# Asymptotic distribution of least square estimators for linear models with dependent errors

**Emmanuel Caron** [1]

http://ecaron.perso.math.cnrs.fr

Directeurs de thèse: Jérôme Dedecker (MAP5), Bertrand Michel (LMJL)

18 Septembre 2018

[1]Ecole Centrale Nantes, Laboratoire de Mathématiques Jean Leray UMR 6629, 1 Rue de la Noë, 44300 Nantes. **Email: emmanuel.caron@ec-nantes.fr**

# Sommaire

# Sommaire

# Linear Regression Model

Linear Regression model:

$$Y = X\beta + \epsilon,$$

- $X$ is a design, random or not, size $[n \times p]$
- $Y$ is a $n$ random vector
- $\beta$ is a $p$ vector of unknown parameters
- $\epsilon$ are the errors, $\epsilon \in \mathbb{R}^n$. The error process is independent of the design $X$.

Usual assumptions:

- the errors are i.i.d.
- $\mathbb{E}(\epsilon) = 0$ and $Var(\epsilon) = \sigma^2 I_n$
- Sometimes, $\epsilon \sim \mathcal{N}(0, \sigma^2 I_n)$

# Least Square Estimator

$$\hat{\beta} = argmin_{\beta \in \mathbb{R}^p} \|Y - X\beta\|_2^2 = (X^t X)^{-1} X^t Y.$$

- $\hat{Y} = X\hat{\beta}$: Orthogonal Projection of $Y$ on $\mathcal{M}_X = Vect\{X_{.,1}, ..., X_{.,p}\}$
- Residual vector: $\hat{\epsilon} = Y - \hat{Y} = Y - X\hat{\beta} \in \mathcal{M}_X^{\perp}$
- $\hat{\sigma}^2 = \frac{\|\hat{\epsilon}\|_2^2}{n-p}$.

Distribution of the LSE in the i.i.d. case:

- Gaussian Case: $\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2 (X^t X)^{-1})$
- Non-Gaussian Case: $D(n)(\hat{\beta} - \beta) \xrightarrow[n \to \infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2 Q^{-1})$.

# Goals and Plan

**Main Goal** : Remove the independence hypothesis and correct the results on the linear regression model in a very general framework

**Plan** :

1. Hannan's Theorem (1973) [4]: convergence of the LSE in the stationary case under very mild conditions
2. Estimation of the covariance matrix
3. Applications with Fisher's tests

# Stationarity

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. $(\epsilon_i)_{i\in\mathbb{Z}}$ is an error process defined on $(\Omega, \mathcal{F}, \mathbb{P})$, supposed strictly stationary, with zero mean, and $\epsilon_0 \in \mathbb{L}^2$.

**Definition : Strict Stationarity**

A stochastic process $(\epsilon_i)_{i\in\mathbb{Z}}$ is said to be strictly stationary if the joint distributions of $(\epsilon_{t_1}, \dots, \epsilon_{t_k})$ and $(\epsilon_{t_1+h}, \dots, \epsilon_{t_k+h})$ are the same for all positive integers $k$ and for all $t_1, \dots, t_k, h \in \mathbb{Z}$.

Let $(\mathcal{F}_i)_{i\in\mathbb{Z}}$ be a non-decreasing filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ (Example: $\mathcal{F}_i = \sigma(\epsilon_k, k \leq i)$).
We always suppose that $\mathcal{F}_{-\infty} = \bigcap_{i\in\mathbb{Z}} \mathcal{F}_i$ is trivial, and $\epsilon_0$ $\mathcal{F}_\infty$-measurable.

# Spectral density

Autocovariance function of the error process:

$$\gamma(k) = \mathrm{Cov}(\epsilon_m, \epsilon_{m+k}) = \mathbb{E}(\epsilon_m \epsilon_{m+k}),$$

and the covariance matrix: $\Gamma_n = [\gamma(j-l)]_{1 \leq j,l \leq n}$.
Let $f$ be the associated spectral density, $\lambda \in [-\pi, \pi]$:

$$\gamma(k) = \int_{-\pi}^{\pi} e^{ik\lambda} f(\lambda) d\lambda,$$

$$f(\lambda) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma(k) e^{ik\lambda}.$$

# Sommaire

# Hannan's condition on the error process

Given the design $X$, Hannan has proved a CLT in the stationary case for the usual LSE $\hat{\beta}$ under very mild conditions.

- $\forall j \in \mathbb{Z}$ and $\forall Z \in \mathbb{L}^2(\Omega)$: $P_j(Z) = \mathbb{E}(Z|\mathcal{F}_j) - \mathbb{E}(Z|\mathcal{F}_{j-1})$.
- **Hannan's condition** on the error process:

$$\sum_{i \in \mathbb{Z}} \|P_0(\epsilon_i)\|_{\mathbb{L}^2} < +\infty.$$

This implies: $\sum_k |\gamma(k)| < +\infty$.

**Hannan's condition is satisfied for most of short-range dependent processes.**

# Examples which verify Hannan's condition

- Linear Processes (Dedecker, Merlevède, Vólny (2007) [2])
- Functions of linear processes ([2])
- Conditions à la Gordin ([2])
- Framework of Wu (Wu (2005) [5])
- Weakly dependent sequences (Dedecker-Prieur (2004) [3], Caron-Dede (2017) [1])

# Hannan's conditions on the design

- Let $X_{.,j}$ be the column $j$ of the matrix $X$, $j \in \{1, \ldots, p\}$:

$$d_j(n) = \|X_{.,j}\|_2 = \sqrt{\sum_{i=1}^{n} x_{i,j}^2},$$

and let $D(n)$ be the diagonal matrix with diagonal term $d_j(n)$.

- **Conditions on the design**:

$$\forall j \in \{1, \ldots, p\}, \qquad \lim_{n\to\infty} d_j(n) = \infty \qquad a.s.,$$

$$\forall j \in \{1, \ldots, p\}, \qquad \lim_{n\to\infty} \frac{\sup_{1\leq i\leq n} |x_{i,j}|}{d_j(n)} = 0 \qquad a.s.,$$

and the following limits exist:

$$\forall j, l \in \{1, \ldots, p\}, \qquad \rho_{j,l}(k) = \lim_{n\to\infty} \sum_{m=1}^{n-k} \frac{x_{m,j} x_{m+k,l}}{d_j(n) d_l(n)} \qquad a.s.$$

We define the $p \times p$ matrix $R(k)$:

$$R(k) = [\rho_{j,l}(k)] = \int_{-\pi}^{\pi} e^{ik\lambda} F_X(d\lambda) \qquad a.s.,$$

where $F_X$ is the spectral measure associated with the matrix $R(k)$.

Moreover, we suppose:

$$R(0) > 0 \qquad a.s.$$

Then let $F$ and $G$ be the matrices:

$$F = \frac{1}{2\pi} \int_{-\pi}^{\pi} F_X(d\lambda) \qquad a.s.,$$

$$G = \frac{1}{2\pi} \int_{-\pi}^{\pi} F_X(d\lambda) \otimes f(\lambda) \qquad a.s.$$

## Theorem (Hannan (1973) [4])

*Under the previous conditions, for all bounded continuous function $f$:*

$$\mathbb{E}\left(f\left(D(n)(\hat{\beta}-\beta)\right)\Big|X\right) \xrightarrow[n\to\infty]{a.s.} \mathbb{E}\left(f(Z)\Big|X\right),$$

*where the distribution of $Z$ given $X$ is: $\mathcal{N}(0, F^{-1}GF^{-1})$.*
*Furthermore we have the convergence of second order moment:*

$$\mathbb{E}\left(D(n)(\hat{\beta}-\beta)(\hat{\beta}-\beta)^t D(n)^t\Big|X\right) \xrightarrow[n\to\infty]{a.s.} F^{-1}GF^{-1},$$

## Remark

*Let us notice that, by the dominated convergence theorem, we have for any bounded continuous function f:*

$$\mathbb{E}\left(f\left(D(n)(\hat{\beta}-\beta)\right)\right) \xrightarrow[n\to\infty]{} \mathbb{E}\left(f(Z)\right),$$

# Sommaire

To obtain confidence regions or test procedures, one needs to estimate the limiting covariance matrix $F^{-1}GF^{-1}$. By Hannan, we have:

$$\mathbb{E}\left(D(n)(\hat{\beta}-\beta)(\hat{\beta}-\beta)^t D(n)^t \middle| X\right) \xrightarrow[n\to\infty]{a.s.} F^{-1}GF^{-1},$$

and:

$$\mathbb{E}\left(D(n)(\hat{\beta}-\beta)(\hat{\beta}-\beta)^t D(n)^t \middle| X\right) = D(n)(X^tX)^{-1}X^t\Gamma_n X(X^tX)^{-1}D(n),$$

with $\Gamma_n = [\gamma(j-l)]_{1\leq j,l\leq n}$ (covariance matrix of the error process). Then we need an estimator of $\Gamma_n$.

Let us first consider a preliminary random matrix:

$$\widehat{\Gamma}_{n,h_n} = \left[ K\left(\frac{j-l}{h_n}\right) \hat{\gamma}_{j-l} \right]_{1 \leq j,l \leq n}$$

with:

$$\hat{\gamma}_k = \frac{1}{n} \sum_{j=1}^{n-|k|} \epsilon_j \epsilon_{j+|k|}, \qquad 0 \leq |k| \leq (n-1).$$

The function $K$ is a kernel such that:

- K is nonnegative, symmetric, and $K(0) = 1$
- K has compact support
- the fourier transform of $K$ is integrable.

The sequence of positive reals $h_n$ is such that $h_n \xrightarrow[n\to\infty]{} \infty$ and $\frac{h_n}{n} \xrightarrow[n\to\infty]{} 0$.

In our context, $(\epsilon_i)_{i\in\{1,\ldots,n\}}$ is not observed. Only the residuals are available:

$$\hat{\epsilon}_i = Y_i - (x_i)^t \hat{\beta} = Y_i - \sum_{j=1}^{p} x_{i,j}\hat{\beta}_j,$$

because only the data $Y$ and the design $X$ are observed. Consequently, we consider the following estimator of $\Gamma_n$:

$$\widehat{\Gamma}^*_{n,h_n} = \left[K\left(\frac{j-l}{h_n}\right)\hat{\gamma}^*_{j-l}\right]_{1\leq j,l\leq n}$$

where:

$$\hat{\gamma}^*_k = \frac{1}{n}\sum_{j=1}^{n-|k|}\hat{\epsilon}_j\hat{\epsilon}_{j+|k|}, \qquad 0\leq |k| \leq (n-1).$$

To estimate the asymptotic covariance matrix $F^{-1}GF^{-1}$, we use the estimator:

$$C_n = D(n)(X^tX)^{-1}X^t\widehat{\Gamma}^*_{n,h_n}X(X^tX)^{-1}D(n).$$

Let us denote by $C$ the matrix $F^{-1}GF^{-1}$ and the coefficients of the matrices $C_n$ and $C$ are respectively denoted by $c_{n,(j,l)}$ and $c_{j,l}$ , for all $j, l$ in $1, ..., p$.

# Consistence

**Theorem (C. (2018), submitted)**

*Let $h_n$ be a sequence of positive reals such that $h_n \to \infty$ as $n$ tends to infinity, and:*

$$h_n \mathbb{E}\left( |\epsilon_0|^2 \left( 1 \wedge \frac{h_n}{n} |\epsilon_0|^2 \right) \right) \xrightarrow[n\to\infty]{} 0.$$

*Then, under the assumptions of Hannan's Theorem, the estimator $C_n$ is consistent, that is for all $j, l$ in $1, ..., p$:*

$$\mathbb{E}\left( \left| c_{n,(j,l)} - c_{j,l} \right| \middle| X \right) \xrightarrow[n\to\infty]{} 0$$

**Corollary**

*Under the same conditions, the estimator $C_n$ converges in probability to $C$ as $n$ tends to infinity.*

## Sketch of the proof

Let $V(X)$ be the matrix $\mathbb{E}\left(D(n)(\hat{\beta}-\beta)(\hat{\beta}-\beta)^t D(n)^t \middle| X\right)$, and let $v_{j,l}$ be its coefficients. By the triangle inequality, $\forall j,l \in \{1,\dots,p\}$:

$$\left|c_{n,(j,l)} - c_{j,l}\right| \leq \left|v_{j,l} - c_{j,l}\right| + \left|c_{n,(j,l)} - v_{j,l}\right|.$$

Thanks to Hannan's Theorem:

$$\lim_{n\to\infty} \mathbb{E}\left(\left|v_{j,l} - c_{j,l}\right| \middle| X\right) = 0, \quad a.s.$$

Then it remains to prove that:

$$\lim_{n\to\infty} \mathbb{E}\left(\left|c_{n,(j,l)} - v_{j,l}\right| \middle| X\right) = 0, \quad a.s.$$

Thanks to the convergence of $D_n(X^tX)^{-1}D_n$ to $R(0)^{-1}$, it is sufficient to consider the matrices:

$$V' = D_n^{-1}X^t\Gamma_n X D_n^{-1}, \qquad C_n' = D_n^{-1}X^t\widehat{\Gamma}^*_{n,h_n} X D_n^{-1}.$$

We know that $\Gamma_n = \sum_{k=-n+1}^{n-1} \gamma(k) J_n^{(k)}$. Thus we have:

$$D(n)^{-1} X^t \Gamma_n X D(n)^{-1} = \sum_{k=-n+1}^{n-1} \gamma(k) B_{k,n}$$

$$D(n)^{-1} X^t \widehat{\Gamma}^*_{n,h_n} X D(n)^{-1} = \sum_{k=-n+1}^{n-1} K\left(\frac{k}{h_n}\right) \hat{\gamma}^*_k B_{k,n}$$

with $B_{k,n} = D(n)^{-1} X^t J_n^{(k)} X D(n)^{-1}$.

$$\left| c'_{n,(j,l)} - v'_{j,l} \right| = \left| \sum_{k=-n+1}^{n-1} \left( K\left(\frac{k}{h_n}\right) \hat{\gamma}^*_k - \gamma(k) \right) b^{k,n}_{j,l} \right|$$

where $b^{k,n}_{j,l}$ is the coefficient $(j,l)$ of the $B_{k,n}$ matrix.

Then:

$$\sum_{k=-n+1}^{n-1}\left(K\left(\frac{k}{h_n}\right)\hat{\gamma}_k^* - \gamma(k)\right)B_{k,n} = \int_{-\pi}^{\pi}\left(f_n^*(\lambda) - f(\lambda)\right)g_n(\lambda)(d\lambda)$$

with:

$$g_n(\lambda) = \frac{1}{2\pi}\sum_{k=-(n-1)}^{n-1} e^{ikx}B_{k,n}.$$

in such a way that the matrices $B_{k,n}$ are the Fourier coefficients of the function $g_n(\lambda)$:

$$B_{k,n} = \int_{-\pi}^{\pi} e^{ik\lambda}g_n(\lambda)d\lambda.$$

We have:

$$\mathbb{E}\left(\left|\int_{-\pi}^{\pi}\left(f_n^*(\lambda) - f(\lambda)\right)[g_n(\lambda)]_{j,l}d\lambda\right|\Big|X\right)$$

$$\leq \sup_{\lambda\in[-\pi,\pi]}\mathbb{E}\left(\left|f_n^*(\lambda) - f(\lambda)\right|\Big|X\right)\int_{-\pi}^{\pi}\left|[g_n(\lambda)]_{j,l}\right|d\lambda$$

# Proof: Spectral density estimate

And:

$$\sup_{\lambda\in[-\pi,\pi]} \mathbb{E}\left(|f_n^*(\lambda)-f(\lambda)|\,\Big|X\right)\int_{-\pi}^{\pi}|[g_n(\lambda)]_{j,l}|\,d\lambda$$
$$\leq \sup_{\lambda\in[-\pi,\pi]} \mathbb{E}\left(|f_n^*(\lambda)-f(\lambda)|\,\Big|X\right).$$

Then consider the following estimator:

$$f_n^*(\lambda)=\frac{1}{2\pi}\sum_{|k|\leq n-1}K\left(\frac{|k|}{h_n}\right)\hat{\gamma}_k^* e^{ik\lambda}, \qquad \lambda\in[-\pi,\pi],$$

where:

$$\hat{\gamma}_k^*=\frac{1}{n}\sum_{j=1}^{n-|k|}\hat{\epsilon}_j\hat{\epsilon}_{j+|k|}, \qquad 0\leq|k|\leq(n-1).$$

## Theorem (C.-Dede (2017) [1], submitted)

*Let $h_n$ be a sequence of positive integers such that $h_n \to \infty$ as $n$ tends to infinity, and:* $h_n \mathbb{E}\left(|\epsilon_0|^2 \left(1 \wedge \frac{h_n}{n} |\epsilon_0|^2\right)\right) \xrightarrow[n\to\infty]{} 0$. *Then, under the assumptions of Hannan's Theorem:*

$$\sup_{\lambda \in [-\pi,\pi]} \|f_n^*(\lambda) - f(\lambda)\|_{\mathbb{L}^1} \xrightarrow[n\to\infty]{} 0.$$

## Remark

*If* $\epsilon_0 \in \mathbb{L}^2$, *then there exists* $h_n \to \infty$ *such that* $h_n \mathbb{E}\left(|\epsilon_0|^2 \left(1 \wedge \frac{h_n}{n} |\epsilon_0|^2\right)\right) \xrightarrow[n\to\infty]{} 0$ *holds.*

This theorem is true for a fixed design $X$. But a quick look to the proof of this theorem suffices to see that:

$$\lim_{n\to\infty} \sup_{\lambda \in [-\pi,\pi]} \mathbb{E}\left(|f_n^*(\lambda) - f(\lambda)| \,\middle|\, X\right) = 0, \quad a.s.$$

# Corollary (Hannan's theorem + Consistence theorem)

**Corollary**

*Under the assumptions of Hannan's Theorem and the previous theorem (Consistence of $C_n$), we get:*

$$C_n^{-\frac{1}{2}} \left( D(n)(\hat{\beta} - \beta) \right) \xrightarrow[n \to \infty]{\mathcal{L}} \mathcal{N}(0, I_p),$$

*where $I_p$ is the $p \times p$ identity matrix.*

Consequently, we can obtain confidence regions and tests for $\beta$ in this dependent context.

# Sommaire

## "Fisher" test: Dependent case

$H_0 : \beta_{j_1} = \ldots = \beta_{j_{p_0}} = 0$, against $H_1 : \exists j_z \in \{j_1, \ldots, j_{p_0}\}$ such that $\beta_{j_z} \neq 0$. If the error process is strictly stationary, we have:

$$C_{n_{p_0}}^{-1/2} \begin{pmatrix} d_{j_1}(n)(\hat{\beta}_{j_1} - \beta_{j_1}) \\ \vdots \\ d_{j_{p_0}}(n)(\hat{\beta}_{j_{p_0}} - \beta_{j_{p_0}}) \end{pmatrix} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0_{p_0\times 1}, I_{p_0}),$$

Then under $H_0$-hypothesis:

$$\begin{pmatrix} Z_{1,n} \\ \vdots \\ Z_{p_0,n} \end{pmatrix} = C_{n_{p_0}}^{-1/2} \begin{pmatrix} d_{j_1}(n)\hat{\beta}_{j_1} \\ \vdots \\ d_{j_{p_0}}(n)\hat{\beta}_{j_{p_0}} \end{pmatrix} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0_{p_0\times 1}, I_{p_0}),$$

and we define the following test statistic:

$$\Xi = Z_{1,n}^2 + \cdots + Z_{p_0,n}^2.$$

Under the $H_0$-hypothesis, $\Xi \xrightarrow[n\to\infty]{\mathcal{L}} \chi^2_{p_0}$.

# One-parameter test

If we have $H_0 : \beta_j = 0$ against $H_1 : \beta_j \neq 0$, for $j$ in $\{1, \ldots, p\}$ ("Student" test), under the H0-hypothesis:

$$d_j(n)\hat{\beta}_j \xrightarrow[n\to\infty]{} \mathcal{N}(0, c_{j,j})$$

Then the test statistic is:

$$T_{j,n} = \frac{d_j(n)\hat{\beta}_j}{\sqrt{c_{n,(j,j)}}}$$

Under the $H_0$-hypothesis, $T_{j,n} \xrightarrow[n\to\infty]{} \mathcal{N}(0,1)$

# Choice of $h_n$

$$C_n = D(n)(X^tX)^{-1}X^t\widehat{\Gamma}^*_{n,h_n}X(X^tX)^{-1}D(n),$$

with:

$$\widehat{\Gamma}^*_{n,h_n} = \left[K\left(\frac{j-l}{h_n}\right)\hat{\gamma}^*_{j-l}\right]_{1\leq j,l\leq n}$$

For the kernel $K$, we shall use:

$$\begin{cases} K(x) = 1 & if\ |x| \leq 0.8 \\ K(x) = 5-5|x| & if\ 0.8 \leq |x| \leq 1 \\ K(x) = 0 & if\ |x| > 1. \end{cases}$$

This kernel verifies the conditions to apply the consistence theorem. It is close to the rectangular kernel (whose Fourier transform is not integrable). Hence, the parameter $h_n$ can be understood as the number of covariance terms that are necessary to obtain a good approximation of $\Gamma_n$. To choose its values, we shall use the graph of the empirical autocovariance of the residuals.

# Example

We first simulate $(Z_1, \ldots, Z_n)$ according to the $AR(1)$ equation $Z_{k+1} = \frac{1}{2}(Z_k + \eta_{k+1})$, where:

- $Z_1$ is uniformly distributed over $[0, 1]$
- $(\eta_i)_{i \geq 2}$ is a sequence of i.i.d. random variables with distribution $\mathcal{B}(1/2)$, independent of $Z_1$.

Let us define:

$$\epsilon_i = F^{-1}_{\mathcal{N}(0,\sigma^2)}(Z_i).$$

By construction, $\epsilon_i$ is $\mathcal{N}(0, \sigma^2)$-distributed (but the sequence $(\epsilon_i)_{i \geq 1}$ is not a Gaussian process). For the simulations, $\sigma^2$ is chosen equal to $25$.

First model simulated:

$$Y_i = \beta_0 + \beta_1(i^2 + X_i) + \epsilon_i, \quad \forall i \in \{1, \dots, n\}$$

with $(X_i)_{i \geq 1}$ a gaussian $AR(1)$ process (the variance is equal to $9$), independent of the process $(\epsilon_i)_{i \geq 1}$.
We test $H_0$: $\beta_1 = 0$, against $H_1$: $\beta_1 \neq 0$, for different choices of $n$ and $h_n$.

- $\beta_0 = 3$.
- Under $H_0$, the same Fischer test is carried out $2000$ times. Then we look at the frequency of rejection of the test (under $H_0$), that is to say the estimated level of the test (we want an estimated level close to $5\%$).

- Case $\beta_1 = 0$ and $h_n = 1$ (no correction):

| $n$ | 200 | 400 | 600 | 800 | 1000 |
|---|---|---|---|---|---|
| Estimated level | 0.203 | 0.195 | 0.183 | 0.205 | 0.202 |

Here, since $h_n = 1$, we do not estimate any of the covariance terms. The result is that the estimated levels are too large. The test will reject the null hypothesis too often.

The parameter $h_n$ may be chosen by analyzing the graph of the empirical autocovariances. For this example, the shape of the empirical autocovariance suggests to keep only $4$ terms. This leads to choose $h_n = 5$.

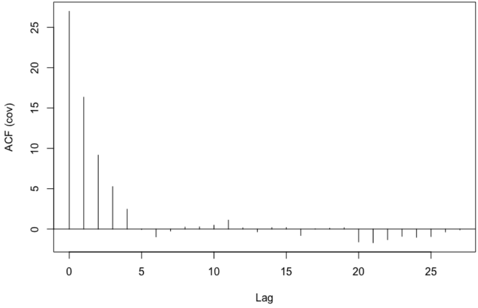


Figure : Empirical autocovariances of the residuals.

- Case $\beta_1 = 0$, $h_n = 5$:

| $n$ | 200 | 400 | 600 | 800 | 1000 |
|---|---|---|---|---|---|
| Estimated level | 0.0845 | 0.065 | 0.0595 | 0.054 | 0.053 |

As suggested by the graph of the empirical autocovariances, the choice $h_n = 5$ gives a better estimated level than $h_n = 1$. If $n = 2000$ and $h_n = 7$, the estimated level is around $0.05$.

- Case $\beta_1 = 0.00001$, $h_n = 5$:

In this example, $H_0$ is not satisfied. We perform the same tests as above ($N = 2000$) to estimate the power of the test.

| $n$ | 200 | 400 | 600 | 800 | 1000 |
|---|---|---|---|---|---|
| Estimated power | 0.1025 | 0.301 | 0.887 | 1 | 1 |

As one can see, the estimated power is always greater than $0.05$, as expected. Still as expected, the estimated power increases with the size of the samples. As soon as $n = 800$, the test always rejects the $H_0$-hypothesis.

# Second model

$$Y_i = \beta_0 + \beta_1(\log(i) + \sin(i) + X_i) + \beta_2 i + \epsilon_i, \quad \forall i \in \{1, ..., n\}$$

We test $H_0$: $\beta_1 = \beta_2 = 0$ against $H_1$: $\beta_1 \neq 0$ or $\beta_2 \neq 0$. The coefficient $\beta_0$ is equal to $3$, and we use the same simulation scheme as above.

- Case $\beta_1 = \beta_2 = 0$ and $h_n = 1$ (no correction):

| $n$ | 200 | 400 | 600 | 800 | 1000 |
|---|---|---|---|---|---|
| Estimated level | 0.348 | 0.334 | 0.324 | 0.3295 | 0.3285 |

As for the first simulation, if $h_n = 1$ the test will reject the null hypothesis too often.

As suggested by the graph of the estimated autocovariances, it suggests to keep only $5$ terms of covariances. This leads to choose $h_n = 6.25$.

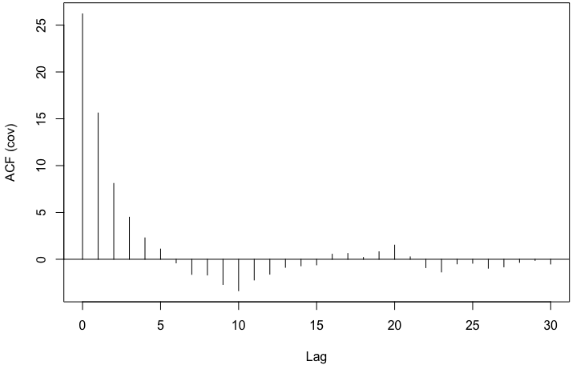


Figure : Empirical autocovariances of the residuals.

- Case $\beta_1 = \beta_2 = 0$, $h_n = 6.25$:

| $n$ | 200 | 400 | 600 | 800 | 1000 |
|---|---|---|---|---|---|
| Estimated level | 0.09 | 0.078 | 0.066 | 0.0625 | 0.0595 |

Here, we see that the choice $h_n = 6.25$ works well. For $n = 1000$, the estimated level is around $0.06$. If $n = 2000$ and $h_n = 6.25$, the estimated level is around $0.05$.

- Case $\beta_1 = 0.2$, $\beta_2 = 0$, $h_n = 6.25$:

Now, we study the estimated power of the test.

| $n$ | 200 | 400 | 600 | 800 | 1000 |
|---|---|---|---|---|---|
| Estimated power | 0.33 | 0.5 | 0.6515 | 0.776 | 0.884 |

As expected, the estimated power increases with the size of the samples, and it is around $0.9$ when $n = 1000$.

# Perspectives

- To develop a data driven criterion for the coefficient $h_n$
- Package R for the applications of these results
- To consider the case where $p$ (number of variables) is greater than $n$ (number of observations)

**Thank you for your attention !**

E. Caron and S. Dede.
Asymptotic distribution of least square estimators for linear models with dependent errors : regular designs.
eprint arXiv:1710.05963, Oct. 2017.

J. Dedecker, F. Merlevède, and D. Volnỳ.
On the weak invariance principle for non-adapted sequences under projective criteria.
*Journal of Theoretical Probability*, 20(4):971–1004, 2007.

J. Dedecker and C. Prieur.
New dependence coefficients. examples and applications to statistics.
*Probability Theory and Related Fields*, 132(2):203–236, 2005.

E. J. Hannan.
Central limit theorems for time series regression.
*Probability theory and related fields*, 26(2):157–170, 1973.

W. B. Wu.
Nonlinear system theory: Another look at dependence.

*Proceedings of the National Academy of Sciences of the United States of America*, 102(40):14150–14154, 2005.