# Généralisation des tests LMC et KPSS : apparition d'une singularité

**Frédéric Proïa**

*Université de Bordeaux, Institut de Mathématiques de Bordeaux, UMR CNRS 5251.*
*351 cours de la libération, 33405 Talence cedex, France.*
`frederic.proia@math.u-bordeaux1.fr`

**Résumé.** Nous abordons la problématique de la stationnarité d'une trajectoire autorégressive munie d'une tendance polynomiale, et nous généralisons sous certains aspects le test LMC, la procédure de Leybourne et McCabe. Après avoir présenté succinctement les généralisations considérées, nous insisterons dans cette communication sur une particularité survenant dans une situation de non stationnarité souvent négligée : lorsque le polynôme autorégressif contient une racine unitaire négative. Nous expliquerons ainsi la raison pour laquelle le test LMC, et par extension le test KPSS, ne rejettent pas l'hypothèse nulle de stationnarité en tendance, à tort, lorsqu'une racine unitaire est située en $-1$. Nous le constaterons également sur des données simulées. Finalement, nous décrirons les processus stochastiques apparaissant dans nos distributions limites.

**Mots-clés.** Test LMC, Test KPSS, Racine unitaire, Test de stationnarité, Tendance polynomiale, Non stationnarité stochastique, Marche aléatoire, Processus ARIMA.

**Abstract.** We tackle the stationarity issue of an autoregressive path with a polynomial trend, and we generalize some aspects of the LMC test, the testing procedure of Leybourne and McCabe. After discribing succinctly the considered generalizations, one shall insist in this talk on a feature occurring in a nonstationarity framework often neglected: when the autoregressive polynomial has a negative unit root. Hence, we will explain the reason why the LMC test, and by extension the KPSS test, do not reject the null hypothesis of trend-stationarity, mistakenly, when the random walk is generated by a unit root located at $-1$. We will also observe it on simulated data. Finally, we will describe the stochastic processes that appear in our limiting distributions.

**Keywords.** LMC test, KPSS test, Unit root, Stationarity testing procedure, Polynomial trend, Stochastic nonstationarity, Random walk, ARIMA process.

**Historique.** Il est bien connu que l'estimateur des moindres carrés du paramètre $\theta$ d'un modèle AR(1) satisfait, sous l'hypothèse nulle $\mathcal{H}_0$ : "$\theta = 1$", la convergence

$$T\left(\widehat{\theta}_T - 1\right) \xrightarrow{\mathcal{L}} \frac{\int_0^1 V(s)\ \mathrm{d}W(s)}{\int_0^1 V^2(s)\,\mathrm{d}s}$$

où le processus $(V(t),\, t \in [0,1])$ dépend de l'ordre de la tendance déterministe estimée et $(W(t),\, t \in [0,1])$ est le processus de Wiener. Ce résultat avait été conjecturé par White [12] en 1958 sans tendance et sous une forme moins élégante, puis progressivement amélioré, dans le cas d'une perturbation gaussienne tout d'abord (Dickey et Fuller [3] en 1979) puis dans le cas d'une perturbation indépendante et identiquement distribuée (Phillips [8] en 1987). Enfin, Chan et Wei [2] en 1988 établissent le résultat lorsque la perturbation forme une différence de martingale possédant certaines propriétés de moments conditionnels. Cette procédure de test est aujourd'hui connue sous la dénomination de *test de Dickey-Fuller* et fut étendue aux processus $\text{ARMA}(p,q)$ par l'intermédiaire de Dickey et Said [4]–[11] dans les années 80, on parle alors de *test de Dickey-Fuller Augmenté*. Nous pouvons également penser au *test de Phillips-Perron* dans lequel Phillips et Perron [9] proposent, en 1988, de corriger la statistique de test de façon non paramétrique afin d'estimer une variance de long terme et d'ainsi s'exempter d'une évaluation nécessairement arbitraire des ordres $p$ et $q$ : on translate en quelque sorte tout phénomène d'autocorrélation dans la perturbation. En parallèle, Kwiatkowski, Phillips, Schmidt et Shin [5] établissent en 1992 une procédure statistique prenant pour hypothèse nulle la stationnarité de la trajectoire autorégressive considérée, alors que Leybourne et McCabe [6] en 1994 en développent une variante paramétrique. Les *test KPSS* et *test LMC* sont aujourd'hui les procédures de test de stationnarité les plus couramment utilisées, et elles reposent sur la statistique

$$\widehat{K}_T = \frac{1}{TQ_T} \sum_{t=1}^{T} S_t^2 \quad \text{où} \quad S_T = \sum_{t=1}^{T} \widehat{\varepsilon}_t \quad \text{et} \quad Q_T = \sum_{t=1}^{T} \widehat{\varepsilon}_t^{\,2}.$$

L'ensemble résiduel $(\widehat{\varepsilon}_t)$ provient de l'estimation éventuelle de la tendance polynomiale couplée à celle de l'autorégression, dans le cas LMC. Il est alors montré que, sous l'hypothèse de stationnarité,

$$\widehat{K}_T \xrightarrow{\mathcal{L}} \int_0^1 B^2(s)\, \mathrm{d}s$$

où le processus $(B(t),\, t \in [0,1])$ décrit une famille de ponts browniens, selon l'ordre de la tendance estimée. Les tests KPSS et LMC ne tiennent compte que d'une tendance constante ou linéaire ($r = 0$ ou $r = 1$ avec les notations que nous utiliserons par la suite).

**Généralisations du test LMC.** Considérons un processus autorégressif muni d'une tendance polynomiale d'ordre $r$ et engendré par une perturbation intégrée. Le processus $(Y_t)$ est donné, pour tout $1 \leq t \leq T$, par

$$\mathcal{A}(L) Y_t = (\alpha_0 + \alpha_1 t_T + \ldots + \alpha_r t_T^r)\, \mathbb{I}_{\{\kappa \neq 0\}} + S_t^{\eta} + \varepsilon_t$$

où, pour tout $z \in \mathbb{C}$, $\mathcal{A}(z) = 1 - \theta_1 z - \ldots - \theta_p\, z^p$ est un polynôme causal, $L$ est l'opérateur retard et $\kappa$ est un indicateur. De plus, pour $|\rho| = 1$,

$$S_t^{\eta} = \rho S_{t-1}^{\eta} + \eta_t$$

et $(\varepsilon_t)$ et $(\eta_t)$ sont deux bruits blancs mutuellement indépendants, de variances respectives $\sigma_\varepsilon^2 \geq 0$ et $\sigma_\eta^2 \geq 0$. Enfin, la tendance est renormalisée par la taille de l'échantillon et nous notons $t_T = t/T$. Il est clair que le processus possède une non stationnarité de nature stochastique sous l'alternative $\mathcal{H}_1$ : "$\sigma_\eta^2 > 0$" alors que, sous l'hypothèse nulle $\mathcal{H}_0$ : "$\sigma_\eta^2 = 0$", $(Y_t)$ est un processus stationnaire en tendance.

Soit alors $\widetilde{\theta}_T$ l'estimateur du maximum de vraisemblance de $\theta$ dans le processus correctement différencié $((1-\rho L)Y_t)$, consistant sous $\mathcal{H}_0$ comme sous $\mathcal{H}_1$ (voir par exemple Brockwell et Davis [1] ou Pötscher [10]). Pour tout $1 \leq t \leq T$, on construit le processus résiduel

$$\widetilde{Y}_t = Y_t - \widetilde{\theta}_1 Y_{t-1} - \ldots - \widetilde{\theta}_p Y_{t-p}$$

et l'on estime $\alpha$ par moindres carrés dans le modèle de régression donné par

$$\widetilde{Y}_t = (\alpha_0 + \alpha_1 t_T + \ldots + \alpha_r t_T^r)\, \mathbb{I}_{\{\kappa \neq 0\}} + \varepsilon_t.$$

Nous montrons alors que, sous $\mathcal{H}_0$,

$$\widehat{K}_T \xrightarrow{\mathcal{L}} \int_0^1 B_r^2(s)\,\mathrm{d}s$$

où $(B_r(t),\, t \in [0,1])$ décrit une famille de ponts browniens généralisés. Nous montrons également que, sous $\mathcal{H}_1$ et sachant que $\rho = 1$, alors

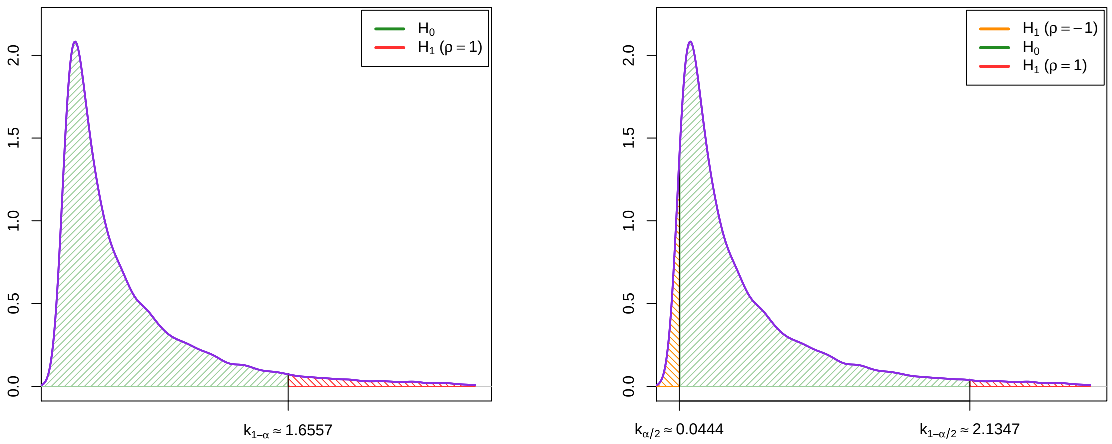
$$\frac{\widehat{K}_T}{T} \xrightarrow{\mathcal{L}} \frac{\int_0^1 C_{r,1}^2(s)\,\mathrm{d}s}{\int_0^1 W_{r,0}^2(s)\,\mathrm{d}s}$$

où $(C_{r,1}(t),\, t \in [0,1])$ et $(W_{r,0},\, t \in [0,1])$ décrivent des familles de ponts browniens intégrés et de processus de Wiener recentrés. Lorsque la perturbation est multi-intégrée d'ordre $d$, et donc que $(Y_t)$ suit la dynamique de modèles ARIMA d'ordres supérieurs, nous établissons les mêmes résultats asymptotiques où les lois limites décrivent des familles de processus plus complexes, mais que nous sommes capables de décrire explicitement grâce aux travaux de MacNeill [7] de 1978, ce que nous nous proposons de faire durant la présentation.

**Une singularité.** Nous montrons enfin que, lorsque $\rho = -1$, alors la statistique de test dégénère vers 0. Nous avons alors, sous $\mathcal{H}_1$ et sachant que $\rho = -1$, la singularité

$$\widehat{K}_T \xrightarrow{\mathcal{P}} 0.$$

Nous établissons de plus le comportement asymptotique de la statistique $\widehat{K}_T$ correctement renormalisée (ici multipliée par $T$) dans le cas particulier où $\kappa = 0$. Ces derniers résultats sont très importants, car ils montrent que les cas de non stationnarité engendrés par une racine unitaire localisée à $-1$ dans le polynôme autorégressif ne sont pas détectés par les tests KPSS et LMC qui ne rejettent donc pas, à tort, l'hypothèse nulle. En effet, la zone de rejet associée à la loi limite sous $\mathcal{H}_0$ est unilatérale à droite. En conséquence, nous proposons de revoir la règle de décision, comme le montre la figure ci-dessous (à gauche : les tests LMC et KPSS, à droite : une procédure tenant compte de la singularité).




Nous terminons l'étude par des résultats de simulations dans lesquels nous illustrons la forme des distibutions asymptotiques sous $\mathcal{H}_0$ comme sous $\mathcal{H}_1$, pour des valeurs grandissantes de $r$. Nous mettons également en pratique notre procédure de test pour évaluer sa puissance empirique et bien contaster qu'elle détecte tous les cas de non stationnarité stochastique, avec un risque d'erreur caractérisant sa zone de rejet.

# References


[1] Brockwell, P. J., and Davis, R. A. *Introduction to Time Series and Forecasting.* Springer-Verlag, New-York, 1996.

[2] Chan, N. H., and Wei, C. Z. Limiting distributions of least squares estimates of unstable autoregressive processes. *Ann. Statis. 16-1* (1988), 367–401.

[3] Dickey, D. A., and Fuller, W. A. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc. 74-366* (1979), 427–431.

[4] Dickey, D. A., and Said, E. S. Testing ARIMA$(p, 1, q)$ versus ARMA$(p+1, q)$. *Proc. Bus. Econ. Statist. Sect., Am. Statist. Assoc.* (1981), 318–322.

[5] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y. Testing the null hypothesis of stationarity against the alternative of a unit root : How sure are we that economic time series have a unit root? *J. Econometrics. 54* (1992), 159–178.

[6] Leybourne, S. J., and McCabe, B. P. M. A consistent test for a unit root. *J. Bus. Econ. Stat. 12-2* (1994), 157–166.

[7] MacNeill, I. B. Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times. *Ann. Statis. 6-2* (1978), 422–433.

[8] Phillips, P. C. B. Time series regression with a unit root. *Econometrica. 55* (1987), 277–302.

[9] Phillips, P. C. B., and Perron, P. Testing for a unit root in time series regression. *Biometrika. 75-2* (1988), 335–346.

[10] Pötscher, B. M. Noninvertibility and pseudo-maximum likelihood estimation of misspecified ARMA models. *Economet. Theor. 7* (1991), 435–449.

[11] Said, E. S., and Dickey, D. A. Testing for unit roots in autoregressive moving average models of unknown order. *Biometrika. 71-3* (1984), 599–607.

[12] White, J. S. The limiting distribution of the serial correlation coefficient in the explosive case. *Ann. Math. Statist. 29* (1958), 1188–1197.