Journal of Multivariate Analysis 106 (2012) 118–146

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Edgeworth expansions for GEL estimators

Gubhinder Kundhi [a,*], Paul Rilstone [b]

[a] *Department of Economics, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6, Canada*

[b] *Department of Economics, York University, 4700 Keele Street, Toronto, Ontario, M3J 1P3, Canada*

ARTICLE INFO

Article history:
Received 31 October 2010
Available online 3 December 2011

AMS subject classifications:
60E05
60E10
62E17
91G70

Keywords:
Higher order asymptotics
Edgeworth expansions
Generalized Empirical Likelihood
Generalized Method of Moments

ABSTRACT

Finite sample approximations for the distribution functions of Generalized Empirical Likelihood (GEL) are derived using Edgeworth expansions. The analytical results obtained are shown to apply to most of the common extremum estimators used in applied work in an i.i.d. sampling context. The GEL estimators considered include the Continuous Updating, Empirical Likelihood and Exponential Tilting estimators. These estimators are popular alternatives to Generalized Method of Moment (GMM) estimators and their finite sample properties are examined. In a Monte Carlo Experiment, higher order analytical corrections provided by Edgeworth approximations work well in comparison to first order approximations and improve inferences in finite samples.

1. Introduction

The higher order asymptotic properties of nonlinear estimators have received substantial attention in the statistics and econometrics literature recently. Most of these estimators are special cases of either Generalized Method of Moments (GMM) or Generalized Empirical Likelihood (GEL) estimators. However, there still remain some substantial gaps in our knowledge of their finite sample properties. Recent papers by Rilstone, Srivastava and Ullah [24, henceforth RSU] and Newey and Smith [19, henceforth NS] have derived the first and second order moments of certain nonlinear estimators, including those focused on in this paper. However, while these results may be useful for bias and dispersion correction and higher order efficiency comparisons, they do not allow for departures from normality in the distribution of the estimators. For this, an Edgeworth or related expansion is necessary, and to derive these requires knowledge of the higher moments of the estimators. The focus of this paper is to derive these expansions for GEL estimators.

The complexity of standard higher order approximations may make them appear quite forbidding. The derivation of Edgeworth expansions for multivariate models, not to mention nonlinear, can be quite intimidating. Most such derivations, although quite elegant, get mired down in a morass of tensor and multivariate notation. One result is that the recent attention of researchers into approximating the distributions of estimators and test statistics has focused primarily on resampling techniques such as the bootstrap and the jackknife. This is ironic, since, to show the validity of resampling techniques, one generally needs to first derive an asymptotic expansion of some kind. Analytical expansions such as the Edgeworth are useful whether one wants to use them directly for improved inferences, or as a device to show the validity of a resampling technique.

The focus for GEL estimators is a $k \times 1$ parameter θ whose true value, θ_0, is the unique solution to an $l \times 1$ set of moment conditions $\mathcal{E}[g_i(\theta_0)] = 0$. With $l > k$ we have the well known over identification problem. From a technical standpoint,

* Corresponding author.
E-mail addresses: gkundhi@connect.carleton.ca (G. Kundhi), pril@yorku.ca (P. Rilstone).

0047-259X/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2011.11.005

this can make the analysis problematic since approximating the estimators thus involves more equations than unknowns. Various remedies to this problem have been proposed by Newey and McFadden [18], RSU [24] and NS [19], who introduce various auxiliary parameters to even out the number of parameters and estimating equations. We do so here. Let $\widehat{\theta}$ denote a generic estimator. For the GEL estimators, as NS [19] have shown, $\widehat{\theta}$ is the solution to a saddle point problem, along with an $l \times 1$ vector of Lagrangian multipliers $\widehat{\lambda}$.

Let $\widehat{\beta}' = (\widehat{\lambda}', \widehat{\theta}')'$ where $'$ denotes transposition. As one simplification we derive approximations to the distribution function of $\tau'\widehat{\beta}$, where τ is an $m \times 1$ vector of constants, rather than $\widehat{\theta}$. Apart from greatly simplifying the analysis there are two other reasons for doing so. First is that, although in principle one may like to derive the k variate distribution of $\widehat{\theta}$, in practice most inferences are focused on linear combinations of the parameter vector. Second, as we show below, since the results hold for arbitrary $\tau \in \Re^m$, it follows that the joint approximate distribution of $\widehat{\beta}$ can be derived from the approximate distribution of $\tau'\widehat{\beta}$. With these results in hand, the critical points from a Cornish–Fisher expansion follow immediately. The expansions we derive are third order accurate in the usual sense that the expansion includes terms up to and including those with a coefficient $1/N$ where N is the sample size. This allows us to correct for higher order bias, variance, skewness and kurtosis.

To name a few papers on the higher order asymptotic properties of estimators and test statistics, Phillips [21] and Bhattacharya and Ghosh [3] have examined the validity of Edgeworth expansions for various estimators. Others, such as Pfanzagl and Wefelmeyer [20], Ghosh et al. [8] have examined higher order efficiency of maximum likelihood estimators. Rothenberg [26], McCullagh [17] and Barndorff-Nielson and Cox [2] provide early surveys. Amemiya [1] examined the logit regression model. RSU [24] and Rilstone and Ullah [25] derived third-order stochastic expansions and the second order bias and mean squared error of a wide class of nonlinear estimators. NS [19] have compared the second order bias and variance of various GEL and GMM estimators.

Wallace [28] developed the Edgeworth expansion for a standardized sample average of observations on a univariate random variable. Extensions to multivariate random variables were done by Chambers [6], Sargan [27], Phillips [21] and Bhattacharya and Rao [4]. Rothenberg [26], Barndorff-Nielson and Cox [2]and Hall [9] provide detailed discussions on the development of Edgeworth and Cornish–Fisher expansions. Bhattacharya and Ghosh [3] derived general conditions under which the Edgeworth expansion provides a valid approximation to various functions of sample moments. They obtained asymptotic expansions for distributions of the Maximum Likelihood Estimators (MLEs) and minimum contrast estimators by application of their results. Phillips and Park [22] derived the Edgeworth expansion for the Wald statistic and Hansen [12] obtained the Edgeworth expansion of the objective function for GMM distance statistic with nonlinear restrictions. Bravo [5] considered the Edgeworth expansion for the maximum dual likelihood estimator and the empirical likelihood ratio statistic. Linton [16] derived the Edgeworth approximation for the semiparametric instrumental variable estimators and associated test statistics.

It is well known that although the optimal (two-step) GMM estimator is asymptotically efficient, its finite sample properties can be problematic. Imbens [14] notes that the estimator is not invariant to linear transformations of the moment conditions. Studies such as Hansen [11] have shown that the optimal GMM estimator may have substantial bias in smaller samples. Partly in response to these considerations, other asymptotically efficient estimators have been proposed. These include the Continuous Updating Estimator (CUE), the Empirical Likelihood (EL) estimator and the Exponential Tilting (ET) estimator. The EL and the ET estimators as discussed by Imbens [14] and Qin and Lawless [23] work well with over identified models and are appealing for their information-theoretic characteristics. Hansen et al. [13] have shown that the CUE estimator can have smaller bias in finite samples than the optimal GMM estimator. NS [19] have shown that all these estimators are GEL estimators.

In this paper stochastic expansions for the GEL estimators are derived and used to obtain their Edgeworth approximations. The Edgeworth approximations for the CUE, EL, ET estimators are subsequently obtained as special cases for comparison purposes. These analytical results are illustrated using an example of an overidentified model from [23].

Analytical expansions are useful for improved finite sample inferences such as lowering coverage errors of confidence intervals. Also, higher order approximate moments of the estimators can be obtained by integrating the Edgeworth expansion, providing more distributional information. These expansions can also enable higher order comparisons such as discussed in [26]. Edgeworth expansions are usually estimated using sample moments and may not provide perfectly accurate numerical approximations. However, even in these cases, the analytical corrections delivered by the higher order terms in the Edgeworth expansions provide an intuitive explanation and measure of the departure of an estimator's actual sampling distribution from its asymptotic distribution.

For each of the GEL estimators considered we proceed as follows. We let λ denote the associated Lagrangian multipliers. Put $\beta' = (\lambda', \theta')$ and $m = k + l$. We consider situations where an $m \times 1$ estimator, $\widehat{\beta}$, solves

$$\frac{1}{N}\sum q_i(\widehat{\beta}) = 0 \tag{1.1}$$

where $q_i(\beta)$ is a known $m \times 1$ vector-valued function of the observable data and $\mathcal{E}[q_i(\beta)] = 0$ only at $\beta = \beta_0$. We elaborate on this in the next section. To sidestep a lengthy itemization of regularity conditions, $\widehat{\beta}$ is assumed to be consistent. We derive the first four moments of $\tau'\widehat{\beta}$ and then specialize for the main cases of interest such as $\tau' = (0_{1\times l}, \alpha')$, where α is an arbitrary $k \times 1$ vector of constants.