CrossRef DOI of original article:

Asymptotic Normality of the Encompassing Test Associated to the Linear Parametric Modelling and the Kernel Method for-Mixing Processes

Received: 1 January 1970 Accepted: 1 January 1970 Published: 1 January 1970

Abstract

Index terms—

1 INTRODUCTION

Model selection is a challenging step in statistical modelling. Modelling any data requires characterization of the associated data generating process (DGP). The DGP is unknown and therefore we face several admissible competing models. Model selection consists on selecting a model, which mimics such unknown DGP, from a set of admissible models according to a criterion. One retains the model which makes such criterion optimal, Tibshirani et al. (2015), Ferraty and Hall (2015) and Li et al. (2017). There exist various model selection criteria in the literature when admissible models have fully parametric specification, such as the Wald test, the likelihood ratio test, the Lagrange multiplier test, the information criteria and so on, see Hamilton (1994), Greene (2003) and Hooten and Hobbs (2015). The other case, when admissible models contain simultaneously parametric and nonparametric specifications, seems underdeveloped, Hendry et al. (2008). Encompassing tests appear to be helpful for the latter situation where an encompassing model is intended to account the salient feature of the encompassed model. Therefore, encompassing test can detect redundant models among the admissible models.

London Journal of Research in Science: Natural and Formal

Encompassing tests are based on two points, that the encompassing model ought to be able to explain the predictions and to predict some mis-specifications of the encompassed model, Hendry et al. (2008). We know that there are various considerations and developments of encompassing tests, we refer readers to Mizon (1984), Hendry and Richard (1989), Gouriéroux and Monfort (1995) and Florens et al. (1996). For an overview on the concept of encompassing tests, see Bontemps and Mizon (2008) and Mizon (2008). Applications of encompassing tests can be found inside the model selection procedure of general to specific (GETS) modelling developed by Hendry and Doornik (1994), Hoover and Perez (1999). For application in real data, see Nazir (2017).

Recently, **??**ontemps et al. (2008) have developed encompassing tests which cover large set of methods such as parametric and nonparametric methods. Among their results, encompassing tests for kernel nonparametric regression method are established. They provide asymptotic normality of the associated encompassing statistics under the independent and identically distributed hypothesis (i.i.d). We extend their results by relaxing the independent hypothesis. We then focus on processes with some dependence structures. This extension lies on the generalization of encompassing test to dependent processes.

The paper is organized as follows. In section 2, we provide an overview of encompassing test. In section 3, we study the asymptotic behaviors of the encompassing test associated to the linear parametric modelling and the kernel nonparametric method. In last section, we conclude. The encompassing statistic is given by the difference between ?M 2 and ?(?M 1) scaled by a coefficient a n . Specification of the encompassing test will depend on the estimation of the regression method: parametric or nonparametric methods.

2 II. ENCOMPASSING TEST

3 London Journal of Research in Science: Natural and Formal

Let S = (Y, X, Z) be a zero mean random process with valued in RxR d xR q where d, q ? N * . For x ? R d and z ? R q , consider the two models M 1 and M 2 as the conditional expectations m(x) and g(z), respectively. These can be defined as follows:M 1 : m(x) = E[Y |X = x] and M 2 : g(z) = E[Y |Z = z] (2.1)

Moreover, the general unrestricted model is given by r(x, z) = E[Y |X = x, Z = z].

Following the encompassing test in **??**ontemps et al. (2008), we are interested in testing the hypothesis that M 1 encompasses M 2 , and then introducing the null hypothesis:H : E[Y |X = x, Z = z] = E[Y |X = x].**(2.2)**

This null states that M 1 is the owner model, and M 2 will be served on validating this statement and is called the rival model. We test this hypothesis H through the following implicit encompassing hypothesis:H * : E[E[Y |X = x]/Z = z] = E[Y |Z = z].**(2.3)**

1 Kullback-Leiber Information Criterion

The following homoskedasticity condition will be assumed all along this work:V ar[Y |X = x, Z = z] = ? 2 .**(2.4)**

Moreover, a necessary condition for the encompassing test relies on the errors of both models where the intended encompassing model M 1 should have smaller standard error than the encompassed model M 2 .

In general, M 1 or M 1 can be estimated using nonparametric or parametric regression method.

We will consider these different situations when the processes (S n) n are dependent. We begin by constructing the encompassing statistic associated to each of these four situations and then discuss their asymptotic behaviors.

4 III. ASYMPTOTIC BEHAVIOR OF THE ENCOMPASSING STATISTIC

We are interested on the asymptotic behavior of the encompassing statistic associated to the null hypothesis M 1 EM 2 . We can encounter the following four situations: M 1 and M 2 are both estimated parametrically, M 1 and M 2 are both estimated nonparametrically, M 1 is estimated nonparametrically and M 2 parametrically and M 1 is estimated parametrically and M 2 nonparametrically. We will consider the kernel regression estimate for nonparametric methods and the linear regression for parametric methods. For both dependent processes, we will study and establish the asymptotic normality of the corresponding four encompassing tests.

5 London Journal of Research in Science: Natural and Formal

Consider a sample S i = (Y i , X i , Z i), i = 1, ..., n, which can be viewed as realization of the random process S = (Y, X, Z) with valued in RxR d xR q where d, q ? N * . We suppose that S i , i = 1, ..., n has a joint density f . Moreover, ?(., .), ?(. | .) and ?(.) will denote the joint, the conditional and the marginal densities of the process (Y, Z), respectively. That is, for y ? R and z ? R q , ?(y, z), ?(y | z) and ?(z) correspond to the density of the following processes (Y, Z) at point (y, z), (Y | Z = z) at point y and Z at point z, respectively. Similarly, h will denote the joint, the conditional and the marginal densities of the process (Y, X), according to the argument that it takes.

To get the asymptotic normality of the associated encompassing statistic, we need the following assumptions from Bosq (1998). The first assumption characterizes the dependence structure.Assumption 3.1. (S t) is ?-mixing with ?(n) = O(n -?) where ? > ? 2 +4

2? for some positive ?.

The next assumption collects regularity conditions on the continuity and on the differentiability of the density functions. with ? (2) denotes any partial derivative of order 2 for ?. Next,Sup t?k ||?(Z 1 , Z t)|| ? < ? and last, ?(.)E[Y 2 1 |Z 1 = .

] is continuous at z.

The last assumption concerns finiteness of the moments of (Y n , Z n) n .Assumption 3.3. ||E[|Y 1 | 4+? |Z 1 = .]|| ? < ?; E[|Z 1 | 4+?] < ? for some positive ?; Sup t?N ||E[Y i 1 Y j t |Z t = ., Z 1 = .]|| ? < ? where i ? 0, j ? 0, i + j = 2.

Throughout this section, we assume the existence of continuous version of the various joint and marginal density functions and of the three conditional means m, g and r. In addition, the square integrability will be assumed.

For more precision, N (µ, v) will denote the Gaussian distribution with mean µ and variance v.

We now consider the first case that is the encompassing test when the two models M 1 and M 2 have parametric specification. London Journal of Research in Science: Natural and Formal

6 Parametric modelling for M1 and M2

Encompassing test for parametric modelling has been developed a lot in the literature. We discuss briefly one parametric encompassing test where models M 1 and M 2 have linear parametric specification. In that case, the two models M 1 and M 2 are given in relation (3.1) with the nesting model r:m(x) = ? x with ? = (E[XX]) -1 E[XY] g(z) = ? z with ? = (E[ZZ]) -1 E[ZY] r(x,z) = ? w with ? = E[W W] -1 E[W Y] and W = (X, Z). (3.1)

We can get the estimates ?, ? and ? of the paramaters ?, ? and ?, respectively, using the sampleS i = (Y i , X i , Z i), i = 1, ..., n. Now, testing M 1 EM 2 corresponds

7 to the test of the null hypothesis

H where the conditional mean is just the linear projection. Therefore, the encompassing statistic of the null M 1 EM 2 can be written as follows.??,? = ? -?L (?),**(3.2)**

where ?L (?) is an estimate of the pseudo-true value ? L (?) associated with ? on H 1 . Remarking that the pseudo-true value is defined by ? L (?) = (E[ZZ]) -1 E[ZX]?,? n ??,? ? N (0, ? 2 ?) in distribution as n ? ?. (3.3)

where? = V ar(Z) -1 E[V ar(Z | X)]V ar(Z) -1 .

For development on this asymptotic behavior of the encompassing statistic, we refer to Gouriéroux Next, we will study the completely nonparametric case.

London Journal of Research in Science: Natural and Formal

8 Nonparametric modelling for M1 and M2

We now consider the case where the two models M 1 and M 2 defined in (2.1) are estimated using nonparametric techniques. To test the hypothesis " M 1 encompasses M 2 ", we build the corresponding encompassing statistic and establish asymptotic property of such statistic.

Considering the functional estimates m n and g n of the unknown functions m and g in relation (2.1) respectively, we define the encompassing statistic as follows:?m,g (z) = g n (z) -?(m n)(z),**(3.4)**

where ?(m n) is an estimate of the pseudo true value G(m) associated with g n on H, which isdefined by G(m) = E[m | Z = z].

Using the sample S i = (Y i , X i , Z i), i = 1, ..., n, the kernel regression estimates m n of the function m, and g n of the function g have the following expressions:m n (x) = 1 nh d 1n n i=1 K 1 (x-X i h 1n)Y i 1 nh d 1n n i=1 K 1 (x-X i h 1n) g n (z) = 1 nh q 2n n i=1 K 2 (z-Z i h 2n)Y i 1 nh q 2n n i=1 K 2 (z-Z i h 2n)**(3.5)**

where h jn and K j , j = 1, 2 are window widths and kernel densities, respectively. The kernel densities satisfyK j (u) ? 0 and K j (u)du = 1 j = 1, 2.**(3.6)**

We provide in the following, a theorem establishing the asymptotic convergence of the encompassing statistic.

Theorem 3.2. Suppose that assumptions 3.1-3.3 hold. Moreover, suppose that relation (3.6) is satisfied. Then, under H, we get:nh q 2n ?m,g (z) ? N (0, ? 2 K 2 2 (u)du ?(z)) in distribution as n ? ?.**(3.7)**

?(z) is the marginal density of the Z at z and? 2 = V ar[Y |X = x, Z = z].

Proof of theorem 3.2 The proof of this theorem will be based on the decomposition of the expression of the encompassing statistic into two parts as follows: London Journal of Research in Science: Natural and Formalnh q 2n ?m,g (z) = nh q 2n (g n (z) -?(m n)(z)) = nh q 2n (n t=1 K 2 (z-Zt h 2n) n t=1 K 2 (z-Zt h 2n) Y t - n t=1 K(z-Zt hn) n t=1 K(z-Zt hn) m n (x t)) = nh q 2n n t=1 K 2 (z-Zt h 2n) n t=1 K 2 (z-Zt h 2n) (Y t -m(x t)) + nh q 2n n t=1 K(z-Zt hn) n t=1 K(z-Zt hn) (m(x t) -m n (x t)) = C 1 + C 2 .

(3.8)

The first part C 1 coincides to the kernel regression of the residuals t = Y t -m(x t) onto Z t .

When assumptions 3.1-3.3 hold, then under H, we achieved the convergence in distribution of the first part to a normal distribution using Rhomari's result in Bosq (1998). The second part C 2 reflects the limit in probability of the supremum of the difference m n (x t) -m(x t) at x t ? R d scaled by nh q n and its convergence can be derived from the rate of covergence of the uniform convergence of the estimate m n (x t) which has been provided by Bosq (1998).

9 Parametric modelling M1 vs nonparametric modelling M2

We consider the case that model M 1 is a linear parametric model and M 2 is estimated by kernel nonparametric technique. Therefore, the hypothesis H will have linear parametric specification.

The encompassing statistic associated to the null M 1 EM 2 can be written as follows:??,g (z) = g n (z) -?L (?)(z),**(3.9)**

where ?L (?) is an estimate of the pseudo-true value G L (?)(z) associated with g n on H, which is defined by GL (?)(z) = ? E[X | Z = z].

For the nonparametric specification of M 2 , we consider the estimate g n as the kernel regression estimate of g given in (2.1). Since the rival model g is estimated using kernel method, the various assumptions on kernel density and window width will be maintained.

Even the process exhibits some dependences, we can still establish the asymptotic normality of the encompassing statistic defined in relation (3.9). London Journal of Research in Science: Natural and Formal with linear specification and when the bandwidth h 2n satisfy kernel regularity condition, we get:nh q 2n ??,g (z) ? N (0, ? 2 K 2 2 (u)du ?(z)

) in distribution as n ? ?.

(3.10) ?(z) is the marginal density of the Z at z.

Proof of theorem 3.3 Using similar techniques as previously, we can write the encompassing statistic as follows: converges in distribution to zero. This completes the proof.nh q 2n ??,g (z) = nh q 2n (g n (z) -?L (?)(z)) = nh q 2n (n t=1 K 2 (z-Zt hn) n t=1 K 2 (z-Zt h 2n) Y t - n t=1 K(z-Zt hn) n t=1 K(z-Zt hn) ? X t) = nh q 2n n t=1 K 2 (z-Zt hn) t=1 K 2 (z-Zt h 2n) (Y t -? X t) + nh q

10 Nonparametric modelling M1 vs parametric modelling M2

We consider the owner model M 1 to be estimated using a nonparametric method and the rival model M 2 to be a linear parametric method. Therefore, the encompassing statistic associated to the null M 1 EM 2 is given by:?m,? = ? -?(m n),**(3.12)**

where ?(m n) is an estimate of the pseudo-true value ?(m) associated with ? on H, which isdefined by ?(m) = (E[ZZ]) -1 E[Zm].

When the estimate of model M 1 is obtained from the kernel regression and the model M 2 is from linear parametric modelling, we summarize the asymptotic results in the following theorem. London Journal of Research in Science: Natural and Formal [1] [2] [3] [4] [5] [6]

32

Figure 1: Assumption 3 . 2 .

Figure 2:

[1] © 2023 Great] Britain Journals Press

[2] © 2023 Great] Britain Journals Press

[3] ©

[4] Great] Britain Journals Press

[5] London Journal of Research in Science: Natural and Formal

[6] © 2023 Great] Britain Journals Press

33

Figure 3: Theorem 3 . 3 .

22

Figure 4: 2 ?(z) K 2 2

Funding statement : this work was performed as part of my employement at the Ministry of Higher Education and Scientific Research of Madagascar. London Journal of Research in Science: Natural and Formal

Theorem 3.4. Assume that relation ??2.4) is satisfied. When the kernel K 1 and the bandwidth h 1n satisfy the usual regularity condition and when we have one of the following points: Assumption 3.1 holds and the kernel regression estimate m n and the process (Y n , X n) n satisfy assumptions 3.2 and 3.3.

Then, under H, we get:

where ? = plim n?? V ar(? n ?m,?).

Proof of theorem 3.4 We split the encompassing statistic ? n ?m,? into two parts. The first part yields

) which gives the asymptotic normality of the theorem, Peligrad and Utev (1997).

The second part is

. Again, we bound this by taking the supremum with respect to x i . Thus, F 2 vanishes to zero from the uniform convergence of m n (x i), Bosq (1998). This completes the proof of theorem 3.4.

We remark that we should be careful about mutual encompassing of both models which concerns the bijection of the pseudo true value function G(.).

.1 IV. CONCLUSION

We have considered encompassing test for functional parameters. As stated in Hendry et al. When using kernel method and linear regression as estimator of conditional expectations, we have established asymptotic normality of the encompassing test for dependent processes. These results would be helpfull for analysing non-nested non-parametric and parametric models.

[Oxford Bulletin of Economics and Statistics] , *Oxford Bulletin of Economics and Statistics* 70 p. .

[Bulletin of Economics and Statistics] , *Bulletin of Economics and Statistics* 70 p. .

[Hooten and Hobbs ()] ‘A guide to bayesian model selection for ecologists’. M B Hooten , N T Hobbs . *Ecological monographs* 2015. 85 (1) p. .

[Ferraty and Hall ()] ‘An algorithm for nonlinear, nonparametric model choice and Prediction’. F Ferraty , P Hall . *J. Comput. Graph. Stat* 2015. 24 p. .

[Bontemps and Mizon ()] C Bontemps , G E Mizon . *Encompassing: concepts and implementation*, (Oxford) 2008.

[Bosq ()] D Bosq . *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction*, 1998.

[Peligrad and Utev ()] ‘Central limit theorem for Linear Processes’. M Peligrad , S A Utev . *The Annals of Probability* 1997. 25 p. .

[Li et al. ()] ‘Cross-validated mixed-datatype bandwidth selection for nonparametric cumulative distribution/survivor functions’. C Li , Q Li , J S Racine . *Econometric Reviews* 2017. 36 (6-9) p. .

[Hoover and Perez ()] ‘Data mining reconsidered: encompassing and the general-to-specific approach to specification search’. K D Hoover , S J Perez . *Econometrics Journal* 1999. 2 p. .

[Greene (ed.) ()] *Econometric analysis*, W H Greene . NJ. Hamilton J.D. (ed.) 2003. 1994. Princeton, NJ: Princeton University Press. (Time series analysis)

[Hendry and Nielsen ()] *Econometric modeling: a Likelihood Approach*, D F Hendry , B Nielsen . 2006. Princeton: Princeton University Press.

[Hendry et al. ()] ‘Encompassing’. D F Hendry , M Marcellino , G E Mizon . *Special Issue: Oxford 15*, 2008. Guest Editor Introduction.

[Mizon ()] ‘Encompassing’. G E Mizon . http://www.dictionaryofeconomics.com/article?id=pde2008E000243 *The New Palgrave Dictionary of Economics, 2nd edn. The New Palgrave Dictionary of Economics Online*, L E Blume , S N Durlauf (eds.) 2008. Palgrave Macmillan.

[Florens et al. ()] ‘Encompassing and specificity’. J P Florens , D F Hendry , J F Richard . *Econometric Theory* 1996. 12 p. .

[Govaerts et al. ()] ‘Encompassing in stationary linear dynamic models’. B Govaerts , D F Hendry , J F Richard . *Journal of Econometrics* 1994. 63 p. .

[Nazir ()] ‘Encompassing of nested and non-nested models: energy-growth models’. S Nazir . https://mpra.ub.uni-muenchen.de/77487/ *MPRA Paper N°77487*, 2017.

[Sawa ()] ‘Information criteria for discriminating among alternative regression models’. T Sawa . *Econometrica* 1978. 46 p. .

[Lecture Notes in Statistics] *Lecture Notes in Statistics*, (Berlin) Springer-Verlag.

[Hendry and Doornik ()] ‘Modelling linear dynamic econometric systems’. D F Hendry , J A Doornik . *Scottish Journal of Political Economy* 1994. 41 p. .

[Bontemps et al. ()] *Parametric and non-parametric encompassing*, C Bontemps , J P Florens , J F Richard . 2008.

[Hendry and Richard ()] 'Recent developments in the theory of encompassing'. D F Hendry , J F Richard . *Contributions to Operations Research and Economics*, B Cornet , H Tulkens (eds.) (Cambridge) 1989. p. .

[Tibshirani et al. ()] *Statistical learning with sparsity : the lasso and generalizations*, R Tibshirani , M Wainwright , T Hastie . 2015. Chapman and Hall.

[Gouriéroux et al. ()] 'Testing nested or non-nested hypotheses'. C Gouriéroux , A Monfort , A Trognon . *Journal of Econometrics* 1983. 21 p. .

[Gouriéroux and Monfort ()] 'Testing, encompassing, and simulating dynamic econometric models'. C Gouriéroux , A Monfort . *Econometric Theory* 1995. 11 p. .

[Mizon ()] 'The encompassing approach in econometrics'. G E Mizon . *Econometrics and Quantitative Economics*, D F Hendry , K F Wallis (eds.) 1984. p. .

[Mizon and Richard ()] 'The encompassing principle and its application to non-nested hypothesis tests'. G E Mizon , J F Richard . *Econometrica* 1986. 54 p. .