Le Test de Breusch-Pagan pour l'Hétéroscédasticité

GSF-6053

Hiver 2025

Table des matières

1 Introduction

Le test de **Breusch-Pagan** est une procédure statistique utilisée pour détecter la présence d'hétéroscédasticité dans un modèle de régression linéaire. L'hétéroscédasticité se produit lorsque la variance des erreurs n'est pas constante à travers les observations, ce qui viole l'une des hypothèses fondamentales du modèle de régression linéaire classique.

2 Hypothèses du Test de Breusch-Pagan

Le test de Breusch-Pagan évalue les hypothèses suivantes :

2.1 Hypothèse Nulle (H_0)

— Il n'y a pas d'hétéroscédasticité dans le modèle. La variance des erreurs est constante, c'est-à-dire $\mathrm{Var}(u_i) = \sigma^2$ pour tout i.

2.2 Hypothèse Alternative (H_1)

— Il existe une relation systématique entre la variance des erreurs et les variables explicatives. La variance des erreurs varie en fonction des variables explicatives, c'est-à-dire $\mathrm{Var}(u_i) = \sigma^2 h(X_i)$ où $h(X_i)$ est une fonction des variables explicatives.

3 Procédure du Test

Le test de Breusch-Pagan se déroule en plusieurs étapes :

3.1 Estimation du Modèle de Régression Principal

On estime le modèle de régression linéaire :

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i$$

en utilisant la méthode des moindres carrés ordinaires (MCO).

3.2 Calcul des Résidus

On calcule les résidus standardisés :

$$\hat{u}_i = Y_i - \hat{Y}_i$$

où $\hat{Y}_i$ est la valeur prédite par le modèle.

3.3 Régression Auxiliaire

On effectue une régression auxiliaire où la variable dépendante est le carré des résidus :

$$\hat{u}_i^2 = \gamma_0 + \gamma_1 Z_{1i} + \gamma_2 Z_{2i} + \cdots + \gamma_m Z_{mi} + \epsilon_i$$

Les variables $Z_{1i}, Z_{2i}, \ldots, Z_{mi}$ sont choisies parmi les variables explicatives originales ou d'autres transformations de celles-ci.

3.4 Calcul de la Statistique de Test

La statistique de test de Breusch-Pagan est donnée par :

$$\text{BP} = \frac{SSR}{2\sigma^4}$$

où SSR est la somme des carrés des résidus de la régression auxiliaire.

Cependant, une approche plus courante consiste à utiliser la statistique basée sur le coefficient de détermination (R^2) de la régression auxiliaire :

$$\text{BP} = \frac{T \cdot R^2}{2}$$

où T est le nombre total d'observations.

3.5 Distribution Sous l'Hypothèse Nulle

Sous l'hypothèse nulle d'homoscédasticité, la statistique de test suit une distribution du chi carré avec m degrés de liberté, où m est le nombre de variables explicatives dans la régression auxiliaire.

4 Interprétation des Résultats

4.1 Décision Statistique

— Si BP $> \chi^2_{m,1-\alpha}$, on rejette l'hypothèse nulle et conclut à la présence d'hétéroscédasticité.

— Sinon, on ne rejette pas l'hypothèse nulle.

4.2 Conséquences de l'Hétéroscédasticité

L'hétéroscédasticité peut entraîner :

— Des estimateurs MCO toujours sans biais, mais inefficaces.

— Des erreurs standard incorrectes, ce qui affecte les tests d'hypothèses et les intervalles de confiance.

5 Exemple Illustratif

Considérons un modèle de régression où l'on souhaite estimer la relation entre le revenu (Y) et le niveau d'éducation (X_1) et l'expérience professionnelle (X_2).

5.1 Étape 1 : Estimation du Modèle Principal

On estime le modèle :

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$$

en utilisant les MCO.

5.2 Étape 2 : Calcul des Résidus

On calcule les résidus $\hat{u}_i = Y_i - \hat{Y}_i$.

5.3 Étape 3 : Régression Auxiliaire

On effectue la régression :

$$\hat{u}_i^2 = \gamma_0 + \gamma_1 X_{1i} + \gamma_2 X_{2i} + \epsilon_i$$

5.4 Étape 4 : Calcul de la Statistique BP

Supposons que $R^2 = 0.05$ et que $T = 100$ observations avec $m = 2$ variables explicatives dans la régression auxiliaire. Alors :

$$\text{BP} = \frac{100 \times 0.05}{2} = 2.5$$

5.5 Étape 5 : Décision

Avec $m = 2$ et un niveau de signification $\alpha = 0.05$, la valeur critique est $\chi^2_{2,0.95} \approx 5.99$. Puisque $2.5 < 5.99$, on ne rejette pas l'hypothèse nulle d'homoscédasticité.

6 Conclusion

Le test de Breusch-Pagan est un outil essentiel pour détecter l'hétéroscédasticité dans les modèles de régression linéaire. En identifiant la présence d'hétéroscédasticité, les chercheurs peuvent ajuster leurs modèles en conséquence, par exemple en utilisant des estimateurs robustes ou en transformant les variables, afin d'obtenir des inférences statistiques fiables.

7 Références

— Breusch, T. S., & Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, 47(5), 1287–1294.

— Gujarati, D. N. (2003). *Basic Econometrics*. McGraw-Hill.

— Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.