Journal of the Korean
Data & Information Science Society
1999, Vol. 10, No. 1, pp. 193 ∼ 198

On the Estimation in Regression Models with Multiplicative Errors

Cheolyong Park[1]

Abstract

The estimation of parameters in regression models with multiplicative errors is usually based on the gamma or log-normal likelihoods. Under reciprocal misspecification, we compare the small sample efficiencies of two sets of estimators via a Monte Carlo study. We further consider the case where the errors are a random sample from a Weibull distribution. We compute the asymptotic relative efficiency of quasi-likelihood estimators on the original scale to least squares estimators on the log-transformed scale and perform a Monte Carlo study to compare the small sample performances of quasi-likelihood and least squares estimators.

Key Words and Phrases: Constant coefficient of variation, Efficiency, Log-normal, Quasi-likelihood, Weibull

1. Introduction

Consider a multiplicative regression model for positive observations, given by

$$Y_i = \mu_i \, \varepsilon_i \quad (i = 1, \ldots, n) \tag{1}$$

with $\log \mu_i = \beta_0 + \sum_{r=1}^{p} \beta_r x_{ir}$ and $\{\varepsilon_i\}$ independently and identically distributed (IID) with $E(\varepsilon_i) = 1$ and $Var(\varepsilon_i) = \phi_1$, where β_0, $\beta_r(r = 1, \ldots, p)$, and ϕ_1 are unknown parameters and $x_{ir}(r = 1, \ldots, p)$ are explanatory variables. This model can be alternatively expressed as an additive model for its logarithm, given by

$$\log Y_i = \nu_i + \eta_i \quad (i = 1, \ldots, n) \tag{2}$$

with $\nu_i = \{\beta_0 + E(\log \varepsilon_i)\} + \sum_{r=1}^{p} \beta_r x_{ir}$ and $\{\eta_i\}$ IID with $E(\eta_i) = 0$ and $Var(\eta_i) = \phi_2$, where ϕ_2 is a unknown parameter. Model (1) has a constant coefficient of variation with logarithmic link and Model (2) has a constant variance with identical

[1]Assistant Professor, Department of Statistics, Keimyung University, Taegu 704-701

link. The non-intercept parameters $\beta_1, \ldots, \beta_p$ can be estimated from either Model (1) or Model (2). To ensure the identifiability of these parameters, we impose $\sum_{i=1}^{n} x_{ir} = 0$ $(r = 1, \ldots, p)$, which is always possible. The intercept parameters cannot be similarly identified.

In order to estimate $\beta_0, \ldots, \beta_p$, we can apply the quasi-likelihood(QL) method, a generalization of the least squares(LS) method, described by Wedderburn (1974). The QL equations are given by

$$\sum_{i=1}^{n} \frac{y_i - \mu_i}{\mu_i} x_{ir} = 0 \quad (r = 0, \ldots, p) \tag{3}$$

for Model (1), which will be denoted by 'OG', and

$$\sum_{i=1}^{n} (\log y_i - \nu_i)\, x_{ir} \quad (r = 0, \ldots, p) \tag{4}$$

for Model (2), which will be denoted by 'LG'. It is well known that the equations (3) are the maximum likelihood(ML) equations when Y_i follows a gamma distribution and the equations (4) are the ML equations when $\log Y_i$ follows a normal distribution. Also we note that the QL equations (4) are the unweighted LS equations.

In Section 2, we first consider the case where the true distribution of ε_i is either gamma or log-normal. We present and compare the asymptotic relative efficiency(ARE) of QL estimators of $\beta_1, \ldots, \beta_p$ relative to ML estimators when the true distribution is reciprocally misspecified. Next we consider the case where the true distribution is Weibull. We calculate and compare the AREs of QL estimators from (3) and (4) relative to ML estimators. In Section 3, we compare the small sample efficiencies of QL estimators relative to ML estimators under reciprocal misspecification and then compare small sample performances of QL estimators from (3) to LS estimators from (4).

2. Asymptotic Relative Efficiencies

Models with Constant Variance

Here we consider Model (2). This model has identical link function $g(\nu_i) = \nu_i$ and a constant variance function $V(\nu_i) = 1$, i.e.

$$E(\log Y_i) = \nu_i = \beta_0^* + \sum_{r=1}^{p} \beta_r x_{ir}, \; Var(\log Y_i) = \phi_2 \quad (i = 1, \ldots, n)$$

with $\beta_0^* = \{\beta_0 + E(\log \varepsilon_i)\}$. The equations LG, given in (4), are used to estimate $\beta_1, \ldots, \beta_p$.

Cox and Hinkley(1968) showed that the ARE of QL estimators of $\beta_1, \ldots, \beta_p$ relative to ML estimators based on the true distribution is given by $(\phi_2 A_\epsilon)^{-1}$, where

$$A_\epsilon = E(-\partial^2 \ell_1 / \partial \nu_1^2)$$

with ℓ_1 being the log likelihood of log Y_1.

Cox and Hinkley(1968) calculated A_ϵ for some distributions and showed that, using the result of Bartlett and Kendall(1946), if ε_i has a gamma distribution with index γ, then the ARE becomes

$$\text{eff}_G(LG) = \{\gamma\psi'(\gamma)\}^{-1}, \tag{5}$$

where $\psi'(\gamma) = \partial^2 \log \Gamma(\gamma)/\partial\gamma^2$ is the trigamma function and $\Gamma(\cdot)$ is the gamma function.

When ε_i has a Weibull distribution with shape c, we can show that $A_\epsilon = c^2$ and $\phi_2 = \psi'(1)/c^2$ and so the ARE becomes

$$\text{eff}_W(LG) = \{\psi'(1)\}^{-1} = 6/\pi^2. \tag{6}$$

Models with Constant Coefficient of Variation

Here we consider Model (1). This model has logarithmic link function $g(\mu_i) = \log \mu_i$ and quadratic variance function $V(\mu_i) = \mu_i^2$, i.e.

$$E(Y_i) = \mu_i,\ \log \mu_i = \beta_0 + \sum_{r=1}^{p} \beta_r x_{ir},\ Var(Y_i) = \phi_1\, \mu_i^2\ \ (i = 1, \ldots, n).$$

The equations OG, given in (3), are used to estimate $\beta_1, \ldots, \beta_p$.

Firth(1987) showed that the ARE of QL estimators of $\beta_1, \ldots, \beta_p$ relative to ML estimators based on the true distribution is given by $(\phi_1 A_\epsilon)^{-1}$, where

$$A_\epsilon = \mu_1^2 E(-\partial^2 \ell_1^* / \partial \mu_1^2)$$

with ℓ_1^* being the log likelihood of Y_1.

Firth (1987) calculated A_ϵ for some distributions and showed that if the true distribution of ε_i is log-normal, the ARE becomes

$$\text{eff}_L(OG) = \log(1 + \phi_2)/\phi_2. \tag{7}$$

When ε_i has a Weibull distribution with shape c, it is easy to show that $A_\epsilon = c^2$ and $\phi_1 = \Gamma(2/c+1)/\Gamma^2(1/c+1) - 1$ and thus the ARE becomes

$$\text{eff}_W(OG) = \left\{c^2[\Gamma(2/c+1)/\Gamma^2(1/c+1) - 1) - 1]\right\}^{-1}. \tag{8}$$

Comparisons

Firth(1988) compared the AREs in equations (5) and (7) by noting the relation that γ^{-1} is equal to ϕ_1, the variance of ε_i. For some five points in the realistic range, 0.1 to 2.0, of ϕ_1, $\mathrm{eff_L(OG)}$ and $\mathrm{eff}_G(LG)$ can be calculated as in Table 1.

Table 1: Asymptotic Relative Efficiencies Under Reciprocal Misspecification

ϕ_1	0.1	0.2	0.5	1.0	2.0
$\mathrm{eff_G(LG)}$	0.951	0.904	0.775	0.608	0.405
$\mathrm{eff_L(OG)}$	0.953	0.912	0.811	0.693	0.549

As noted in Firth(1988), over the range considered, $\mathrm{eff_G(LN)}$ is slightly bigger than $\mathrm{eff_L(OG)}$ and the difference increases in ϕ_1. Thus QL estimators on the original scale performs a little better than LS estimators on the log-transformed scale under reciprocal misspecification.

We now consider the case where ε_i has a Weibull distribution with shape c. Since

$$\mathrm{eff_W(LG)} = 6/\pi^2 \text{ and } \mathrm{eff_W(OG)} = \{\mathrm{c}^2[\Gamma(2/\mathrm{c}+1)/\Gamma^2(1/\mathrm{c}+1) - 1]\}^{-1},$$

the ARE of QL estimators on the original scale to LS estimators on the logarithmic scale is given by

$$\mathrm{eff_W(OG,LG)} = \pi^2 / \left\{6\mathrm{c}^2[\Gamma(2/\mathrm{c}+1)/\Gamma^2(1/\mathrm{c}+1) - 1]\right\}. \tag{9}$$

$\mathrm{eff}_W(OG, LG)$ increases in c on the range (0,1) and then decreases in c on the range $(1, \infty)$. The maximum value of $\mathrm{eff_W(OG,LG)}$ is $\pi^2/6$ at $c = 1$, and $\mathrm{eff}_W(OG, LG)$ approaches to 0 as c goes to 0 whereas it approaches to 1 as c goes to ∞. It is easy to show that $\mathrm{eff_W(OG,LG)} = 1$ is attained at c approximately equal to $c^* \equiv 0.3881$. Therefore $\mathrm{eff_W(OG,LG)} < 1$ for $0 < c < c^*$ and $\mathrm{eff_W(OG,LG)} > 1$ for $c > c^*$. To calculate the ARE in terms of ϕ_1, the variance of ε_i, we can use the relation between ϕ_1 and c, $\phi_1 = \Gamma(2/c+1)/\Gamma^2(1/c+1) - 1$, which is given just prior to (6). For the five points in the realistic range, 0.1 to 2.0, of ϕ_1, $\mathrm{eff_W(OG,LG)}$ are calculated and summarized in Table 2.

Since $c > c^*$ is approximately equivalent to $\phi_1 < 10.92$, over the range considered, the QL estimation on the original scale performs much better than the LS estimation on the logarithmic scale.

3. Monte Carlo Study

In this section, we first calculate and compare small sample efficiencies of QL estimators to ML estimators under reciprocal misspecification when the true distribution

Table 2: ARE of QL on the Original Scale to LS on the Logarithmic Scale

ϕ_1	0.1	0.2	0.5	1.0	2.0
$\text{eff}_W(\text{OG}, \text{LG})$	1.341	1.452	1.596	1.645	1.583

of ε_i is either gamma or log-normal. We then compare small sample performances of QL estimators from (3) to LS estimators from (4) when the true distribution is Weibull.

Since the ARE of QL estimators of non-intercept parameters $\beta_1, \ldots, \beta_p$ to ML estimators is the same for all estimators, we only consider the simple regression model with intercept, i.e. $p = 1$. We consider the sample sizes such as $n = 10, 20, 50$ and, for each sample size, we generated the explanatory variable x_{i1} $(i = 1, \ldots, n)$ from the uniform distribution from 0 to 1 and then centered it so that $\sum x_{i1} = 0$. As noted in Section 1, this centering is taken to ensure the identifiability of β_1. When we compare two methods of estimation, we generate 500 random samples from the true distribution and then calculate 500 estimates of β_1 by each method. The small sample efficiency of method 1 to method 2 is just given by $\sum(b_i^{(2)} - \beta_1)^2 / \sum(b_i^{(1)} - \beta_1)^2$, where $b_i^{(1)}$ and $b_i^{(2)}$ are the estimators of β_1 based on method 1 and method 2, respectively.

We first consider gamma and log-normal distributions as true distributions of ε_i. For the five points of ϕ_1 considered in Section 2, small sample efficiencies, $\hat{\text{eff}}_G(LG)$ and $\hat{\text{eff}}_L(OG)$, of QL estimator of β_1 to ML estimator under reciprocal misspecification are calculated and summarized in Table 3.

Table 3: Small Sample Efficiencies under Reciprocal Misspecification

ϕ_1		0.1	0.2	0.5	1.0	2.0
$n = 10$	$\hat{\text{eff}}_G(LG)$	0.960	0.941	0.848	0.776	0.617
	$\hat{\text{eff}}_L(OG)$	0.970	0.936	0.892	0.816	0.814
$n = 20$	$\hat{\text{eff}}_G(LG)$	0.963	0.925	0.789	0.643	0.531
	$\hat{\text{eff}}_L(OG)$	0.964	0.931	0.889	0.835	0.723
$n = 50$	$\hat{\text{eff}}_G(LG)$	0.967	0.922	0.782	0.708	0.423
	$\hat{\text{eff}}_L(OG)$	0.955	0.926	0.848	0.807	0.645

We find that the difference in $\hat{\text{eff}}_L(OG)$ and $\hat{\text{eff}}_G(LG)$ is small for $\phi_1 = 0.1, 0.2, 0.5$

and becomes large for $\phi_1 = 1, 2$. We also find that there is a trend that the difference increases in ϕ_1. Therefore, as expected from the comparison of the AREs in Section 2, we can conclude that QL estimator on the original scale performs a little better than LS estimator on the logarithmic scale under reciprocal misspecification.

Next we consider the case where the true distribution of ε_i is Weibull. Small sample efficiency $\hat{\text{eff}}_W(OG, LG)$ of QL estimator of β_1 on the original scale to LS estimator on the logarithmic scale is calculated for the five points of ϕ_1 and summarized in Table 4.

Table 4: Small Sample Efficiency $\hat{\text{eff}}_W(OG, LG)$

ϕ_1	0.1	0.2	0.5	1.0	2.0
$n = 10$	1.160	1.227	1.290	1.448	1.417
$n = 20$	1.272	1.381	1.544	1.685	1.489
$n = 50$	1.336	1.400	1.535	1.536	1.421

We find that $\hat{\text{eff}}_W(OG, LG) > 1$ for all values of ϕ_1 and that $\hat{\text{eff}}_W(OG, LG)$ increases in ϕ_1 on the range [0.1, 1] and drops a little bit at $\phi_1 = 2$. These findings are well expected from the study on ARE in Section 2 and we can conclude that QL estimator on the original scale performs better than LS estimator on the logarithmic scale.

References

1. Bartlett, M.S. and Kendall, D.G. (1946). The statistical analysis of variance-heterogeneity and the logarithmic transformation. *Journal of the Royal Statistical Society, Suppl.*, 8, 128-138.

2. Cox, D.R. and Hinkley, D.V. (1968). A note on the efficiency of least square estimates. *Journal of the Royal Statistical Society, B*, 30, 284-289.

3. Firth, D. (1987). On the efficiency of quasi-likelihood estimation. *Biometrika*, 74, 233-245.

4. Firth, D. (1988). Multiplicative Errors: Log-Normal or Gamma? *Journal of the Royal Statistical Society, B*, 50, 266-268.

5. Wedderburn, R.W.M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika*, 61, 439-447.