Theory of Stochastic Processes
Vol.13 (29), no.1-2, 2007, pp.122-131

ANDRII MALENKO

EFFICIENCY COMPARISON OF TWO CONSISTENT ESTIMATORS IN NONLINEAR REGRESSION MODEL WITH SMALL MEASUREMENT ERRORS

We study a nonlinear measurement model where the response variable has a density belonging to the exponential family. We consider two consistent estimators: Corrected Score (CS) and Quasi Score (QS) ones. Their relative efficiency is compared with respect to asymptotic covariance matrices. We derive expansions of these matrices for small error variances. It is shown that the QS estimator is more efficient than the CS one. The polynomial and Poisson regression models are studied in more detail.

1. Introduction

In this paper we consider general nonlinear regression model with errors in the variables, where the response variable has a density belonging to the exponential family. It is well known that ignoring measurement error leads to inconsistent estimators. We consider two *consistent* estimators: the Corrected Score (CS) one and the Quasi Score (QS) one.

There is a number of papers dealing with these estimators. Kukush *et al.* (2006) prove that asymptotic covariance matrix (ACM) of QS estimator is not greater than ACM of CS one in Lowener order, and give conditions for strict inequality. In Kukush and Schneeweiss (2005) it is proved that the ACMs are equal up to $O(\sigma_\delta^4)$, where σ_δ^2 is error variance tending to zero. The goal of this paper is to compare the terms of expansions of order σ_δ^4.

We denote by $\mathbf{E}$ the expectation of random values, vectors, or matrices, $\mathbf{V}$ stands for the variance. The expectation $\mathbf{E}f(z,\beta)$ is taken under the same parameter β of the distribution of z as the β of the argument of f unless otherwise specified. Derivatives are denoted as subindexes, vector derivatives are column vectors of partial derivatives. The sign ^{t} means transposition, the symmetrization operation $[A]_S := A + A^t$ makes sense for

2000 *Mathematics Subject Classification:* 62J10, 62J02, 62J12, 62F12.
Key words and phrases. Errors-in-variables models, corrected score, quasi score.

square matrices. We denote convergence in distribution of random vectors by $\xrightarrow{d}$.

The paper is organized as follows. In the next section the model is described. Section 3 introduces the estimators. In Section 4 we derive expansions of the difference of ACMs. In Section 5 and 6 we consider two particular models and Section 7 concludes.

2. General model

Let $(\Omega,\ \mathcal{F},\ \mathbf{E})$ be a probability space. We study a nonlinear errors-in-variables model, as considered in Kukush and Schneeweiss (2005). Let ν be a σ-finite measure on Borel σ-field on $\mathbb{R}$. We observe a random variable y with conditional density $f(y|\eta)$ with respect to the measure ν. The density belongs to an exponential family,

$$f(y|\eta) = \exp\left\{\frac{y\eta - C(\eta)}{\varphi} + c(y,\varphi)\right\}. \tag{1}$$

The $C(\cdot)$ function is smooth enough, $C'' > 0$, and $c(y,\varphi)$ is measurable and does not depend of η. The parameter $\varphi > 0$ is the dispersion parameter of y, it is supposed to be known.

Assume that $\eta = \eta(\xi,\beta)$, where ξ is a random latent regressor, and β is unknown parameter vector. We observe noisy variable $x = \xi + \delta$, where ξ and δ are independent. δ is called measurement error.

Let for $i = 1,\dots,n$ random vectors (y_i,ξ_i,δ_i) be i.i.d., $\xi_i \sim \mathcal{N}(\mu_\xi,\,\sigma_\xi^2)$, $\delta_i \sim \mathcal{N}(0,\,\sigma_\delta^2)$, where parameters μ_ξ, σ_ξ and σ_δ are known, $\sigma_\xi > 0$, $\sigma_\delta > 0$, and ξ_i, δ_i are independent.

Suppose that $\beta \in \Theta$, where Θ is a compact set in $\mathbb{R}^k$. Vector β is to be estimated based on observations $(y_i, x_i),\ i = 1,\dots,n$.

Introduce the following smoothness assumptions.

(i) The true value of β is an interior point of the set Θ.

(ii) $C(\cdot) \in C^{(6)}(\mathbb{R})$, and there exist constants $A, B > 0$ such that

$$\forall\xi\in\mathbb{R}\ \forall\beta\in\Theta:\quad \left|C^{(i)}(\eta(\xi,\beta))\right| \le A\cdot e^{B|\xi|},\quad i = 1,\dots,6.$$

(iii) $\eta(\cdot,\,\cdot) \in C^{(4,1)}(\mathbb{R}\times\Theta)$, and there exist constants $A, B > 0$ such that

$$\forall\xi\in\mathbb{R}\ \forall\beta\in\Theta:\ \left\|\frac{\partial^{i+j}}{\partial\xi^i\,\partial\beta^j}\eta(\xi,\beta)\right\| \le A\cdot e^{B|\xi|},\ i = 0,\dots,4,\ j = 0,1.$$

3. Estimators

Several consistent estimators of β are proposed in the literature, see Carroll *et al.* (1995). We will consider and compare the Corrected Score (CS) and the Quasi Score (QS) ones.

The Quasi Score method is based on conditional expectation and conditional variance of response variable y given x:

$$m(x,\beta) := \mathbf{E}(y|x) = \mathbf{E}[C'(\eta(\xi,\beta))|x],$$

$$v(x,\beta) := \mathbf{V}(y|x) = \mathbf{V}[C'(\eta(\xi,\beta))|x] + \varphi\mathbf{E}[C''(\eta(\xi,\beta))|x].$$

Estimator $\hat{\beta}_Q$ is defined as a measurable solution to the equation

$$\sum_{i=1}^{n} \frac{y_i - m(x_i,\beta)}{v(x_i,\beta)} \cdot m_\beta(x_i,\beta) = 0. \tag{2}$$

We will say that for a sequence of random variables $\{U_n : \ n \geq 1\}$ a sequence of statements $A_n(U_n(\omega))$, $\omega \in \Omega$, holds *eventually*, if

$$\exists\,\Omega_0 \subset \Omega,\ \mathbf{P}(\Omega_0) = 1,\ \forall\,\omega \in \Omega_0\ \exists\,N = N(\omega)\ \forall\,n \geq N : \quad A_n(U_n(\omega)) \text{ holds.}$$

Consider the following assumptions.

(iv) For some $A, B > 0$, $\forall \xi \in \mathbb{R}\ \forall \beta \in \Theta : C''(\eta(\xi,\beta)) \geq A \cdot e^{-B|\xi|}$.

(v) The equation $\mathbf{E}[v^{-1}(m_0 - m)m_\beta] = 0$, $\beta \in \Theta$, has the only solution $\beta = \beta_0$. Here β_0 is the true value of parameter β, $m_0 := m(x,\beta_0)$, $m = m(x,\beta)$ and $v = v(x,\beta)$.

(vi) The matrix $\mathbf{E}\, m_\beta m_\beta^t$ is positive definite at the true point $\beta = \beta_0$.

Theorem 1. *Let conditions* (i) *to* (vi) *hold true. Then:*

a) eventually, the equation (2) has a solution $\hat{\beta}_Q \in \Theta$*;*

b) eventually, the solution to the equation (2) is unique;

c) the estimator $\hat{\beta}_Q$ *is strictly consistent, i.e.*

$$\hat{\beta}_Q \to \beta \text{ a.s., as } n \to \infty.$$

The theorem is proved in Kukush and Schneeweiss (2005). The next statement about the asymptotic normality is also proved there.

Theorem 2. *Let conditions* (i) *to* (vi) *hold. Then* $\hat{\beta}_Q$ *is asymptotically normal with ACM* $\Sigma_Q = \Phi^{-1}$*, where*

$$\Phi = \mathbf{E}\,\frac{m_\beta(x,\beta)m_\beta^t(x,\beta)}{v(x,\beta)}.$$

To define the Corrected Score we consider the likelihood score function in the error-free model. Denote

$$\psi(y,\xi,\beta) = y\eta_\beta - C'(\eta)\eta_\beta, \tag{3}$$

where η and the derivative η_β are taken at the point (ξ,β). To find the ML estimator of β by observations (y_i,ξ_i), $i=1,\dots,n$, one should solve the equation

$$\frac{1}{n}\sum_{i=1}^{n}\psi(y_i,\xi_i,\beta) = 0, \quad \beta\in\Theta.$$

Consider the limit equation

$$\mathbf{E}\left[(C'(\eta(\xi,\beta_0)) - C'(\eta(\xi,\beta)))\eta(\xi,\beta)\right] = 0, \quad \beta\in\Theta, \tag{4}$$

where β_0 is the true value of parameter β.

Assume the following identifiability condition for the error-free model.

(vii) The equation (4) has unique solution $\beta=\beta_0$.

We introduce the corrected score function $\psi_c(y,x,\beta)$ such that

$$\mathbf{E}(\psi_c(y,x,b)|y,\xi) = \psi(y,\xi,b), \quad b\in\Theta.$$

Denote $f_1(x,\beta)=\eta_\beta(x,\beta)$, $f_2(x,\beta) = C'(\eta(x,\beta))\eta_\beta(x,\beta)$. We search for such functions $f_{ic}(x,\beta)$, $i=1,2$, that

$$\mathbf{E}(f_{ic}(x,\beta)|\xi) = f_i(\xi,\beta), \quad i=1,2. \tag{5}$$

Then $\psi_c(y,x,\beta) = yf_{1c}(x,\beta) - f_{2c}(x,\beta)$.

Suppose that

(viii) Functions f_{ic} in (5) are defined in a neighborhood of Θ.

(ix) For small enough σ_δ the following relations hold true:

$$\left\|\frac{\partial^j}{\partial\beta^j}f_{ic} - \left(\frac{\partial^j}{\partial\beta^j}f_i - \frac{1}{2}\sigma_\delta^2\frac{\partial^j}{\partial\beta^j}(f_i)_{xx} + \frac{1}{8}\sigma_\delta^4\frac{\partial^j}{\partial\beta^j}(f_i)_{x^4}\right)\right\| \le C\cdot e^{A|x|}\sigma_\delta^6,$$

for $i=1,2$ and $j=0,1$ and for some fixed $A>0$, $C=const$.

The last condition holds true for the polynomial and Poisson models. It is closely related to a series expansion of the solution to the deconvolution problem like (5), which is presented in Stefanski (1989).

The Corrected Score estimator $\hat\beta_C$ is defined as a solution to the equation

$$\frac{1}{n}\sum_{i=1}^{n}\psi_c(y_i,x_i,\beta) = 0, \quad \beta\in\Theta.$$

Under $n \to \infty$ we get exactly the equation (4).

Asymptotic properties of $\hat{\beta}_C$ are studied in Kukush and Schneeweiss (2005). Under conditions (vii) to (ix), $\hat{\beta}_C$ is strictly consistent and asymptotically normal, its ACM is given by the sandwich formula

$$\Sigma_C = A_c^{-1} \cdot B_c \cdot A_c^{-1},$$

where matrices A_c and B_c are

$$A_c = \mathbf{E}\, C''(\eta)\eta_\beta\eta_\beta^t, \quad \eta = \eta(\xi,\beta), \quad B_c = \mathbf{E}\, \psi_c(y,x,\beta)\psi_c^t(y,x,\beta).$$

4. Approximation of Σ_C and Σ_Q

A reader can find the exact comparison of Σ_Q and Σ_C in Kukush *et al.* (2006). In Kukush and Schneeweiss (2005) it is proved that under conditions (i) to (ix),

$$\Sigma_Q - \Sigma_C = O(\sigma_\delta^4), \text{ as } \sigma_\delta^2 \to 0.$$

That is, for small σ_δ^2 the asymptotic efficiency of these estimators is approximately equal up to $O(\sigma_\delta^4)$.

Under stronger conditions on $C(\eta)$ and $\eta(\xi,\beta)$, we can find further terms of expansion of Σ_Q and Σ_C.

Theorem 3. *Let conditions* (i) *to* (ix) *hold and the next condition holds as well.*

(x) *The matrix* $S = \mathbf{E}C''\eta_\beta\eta_\beta^t$ *is positive definite.*

Then under $\sigma_\delta^2 \to 0$ *we have*

$$\Sigma_C - \Sigma_Q = \sigma_\delta^4\, S^{-1} \cdot \Delta \cdot S^{-1} + O(\sigma_\delta^6), \tag{6}$$

where

$$\begin{aligned}
\varphi\Delta &= \mathbf{E}C''^3\eta_x^4\eta_\beta\eta_\beta^t - \mathbf{E}[C''^2\eta_x^2\eta_\beta\eta_\beta^t]S^{-1}\mathbf{E}[C''^2\eta_x^2\eta_\beta\eta_\beta^t] \\
&+ \varphi\mathbf{E}\Big(\frac{1}{\sigma_x^2}C''^2\eta_x^2\eta_\beta\eta_\beta^t + {C^{(3)}}^2\eta_x^4\eta_\beta\eta_\beta^t + 2C''C^{(3)}\eta_x^3[\eta_{x\beta}\eta_\beta^t]_S \\
&+ 3C''^2\eta_x^2\eta_{x\beta}\eta_{x\beta}^t + 2C''C^{(3)}\eta_x^2\eta_{xx}\eta_\beta\eta_\beta^t + 3C''^2\eta_x\eta_{xx}[\eta_{x\beta}\eta_\beta^t]_S \\
&+ C''^2\eta_{xx}^2\eta_\beta\eta_\beta^t - \mathbf{E}[C''^2\eta_x^2\eta_\beta\eta_\beta^t]S^{-1}\mathbf{E}[C''\eta_{x\beta}\eta_{x\beta}^t] \\
&- \mathbf{E}[C''\eta_{x\beta}\eta_{x\beta}^t]S^{-1}\mathbf{E}[C''^2\eta_x^2\eta_\beta\eta_\beta^t]\Big) \\
&+ \varphi^2\mathbf{E}\Big(\frac{1}{\sigma_x^2}C''\eta_{x\beta}\eta_{x\beta}^t - C^{(3)}\eta_{xx}\eta_{x\beta}\eta_{x\beta}^t + C''\eta_{xx\beta}\eta_{xx\beta}^t \\
&+ \left(\frac{{C^{(3)}}^2}{C''} - C^{(4)}\right)\eta_x^2\eta_{x\beta}\eta_{x\beta}^t - \mathbf{E}[C''\eta_{x\beta}\eta_{x\beta}^t]S^{-1}\mathbf{E}[C''\eta_{x\beta}\eta_{x\beta}^t]\Big).
\end{aligned} \tag{7}$$

Here we calculate the function η and its derivatives at the point (x,β), and the function C and its derivatives at the point $\eta(x,\beta)$, where β is the true parameter.

Proof. The idea of the proof is to approximate each ACM with summands of similar structure. These will be the expectations of products of functions C, η and their derivatives at the point (x, β).

4.1°. *Approximation of* Σ_Q. We approximate functions $m(x, \beta)$ and $v(x, \beta)$. We note that they can be expressed in terms of summands like $\mathbf{E}[f(\xi, \beta)|x]$, where function $f(\xi, \beta)$ and its derivatives are bounded by $Ce^{B|\xi|}$ uniformly for all $\beta \in \Theta$ because of conditions (ii) and (iii).

Since $\xi|x \sim \mathcal{N}(\mu(x), \tau^2)$, where $\mu(x) = x - \frac{\sigma_\delta^2}{\sigma_x^2}(x - \mu)$, $\tau^2 = \sigma_\delta^2 - \frac{\sigma_\delta^4}{\sigma_x^2}$, we have for $\gamma \sim \mathcal{N}(0, 1)$, γ is independent of x, that

$$\mathbf{E}[f(\xi, \beta)|x] = \mathbf{E}[f(\mu(x) + \tau\gamma, \beta)|x] = \mathbf{E} f(\mu(t) + \tau\gamma, \beta)\Big|_{t=x}.$$

Denote $\alpha = (x - \mu)\sigma_x^{-2}$. Expanding the function $f(\mu(x) + \tau\gamma, \beta)$ into the Taylor series near the x point and taking expectation w.r.t. γ, we have:

$$\begin{aligned}&\mathbf{E}[f(\mu(x) + \tau\gamma, \beta)|x] = f(x, \beta) - \alpha\sigma_\delta^2 f_x(x, \beta)\\ &+\frac{\tau^2 + \alpha^2\sigma_\delta^4}{2} f_{xx}(x, \beta) - \frac{\alpha\sigma_\delta^4}{2} f_{xxx}(x, \beta) + \frac{\sigma_\delta^4}{8} f_{x^4}(x, \beta) + r(x, \beta, \sigma_\delta),\end{aligned} \tag{8}$$

and there exists constant A such that for all $\beta \in \Theta$ and for small enough σ_δ^2, $\mathbf{E}\,|r(x, \beta, \sigma_\delta)| \le A\sigma_\delta^6$. Here the expectation is taken w.r.t. $x \sim N(\mu, \sigma_x^2)$.

To approximate $m(x, \beta)$ we use (8) with $f(x, \beta) = C'(\eta)$, $\eta = \eta(x, \beta)$. We rewrite $v(x, \beta) = A_1(x, \beta) + \varphi A_2(x, \beta)$. Here

$$A_1(x, \beta) = \mathbf{E}[C'^2(\eta(\xi, \beta))|x] - m^2(x, \beta), \quad A_2(x, \beta) = \mathbf{E}[C''(\eta(\xi, \beta))|x].$$

We use (8) with $f(x, \beta) = C'^2(\eta)$ and $f(x, \beta) = C''(\eta)$, respectively.

Because of (iv) the random variable $v^{-1}(x, \beta)$ is well-defined and uniformly in $\beta \in \Theta$ bounded from above by $const \cdot e^{B|x|}$. The random matrix $m_\beta(x, \beta)m_\beta^t(x, \beta)$ is also majorized by $const \cdot e^{B|x|}$ uniformly in β.

In approximation of Φ we have summands of the form $\mathbf{E}\alpha^k h(x, \beta)$, $k = 1, 2$. Here the function $h(x, \beta)$ satisfies conditions (ii) and (iii). To transform these summands we use the partial integration formulae:

$$\mathbf{E}\alpha h(x, \beta) = \mathbf{E} h_x(x, \beta), \quad \mathbf{E}\alpha^2 h(x, \beta) = \sigma_x^{-2}\mathbf{E} h(x, \beta) + \mathbf{E} h_{xx}(x, \beta).$$

Summarizing we have:

$$\Phi = \varphi S + \frac{\sigma_\delta^2}{2} Q + \frac{\sigma_\delta^4}{8} T + O(\sigma_\delta^6), \text{ as } \sigma_\delta^2 \to 0.$$

To invert Φ we use the following expansion: as $\delta \to 0$,

$$(A - \delta B + \delta^2 C)^{-1} = A^{-1} + \delta A^{-1}BA^{-1} + \delta^2 A^{-1}(BA^{-1}B - C)A^{-1} + O(\delta^3), \tag{9}$$

which holds true for all square matrices A, B, C of the same size, where A is nonsingular. Based on (x), we use (9) for $A = \varphi^{-1}S$, $B = \frac{1}{2}Q$, $C = \frac{1}{8}T$, $\delta = \sigma_\delta^2$. We have

$$\Sigma_Q = \varphi S^{-1} + \frac{\sigma_\delta^2}{2}\varphi^2 S^{-1}QS^{-1} + \frac{\sigma_\delta^4}{8}\varphi^2 S^{-1}(2\varphi QS^{-1}Q - T)S^{-1} + O(\sigma_\delta^6).$$

4.2°. *Approximation of* Σ_C. To expand A_c we use the next general result. Let the function $g(x,\beta)$ satisfy the condition

$$\left|\frac{\partial^i}{\partial x^i} g(x,\beta)\right| \le const \cdot e^{C_2|x|}, \quad i = 0, \dots, 6,$$

with some positive constant C_2, which may depend on β, $\xi \sim N(\mu, \sigma_\xi^2)$ and $\delta \sim N(0, \sigma_\delta^2)$ are independent, $x = \xi + \delta$. Then

$$\mathbf{E}g(\xi,\beta) = \mathbf{E}g(x,\beta) - \frac{\sigma_\delta^2}{2}\mathbf{E}g_{xx}(x,\beta) + \frac{\sigma_\delta^4}{8}\mathbf{E}g_{x^4}(x,\beta) + O(\sigma_\delta^6), \text{ as } \sigma_\delta \to 0.$$

We set $g(x,\beta) = C''(\eta(x,\beta))\eta_\beta(x,\beta)\eta_\beta^t(x,\beta)$. Then

$$A_c = S - \frac{1}{2}\sigma_\delta^2 A_{c2} + \frac{1}{8}\sigma_\delta^4 A_{c4} + O(\sigma_\delta^6),$$

where

$$A_{c2} = \mathbf{E}(C''\eta_\beta\eta_\beta^t)_{xx}(x,\beta), \quad A_{c4} = \mathbf{E}(C''\eta_\beta\eta_\beta^t)_{x^4}(x,\beta).$$

We apply (9):

$$A_c^{-1} = S^{-1} + \frac{1}{2}\sigma_\delta^2 S^{-1}A_{c2}S^{-1} + \frac{1}{8}\sigma_\delta^4 S^{-1}(2A_{c2}S^{-1}A_{c2} - A_{c4})S^{-1} + O(\sigma_\delta^6).$$

To approximate the matrix B_c we use condition (vi):

$$\psi_c(y,x,\beta) \approx yf_1 - f_2 - \frac{1}{2}\sigma_\delta^2(yf_1 - f_2)_{xx} + \frac{1}{8}\sigma_\delta^4(yf_1 - f_2)_{x^4}.$$

The remainder in the last approximate equality is bounded by

$$const \cdot (|y| + 1)e^{A|x|}\sigma_\delta^6.$$

Since $yf_1 - f_2 = (y - C'(\eta))\eta_\beta$, then in the approximation of $\psi_c\psi_c^t$ we can find terms like $(y - C'(\eta))^k$, $k = 1, 2$. We get rid of them and finally obtain

$$B_c = \varphi S - \frac{1}{2}\sigma_\delta^2 B_{c2} + \frac{1}{8}\sigma_\delta^4 B_{c4}.$$

We have

$$\begin{aligned}\Sigma_C &= \varphi S^{-1} + \tfrac{1}{2}\sigma_\delta^2 S^{-1}(2\varphi A_{c2} - B_{c2})S^{-1} \\ &+ \frac{1}{8}\sigma_\delta^4 S^{-1}(B_{c4} - 2\varphi A_{c4} + (6\varphi A_{c2} - 2B_{c2})S^{-1}A_{c2} - 2A_{c2}S^{-1}B_{c2}).\end{aligned}$$

At last we write the difference between Σ_C and Σ_Q and simplify it. □

Lemma. *Let F and G be two random matrices of the same size such that $\mathbf{E}\,GG^t$ is positive definite. Then $\mathbf{E}\,FF^t - \mathbf{E}\,FG^t\,(\mathbf{E}\,GG^t)^{-1}\,\mathbf{E}\,GF^t$ is positive semidefinite matrix. Moreover, it is zero matrix if, and only if, $F = HG$, $H = \mathbf{E}\,FG^t\,(\mathbf{E}\,GG^t)^{-1}$,* a.s.

Proof. Consider the matrix $A = F - HG$. The matrix AA^t is positive semidefinite a.s. Its expectation equals

$$\mathbf{E}\,(F-HG)(F-HG)^t = \mathbf{E}\,FF^t - \mathbf{E}\,HGF^t - \mathbf{E}\,FG^tH^t - \mathbf{E}\,HGG^tH^t \geq 0.$$

After substitution of H we have the lemma proved. □

We rewrite Δ from (7) in the form

$$\varphi\Delta = \mathbf{E}\,FF^t - \mathbf{E}\,FG^t\left(\mathbf{E}\,GG^t\right)^{-1}\mathbf{E}\,GF^t + \varphi L + \varphi^2 M,$$

where

$$F = (C'')^{3/2}\,\eta_x^2\eta_\beta, \quad G = (C'')^{1/2}\,\eta_\beta.$$

Here the function $C'' = C''(\eta)$, and $\eta = \eta(x,\beta)$. By Lemma the intercept term of matrix polynomial $\varphi\Delta$ (it is polynomial w.r.t. φ) is positive semidefinite. It can be zero if, and only if,

$$\left(C''\eta_x^2 \cdot I - \mathbf{E}\left[C''^2\eta_x^2\eta_\beta\eta_\beta^t\right]\cdot S^{-1}\right)\eta_\beta = 0 \quad \text{a.s.,} \tag{10}$$

where S comes from condition (x), and I is the identity matrix. Thus

$$\lim_{\varphi\to 0^+}\left[\varphi \lim_{\sigma_\delta^2\to 0^+}\sigma_\delta^{-4}(\Sigma_C - \Sigma_Q)\right]$$

is a positive semidefinite matrix. It is zero iff the condition (10) holds.

5. Polynomial model

Polynomial measurement error model has a form

$$\begin{cases} y_i = \beta_0 + \beta_1\xi_i + \ldots + \beta_m\xi_i^m + \varepsilon_i, \\ x_i = \xi_i + \delta_i, \end{cases} \qquad i = 1,\ldots,n. \tag{11}$$

Here $m \geq 1$, ε_i are i.i.d., $\varepsilon_i \sim N(0,\sigma_\varepsilon^2)$, and ε_i are independent of ξ_i and δ_i, $\{\xi_i\}$ and $\{\delta_i\}$ are the same as in Section 2.

The model (11) belongs to the exponential family (1) with functions $C(\eta) = \eta^2/2$, $\eta(\xi,\beta) = \beta_0 + \beta_1\xi_i + \ldots + \beta_m\xi_i^m$, and $\varphi = \sigma_\varepsilon^2$. The unknown parameter $\beta = (\beta_0,\ldots,\beta_m)^t$.

The conditions (ii) to (iv) are fulfilled. The conditions (v) to (ix) are explained in Kukush and Schneeweiss (2005). The matrix S from condition

(x) is Gram matrix for random vector $\zeta(x) = (1, x, \ldots, x^m)^t$, $S = \mathbf{E}\,\zeta\zeta^t$, and therefore it is positive definite.

We apply Theorem 3 and have

$$\begin{aligned}
&\varphi\Delta = \mathbf{E}\,\eta_\xi^4 K + \varphi\,\mathbf{E}(\sigma_x^{-2}\eta_\xi^2 K + 3\eta_\xi^2 K_1 + 3\eta_\xi\eta_{\xi\xi}K_s + \eta_{\xi\xi}^2 K) + \\
&+\varphi^2\,\mathbf{E}(\sigma_x^{-2}K_1 + K_2) - \mathbf{E}(\eta_\xi^2 K + \varphi K_1)\cdot S^{-1}\cdot \mathbf{E}(\eta_\xi^2 K + \varphi K_1),\\
&K := \zeta\zeta^t, \quad K_1 := \zeta'\zeta'^t, \quad K_2 := \zeta''\zeta''^t, \quad K_s := [\zeta'\zeta^t]_S.
\end{aligned}$$

The senior term of $\varphi\Delta$ is zero, i.e.

$$\mathbf{E}(\sigma_x^{-2}K_1 + K_2) = \mathbf{E}\,K_1\cdot S^{-1}\cdot \mathbf{E}\,K_1. \tag{12}$$

To prove this fact we use the orthonormal Hermite polynomials

$$h_i(x) = \frac{(-\sigma_x)^i}{\sqrt{i!}}\exp\left\{\frac{(x-\mu)^2}{2\sigma_x^2}\right\}\frac{d^i}{dx^i}\exp\left\{-\frac{(x-\mu)^2}{2\sigma_x^2}\right\},\ \mathbf{E}\,h_i(x)h_j(x) = \delta_{ij}.$$

Denote $h = (h_0(x), \ldots, h_m(x))^t$. Then there exists lower triangular nonsingular matrix B such that $\zeta = Bh$. We have $h_i'(x) = h_{i-1}(x)\sqrt{i}/\sigma_x$, $i \ge 1$, therefore $h' = Dh$, where the matrix D has zero components except $d_{i,i-1} = \sqrt{i}/\sigma_x$. We substitute in (12) the following expressions

$$\zeta' = B\cdot D\cdot h, \quad \zeta'' = B\cdot D^2\cdot h, \quad \mathbf{E}\,hh^t = I.$$

Then (12) holds due to the equality $DD^t + \sigma^2 D^2 D^{2t} = \sigma^2 DD^t DD^t$.

We have that $\varphi\Delta$ is linear in φ, i.e., $\varphi\Delta = A + \varphi B$. By Lemma we have $A \ge 0$. But under $\beta_m \ne 0$ we can easily prove that A is positive definite. Next, it was proved in Kukush *et al.* (2006) that $\Sigma_C \ge \Sigma_Q$. Thus B is positive semidefinite. So we can summarize that under $\beta_m \ne 0$,

$$\lim_{\sigma_\delta^2\to 0+}\sigma_\delta^{-4}(\Sigma_C - \Sigma_Q)$$

is positive definite matrix. Thus QS is more efficient than CS for small measurement error variance.

6. Poisson measurement error model

In Poisson model the conditional distribution $y|\eta$ belongs to the exponential family (1) with functions $C(\eta) = e^\eta$ and $\eta(\xi,\beta) = \beta_0 + \beta_1\xi$, and constant $\varphi = 1$. Here ν is a counting measure, $\nu(A) = \#(A \cap \{0,1,2,\ldots\})$, $A \in \mathcal{B}(\mathbb{R})$. The unknown parameter $\beta = (\beta_0, \beta_1)^t$.

It is easy to check that the conditions (ii), (iii), (iv), (vi), and (vii) hold true. The matrix $S = \mathbf{E}[e^\eta\eta_\beta\eta_\beta^t]$ is positive definite.

The matrix Δ is equal to

$$\Delta = \beta_1^2\sigma_x^{-2}\mathbf{E}e^{2\eta}\eta_\beta\eta_\beta^t + \beta_1^4 A_1 + A_2,$$

where A_1 is positive semidefinite and A_2 is positive definite under $\beta_1 \neq 0$. Then under $\beta_1 \neq 0$ we have $\Sigma_C - \Sigma_Q$ is positive definite for small σ_δ^2, and the first positive definite term of expansion of this difference is the term of order σ_δ^4. Under $\beta_1 = 0$ we have $\Delta = 0$ and $\Sigma_C = \Sigma_Q + O(\sigma_\delta^6)$.

7. Conclusions

In this paper we considered the nonlinear regression model with normal measurement errors and compared the efficiency of two consistent estimators of unknown parameter. All nuisance parameters, that is measurement error variance σ_δ^2, parameters of distribution of latent variable μ_ξ and σ_ξ^2, and dispersion parameter φ, were supposed to be known.

We considered two consistent estimators, the Quasi Score (QS) and the Corrected Score (CS) ones. We found expansions of their ACMs up to $O(\sigma_\delta^6)$ and proved that in polynomial and Poisson regression models the difference between ACMs of CS and QS is positive definite for small measurement error. Kukush and Schneeweiss (2005) proved, that choosing CS estimator instead of QS one results into negligible loss of efficiency (up to the order $O(\sigma_\delta^4)$). In this paper we showed that QS is more efficient than CS up to $O(\sigma_\delta^6)$. This result can be useful for selection of estimator if one knows a priori that the measurement error variance is small.

The author is grateful to Prof. A. Kukush for the problem statement and discussions.

Bibliography

1. Carroll, R. J., Ruppert, D., and Stefanski, L. A. (1995). *Measurement Error in Nonlinear Models.* – Chapman and Hall, London.
2. Kukush, A., and Schneeweiss, H. (2005). *Comparing different estimators in a nonlinear measurement error model. I. Mathematical Methods of Statistics,* **14**, *53-79.*
3. Kukush, A., Malenko, A., and Schneeweiss, H. (2006). *Optimality of the quasi-score-like estimator in a mean-variance model. Discussion paper 384. SFB 386, University of Munich.*
4. Stefanski, L. A. (1989). *Unbiased estimation of a nonlinear function of a normal mean with application to measurement error models.* Communication in Statistics, Series A, **18**, 4335-4358.

Department of Probability Theory and Mathematical Statistics, Kyiv National Taras Shevchenko University, Kyiv, Ukraine