

# Rapport mathématique sur le test de stationnarité de Phillips-Perron (PP)

DIALLO Souleymane  
HINIMDOU Morsia Guitdam  
NANKOULI Marc Thierry

April 11, 2025

Exposé  
Probabilité et Tests Statistiques pour la Data Science

ENSAM

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Contexte théorique</b>	<b>2</b>
2.1	Définition de la stationnarité . . . . .	2
2.2	Racine unitaire et intégration . . . . .	2
2.3	Problèmes liés à la non-stationnarité . . . . .	3
<b>3</b>	<b>Organigramme du test de Phillips-Perron</b>	<b>3</b>
<b>4</b>	<b>La procédure du test de Phillips-Perron</b>	<b>3</b>
4.1	Modèle et hypothèses . . . . .	3
4.2	Estimation initiale . . . . .	4
4.3	Correction pour l'autocorrélation et l'hétéroscédasticité . . . . .	5
4.4	Statistiques ajustées . . . . .	5
4.5	Distribution asymptotique . . . . .	6
<b>5</b>	<b>Avantages et limites</b>	<b>6</b>
5.1	Avantages . . . . .	6
5.2	Limites . . . . .	6
5.3	Comparaison avec d'autres tests . . . . .	6
<b>6</b>	<b>Applications pratiques et interprétation</b>	<b>7</b>
6.1	Exemple numérique détaillé . . . . .	7
6.2	Applications réelles . . . . .	7
<b>7</b>	<b>Conclusion</b>	<b>7</b>
<b>8</b>	<b>Références</b>	<b>7</b>

# 1 Introduction

Le test de Phillips-Perron (PP), développé par Phillips (1987) et Phillips et Perron (1988), est une méthode statistique fondamentale pour tester la stationnarité d'une série temporelle, un concept essentiel en économétrie et en statistique mathématique. Une série temporelle est stationnaire si ses propriétés statistiques (moyenne, variance, autocovariance) restent constantes dans le temps. Le test PP vise à détecter la présence d'une racine unitaire dans une série, c'est-à-dire à vérifier si la série est intégrée d'ordre 1 (I(1), non stationnaire) ou d'ordre 0 (I(0), stationnaire). Ce test se distingue par sa capacité à corriger les biais liés à l'autocorrélation et à l'hétéroscédasticité des résidus, en adoptant une approche non paramétrique.

## 2 Contexte théorique

### 2.1 Définition de la stationnarité

Une série temporelle  $\{y_t\}_{t=1}^T$  est dite stationnaire au sens faible (ou stationnaire d'ordre 2) si elle satisfait les conditions suivantes :

- Moyenne constante :

$$E[y_t] = \mu, \quad \forall t, \quad (1)$$

- Variance constante :

$$Var(y_t) = \sigma^2 < \infty, \quad \forall t, \quad (2)$$

- Autocovariance dépendant de l'écart temporel :

$$Cov(y_t, y_{t-k}) = \gamma_k, \quad \forall t, k. \quad (3)$$

Une série est stationnaire stricte si sa distribution jointe  $F(y_{t_1}, y_{t_2}, \dots, y_{t_n})$  est invariante par translation dans le temps, une condition plus forte. Le test PP se concentre sur la stationnarité faible.

### 2.2 Racine unitaire et intégration

Une série temporelle  $y_t$  suit un processus autorégressif d'ordre 1 (AR(1)) si :

$$y_t = \rho y_{t-1} + \epsilon_t, \quad (4)$$

où  $\epsilon_t \sim \text{i.i.d.}(0, \sigma^2)$  est un bruit blanc. Si  $\rho = 1$ , la série a une racine unitaire et est non stationnaire. Réécrivons l'équation sous la forme :

$$\Delta y_t = (\rho - 1)y_{t-1} + \epsilon_t, \quad (5)$$

où  $\Delta y_t = y_t - y_{t-1}$ . Si  $\rho = 1$ , alors :

$$\Delta y_t = \epsilon_t, \quad (6)$$

ce qui correspond à une marche aléatoire, une série intégrée d'ordre 1 (I(1)). Si  $|\rho| < 1$ , la série est stationnaire (I(0)).

Le test PP vise à tester :

- $H_0 : \rho = 1$  (racine unitaire, série non stationnaire),
- $H_1 : |\rho| < 1$  (série stationnaire).

La variance d'une marche aléatoire ( $\rho = 1$ ) est :

$$Var(y_t) = Var\left(\sum_{s=1}^t \epsilon_s\right) = t\sigma^2, \quad (7)$$

ce qui montre que la variance croît avec le temps, un signe de non-stationnarité.

## 2.3 Problèmes liés à la non-stationnarité

Une série non stationnaire pose des problèmes mathématiques :

- **Régressions fallacieuses** : Si  $y_t$  et  $x_t$  sont  $I(1)$  et non liées, une régression  $y_t = \beta x_t + \epsilon_t$  peut produire un  $\hat{\beta}$  significatif (test  $t$  élevé) même sans relation réelle.
- **Distributions non standards** : Les estimateurs MCO dans une régression impliquant des séries  $I(1)$  convergent vers des distributions non normales, souvent des fonctionnelles de processus de Wiener.

## 3 Organigramme du test de Phillips-Perron

Voici un organigramme détaillant les étapes mathématiques du test PP :

## 4 La procédure du test de Phillips-Perron

### 4.1 Modèle et hypothèses

Le test PP repose sur la régression suivante :

$$y_t = \alpha + \beta t + \rho y_{t-1} + \epsilon_t, \quad (8)$$

où  $\alpha$ ,  $\beta$ , et  $\rho$  sont les paramètres à estimer, et  $\epsilon_t$  est le terme d'erreur. Les modèles possibles sont :

- Sans constante ni tendance ( $\alpha = 0, \beta = 0$ ) :

$$y_t = \rho y_{t-1} + \epsilon_t, \quad (9)$$

- Avec constante ( $\beta = 0$ ) :

$$y_t = \alpha + \rho y_{t-1} + \epsilon_t, \quad (10)$$

- Avec constante et tendance :

$$y_t = \alpha + \beta t + \rho y_{t-1} + \epsilon_t. \quad (11)$$

Les hypothèses du test sont :

- $H_0 : \rho = 1$  (racine unitaire, série non stationnaire),
- $H_1 : |\rho| < 1$  (série stationnaire).

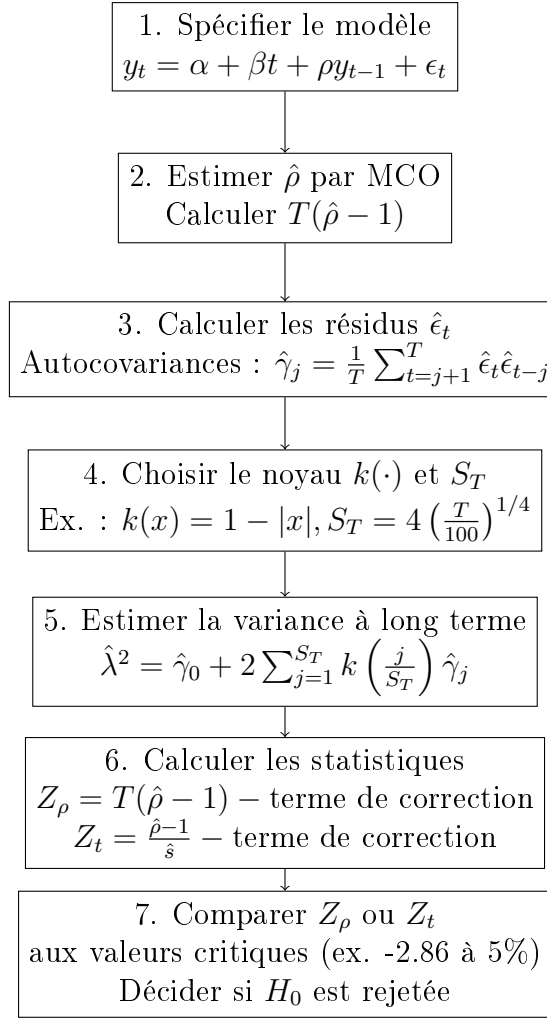


Figure 1: Organigramme mathématique des étapes du test de Phillips-Perron

## 4.2 Estimation initiale

On estime  $\rho$  par les moindres carrés ordinaires (MCO), obtenant  $\hat{\rho}$ . La statistique de base pour tester  $H_0$  est :

$$T(\hat{\rho} - 1), \quad (12)$$

où  $T$  est la taille de l'échantillon. Le t-statistique associé est :

$$t_{\hat{\rho}=1} = \frac{\hat{\rho} - 1}{\hat{s}}, \quad (13)$$

où  $\hat{s}$  est l'écart-type estimé de  $\hat{\rho}$ , donné par :

$$\hat{s} = \sqrt{\frac{\hat{\sigma}^2}{\sum_{t=2}^T (y_{t-1} - \bar{y}_{t-1})^2}}, \quad (14)$$

avec  $\hat{\sigma}^2 = \frac{1}{T-k} \sum_{t=1}^T \hat{\epsilon}_t^2$ , et  $k$  le nombre de paramètres estimés (par exemple,  $k = 2$  pour un modèle avec constante et tendance).

### 4.3 Correction pour l'autocorrélation et l'hétéroscédasticité

Le test PP ajuste les statistiques pour tenir compte des résidus non i.i.d. On calcule d'abord les résidus :

$$\hat{\epsilon}_t = y_t - \hat{\alpha} - \hat{\beta}t - \hat{\rho}y_{t-1}, \quad (15)$$

puis les autocovariances :

$$\hat{\gamma}_j = \frac{1}{T} \sum_{t=j+1}^T \hat{\epsilon}_t \hat{\epsilon}_{t-j}, \quad j = 0, 1, \dots, S_T. \quad (16)$$

La variance à long terme  $\hat{\lambda}^2$  est estimée avec une fonction de noyau  $k(\cdot)$  et un paramètre de troncature  $S_T$  :

$$\hat{\lambda}^2 = \hat{\gamma}_0 + 2 \sum_{j=1}^{S_T} k\left(\frac{j}{S_T}\right) \hat{\gamma}_j. \quad (17)$$

Un choix courant pour le noyau est le noyau de Bartlett :

$$k(x) = \begin{cases} 1 - |x|, & \text{si } |x| \leq 1, \\ 0, & \text{sinon,} \end{cases} \quad (18)$$

et pour  $S_T$ , une règle empirique est :

$$S_T = 4 \left( \frac{T}{100} \right)^{1/4}. \quad (19)$$

Une alternative est le noyau de Parzen :

$$k(x) = \begin{cases} 1 - 6x^2 + 6|x|^3, & \text{si } 0 \leq |x| \leq \frac{1}{2}, \\ 2(1 - |x|)^3, & \text{si } \frac{1}{2} \leq |x| \leq 1, \\ 0, & \text{sinon.} \end{cases} \quad (20)$$

La variance à court terme est :

$$\hat{\gamma}_0 = \frac{1}{T} \sum_{t=1}^T \hat{\epsilon}_t^2. \quad (21)$$

On définit également la variance empirique de la série :

$$S^2 = \frac{1}{T} \sum_{t=1}^T (y_t - \bar{y})^2, \quad (22)$$

où  $\bar{y} = \frac{1}{T} \sum_{t=1}^T y_t$ .

### 4.4 Statistiques ajustées

Les statistiques ajustées  $Z_\rho$  et  $Z_t$  sont données par :

$$Z_\rho = T(\hat{\rho} - 1) - \frac{1}{2} \left( \frac{\hat{\lambda}^2 - \hat{\gamma}_0}{\hat{\lambda}^2} \right) \left( \frac{T^2 \hat{\sigma}^2}{S^2} \right), \quad (23)$$

$$Z_t = \left( \frac{\hat{\rho} - 1}{\hat{s}} \right) - \frac{1}{2} \left( \frac{\hat{\lambda}^2 - \hat{\gamma}_0}{\hat{\lambda}^2} \right) \left( \frac{T \hat{\sigma}}{\hat{s} S} \right). \quad (24)$$

Ces termes de correction ajustent les biais dus à l'autocorrélation et à l'hétéroscédasticité.

## 4.5 Distribution asymptotique

Sous  $H_0$ , les statistiques  $Z_\rho$  et  $Z_t$  convergent vers des distributions non standards, similaires à celles du test de Dickey-Fuller. Par exemple,  $Z_t$  suit une distribution de type Dickey-Fuller, avec des valeurs critiques tabulées. Pour un modèle avec constante, les valeurs critiques à 5% sont environ -2.86 pour  $Z_t$ .

## 5 Avantages et limites

### 5.1 Avantages

- **Robustesse** : La correction non paramétrique via  $\hat{\lambda}^2$  permet de gérer l'autocorrélation et l'hétéroscédasticité sans modélisation explicite.
- **Flexibilité** : Pas besoin de spécifier un nombre de retards, contrairement au test ADF.

### 5.2 Limites

- **Choix de  $S_T$**  : La statistique dépend de  $S_T$ . Une règle comme  $S_T = 4 \left(\frac{T}{100}\right)^{1/4}$  peut être sous-optimale.
- **Petites tailles d'échantillon** : Pour  $T < 50$ , l'estimation de  $\hat{\lambda}^2$  est imprécise.
- **Ruptures structurelles** : Le test peut être biaisé en présence de ruptures.

### 5.3 Comparaison avec d'autres tests

- **Test de Dickey-Fuller (DF)** : Le test DF utilise la statistique :

$$\text{DF statistic} = T(\hat{\rho} - 1), \quad (25)$$

mais suppose des résidus i.i.d., ce qui est restrictif.

- **Test ADF** : L'ADF ajoute des termes retardés :

$$\Delta y_t = (\rho - 1)y_{t-1} + \sum_{j=1}^p \phi_j \Delta y_{t-j} + \epsilon_t, \quad (26)$$

mais nécessite un choix de  $p$ .

- **Test KPSS** : Teste  $H_0$  : stationnarité, avec la statistique :

$$\text{KPSS} = \frac{1}{T^2} \sum_{t=1}^T S_t^2 / \hat{\lambda}^2, \quad (27)$$

où  $S_t = \sum_{i=1}^t \hat{\epsilon}_i$ .

## 6 Applications pratiques et interprétation

### 6.1 Exemple numérique détaillé

Prenons une série simulée  $y_t = 0.95y_{t-1} + \epsilon_t$ , avec  $\epsilon_t \sim N(0, 1)$ ,  $T = 100$ .

1. **Régression** : Modèle  $y_t = \alpha + \rho y_{t-1} + \epsilon_t$ . Résultats MCO :  $\hat{\alpha} = 0.1$ ,  $\hat{\rho} = 0.948$ , donc :

$$T(\hat{\rho} - 1) = 100 \times (0.948 - 1) = -5.2. \quad (28)$$

Le t-statistique est :

$$t_{\hat{\rho}=1} = \frac{0.948 - 1}{\hat{s}}, \quad (29)$$

avec  $\hat{s} = 0.021$ , donc  $t_{\hat{\rho}=1} = -2.48$ .

2. **Résidus et autocovariances** : Résidus  $\hat{\epsilon}_t$ , autocovariances :  $\hat{\gamma}_0 = 0.98$ ,  $\hat{\gamma}_1 = 0.3$ ,  $\hat{\gamma}_2 = 0.1$ ,  $\hat{\gamma}_j \approx 0$  pour  $j \geq 3$ .

3. **Variance à long terme** : Avec  $S_T = 4$ , noyau de Bartlett :

$$\hat{\lambda}^2 = 0.98 + 2 \left[ \left(1 - \frac{1}{4}\right) \times 0.3 + \left(1 - \frac{2}{4}\right) \times 0.1 \right] = 1.53. \quad (30)$$

4. **Statistiques** : Supposons  $S^2 = 10$ ,  $\hat{\sigma}^2 = 1.0$ . Alors :

$$Z_\rho = -5.2 - \frac{1}{2} \left( \frac{1.53 - 0.98}{1.53} \right) \left( \frac{100^2 \times 1.0}{10} \right). \quad (31)$$

Après calcul, disons  $Z_\rho = -6.5$ . Pour  $Z_t$  :

$$Z_t \approx -2.48 - \text{terme de correction} \approx -3.0. \quad (32)$$

5. **Décision** : Valeur critique à 5% : -2.86. Puisque  $Z_t = -3.0 < -2.86$ , on rejette  $H_0$ .

### 6.2 Applications réelles

- **PIB nominal** : Vérifier si  $\text{PIB}_t$  suit une marche aléatoire.
- **Cours boursiers** : Analyser S&P 500<sub>t</sub> pour des modèles de prévision.

## 7 Conclusion

Le test PP est un outil mathématique rigoureux pour tester la stationnarité, avec des formules bien définies pour corriger les biais. Ce rapport a détaillé toutes les équations, de la régression initiale aux statistiques ajustées, et a illustré leur application. Il reste limité par le choix de  $S_T$  et les petites tailles d'échantillon.

## 8 Références

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