

Dossier 6 - Problème d'hétéroscédasticité des résidus

Thomas Chuffart - thomas.chuffart@univ-amu.fr

1 Sujet

Considérez le modèle de régression

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + u.$$

Si $\beta_1 = 0$, et sous certaines conditions sur les aléas u , la statistique de Student associée à l'explicative x_1 suit la loi de Student à $n - 3$ degrés de liberté, $n =$ taille de l'échantillon. Si les conditions sur les aléas ne sont pas vérifiées, la statistique ne suit pas une loi de Student, même si $\beta_1 = 0$. Ce fait signifie que la probabilité de rejeter l'hypothèse (vraie) que $\beta_1 = 0$ par un test de Student n'est pas égale au niveau nominal (α) du test.

Mettez en oeuvre une expérience de simulation afin d'estimer la probabilité de rejet au niveau nominal $\alpha = 5\%$ pour le processus caractérisé par les variables x_1 et x_2 trouvées dans le fichier de données, $\alpha = 1$, $\beta_1 = 0$, $\beta_2 = -2$, $n = 20$ et les aléas IID.

Pour vous aider, il va vous falloir simuler des y à l'aide de x_1 et x_2 et des aléas normaux d'espérance nulle et de variance unitaire pour chaque observation (la fonction `normrnd` sur Matlab fait ça très bien). Créez de l'hétéroscédasticité, c'est à dire que pour chaque observations, la variance de la loi normale sera différente de 1. Refaites votre expérience de simulation et calculez à nouveau la probabilité de rejet. Qu'observez vous?

Lors de la présentation, présentez bien les conditions pour que la statistique de Student soit distribuée selon une Student $n - 3$. Pour l'expérience, l'idée va être de répéter le test un certain nombre de fois (10 000) afin de calculer la p-value pour chaque test. On rejettera quand cette p-value sera inférieure à 0.05. Vous pouvez donc calculer le nombre moyen de rejet et le comparer au niveau nominal puis conclure.

Quel test proposez vous d'utiliser dans ce cas là? Présenter le brièvement.