TEST DE CHOW

Dans ce chapitre on va aborder la critique de l'hypothèse: H_0^3 les coefficients des MCO sont des constantes inconnues.

1 Théorie

1.1 Présentation

Ce test étudie une forme particulière de l'instabilité des coefficients théoriques du modèle. Plusieurs forment d'instabilité peuvent se produire.

- A chaque valeur de t les coefficients varient, c'est la forme générale de l'instabilité.

$$Y_t = a_{1t}X_{1t} + ... + a_{kt}X_{kt} + \epsilon_t$$

  Il est bien entendu impossible d'estimer tous ces coefficients il y en a k*n donc beaucoup trop, dans ces modèles que l'on appelle modèles à coefficients variables on est amené à poser des hypothèses sur ces coefficients afin d'estimer un nombre beaucoup moins important de coefficients. Nous aborderons dans un chapitre ultérieur ce problème.

- Chow aborde le problème beaucoup plus simple d'un seul point de rupture. On va appeler **point de rupture** un point avant lequel les coefficients des MCO prennent une valeur inconnue mais fixe et à partir duquel les coefficients sont à nouveau fixes mais avec une autre valeur. On note t_0 le point de rupture, avant ce point les coefficients sont les b_i et à partir de ce point les coefficients sont c_i.

$$\begin{array}{rcll} & & 1\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_t_0 & \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_n \\ Y_t & = & b_1X_{1t} + ... + b_kX_{kt} + \epsilon_t & Y_t = c_1X_{1t} + ... + c_kX_{kt} + \epsilon_t \end{array}$$

On va construire le test de CHOW

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

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

1.2 Loi de Fisher

Si on fait les **hypothèses classiques** sur les erreurs de MCO, on va être conduit à effectuer un test de Fisher classique. Sous l'hypothèse H0 de stabilité les coefficients $b_i = c_i = a_i$ en nommant a_i la valeur commune des coefficients.

1.2.1 Sous l'hypothèse H0

Pour estimer les coefficients sous H0 il suffit d'estimer leurs valeurs communes les a_i dans le modèle $Y_t = a_1 X_{1t} + ... + a_k X_{kt} + \epsilon_t$ sur tout l'échantillon de taille donc n.

$$\overrightarrow{Y} = \underset{(n,k)}{X} \overrightarrow{a} + \overrightarrow{\epsilon}$$

Les MCO donnent $\overrightarrow{\widehat{a}}$, la somme des carrés des résidus notée SCR_0, le nombre de degrés de liberté est (n-k) et SCR_0/σ^2 suit une loi du χ^2_{n-k} à (n-k) degrés de liberté.

1.2.2 Sous l'hypothèse H1

Dans cette hypothèse il faut partager le modèle en deux parties:

- Dans l'échantillon de 1 à $\text{t}_0 - 1$ la taille est $n_1 = (t_0 - 1)$

  Le modèle s'écrit dans cet échantillon $\overrightarrow{Y_1} = X_1 \overrightarrow{b} + \overrightarrow{\epsilon_1}$ avec $\underset{(t_0-1,1)}{\overrightarrow{Y_1}}$ correspondant aux $(\text{t}_0 - 1)$ premières valeurs du vecteur $\overrightarrow{Y}$, $\underset{(t_0-1,k)}{X_1}$ correspondant aux $(\text{t}_0 - 1)$ premières lignes de la matrice X, $\overrightarrow{b}$ aux vecteur des b_i

  On fait les MCO, on estime donc les b_i, on obtient $\overrightarrow{\widehat{b}}$, la somme des carrés des résidus notée SCR_1, le nombre de degrés de liberté est $(n_1 - k)$ et SCR_1/σ^2 suit une loi du $\chi^2_{n_1-k}$ à $(\text{n}_1\text{-k})$ degrés de liberté.

- Dans la seconde partie de l'échantillon de t_0 à n la taille est $n_2 = n - (t_0 - 1)$. Il est évident que $n_1 + n_2 = n$.

  Le modèle s'écrit dans cet échantillon $\overrightarrow{Y_2} = X_2 \overrightarrow{c} + \overrightarrow{\epsilon_2}$ avec $\underset{(n-t_0+1,1)}{\overrightarrow{Y_2}}$ correspondant aux $(n - t_0 + 1)$ dernières valeurs du vecteur $\overrightarrow{Y}$, $\underset{(n-t_0+1,k)}{X_2}$ correspondant aux $(n - t_0 + 1)$ dernières lignes de la matrice X, $\overrightarrow{c}$ aux vecteur des c_i

  On fait les MCO, on estime donc les c_i, on obtient $\overrightarrow{\widehat{c}}$, la somme des carrés des résidus notée SCR_2, le nombre de degrés de liberté est $(n_2 - k)$ et SCR_2/σ^2 suit une loi du $\chi^2_{n_2-k}$ à $(\text{n}_2\text{-k})$ degrés de liberté.

- Globalement sous H1 le modèle peut s'écrire

$$\overrightarrow{Y} = \begin{pmatrix} \underset{(n_1,k)}{X_1} & \underset{(n_1,k)}{0} \\ \underset{(n2,k)}{0} & \underset{(n_2,k)}{X_2} \end{pmatrix} \begin{pmatrix} \overrightarrow{b} \\ \overrightarrow{c} \end{pmatrix} + \begin{pmatrix} \overrightarrow{\epsilon_1} \\ \overrightarrow{\epsilon_2} \end{pmatrix}$$

  Les matrices $\underset{(n_1,k)}{0}$ et $\underset{(n2,k)}{0}$ montrent la non corrélation entre les deux sous-échantillons. Ce qui conduit au fait que les résultats sur l'ensemble de l'échantillon sont la somme des résultats de la première sous-partie et de la seconde. En conséquence, si on effectue la méthode des MCO sur ce modèle il est évident que la somme des carrés des résidus sur ce dernier modèle $SCR_a = SCR_1 + SCR_2$ avec $(\text{n}_1 - k + n_2 - k)$ degrés de liberté soit (n-2k) degrés de liberté, la variable SCR_a/σ^2 suit un χ^2 à n-2k degrés de liberté.

1.2.3 Construction du test de CHOW

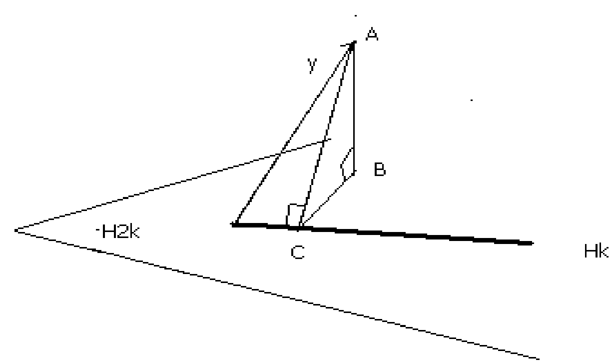
Si H0 est l'hypothèse correcte alors $\overrightarrow{\widehat{b}}$ et $\overrightarrow{\widehat{c}}$ sont proches de $\overrightarrow{\widehat{a}}$ et SCR_a proche de SCR_0, on regarde donc leur écart , sous H0 vrai ce écart doit être faible. Loi de cet écart sous H0 vrai :
En regardant l'écriture du modèle global sous H1 on voit que la variété linéaire est formée de 2k vecteurs elle est notée H_{2k}: les k premiers vecteurs sont formés des k variables explicatives jusqu'en $(t_0 - 1) = n_1$ puis ensuite des 0 jusqu'à n; les k derniers sont formés de 0 de 1 à $(t_0 - 1)$ puis des valeurs des variables explicatives jusqu'à n. On a donc une variété linéaire formée de 2k vecteurs sous H1.
Par contre sous H0

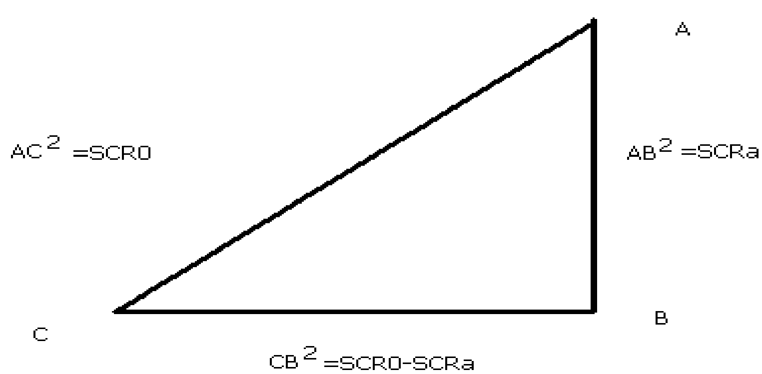
$$\overrightarrow{Y} = \underset{(n,k)}{X} \overrightarrow{a} + \overrightarrow{\epsilon}$$

la variété linéaire est formée des k variables sur tout l'échantillon, elle est notée H_k. On constate que

$$\begin{pmatrix} X_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ X_2 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = X$$

Les vecteurs qui définissent H_k sont donc la somme de vecteurs qui définissent H_{2k} par conséquent le sous-ensemble de R^n H_k est inclus dans H_{2k}.





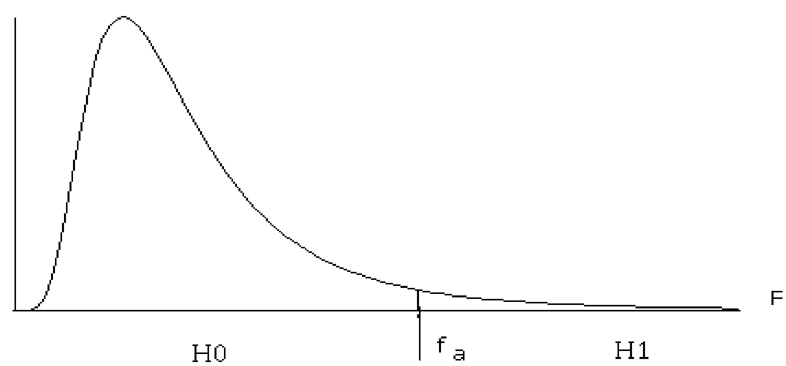
AB est le vecteur des résidus pour la projection de Y sur H_{2k} (ce qui correspond à l'hypothèse H1), AB est orthogonal à H_{2k}. Comme on l'a vu H_k est inclus dans H_{2k} (donc ici graphiquement une droite), AC est le vecteur des résidus pour la projection de Y sur H_k(ce qui correspond à l'hypothèse H0). Le carré de la longueur de ces vecteurs est la somme des carrés des résidus. Donc $AB^2 = SCRa$ et $AC^2 = SCR_0$,et alors d'après Pythagore $CB^2 = AB^2 - AC^2 = SCR_0 - SCRa$.

$$\begin{aligned} \text{on a vu que } \frac{SCR_0}{\sigma^2} &= \chi^2_{n-k} \text{ et } \frac{SCRa}{\sigma^2} = \chi^2_{n-2k} \\ \text{donc } \frac{SCR_0 - SCRa}{\sigma^2} &= \chi^2_{n-k-(n-2k)} = \chi^2_k \end{aligned}$$

Pour éliminer la variance σ^2 il suffit de faire le rapport de deux χ^2 indépendants donc de supports orthogonaux. Le graphique montre que AB est orthogonal à CB (donc $SCR_0 - SCRa$ est indépendant de $SCRa$).
On passe donc à la loi de FISHER rapport de deux χ^2 indépendants divisés par leur nombre de degrés de liberté:

$$\frac{\frac{SCR_0 - SCRa}{\sigma^2}}{\frac{SCRa}{\sigma^2}} . \frac{n-2k}{k} = \frac{SCR_0 - SCRa}{SCRa} . \frac{n-2k}{k} = F_{k,n-2k}$$

Sous H0 vrai le Fisher est petit car SCR_0 est proche de SCRa. Le test se construit donc



f_α étant la borne du Fisher au risque α

2 Application " manuelle"

Avant de voir comment utiliser le sous-programme qui fera ce test automatiquement, nous allons détailler manuellement un exemple.

2.1 Récupération des éléments

Le programme se trouve dans chow.prg et les données dans mco2.rat
end xxx

cal 1971 1 4
all 2005:4
open data mco2.xls
data(for=excel,org=cols) /
com start = 1971:1
com end = 2005:4
smpl start end
lin Y2 / residus1
 # constant X1 X2

SUPPOSONS QUE LES ERREURS ONT LES BONNES PROPRIETES. On pourra le vérifier avec les autres tests du chapitre 3.

```
Linear Regression - Estimation by Least Squares
Dependent Variable Y2
Quarterly Data From 1971:01 To 2005:04
Usable Observations    140      Degrees of Freedom   137
Centered R**2     0.999202      R Bar **2   0.999190
Uncentered R**2   0.999938      T x R**2     139.991
Mean of Dependent Variable      687633.34015
Std Error of Dependent Variable 200772.53945
Standard Error of Estimate        5713.37876
Sum of Squared Residuals        4472049468.0
Regression F(2,137)               85755.2323
Significance Level of F           0.00000000
Log Likelihood                   -1408.21431
Durbin-Watson Statistic             1.571691

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                13694.974760  2010.146553      6.81292  0.00000000
2.  X1                          0.820824     0.056480     14.53303  0.00000000
3.  X2                          0.483999     0.004052    119.45363  0.00000000
```

Comme on l'a vu dans la théorie, Chow a construit son test en supposant connu l'éventuel point de rupture, nous discuterons de ce problème dans le paragraphe suivant. Ici supposons que l'on se demande si le point 1991:1 (premier trimestre de 1991) n'est pas un point de rupture.
On va donc comme dans la théorie partager l'échantillon en deux parties, avant le point de rupture , puis après le point de rupture celui-ci étant toujours dans la seconde partie.
Modèle avant le point de rupture:
smpl start 1990:4
lin Y2 / residus2
 # constant X1 X2

```
Linear Regression - Estimation by Least Squares
Dependent Variable Y2
Quarterly Data From 1971:01 To 1990:04
Usable Observations     80      Degrees of Freedom    77
Centered R**2     0.998358      R Bar **2   0.998316
Uncentered R**2   0.999933      T x R**2      79.995
Mean of Dependent Variable      539018.24545
Std Error of Dependent Variable 111798.30983
Standard Error of Estimate        4588.00835
Sum of Squared Residuals        1620836186.8
Regression F(2,77)                23415.6216
Significance Level of F           0.00000000
```

```
Log Likelihood                    -786.48234
Durbin-Watson Statistic             2.053834

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                 -122.760136  2546.009743     -0.04822  0.96166840
2.  X1                          0.847217     0.099387      8.52445  0.00000000
3.  X2                          0.497543     0.009204     54.05819  0.00000000
```

smpl 1991:1 end
lin Y2 / residus3
 # constant X1 X2

```
Linear Regression - Estimation by Least Squares
Dependent Variable Y2
Quarterly Data From 1991:01 To 2005:04
Usable Observations     60      Degrees of Freedom    57
Centered R**2     0.997175      R Bar **2   0.997076
Uncentered R**2   0.999971      T x R**2      59.998
Mean of Dependent Variable      885786.79976
Std Error of Dependent Variable  91394.35165
Standard Error of Estimate        4942.11155
Sum of Squared Residuals        1392194595.9
Regression F(2,57)                10060.2101
Significance Level of F           0.00000000
Log Likelihood                    -593.93039
Durbin-Watson Statistic             2.642511

   Variable                     Coeff       Std Error      T-Stat     Signif
*******************************************************************************
1.  Constant                 239.2926208 6286.0092146      0.03807  0.96976688
2.  X1                         0.5779174    0.0628367      9.19713  0.00000000
3.  X2                         0.5107492    0.0063231     80.77544  0.00000000
```

On remarque que la somme des deux parties redonne bien l'échantillon global (80+60=140), aucun point ne doit être éliminé pour faire la comparaison des deux sous-parties avec le modèle global.

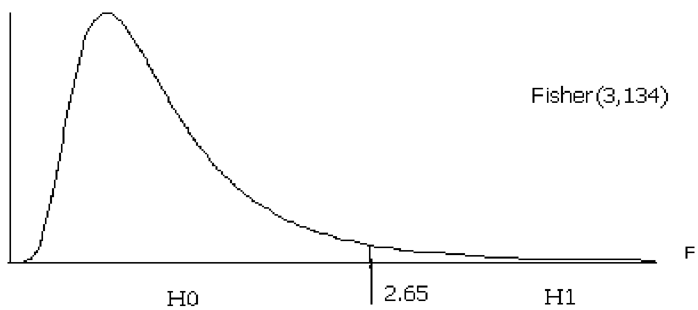
2.2 Calcul du test

Sous l'hypothèse H0 de stabilité du modèle SCR0= 4472049468.0 et le nombre de degrés de liberté est n-k=140-3=137 (résultats dans la régression sur tout l'échantillon).
Sous l'hypothèse H1 la somme des carrés des résidus pour tout l'échantillon est la somme de cette somme des carrés des résidus dans la première sous-partie et dans la seconde.
SCRa=SCR1+SCR2=1620836186.8+1392194595.9=3.01303e+09 (valeur très grande en 10 puissance 9 (10^9) et le nombre de degrés de liberté n1-k+n2-k=77+57= 134

$$\frac{SCR_0 - SCRa}{SCRa}.\frac{n-2k}{k} = F_{k,n-2k} = F_{3,134}$$

$$\frac{4472049468.0 - (1620836186.8 + 1392194595.9)}{(1620836186.8 + 1392194595.9)}.\frac{134}{3} = 21.629$$

La borne du Fisher $F_{3,134}$ est d'environ 2.65 nettement inférieure à 21.629, on décide donc H1 le modèle est instable.



3 Sous-programme chow.src

Dans cet exemple on a choisi la valeur de l'éventuel point de rupture, mais en pratique, comment trouver ces éventuels points de rupture? Souvent une raison économique peut donner la solution, changement dans l'entreprise...... Il existe aussi une méthode plus brutale, elle consiste à faire un balayage sur tout l'échantillon et de voir les points qui donnent un mauvais Fisher, le plus souvent on va trouver toute une période qui donne un mauvais ou plusieurs périodes qui donnent un mauvais Fisher, dans ces deux derniers cas on prendra pour point de rupture le point qui donne le Fisher le plus grand.
Deux cas donc:

- Tous les points de l'échantillon donnent un Fisher inférieur à la borne f_α: on dira alors que le modèle est stable, il n'y a pas de point de rupture.

- Certains points donnent des Fisher supérieurs à la borne f_α: on aura alors au moins un point de rupture. On prendra comme premier point de rupture le point qui a le Fisher le plus grand (ce Fisher sera bien sur supérieur à f_α). Dans ce cas on partage l'échantillon en deux parties avant le point de rupture et la partie suivante et on recommence dans chacune de ces parties le test de Chow pour savoir s'il y a plusieurs points de rupture.

  REMARQUE: si vous trouvez un grand nombre de points de rupture, c'est que l'on se trouve dans le cas d'un modèle à coefficients variables, chaque coefficient théorique bouge avec le temps, comme on l'a vu au début de ce chapitre. C'est peut-être parce que vous avez oublié une variable qui tenait compte de ces variations.

3.1 Utilisation de chow.src dans l'exemple

On continue l'exemple précédent avec la recherche du point du rupture par chow.src. Le programme fait donc le test de Chow pour tous les points de l'échantillon, en fait il ne peut le faire pour tous car si on fait les MCO sur les deux sous-parties il faut par définition même des MCO que la taille de l'échantillon soit supérieure à k; donc il faut toujours avoir $n1>k$ et $n2>k$ ce qui élimine le début et la fin de l'échantillon, ce qui conduit à la remarque de bon sens que si une rupture se produit au début ou en fin d'échantillon, on n'a pas assez d'éléments pour le démontrer.
Reprise de l'exemple précédent:

source chow.src
@chow y2 start end
 # constant X1 X2
Je donne une sous-partie des résultats pour ne pas augmenter inutilement le texte

```
**************************************************************
sous-periodes utilisees 1971:01 1972:01 1972:02 2005:04
F(3,134)=      0.34097 with Significance Level 0.79572961
**************************************************************
sous-periodes utilisees 1971:01 1972:02 1972:03 2005:04
F(3,134)=      1.04938 with Significance Level 0.37300429
**************************************************************
sous-periodes utilisees 1971:01 1972:03 1972:04 2005:04
F(3,134)=      0.60277 with Significance Level 0.61430538
**************************************************************
sous-periodes utilisees 1971:01 1972:04 1973:01 2005:04
F(3,134)=      0.66628 with Significance Level 0.57418128
**************************************************************
sous-periodes utilisees 1971:01 1973:01 1973:02 2005:04
F(3,134)=      0.50174 with Significance Level 0.68172095
**************************************************************
sous-periodes utilisees 1971:01 1973:02 1973:03 2005:04
F(3,134)=      0.89413 with Significance Level 0.44607672
**************************************************************
............

sous-periodes utilisees 1971:01 1975:02 1975:03 2005:04
F(3,134)=      1.78989 with Significance Level 0.15211901
**************************************************************
sous-periodes utilisees 1971:01 1975:03 1975:04 2005:04
F(3,134)=      2.21767 with Significance Level 0.08898739
**************************************************************
sous-periodes utilisees 1971:01 1975:04 1976:01 2005:04
F(3,134)=      2.11681 with Significance Level 0.10104945
**************************************************************
sous-periodes utilisees 1971:01 1976:01 1976:02 2005:04
F(3,134)=      3.78447 with Significance Level 0.01206446
**************************************************************
sous-periodes utilisees 1971:01 1976:02 1976:03 2005:04
F(3,134)=      3.26005 with Significance Level 0.02360294
**************************************************************
sous-periodes utilisees 1971:01 1976:03 1976:04 2005:04
F(3,134)=      3.00721 with Significance Level 0.03261055
**************************************************************
sous-periodes utilisees 1971:01 1976:04 1977:01 2005:04
F(3,134)=      3.86107 with Significance Level 0.01093825
**************************************************************
sous-periodes utilisees 1971:01 1977:01 1977:02 2005:04
F(3,134)=      4.13422 with Significance Level 0.00771364
**************************************************************
sous-periodes utilisees 1971:01 1977:02 1977:03 2005:04
F(3,134)=      4.51965 with Significance Level 0.00471586
**************************************************************
sous-periodes utilisees 1971:01 1977:03 1977:04 2005:04
F(3,134)=      4.30404 with Significance Level 0.00620936
**************************************************************
sous-periodes utilisees 1971:01 1977:04 1978:01 2005:04
F(3,134)=      3.84765 with Significance Level 0.01112757
**************************************************************
sous-periodes utilisees 1971:01 1978:01 1978:02 2005:04
F(3,134)=      4.30866 with Significance Level 0.00617281
**************************************************************
sous-periodes utilisees 1971:01 1978:02 1978:03 2005:04
F(3,134)=      3.86440 with Significance Level 0.01089170
**************************************************************
sous-periodes utilisees 1971:01 1978:03 1978:04 2005:04
F(3,134)=      3.58652 with Significance Level 0.01554279
```

```
**************************************************************
sous-periodes utilisees 1971:01 1978:04 1979:01 2005:04
F(3,134)=      3.62049 with Significance Level 0.01488139
**************************************************************
sous-periodes utilisees 1971:01 1979:01 1979:02 2005:04
F(3,134)=      3.47423 with Significance Level 0.01794501
**************************************************************
sous-periodes utilisees 1971:01 1979:02 1979:03 2005:04
F(3,134)=      3.47813 with Significance Level 0.01785549
**************************************************************
sous-periodes utilisees 1971:01 1979:03 1979:04 2005:04
F(3,134)=      3.51260 with Significance Level 0.01708500
**************************************************************

..........

**************************************************************
sous-periodes utilisees 1971:01 1989:04 1990:01 2005:04
F(3,134)=     11.62558 with Significance Level 0.00000081
**************************************************************
sous-periodes utilisees 1971:01 1990:01 1990:02 2005:04
F(3,134)=     14.28644 with Significance Level 0.00000004
**************************************************************
sous-periodes utilisees 1971:01 1990:02 1990:03 2005:04
F(3,134)=     15.43224 with Significance Level 0.00000001
**************************************************************
sous-periodes utilisees 1971:01 1990:03 1990:04 2005:04
F(3,134)=     20.44297 with Significance Level 0.00000000
**************************************************************
sous-periodes utilisees 1971:01 1990:04 1991:01 2005:04
F(3,134)=     21.62922 with Significance Level 0.00000000
**************************************************************
sous-periodes utilisees 1971:01 1991:01 1991:02 2005:04
F(3,134)=     20.94938 with Significance Level 0.00000000
**************************************************************
sous-periodes utilisees 1971:01 1991:02 1991:03 2005:04
F(3,134)=     17.83919 with Significance Level 0.00000000
**************************************************************
sous-periodes utilisees 1971:01 1991:03 1991:04 2005:04
F(3,134)=     17.12399 with Significance Level 0.00000000
**************************************************************
sous-periodes utilisees 1971:01 1991:04 1992:01 2005:04
F(3,134)=     19.83148 with Significance Level 0.00000000
**************************************************************
sous-periodes utilisees 1971:01 1992:01 1992:02 2005:04
F(3,134)=     15.74323 with Significance Level 0.00000001
**************************************************************
sous-periodes utilisees 1971:01 1992:02 1992:03 2005:04
F(3,134)=     14.27687 with Significance Level 0.00000004
**************************************************************
sous-periodes utilisees 1971:01 1992:03 1992:04 2005:04
F(3,134)=     12.06996 with Significance Level 0.00000048
**************************************************************
sous-periodes utilisees 1971:01 1992:04 1993:01 2005:04
F(3,134)=     10.95095 with Significance Level 0.00000177
**************************************************************
sous-periodes utilisees 1971:01 1993:01 1993:02 2005:04
F(3,134)=      8.83261 with Significance Level 0.00002196
**************************************************************
sous-periodes utilisees 1971:01 1993:02 1993:03 2005:04
F(3,134)=     10.29838 with Significance Level 0.00000381
**************************************************************
sous-periodes utilisees 1971:01 1993:03 1993:04 2005:04
F(3,134)=      9.59766 with Significance Level 0.00000876
**************************************************************
sous-periodes utilisees 1971:01 1993:04 1994:01 2005:04
F(3,134)=      9.58888 with Significance Level 0.00000885
**************************************************************
sous-periodes utilisees 1971:01 1994:01 1994:02 2005:04
F(3,134)=      9.59423 with Significance Level 0.00000879
```

```
*************************************************************
sous-periodes utilisees 1971:01 1994:02 1994:03 2005:04
F(3,134)=      9.04020 with Significance Level 0.00001710
*************************************************************
sous-periodes utilisees 1971:01 1994:03 1994:04 2005:04
F(3,134)=      9.59275 with Significance Level 0.00000881
*************************************************************
sous-periodes utilisees 1971:01 1994:04 1995:01 2005:04
F(3,134)=      9.25666 with Significance Level 0.00001318
*************************************************************
sous-periodes utilisees 1971:01 1995:01 1995:02 2005:04
F(3,134)=      8.99087 with Significance Level 0.00001814
*************************************************************
sous-periodes utilisees 1971:01 1995:02 1995:03 2005:04
F(3,134)=      8.99962 with Significance Level 0.00001795
*************************************************************
sous-periodes utilisees 1971:01 1995:03 1995:04 2005:04
F(3,134)=      9.02149 with Significance Level 0.00001749
*************************************************************
sous-periodes utilisees 1971:01 1995:04 1996:01 2005:04
F(3,134)=      8.99689 with Significance Level 0.00001801
*************************************************************
sous-periodes utilisees 1971:01 1996:01 1996:02 2005:04
F(3,134)=      8.95724 with Significance Level 0.00001889
*************************************************************
sous-periodes utilisees 1971:01 1996:02 1996:03 2005:04
F(3,134)=      9.04913 with Significance Level 0.00001691
*************************************************************
sous-periodes utilisees 1971:01 1996:03 1996:04 2005:04
F(3,134)=      9.19792 with Significance Level 0.00001414
*************************************************************
sous-periodes utilisees 1971:01 1996:04 1997:01 2005:04
F(3,134)=      8.83592 with Significance Level 0.00002188
*************************************************************
sous-periodes utilisees 1971:01 1997:01 1997:02 2005:04
F(3,134)=      8.87823 with Significance Level 0.00002079
*************************************************************
sous-periodes utilisees 1971:01 1997:02 1997:03 2005:04
F(3,134)=      7.77899 with Significance Level 0.00007928
*************************************************************
sous-periodes utilisees 1971:01 1997:03 1997:04 2005:04
F(3,134)=      8.95216 with Significance Level 0.00001901
*************************************************************
sous-periodes utilisees 1971:01 1997:04 1998:01 2005:04
F(3,134)=      7.97079 with Significance Level 0.00006267
*************************************************************
sous-periodes utilisees 1971:01 1998:01 1998:02 2005:04
F(3,134)=      8.28243 with Significance Level 0.00004283
*************************************************************
sous-periodes utilisees 1971:01 1998:02 1998:03 2005:04
F(3,134)=      7.73803 with Significance Level 0.00008337
*************************************************************
sous-periodes utilisees 1971:01 1998:03 1998:04 2005:04
F(3,134)=      7.85180 with Significance Level 0.00007251
*************************************************************
sous-periodes utilisees 1971:01 1998:04 1999:01 2005:04
F(3,134)=      5.46066 with Significance Level 0.00142726
*************************************************************
sous-periodes utilisees 1971:01 1999:01 1999:02 2005:04
F(3,134)=      4.23759 with Significance Level 0.00675924
*************************************************************
sous-periodes utilisees 1971:01 1999:02 1999:03 2005:04
F(3,134)=      3.35727 with Significance Level 0.02084215
*************************************************************
sous-periodes utilisees 1971:01 1999:03 1999:04 2005:04
F(3,134)=      2.75899 with Significance Level 0.04476507
*************************************************************
sous-periodes utilisees 1971:01 1999:04 2000:01 2005:04
F(3,134)=      2.95843 with Significance Level 0.03470705
*************************************************************
```

```
sous-periodes utilisees 1971:01 2000:01 2000:02 2005:04
F(3,134)=      3.29618 with Significance Level 0.02253679
*************************************************************
sous-periodes utilisees 1971:01 2000:02 2000:03 2005:04
F(3,134)=      3.38060 with Significance Level 0.02022932
*************************************************************
sous-periodes utilisees 1971:01 2000:03 2000:04 2005:04
F(3,134)=      3.45775 with Significance Level 0.01832742
*************************************************************
sous-periodes utilisees 1971:01 2000:04 2001:01 2005:04
F(3,134)=      2.18464 with Significance Level 0.09277351
*************************************************************
sous-periodes utilisees 1971:01 2001:01 2001:02 2005:04
F(3,134)=      2.22673 with Significance Level 0.08797596
*************************************************************
sous-periodes utilisees 1971:01 2001:02 2001:03 2005:04
F(3,134)=      1.68485 with Significance Level 0.17327023
*************************************************************
sous-periodes utilisees 1971:01 2001:03 2001:04 2005:04
F(3,134)=      1.86008 with Significance Level 0.13939121
*************************************************************

..........

*************************
 LE MODELE EST INSTABLE
 le plus grand fisher f= 21.629 pour un niveau de significativit\U{e9} ns= 0.000000
 le point  correspondant \U{e0} ce plus grand fisher est 1991:01
*************************
```

Comme vous le voyez je donne la réponse en fin de résultats, si on reprend tous les Fisher, on constate qu'ils sont mauvais (ns>0.05) à partir des sous-périodes 1971:1 , 1976:1 ; 1976:2 2005:4 jusqu'à 1971:1, 2000:4 ; 2001:1 2005:4 . Le plus grand Fisher est de 21.629 il correspond au point de rupture 1991:1 premier trimestre de 1991.

3.2 Y a-t-il d'autres points de rupture ?

Pour le savoir il faut reprendre chacun des sous-échantillons et leur appliquer le test de CHOW. Dans cet exemple on reprend :

- modèle de 1971:1 à 1990:4

  smpl start 1990:4

  lin Y2 / residus2

  # constant X1 X2

  @chow y2 start 1990:4

  # constant X1 X2

  ****

  LE MODELE EST INSTABLE

  le plus grand fisher f= 2.760 pour un niveau de significativité ns= 0.048113

  le point correspondant à ce plus grand Fisher est 1983:04

  ****

  On constate dans cette première partie un niveau de significativité mauvais mais très proche de 5%, on peut donc penser qu'il y a une rupture mais beaucoup plus faible que celle constatée en 1991:1; en général on néglige ces faibles ruptures.

- modèle de 1991:1 à 2005:4

  smpl 1991:1 end

  lin Y2 / residus3

  # constant X1 X2

  @chow y2 1991:1 end

  # constant X1 X2

  ****

  le modèle est stable

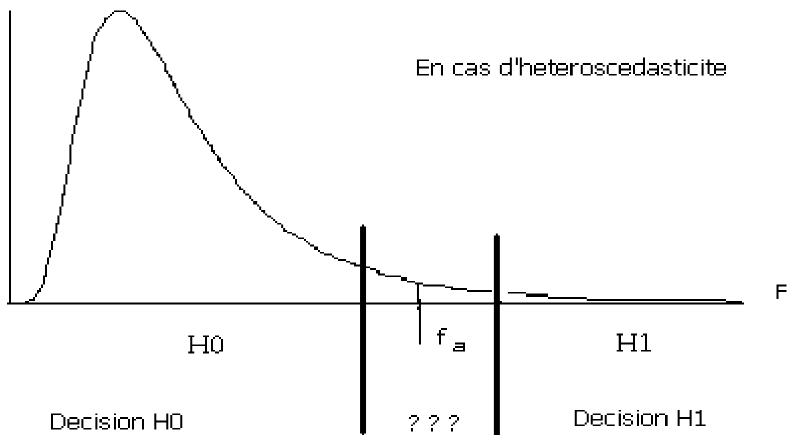
  ****

4 Que faire quand un modèle est instable ?

Si dans un modèle on a trouvé un ou plusieurs points de rupture, l'estimation des coefficients ne peut plus se faire sur le modèle global, les coefficients théoriques n'étant plus les mêmes sur toute la période. La solution est donc le partage de l'échantillon en deux (ou plus) parties. En particulier si on veut utiliser ce modèle en prévision, seule la dernière sous-partie sera utile. Cela pose souvent des problèmes car il est évident que sauf miracle, les deux parties ne sont pas égales et l'une d'entre elles peut se révéler trop petite pour avoir une estimation correcte des coefficients (n'oublions pas qu'une bonne estimation demande une taille d'échantillon importante).
Parfois un ou plusieurs points de rupture sont le signe d'un modèle pas assez précis, c'est-à-dire d'un manque de variables explicatives. On constatera souvent que si on trouve un modèle statique (toutes les variables en fonction de t seulement) instable, le modèle dynamique (des variables endogène et exogènes fonction de t mais aussi de t-1 ....) correspondant sera lui stable. En conséquence si vous trouvez un modèle statique instable, pour la construction du modèle dynamique il faut reprendre l'étude sur l'ensemble de l'échantillon, quitte à redécouper s'il est lui même instable. (Voir des exemples dans les chapitres qui seront faits sur les modèles dynamiques, dans tous ces premiers chapitres les modèles sont supposés statiques).Une remarque ici cependant sur les modèles dynamiques instables: si vous trouvez un modèle dynamique instable il faut reprendre le choix des retards sur les variables dans les deux sous-parties car les retards n'ont aucune raison d'être identiques.

5 Remarque importante sur l' hétéroscédasticité

On a indiqué au départ que comme dans tous ces chapitres de base de l'économétrie on supposait que les erreurs avaient les bonnes propriétés: Normalité, non autocorrélation et homoscédasticité. Or en particulier pour cette dernière propriété il se passe parfois le problème suivant: une instabilité entraîne l'hétéroscédasticité du modèle global alors que les deux sous-modèles sont bien homoscédastiques. On se trouve alors devant un problème insoluble, on ne peut faire les tests de Chow car il y a hétéroscédasticité, mais cette hétéroscédasticité est due à la rupture. Il faut donc trouver un moyen de résoudre ce problème. On le trouve dans un article de TOYODA qui à l'aide de simulations a montré lorsque qu'il y a hétéroscédasticité dans le modèle global, que lorsque la valeur du Fisher est assez loin de fα vers la gauche on décide bien H0 le modèle est stable et que lorsque le Fisher calculé est très à droite de fα on décide sans remord H1 le modèle est instable, mais si le Fisher calculé est proche de fα là le test de Chow ne peut rien décider. Mais ne nous faisons pas trop de soucis car le plus souvent le Fisher calculé est très loin de fα.



En gros si vous trouvez pour le Fisher $F_{k,n-2k}$ un niveau de significativité nettement inférieur à 5% , soit inférieur à environ 3% vous décidez H1 si le Fisher est nettement supérieur à 5%, soit supérieur à 15% ou 20% vous décidez H0 , mais entre les deux le test n'est pas applicable. Deux solutions alors:

- soit en décide de prendre la décision la plus problématique et on prend la décision H1
- soit on essaie de transformer ce modèle pour atténuer l'hétéroscédasticité.

Nous verrons dans le chapitre suivant comment tester l'homoscédasticité. A titre d'exemple vous pourrez alors tester l'hétéroscédasticité dans le modèle global et les sous-échantillons.