Économétrie

Hétéroscédasticité

Stéphane Adjemian

stephane.adjemian@univ-lemans.fr

Novembre 2025

Le modèle linéaire classique

Considérons le modèle de régression linéaire :

$$Y_i = X_i'\beta + \varepsilon_i, \quad i = 1, \ldots, n$$

Hypothèses classiques :

- $\mathbb{E}[\varepsilon_i | X_i] = 0$ (exogénéité)
- $\text{Var}(\varepsilon_i | X_i) = \sigma^2$ (homoscédasticité)
- $\text{Cov}(\varepsilon_i, \varepsilon_j | X_i, X_j) = 0$ pour $i \neq j$ (absence d'autocorrélation)

Qu'est-ce que l'hétéroscédasticité ?

Définition : L'hétéroscédasticité se produit lorsque la variance des erreurs n'est pas constante :

$$\text{Var}(\varepsilon_i | X_i) = \sigma_i^2 \neq \sigma^2$$

Conséquences :

- Les estimateurs MCO restent **sans biais** et **convergents**
- Mais ils ne sont plus **efficients** (BLUE)
- Les écarts-types usuels sont **biaisés**
- Les tests de Student et de Fisher sont **invalides**

Illustration graphique

Homoscédasticité vs Hétéroscédasticité

Homoscédasticité

Variance constante σ^2
Nuage de points avec dispersion uniforme autour de la droite de régression

Hétéroscédasticité

Variance σ_i^2 variable
Dispersion croissante (ou décroissante) le long de la droite

Préliminaires pour les démonstrations

Hypothèses générales :

1. $\mathbb{E}[\varepsilon_i | X_i] = 0$
2. Sous H_0 : $\text{Var}(\varepsilon_i | X_i) = \sigma^2$
3. Les observations sont i.i.d.
4. $\mathbb{E}[X_i X_i']$ existe et est définie positive

Estimateur MCO :

$$\hat{\beta} = \left(\sum_{i=1}^{n} X_i X_i' \right)^{-1} \sum_{i=1}^{n} X_i Y_i$$

Résidus MCO :

$$\hat{\varepsilon}_i = Y_i - X_i'\hat{\beta} = \varepsilon_i - X_i'(\hat{\beta} - \beta)$$

Test de Breusch-Pagan : Principe

Idée : Tester si la variance des résidus est liée aux variables explicatives.

Hypothèses :

- H_0 : $\sigma_i^2 = \sigma^2$ (homoscédasticité)
- H_1 : $\sigma_i^2 = h(Z_i'\alpha)$ où Z_i contient les régresseurs

Forme spécifique :

$$\sigma_i^2 = \alpha_0 + \alpha_1 Z_{i1} + \cdots + \alpha_p Z_{ip}$$

Test de Breusch-Pagan : Procédure

Étapes :

1. Estimer le modèle par MCO et obtenir les résidus $\hat{\varepsilon}_i$
2. Calculer les résidus au carré $\hat{\varepsilon}_i^2$
3. Régresser $\hat{\varepsilon}_i^2$ sur Z_i (généralement $Z_i = X_i$) :

$$\hat{\varepsilon}_i^2 = \alpha_0 + \alpha_1 Z_{i1} + \cdots + \alpha_p Z_{ip} + u_i$$

4. Récupérer le R^2 de cette régression auxiliaire
5. Calculer la statistique de test :

$$LM = n \cdot R^2 \xrightarrow{H_0} \chi_p^2$$

Test de Breusch-Pagan : Décision

Règle de décision :

- On rejette H_0 si $LM > \chi^2_{p,\alpha}$ (valeur critique)
- Ou si la p-valeur $< \alpha$ (seuil de significativité)

Avantages :

- Simple à mettre en œuvre
- Test basé sur le multiplicateur de Lagrange

Limites :

- Suppose une forme spécifique pour σ_i^2
- Sensible aux hypothèses de normalité

Démonstration Breusch-Pagan (1/4)

Modèle de variance :

$$\sigma_i^2 = Z_i'\alpha$$

où Z_i contient les variables explicatives (souvent $Z_i = X_i$).

Hypothèses :

- $H_0 : \alpha_1 = \cdots = \alpha_p = 0$ (seul $\alpha_0 = \sigma^2$)
- H_1 : au moins un $\alpha_j \neq 0$ pour $j \geq 1$

Régression auxiliaire :

$$\hat{\varepsilon}_i^2 = Z_i'\hat{\alpha} + u_i$$

Démonstration Breusch-Pagan (2/4)

Étape 1 : Résidus MCO
Les résidus MCO vérifient :

$$\hat{\varepsilon}_i = \varepsilon_i - X_i'(\hat{\beta} - \beta)$$

Sous H_0 et par le TCL :

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} \mathcal{N}(0, \sigma^2 Q^{-1})$$

où $Q = \mathbb{E}[X_i X_i']$.

Donc :

$$\hat{\varepsilon}_i^2 = \varepsilon_i^2 - 2\varepsilon_i X_i'(\hat{\beta} - \beta) + O_p(n^{-1})$$

Démonstration Breusch-Pagan (3/4)

Étape 2 : Régression auxiliaire
La somme des carrés expliquée de la régression de $\hat{\varepsilon}_i^2$ sur Z_i est :

$$SCE = \hat{\alpha}' \sum_{i=1}^{n} Z_i Z_i' \hat{\alpha}$$

où $\hat{\alpha}$ est l'estimateur MCO de la régression auxiliaire.

Sous H_0 : $\mathbb{E}[\hat{\varepsilon}_i^2] = \sigma^2$ et :

$$\hat{\alpha} \approx \left(\sum_{i=1}^{n} Z_i Z_i' \right)^{-1} \sum_{i=1}^{n} Z_i (\varepsilon_i^2 - \sigma^2)$$

Démonstration Breusch-Pagan (4/4)

Étape 3 : Distribution asymptotique
Par le TCL, sous H_0 :

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i(\varepsilon_i^2 - \sigma^2) \xrightarrow{d} \mathcal{N}(0, \Omega)$$

où $\Omega = \text{Var}(Z_i(\varepsilon_i^2 - \sigma^2))$.

La statistique nR^2 de la régression auxiliaire vérifie :

$$LM = n \cdot R^2 = \frac{SCE}{\hat{\sigma}^4} \xrightarrow{d} \chi_p^2$$

où p est le nombre de régresseurs dans Z_i (hors constante).

Conclusion : Sous H_0, $LM \sim \chi_p^2$ asymptotiquement.

Test de White : Principe

Idée : Version plus générale du test de Breusch-Pagan, sans spécifier de forme fonctionnelle.

Hypothèses :

- H_0 : Homoscédasticité
- H_1 : Toute forme d'hétéroscédasticité

Caractéristique : Utilise tous les régresseurs, leurs carrés et leurs produits croisés.

Test de White : Procédure

Étapes :

1. Estimer le modèle par MCO : $Y = X\beta + \varepsilon$
2. Calculer les résidus $\hat{\varepsilon}_i$ et leurs carrés $\hat{\varepsilon}_i^2$
3. Régresser $\hat{\varepsilon}_i^2$ sur :
   - Tous les régresseurs X_{ij}
   - Leurs carrés X_{ij}^2
   - Leurs produits croisés $X_{ij}X_{ik}$ pour $j \neq k$
4. Obtenir le R^2 de cette régression auxiliaire
5. Calculer :

   $$W = n \cdot R^2 \xrightarrow{H_0} \chi^2_q$$

   où q est le nombre de régresseurs dans la régression auxiliaire (moins 1)

Test de White : Décision et remarques

Règle de décision :

- Rejeter H_0 si $W > \chi^2_{q,\alpha}$ ou si p-valeur $< \alpha$

Avantages :

- Ne suppose aucune forme spécifique pour l'hétéroscédasticité
- Test très général et robuste

Limites :

- Peut inclure beaucoup de régresseurs (problème si n petit)
- Perte de degrés de liberté
- Possibilité de multicolinéarité dans la régression auxiliaire

Démonstration Test de White (1/3)

Principe du test :
Le test de White est une extension du test de Breusch-Pagan sans hypothèse spécifique sur la forme de l'hétéroscédasticité.

Régression auxiliaire :

$$\hat{\varepsilon}_i^2 = \gamma_0 + \sum_{j=1}^{k} \gamma_j X_{ij} + \sum_{j=1}^{k} \gamma_{jj} X_{ij}^2 + \sum_{j<\ell} \gamma_{j\ell} X_{ij} X_{i\ell} + u_i$$

Le vecteur γ contient tous les coefficients à tester.

Hypothèses :

- H_0 : $\gamma_j = 0$ pour tout $j \geq 1$
- H_1 : au moins un $\gamma_j \neq 0$

Démonstration Test de White (2/3)

Étape 1 : Développement de Taylor
Sous une forme générale d'hétéroscédasticité :

$$\sigma_i^2 = h(X_i)$$

On peut approximer h par un développement de Taylor du second ordre :

$$h(X_i) \approx h(0) + \nabla h(0)' X_i + \frac{1}{2} X_i' H(0) X_i$$

où $H(0)$ est la matrice hessienne en 0.

Cela justifie l'inclusion des termes quadratiques et croisés dans la régression auxiliaire.

Démonstration Test de White (3/3)

Étape 2 : Distribution asymptotique
Soit W_i le vecteur contenant $(1, X_{i1}, \ldots, X_{ik}, X_{i1}^2, \ldots, X_{ij}X_{i\ell}, \ldots)$.
Sous H_0, la statistique de test est :

$$W = n \cdot R^2$$

où R^2 provient de la régression de $\hat{\varepsilon}_i^2$ sur W_i.

Par un raisonnement similaire au test de Breusch-Pagan :

$$W = n \cdot R^2 \xrightarrow{d} \chi_q^2$$

où q est le nombre de régresseurs dans W_i (hors constante).

Remarque : $q = k + \frac{k(k+1)}{2}$ si on inclut tous les termes.

Test de Goldfeld-Quandt : Principe

Idée : Comparer les variances des résidus dans deux sous-échantillons.

Contexte : Utile quand on suspecte que la variance dépend d'une variable spécifique.

Hypothèses :

- $H_0 : \sigma_1^2 = \sigma_2^2$ (variances égales)
- $H_1 : \sigma_1^2 \neq \sigma_2^2$ (variances différentes)

Test de Goldfeld-Quandt : Procédure

Étapes :

1. Ordonner les observations selon la variable suspectée de causer l'hétéroscédasticité
2. Diviser l'échantillon en 3 parties :
   - Premier tiers : n_1 observations
   - Observations centrales : à exclure (environ c observations)
   - Dernier tiers : n_2 observations
3. Estimer le modèle séparément sur les deux sous-échantillons
4. Calculer les sommes des carrés des résidus SCR_1 et SCR_2
5. Calculer la statistique de test :

$$F = \frac{SCR_2/(n_2 - k)}{SCR_1/(n_1 - k)} \xrightarrow{H_0} F_{n_2-k,n_1-k}$$

où k est le nombre de paramètres

Test de Goldfeld-Quandt : Décision

Règle de décision :

- On rejette H_0 si $F > F_{n_2-k,n_1-k,\alpha}$
- Ou si la p-valeur $< \alpha$

Avantages :

- Simple à comprendre et à mettre en œuvre
- Ne nécessite pas d'hypothèse de normalité stricte
- Utile pour des formes monotones d'hétéroscédasticité

Limites :

- Nécessite de choisir une variable d'ordonnancement
- Perte d'observations (celles du milieu)
- Moins puissant que d'autres tests dans certains cas

Démonstration Test de Goldfeld-Quandt (1/2)

Principe :
On divise l'échantillon en deux groupes selon une variable Z_i :

- Groupe 1 : n_1 observations avec Z_i faible
- Groupe 2 : n_2 observations avec Z_i élevé

On estime le modèle séparément sur chaque groupe et on obtient :

- $SCR_1 = \sum_{i \in \text{Groupe 1}} \hat{\varepsilon}_i^2$
- $SCR_2 = \sum_{i \in \text{Groupe 2}} \hat{\varepsilon}_i^2$

Hypothèses :

- $H_0 : \sigma_1^2 = \sigma_2^2$
- $H_1 : \sigma_1^2 \neq \sigma_2^2$

Démonstration Test de Goldfeld-Quandt (2/2)

Distribution sous H_0 :
Sous l'hypothèse de normalité des erreurs et sous H_0 :

$$\frac{SCR_j}{\sigma^2} \sim \chi^2_{n_j - k}$$

pour $j = 1, 2$, où k est le nombre de paramètres.

Par conséquent, le rapport :

$$F = \frac{SCR_2/(n_2 - k)}{SCR_1/(n_1 - k)} = \frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2} \sim F_{n_2 - k, n_1 - k}$$

sous H_0.

Remarque : Ce test nécessite l'hypothèse de normalité, contrairement aux tests LM.

Maximum de vraisemblance sous hétéroscédasticité

Modèle avec hétéroscédasticité :

$$Y_i = X_i'\beta + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma_i^2)$$

où $\sigma_i^2 = \sigma_i^2(Z_i, \alpha)$ dépend de paramètres α à estimer.

Log-vraisemblance :

$$\ell(\beta, \alpha) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\log(\sigma_i^2(Z_i, \alpha)) - \frac{1}{2}\sum_{i=1}^{n}\frac{(Y_i - X_i'\beta)^2}{\sigma_i^2(Z_i, \alpha)}$$

Objectif : Maximiser $\ell(\beta, \alpha)$ par rapport à β et α conjointement.

Conditions de premier ordre

Dérivée par rapport à β :

$$\frac{\partial \ell}{\partial \beta} = \sum_{i=1}^{n} \frac{X_i(Y_i - X_i'\beta)}{\sigma_i^2(Z_i, \alpha)} = 0$$

Cela donne :

$$\hat{\beta}_{MV} = \left(\sum_{i=1}^{n} \frac{X_i X_i'}{\sigma_i^2(Z_i, \hat{\alpha})} \right)^{-1} \sum_{i=1}^{n} \frac{X_i Y_i}{\sigma_i^2(Z_i, \hat{\alpha})}$$

Dérivée par rapport à α :

$$\frac{\partial \ell}{\partial \alpha} = -\frac{1}{2} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \frac{\partial \sigma_i^2}{\partial \alpha} + \frac{1}{2} \sum_{i=1}^{n} \frac{(Y_i - X_i'\beta)^2}{(\sigma_i^2)^2} \frac{\partial \sigma_i^2}{\partial \alpha} = 0$$

Les deux conditions doivent être satisfaites simultanément.

Lien avec les MCG

Rappel : MCG (Moindres Carrés Généralisés)
Si $\Omega = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$ est **connu**, l'estimateur MCG est :

$$\hat{\beta}_{MCG} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y = \left(\sum_{i=1}^{n} \frac{X_i X_i'}{\sigma_i^2}\right)^{-1} \sum_{i=1}^{n} \frac{X_i Y_i}{\sigma_i^2}$$

Question : $\hat{\beta}_{MV} = \hat{\beta}_{MCG}$?

Réponse : NON en général !

- MCG suppose Ω **connu** (ou estimé indépendamment)
- MV estime β et α (donc σ_i^2) **conjointement**
- Les deux approches coïncident seulement si α est connu a priori

Lien avec les MCQG (MCG Faisables)

Rappel : MCQG/FGLS (Feasible GLS)
Procédure en deux étapes :

1. Estimer $\hat{\alpha}$ par une méthode préliminaire (ex: régression de $\hat{\varepsilon}_i^2$ sur Z_i)
2. Calculer $\hat{\sigma}_i^2(Z_i, \hat{\alpha})$ et appliquer MCG :

$$\hat{\beta}_{FGLS} = \left(\sum_{i=1}^{n} \frac{X_i X_i'}{\hat{\sigma}_i^2} \right)^{-1} \sum_{i=1}^{n} \frac{X_i Y_i}{\hat{\sigma}_i^2}$$

Question : $\hat{\beta}_{MV} = \hat{\beta}_{FGLS}$?

Réponse : NON en général !

- FGLS : estimation séquentielle (α puis β)
- MV : estimation conjointe et optimale
- Asymptotiquement, sous certaines conditions, les deux sont équivalents

Comparaison des trois approches

Critère	**MCG**	**FGLS**	**MV**
Ω connu ?	Oui	Non	Non
Estimation	Une étape	Deux étapes	Conjointe
Efficience	Optimale	Asympt. opt.	Optimale
Complexité	Faible	Moyenne	Élevée
Normalité	Non requise	Non requise	Requise

Remarques :

- Si Ω est connu : MCG est le meilleur (BLUE)
- Si Ω est inconnu : MV est asymptotiquement efficace sous normalité
- FGLS est une approximation pratique du MV
- Sous normalité et avec $n \to \infty$: $\hat{\beta}_{MV} \approx \hat{\beta}_{FGLS}$

Conditions d'équivalence MV et MCG

Question : Quand a-t-on $\hat{\beta}_{MV} = \hat{\beta}_{MCG}$?

Réponse : L'équivalence exacte requiert :

1. **Connaissance parfaite de Ω :**
$$\sigma_i^2 = \sigma_i^2(\alpha) \text{ avec } \alpha \text{ connu}$$
2. **Indépendance des paramètres :**
Les paramètres de variance α ne dépendent pas de β
3. **Normalité des erreurs :**
$$\varepsilon_i | X_i \sim \mathcal{N}(0, \sigma_i^2)$$

Remarque : Si Ω est parfaitement connu, alors :

$$\hat{\beta}_{MV} = \hat{\beta}_{MCG} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y$$

Conditions d'équivalence asymptotique MV et MCQG

Question : Quand a-t-on $\hat{\beta}_{MV} \approx \hat{\beta}_{FGLS}$ asymptotiquement ?

Conditions pour l'équivalence asymptotique :

1. **Convergence de l'estimateur préliminaire :**

$$\sqrt{n}(\hat{\alpha} - \alpha) = O_p(1)$$

   L'estimateur $\hat{\alpha}$ doit être $\sqrt{n}$-convergent

2. **Spécification correcte du modèle de variance :**
   La forme fonctionnelle $\sigma_i^2(Z_i, \alpha)$ doit être correctement spécifiée

3. **Conditions de régularité :**
   - $\mathbb{E}[X_i X_i']$ existe et est définie positive
   - $\mathbb{E}[\varepsilon_i^4] < \infty$ (moments d'ordre 4 finis)
   - Conditions de continuité sur $\sigma_i^2(\cdot)$

Résultat d'équivalence asymptotique

Théorème : Sous les conditions précédentes, on a :

$$\sqrt{n}(\hat{\beta}_{MV} - \beta) - \sqrt{n}(\hat{\beta}_{FGLS} - \beta) \xrightarrow{p} 0$$

Conséquences :

1. Les deux estimateurs ont la même distribution asymptotique :

$$\sqrt{n}(\hat{\beta}_{MV} - \beta) \xrightarrow{d} \mathcal{N}(0, (X'\Omega^{-1}X)^{-1})$$

$$\sqrt{n}(\hat{\beta}_{FGLS} - \beta) \xrightarrow{d} \mathcal{N}(0, (X'\Omega^{-1}X)^{-1})$$

2. Les deux estimateurs sont asymptotiquement efficaces
3. En pratique, MCQG est souvent préféré pour sa simplicité computationnelle

Attention : En échantillon fini, MV peut être plus efficace que MCQG si la normalité est vérifiée.

Cas particulier : Variance linéaire

Modèle de variance linéaire :

$$\sigma_i^2 = Z_i'\alpha$$

Estimation MCQG :

1. Estimer β par MCO : $\hat{\beta}_{MCO}$
2. Régresser $\hat{\varepsilon}_i^2$ sur Z_i : $\hat{\alpha} = (Z'Z)^{-1}Z'\hat{\varepsilon}^2$
3. Calculer $\hat{\sigma}_i^2 = Z_i'\hat{\alpha}$
4. Appliquer MCG : $\hat{\beta}_{FGLS} = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}Y$

Résultat : Si $n \to \infty$ et sous normalité :

$$\hat{\beta}_{MV} \approx \hat{\beta}_{FGLS}$$

avec une différence d'ordre $O_p(n^{-1})$.

Intuition : L'erreur d'estimation de α devient négligeable asymptotiquement.

Exemple : Hétéroscédasticité multiplicative

Spécification courante :

$$\sigma_i^2 = \sigma^2 \exp(Z_i'\gamma)$$

Approche MV :
La log-vraisemblance devient :

$$\ell(\beta, \sigma^2, \gamma) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2}\sum_{i=1}^{n} Z_i'\gamma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\frac{(Y_i - X_i'\beta)^2}{\exp(Z_i'\gamma)}$$

On maximise conjointement en $(\beta, \sigma^2, \gamma)$.

Approche FGLS :

1. Estimer par MCO et obtenir $\hat{\varepsilon}_i$
2. Régresser $\log(\hat{\varepsilon}_i^2)$ sur Z_i pour obtenir $\hat{\gamma}$
3. Calculer $\hat{\sigma}_i^2 = \exp(\hat{\delta}_0 + Z_i'\hat{\gamma})$
4. Appliquer MCG avec ces poids

Efficience asymptotique

Résultat théorique :
Sous normalité et conditions de régularité, l'estimateur du MV atteint la borne de Cramér-Rao :

$$\text{Var}(\hat{\theta}_{MV}) = \mathscr{I}(\theta)^{-1}$$

où $\theta = (\beta', \alpha')'$ et $\mathscr{I}(\theta)$ est la matrice d'information de Fisher.

Pour FGLS :

- Si $\hat{\alpha}$ est $\sqrt{n}$-convergent, alors $\hat{\beta}_{FGLS}$ est asymptotiquement équivalent à $\hat{\beta}_{MV}$
- $\sqrt{n}(\hat{\beta}_{FGLS} - \beta) \xrightarrow{d} \mathscr{N}(0, (X'\Omega^{-1}X)^{-1})$
- Même distribution asymptotique que MCG avec Ω connu

En pratique : FGLS est souvent préféré pour sa simplicité, malgré une légère perte d'efficience en échantillon fini.

Motivation

Problème : Sous hétéroscédasticité, la variance usuelle des MCO est biaisée :

$$\text{Var}_{\text{usuelle}}(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

Cette formule suppose $\text{Var}(\varepsilon_i | X_i) = \sigma^2$ (homoscédasticité).

Solution : Utiliser des estimateurs de variance **robustes** qui restent valides sous hétéroscédasticité.

Avantage :

- Garder l'estimateur MCO (simple à calculer)
- Corriger uniquement les écarts-types
- Tests et intervalles de confiance valides

Vraie variance des MCO sous hétéroscédasticité

Expression générale :
Sous hétéroscédasticité, la vraie variance de $\hat{\beta}$ est :

$$\mathsf{Var}(\hat{\beta}) = (X'X)^{-1}X'\Omega X(X'X)^{-1}$$

où $\Omega = \mathsf{diag}(\sigma_1^2, \ldots, \sigma_n^2)$.

En notation par observation :

$$\mathsf{Var}(\hat{\beta}) = \left(\sum_{i=1}^{n} X_i X_i'\right)^{-1} \left(\sum_{i=1}^{n} \sigma_i^2 X_i X_i'\right) \left(\sum_{i=1}^{n} X_i X_i'\right)^{-1}$$

C'est la formule "sandwich" : $(X'X)^{-1}$ entoure le terme central.

Estimateur de White (HC0)

Idée : Estimer σ_i^2 par $\hat{\varepsilon}_i^2$.

Estimateur de White (1980) :

$$\widehat{\text{Var}}_{HC0}(\hat{\beta}) = (X'X)^{-1} \left(\sum_{i=1}^{n} \hat{\varepsilon}_i^2 X_i X_i' \right) (X'X)^{-1}$$

ou en notation matricielle :

$$\widehat{\text{Var}}_{HC0}(\hat{\beta}) = (X'X)^{-1} X' \hat{\Omega} X (X'X)^{-1}$$

où $\hat{\Omega} = \text{diag}(\hat{\varepsilon}_1^2, \ldots, \hat{\varepsilon}_n^2)$.

Propriétés :

- Convergent : $\widehat{\text{Var}}_{HC0}(\hat{\beta}) \xrightarrow{p} \text{Var}(\hat{\beta})$
- Valide sous hétéroscédasticité arbitraire
- Peut être biaisé à la baisse en petit échantillon

Variantes : HC1, HC2, HC3

Pour améliorer les propriétés en petit échantillon, plusieurs corrections ont été proposées :

HC1 (correction des degrés de liberté) :

$$\widehat{\mathsf{Var}}_{HC1}(\hat{\beta}) = \frac{n}{n-k}(X'X)^{-1}\left(\sum_{i=1}^{n}\hat{\varepsilon}_i^2 X_i X_i'\right)(X'X)^{-1}$$

HC2 (MacKinnon & White, 1985) :

$$\widehat{\mathsf{Var}}_{HC2}(\hat{\beta}) = (X'X)^{-1}\left(\sum_{i=1}^{n}\frac{\hat{\varepsilon}_i^2}{1-h_{ii}} X_i X_i'\right)(X'X)^{-1}$$

où h_{ii} est l'élément diagonal de la matrice chapeau $H = X(X'X)^{-1}X'$.
HC3 (Davidson & MacKinnon, 1993) :

$$\widehat{\mathsf{Var}}_{HC3}(\hat{\beta}) = (X'X)^{-1}\left(\sum_{i=1}^{n}\frac{\hat{\varepsilon}_i^2}{(1-h_{ii})^2} X_i X_i'\right)(X'X)^{-1}$$

Comparaison des estimateurs robustes

Estimateur	Biais petit échantillon	Usage
HC0 (White)	Plus élevé	Grands échantillons
HC1	Moyen	Usage général
HC2	Faible	Recommandé
HC3	Très faible	Petits échantillons

Recommandations :

- **HC1** : le plus utilisé en pratique (bon compromis)
- **HC3** : recommandé pour $n < 250$ (Long & Ervin, 2000)
- **HC0** : acceptable si n est grand ($n > 500$)

Illustration : Matrice chapeau

Rappel : La matrice chapeau est définie par :

$$H = X(X'X)^{-1}X'$$

Propriétés :

- $\hat{Y} = HY$ (valeurs ajustées)
- $\hat{\varepsilon} = (I - H)\varepsilon$ (résidus)
- $0 \leq h_{ii} \leq 1$ pour tout i
- $\sum_{i=1}^{n} h_{ii} = k$ (nombre de paramètres)

Interprétation de h_{ii} :

- Mesure le "levier" de l'observation i
- h_{ii} élevé $\Rightarrow$ observation influente
- HC2 et HC3 donnent plus de poids aux observations à fort levier

Démonstration : Convergence de l'estimateur de White

Objectif : Montrer que $\widehat{\text{Var}}_{HC0}(\hat{\beta}) \xrightarrow{p} \text{Var}(\hat{\beta})$.

Étape 1 : La vraie variance est

$$\text{Var}(\hat{\beta}) = \left(\sum_{i=1}^{n} X_i X_i'\right)^{-1} \left(\sum_{i=1}^{n} \sigma_i^2 X_i X_i'\right) \left(\sum_{i=1}^{n} X_i X_i'\right)^{-1}$$

Étape 2 : L'estimateur de White utilise $\hat{\varepsilon}_i^2$ à la place de σ_i^2 :

$$\widehat{\text{Var}}_{HC0}(\hat{\beta}) = \left(\sum_{i=1}^{n} X_i X_i'\right)^{-1} \left(\sum_{i=1}^{n} \hat{\varepsilon}_i^2 X_i X_i'\right) \left(\sum_{i=1}^{n} X_i X_i'\right)^{-1}$$

Démonstration : Convergence de l'estimateur de White (suite)

Étape 3 : Il suffit de montrer que

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i^2 X_i X_i' \xrightarrow{p} \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2 X_i X_i'$$

Étape 4 : On a $\hat{\varepsilon}_i = \varepsilon_i - X_i'(\hat{\beta} - \beta)$, donc :

$$\hat{\varepsilon}_i^2 = \varepsilon_i^2 - 2\varepsilon_i X_i'(\hat{\beta} - \beta) + (\hat{\beta} - \beta)' X_i X_i'(\hat{\beta} - \beta)$$

Par la loi des grands nombres et le fait que $\hat{\beta} \xrightarrow{p} \beta$:

$$\frac{1}{n}\sum_{i=1}^{n}(\hat{\varepsilon}_i^2 - \varepsilon_i^2) X_i X_i' \xrightarrow{p} 0$$

Et $\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2 X_i X_i' \xrightarrow{p} \mathbb{E}[\sigma_i^2 X_i X_i']$.

Conclusion : L'estimateur de White est convergent.

Tableau comparatif

Critère	**Breusch-Pagan**	**White**	**Goldfeld-Quandt**
Forme fonctionnelle	Linéaire	Générale	Aucune
Puissance	Moyenne	Élevée	Variable
Simplicité	Élevée	Moyenne	Élevée
Normalité requise	Oui	Non	Oui
Taille échantillon	Moyenne	Grande	Moyenne

Recommandation : Le test de White est généralement le plus utilisé en pratique pour sa généralité.

Que faire en cas d'hétéroscédasticité ?

Si l'hétéroscédasticité est détectée, plusieurs solutions :

1. **Écarts-types robustes (White/Huber)**
   - Corriger les écarts-types sans changer les estimateurs
   - Solution la plus courante en pratique
2. **Moindres Carrés Généralisés (MCG)**
   - Si la forme de l'hétéroscédasticité est connue
   - Estimateurs plus efficaces
3. **Transformation des variables**
   - Logarithme, racine carrée, etc.
   - Peut stabiliser la variance

Exemple sous Python

```
import pandas as pd
import statsmodels.api as sm
from statsmodels.stats.diagnostic import het_breuschpagan
from statsmodels.stats.diagnostic import het_white

# Charger les données
data = pd.read_csv("donnees.csv")

# Estimer le modèle
X = sm.add_constant(data[['X1', 'X2', 'X3']])
model = sm.OLS(data['Y'], X).fit()

# Test de Breusch-Pagan
bp_test = het_breuschpagan(model.resid, X)
print(f"BP = {bp_test[0]:.2f}, p-value = {bp_test[1]:.4f}")

# Test de White
white_test = het_white(model.resid, X)
print(f"White = {white_test[0]:.2f}, p-value = {white_test[1]:.4f}")
```

Exemple Python : Estimateurs robustes

```
# Écarts-types robustes (différentes variantes)
model_hc0 = model.get_robustcov_results(cov_type='HC0')
model_hc1 = model.get_robustcov_results(cov_type='HC1')
model_hc2 = model.get_robustcov_results(cov_type='HC2')
model_hc3 = model.get_robustcov_results(cov_type='HC3')

print("Écarts-types standard:", model.bse)
print("Écarts-types HC1:", model_hc1.bse)

# MCG faisables (FGLS)
# Spécifier un modèle pour la variance
import numpy as np
log_resid_sq = np.log(model.resid**2 + 1e-8)
variance_model = sm.OLS(log_resid_sq, X).fit()
weights = 1 / np.exp(variance_model.fittedvalues)

# Estimation par MCG
model_gls = sm.WLS(data['Y'], X, weights=weights).fit()
print(model_gls.summary())
```

Interprétation des résultats

Exemple de sortie :

```
BP = 12.45, p-value = 0.0060
White = 15.32, p-value = 0.0180
```

Interprétation :

- Les statistiques BP et White suivent des lois χ^2
- p-valeurs < 0.05 pour les deux tests
- On rejette H_0 au seuil de 5%
- Conclusion : présence d'hétéroscédasticité
- Action : utiliser des écarts-types robustes

Points clés à retenir

1. L'hétéroscédasticité affecte l'efficience des estimateurs MCO et invalide les tests usuels
2. Trois tests principaux pour détecter l'hétéroscédasticité :
   - **Breusch-Pagan :** forme linéaire spécifique
   - **White :** forme générale, le plus utilisé
   - **Goldfeld-Quandt :** comparaison de sous-échantillons
3. Deux approches principales pour traiter l'hétéroscédasticité :
   - **Estimateurs robustes** (HC0, HC1, HC2, HC3) : corriger les écarts-types
   - **MCG/FGLS** : ré-estimer le modèle avec pondération optimale
4. En pratique : utiliser HC1 ou HC3 pour les écarts-types robustes

Références

- Greene, William H. (2017). *Econometric analysis*. Pearson.
- Schmidt, Peter (1976). *Econometrics*. Taylor & Francis.