Économétrie — TD 4

Les tests d'hypothèses économétriques

Pierre Beaucoral

Introduction

Ce document reprend les **slides du TD 4** sous forme **cours**, avec des explications supplémentaires pour guider la lecture et la pratique sous **EViews**. L'objectif est de comprendre **pourquoi** on applique ces tests, comment **les mettre en œuvre**, et **comment interpréter** les résultats.

> Note
>
> **À retenir.** Les tests présentés concernent les hypothèses sur les **erreurs** du modèle MCO. Ils ne modifient pas la structure du modèle mais la **validation des inférences** que l'on effectue (significativité, intervalles de confiance).

Rappel : hypothèses MCO et propriétés

Les hypothèses des estimations MCO

- L'estimateur des **moindres carrés ordinaires (MCO)** est le **BLUE** (*Best Linear Unbiased Estimator*) **si** certaines hypothèses sont respectées.
- Ces hypothèses portent sur le terme d'erreur ε :
  1. **Normalité** : $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$
  2. **Espérance nulle** : $\mathbb{E}[\varepsilon_i] = 0$
  3. **Homoscédasticité** : $\mathrm{Var}(\varepsilon_i) = \sigma^2$ (constante)
  4. **Indépendance sérielle** : $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0 \ \forall i \neq j$
  5. **Orthogonalité** : $\mathrm{Cov}(x_i, \varepsilon_i) = 0$ (exogénéité)

Caution

Sensibilité. L'estimateur MCO est sensible aux **observations extrêmes** : soyez vigilants aux **valeurs aberrantes** et aux **leviers**.

Propriétés des estimateurs (rappels sémantiques)

- **Petit échantillon**
  - *Sans biais* : $\mathbb{E}[\widehat{\beta}] = \beta$
  - *Variance minimale* : $\mathrm{Var}(\hat{\beta}) \leq \mathrm{Var}(\tilde{\beta})$
  - *Efficace* : sans biais **et** à variance minimale
- **Grand échantillon**
  - *Convergent* : $\mathrm{Var}(\hat{\beta}) \to 0$ lorsque $N \to \infty$

Hypothèses, problèmes et corrections

Propriété / Hypothèse	Problème si non respectée	Test(s) associé(s)	Méthode(s) de correction
Absence de biais (orthogonalité)	Biais ; non-convergence	—	**Instrumentation** (variables instrumentales)
Efficience (sphéricité des erreurs)	Estimation non efficace (mais pas de biais)	**Homoscédasticité** : Breusch–Pagan, White ; **Indépendance sérielle**: Durbin–Watson, Breusch–Godfrey	Erreurs standards **White** (robustes) ; **HAC** (Newey–West)

Important

Focus du TD. Nous traitons surtout les hypothèses d'**efficience** (3 et 4) : homoscédasticité et indépendance sérielle.

Les tests d'hypothèses : principe

- **Objectif** : vérifier que le **modèle est économétriquement correct** au sens des hypothèses MCO (afin d'interpréter correctement t, F, IC…).
- **Démarche** : on applique des **tests sur les résidus** et on adapte les **erreurs standards** (ou le modèle) si nécessaire.

Hypothèse	Test(s) associé(s)	Traité ?
H1 : Normalité	Bera–Jarque	
H3 : Homoscédasticité	Goldfeld–Quandt	
	Breusch–Pagan	
	White	
H4 : Indépendance sérielle	Durbin–Watson	
	Breusch–Godfrey	

Normalité des erreurs : test de Bera–Jarque

Intuition et indicateurs

- La normalité facilite certains tests de **sphéricité**.
- **Bera–Jarque (BJ)** combine :
  - **Skewness ( )** : asymétrie (= 0 sous normalité)
  - **Kurtosis ( )** : aplatissement (= 3 sous normalité)

Statistique, hypothèses, décision

- **Statistique** : $BJ = N\left[\frac{\eta^2}{6} + \frac{(\upsilon-3)^2}{24}\right]$
  Sous H_0, BJ suit approximativement une χ^2.
- **Hypothèses** :
  - H_0 : normalité $BJ \approx 0$
  - H_1 : non-normalité $BJ \neq 0$
- **Règle** (5 %) : si $BJ > \chi^2_{2;0,95} \approx 6$ **rejeter** H_0.

Mise en œuvre dans EViews

1. Ouvrir la **fenêtre de l'équation**.
2. **View** → **Residual Diagnostic** → **Histogram – Normality Test**.
3. Lire **BJ**, *p-value*, skewness et kurtosis ; conclure.

> Tip
>
> **Lecture** : une *p-value* faible (5 %) → rejet de la normalité. On peut poursuivre les tests d'hétéroscédasticité / autocorrélation même si la normalité est discutée.

Homoscédasticité (variance constante)

- L'homoscédasticité : $\text{Var}(\varepsilon_i) = \sigma^2$ **constante**.
- L'hétéroscédasticité survient souvent quand la **dispersion des erreurs** croît avec le **niveau d'une variable** explicative (ou de la variable expliquée).

> ! Important
>
> Si l'hypothèse est violée, les MCO restent **sans biais**, mais ne sont plus **efficients** → erreurs standards faussées (t, F, IC non fiables).

Test de Breusch–Pagan (BP)

- **Idée** : la **variance** des erreurs dépend des **variables explicatives**.
- **Modèle** estimé : $Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \varepsilon_i$ et $\hat{\varepsilon}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i - \hat{\beta}_2 Z_i$.
- **Hypothèses** :
  - H_0 : $\text{Var}(\hat{\varepsilon}_i) = \sigma^2$ (indépendante de X, Z)
  - H_1 : $\text{Var}(\hat{\varepsilon}_i) = \theta_0 + \theta_1 X_i + \theta_2 Z_i + \omega_i$
- **Procédure** :
  1. Estimer le modèle MCO ; calculer $\hat{\varepsilon}_i^2$.
  2. Régression de test : $\hat{\varepsilon}_i^2 = \theta_0 + \theta_1 X_i + \theta_2 Z_i + \omega_i$.
  3. $BP = N \times R^2$ de l'équation de test → $\chi^2(K-1)$;
     si $BP \geq \chi_t^2 h$, **rejeter** H_0.

ℹ Note

Intuition : sous homoscédasticité, **R^2 → 0** dans l'équation auxiliaire : X et Z n'expliquent pas la **variance** des résidus.

Test de White

- Même logique que BP mais **plus flexible** (termes au carré et produits croisés).
- **Hypothèses** :
  - $H_0 : \text{Var}(\hat{\varepsilon}_i) = \sigma^2$
  - $H_1 : \text{Var}(\hat{\varepsilon}_i) = \theta_0 + \theta_1 X_i + \theta_2 Z_i + \theta_3 X_i^2 + \theta_4 X_i Z_i + \theta_5 Z_i^2 + \omega_i$
- **Statistiques** :
  - $W = N \times R^2 \to \chi^2(K-1)$ (K = nb de paramètres de l'équation auxiliaire)
  - Variante petits échantillons (F-test) : $W = \left(\frac{SCR_r - SCR_{nr}}{SCR_r}\right) \frac{N-k}{k-1} \to F(k-1, N-k)$

Mise en œuvre dans EViews

- **View → Residual Diagnostic → Heteroskedasticity Tests**
  - **Breusch–Pagan–Godfrey** (BP)
  - **White**

Indépendance sérielle (autocorrélation)

- Absence d'autocorrélation : **Cov(,) = 0** pour **t s**.
- Problème classique en **séries temporelles** ; fausse les erreurs standards.

Test de Durbin–Watson (DW)

- Conditions : constante incluse ; $n > 15$; pas de variable dépendante retardée ; pas de données manquantes ; processus AR(1) uniquement.
- **Statistique** : $DW = \frac{\sum_{t=2}^{T} (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{T} \hat{\varepsilon}_t^2} \approx 2(1 - \hat{\rho})$
- **Lecture** via bornes **D__L** et **D__U** (zones : rejet H , incertitude, acceptation).

Caution

Limites DW : **zone d'incertitude**, conditions restrictives → préférer souvent **Breusch–Godfrey** quand les conditions DW ne sont pas réunies.

Test de Breusch–Godfrey (BG)

- Plus **flexible** (AR(p) avec $p \geq 1$).
- Équation de test (ex. AR(2)) : $\hat{\varepsilon}_t = \rho 1 \hat{\varepsilon} t - 1 + \rho 2 \hat{\varepsilon} t - 2 + \theta_1 X_t + \theta_2 Z_t + \omega_t$
- **Statistique** : $BG = T \times R^2 \to \chi^2(p)$; rejeter H_0 si $BG \geq \chi^2_t h.$
- **EViews** : **View** → **Residual Diagnostic** → **Serial correlation LM test** (choisir les lags).

Corriger pour des inférences valides

- Si **hétéroscédasticité** : erreurs standards **White** (robustes).
- Si **autocorrélation** (et/ou hétéroscédasticité) : erreurs standards **HAC** (Newey–West).
- Si l'autocorrélation est structurelle : envisager une **modélisation explicite** (par ex. **Cochrane–Orcutt**, AR(1) sur l'erreur) plutôt que de corriger seulement les erreurs standards.

Tip

Effet des corrections. Les **coefficients** MCO restent en général inchangés ; ce sont les **écarts-types** (et donc t, p-values, IC) qui sont ajustés.

Questions – Réponses (TD4)

Importez le fichier de travail sur les compagnies aériennes.

Estimez l'équation suivante par les MCO :

$log(Pass_i) = \beta_0 + \beta_1 Fatal_Passagers_i + \beta_2 NonFatal_Passagers_i + \beta_3 Low_cost_i + \beta_4 Public_i + \beta_5 Inter_i + \beta_6 Age_i + \beta_7 Trafic_nat_i + \beta_8 Trafic_dest_i + \varepsilon_i$

Homoscédasticité

· Qu'est-ce que l'homoscédasticité et quel problème induit son non-respect pour les MCO ?

Afficher la réponse

Homoscédasticité = la variance de l'erreur est **constante** pour toutes les valeurs des régressseurs :

$\mathrm{Var}(u_i \mid X) = \sigma^2$ pour tout i.

Si cette hypothèse est violée (hétéroscédasticité) :

- Les estimateurs MCO $\hat{\beta}$ restent **sans biais** et **consistants** si $E[u \mid X] = 0$ tient, mais ils **ne sont plus efficaces** (plus BLUE) : il existe de meilleurs estimateurs (GLS/WLS).
- Les **écarts-types MCO "classiques"** sont **faussés** **tests t/F et IC** peuvent être trompeurs (trop optimistes ou trop prudents).
- Conséquence pratique majeure : **mauvaise inférence**.

Que faire ?

- Utiliser des **erreurs-types robustes à l'hétéroscédasticité** (HC0–HC3/"White").

Homoscédasticité

· A l'aide des tests de Goldfeld et Quandt, de Breusch-Pagan-Koenker et de White, que peut-on conclure quant à l'homoscédasticité du terme d'erreurs ?

Afficher la réponse

Correction(s)

· En fonction des résultats des divers tests, proposez une correction le cas échéant.

Afficher la réponse

Correction(s)

· Vos conclusions quant à l'effet des accidents mortels et non mortels sont-elles modifiées ?

Afficher la réponse