.2) $$\lim_{N\to\infty} P_{\Delta^0}\left\{\max_{\|\Delta-\Delta^0\|\le C} |M_j(X_N, \Delta) - M_j(X_N, \Delta^0) + \omega(\Delta - \Delta^0)\sigma^{(j)}| \ge \varepsilon\right\} = 0$$

holds for any $C > 0$, $\varepsilon > 0$ and $j = 1,\dots, p$; $\sigma^{(j)}$ is the j-th column of Σ .

For the proof of Theorem 3.1, the non-uniform version of (3.2) is proved for any fixed Δ by approximating ψ by a sequence of bounded functions; the uniformity is then a consequence of Lemma 3.1.

Theorem 3.1 has an easy corollary which yields an approximation of $\hat{\Delta}_1$ by a sum of independent random variables.

Corollary 1. Let $\tilde{\Delta}_N$ be any sequence of random vectors such that $\tilde{\Delta}_N - \Delta^0$ are bounded in probability. Then under the assumptions 1^o and 2^o

(3.3) $$\lim_{N\to\infty} P_{\Delta^0}\left\{|M_j(X_N, \tilde{\Delta}_N) - M_j(X_N, \Delta^0) + \omega(\tilde{\Delta}_N - \Delta^0)\sigma^{(j)}| \ge \varepsilon\right\} = 0$$

holds for $j = 1,\dots, p$ and any $\varepsilon > 0$.

For being able to apply the corollary to $\tilde{\Delta}_N = \hat{\Delta}_1$,we need the following lemma which may be proved by help of Theorem 3.1:

Lemma 3.3. Under 1^o and 2^o, there exist $C^* > 0$, $\eta > 0$ and a positive integer N_0 corresponding to any fixed $\varepsilon > 0$, such that

(3.4) $$P_{\Delta^0}\left\{\min_{\|\Delta-\Delta^0\|\ge C^*} \|M(X_N, \Delta)\| < \eta\right\} < \varepsilon$$

holds for $N > N_0$, where $\mathbf{M}(\mathcal{X}_N, \Delta) = (M_1(\mathcal{X}_N, \Delta), \ldots$
$\ldots, M_p(\mathcal{X}_N, \Delta))$.

The lemmas and corollaries then enable to approximate the differences $\hat{\Delta}_1 - \hat{\Delta}_2$ by random vectors to which the classical central limit theorem is easily applicable. This completes the proof of Theorem 2.1.

As a consequence of the proof of Theorem 2.1, we get the following corollary

Corollary 2. Under 1^o and 2^o, the estimate $\hat{\Delta}_1$ is for $N \to \infty$ asymptotically normal $\mathcal{N}_p(\Delta^0, \omega^{-2}\rho^2 \Sigma^{-1})$.

4. Examples.

4.1. Let $f(x) = \frac{1}{2} \exp\{-|x|\}$, $x \in R^1$ (double exponential distribution)

$g(x) = (1/\sqrt{2\pi}) \exp\{-\frac{x^2}{2}\}$, $x \in R^1$ (standard normal distribution)

Then $\hat{\Delta}_1 - \hat{\Delta}_2 \longrightarrow 0$ in probability for $N \to \infty$.

Actually, we have

$$\psi(x) = \operatorname{sign} x$$

$$F(x) = 1 - \frac{1}{2} e^{-x} \quad \text{if} \quad x \geq 0$$

$$= 1 - F(-x) \quad \text{if} \quad x < 0$$

$$F^{-1}(u) = -\log(2(1-u)) \quad \text{if} \quad \frac{1}{2} \leq u < 1$$

$$= -F^{-1}(1-u) \quad \text{if} \quad 0 < u < \frac{1}{2}$$

so that

$$\omega = (1/\sqrt{2\pi})\int_{-\infty}^{\infty} |x| \exp\{-\frac{x^2}{2}\}\, dx = \sqrt{\frac{2}{\pi}}$$

$$\gamma = (1/\sqrt{2\pi})\int_{-\infty}^{\infty} x \operatorname{sign}(2F(x)-1) \exp\{-\frac{x^2}{2}\}\, dx = \sqrt{\frac{2}{\pi}}$$

and the factor in the asymptotic covariance matrix is equal to

$$\int_{-\infty}^{\infty} [\omega^{-1}\psi(x) - \gamma^{-1}\psi(F^{-1}(G(x)))]^2\, g(x)\, dx =$$

$$= (1/\sqrt{2\pi}) \cdot \frac{\pi}{2}\int_{-\infty}^{\infty} [\operatorname{sign} x -$$

$$- \operatorname{sign}(2F(x)-1)]^2 \exp\{-\frac{x^2}{2}\}\, dx = 0 .$$

4.2. Let $F(x) = \Phi(\frac{x}{\sigma_1})$ and $G(x) = \Phi(\frac{x}{\sigma_2})$ where Φ is the standard normal distribution function and $\sigma_1, \sigma_2 > 0$. Then $\hat{\Delta}_1 - \hat{\Delta}_2 \rightarrow 0$ in probability for $N \rightarrow \infty$.

Actually, $\omega = \sigma_1^{-2}$, $\gamma = (\sigma_1\ \sigma_2)^{-1}$ and $\psi(x) = x\,\sigma_1^{-2}$ so that

$$\int_{-\infty}^{\infty} [\omega^{-1}\psi(x) - \gamma^{-1}\psi(F^{-1}(G(x)))]^2\, g(x)\, dx = 0 .$$

4.3. Let $f(x) = (1/\sqrt{2\pi}) \exp\{-\frac{x^2}{2}\}$, $x \in R^1$

$$g(x) = 1 \quad \text{if } |x| \leq \frac{1}{4}$$

$$= 1/(16\,x^2) \quad \text{if } |x| > \frac{1}{4} .$$

Then $\varrho^2 = \int x^2 g(x)\, dx = \infty$ so that $\hat{\Delta}_1$ has infinite asymptotic variance. On the other hand, the asymptotic variance of $\hat{\Delta}_2$ is always finite (see Jurečková [3]).

R e f e r e n c e s

[1] J. HÁJEK, Z. ŠIDÁK: Theory of rank tests, Academia, Praha,1967.

[2] J. JUREČKOVÁ: Asymptotic linearity of a rank statistic in regression parameter, Ann.Math.Statist.40 (1969),1889-1900.

[3] J. JUREČKOVÁ: Nonparametric estimate of regression coefficients , Ann.Math.Statist.42(1971),1328-1338.

[4] E.L. LEHMANN: Some concepts of dependence, Ann.Math. Statist.37(1966),1137-1153.

Matematicko-fyzikální fakulta

Karlova universita

Sokolovská 83, 18600 Praha 8

Československo

(Oblatum 19.12.1974)