General considerations on finite-sample inference in econometrics and statistics [*]

Jean-Marie Dufour [†]

First version: February 2002
This version: March 29, 2002, 5:10pm

[*] This work was supported by the Canada Research Chair Program (Chair in Econometrics, Université de Montréal), the Canadian Network of Centres of Excellence [program on *Mathematics of Information Technology and Complex Systems* (MITACS)], the Canada Council for the Arts (Killam Fellowship), the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Fonds FCAR (Government of Québec).

[†] Canada Research Chair Holder (Econometrics). Centre de recherche et développement en économique (C.R.D.E.), Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Département de sciences économiques, Université de Montréal. Mailing address: Département de sciences économiques, Université de Montréal, C.P. 6128 succursale Centre-ville, Montréal, Québec, Canada H3C 3J7. TEL: 1 514 343 2400; FAX: 1 514 343 5831; e-mail: jean.marie.dufour@umontreal.ca. Web page: http://www.fas.umontreal.ca/SCECO/Dufour .

Contents

1. Valid tests 1

2. Approaches to deal with nuisance parameters 3

2.1. **Bounding** 4

2.2. Conditioning 4

2.3. Transforming 4

1. Valid tests

$$Y_n : \text{Vector of observations } (n \times 1)$$
$$Y_n \sim P_\theta$$

θ : Vector of unknown parameters
Finite (Parametric problems)
or
Infinite (Non parametric problem)

$$\theta \in \Omega \tag{1.1}$$

$$H_0 : \theta \in \Omega_0 \tag{1.2}$$

$$\phi \neq \Omega_0 \subset \Omega \tag{1.3}$$

Consider critical region of the form:

$$T_n \equiv T_n(Y_n) > c_n(\alpha) \tag{1.4}$$

We reject H_0 if

$$T_n \geq c_n(\alpha)\,,\ 0 \leq \alpha \leq 1\,. \tag{1.5}$$

The test is valid at **level** α if

$$P_\theta\left[T_n > c_n(\alpha)\right] \leq \alpha \ \ ,\ \forall \theta \in \Omega_0\,.$$

The test has **size** α if

$$\sup_{\theta \in \Omega_0} P_0 \left[T_n > c_n(\alpha) \right] = \alpha \, ; \tag{1.6}$$

see Lehmann (1986, Section 3.1).
In general,

$P_0 \left[T_n > c_n(\alpha) \right]$ depends on $\theta \in \Omega_0$,
if $P_\theta \left[T_n \geq c_n(\alpha) \right]$ = constant similar test.

If no data-dependent critical region which satisfies the level constraint can be found,

$\theta \in \Omega_0$ is not testable at level α.
We have a non-sensical problem.
We must restrict the hypothesis model up to the point where it becomes testable.

In many econometric and statistical problems appropriate $c_n(\alpha)$ difficult to determine. The usual strategy consists in finding $c_n'(\alpha)$ such that

$$\lim_{n \to \infty} P_0 \left[T_n > c_n'(\alpha) \right] = \alpha \, , \quad \forall \theta \in \Omega_0 \, . \tag{1.7}$$

But this does not entail that

$$\lim_{n \to \infty} \left\{ \sup_{\theta \in \Omega_0} P \left[T_n > c_n'(\alpha) \right] \right\} = \alpha \, . \tag{1.8}$$

Possible that (1.7) be satisfied, while

$$\sup_{\theta \in \Omega_0} P\left[T_n > c_n'(\alpha)\right] = 1 \;, \quad \forall n \geq 1 \,.$$

In such a case, the test is "asymptotically valid" in the usual sense but invalid for any finite sample size n.

1.1 Example Linear regression with AR(1) errors:

$$\begin{aligned} y_t &= x_t'\beta + u_t && (1.9)\\ u_t &= \rho u_{t-1} + \varepsilon_t \;, \quad |\rho| < 1, && (1.10)\\ &\varepsilon_t \overset{i.i.d}{\sim} N\left(0, \sigma^2\right) \end{aligned}$$

In such a case, the usual uncorrected F-test for linear restrictions on β is asymptotically valid but invalid for any finite sample [Kiviet (1980, JASA, 353-359)].

2. Approaches to deal with nuisance parameters

3 main approaches:

1. Bounding;
2. Conditioning;
3. Transforming.

2.1. Bounding

Bound the distribution of T_n over $\theta \in \Omega_0$

- Dufour (1989, Econometrica, 335-355)
- Dufour (1990, Econometrica, 475-494)
- Dufour (1991, Hackl and A. Westlund, eds, 49-57)
- Kiviet (1991, J. Gruber, ed., 464-518)

2.2. Conditioning

Consider the conditional distribution of T_n given some appropriate statistic, so that conditional distribution has no nuisance parameter, e.g.

- Regressors (linear regression)
- Sufficient statistic

$$\left.\begin{array}{l}\text{- Order statistics}\\ \text{Absolute values}\end{array}\right\}\text{Nonparametric statistic}$$

2.3. Transforming

1. Transform the observations so that the transformed observations have distributions with no nuisance parameter

- t-tests in multiple linear regression
- Ranks
- Signs

2. Choice of transformations based on invariance principles
3. Sign and signed rank tests based on techniques conditioning and transformations
4. Transforming in an appropriate way

$$\text{eliminates an infinity of nuisance parameters.} \tag{2.11}$$

References

DUFOUR, J.-M. (1989): "Nonlinear Hypotheses, Inequality Restrictions, and Non-Nested Hypotheses: Exact Simultaneous Tests in Linear Regressions," *Econometrica*, 57, 335–355.

——— (1990): "Exact Tests and Confidence Sets in Linear Regressions with Autocorrelated Errors," *Econometrica*, 58, 475–494.

——— (1991): "Kimball's Inequality and Bounds Tests for Comparing Several Regressions under Heteroskedasticity," in *Economic Structural Change. Analysis and Forecasting*, ed. by P. Hackl, and A. Westlund, pp. 49–57. Springer-Verlag, Berlin.

KIVIET, J. F. (1980): "Effects of ARMA Errors on Tests for Regression Coefficients: Comments on Vinod's Article; Improved and Additional Results," *Journal of the American Statistical Association*, 75, 353–358.

LEHMANN, E. L. (1986): *Testing Statistical Hy-*

potheses, 2nd edition. John Wiley & Sons, New York.