Supplementary material for
“Evaluating the Relative Merits of Competing Models Based on Empirical Likelihood Ratio Test”

Yan Fan\textsuperscript{a} and Yukun Liu\textsuperscript{b}\textsuperscript{*}

\textsuperscript{a} School of Business Information, Shanghai University of International Business, and Economics, 1900 Wenzhang Road, Shanghai 201620, China

\textsuperscript{b} School of Finance and Statistics, East China Normal University, 500 Dongchuan Road, Shanghai, 200241, China

(Received 00 Month 200x; in final form 00 Month 200x)

This supplementary material gives a detailed proof of Theorem 1 of the corresponding paper.

1. Assumptions and Theorem

For clarity and completeness, we begin with restating the necessary assumptions and Theorem 1. Let

\[ A(\phi) = \mathbb{E}\{\nabla^2 \Lambda_i(\phi^*)\}, \quad B(\phi) = \mathbb{E}[\nabla \Lambda_i(\phi^*)\{\nabla \Lambda_i(\phi^*)\}^\tau], \]

where \( \nabla \) is the differentiation operator \( \partial/\partial \phi \) with \( \phi^\tau = (\alpha^\tau, \beta^\tau) \). We make the following assumptions on the competing models under consideration.

(C1) The parameter spaces \( A \subset \mathbb{R}^{d_1} \) and \( B \subset \mathbb{R}^{d_2} \) are both compact.

(C2) (Differentiability and integrability)

(i) For all \((y, x)\) on their supports, \( f(y|\alpha) \) and \( g(y|\beta) \) are three times differentiable with respect to \( \alpha \) and \( \beta \) respectively.

(ii) There exists a nonnegative function \( H(y, x) \) satisfying \( \mathbb{E}\{H(Y, X)\} < \infty \) such that \( |\log f(y|\alpha)|, |\log g(y|\beta)| \) and all their first three orders of derivatives with respect to \( \alpha \) and \( \beta \) are controlled by \( H(y, x) \).

(C3) As a function of \( \alpha \), \( \mathbb{E}[\log f(Y|X, \alpha)] \) has a unique maximum on \( A \) at an interior point \( \alpha^* \); And as a function of \( \beta \), \( \mathbb{E}[\log g(Y|X, \beta)] \) has a unique maximum on \( B \) at an interior point \( \beta^* \).

(C4) \( A(\phi^*) \) is nonsingular.

Let \( \Sigma = A^{-1}BA^{-1} \) and \( \Omega = -\Sigma^{1/2}A\Sigma^{1/2} \) with \( A = A(\phi^*) \) and \( B = B(\phi^*) \).

\textbf{THEOREM 1.1} \textit{Assume the conditions in (C1-C4). If case (a) is true,}

\( \text{ELR} \xrightarrow{d} \chi_1^2 \) \textit{where} \( \xrightarrow{d} \) \textit{denotes convergence in distribution. If case (b) is true,}

\( \text{ELR} \xrightarrow{d} (\xi^\tau \Omega \xi)/((\xi^\tau \Omega \xi)^2 \xi) \), \textit{where} \( \xi \), \textit{of the same length as} \( (\alpha^*, \beta^*) \), \textit{is a standard multivariate normal random vector.}

\textsuperscript{*}Corresponding author. Email: ykliu@sfs.ecnu.edu.cn

ISSN: 0266-4763 print/ISSN 1360-0532 online
© 200x Taylor & Francis
DOI: 10.1080/0266476XYxxxxxxxx
http://www.informaworld.com
Empirical likelihood ratio model selection test

2. Proof

Note that

\[ Z_i = \Lambda_i(\hat{\phi}) = \log\{f(Y_i|X_i; \hat{\alpha})\} - \log\{g(Y_i|X_i; \hat{\beta})\}. \]

Under regular conditions,

\[ \hat{\phi} = \phi^* + O_p(n^{-1/2}). \]

According to the definition of \( \hat{\phi} \), we have

\[
0 = \frac{1}{n} \sum_{i=1}^{n} \nabla \Lambda_i(\hat{\phi}) = \frac{1}{n} \sum_{i=1}^{n} \nabla \Lambda_i(\phi^*) + \frac{1}{n} \sum_{i=1}^{n} \nabla^2 \Lambda_i(\phi^*)(\hat{\phi} - \phi^*) + o_p(n^{-1/2}).
\]

This leads to

\[
\hat{\phi} - \phi^* = -A^{-1} \frac{1}{n} \sum_{i=1}^{n} \nabla \Lambda_i(\phi^*) + o_p(n^{-1/2})
\]

since \( A = \mathbb{E}\{\nabla^2 \Lambda_i(\phi^*)\} \).

The second-order Taylor expansion gives

\[
Z_i = \Lambda_i(\hat{\phi}) = \Lambda_i(\phi^*) + (\hat{\phi} - \phi^*)^T \nabla \Lambda_i(\phi^*) + \frac{1}{2} (\hat{\phi} - \phi^*)^T \nabla^2 \Lambda_i(\phi^*)(\hat{\phi} - \phi^*) + o_p(n^{-1}).
\]

We consider the limiting distributions of ELR in cases (a) and (b) of the null hypothesis, respectively.

2.1 Limiting distribution of ELR in case (a)

It can be seen that

\[
\max_i Z_i = o_p(n^{1/2}), \quad \frac{1}{n} \sum_{i=1}^{n} Z_i = O_p(n^{-1/2}), \quad \frac{1}{n} \sum_{i=1}^{n} Z_i^2 = O_p(1),
\]

which means that the Lagrange multiplier \( \lambda \) satisfies

\[
\lambda = \left( \sum_{i=1}^{n} Z_i^2 \right)^{-1} \sum_{i=1}^{n} Z_i + o_p(n^{-1/2}) = O_p(n^{-1/2}).
\]

The fact that \( \max_i |\lambda Z_i| = o_p(1) \) implies that

\[
\text{ELR} = \left( \sum_{i=1}^{n} Z_i^2 \right)^{-1} \left( \sum_{i=1}^{n} Z_i \right)^2 + o_p(1).
\]

The limiting distribution of ELR is determined by the asymptotic properties of
\[(1/n) \sum_{i=1}^{n} Z_i \text{ and } (1/n) \sum_{i=1}^{n} Z_i^2. \text{ Note that} \]
\[
\frac{1}{n} \sum_{i=1}^{n} Z_i = \frac{1}{n} \sum_{i=1}^{n} \Lambda_i(\phi^*) + \frac{1}{n} \sum_{i=1}^{n} (\hat{\phi} - \phi^*)^T \nabla \Lambda_i(\phi^*) + o_p(n^{-1/2})
\]

According to the definition of \(\alpha^*\) and \(\beta^*\), it follows that

\[
\nabla \mathbb{E} \left[ \log \{ f(Y_i | X_i, \alpha^*) \} \right] = 0, \quad \nabla \mathbb{E} \left[ \log \{ g(Y_i | X_i, \beta^*) \} \right] = 0.
\]

Under regularity conditions, \(\mathbb{E} \{ \nabla \Lambda_i(\phi^*) \} = 0\), which means

\[
\frac{1}{n} \sum_{i=1}^{n} \nabla \Lambda_i(\phi^*) = O_p(n^{-1/2}).
\]

Since \((\hat{\phi} - \phi^*) = O_p(n^{-1/2})\), it then follows that

\[
\frac{1}{n} \sum_{i=1}^{n} Z_i = \frac{1}{n} \sum_{i=1}^{n} \Lambda_i(\phi^*) + o_p(n^{-1/2})
\]

and

\[
n \text{Var} \left( \frac{1}{n} \sum_{i=1}^{n} Z_i \right) = \text{Var} \{ \Lambda_i(\phi^*) \} + o(1).
\]

Meanwhile

\[
\frac{1}{n} \sum_{i=1}^{n} Z_i^2 = \frac{1}{n} \sum_{i=1}^{n} \left\{ \Lambda_i(\phi^*) + (\hat{\phi} - \phi^*)^T \nabla \Lambda_i(\phi^*) + o_p(n^{-1/2}) \right\}^2
\]
\[
= \frac{1}{n} \sum_{i=1}^{n} \{ \Lambda_i(\phi^*) \}^2 + o_p(1)
\]
\[
= \text{Var} \{ \Lambda_i(\phi^*) \} + o_p(1) \quad \text{If } H_0 \text{ is true.}
\]

Consequently, if the null hypothesis is true, the empirical likelihood ratio

\[
l(0) = \left( \sum_{i=1}^{n} Z_i^2 \right)^{-1} \left( \sum_{i=1}^{n} Z_i \right)^2 + o_p(1)
\]
\[
= [\text{Var} \{ \Lambda_i(\phi^*) \}]^{-1} \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Lambda_i(\phi^*) \right\}^2 + o_p(1)
\]

is asymptotically distributed as \(\chi^2_1\).

### 2.2 Limiting distribution of ELR in case (b)

If \(f(\cdot, \alpha^*) = g(\cdot, \beta^*)\) almost surely, then

\[
Z_i = (\hat{\phi} - \phi^*)^T \nabla \Lambda_i(\phi^*) + \frac{1}{2} (\hat{\phi} - \phi^*)^T \nabla^2 \Lambda_i(\phi^*)(\hat{\phi} - \phi^*) + o_p(n^{-1}). \tag{2}
\]
Recall that $A = \mathbb{E}\{\nabla^2 \Lambda_i(\phi^*)\}$ and $B = \mathbb{E}[\nabla \Lambda_i(\phi^*)\{\nabla \Lambda_i(\phi^*)\}^\top]$. It follows from (1) that $\sqrt{n}(\hat{\phi} - \phi^*)$ has a jointly normal limiting distribution with mean zero and covariance matrix $\Sigma = A^{-1}BA^{-1}$. Meanwhile, combining equations (1) and (2), we have

$$\sum_{i=1}^n Z_i = -\frac{n}{2}(\hat{\phi} - \phi^*)^\top A(\hat{\phi} - \phi^*) + o_p(1)$$

and

$$\sum_{i=1}^n Z_i^2 = \sum_{i=1}^n (\hat{\phi} - \phi^*)^\top \nabla \Lambda_i(\phi^*)]^2 + o_p(1) = n(\hat{\phi} - \phi^*)^\top B(\hat{\phi} - \phi^*) + o_p(1).$$

It can be seen that

$$\max_i |Z_i| = o_p(1), \quad \frac{1}{n} \sum_{i=1}^n Z_i = O_p(n^{-1}), \quad \frac{1}{n} \sum_{i=1}^n Z_i^2 = O_p(n^{-1}),$$

which means that the Lagrange multiplier $\lambda$ satisfies

$$\lambda = \left(\sum_{i=1}^n Z_i^2\right)^{-1} \sum_{i=1}^n Z_i + o_p(n^{-1/2}) = O_p(1).$$

It then follows that $\max_i |\lambda Z_i| = o_p(1)$ and therefore

$$\text{ELR} = \left(\sum_{i=1}^n Z_i^2\right)^{-1} \left(\sum_{i=1}^n Z_i\right)^2 + o_p(1) = \frac{1}{4} \frac{-\{n(\hat{\phi} - \phi^*)^\top A(\hat{\phi} - \phi^*)\}^2}{n(\hat{\phi} - \phi^*)^\top B(\hat{\phi} - \phi^*)} + o_p(1)$$

Since $\sqrt{n}(\hat{\phi} - \phi^*) \xrightarrow{d} \Sigma^{1/2} \xi$ where $\xi$ is a standard multivariate normal variable, it follows that

$$n(\hat{\phi} - \phi^*)^\top A(\hat{\phi} - \phi^*) \xrightarrow{d} \xi^\top \Omega \xi, \quad n(\hat{\phi} - \phi^*)^\top B(\hat{\phi} - \phi^*) \xrightarrow{d} \xi^\top \Omega_1 \xi,$$

where $\Omega = -\Sigma^{1/2}A\Sigma^{1/2}$ and $\Omega_1 = \Sigma^{1/2}B\Sigma^{1/2} = \Sigma^{1/2}A\Sigma\Sigma^{1/2} = \Omega^2$. Therefore

$$\text{ELR} \xrightarrow{d} \frac{(\xi^\top \Omega \xi)^2}{\xi^\top \Omega_1 \xi} + o_p(1).$$

This completes the proof of Theorem 1.