# Les RESIDUS RECURSIFS

Nous allons définir de nouveaux résidus et montrer qu'ils ont les mêmes propriétés que les erreurs contrairement aux résidus classiques. Ils vont nous aider à construire de meilleurs tests.
Pourquoi n'a-t-on pas utilisé directement les résidus récursifs ? Parce qu'ils sont plus difficiles à calculer que les résidus classiques. Comme avant les années 70 nous ne disposions pas d'ordinateur puissant, toute la théorie économétrique s'est construite avec les résidus classiques. Mais tous les tests peuvent s'effectuer sur les résidus récursifs, certains sont même beaucoup plus simples qu'avec les résidus classiques, comme le test d'autocorrélation d'ordre 1 et le test de Stabilité.
Après avoir défini les résidus récursifs qui sont basés sur les résidus de prévision, nous définirons de nouveaux test comme celui de Von Neumann ou celui de Brown, Durbin et Evans.

# 1 Les résidus récursifs

## 1.1 Définition des résidus récursifs

### 1.1.1 Notations :

X matrice des k variables exogènes non aléatoires dans l'échantillon de taille n.
$X_{(t)}$ matrice dans l'échantillon de taille t des t premières lignes de X
$\overrightarrow{y}_{(t)}$ vecteur des t premières valeurs de $\overrightarrow{y}$
$\overrightarrow{\epsilon_{(t)}}$ vecteur des t premières valeurs de $\overrightarrow{\epsilon}$
$\overrightarrow{\widehat{a_{(t)}}}$ estimateur des MCO obtenu avec les t premières valeurs de l'échantillon
${}^t\overrightarrow{x_{(t)}}$ vecteur transposé de la ligne t de X
Les erreurs sont homoscédastiques non-autocorrélées et suivent une loi Normale.

### 1.1.2 Définition des résidus de prévision

A l'aide de l'échantillon de taille t-1 ( de 1 à t-1) on estime le vecteur $\overrightarrow{a}$ par $\overrightarrow{\widehat{a_{(t-1)}}} = ({}^tX_{(t-1)}X_{(t-1)})^{-1}\,{}^tX_{(t-1)}\overrightarrow{y}_{(t-1)}$
Le résidu de prévision en t utilise l'estimation de $\overrightarrow{a}$ en t-1 soit $\overrightarrow{\widehat{a_{(t-1)}}}$, pour t allant de k+1 à n. Il y a donc n-k résidus de prévision possibles car t doit être >k pour pouvoir faire les MCO.
Notons $u_t$ le résidu de prévision en t

$$u_t = y_t - {}^t\overrightarrow{x_{(t)}}\overrightarrow{\widehat{a_{(t-1)}}}$$
$$u_t = y_t - {}^t\overrightarrow{x_{(t)}}({}^tX_{(t-1)}X_{(t-1)})^{-1}\,{}^tX_{(t-1)}\overrightarrow{y}_{(t-1)}$$

$$u_t = {}^t\overrightarrow{x_{(t)}}\overrightarrow{a} + \epsilon_t - {}^t\overrightarrow{x_{(t)}}\overrightarrow{\widehat{a_{(t-1)}}} = {}^t\overrightarrow{x_{(t)}}(\overrightarrow{a} - \overrightarrow{\widehat{a_{(t-1)}}}) + \epsilon_t$$

$$\overrightarrow{\widehat{a_{(t-1)}}} = ({}^tX_{(t-1)}X_{(t-1)})^{-1t}X_{(t-1)}\overrightarrow{y}_{(t-1)} = ({}^tX_{(t-1)}X_{(t-1)})^{-1t}X_{(t-1)}(X_{(t-1)}\overrightarrow{a} + \overrightarrow{\epsilon_{(t-1)}})$$

$$\overrightarrow{\widehat{a_{(t-1)}}} - \overrightarrow{a} = ({}^tX_{(t-1)}X_{(t-1)})^{-1t}X_{(t-1)}\overrightarrow{\epsilon_{(t-1)}})$$

$$u_t = \epsilon_t -^t \overrightarrow{x_{(t)}}({}^tX_{(t-1)}X_{(t-1)})^{-1t}X_{(t-1)}\overrightarrow{\epsilon_{(t-1)}})$$

### 1.1.3 Propriétés des résidus de prévision

On a E($u_t$) $= 0$ car les espérances des erreurs $\epsilon$ sont nulles.
Calcul de la covariance Cov(u$_t, u_s$) $= E(u_tu_s)$ avec s>t

$$E(u_tu_s) = E[(\epsilon_t - {}^t\overrightarrow{x_{(t)}}({}^tX_{(t-1)}X_{(t-1)})^{-1t}X_{(t-1)}\overrightarrow{\epsilon_{(t-1)}}))(\epsilon_s - {}^t\overrightarrow{x_{(s)}}({}^tX_{(s-1)}X_{(s-1)})^{-1t}X_{(s-1)}\overrightarrow{\epsilon_{(s-1)}}))]$$

En remarquant que chaque expression est égale à sa transposée on est conduit à calculer les espérances suivantes
On E($\epsilon_t\epsilon_s$) $= 0$ car s>t
de même E($\epsilon_t$ ${}^t\overrightarrow{\epsilon_{(s-1)}}$) $= (E(\epsilon_t\epsilon_1)$ ...$E(\epsilon_t\epsilon_i)$...$E(\epsilon_t\epsilon_{s-1}) = (0\ 0...\sigma^2...0)$ nul sauf pour i=t
E($\epsilon_s$ ${}^t\overrightarrow{\epsilon_{(t-1)}}$) $= (0...0)$ toujours nul car s>t-1
E($\overrightarrow{\epsilon_{(t-1)}}{}^t\overrightarrow{\epsilon_{(s-1)}} = \sigma^2[I_{(t-1)}0_{(t-1,s-1)}]$ car on ne retrouve $\sigma^2$ que pour les valeurs de i $\leq (t-1)$

$$E(u_tu_s) = 0 - \sigma^{2t}\overrightarrow{x_{(t)}}({}^tX_{(s-1)}X_{(s-1)})^{-1}\overrightarrow{x_{(s)}} + 0 + \sigma^{2t}\overrightarrow{x_{(t)}}({}^tX_{(s-1)}X_{(s-1)})^{-1}\overrightarrow{x_{(s)}} = 0$$

Ainsi les covariances sont nulles
On refait le même calcul cette fois pour s=t et on obtient le résultat suivant

$$v(u_t) = \sigma^2(1 +^t \overrightarrow{x_{(t)}}({}^tX_{(s-1)}X_{(s-1)})^{-1}\overrightarrow{x_{(s)}}) = \sigma^2h(t)$$

en posant

$$h(t) = (1 +^t \overrightarrow{x_{(t)}}({}^tX_{(s-1)}X_{(s-1)})^{-1}\overrightarrow{x_{(s)}})$$

Chaque u$_t$ est une combinaison linéaire des $\epsilon_i$ il suit donc également une loi Normale, comme ils ont des Cov nulles le vecteur des u$_i$ de dimension (n-k) ( car nous avons montré que l'on ne peut construire que (n-k) résidus de prévision ) suit une loi N$_{n-k}$.

### 1.1.4 Définition et propriétés des résidus récursifs

On définit les residus récursifs w$_t$ à partir des résidus de prévision u$_t$ pour t variant de k+1 à n par

$$w_t = \frac{u_t}{\sqrt{h(t)}}$$

Si les variables explicatives sont non aléatoires on a donc

- E(w$_t$) $= 0$ Cov(w$_t, w_s$) $= 0$ et V(w$_t$) $= \sigma^2$
- Les w$_t$ suivent une loi Normale et le vecteur $\overrightarrow{W}$ des w$_t$ une loi N$_{n-k}$.

On constate donc que les w ont les mêmes propriétés que les erreurs contrairement aux résidus ( voir chapitre2 sur les résidus).

- on peut démontrer aussi qu'en notant $\mathrm{SCR}_t$ la somme des carrés des résidus classiques jusqu'à t on a
$$SCR_t = SCR_{t-1} + w_t^2 \text{ pour t=k+1, ....,n}$$
- on en déduit que
$$SCR_n = \sum_{t=k+1}^{n} w_t^2$$
- La moyenne des résidus récursifs n'est pas nulle $\overline{w} \neq 0$ mais on a bien sur $\mathrm{E}(\overline{w})=0$

# 2 Les tests classiques revus avec les résidus récursifs

## 2.1 Les tests identiques avec les deux résidus

- Les tests de Normalité Skewness, Kurtosis et Jarque et Berra s'effectuent de la même façon
- Les tests d'hétéroscédasticité également
- Les tests d'autocorrélation d'ordre supérieur à 1, nous n'y reviendrons donc pas.

## 2.2 le test de von Neumann

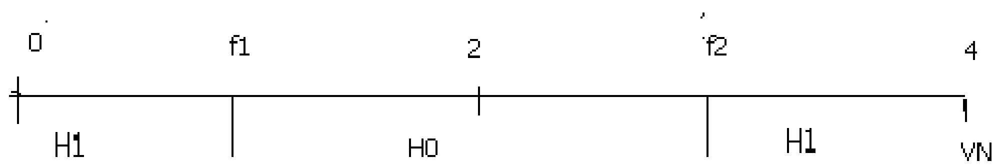
Il remplace le test de Durbin et Watson.
Nous avons déjà parlé du test de von Neumann lors de la définition de la statistique DW.
Von Neumann a testé le cas suivant: prenons une variable $z_t = \rho z_{t-1} + u_t$ où u est un bruit blanc

H0 : il y a non autocorrélation $\iff$ $\rho$ =0
H1 : il y a autocorrélation $\iff$ $\rho \neq$0 avec toujours $|\rho| < 1$

Von Neumann a montré que sous l'hypothèse H0 la statistique

$$VN = \frac{\sum_{t=2}^{n}(z_t - z_{t-1})^2}{\sum_{t=1}^{n}(z_t - \overline{z})^2}$$

suit asymptotiquement une loi Normale N(2,4/n)
Cette statistique est comprise entre 0 et 4 car la loi est symétrique avec une espérance égale à 2 et de plus VN est positif, les bornes du test sont f1 et f2, si VN est entre f1 et f2 on décide H0 il y a non autocorrélation des z, si VN est <f1 ou VN>f2 on décide H1.
Nous n'avons pas pu utiliser ce test avec les résidus classique car sous $\mathrm{H}_0$ les résidus sont autocorrélés.
Par contre dans le cadre des résidus récursifs ceux-ci sont indépendants sous $\mathrm{H}_0$, On va donc appliquer le test de von Neumann pour ces résidus




$$VN = \frac{\sum_{t=k+1}^{n} (w_t - w_{t-1})^2}{\sum_{t=k+1}^{n} (w_t - \overline{w})^2}$$

- qui suivra asymptotiquement une loi N(2,4/n) avec le même intervalle de confiance que ci-dessus.
- si n est plus petit (n-k<60) et à condition que n soit nettement supérieur à k (n-k>10 au moins), Press et Brooks (1969) ( voir la table aussi dans l'ouvrage de Harvey The econometric analysis of time series) ont donné des estimations pour les bornes f1 et f2.

Par contre, comme dans le cas des résidus classiques, ce test n'est pas valable dans le cadre des modèles autorégressifs, le test de Durbin peut être utilisé avec les résidus récursifs.

## 2.3 Les tests de stabilité de Brown, Durbin et Evans

Voir l'article de Brown, Durbin et Evans dans the Journal of the Royal Statistical Society Serie B Vol. 37 n°2 de 1975
Le test de Chow peut être utilisé sur les résidus récursifs . Mais on utilise le plus souvent le test propre aux résidus récursifs:
Le test de Brown, Durbin et Evans (B.D.E.) est un test de stabilité des coefficients très différent du test de CHOW pour plusieurs raisons:

- Il utilise non les résidus classiques mais les résidus récursifs
- Il permet en une seule fois de rechercher plusieurs points de rupture.
- Il n'est pas nécessaire de connaitre ces éventuels points de rupture.

Il nécessite cependant comme le test de Chow, l'**homoscédasticité** des erreurs. Il a comme défaut d'être très sensible à cette hypothèse. Quand on se trouve avec un modèle fortement hétéroscédastique alors le test BDE n'est plus applicable.
Hypothèses du test :

$$\begin{array}{rcll} & & k+1\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ & t_0\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ n \\ Y_t & = & b_1 X_{1t} + .... + b_k X_{kt} + \epsilon_t & Y_t = c_1 X_{1t} + .... + c_k X_{kt} + \epsilon_t \end{array}$$

- H0: le modèle est stable : $b_i = c_i$ pour tout i=1,.....k

  H1: le modèle est instable : l'un au moins des $b_i \neq c_i$

Il existe deux tests de stabilité de BDE :

### 2.3.1 Le test des CUSUM

ou test des sommes cummulées.
Si un coefficient $a_i$ du modèle est constant jusqu'en $t_0$ puis prend une autre valeur ensuite cela signifie lorsque l'on fait les MCO sur le modèle que de k+1 jusqu'à $(t_0 - 1)$ on a $E(w_i) = 0$ et qu'ensuite de $t_0$ à n $E(w_i) \neq 0$.Ce qui suggère d'examiner l'évolution des sommes cumulées les

$$Cusum_t = \sum_{i=k+1}^{t} w_i/s$$

avec $s^2 = \frac{\sum w_i^2}{(n-k)}$ estimation de $\sigma^2$
Sous l'hypothèse $H_0$ on a (sans démonstration) $E(cusum_t) = 0$ , $V(cusum_t) = t - k$ et cov(t,t')=t-k si t<t'
De plus, $cusum_t$ suit asymptotiquement une loi Normale (0,t-k)
Les auteurs ont démontré que sous l'hypothèse $H_0$, les valeurs de $cusum_t$ sont à l'intérieur de l'intervalle défini par les deux droites en fonction de t

$$y = \pm \frac{a(n + 2t - 3k)}{\sqrt{n - k}}$$

La valeur de a est a=1.14 si le risque $\alpha = 1\%$ , a=0.95 si $\alpha = 5\%$ et 0.85 pour $\alpha = 10\%$
Si un point ou plusieurs sortent de l'intervalle ils seront considérés comme points de rupture.
On constate donc que ce test fournit les éventuels points de rupture.
Quand un groupe de points ressortent de l'intervalle de confiance, on prend comme point de rupture celui qui s'éloigne le plus.

### 2.3.2 Le test des CUSUM of Square

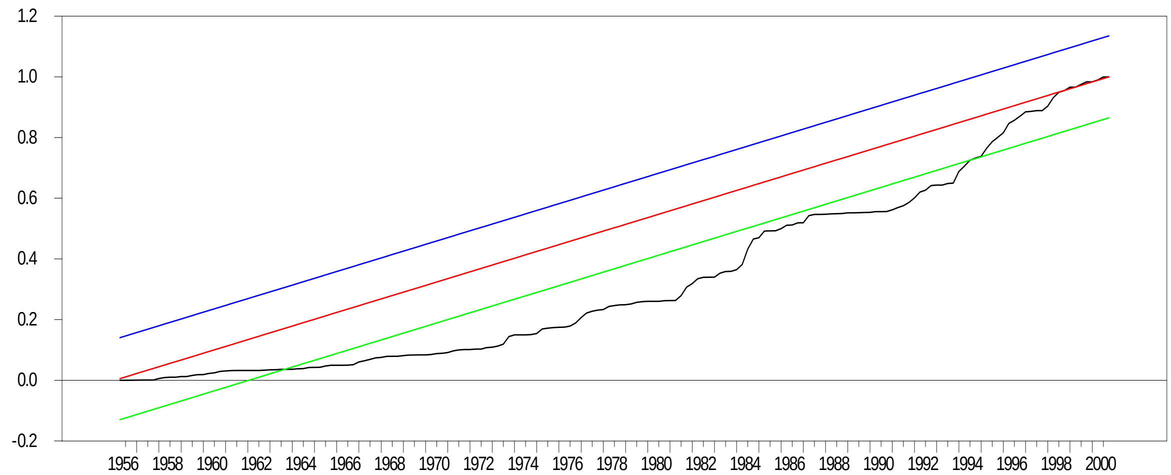
ou test des sommes des carrés.
Il est basé sur la statistique $S_t$

$$S_t = \frac{\sum_{i=k+1}^{t} w_i^2}{\sum_{j=k+1}^{n} w_i^2}$$

On constate que c'est une suite monotone croissante entre 0 pour t=k et 1 pour t=n.
On montre que sous l'hypothèse $H_0$ on a les propriétés suivantes :

$$E(S_t) = \frac{t - k}{n - k} \text{ et S suit sous } H_0 \text{ une loi } \beta$$

Le test est basé comme le précédent : si $H_0$ est vrai alors $S_t$ est proche de la droite en t $\frac{t-k}{n-k}$ si $H_1$ est vrai il s'en éloigne.




On construit donc un intervalle de confiance limité par les deux droites

$$d_1 = \frac{t-k}{n-k} + c_0 \text{ et } d_2 = \frac{t-k}{n-k} - c_0$$

Les valeurs de $c_0$ sont données dans la table de Durbin dans son article dans Biometrika de 1969. par exemple pour 0.5(n-k)-1=100 on a $c_0 = 0.1$
**Exemple**:
Test BDE pour la consommation des ménages aux USA. Nous sommes dans le cas d'un modèle fortement hétéroscédastique comme nous l'avons vu dans les révisions du chapitre 3, le test BDE est ici déconseillé. Il pourra être remplacé par Chow sur ces résidus récursifs. Nous allons donner cependant le résultat juste pour illustrer ce test.
Ici 0.5(n-k)-1=0.5(184-5)-1=88.5 la valeur de c0 pour un test bilatéral donc $\alpha = .025$ est approximaativement de c0=0.135
Graphique du résultat: (voir ci-dessus)

On constate que dès 1964 les CUSUMS of square sortent de l'intervalle de confiance, pour n'y revenir que vers la fin contraints et forcés car la somme finale doit être égale à 1. Ceci nous permet de remarquer que ce test n'est plus valable près de la fin d'échantillon tout comme le test de Chow. Il semble donc que le modèle soit toujours instable. Mais n'oublions pas que nous l'avons utilisé avec les mauvaises hypothèses d'hétéroscédasticité. Qu'elle est la part de l'instabilité et celle de l'hétéroscédasticité ? Mystère !
En conclusion nous pouvons dire que ce test est très bien lorsqu'il y a homoscédasticité ou une faible hétéroscédasticité, car on visualise les points de rupture et qu'alors on peut présenter un beau graphique pour illustrer ce problème. Sinon le test de Chow est meilleur.

Nous venons de voir que l'on peut aussi bien utiliser les résidus récursifs que les résidus

classiques et que le test d'autocorrélation d'ordre 1 est bien meilleur dans le cadre des résidus récursifs.