



Université d'Orléans

MASTER ECONOMETRIE ET STATISTIQUE APPLIQUEE (ESA)

Université d'Orléans

Site Value-at-Risk

http://193.49.79.89/esa_prof/index.php

Sous la Direction de

Christophe Hurlin

Année Universitaire 2006-2007

Master Econométrie et Statistique Appliquée (ESA)
Université d'Orléans
Faculté de Droit, d'Economie et de Gestion
Bureau A 224
Rue de Blois – BP 6739
45067 Orléans Cedex 2
www.univ-orleans.fr/deg/masters/ESA/

Liste des méthodes - Windows Internet Explorer

http://193.49.79.89/esa_prof/methodes.php?PHPSESSID=fc4dadb72216221bfbbf85da1b361553

Fichier Edition Affichage Favoris Outils ?

Liste des méthodes

Value-at-Risk
Prévisions et Tests de Validation

UNIVERSITÉ
ESA
Université d'Orléans

ACCUEIL->Méthodes Paramétriques/ Non Paramétriques:: Janvier 10, 2007

OBJECTIF : A partir de cette page vous pourrez **executer en ligne différents programmes SAS** permettant de réaliser des prévisions de **Value-at Risk (VaR)** associée à la détention d'un actif financier à partir de la série historique de ses rendements passés.
[Plus de détails](#)

[Présentation Générale](#) [Graphique des Rendements](#) [Statistiques Descriptives](#)

METHODES PARAMETRIQUES DE CALCUL DE LA VaR

<p>Modèle GARCH</p> <p>Distribution Conditionnelle Loi Normale</p> <p> Loi de Student</p>  <p>Loi GED</p> <p>Loi Skewed Student</p>	<p>Modèle EGARCH</p> <p>Distribution Conditionnelle Loi Normale</p> <p> Loi de Student</p>  <p>Loi GED</p> <p>Loi Skewed Student</p>	<p>Modèle QGARCH</p> <p>Distribution Conditionnelle Loi Normale</p> <p> Loi de Student</p>  <p>Loi GED</p> <p>Loi Skewed Student</p>
<p>Modèle IGARCH</p> <p>Distribution Conditionnelle Loi Normale</p> <p> Loi de Student</p> 	<p>Modèle GJR-GARCH</p> <p>Distribution Conditionnelle Loi Normale</p> <p> Loi de Student</p> 	<p>Modèle GARCH-M</p> <p>Distribution Conditionnelle Loi Normale</p> <p> Loi de Student</p> 

Fichiers

Démonstrations
Nasdaq [Aide](#)

Fiches

1. [Définition de la VaR](#)
2. [Distribution de Pertes et Profits](#)
3. [Prévisions de la VaR et Modèles GARCH](#)
4. [Violations de la VaR et Expected Shortfall](#)
5. [Méthodes de Calcul de la VaR](#)
6. [Modèle GARCH](#)
7. [Modèle EGARCH](#)
8. [Modèle QGARCH](#)
9. [Modèle IGARCH](#)
10. [Modèle GJR-GARCH](#)
11. [Modèle GARCH-M](#)
12. [Distributions](#)

Démarrer Eudora - [In] C:\Chris\Cours\Projet_V... Liste des méthodes - ... Document1 - Microsoft ... 06:56

Master Econométrie et Statistique Appliquée

Université d'Orléans

Site Value-at-Risk

http://193.49.79.89/esa_prof/index.php



Fiche 1. Définition de la Value-at-Risk

La notion de Value-at-Risk (**VaR**) est apparue pour la première fois dans le secteur de l'assurance. A la fin des années 1980, la banque Bankers Trust fut l'une des premières institutions à utiliser cette notion sur les marchés financiers aux Etats-Unis, mais c'est principalement la banque JP Morgan qui dans les années 90 a popularisée ce concept notamment grâce à son système RiskMetrics (pour un historique complet de la notion de Value-at-Risk et de sa diffusion se reporter au livre de [Dowd, 2005](#)). La Value-at-Risk est ensuite devenue, en moins d'une dizaine d'années, une mesure de référence du risque sur les marchés financiers, consacrée notamment par la réglementation prudentielle définie dans le cadre des accords de Bâle II.

De façon générale, la Value-at-Risk est définie comme la perte maximale potentielle qui ne devrait être atteinte qu'avec une probabilité donnée sur un horizon temporel donné ([Engle et Manganelli, 2001](#)). La Value at Risk est donc la pire perte attendue sur un horizon de temps donné pour un niveau de confiance donné. Cette définition très simple constitue l'un des principaux attraits de la Value-at-Risk : il est en effet très facile de communiquer sur la VaR et de ainsi proposer une mesure homogène et générale (quelque soit la nature de l'actif, la composition du portefeuille etc.) de l'exposition au risque.

Ainsi, la Value-at-Risk n'est rien d'autre qu'un fractile de la distribution de perte et profit associée à la détention d'un actif ou d'un portefeuille d'actifs sur une période donnée. La mesure de Value-at-Risk ne fait que refléter l'information contenue dans la queue gauche (associée aux pertes) de la distribution des rendements d'un actif. Si l'on considère un taux de couverture de α % (ou de façon équivalente un niveau de confiance de $1-\alpha$ %) la Value-at-Risk correspond

tout simplement au fractile de niveau $\alpha\%$ de la distribution de perte et profit valable sur la période de détention de l'actif :

$$VaR(\alpha) = F^{-1}(\alpha)$$

où $F(\cdot)$ désigne la fonction de répartition associée à la distribution de perte et profit. De cette définition générale découlent plusieurs définitions techniques tout aussi simples.

1. Définitions

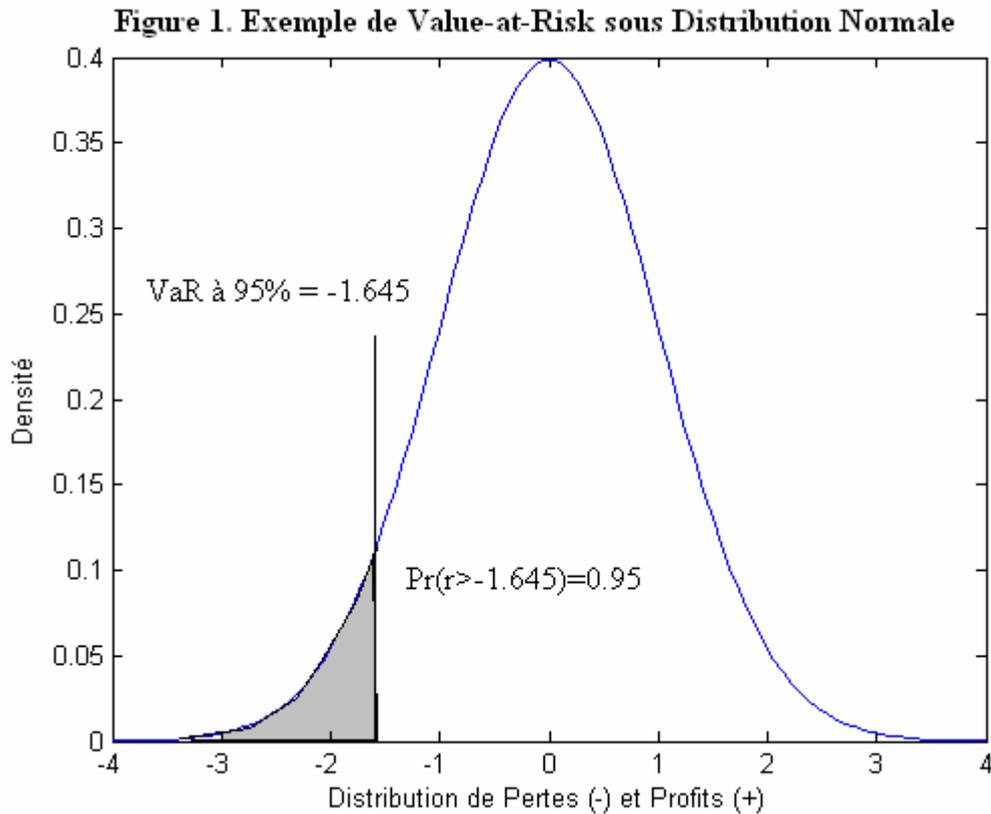
Ainsi, la Value-at-Risk dépend de trois éléments : (i) la [distribution des pertes et profits](#) du portefeuille valable pour la période de détention (ii) le niveau de confiance (ou de façon équivalente le taux de couverture égal à un moins le niveau de confiance) et (iii) la période de détention de l'actif.

1.1. Taux de couverture et Niveau de Confiance

Le niveau de confiance choisi est un paramètre compris entre 0 et 1 (95% ou 99% en général) qui permet de contrôler la probabilité que l'on obtienne un rendement supérieur ou égale à la Value-at-Risk. Supposons que la distribution des pertes et profits associée à la détention d'un actif sur une période corresponde à une distribution normale standard. Sur la Figure 1 est reproduite cette distribution de perte et profit supposée normale : sur la partie gauche de l'axe des abscisses figurent les rendements négatifs (pertes) tandis qu'à droite figure les rendements positifs (profits). Dans ce cas, la Value-at-Risk définie pour un niveau de confiance de 95% ($\alpha = 5\%$) est égale tout simplement à -1.645. Dit autrement, dans cet exemple il y a 95% de chances que le rendement de l'actif, noté r , soit au moins égal à -1.645 sur la période de détention.

$$\Pr[r < VaR(0.05)] = \Pr[r < -1.645] = 0.05$$

De la même façon, la Value-at-Risk définie pour un niveau de confiance de 99% ($\alpha = 1\%$) est égale à -2.326.



Ainsi, la Value-at-Risk correspond généralement à une perte (valeur négative). Toutefois, on trouve souvent une Value-at-Risk définie non pas à partir de la distribution de perte (-) et profit (+), mais à partir au contraire d'une distribution de profit(-) et perte(+). Dit autrement, une telle définition revient à omettre le signe moins devant la perte et donc à afficher une Value-at-Risk positive. Dans ce cas, la définition de la Value-at-Risk correspond à l'opposé du fractile de la distribution de perte et profit :

$$VaR(\alpha) = -F^{-1}(\alpha)$$

Si l'on reprend notre exemple de distribution normale, on affiche alors une Value-at-Risk pour un niveau de confiance de 95% ($\alpha = 5\%$) égale à 1.645. Cela signifie qu'il y a 95% de chances que la perte associée à la détention de l'actif n'excède pas 1.645.

1.2. Horizon de détention

Le deuxième élément fondamental dans le calcul de la Value-at-Risk est la période de détention de l'actif ou du portefeuille d'actifs. La formule de calcul de la Value-at-Risk doit alors être ajustée de façon à tenir compte de la composition des rendements. Il n'existe aucune règle quant au choix de la période de détention dans le calcul de la Value-at-Risk puisque ce choix dépend fondamentalement de l'horizon de reporting ou d'investissement des opérateurs.

Toutefois, les autorités de régulation peuvent spécifier des horizons de détention spécifiques notamment dans le cadre des procédures de validation de la Value-at-Risk.

Dans le cadre de ce site, tous les calculs proposés portent sur une Value-at-Risk définie à partir de la distribution de pertes (-) et profits (+) (ce qui implique une valeur négative de la Value-at-Risk) et pour un horizon de détention d'une période (une journée, un mois etc. suivant les données utilisées pour les calculs).

2. Value-at-Risk conditionnelle

Il est possible de distinguer deux types de distribution de pertes et profits : la distribution conditionnelle et la distribution non conditionnelle. Dès lors, on peut naturellement définir une mesure de Value-at-Risk conditionnelle à un ensemble d'information. Soit R le rendement d'un actif. On suppose que le rendement est une variable aléatoire réelle de densité (distribution de perte et profit) $f_R(r) \forall r \in \mathbb{R}$. Naturellement, pour cette variable aléatoire il est possible de définir une densité conditionnelle à un certain ensemble d'information, noté Ω . Soit $f_R(r|\Omega) \forall r \in \mathbb{R}$ la densité conditionnelle associée au rendement (densité conditionnelle de perte et profit).

La Value-at-Risk conditionnelle à l'ensemble d'information Ω , associée à un taux de couverture de $\alpha\%$, correspond au fractile d'ordre α de la distribution conditionnelle de pertes et profits.

$$VAR(\alpha) = F_R^{-1}(\alpha|\Omega)$$

Cette notion de distribution conditionnelle prend toute son importance dans une dimension temporelle. En effet, jusqu'à présent nous avons considéré les rendements d'un actif ou d'un portefeuille sans indication de date. Or, on peut reprendre le raisonnement en introduisant de façon explicite le temps dans la détermination (et donc par conséquent la [prévision](#)) de la Value-at-Risk.

Soit R_t le rendement à la date t et soit $f_{R_t}(r) \forall r \in \mathbb{R}$ la distribution des pertes et profits pour cette même date. Cette densité peut être différente d'une date à l'autre, et sans là sans doute que réside la difficulté majeure de l'évaluation d'une Value-at-Risk non conditionnelle. De la même façon on peut définir une densité conditionnelle à un ensemble d'information disponible à la date t , noté Ω_t (distribution conditionnelle de perte et profit). Cette densité conditionnelle, notée

$f_{R_t}(r|\Omega_t) \forall r \in \mathbb{R}$, peut elle aussi être différente d'une date à l'autre, mais généralement on se restreint à des densités conditionnelles invariantes dans le temps, i.e. $f_R(r|\Omega_t) \forall r \in \mathbb{R}$. Cela revient à supposer que conditionnellement à un ensemble d'information Ω_t (ou Ω_{t-1} lorsque l'on cherche à prévoir la Value-at-Risk), les rendements sont identiquement distribués. C'est précisément cette hypothèse qui permet de réaliser une [prévision de Value-at-Risk](#) dans le cas des modèles paramétriques ([modèles GARCH](#) par exemple). La Value-at-Risk à la date t obtenue conditionnellement à l'ensemble d'information Ω_t s'écrit sous la forme :

$$VAR_t(\alpha) = F_{R_t}^{-1}(\alpha|\Omega_t)$$

Pour une discussion sur les avantages et les limites de la Value-at-Risk, on renvoie à l'ouvrage de [Dowd \(2005\)](#). Retenons simplement que la Value-at-Risk n'est pas une mesure de risque cohérente au sens de [Artzner et al. \(1997\)](#), parce que notamment la Value-at-Risk n'est pas de façon générale subadditive. Une mesure de risque, notée ρ , est subadditive si et seulement pour deux actifs A et B on a $\rho(A+B) \leq \rho(A) + \rho(B)$. C'est un problème fondamental car cela implique que la Value-at-Risk ne peut pas être considérée comme une mesure « propre » du risque rattachée à une théorie et un ensemble d'axiomes permettant de définir ce qu'est une mesure du risque. La Value-at-Risk n'est rien d'autre qu'un fractile et doit être considérée en tant que telle.

Rédacteurs : [Chevreau Antoine](#), [Godin Sylvain](#), [Ivanof Mihaela](#) et [Patin Antoine](#)

Correction : [Hurlin Christophe](#)

Master ESA, Université d'Orléans : Novembre 2006

Master Économétrie et Statistique Appliquée

Université d'Orléans

Site Value-at-Risk

http://193.49.79.89/esa_prof/index.php



Fiche 2. Distribution de Pertes et Profits

Mesures de Rendements

Les données à partir desquelles on calcule une Value-at-Risk peuvent prendre différentes formes ([Dowd, 2005](#)). Rappelons que la Value-at-Risk se définit comme un fractile de la distribution de pertes et profits associée à la détention d'un actif ou d'un portefeuille d'actifs sur une période donnée. Dès lors, la forme la plus simple consiste à présenter les données en termes de profits et pertes. On note P_t la valeur d'un actif (ou d'un portefeuille) à la fin de la période t et D_t l'ensemble des paiements intermédiaires obtenus entre les dates $t-1$ et t . Les pertes et profits associés à la détention de l'actif (ou du portefeuille) sont alors définis par la différence :

$$P/L = P_t + D_t - P_{t-1}$$

Dans ce cas, une valeur positive correspond à un profit et une valeur négative à une perte. Généralement, ces pertes et profits sont exprimés sous la forme d'un rendement géométrique noté R_t :

$$R_t = \ln\left(\frac{P_t + D_t}{P_{t-1}}\right)$$

Les bases d'exemples fournies sur ce site (Nasdaq, SP500, Nikkei, CAC40) sont constituées de séries historiques de rendements géométriques (pertes (-) et profits (+)).

On suppose que le rendement à la date t , i.e. R_t , est une variable aléatoire réelle. On appelle **fonction de distribution de pertes et profits**, la fonction de densité associée à cette variable aléatoire réelle, notée :

$$f_{R_t}(r) \forall r \in \mathbb{R}$$

Cette densité peut être différente d'une date à l'autre, et sans là sans doute que réside la difficulté majeure de l'évaluation d'une Value-at-Risk (non conditionnelle).

Parallèlement, on peut définir une **distribution de pertes et profits conditionnelle**, c'est-à-dire une fonction de densité conditionnelle à un ensemble d'information disponible à la date t , noté Ω_t . Cette densité conditionnelle est notée :

$$f_{R_t}(r|\Omega_t) \quad \forall r \in \mathbb{R}$$

Cette densité conditionnelle peut elle aussi être différente d'une date à l'autre, mais généralement on se restreint à des densités conditionnelles invariantes dans le temps, i.e. telles que :

$$f_{R_t}(r|\Omega_t) = f_R(r|\Omega_t) \quad \forall r \in \mathbb{R}, \forall t$$

Cela revient à supposer que conditionnellement à un ensemble d'information Ω_t (ou Ω_{t-1} lorsque l'on cherche à prévoir la Value-at-Risk), les rendements sont identiquement distribués. C'est précisément cette hypothèse qui permet de réaliser une [prévision de Value-at-Risk](#) dans le cas des modèles paramétriques ([modèles GARCH](#) par exemple). La Value-at-Risk à la date t obtenue conditionnellement à l'ensemble d'information Ω_t s'écrit alors sous la forme :

$$VAR_t(\alpha) = F_{R_t}^{-1}(\alpha|\Omega_t)$$

Rédacteurs : [Chevreau Antoine](#), [Godin Sylvain](#), [Ivanof Mihaela](#) et [Patin Antoine](#)

Correction : [Hurlin Christophe](#)

Master ESA, Université d'Orléans : Novembre 2006

Master Econométrie et Statistique Appliquée

Université d'Orléans

Site Value-at-Risk

http://193.49.79.89/esa_prof/index.php



Fiche 3. Prévisions de Value-at-Risk et Modèle GARCH

Sur ce site peuvent être exécutés en ligne différents programmes permettant de réaliser des **prévisions out of sample de Value-at-Risk** à l'horizon d'une période.

Une **prévision de Value-at-Risk pour un niveau de confiance de $1 - \alpha\%$ pour la date $t+1$** correspond à simplement au **fractile de niveau $\alpha\%$ de la distribution conditionnelle de pertes et profits** (voir Fiche [Distribution de Pertes et Profits](#)). Formellement on définit la prévision $VAR_{t+1|t}(\alpha)$ sous la forme suivante :

$$VAR_{t+1|t}(\alpha) = F_{R_{t+1}}^{-1}(\alpha | \Omega_t)$$

où $F_{R_{t+1}}(\alpha | \Omega_t)$ désigne la fonction de répartition associée à la fonction de distribution des rendements à la date $t+1$, notés R_{t+1} , conditionnelle à l'ensemble d'information Ω_t disponible à la date t .

Afin d'illustrer cette définition, considérons le cas des prévisions de Value-at-Risk obtenues à partir des modèles paramétriques de type GARCH univariés.

1. Prévisions de Value-at-Risk et Modèles GARCH

Comment obtenir une prévision de Value-at-Risk à partir d'un modèle GARCH ? La démarche est indirecte : dans un premier temps, on fait une hypothèse sur la distribution conditionnelle des rendements de l'actif, puis l'on estime les paramètres du modèle GARCH sur les observations de la période 1 à T, généralement par une procédure de type maximum de vraisemblance. Dans une seconde étape, on déduit du modèle GARCH estimé une prévision de variance conditionnelle, qui couplée à l'hypothèse retenue sur la distribution des rendements,

permet de construire une prévision sur le fractile de la distribution de pertes et profits valable pour $T+1$.

Considérons l'exemple d'un modèle GARCH sous hypothèse de distribution de Student. On suppose ainsi que les rendements d'un actif, notés r_t , satisfont le modèle suivant :

$$\begin{aligned} r_t &= c + \varepsilon_t \\ \varepsilon_t &= z_t \sqrt{h_t} \\ h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \\ z_t & \text{ i.i.d. } Student(v) \end{aligned}$$

où z_t désigne un bruit blanc faible homoscédastique et où les paramètres $\alpha_0, \alpha_1, \beta_1, v, c$ sont des réels vérifiant les contraintes suivantes $\alpha_0 > 0, \alpha_1 \geq 0, \beta_1 \geq 0$ et $v > 2$. Le terme $h_t = E[\varepsilon_t | \varepsilon_{t-1}]$ désigne (à un facteur constant près) la variance conditionnelle du résidu ε_t et donc des rendements r_t . On suppose que l'on dispose de la série historique des rendements $\{r_1, \dots, r_T\}$ observés entre les dates $t=1$ et $t=T$.

Soient $\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_1, \hat{v}, \hat{c}$ des estimateurs convergents (du maximum de vraisemblance par exemple) des paramètres $\alpha_0, \alpha_1, \beta_1, v, c$ (voir fiche [Modèle GARCH](#)). Soit h_1 une condition initiale sur le processus de variance conditionnelle h_t (dans le cadre de nos programmes cette condition initiale est fixée de façon arbitraire au niveau de la variance non conditionnelle). A partir de ces différents éléments, il est possible de prévoir la variance conditionnelle des rendements pour la date $T+1$ de la façon suivante :

$$\hat{h}_{T+1} = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{\varepsilon}_T^2 + \hat{\beta}_1 \hat{h}_T \quad \text{avec } h_1 \text{ donné}$$

soit encore

$$\hat{h}_{T+1} = \hat{\alpha}_0 + \hat{\alpha}_1 (r_T - \hat{c})^2 + \hat{\beta}_1 \hat{h}_T \quad \text{avec } h_1 \text{ donné}$$

On note $VAR_{T+1|T}(\alpha)$ la prévision de la Value-at-Risk de niveau $1-\alpha$ anticipée pour la date $T+1$ conditionnellement à l'information disponible à la date T . Par définition de la Value-at-Risk (cf. [Fiche Définition de la Value-at-Risk](#)), on a :

$$\Pr[r_{T+1} < VAR_{T+1|T}(\alpha) | \Omega_T] = \alpha$$

On en déduit immédiatement que :

$$\Pr \left[z_{T+1} < \frac{VAR_{T+1|T}(\alpha) - c}{\sqrt{h_{T+1}}} \middle| \Omega_T \right] = \alpha$$

En remplaçant le paramètre c par sa valeur estimée et la variance conditionnelle h_{T+1} par sa valeur anticipée \hat{h}_{T+1} , il vient :

$$\Pr \left[z_{T+1} < \frac{VAR_{T+1|T}(\alpha) - \hat{c}}{\sqrt{\hat{h}_{T+1}}} \middle| \Omega_T \right] = \alpha$$

Sachant que l'on suppose que la distribution conditionnelle des rendements est une distribution de Student à ν degrés de libertés de variance égale à $\nu/(\nu-2)\nu$ (cf. [Fiche Distribution Conditionnelle](#)), on peut en déduire immédiatement une prévision du fractile de la distribution conditionnelle valable pour la date $T+1$, c'est-à-dire de la Value-at-Risk.

Soit $G(x; \nu)$ la fonction de répartition de la loi de Student à ν degrés de liberté. La prévision de Value-at-Risk est finalement définie par la quantité suivante :

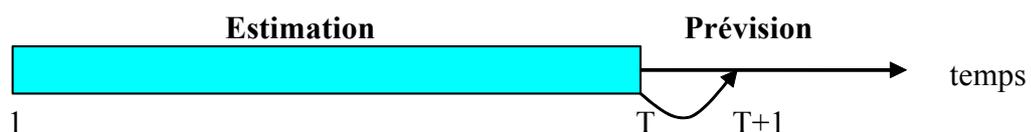
$$VAR_{T+1|T}(\alpha) = \sqrt{\hat{h}_{T+1}} G^{-1}(\alpha; \hat{\nu}) + \hat{c}$$

où $\hat{\nu}$ désigne la valeur estimée du paramètre ν .

2. Prévisions out of sample

La séquence décrite dans la section précédente permet d'obtenir une **prévision out of sample** de Value-at-Risk, dans le sens où la prévision est réalisée pour une période qui se situe à l'extérieur de l'échantillon utilisé pour l'estimation des paramètres du modèle GARCH. En effet, comme l'indique la Figure 1, les paramètres du modèle sont estimés à partir des observations des rendements $\{r_1, \dots, r_T\}$, tandis que la prévision porte sur la Value-at-Risk de la période $T+1$.

Figure 1. Prévision Out-of-Sample



Ainsi, à l'issue de l'exécution de nos programmes, pour chaque modèle GARCH, est reportée la prévision out of sample de la Value-at-Risk pour la date $T+1$ à partir des données historiques de rendements fournies par l'utilisateur pour les date 1 à $T+1$.

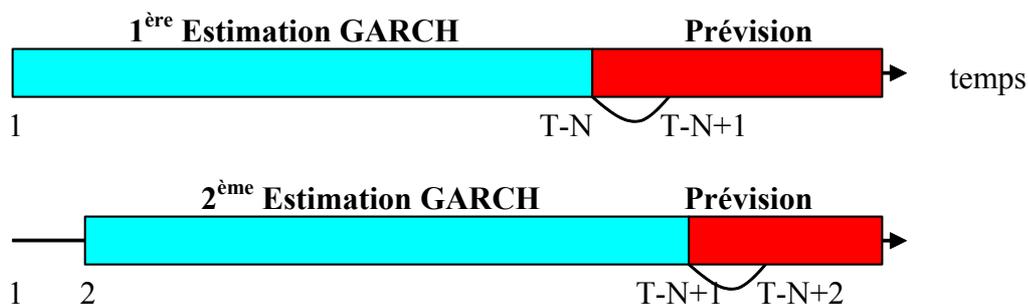
Afin de valider un modèle de calcul et / ou de prévision de Value-at-Risk, il est courant de séparer l'échantillon initial en deux sous échantillons : un échantillon d'estimation et un échantillon de prévision. Cela permet ainsi d'obtenir une séquence de prévisions de Value-at-Risk et de les comparer aux rendements historiques afin notamment de mettre en œuvre des tests de validation (cf. [Fiche Tests de Validation](#)).

Supposons que l'on dispose au total de T observations des rendements et que l'on partage l'échantillon en deux sous échantillons : le second échantillon, dit échantillon de prévision, comprenant N observations. **Dans nos programmes, l'échantillon de prévision correspond aux 25% des observations les plus récentes :**

$$N = \frac{T}{4}N$$

A partir des $T - N$ observations du premier échantillon (échantillon d'estimation), on estime le modèle GARCH et l'on construit une prévision out of sample pour la première observation de l'échantillon de prévision. On obtient ainsi une prévision de la Value-at-Risk pour la date $T-N+1$, notée $VAR_{T-N+1|T-N}(\alpha)$. On peut alors reproduire cette procédure afin d'obtenir une séquence de N prévisions, notées $\{VAR_{t+1|t}(\alpha)\}_{t=T-N}^{T-1}$. Pour cela différentes approches peuvent être utilisées. La plus évidente consiste à ré-estimer le modèles des dates 2 à $T-N+1$ inclus, puis à réaliser une prévision pour la date $T-N+2$ et ainsi de suite (**rolling estimate ou estimation par fenêtre glissante**) comme l'indique la Figure 2.

Figure 2. Prévisions Out-of-Sample et Estimation Glissante



L'inconvénient de cette méthode est qu'elle nécessite d'estimer successivement N modèles GARCH, ce qui peut prendre beaucoup de temps pour des modèles plus compliqués.

C'est pourquoi, afin de limiter le temps de calcul sur nos serveurs, nos applications en ligne permettent de construire un ensemble de prévisions out of sample sans ré-estimer les paramètres des modèles GARCH. On estime une fois pour toute les paramètres sur l'échantillon d'estimation (soient les 75% d'observations des dates 1 à $T - N$) et l'on construit ensuite la séquence des variances conditionnelles sur l'échantillon de prévision à partir desquelles sont déduites les prévisions de Value-at-Risk.

Formellement, dans le cas d'un modèle GARCH sous loi de Student, si l'on note $\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_1, \hat{\nu}, \hat{c}$ les estimateurs des paramètres $\alpha_0, \alpha_1, \beta_1, \nu, c$ obtenus sur l'échantillon d'estimation, on construit alors la séquence des variances conditionnelles $\{\hat{h}_t\}_{t=T-N+1}^T$ comme suit :

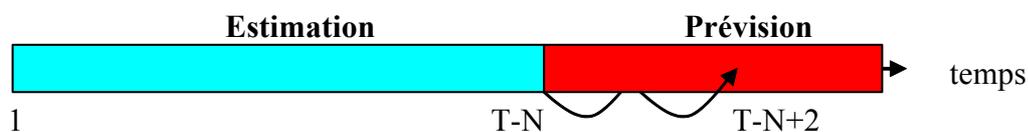
$$\hat{h}_{t+1} = \hat{\alpha}_0 + \hat{\alpha}_1 (r_t - \hat{c})^2 + \hat{\beta}_1 \hat{h}_t \quad t = T - N + 1, \dots, T$$

et l'on en déduit les prévisions successives de Value-at-Risk :

$$VAR_{t+1|t}(\alpha) = \sqrt{\hat{h}_{t+1}} G^{-1}(\alpha; \hat{\nu}) + \hat{c} \quad t = T - N + 1, \dots, T$$

où $G(x; \nu)$ la fonction de répartition de la loi de Student à ν degrés de liberté. Ainsi on obtient, sans ré-estimation du modèle GARCH, une séquence de N prévisions de Value-at-Risk aux date $T - N + 1, \dots, T$ comme l'indique la Figure 3.

Figure 3. Séquence de Prévisions Out-of-Sample



Naturellement une telle démarche suppose que le modèle GARCH estimé sur la période d'estimation soit stable sur la période de prévision. Mais cette hypothèse de stabilité du modèle sur les 25% dernières observations est le prix à payer pour accélérer le traitement informatique de nos applications.

Rédacteurs : [Chevreau Antoine](#), [Godin Sylvain](#), [Ivanof Mihaela](#) et [Patin Antoine](#)

Correction : [Hurlin Christophe](#) Master ESA : Janvier 2007

Master Econométrie et Statistique Appliquée

Université d'Orléans

Site Value-at-Risk

http://193.49.79.89/esa_prof/index.php



Fiche 4. Violations de la Value-at-Risk et Expected Shortfall

A partir de la fin des années 90, plusieurs tests ont été proposés afin d'évaluer la validité des mesures et des prévisions de Value-at-Risk. La plupart de ces tests de validation sont fondés sur les occurrences de violations de la Value-at-Risk.

1. Violations de la VaR

Soit $VaR_{t|t-1}(\alpha)$ la Value-at-Risk pour un taux de couverture de $\alpha\%$ prévue pour la date t conditionnellement à un ensemble d'information disponible à la date $t-1$, noté Ω_{t-1} . Formellement, on note :

$$VaR_{t|t-1}(\alpha) = F_{R_t}^{-1}(\alpha | \Omega_{t-1})$$

où $F_{R_t}(u | \Omega_{t-1})$ désigne la fonction de répartition associée à la distribution conditionnelle de pertes et profits des rendements de l'actif financier (ou portefeuille), notés R_t .

On définit la violation comme une situation dans laquelle on observe *ex-post* une perte plus importante en valeur absolue que la VaR prévue *ex-ante*. Formellement, il y a violation si et seulement si :

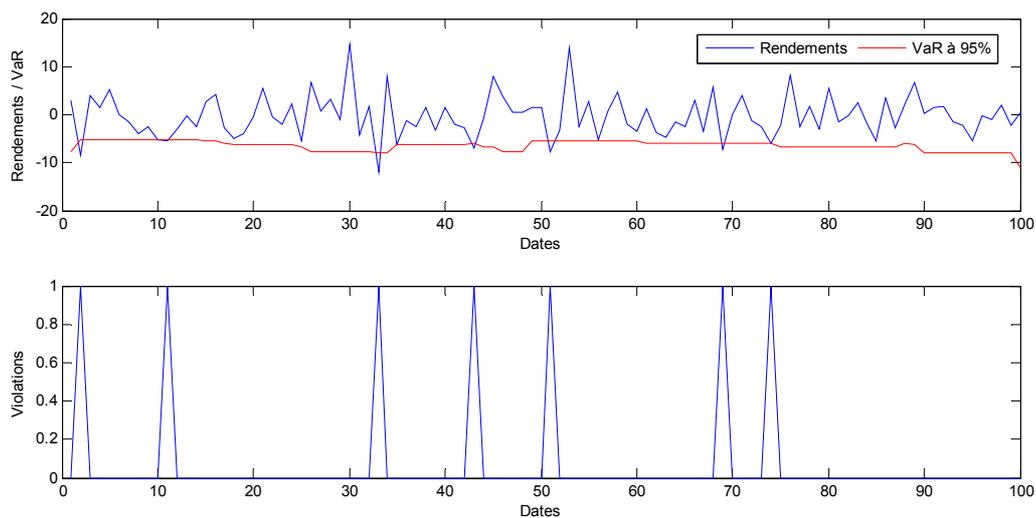
$$R_t < VaR_{t|t-1}(\alpha)$$

Dans la pratique, on définit souvent une variable dichotomique associée à l'occurrence d'une violation, appelée elle aussi violation ou **Hit function** ([Campbell, 2006](#) ; [Engle et Manganelli, 2004](#)) :

$$I_t(\alpha) = \begin{cases} 1 & \text{si } R_t < VaR_{t|t-1}(\alpha) \\ 0 & \text{sinon} \end{cases}$$

La séquence de violations ou Hit (prenant successivement les valeurs 0 ou 1) identifie les périodes pour lesquelles on observe une perte plus importante en valeur absolue que la VaR anticipée, comme le montre l'exemple de la Figure 1. Dans cet exemple, un rendement fictif a été tiré dans une loi normale de moyenne égale à 2% et de variance égale à 4. A partir de ces tirages, on a estimé une Value-at-Risk par une méthode non paramétrique dite [Méthode Hybride](#), en utilisant un paramètre de lissage égal à 0.95 (voir fiche [Méthode Hybride](#)). Comme on peut le constater sur le graphique de la partie haute de la Figure 1, la VaR ainsi prévue n'est pas constante (alors qu'elle devrait théoriquement l'être, la distribution des pertes et profits étant invariante dans le temps). Sur le graphique de la partie basse, sont reportées les valeurs de la fonction de Hit prenant des valeurs 1 ou 0. Lorsque la variable de Hit prend la valeur 1, cela indique une période de violation.

Figure 1. Graphique de Violations (Hit Function)



2. Expected Shortfall

Il est évident qu'en cas de violation, il peut apparaître réducteur de ne conserver au final comme information qu'une variable indicatrice prenant la valeur 1. Naturellement, de nombreuses autres mesures ont développés permettant de prendre en compte l'ampleur des pertes au-delà de la Value-at-Risk obtenues en cas de violations. La première mesure évidente est la **TCE (Tail Conditional Expectation)** définie par l'espérance conditionnelle de la perte en cas de violation :

$$TCE_{\alpha}(R) = E[R | R < VaR(\alpha)]$$

Une mesure alternative est la mesure dite de **l'Expected Shortfall (ES)**. Cette mesure correspond à la moyenne des pertes extrêmes telle que :

$$ES_{\alpha}(R) = \frac{1}{\alpha} \int_0^{\alpha} F^{-1}(p) dp$$

où $F(\cdot)$ désigne la fonction de répartition associée à la distribution de pertes et profits. Les deux mesures TCE et ES coïncident dans le cas d'une distribution continue. Les deux statistiques donnent finalement une mesure de ce peuvent être les pertes dans les pires états du système financier. Un des avantages de l'ES par rapport à la Value-at-Risk en tant que mesure de risque est qu'il s'agit d'une mesure cohérente de risque au sens de [Artzner et al. \(1997, 1999\)](#). C'est en particulier une mesure qui vérifie l'axiome de subadditivité, contrairement à la Value-at-Risk (voir fiche [Définition de la Value-at-Risk](#)).

Rédacteurs : [Chevreau Antoine](#), [Godin Sylvain](#), [Ivanof Mihaela](#) et [Patin Antoine](#)

Correction : [Hurlin Christophe](#)

Master ESA, Université d'Orléans : Janvier 2007

Master Econométrie et Statistique Appliquée

Université d'Orléans

Site Value-at-Risk

http://193.49.79.89/esa_prof/index.php



Fiche 5. Méthodes de Calcul de la Value-at-Risk

On dénombre trois grandes classes de méthodes de calcul de la VaR :

- Méthodes Non-paramétriques (Historical Simulation, Weighted Historical Simulation, Monte Carlo Simulation, Filtered Historical Simulation...).
- Méthodes Semi-paramétriques (CAViaR, théorie des extrêmes).
- Méthodes Paramétriques Paramétrique (ARCH, GARCH univarié, GARCH multivarié, RiskMetrics).

Cette présentation est inspirée de celle proposée par [Dowd \(2005\)](#) ou [Engle et Manganelli \(2001\)](#).

1. Les Méthodes Non Paramétriques

Le principe général des méthodes non paramétriques d'estimation / prévision de la Value-at-Risk est que l'on impose a priori aucune distribution paramétrique de pertes et profits. Au delà de ce point commun, il existe une grande variété de méthodes non paramétriques de calcul de la Value-at-Risk.

1.1 Historical Simulation (HS)

La simulation historique (*Historical Simulation*, ou HS) est une méthode très simple d'estimation des mesures de risque fondée sur la distribution empirique des données historiques de rendements. Formellement, la VaR est estimée simplement par lecture directe des fractiles empiriques des rendements passés. Si l'on considère par exemple un niveau de confiance de 95% et que l'on dispose d'un échantillon de 1000 observations historiques de rendements, la VaR est donnée par la valeur du rendement qui correspond à la 50ème forte de perte.

1.2. Bootstrapped Historical Simulation

Une amélioration simple de la méthode HS consiste à estimer la VaR à partir de données simulées par Bootstrap. Le Bootstrap consiste à ré-échantillonner les données historiques de rendements avec

remise. Plus précisément, dans notre contexte, la procédure consiste à créer un grand nombre d'échantillons de rendements simulés, où chaque observation est obtenue par tirage au hasard à partir de l'échantillon original. Chaque nouvel échantillon constitué de la sorte permet d'obtenir une estimation de la VaR par la méthode HS standard, et l'on définit au final une estimation en faisant la moyenne de ces estimations basées sur les ré-échantillonnages.

1.3 Simulation Historique et Estimation Non Paramétrique de Densité

Une autre amélioration possible de la basic HS est d'utiliser une estimation non paramétrique de la distribution conditionnelle de pertes et profits. On sait en effet que l'histogramme associé aux réalisations historiques des rendements n'est pas un bon estimateur d'une fonction de densité. Des estimateurs obtenus par lissage, comme les estimateurs à noyau, présentent généralement de meilleures propriétés ([Yatchew, 2001](#)). Dans cette perspective, la méthode HS étendue consiste à estimer par une méthode de noyau la densité conditionnelle de pertes et profits, puis de calculer à partir de cette densité estimée le fractile correspondant à la Value-at-Risk ([Butler et Schachter, 1998](#)). Cela nous permet notamment d'estimer la Value-at-Risk pour n'importe quel niveau de confiance (et ainsi d'éviter les problèmes dus aux contraintes imposées sur la taille des échantillons).

1.4. Weighted Historical Simulation (WHS)

La caractéristique essentielle de la méthode HS traditionnelle est que l'on accorde le même poids aux observations historiques, quelles soient relativement récentes ou au contraire très anciennes. Concrètement, si l'on considère une estimation HS de la Var à 5% à partir d'une fenêtre glissante de 1000 observations, cela revient à prendre le 50^{ème} rendement le plus faible parmi les 1000 observations les plus récentes. Dès lors, dans cette estimation toutes les observations historiques de rendement datées de plus de 1000 périodes n'interviennent pas dans le calcul de la VaR, tout comme toutes les observations de moins de 1000 périodes interviennent avec la même poids dans la construction de l'estimation.

Une approche alternative consiste à attribuer aux observations de rendements des poids en fonction soit de leur ancienneté, de la volatilité observée des marchés, ou de tout autre facteur. Cette approche, qualifiée par le terme générique de WHS (*Weighted Historical Simulation*) recouvre notamment :

- La méthode *Aged-weighted* HS où les poids dépendent de l'ancienneté des observations ([Boudoukh, Richardson et Whitelaw, 1998](#)).
- La méthode *Volatility-weighted* HS où les poids dépendent de la volatilité. L'idée de base, suggérée

par [Hullet et White \(1998\)](#) est de prendre en compte les changement récents de volatilité.

- La méthode *Correlation-weighted* HS où l'on ajuste les rendements passés de façon à ce qu'ils reflètent les changements entre les corrélations passées et futures.

1.5. Filtered Historical Simulation (FHS)

La méthode FHS est une forme de Bootstrap semi-paramétrique qui vise à combiner les avantages de la simulation historique avec la puissance et la flexibilité des modèles à volatilité conditionnelle tel que le modèle GARCH. Elle consiste à faire un Bootstrap sur les rendements dans un cadre de volatilité conditionnelle, le Bootstrap préservant la nature non paramétrique de la simulation historique, et le modèle à volatilité conditionnelle donnant un traitement sophistiqué de la volatilité.

1.6. La méthode de Monte Carlo

La méthode de Monte Carlo consiste à simuler un grand nombre de fois les comportements futurs possibles des facteurs de risque selon un certain nombre d'hypothèses, et d'en déduire une distribution des pertes et profits à partir de laquelle on estime finalement un fractile. Si cette approche peut s'appliquer, en théorie, quelles que soient les lois de probabilité suivies par les facteurs de risque, elle est couramment utilisée en pratique, pour des raisons techniques, en supposant que les variations relatives des paramètres de marché suivent des lois normales. Cette méthode convient également à tous les types d'instruments, y compris optionnels, et permet de tester de nombreux scénarios et d'y inclure explicitement des queues de distribution épaisses (événements extrêmes pris en compte dans une certaine mesure).

2. Les Méthodes Semi-Paramétriques

2.1. Théorie des Valeurs Extrêmes

Parmi les méthodes semi-paramétriques figurent tout d'abord l'ensemble des méthodes et approches qui relèvent de la théorie des extrêmes (EVT) qui diffère de la théorie statistique habituelle fondée pour l'essentiel sur des raisonnements de type "tendance centrale". Les extrêmes sont en effet gouvernés par des théorèmes spécifiques qui permettent d'établir sous différentes hypothèses la distribution suivie par ces extrêmes. Il existe deux principales branches de la théorie des valeurs extrêmes : la *théorie des valeurs extrêmes généralisée* et la *loi de Pareto généralisée* (ou l'approche POT - "*peaks-over-threshold*"). L'approche POT permet l'étude de la distribution des pertes excessives au dessus d'un seuil (élevé), tandis que la théorie des valeurs extrêmes généralisée permet de modéliser le maximum ou le minimum d'un très grand échantillon.

2.2. L'approche par régression sur quantiles – CAViaR

Une seconde grande catégorie de méthodes semi-paramétriques utilisées actuellement pour le calcul et la prévision de la Value-at-Risk relève plus généralement de l'approche de la régression quantile. L'idée est la suivante : plutôt que de modéliser une distribution et d'en déduire un quantile (la Value-at-Risk), cette approche consiste à modéliser directement le quantile lui-même en utilisant des méthodes de régression quantile. Un exemple de ces méthodes est le modèle *Conditional Autoregressive Value at Risk* (CAViaR) de [Engle et Manganelli \(2004\)](#), qui spécifie comment la dynamique autorégressive pour le quantile conditionnel.

3. Les Méthodes Paramétriques

La détermination de la VaR paramétrique se fait au moyen d'un calcul analytique relativement aisé en pratique mais sous des hypothèses théoriques assez contraignantes, l'exemple le plus connu d'un tel modèle étant sans doute *RiskMetrics*. Les principales hypothèses simplificatrices consistent à supposer, d'une part, que les lois de probabilité qui régissent les distributions des variations des prix de marché sont normales et, d'autre part, que les instruments présentent un profil de risque linéaire. Sous ces hypothèses, la matrice de variances/covariances peut être appliquée assez directement aux positions détenues pour calculer la VaR. Les calculs utilisés dans la méthode RiskMetrics sont rapides et simples, et requièrent uniquement la connaissance de la matrice des variances/covariances des rendements du portefeuille. Néanmoins, cette méthode s'avère être inadaptée aux portefeuilles non linéaires (instruments optionnels), et théoriquement peu adaptée aux queues de distribution épaisses et aux distributions non normales des rendements.

Enfin, figurent parmi les méthodes paramétriques l'ensemble des méthodes de calcul et de prévision de la VaR fondées sur des [modèles GARCH](#) univariés ou multivariés (Engle, 2001). Ces modèles permettent de modéliser et de prévoir la variance conditionnelle de la distribution de pertes et profits, ce qui permet dans un second temps d'en déduire une modélisation ou une prévision de la Value-at-Risk sous un certain nombre d'hypothèse sur la distribution conditionnelle des rendements.

Rédacteurs : [Albayrak Adem](#) et [Arnoult Benoit](#)

Master ESA, Université d'Orléans : Janvier 2007

Master Econométrie et Statistique Appliquée

Université d'Orléans

Site Value-at-Risk

http://193.49.79.89/esa_prof/index.php



Fiche 6. Modèle GARCH

Le modèle GARCH (General Autoregressive Conditional Heteroskedasticity) a été introduit par [Bollerslev \(1986\)](#). C'est une extension du modèle ARCH initialement développé par Engle (2002). Ces modèles permettent une représentation autogressive de la variance conditionnelle d'un processus, ce qui permet de les utiliser notamment à des fins de prévisions de la volatilité sur les marchés financiers.

1. Présentation du Modèle GARCH

L'écriture du modèle GARCH porte sur la variance conditionnelle du processus considéré. Soit un processus y_t , d'espérance $E(y_t) = 0$, satisfaisant une représentation de type GARCH (p,q). Ce processus s'écrit sous la forme suivante :

$$\begin{aligned}y_t &= c + \varepsilon_t \\ \varepsilon_t &= z_t \sqrt{h_t} \\ h_t &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i}\end{aligned}$$

où z_t désigne un bruit blanc faible homoscédastique tel que $E(z_t) = 0$ et $Var(z_t) = \sigma_z^2$ et où les paramètres α_i, β_i sont des réels. De façon usuelle, la quantité h_t désigne la variance conditionnelle du processus y_t telle que $V(y_t | \underline{y}_{t-1}) = V(\varepsilon_t | \underline{\varepsilon}_{t-1}) = h_t$ où \underline{y}_{t-1} désigne l'ensemble des valeurs passées $\{y_{t-1}, \dots, y_0\}$. Afin de garantir la positivité de la variance conditionnelle, on suppose que $\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, q, \beta_i \geq 0, i = 1, \dots, p$.

2. Estimation des Paramètres

Les paramètres du modèle GARCH peuvent être estimés selon différentes méthodes : maximum de vraisemblance, pseudo maximum de vraisemblance, méthode des moments etc. (pour plus de détails, voir [Gouriéroux, 1997](#)). Les méthodes généralement retenues sont celles du maximum de vraisemblance (MV) ou du pseudo maximum de vraisemblance (PMV). L'avantage du PMV réside dans le fait que l'estimateur obtenu converge malgré une mauvaise spécification (supposée normale) de la distribution conditionnelle des résidus, à condition que la loi spécifiée appartienne à la famille des lois exponentielles ([Gouriéroux et Montfort, 1989](#)). Ainsi, l'estimateur du MV obtenu sous l'hypothèse de normalité des résidus et l'estimateur du PMV sont identiques, seules leurs lois asymptotiques respectives diffèrent. Toutefois dans les deux cas (MV ou PMV), sous les hypothèses standards, l'estimateur est asymptotiquement convergent et asymptotiquement normal.

Considérons le cas d'un modèle GARCH(1,1) sous l'hypothèse de normalité des innovations.

$$\begin{aligned}y_t &= c + \varepsilon_t \\ \varepsilon_t &= z_t \sqrt{h_t} \\ h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \\ z_t & \text{ i.i.d. } N(0,1)\end{aligned}$$

où les variables z_t sont identiquement et indépendamment distribuées (*i.i.d.*) selon une loi normale centrée et réduite. La fonction de log-vraisemblance associée à un échantillon de T observations $\{y_1, \dots, y_T\}$ obtenue sous l'hypothèse de normalité de la loi conditionnelle de y_t sachant son propre passé s'écrit :

$$\log L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(h_t(\theta)) - \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\theta)]^2}{h_t(\theta)}$$

où θ désigne l'ensemble des paramètres du modèle, $m_t(\theta)$ désigne l'espérance conditionnelle et $h_t(\theta)$ désigne la variance conditionnelle. Dans le cas du modèle GARCH(1,1) présenté ci-dessus :

$$\begin{aligned}m_t(\theta) &= c \\ h_t(\theta) &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}\end{aligned}$$

Les estimateurs du maximum de vraisemblance sont alors obtenus par résolution analytique d'un système de $K = p + q + 2$ (nombre de paramètres à estimer) équations non linéaires :

$$\left. \frac{\partial \log L(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} = 0$$

$$\left. \frac{\partial^2 \log L(\theta)}{\partial \theta^2} \right|_{\theta=\hat{\theta}} < 0$$

Dans le cas général du PMV, l'estimateur du PMV est asymptotiquement convergent et normal ([Gouriéroux, 1997](#)).

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow[T \rightarrow \infty]{d} N(0, J^{-1} I J^{-1})$$

$$J = E_0 \left[- \frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \right]$$

$$I = E_0 \left[\frac{\partial \log L(\theta)}{\partial \theta} \frac{\partial \log L(\theta)}{\partial \theta'} \right]$$

où E_0 désigne l'espérance prise par rapport à la vraie loi. Si la vraie distribution des erreurs est une loi normale (cas du MV) alors $I = J$.

3. Problèmes de convergence et choix des conditions initiales

Il n'existe pas de solution analytique à la maximisation de la vraisemblance dans le cas général des modèles GARCH. On doit donc recourir à une optimisation numérique de la vraisemblance ou de la pseudo vraisemblance suivant les cas. Dans la pratique, se pose alors le problème de la convergence de l'algorithme d'optimisation retenu et plus spécifiquement du problème du choix de conditions initiales.

En ce qui concerne l'algorithme d'optimisation, deux choix sont disponibles dans le cadre de la procédure MODEL sous SAS utilisée pour estimer les modèles GARCH : la méthode de Gauss Newton Raphson et la méthode de Marquardt –Levenberg. C'est cette dernière qui est utilisée dans nos programmes. Pour cet algorithme, nous avons retenu pour conditions initiales sur les paramètres α_i et β_i les valeurs estimées dans le cas d'un modèle GARCH sous hypothèse de normalité.

4. Exemple

Considérons l'exemple d'un modèle GARCH(1,1) estimé à partir de la série des rendements quotidien du Dow Jones sur la période 01 octobre 1996 au 16 novembre 2006. On considère ici une procédure de maximum de vraisemblance sous une hypothèse de distribution conditionnelle normale. On note rdt la série de rendement. Soit le modèle suivant.

$$rdt_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

$$z_t \text{ i.i.d. } N(0,1)$$

La sortie SAS issue de la procédure MODEL retenue ([Hurlin, 2006](#)) débute par un ensemble d'informations fournies sur la structure du modèle.

The Equation to Estimate is	
rdt =	F(intercept(1))
VAR(rdt) =	H(alpha0, alpha1, beta1, intercept)

Il s'agit d'informations portant sur l'écriture du modèle qui rappelle que l'espérance du rendement est une constante et que la variance conditionnelle du rendement, c'est-à-dire le processus h_t , s'écrit comme une fonction des paramètres α_0, α_1 et β_1 .

NOTE: At FIML Iteration 27 CONVERGE=0.001 Criteria Met.

Puis, la sortie indique que l'algorithme d'optimisation numérique a convergé au bout de 27 itérations, selon un critère d'arrêt de la convergence de 0.001. Enfin, différentes informations sont fournies quant à l'ajustement du modèle (voir PROC MODEL).

Final Convergence Criteria	
R	0.000029
PPC(arch1)	0.00031
RPC(arch1)	0.009658
Object	4.12E-6
Trace(S)	0.000124
Gradient norm	2.008953
Log likelihood	8147.033

La log vraisemblance est ici égale à -8147, puisqu'il est important de préciser que SAS minimise l'opposé de la fonction de vraisemblance. Ce tableau fournit en outre parmi un ensemble d'informations, la valeur de la norme du gradient de la log-vraisemblance au point optimal.

Enfin, figurent les résultats à proprement parler de la phase d'estimation des paramètres du modèle GARCH.

Nonlinear FIML Parameter Estimates				
Parameter	Estimate	Approx Std Err	t Value	Approx Pr > t
intercept	-0.0006	0.000170	-3.51	0.0005
arch0	1.553E-6	3.461E-7	4.49	<.0001
arch1	0.083347	0.0104	8.00	<.0001

Nonlinear FIML Parameter Estimates				
Parameter	Estimate	Approx Std Err	t Value	Approx Pr > t
garch1	0.905527	0.0108	83.94	<.0001

Dans ce tableau, figure en première la réalisation des estimateurs du MV de chacun des paramètres. Dans la seconde colonne, figure la réalisation de l'estimateur de l'écart type asymptotique de l'estimateur du MV, à partir de laquelle est construite la valeur de la statistique de Student associée au test du nullité du paramètre considéré (troisième colonne). Enfin, figure dans la dernière colonne la p-value associée à la loi asymptotique de la statistique de Student. Rappelons que pour un risque de première espèce de $\alpha\%$, si la p-value est inférieure au seuil de $\alpha\%$, on est conduit à rejeter l'hypothèse nulle de nullité du coefficient. Le coefficient est donc significativement différent de zéro. On observe que dans le cas de notre exemple, tous les paramètres α_0, α_1 et β_1 sont significativement différents de zéro au seuil de risque de 1%.

A partir de ce modèle, on peut enfin proposer une estimation de la variance conditionnelle, selon l'équation :

$$\hat{h}_t = \hat{\alpha}_0 + \hat{\alpha}_1 \varepsilon_{t-1}^2 + \hat{\beta}_1 \hat{h}_{t-1}$$

ou encore

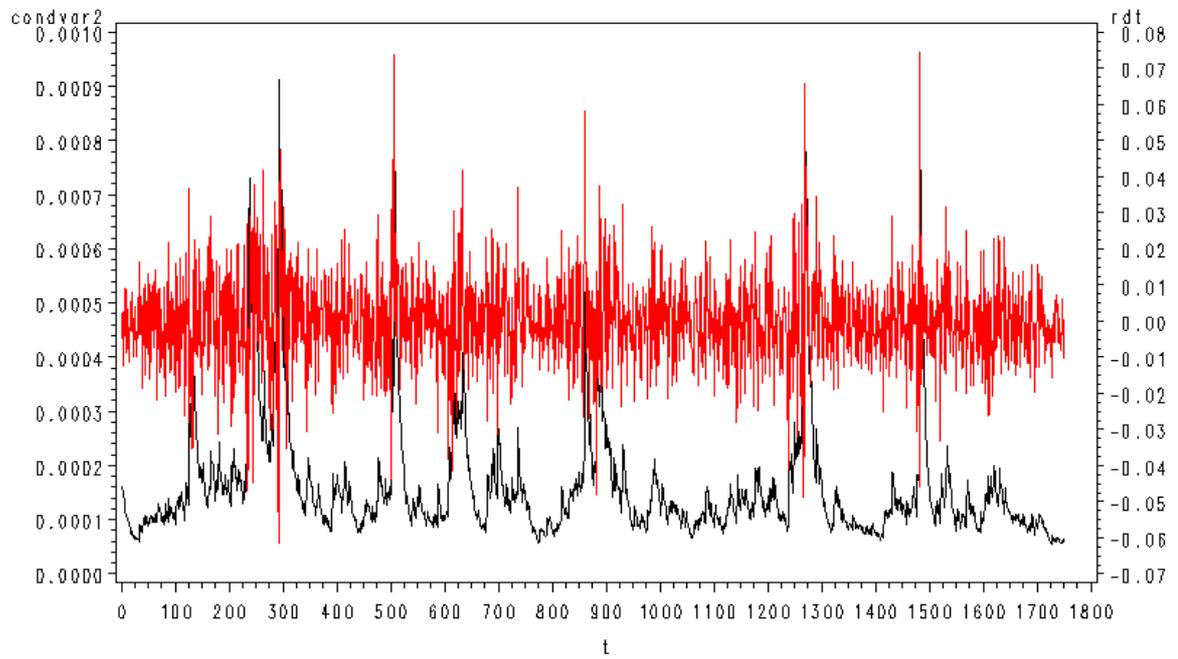
$$\hat{h}_t = \hat{\alpha}_0 + \hat{\alpha}_1 (rdt_{t-1} - \hat{c})^2 + \hat{\beta}_1 \hat{h}_{t-1}$$

avec \hat{h}_1 donné

où $\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_1$ désignent les estimateurs du MV. Les valeurs estimées de la variance conditionnelle obtenues dans notre exemple sont reproduites sur la Figure suivante et rendent bien compte des clusters de volatilité, c'est-à-dire de la succession de périodes de turbulences et de calme sur le marché considéré.

Variance conditionnelle et Rendements estimés

Exemple du GARCH(1,1)



Rédacteurs : [Edwige Balliet](#), [Carole Njoya](#) et [Sidina Medani](#)

Correction : [Hurlin Christophe](#)

[Master ESA](#). 14 Novembre 2006

Master Econométrie et Statistique Appliquée

Université d'Orléans

Site Value-at-Risk

http://193.49.79.89/esa_prof/index.php



Fiche 7. Modèle EGARCH (Exponential GARCH)

Le modèle EGARCH (*Exponential General Autoregressive Conditional Heteroscedasticity*) a été introduit par [Nelson \(1991\)](#). C'est un modèle ARCH non linéaire qui permet de rendre compte de l'asymétrie dans la réponse de la variance conditionnelle à une innovation. Plus précisément, le modèle EGARCH autorise une forme d'asymétrie qui dépend non seulement du signe positif ou négatif de l'innovation, mais aussi de l'amplitude de ce choc. Par ailleurs le modèle EGARCH présente en outre l'avantage, par rapport au modèle GARCH standard, de ne nécessiter aucune restriction de non négativité sur les paramètres afin de garantir la positivité de la variance conditionnelle.

1. Présentation du Modèle EGARCH

L'écriture du modèle EGARCH porte sur le logarithme de la variance conditionnelle du processus considéré. Soit un processus y_t , d'espérance $E(y_t) = c$, satisfaisant une représentation de type EGARCH(p,q). Ce processus s'écrit sous la forme suivante :

$$y_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$\log(h_t) = \alpha_0 + \sum_{i=1}^q \alpha_i g(z_{t-i}) + \sum_{i=1}^p \beta_i \log(h_{t-i})$$

$$g(z_{t-i}) = \theta z_{t-i} + \gamma (|z_{t-i}| - E|z_{t-i}|)$$

où z_t désigne un bruit blanc faible homoscédastique tel que $E(z_t) = 0$ et $Var(z_t) = \sigma_z^2$ et où les paramètres $\alpha_i, \beta_i, \gamma$ et θ sont des réels. De façon usuelle, la quantité h_t désigne la variance

conditionnelle du processus y_t telle que $V(y_t | \underline{y}_{t-1}) = V(\varepsilon_t | \varepsilon_{t-1}) = h_t$ où \underline{y}_{t-1} désigne l'ensemble des valeurs passées $\{y_{t-1}, \dots, y_0\}$.

La valeur de $E|z_{t-i}|$ dépend naturellement de la loi supposée de z_t . Pour les lois usuelles retenues dans nos applications, cette quantité est définie de la façon suivante :

$$\text{Loi Normale: } E|z_t| = \sqrt{2/\pi}$$

$$\text{Loi de Student (v) : } E|z_t| = 2 \frac{\Gamma(v/2)\sqrt{v-2}}{\sqrt{\pi}(v-1)\Gamma(v/2)}$$

$$\text{Loi GED (v) : } E|z_t| = \frac{\Gamma(2/v)}{\sqrt{\Gamma(1/v)\Gamma(3/v)}}$$

où $\Gamma(\cdot)$ désigne la fonction gamma.

Dans le cas du modèle EGARCH, il est inutile d'imposer des contraintes sur les paramètres pour satisfaire la condition de positivité de la variance conditionnelle h_t en raison de son écriture logarithmique. Le paramètre θ permet de modéliser un effet asymétrique lié au signe de l'innovation z_t . Si $\theta > 0$ (respectivement si $\theta < 0$), un choc positif sur la variance conditionnelle à la date t se traduira à la date $t+1$ par une augmentation (respectivement une diminution) de la variance conditionnelle, c'est-à-dire de la volatilité, du processus y_t . Le paramètre γ permet de prendre en compte une asymétrie liée à l'amplitude de l'innovation z_t mesurée par l'écart $|z_t| - E|z_t|$. Si $\gamma = 0$, alors une innovation positive aura le même effet (en valeur absolue) sur la variance conditionnelle qu'une innovation négative. En revanche, si $\gamma > 0$ un choc de forte d'amplitude aura relativement plus d'effet (en valeur absolue) sur la variance conditionnelle qu'un choc de faible ampleur.

2. Estimation des Paramètres

Les paramètres du modèle EGARCH peuvent être estimés selon différentes méthodes : maximum de vraisemblance, pseudo maximum de vraisemblance, méthode des moments etc. (pour

plus de détails, voir [Gouriéroux, 1997](#)). Les méthodes généralement retenues sont celles du maximum de vraisemblance (MV) ou du pseudo maximum de vraisemblance (PMV). L'avantage du PMV réside dans le fait que l'estimateur obtenu converge malgré une mauvaise spécification (supposée normale) de la distribution conditionnelle des résidus, à condition que la loi spécifiée appartienne à la famille des lois exponentielles ([Gouriéroux et Montfort, 1989](#)). Ainsi, l'estimateur du MV obtenu sous l'hypothèse de normalité des résidus et l'estimateur du PMV sont identiques, seules leurs lois asymptotiques respectives diffèrent. Toutefois dans les deux cas (MV ou PMV), sous les hypothèses standards, l'estimateur est asymptotiquement convergent et asymptotiquement normal.

Considérons le cas d'un modèle EGARCH(1,1) sous l'hypothèse de normalité des innovations.

$$\begin{aligned} y_t &= c + \varepsilon_t \\ \varepsilon_t &= z_t \sqrt{h_t} \\ h_t &= \exp(\alpha_0 + \alpha_1 \theta z_{t-1} + \alpha_1 (|z_{t-1}| - \sqrt{2/\pi}) + \beta_1 \log(h_{t-1})) \\ z_t & \text{ i.i.d. } N(0,1) \end{aligned}$$

où les variables z_t sont identiquement et indépendamment distribuées (*i.i.d.*) selon une loi normale centrée et réduite. Le paramètre γ a été normalisé à 1 puisque dans le cas du EGARCH(1,1) il y a 4 régresseurs et 5 paramètres à estimer. La fonction de log-vraisemblance associée à un échantillon de T observations $\{y_1, \dots, y_T\}$ obtenue sous l'hypothèse de normalité de la loi conditionnelle de y_t sachant son propre passé s'écrit :

$$\log L(\rho) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(h_t(\rho)) - \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\rho)]^2}{h_t(\rho)}$$

où ρ désigne l'ensemble des paramètres du modèle, $m_t(\rho)$ désigne l'espérance conditionnelle et $h_t(\rho)$ désigne la variance conditionnelle. Dans le cas du modèle EGARCH(1,1) présenté ci-dessus :

$$\begin{aligned} m_t(\rho) &= c \\ h_t(\rho) &= \exp(\alpha_0 + \alpha_1 z_{t-1} + \alpha_1 (|z_{t-1}| - \sqrt{2/\pi}) + \beta_1 \log(h_{t-1})) \end{aligned}$$

Les estimateurs du maximum de vraisemblance sont alors obtenus par résolution analytique d'un système de $K = p + q + 2$ (nombre de paramètres à estimer) équations non linéaires :

$$\left. \frac{\partial \log L(\rho)}{\partial \rho} \right|_{\rho=\hat{\rho}} = 0$$

$$\left. \frac{\partial^2 \log L(\rho)}{\partial \rho^2} \right|_{\rho=\hat{\rho}} < 0$$

Dans le cas général du PMV, l'estimateur du PMV est asymptotiquement convergent et normal ([Gouriéroux, 1997](#)).

$$\sqrt{T}(\hat{\rho} - \rho) \xrightarrow[T \rightarrow \infty]{d} N(0, J^{-1}IJ^{-1})$$

$$J = E_0 \left[- \frac{\partial^2 \log L(\rho)}{\partial \rho \partial \rho'} \right]$$

$$I = E_0 \left[\frac{\partial \log L(\rho)}{\partial \rho} \frac{\partial \log L(\rho)}{\partial \rho'} \right]$$

où E_0 désigne l'espérance prise par rapport à la vraie loi. Si la vraie distribution des erreurs est une loi normale (cas du MV) alors $I = J$.

3. Problèmes de convergence et choix des conditions initiales

Il n'existe pas de solution analytique à la maximisation de la vraisemblance dans le cas général des modèles GARCH. On doit donc recourir à une optimisation numérique de la vraisemblance ou de la pseudo vraisemblance suivant les cas. Dans la pratique, se pose alors le problème de la convergence de l'algorithme d'optimisation retenu et plus spécifiquement du problème du choix des conditions initiales.

En ce qui concerne l'algorithme d'optimisation, deux choix sont disponibles dans le cadre de la procédure MODEL sous SAS utilisée pour estimer les modèles GARCH : la méthode de Gauss Newton Raphson et la méthode de Marquardt –Levenberg. C'est cette dernière qui est utilisée dans nos programmes. Pour cet algorithme, nous avons retenu pour conditions initiales sur les paramètres α_i et β_i les valeurs estimées dans le cas d'un modèle GARCH sous hypothèse de normalité. Concernant le paramètre θ , non définis dans un modèle GARCH, nous considérons une condition initiale fixée à 0, initialisant ainsi l'algorithme d'optimisation sur un modèle symétrique.

4. Exemple

Considérons l'exemple d'un modèle EGARCH(1,1) estimé à partir de la série des rendements quotidiens du Nasdaq sur la période allant du 1^{er} janvier 1996 au 31 octobre 2006. On considère ici une procédure de maximum de vraisemblance sous une hypothèse de distribution conditionnelle normale. On note rdt la série de rendement. Soit le modèle suivant.

$$rdt_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t = \exp(\alpha_0 + \alpha_1 \theta z_{t-1} + \alpha_1 (|z_{t-1}| - \sqrt{2/\pi}) + \beta_1 \log(h_{t-1}))$$

$$z_t \text{ i.i.d. } N(0,1)$$

La sortie SAS issue de la procédure MODEL retenue ([Hurlin, 2006](#)) débute par un ensemble d'informations fournies sur la structure du modèle.

The Equation to Estimate is	
rdt =	F(intercept(1))
VAR(rdt) =	H(alpha0, alpha1, beta1, theta, intercept)

Il s'agit d'informations portant sur l'écriture du modèle qui rappelle que l'espérance du rendement est une constante et que la variance conditionnelle du rendement, c'est-à-dire le processus h_t , s'écrit comme une fonction des paramètres $\alpha_0, \alpha_1, \theta$ et β_1 .

NOTE: At FIML Iteration 40 CONVERGE=0.001 Criteria Met. A warning in the log indicates a possible problem with the model.

Puis, la sortie indique que l'algorithme d'optimisation numérique a convergé au bout de 40 itérations, selon un critère d'arrêt de la convergence de 0.001. Enfin, différentes informations sont fournies quant à l'ajustement du modèle (voir PROC MODEL).

Final Convergence Criteria	
R	0.000677

Final Convergence Criteria	
PPC(intercept)	0.026255
RPC(intercept)	0.041207
Object	2.589 ^E -6
Trace(S)	0.000343
Gradient norm	0.068603
Log likelihood	7662.809
Lambda	1E-7

La log vraisemblance est ici égale à -7662 puisqu'il est important de préciser que SAS minimise l'opposé de la fonction de vraisemblance. Ce tableau fournit en outre parmi un ensemble d'informations, la valeur de la norme du gradient de la log-vraisemblance au point optimal.

Enfin, figurent les résultats à proprement parler de la phase d'estimation des paramètres du modèle EGARCH.

Estimations FIML Parameter non linéaires				
Parameter	Estimation	Erreur standard appr.	Valeur du test t	Approx Pr > t
alpha0	-0.1217	0.0285	-4.26	<.0001
alpha1	0.06046	0.00968	6.25	<.0001
beta1	0.985106	0.00339	290.93	<.0001
theta	3.198596	0.5900	5.42	<.0001
intercept	0.000413	0.000225	1.84	0.0665

Dans ce tableau, figure en deuxième colonne la réalisation des estimateurs du MV de chacun des paramètres. Dans la troisième colonne, figure la réalisation de l'estimateur de l'écart type

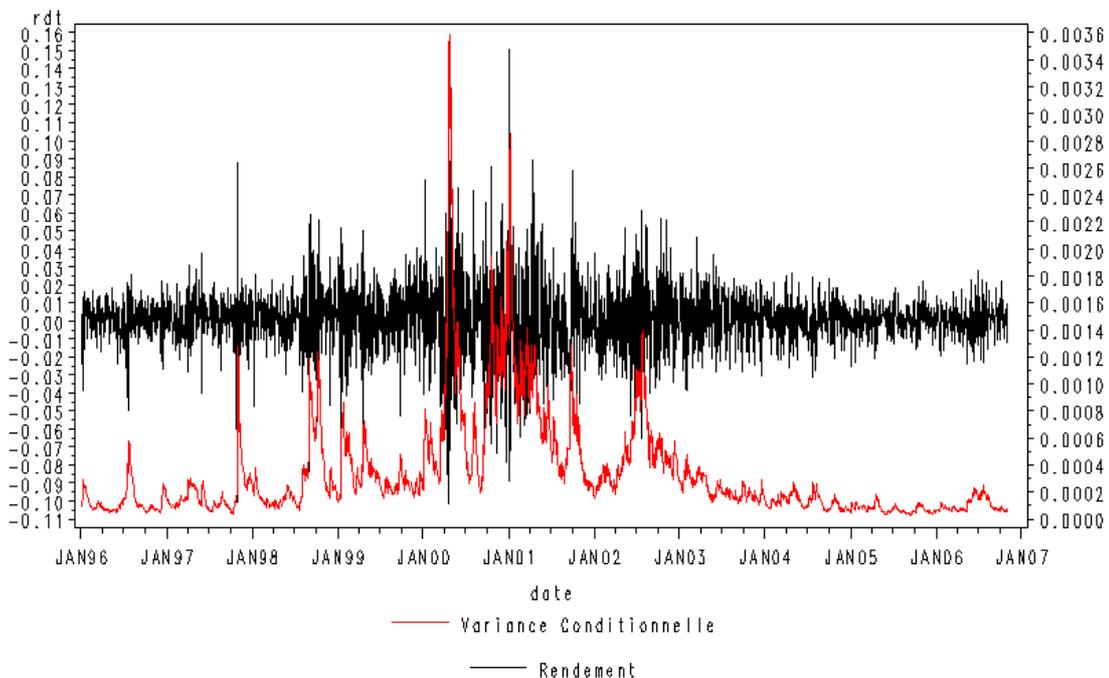
asymptotique de l'estimateur du MV, à partir de laquelle est construite la valeur de la statistique de Student associée au test du nullité du paramètre considéré (quatrième colonne). Enfin, figure dans la dernière colonne la p-value associée à la loi asymptotique de la statistique de Student. Rappelons que pour un risque de première espèce de $\alpha\%$, si la p-value est inférieure au seuil de $\alpha\%$, on est conduit à rejeter l'hypothèse nulle de nullité du coefficient. Le coefficient est donc significativement différent de zéro. On observe que dans le cas de notre exemple, les paramètres $\alpha_0, \alpha_1, \beta_1$ et \mathcal{G} sont significativement différents de zéro au seuil de risque de 1%. La constante est quand à elle significative pour un risque de première espèce de 7%.

A partir de ce modèle, on peut enfin proposer une estimation de la variance conditionnelle, selon l'équation :

$$\hat{h}_t = \exp(\hat{\alpha}_0 + \hat{\alpha}_1 \hat{\theta} z_{t-1} + \hat{\alpha}_1 (|z_{t-1}| - \sqrt{2/\pi}) + \hat{\beta}_1 \log(\hat{h}_{t-1}))$$

où $\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_1$ et $\hat{\mathcal{G}}$ désignent les estimateurs du MV et où \hat{h}_t est donnée. Les valeurs estimées de la variance conditionnelle obtenues dans notre exemple sont reproduites sur la Figure 1 et rendent bien compte des clusters de volatilité, c'est-à-dire de la succession de périodes de turbulences et de calme sur le marché considéré.

Variance Conditionnelle Estimée – Nasdaq



Rédacteurs : [Outreville Séverine](#) et [Guillaumin Claire](#)

Correction : [Hurlin Christophe](#)

Master Econométrie et Statistique Appliquée

Université d'Orléans

Site Value-at-Risk

http://193.49.79.89/esa_prof/index.php



Fiche 8. Modèle QGARCH (Quadratic GARCH)

Le processus QGARCH introduit par [Engle et Ng \(1993\)](#) et [Sentana \(1995\)](#) permet de prendre en compte des asymétries dans la réponse de la volatilité conditionnelle aux innovations. Sentana (1995) a montré que les propriétés des modèles GARCH(1,1) et QGARCH (1,1) sont très similaires : les conditions de stationnarité sont identiques à celles dérivées dans le cas du modèle GARCH et comme le résidu est un processus centré, l'expression de son espérance non conditionnelle est également identique à celle obtenue avec un GARCH. En revanche, la Kurtosis dans les modèles QGARCH est toujours plus grande que dans les modèles GARCH symétriques correspondants, ce qui explique que le QGARCH domine souvent empiriquement le GARCH.

1. Présentation du Modèle QGARCH

L'écriture du modèle QGARCH porte sur la variance conditionnelle du processus considéré. Soit un processus y_t , d'espérance $E(y_t) = c$, satisfaisant une représentation de type QGARCH(1,1). Ce processus s'écrit sous la forme suivante :

$$y_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

où z_t désigne un bruit blanc faible homoscédastique tel que $E(z_t) = 0$ et $Var(z_t) = \sigma_z^2$ et où les paramètres α_i, β_i sont des réels. De façon usuelle, la quantité h_t désigne la variance conditionnelle

du processus y_t telle que $V(y_t | y_{t-1}) = V(\varepsilon_t | \varepsilon_{t-1}) = h_t$ où y_{t-1} désigne l'ensemble des valeurs passées $\{y_{t-1}, \dots, y_0\}$. Le paramètre α_2 permet de modéliser l'effet asymétrique.

Dans le cas du modèle QGARCH, il est utile d'imposer des contraintes sur les paramètres pour satisfaire la condition de positivité de la variance conditionnelle h_t . Ces contraintes se ramènent à $\alpha_0, \alpha_2, \beta_1 > 0$ et $\alpha_1^2 \leq 4\alpha_2\alpha_0$. La stationnarité est assurée si $\alpha_2 + \beta_1 < 1$. Notons que la forme quadratique $f(\varepsilon_{t-1}) = \alpha_1\varepsilon_{t-1} + \alpha_2\varepsilon_{t-1}^2$ est minimale en $-\alpha_1/(2\alpha_2)$. Par conséquent, la symétrie de la réponse n'est pas obtenue en zéro mais en ce point : à amplitude donnée de l'innovation passée, on a bien un impact sur h_t différent selon le signe de ε_{t-1} .

[Sentana \(1995\)](#) propose une généralisation à un modèle QGARCH(p,q) telle que :

$$h_t = \theta + \sum_{i=1}^q \psi_i \varepsilon_{t-i} + \sum_{i=1}^q a_{ii} \varepsilon_{t-i}^2 + 2 \sum_{i=1}^q \sum_{j=i+1}^q a_{ij} \varepsilon_{t-i} \varepsilon_{t-j}$$

Ce modèle permet de tester simplement les asymétries dynamiques dans la fonction de variance conditionnelle de certaines séries financières. De plus, la formulation du QGARCH peut aussi être interprétée comme une approximation de Taylor au second ordre de la fonction de variance ou comme la projection quadratique de l'innovation au carré sur l'ensemble d'information. En fait, cette projection coïncide avec la projection de la vraie variance conditionnelle sur un « ensemble d'information quadratique », donc nous pouvons ainsi comprendre le modèle QGARCH comme produisant des estimateurs filtrés et lissés de la variance conditionnelle.

2. Estimation des Paramètres

Les paramètres du modèle QGARCH peuvent être estimés selon différentes méthodes : maximum de vraisemblance, pseudo maximum de vraisemblance, méthode des moments etc. (pour plus de détails, voir [Gouriéroux, 1997](#)). Les méthodes généralement retenues sont celles du maximum de vraisemblance (MV) ou du pseudo maximum de vraisemblance (PMV). L'avantage du PMV réside dans le fait que l'estimateur obtenu converge malgré une mauvaise spécification (supposée normale) de la distribution conditionnelle des résidus, à condition que la loi spécifiée

appartienne à la famille des lois exponentielles ([Gouriéroux et Montfort, 1989](#)). Ainsi, l'estimateur du MV obtenu sous l'hypothèse de normalité des résidus et l'estimateur du PMV sont identiques, seules leurs lois asymptotiques respectives diffèrent. Toutefois dans les deux cas (MV ou PMV), sous les hypothèses standards, l'estimateur est asymptotiquement convergent et asymptotiquement normal.

Considérons le cas d'un modèle QGARCH(1,1) sous l'hypothèse de normalité des innovations.

$$y_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \varepsilon_{t-1} + \beta_1 h_{t-1}$$

$$z_t \text{ i.i.d. } N(0,1)$$

où les variables z_t sont identiquement et indépendamment distribuées (*i.i.d.*) selon une loi normale centrée et réduite.

La fonction de log-vraisemblance associée à un échantillon de T observations $\{y_1, \dots, y_T\}$ obtenue sous l'hypothèse de normalité de la loi conditionnelle de y_t sachant son propre passé s'écrit :

$$\log L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log[h_t(\theta)] - \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\theta)]^2}{h_t(\theta)}$$

où θ désigne l'ensemble des paramètres du modèle, $m_t(\theta)$ désigne l'espérance conditionnelle et $h_t(\theta)$ désigne la variance conditionnelle. Dans le cas du modèle QGARCH(1,1) présenté ci-dessus :

$$m_t(\theta) = c$$

$$h_t(\theta) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \varepsilon_{t-1} + \beta_1 h_{t-1}$$

Les estimateurs du maximum de vraisemblance sont alors obtenus par résolution analytique d'un système de $K = p + q + 2$ (nombre de paramètres à estimer) équations non linéaires :

$$\left. \frac{\partial \log L(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} = 0$$

$$\left. \frac{\partial^2 \log L(\theta)}{\partial \theta^2} \right|_{\theta=\hat{\theta}} < 0$$

Dans le cas général du PMV, l'estimateur du PMV est asymptotiquement convergent et normal ([Gouriéroux, 1997](#)).

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow[T \rightarrow \infty]{d} N(0, J^{-1} I J^{-1})$$

$$J = E_0 \left[- \frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \right]$$

$$I = E_0 \left[\frac{\partial \log L(\theta)}{\partial \theta} \frac{\partial \log L(\theta)}{\partial \theta'} \right]$$

où E_0 désigne l'espérance prise par rapport à la vraie loi. Si la vraie distribution des erreurs est une loi normale (cas du MV) alors $I = J$.

3. Problèmes de convergence et choix des conditions initiales

Il n'existe pas de solution analytique à la maximisation de la vraisemblance dans le cas général des modèles GARCH. On doit donc recourir à une optimisation numérique de la vraisemblance ou de la pseudo vraisemblance suivant les cas. Dans la pratique, se pose alors le problème de la convergence de l'algorithme d'optimisation retenu et plus spécifiquement du problème du choix de conditions initiales.

En ce qui concerne l'algorithme d'optimisation, deux choix sont disponibles dans le cadre de la procédure MODEL sous SAS utilisée pour estimer les modèles GARCH : la méthode de Gauss Newton Raphson et la méthode de Marquardt –Levenberg. C'est cette dernière qui est utilisée dans nos programmes. Pour cet algorithme, nous avons retenu pour conditions initiales sur les paramètres α_i et β_i les valeurs estimées dans le cas d'un modèle GARCH sous hypothèse de normalité. Concernant le paramètre γ , non défini dans un modèle GARCH, nous retenons une condition initiale fixée à 0, initialisant ainsi l'algorithme d'optimisation sur un modèle symétrique.

4. Exemple

Considérons l'exemple d'un modèle QGARCH(1,1) estimé à partir de la série des rendements quotidiens du Nasdaq sur la période allant du 1^{er} janvier 1996 au 31 octobre 2006. On considère ici une procédure de maximum de vraisemblance sous une hypothèse de distribution conditionnelle normale. On note rdt la série de rendement. Soit le modèle suivant.

$$rdt_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \varepsilon_{t-1} + \beta_1 h_{t-1}$$

$$z_t \text{ i.i.d. } N(0,1)$$

La sortie SAS issue de la procédure MODEL retenue ([Hurlin, 2006](#)) débute par un ensemble d'informations fournies sur la structure du modèle.

The Equation to Estimate is	
rdt =	F(intercept(1))
VAR(rdt) =	H(alpha0, alpha1, beta1, gamma1, intercept)

Il s'agit d'informations portant sur l'écriture du modèle qui rappellent que l'espérance du rendement est une constante et que la variance conditionnelle du rendement, c'est-à-dire le processus h_t , s'écrit comme une fonction des paramètres $\alpha_0, \alpha_1, \gamma_1$ et β_1 .

NOTE: At FIML Iteration 22 CONVERGE=0.001 Criteria Met. A warning in the log indicates a possible problem with the model.

Puis, la sortie indique que l'algorithme d'optimisation numérique a convergé au bout de 22 itérations, selon un critère d'arrêt de la convergence de 0.0001. Enfin, différentes informations sont fournies quant à l'ajustement du modèle (voir PROC MODEL).

Final Convergence Criteria	
R	0.000542
PPC(intercept)	0.010994

Final Convergence Criteria	
RPC(theta)	0.014533
Object	1.382E-6
Trace(S)	0.00021
Gradient norm	39.32242
Log likelihood	16973.68
Lambda	0.006872

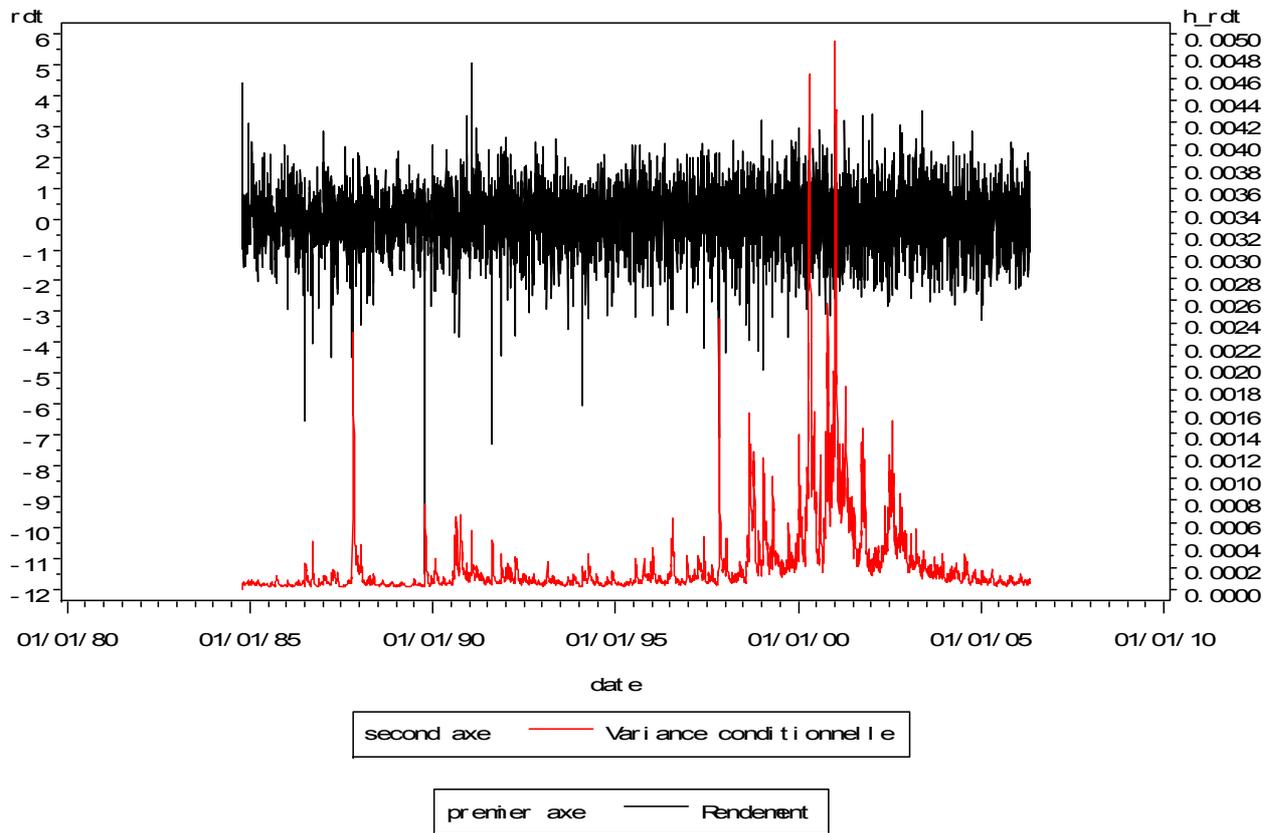
La log vraisemblance est ici égale à -16973,68, puisqu'il est important de préciser que SAS minimise l'opposé de la fonction de vraisemblance. Ce tableau fournit en outre parmi un ensemble d'informations sur la convergence de l'algorithme d'optimisation parmi lesquelles figure la valeur de la norme du gradient de la log-vraisemblance au point optimal.

Enfin, figurent les résultats à proprement parler de la phase d'estimation des paramètres du modèle QGARCH.

Estimations FIML Parameter non linéaires				
Parameter	Estimation	Erreur standard appr.	Valeur du test t	Approx Pr > t
alpha0	3.213E-6	3.794E-7	8.47	<.0001
alpha1	0.140495	0.0117	12.00	<.0001
beta1	0.845815	0.0112	75.77	<.0001
gamma1	-0.00054	0.000110	-4.96	<.0001
intercept	0.000645	0.000121	5.33	<.0001

Dans ce tableau, figure en première colonne la réalisation des estimateurs du MV de chacun des paramètres. Dans la seconde colonne, figure la réalisation de l'estimateur de l'écart type asymptotique de l'estimateur du MV, à partir de laquelle est construite la valeur de la statistique de Student associée au test de nullité du paramètre considéré (troisième colonne). Enfin, figure dans la dernière colonne la p-value associée à la loi asymptotique de la statistique de Student. Rappelons que pour un risque de première espèce de $\alpha\%$, si la p-value est inférieure au seuil de $\alpha\%$, on est conduit à rejeter l'hypothèse nulle de nullité du coefficient. Le coefficient est donc significativement différent de zéro. On observe que dans le cas de notre exemple, tous les paramètres sont significativement différents de zéro au seuil de risque de 1%.

Variance Conditionnelle Estimée — Nasdaq



A partir de ce modèle, on peut enfin proposer une estimation de la variance conditionnelle, selon l'équation :

$$\hat{h}_t = \hat{\alpha}_0 + \hat{\alpha}_1 \varepsilon_{t-1}^2 + \hat{\gamma}_1 \varepsilon_{t-1} + \hat{\beta}_1 \hat{h}_{t-1}$$

\hat{h}_1 donné

où $\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_1, \hat{\gamma}_1$ désignent les estimateurs du MV. Les valeurs estimées de la variance conditionnelle obtenues dans notre exemple sont reproduites sur la Figure 1 et rendent bien compte des clusters de volatilité, c'est-à-dire de la succession de périodes de turbulences et de calme sur le marché considéré.

Rédacteurs : [Doublet Emilie](#) et [Renvoisé Amélie](#)

Correction : [Hurlin Christophe](#)

[Master ESA](#). 15 Novembre 2006

Master Econométrie et Statistique Appliquée

Université d'Orléans

Site Value-at-Risk

http://193.49.79.89/esa_prof/index.php



Fiche 9. Modèle IGARCH (Integrated GARCH)

Le modèle IGARCH (*Integrated General Autoregressive Conditional Heteroskedasticity*) correspond au cas d'une racine unitaire dans la variance conditionnelle. C'est un modèle ARCH non linéaire caractérisé par un effet de persistance dans la variance. C'est-à-dire qu'un choc sur la variance conditionnelle actuelle se répercute sur toutes les valeurs futures prévues.

1. Présentation du Modèle IGARCH

L'écriture du modèle IGARCH porte sur la non stationnarité de son processus de variance conditionnelle et par une variance non conditionnelle infinie. Soit un processus y_t , d'espérance $E(y_t) = 0$, satisfaisant une représentation de type IGARCH(p,q). Ce processus s'écrit sous la forme suivante :

$$y_t = z_t \sqrt{h_t}$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p (1 - \alpha_i) h_{t-i}$$

où z_t désigne un bruit blanc faible homoscédastique tel que $E(z_t) = 0$ et $Var(z_t) = 1$ et où les paramètres α_0, α_i sont des réels tels que:

$$\sum_{i=1}^q \alpha_i + \sum_{i=1}^p (1 - \alpha_i) = 1$$

De façon usuelle, la quantité h_t désigne la variance conditionnelle du processus y_t telle que $V(y_t | y_{t-1}) = V(\varepsilon_t | \varepsilon_{t-1}) = h_t$ où y_{t-1} désigne l'ensemble des valeurs passées $\{y_{t-1}, \dots, y_0\}$.

Les prévisions de la variance conditionnelle aux différents horizons k sont de la forme :

$$E(h_{t+k} / \varepsilon_t) = (\alpha_1 + \beta_1)^k h_t + \alpha_0 \sum_{i=0}^{k-1} (\alpha_1 + \beta_1)^i$$

Lorsque $\alpha_1 + \beta_1 < 1$, le processus ε_t est stationnaire et l'influence du choc sur la variance conditionnelle est décroissante et asymptotiquement négligeable sur h_{t+k} . Dans ce cas, on a $\beta_1 \neq (1 - \alpha_1)$. Lorsque $\alpha_1 + \beta_1 = 1$ avec $\beta_1 = (1 - \alpha_1)$ on a :

$$E(h_{t+k} / \varepsilon_t) = h_t + \alpha_0 k$$

en présence d'un terme constant, $E(h_{t+k} / \varepsilon_t)$ diverge avec k . Ce qui traduit alors la persistance des chocs passés.

2. Estimation des Paramètres

Les paramètres du modèle GARCH peuvent être estimés selon différentes méthodes : maximum de vraisemblance, pseudo maximum de vraisemblance, méthode des moments etc. (pour plus de détails, voir [Gouriéroux, 1997](#)). Les méthodes généralement retenues sont celles du maximum de vraisemblance (MV) ou du pseudo maximum de vraisemblance (PMV). L'avantage du PMV réside dans le fait que l'estimateur obtenu converge malgré une mauvaise spécification (supposée normale) de la distribution conditionnelle des résidus, à condition que la loi spécifiée appartienne à la famille des lois exponentielles ([Gouriéroux et Montfort, 1989](#)). Ainsi, l'estimateur du MV obtenu sous l'hypothèse de normalité des résidus et l'estimateur du PMV sont identiques, seules leurs lois asymptotiques respectives diffèrent. Toutefois dans les deux cas (MV ou PMV), sous les hypothèses standards, l'estimateur est asymptotiquement convergent et asymptotiquement normal.

Considérons le cas d'un modèle IGARCH(1,1) sous l'hypothèse de normalité des innovations.

$$y_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + (1 - \alpha_1) h_{t-1}$$

$$z_t \text{ i.i.d. } N(0,1)$$

où les variables z_t sont identiquement et indépendamment distribuées (*i.i.d.*) selon une loi normale centrée et réduite. La fonction de log-vraisemblance associée à un échantillon de T observations $\{y_1, \dots, y_T\}$ obtenue sous l'hypothèse de normalité de la loi conditionnelle de y_t sachant son propre passé s'écrit :

$$\log L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(h_t(\theta)) - \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\theta)]^2}{h_t(\theta)}$$

où θ désigne l'ensemble des paramètres du modèle, $m_t(\theta)$ désigne l'espérance conditionnelle et $h_t(\theta)$ désigne la variance conditionnelle. Dans le cas du modèle IGARCH(1,1) présenté ci-dessus :

$$m_t(\theta) = c$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + (1 - \alpha_1) h_{t-1}$$

Les estimateurs du maximum de vraisemblance sont alors obtenus par résolution analytique d'un système de $K = p + q + 2$ (nombre de paramètres à estimer) équations non linéaires :

$$\left. \frac{\partial \log L(\theta)}{\partial \theta} \right|_{\theta = \hat{\theta}} = 0$$

$$\left. \frac{\partial^2 \log L(\theta)}{\partial \theta^2} \right|_{\theta = \hat{\theta}} < 0$$

Dans le cas général du PMV, l'estimateur du PMV est asymptotiquement convergent et normal ([Gouriéroux, 1997](#)).

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow[T \rightarrow \infty]{d} N(0, J^{-1} I J^{-1})$$

$$J = E_0 \left[-\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \right]$$

$$I = E_0 \left[\frac{\partial \log L(\theta)}{\partial \theta} \frac{\partial \log L(\theta)}{\partial \theta'} \right]$$

où E_0 désigne l'espérance prise par rapport à la vraie loi. Si la vraie distribution des erreurs est une loi normale (cas du MV) alors $I = J$.

3. Problèmes de convergence et choix des conditions initiales

Il n'existe pas de solution analytique à la maximisation de la vraisemblance dans le cas général des modèles GARCH. On doit donc recourir à une optimisation numérique de la vraisemblance ou de la pseudo vraisemblance suivant les cas. Dans la pratique, se pose alors le problème de la convergence de l'algorithme d'optimisation retenu et plus spécifiquement du problème du choix de conditions initiales.

En ce qui concerne l'algorithme d'optimisation, deux choix sont disponibles dans le cadre de la procédure AUTOREG sous SAS utilisée pour estimer les modèles GARCH : la méthode de Gauss Newton Raphson et la méthode de Marquardt –Levenberg. C'est cette dernière qui est utilisée dans nos programmes. Pour cet algorithme, nous avons retenu pour conditions initiales sur les paramètres α_i et β_i les valeurs estimées dans le cas d'un modèle IGARCH sous hypothèse de normalité.

4. Exemple

Considérons l'exemple d'un modèle IGARCH(1,1) estimé à partir de la série des rendements quotidien du Dow Jones sur la période 01 octobre 1996 au 16 novembre 2006. On considère ici une procédure de maximum de vraisemblance sous une hypothèse de distribution conditionnelle normale. On note rdt la série de rendement. Soit le modèle suivant.

$$rdt_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t(\theta) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

$$z_t \text{ i.i.d. } N(0,1)$$

La sortie SAS issue de la procédure AUTOREG retenue ([Hurlin, 2006](#)) débute par un ensemble d'informations fournies sur la structure du modèle.

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > t
Intercept	1	-0.000547	0.000175	-3.12	0.0018
ARCH0	1	7.9516E-7	1.9997E-7	3.98	<.0001
ARCH1	1	0.0899	0.008099	11.11	<.0001

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > t
GARCH1	1	0.9101	0.008099	112.36	<.0001

Dans ce tableau, figure en premier le degré de liberté de la statistique de Student, ensuite réalisation des estimateurs du MV de chacun des paramètres. Dans la troisième colonne, figure la réalisation de l'estimateur de l'écart type asymptotique de l'estimateur du MV, à partir duquel est construit la valeur de la statistique de Student associée au test de nullité du paramètre considéré (troisième colonne). Enfin, figure dans la dernière colonne la p-value associée à la loi asymptotique de la statistique de Student. Rappelons que pour un risque de première espèce de $\alpha\%$, si la p-value est inférieure au seuil de $\alpha\%$, on est conduit à rejeter l'hypothèse nulle de nullité du coefficient. Le coefficient est donc significativement différent de zéro. On observe que dans le cas de notre exemple, les paramètres α_0 , α_1 , β_1 et la constante sont significativement différents de zéro au seuil de risque de 1%.

A partir de ce modèle, on peut enfin proposer une estimation de la variance conditionnelle, selon l'équation :

$$\hat{h}_t = \hat{\alpha}_0 + \hat{\alpha}_1 \varepsilon_{t-1}^2 + \left(1 - \hat{\alpha}_1\right) \hat{h}_{t-1}$$

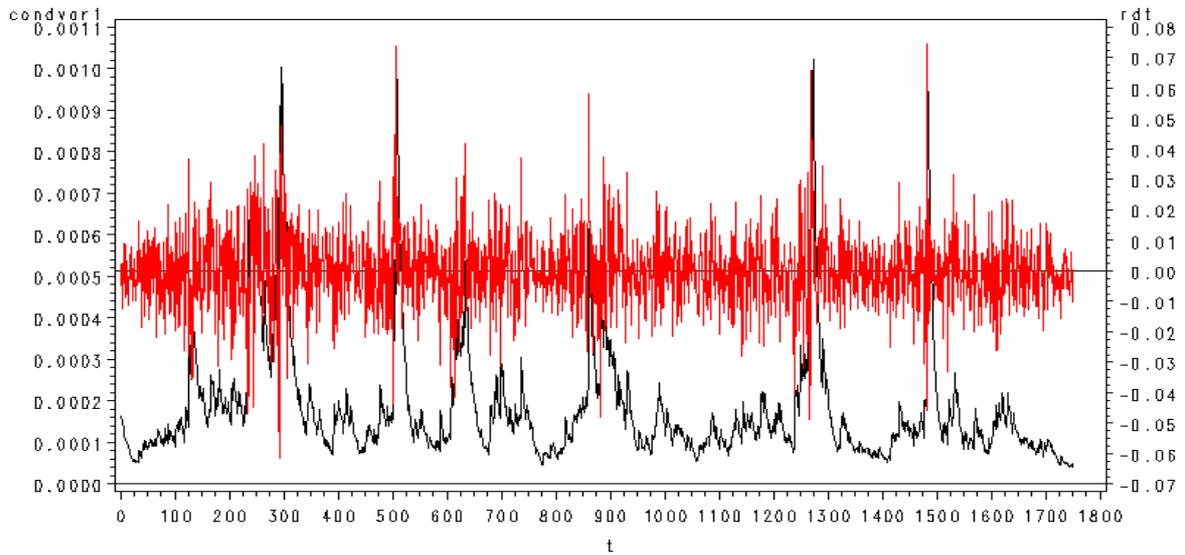
$$\text{ou } \hat{h}_t = \hat{\alpha}_0 + \hat{\alpha}_1 (rdt - \hat{c})_{t-1}^2 + \left(1 - \hat{\alpha}_1\right) \hat{h}_{t-1}$$

où $\hat{\alpha}_0, \hat{\alpha}_1$ sont les coefficients estimés par MV et où \hat{h}_1 est donnée.

Les valeurs estimées de la variance conditionnelle obtenues dans notre exemple sont reproduites sur la Figure suivante et rendent bien compte des clusters de volatilité, c'est-à-dire de la succession de périodes de turbulences et de calme sur le marché considéré.

Variance conditionnelle et Rendements estimés

Exemple du I-GARCH(1,1)



Rédacteurs : [Edwige Balliet](#), [Carole Njoya](#) et [Sidina Medani](#)

Correction : [Hurlin Christophe](#)

[Master ESA](#). 14 Novembre 2006

Master Econométrie et Statistique Appliquée

Université d'Orléans

Site Value-at-Risk

http://193.49.79.89/esa_prof/index.php



Université d'Orléans

Fiche 10. Modèle GJR-GARCH

Le modèle GJR-GARCH a été introduit par [Glosten, Jagannathan et Runkle \(1993\)](#). C'est un modèle GARCH non linéaire qui permet de rendre compte de l'asymétrie dans la réponse de la variance conditionnelle à une innovation. La logique de ce modèle est similaire à celle des modèles à changements de régimes et plus spécifiquement des modèles à seuils ([Tong, 1990](#)). Le principe du modèle GJR-GARCH est que la dynamique de la variance conditionnelle admet un changement de régime qui dépend du signe de l'innovation passée.

1. Présentation du Modèle GJR-GARCH

Soit un processus y_t , d'espérance $E(y_t) = c$, satisfaisant une représentation de type GJR-GARCH(p,q). Ce processus s'écrit sous la forme suivante :

$$y_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \sum_{i=1}^q (\alpha_i \varepsilon_{t-i}^2 + \varphi_i I_{\varepsilon_{t-i} < 0} \varepsilon_{t-i}^2) + \sum_{i=1}^p \beta_i h_{t-i}$$

où z_t désigne un bruit blanc faible homoscédastique tel que $E(z_t) = 0$ et $Var(z_t) = \sigma_z^2$, où les paramètres α_i, β_i et φ_i sont des réels et où $I_{\varepsilon_{t-i} < 0}$ désigne la fonction indicatrice telle que $I_{\varepsilon_{t-i} < 0} = 1$ si $\varepsilon_{t-i} < 0$ et $I_{\varepsilon_{t-i} < 0} = 0$ sinon. De façon usuelle, la quantité h_t désigne la variance conditionnelle du processus y_t telle que $V(y_t | \underline{y}_{t-1}) = V(\varepsilon_t | \underline{\varepsilon}_{t-1}) = h_t$ où \underline{y}_{t-1} désigne l'ensemble des valeurs passées $\{y_{t-1}, \dots, y_0\}$.

Dans le cas du modèle GJR-GARCH, il est utile d'imposer des contraintes sur les paramètres pour satisfaire la condition de positivité de la variance conditionnelle h_t : $\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, q$ et $\beta_i \geq 0, i = 1, \dots, p$. Les paramètres φ_i permettent de modéliser un effet asymétrique lié aux signes des innovations passées ε_{t-i} . Si $\varphi_i > 0$ (respectivement si $\varphi_i < 0$), un choc positif sur les innovations passées à la $t-i$ se traduira à la date t par une augmentation (respectivement une diminution) de la variance conditionnelle, c'est-à-dire de la volatilité, du processus y_t .

2. Estimation des Paramètres

Les paramètres du modèle GJR-GARCH peuvent être estimés selon différentes méthodes : maximum de vraisemblance, pseudo maximum de vraisemblance, méthode des moments etc. (pour plus de détails, voir [Gouriéroux, 1997](#)). Les méthodes généralement retenues sont celles du maximum de vraisemblance (MV) ou du pseudo maximum de vraisemblance (PMV). L'avantage du PMV réside dans le fait que l'estimateur obtenu converge malgré une mauvaise spécification (supposée normale) de la distribution conditionnelle des résidus, à condition que la loi spécifiée appartienne à la famille des lois exponentielles ([Gouriéroux et Montfort, 1989](#)). Ainsi, l'estimateur du MV obtenu sous l'hypothèse de normalité des résidus et l'estimateur du PMV sont identiques, seules leurs lois asymptotiques respectives diffèrent. Toutefois dans les deux cas (MV ou PMV), sous les hypothèses standards, l'estimateur est asymptotiquement convergent et asymptotiquement normal.

Considérons le cas d'un modèle GJR-GARCH(1,1) sous l'hypothèse de normalité des innovations.

$$\begin{aligned}
 y_t &= c + \varepsilon_t \\
 \varepsilon_t &= z_t \sqrt{h_t} \\
 h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \varphi I_{\varepsilon_{t-1} < 0} \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \\
 z_t & \text{ i.i.d. } N(0,1)
 \end{aligned}$$

où les variables z_t sont identiquement et indépendamment distribuées (*i.i.d.*) selon une loi normale centrée et réduite. La fonction de log-vraisemblance associée à un échantillon de T observations

$\{y_1, \dots, y_T\}$ obtenue sous l'hypothèse de normalité de la loi conditionnelle de y_t sachant son propre passé s'écrit :

$$\log L(\rho) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(h_t(\rho)) - \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\rho)]^2}{h_t(\rho)}$$

où ρ désigne l'ensemble des paramètres du modèle, $m_t(\rho)$ désigne l'espérance conditionnelle et $h_t(\rho)$ désigne la variance conditionnelle. Dans le cas du modèle GJR-GARCH(1,1) présenté ci-dessus :

$$m_t(\rho) = c$$

$$h_t(\rho) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \varphi I_{\varepsilon_{t-1} < 0} \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

Les estimateurs du maximum de vraisemblance sont alors obtenus par résolution analytique d'un système de $K = p + q + 2$ (nombre de paramètres à estimer) équations non linéaires :

$$\left. \frac{\partial \log L(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} = 0$$

$$\left. \frac{\partial^2 \log L(\theta)}{\partial \theta^2} \right|_{\theta=\hat{\theta}} < 0$$

Dans le cas général du PMV, l'estimateur du PMV est asymptotiquement convergent et normal ([Gouriéroux, 1997](#)).

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow[T \rightarrow \infty]{d} N(0, J^{-1} I J^{-1})$$

$$J = E_0 \left[-\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \right]$$

$$I = E_0 \left[\frac{\partial \log L(\theta)}{\partial \theta} \frac{\partial \log L(\theta)}{\partial \theta'} \right]$$

où E_0 désigne l'espérance prise par rapport à la vraie loi. Si la vraie distribution des erreurs est une loi normale (cas du MV) alors $I = J$.

3. Problèmes de convergence et choix des conditions initiales

Il n'existe pas de solution analytique à la maximisation de la vraisemblance dans le cas général des modèles GARCH. On doit donc recourir à une optimisation numérique de la vraisemblance ou de la pseudo vraisemblance suivant les cas. Dans la pratique, se pose alors le problème de la convergence de l'algorithme d'optimisation retenu et plus spécifiquement du problème du choix de conditions initiales.

En ce qui concerne l'algorithme d'optimisation, deux choix sont disponibles dans le cadre de la procédure MODEL sous SAS utilisée pour estimer les modèles GARCH : la méthode de Gauss Newton Raphson et la méthode de Marquardt –Levenberg. C'est cette dernière qui est utilisée dans nos programmes. Pour cet algorithme, nous avons retenu pour conditions initiales sur les paramètres α_i et β_i les valeurs estimées dans le cas d'un modèle GARCH sous hypothèse de normalité. Concernant les paramètres φ_i , non définis dans un modèle GARCH, nous considérons des conditions initiales fixées à 0.0001 permettant ainsi d'évaluer la pertinence de notre modèle asymétrique.

4. Exemple

Considérons l'exemple d'un modèle GJR-GARCH(1,1) estimé à partir de la série des rendements quotidien du Nasdaq sur la période allant du 1^{er} janvier 1996 au 31 octobre 2006. On considère ici une procédure de maximum de vraisemblance sous une hypothèse de distribution conditionnelle normale. On note rdt la série de rendement. Soit le modèle suivant.

$$rdt_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \varphi I_{\varepsilon_{t-1} < 0} \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

$$z_t \text{ i.i.d. } N(0,1)$$

La sortie SAS issue de la procédure MODEL retenue ([Hurlin, 2006](#)) débute par un ensemble d'informations fournies sur la structure du modèle.

The Equation to Estimate is	
rdt =	F(intercept(1))

The Equation to Estimate is	
VAR(rdt) =	H(alpha0, alpha1, beta1, phi, intercept)

Il s'agit d'informations portant sur l'écriture du modèle qui rappelle que l'espérance du rendement est une constante et que la variance conditionnelle du rendement, c'est-à-dire le processus h_t , s'écrit comme une fonction des paramètres α_0, α_1, ρ et β_1 .

NOTE: At FIML Iteration 15 CONVERGE=0.001 Criteria Met. A warning in the log indicates a possible problem with the model.

Puis, la sortie indique que l'algorithme d'optimisation numérique a convergé au bout de 15 itérations, selon un critère d'arrêt de la convergence de 0.001. Enfin, différentes informations sont fournies quant à l'ajustement du modèle.

Final Convergence Criteria	
R	0.000981
PPC(alpha0)	0.013881
RPC(phi)	0.048459
Object	5.664E-7
Trace(S)	0.000343
Gradient norm	9.905341
Log likelihood	7644.754
Lambda	0.134218

La log-vraisemblance est ici égale à -7644,7, puisqu'il est important de préciser que SAS minimise l'inverse de la fonction de vraisemblance. Ce tableau fournit en outre parmi un ensemble d'informations, la valeur de la norme du gradient de la log-vraisemblance au point optimal.

Enfin, figurent les résultats à proprement parler de la phase d'estimation des paramètres du modèle GJR-GARCH.

Estimations FIML Parameter non linéaires				
Parameter	Estimation	Erreur standard appr.	Valeur du test t	Approx Pr > t
alpha0	4.094E-6	7.97E-7	5.14	<.0001
alpha1	0.174515	0.0190	9.17	<.0001
beta1	0.866692	0.0136	63.83	<.0001
phi	-0.10213	0.0180	-5.67	<.0001
intercept	0.000482	0.000232	2.08	0.0375

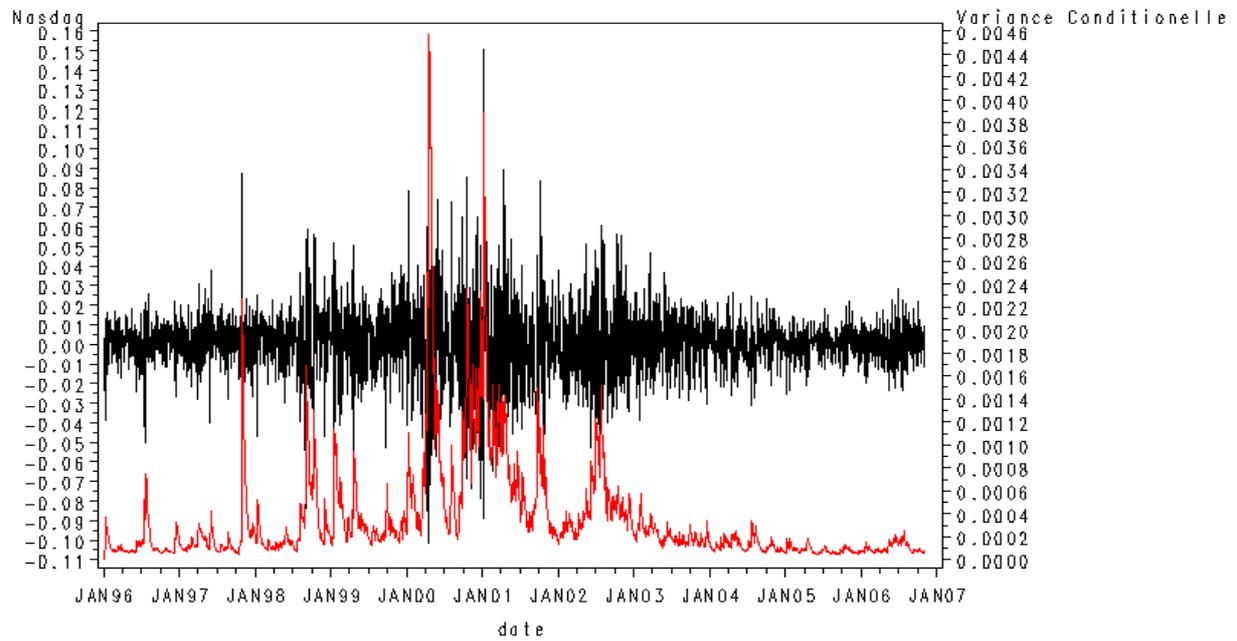
Dans ce tableau, figure en première colonne la réalisation des estimateurs du MV de chacun des paramètres. Dans la seconde colonne, figure la réalisation de l'estimateur de l'écart type asymptotique de l'estimateur du MV, à partir de laquelle est construite la valeur de la statistique de Student associée au test du nullité du paramètre considéré (troisième colonne). Enfin, figure dans la dernière colonne la p-value associée à la loi asymptotique de la statistique de Student. Rappelons que pour un risque de première espèce de $\alpha\%$, si la p-value est inférieure au seuil de $\alpha\%$, on est conduit à rejeter l'hypothèse nulle de nullité du coefficient. Le coefficient est donc significativement différent de zéro. On observe que dans le cas de notre exemple, les paramètres $\alpha_0, \alpha_1, \varphi$ et β_1 sont significativement différents de zéro au seuil de risque de 1%. La constante est quand à elle significative pour un risque de première espèce de 4%.

A partir de ce modèle, on peut enfin proposer une estimation de la variance conditionnelle, selon l'équation :

$$\hat{h}_t = \hat{\alpha}_0 + \hat{\alpha}_1 \varepsilon_{t-1}^2 + \hat{\varphi} I_{\varepsilon_{t-1} < 0} \varepsilon_{t-1}^2 + \hat{\beta}_1 h_{t-1}$$

où $\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_1$ et $\hat{\varphi}$ désignent les estimateurs du MV. Les valeurs estimées de la variance conditionnelle obtenues dans notre exemple sont reproduites sur la Figure suivante et rendent bien compte de la hausse très nette de la volatilité observée sur l'indice Nasdaq à partir de la fin des années 2000.

Variance Conditionnelle Estimée — Nasdaq



Rédacteurs : [Outreville Séverine](#) et [Guillaumin Claire](#)

Correction : [Hurlin Christophe](#)

[Master ESA](#), 14 Novembre 2006

Master Econométrie et Statistique Appliquée

Université d'Orléans

Site Value-at-Risk

http://193.49.79.89/esa_prof/index.php



Fiche 11. Modèle GARCH-M (GARCH in Mean)

Le modèle GARCH-M (*General Autoregressive Conditional Heteroskedasticity in Mean*) a été introduit par [Engle-Lilien-Robbins \(1987\)](#). C'est un modèle ARCH qui permet de mesurer l'influence du rendement des titres sur la volatilité conditionnelle. Comme les modèles ARCH linéaires, le modèle GARCH-M repose sur une spécification quadratique de la variance conditionnelle des perturbations.

1. Présentation du Modèle GARCH-M

L'écriture du modèle GARCH-M porte sur la non stationnarité de son processus de variance conditionnelle et par une variance non conditionnelle infinie. Soit un processus y_t , d'espérance $E(y_t) = 0$, satisfaisant une représentation de type GARCH-M (p,q). Ce processus s'écrit sous la forme suivante :

$$\begin{aligned}y_t &= x_t b + \delta h_t + \varepsilon_t \\ \varepsilon_t &= z_t \sqrt{h_t} \\ h_t &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i}\end{aligned}$$

où z_t désigne un bruit blanc faible homoscédastique tel que $E(z_t) = 0$ et $Var(z_t) = \sigma_z^2$ et où les paramètres α_i, β_i sont des réels. De façon usuelle, la quantité h_t désigne la variance conditionnelle du processus y_t telle que $V(y_t | \underline{y}_{t-1}) = V(\varepsilon_t | \underline{\varepsilon}_{t-1}) = h_t$ où \underline{y}_{t-1} désigne l'ensemble des valeurs passées $\{y_{t-1}, \dots, y_0\}$. Afin de garantir la positivité de la variance conditionnelle, les conditions de positivité des paramètres suffisent, on suppose : $\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, q, \beta_i \geq 0, i = 1, \dots, p$.

Le paramètre δ mesure l'impact de la variance conditionnelle sur le rendement ; il correspond également au coefficient de l'aversion relative au risque. Les variables z_t sont indépendamment et identiquement distribuées (*i.i.d.*).

En plus de la forme linéaire de l'écriture de y_t ci-dessus des variantes sont possibles ; c'est-à-dire les formes log-linéaire et racine carré.

$$y_t = x_t b + \delta \log(h_t) + \varepsilon_t \quad \text{Forme Log-Linéaire}$$

$$y_t = x_t b + \delta \sqrt{h_t} + \varepsilon_t \quad \text{Forme Racine Carré}$$

2. Estimation des Paramètres

Les paramètres du modèle GARCH peuvent être estimés selon différentes méthodes : maximum de vraisemblance, pseudo maximum de vraisemblance, méthode des moments etc. (pour plus de détails, voir [Gouriéroux, 1997](#)). Les méthodes généralement retenues sont celles du maximum de vraisemblance (MV) ou du pseudo maximum de vraisemblance (PMV). L'avantage du PMV réside dans le fait que l'estimateur obtenu converge malgré une mauvaise spécification (supposée normale) de la distribution conditionnelle des résidus, à condition que la loi spécifiée appartienne à la famille des lois exponentielles ([Gouriéroux et Montfort, 1989](#)). Ainsi, l'estimateur du MV obtenu sous l'hypothèse de normalité des résidus et l'estimateur du PMV sont identiques, seules leurs lois asymptotiques respectives diffèrent. Toutefois dans les deux cas (MV ou PMV), sous les hypothèses standards, l'estimateur est asymptotiquement convergent et asymptotiquement normal.

Considérons le cas d'un modèle GARCH-M (1,1) sous l'hypothèse de normalité des innovations.

$$y_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

$$z_t \text{ i.i.d. } N(0,1)$$

où les variables z_t sont identiquement et indépendamment distribuées (*i.i.d.*) selon une loi normale centrée et réduite. La fonction de log-vraisemblance associée à un échantillon de T observations

$\{y_1, \dots, y_T\}$ obtenue sous l'hypothèse de normalité de la loi conditionnelle de y_t sachant son propre passé s'écrit :

$$\log L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(h_t(\theta)) - \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\theta)]^2}{h_t(\theta)}$$

où θ désigne l'ensemble des paramètres du modèle, $m_t(\theta)$ désigne l'espérance conditionnelle et $h_t(\theta)$ désigne la variance conditionnelle. Dans le cas du modèle GARCH-M (1,1) :

$$m_t(\theta) = c$$

$$h_t(\theta) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

Les estimateurs du maximum de vraisemblance sont alors obtenus par résolution analytique d'un système de $K = p + q + 2$ (nombre de paramètres à estimer) équations non linéaires :

$$\left. \frac{\partial \log L(\theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} = 0$$

$$\left. \frac{\partial^2 \log L(\theta)}{\partial \theta^2} \right|_{\theta=\hat{\theta}} < 0$$

Dans le cas général du PMV, l'estimateur du PMV est asymptotiquement convergent et normal ([Gouriéroux, 1997](#)).

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow[T \rightarrow \infty]{d} N(0, J^{-1} I J^{-1})$$

$$J = E_0 \left[-\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \right]$$

$$I = E_0 \left[\frac{\partial \log L(\theta)}{\partial \theta} \frac{\partial \log L(\theta)}{\partial \theta'} \right]$$

où E_0 désigne l'espérance prise par rapport à la vraie loi. Si la vraie distribution des erreurs est une loi normale (cas du MV) alors $I = J$.

3. Problèmes de convergence et choix des conditions initiales

Il n'existe pas de solution analytique à la maximisation de la vraisemblance dans le cas général des modèles GARCH. On doit donc recourir à une optimisation numérique de la vraisemblance ou de la pseudo vraisemblance suivant les cas. Dans la pratique, se pose alors le problème de la convergence

de l'algorithme d'optimisation retenu et plus spécifiquement du problème du choix de conditions initiales.

En ce qui concerne l'algorithme d'optimisation, deux choix sont disponibles dans le cadre de la procédure AUTOREG sous SAS utilisée pour estimer les modèles GARCH : la méthode de Gauss Newton Raphson et la méthode de Marquardt –Levenberg. C'est cette dernière qui est utilisée dans nos programmes. Pour cet algorithme, nous avons retenu pour conditions initiales sur les paramètres α_i et β_i les valeurs estimées dans le cas d'un modèle GARCH-M sous hypothèse de normalité.

4. Exemple

Considérons l'exemple d'un modèle GARCH-M (1,1) estimé à partir de la série des rendements quotidien du Dow Jones sur la période 01 octobre 1996 au 16 novembre 2006. On considère ici une procédure de maximum de vraisemblance sous une hypothèse de distribution conditionnelle normale. On note rdt la série de rendement. Soit le modèle suivant.

$$rdt_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t(\theta) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

$$z_t \text{ i.i.d. } N(0,1)$$

La sortie SAS issue de la procédure AUTOREG retenue ([Hurlin, 2006](#)) débute par un ensemble d'informations fournies sur la structure du modèle.

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > t
Intercept	1	-0.000547	0.000175	-3.12	0.0018
ARCH0	1	7.9516E-7	1.9997E-7	3.98	<.0001
ARCH1	1	0.0899	0.008099	11.11	<.0001
GARCH1	1	0.9101	0.008099	112.36	<.0001

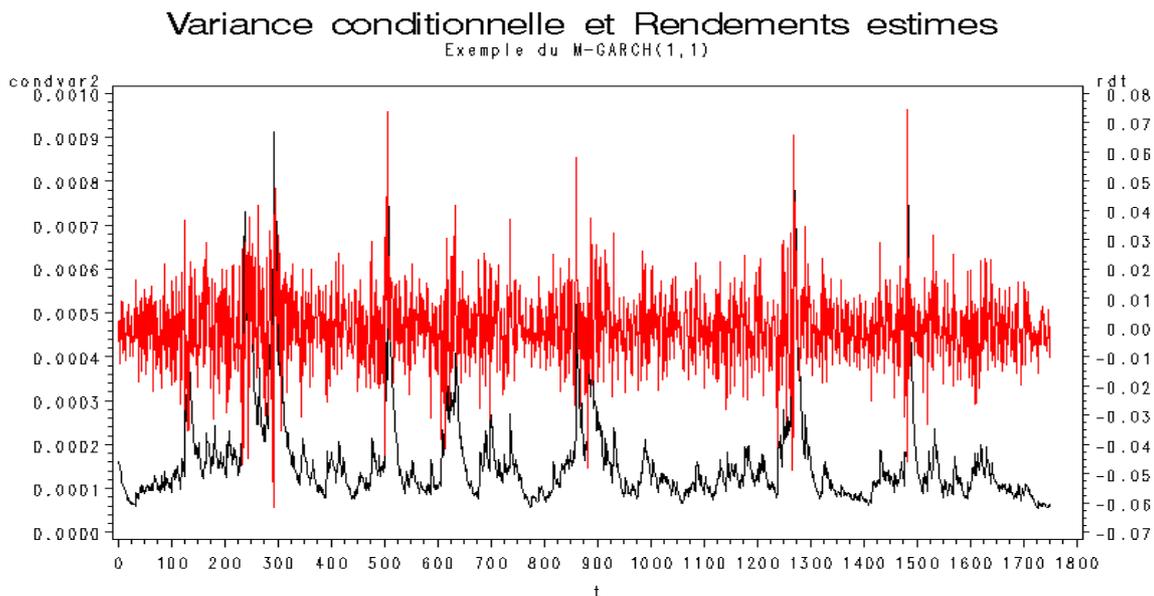
Dans ce tableau, figure en premier le degré de liberté de la statistique de Student, ensuite réalisation des estimateurs du MV de chacun des paramètres. Dans la troisième colonne, figure la réalisation de l'estimateur de l'écart type asymptotique de l'estimateur du MV, à partir duquel est construit la valeur de la statistique de Student associée au test de nullité du paramètre considéré (troisième colonne). Enfin, figure dans la dernière colonne la p-value associée à la loi asymptotique de la statistique de Student. Rappelons que pour un risque de première espèce de $\alpha\%$, si la p-value est inférieure au seuil de $\alpha\%$, on est conduit à rejeter l'hypothèse nulle de nullité du coefficient. Le coefficient est donc significativement différent de zéro. On observe que dans le cas de notre exemple, les paramètres α_0 , α_1 , β_1 et la constante sont significativement différents de zéro au seuil de risque de 1%.

A partir de ce modèle, on peut enfin proposer une estimation de la variance conditionnelle, selon l'équation :

$$\hat{h}_t = \hat{\alpha}_0 + \hat{\alpha}_1 \varepsilon_{t-1}^2 + \left(1 - \hat{\alpha}_1\right) \hat{h}_{t-1}$$

$$\text{ou } \hat{h}_t = \hat{\alpha}_0 + \hat{\alpha}_1 (rdt - \hat{c})_{t-1}^2 + \left(1 - \hat{\alpha}_1\right) \hat{h}_{t-1}$$

où $\hat{\alpha}_0, \hat{\alpha}_1$ sont les coefficients estimés par MV et où \hat{h}_1 est donnée.



Les valeurs estimées de la variance conditionnelle obtenues dans notre exemple sont reproduites sur la Figure 1 et rendent bien compte des clusters de volatilité, c'est-à-dire de la succession de périodes de turbulences et de calme sur le marché considéré.

Rédacteurs : [Edwige Balliet](#), [Carole Njoya](#) et [Sidina Medani](#)

Correction : [Hurlin Christophe](#)

[Master ESA](#). 14 Novembre 2006

Master Econométrie et Statistique Appliquée

Université d'Orléans

Site Value-at-Risk

http://193.49.79.89/esa_prof/index.php



Fiche 12. Distributions Conditionnelles - Modèles GARCH

Introduction

Différentes lois peuvent être utilisées dans le cadre des procédures de maximum de vraisemblance pour l'estimation des paramètres du modèle GARCH. Nous proposons ici une présentation des propriétés des lois les plus généralement utilisées à savoir :

- la [distribution Normale](#)
- la [distribution de Student](#)
- la [distribution General Error Distribution \(GED\)](#)
- la [distribution Skewed Student](#)
- la [distribution Skewed GED](#)

1. Loi Normale

La densité normale a longtemps été utilisée dans la littérature du fait de sa simplicité. Pour autant, il est aujourd'hui reconnu que les propriétés de la loi normale, symétrique et mésokurtique, ne sont pas compatibles avec les faits stylisés (distribution conditionnelle leptokurtique et asymétrique) observés généralement sur les séries de rendements des actifs financiers ([Cont, 2001](#)). Notons toutefois, que supposer à tort la normalité peut tout de même conduire à des estimations convergentes des paramètres d'un modèle GARCH (principe du pseudo-maximum de vraisemblance), bien que non efficaces. Mais, dans le cas d'une application à la VaR, la spécification de la loi conditionnelle ne concerne pas uniquement le problème de l'estimation des paramètres du modèle GARCH et affecte plus directement la détermination du fractile de la distribution conditionnelle. Le choix à tort d'une spécification normale peut dès lors avoir des conséquences notables sur les estimations et les prévisions de VaR.

Les formules utiles dans une perspective de calcul de la VaR associées à la loi normale centrée réduite sont les suivantes :

$$f(z) = \phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$$

$$F(z) = \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{t^2}{2}\right) dt$$

$$F^{-1}(\alpha) = \Phi^{-1}(\alpha)$$

$$E|z| = \sqrt{2/\pi}$$

$$\gamma_3 = E[(z - E(z))^3] = 0$$

$$\gamma_4 = E[(z - E(z))^4] = 3$$

Où γ_3 et γ_4 désignent respectivement la Skewness et la Kurtosis.

La log-vraisemblance associée à une observation z_t sous l'hypothèse de normalité s'écrit :

$$\text{Log } L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log[h_t(\theta)] - \frac{1}{2} \sum_{t=1}^T \frac{[z_t - m_t(\theta)]^2}{h_t(\theta)}$$

où $h_t(\theta)$ désigne la variance conditionnelle d'un processus GARCH et $m_t(\theta)$ fait référence à l'espérance conditionnelle.

2. Loi de Student

La distribution de Student permet de modéliser des queues de distribution plus épaisses que celles de la loi normale (distribution leptokurtique). Plus précisément, la Kurtosis de la distribution de Student est déterminée par le degré liberté ν . Dès lors, dans le cadre des modèles GARCH, ce paramètre estimé permet de capturer l'excès de Kurtosis qui ne peut pas être expliqué par le modèle GARCH lui-même. La distribution de Student standardisée est symétrique et la Skewness est nulle à condition que $\nu > 3$. La distribution est leptokurtique dès lors que $\nu > 4$.

Les formules utiles dans une perspective de calcul de la VaR associées à une loi de Student à ν degrés de liberté sont les suivantes ([Bao, Lee et Saltogu, 2004](#)).

$$f(z; \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\pi\nu}} \left(1 + \frac{z^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \nu > 2$$

$$F(z; \nu) = G(z; \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\pi\nu}} \int_{-\infty}^x \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} dt, \nu > 2$$

$$F^{-1}(\alpha) = \sqrt{\frac{(\nu-2)}{\nu}} G^{-1}(\alpha; \nu)$$

$$E|z| = \frac{4\Gamma\left(\frac{1+\nu}{2}\right)\sqrt{\nu-2}}{1 + \sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)(\nu-1)}$$

$$\gamma_3 = 0, \nu > 3$$

$$\gamma_4 = \frac{3\nu^2}{(\nu-2)(\nu-4)}, \nu > 4$$

où γ_3 et γ_4 désignent respectivement la Skewness et la Kurtosis.

La log-vraisemblance associée à une observation z_t sous l'hypothèse de Student s'écrit :

$$\log L(\theta, \varepsilon_t) = \log \left[\Gamma\left(\frac{\nu+1}{2}\right) \right] - \log \left[\Gamma\left(\frac{\nu}{2}\right) \right] - 0.5 \left\{ \left[\log[\pi(\nu-2)] + \log(h_t) + (1+\nu) \log \left(1 + \frac{z_t^2}{\nu-2} \right) \right] \right\}$$

où $\Gamma(\cdot)$ désigne la fonction gamma.

3. Loi GED (General Error Distribution)

La distribution à erreurs généralisées (GED) est également symétrique mais est assez flexible dans les queues grâce au paramètre ν . En effet, quand $\nu = 2$, nous tombons sur une distribution normale standard. Lorsque $\nu < 2$, cette distribution a des queues plus épaisses que celle de la normale (distribution leptokurtique). Pour $\nu = 1$, la distribution GED dégénère vers une distribution double exponentielle. Lorsque ce paramètre ν est supérieur à 2, la distribution GED a des queues plus fines que celles de la normale (distribution platykurtique). Enfin, quand ν tend vers l'infini, on obtient une distribution uniforme. L'excès de Kurtosis est donné par $\Gamma(1/\nu) \Gamma(5/\nu) / [\Gamma(3/\nu)]^2$, ([Bao, Lee et Saltogu, 2004](#)).

Les formules utiles dans une perspective de calcul de la VaR associées à une loi GED à ν degrés de liberté sont les suivantes.

$$f(z; \nu) = \frac{\nu \exp(-0.5|z/\lambda|^\nu)}{\lambda 2^{(1+1/\nu)} \Gamma(1/\nu)}, \nu > 0$$

$$F(z; \nu) = \begin{cases} \frac{\Gamma\left(1/\nu, \frac{1}{2}\left(-\frac{z}{\lambda}\right)^\nu\right)}{2\Gamma(1/\nu)} & \text{si } z \leq 0 \\ \frac{1}{2} + \frac{\gamma\left(1/\nu, \frac{1}{2}\left(\frac{z}{\lambda}\right)^\nu\right)}{2\Gamma(1/\nu)} & \text{si } z > 0 \end{cases}$$

$$F^{-1}(\alpha) = \begin{cases} -2\lambda\Gamma^{-1}(1/\nu, 2\alpha\Gamma(1/\nu)) & \text{si } \alpha \leq 0.5 \\ 2\lambda\gamma^{-1}(1/\nu, 2(\alpha - \frac{1}{2})\Gamma(1/\nu)) & \text{si } \alpha > 0.5 \end{cases}$$

$$E|z| = \lambda 2^{1/\nu} \frac{\Gamma(2/\nu)}{\Gamma(1/\nu)}$$

$$\gamma_3 = 0$$

$$\gamma_4 = \frac{\Gamma(1/\nu)\Gamma(5/\nu)}{[\Gamma(3/\nu)]^2}$$

où γ_3 et γ_4 désignent respectivement la Skewness et la Kurtosis et λ une constante telle que :

$$\lambda = \sqrt{2^{(-2/\nu)} \Gamma(1/\nu) / \Gamma(3/\nu)}$$

La log-vraisemblance associée à une observation z_t sous l'hypothèse GED s'écrit :

$$\log L(\theta, z_t) = \log(\nu/\lambda) - 0.5 \left| \frac{z_t}{\lambda} \right|^\nu - (1 + \nu^{-1}) \log(2) - \log \left[\Gamma\left(\frac{1}{\nu}\right) \right] - 0.5 \log(h_t)$$

4. Loi Skewed Student

Bien que la GED permette une flexibilité dans la modélisation de la Kurtosis, elle ne peut pas modéliser la Skewness observée dans un grand nombre de séries financières. Or, négliger la Skewness peut conduire à une inférence biaisée du risque.

[Fernandez et Steel \(1998\)](#) ont été les premiers à proposer la Skewed Student comme une distribution à quatre paramètres où les paramètres régissant la location, la dispersion, la Skewness et la Kurtosis ont une interprétation intuitive.

Les formules utiles dans une perspective de calcul de la VaR associées à une loi de Skewed Student à v degrés de liberté sont les suivantes.

$$f(z; \xi, v) = \begin{cases} \frac{2s}{\xi + \xi^{-1}} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{\pi(v-2)}} \left(1 + \frac{\xi^2(sz+m)^2}{v-2}\right)^{-\frac{v+1}{2}} & \text{si } z \leq -\frac{m}{s} \\ \frac{2s}{\xi + \xi^{-1}} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \frac{1}{\sqrt{\pi(v-2)}} \left(1 + \frac{\xi^{-2}(sz+m)^2}{v-2}\right)^{-\frac{v+1}{2}} & \text{si } z > -\frac{m}{s} \end{cases}$$

$$F(z; \xi, v) = \begin{cases} \frac{2}{\xi^2 + 1} G(\xi(sz+m)\sqrt{\frac{v}{v-2}}; v) & \text{si } z \leq -\frac{m}{s} \\ 1 - \frac{2}{\xi^{-2} + 1} G(-\xi^{-1}(sz+m)\sqrt{\frac{v}{v-2}}; v) & \text{si } z > -\frac{m}{s} \end{cases}$$

$$F^{-1}(\alpha) = \begin{cases} \frac{\frac{1}{\xi} G^{-1}\left(\frac{\alpha}{2}(1+\xi^2); v\right) - m}{s} & \text{si } \alpha \leq \frac{1}{1+\xi^2} \\ \frac{-\xi G^{-1}\left(\frac{1-\alpha}{2}(1+\xi^{-2}); v\right) - m}{s} & \text{si } \alpha > \frac{1}{1+\xi^2} \end{cases}$$

$$E|z| = \frac{4\xi^2\Gamma\left(\frac{1+v}{2}\right)\sqrt{v-2}}{\xi + \frac{1}{\xi}\sqrt{\pi}\Gamma\left(\frac{v}{2}\right)(v-1)}$$

$$\gamma_3 = \frac{(v-2)^{3/2}(\xi^2-1)(\xi^4+1)\Gamma[(v-3)/2]}{\sqrt{\pi}\xi^3s^3\Gamma(v/2)} - \frac{m(m^2+3s)}{s^3}$$

$$\gamma_4 = \frac{3(\xi^5 + \xi^{-5})(v-2)}{(\xi + \xi^{-1})(v-4)s^4} + \frac{3m^2(m^2+2s)}{s^4} - \frac{4m(v-2)^{3/2}(\xi^2-1)(\xi^4+1)\Gamma[(v-3)/2]}{\sqrt{\pi}\xi^3s^4\Gamma(v/2)}$$

où $\xi > 0, v > 2, m = \Gamma((v-1)/2)\sqrt{(v-2)/\pi}(\xi - 1/\xi)/\Gamma(v/2), s = \sqrt{(\xi^2 + 1/\xi^2 - 1) - m^2}$

γ_3 et γ_4 désignent respectivement la Skewness et la Kurtosis si v est respectivement supérieur à 3 et 4. $G(\cdot; v)$ dénote la fonction de répartition d'une distribution de Student standard. ξ est un

indicateur d'asymétrie tel que lorsque $\xi = 1$, la distribution Skewed Student est égale à la distribution de Student.

La log-vraisemblance associée à une observation z_t sous l'hypothèse Skewed Student s'écrit :

$$\log(\theta, z_t) = T \left[\log \Gamma\left(\frac{\nu+1}{2}\right) - \log\left(\frac{\nu}{2}\right) - \frac{1}{2} \log(\pi(\nu-2)) + \log\left(\frac{2}{\xi + \frac{1}{\xi}}\right) + \log(s) \right] - \frac{1}{2} \sum_{t=1}^T \left[\log(\sigma_t^2) + (1+\nu) \log\left(1 + \frac{(sz_t + m)^2}{(\nu-2)\xi^{-2t}}\right) \right]$$

5. Loi Skewed GED

[Theodossiou \(2000\)](#) donne la version Skewed de la GED avec deux paramètres contrôlant respectivement la Skewness et la Kurtosis. La Skewed GED était utilisé pour examiner la distribution empirique des rentabilités de plusieurs actifs financiers à des fréquences différentes et les résultats ont montré qu'elle est appropriée ([Bao, Lee et Saltoglu, 2004](#)).

Les formules utiles dans une perspective de calcul de la VaR associées à une loi de Skewed GED à ν degrés de liberté sont les suivantes.

$$f(z; k, \lambda) = \begin{cases} c \exp(-w_1 |z - \delta|^k) & \text{si } z \leq \delta \\ c \exp(-w_2 |z - \delta|^k) & \text{si } z > \delta \end{cases}$$

$$F(z; k, \lambda) = \begin{cases} \frac{(1-\lambda)\Gamma(1/k, w_1(\delta - z)^k)}{2\Gamma(1/k)} & \text{si } z \leq \delta \\ \frac{1-\lambda}{2} + \frac{(1+\lambda)\gamma(1/k, w_2(z - \delta)^k)}{2\Gamma(1/k)} & \text{si } z > \delta \end{cases}$$

$$F^{-1}(\alpha) = \begin{cases} \delta - \frac{1-\lambda}{\phi} \left[\Gamma^{-1}\left(\frac{1}{k}, \frac{2\alpha\Gamma(1/k)}{1-\lambda}\right) \right]^{1/k} & \text{si } \alpha \leq \frac{1-\lambda}{2} \\ \delta + \frac{1+\lambda}{\phi} \left[\gamma^{-1}\left(\frac{1}{k}, \frac{2\left(\alpha - \frac{1-\lambda}{2}\right)\Gamma(1/k)}{1+\lambda}\right) \right]^{1/k} & \text{si } \alpha > \frac{1-\lambda}{2} \end{cases}$$

$$E|z| = \begin{cases} c(-A_1 - B_1 + C_1) & \text{si } \delta \leq 0 \\ c(-A_2 + B_2 + C_2) & \text{si } \delta > 0 \end{cases}$$

$$\gamma_3 = \frac{4\lambda(1 + \lambda^2)\Gamma(4/k)}{\phi^3\Gamma(1/k)} + 3\delta + \delta^3$$

$$\gamma_4 = \frac{(1 + 10\lambda^2 + 5\lambda^4)\Gamma(5/k)}{\phi^4\Gamma(1/k)} + \frac{16\lambda\delta(1 + \lambda^2)\Gamma(4/k)}{\phi^3\Gamma(1/k)} + 6\delta^2 + 3\delta^4 - 3$$

où $k > 0$, $1 > \lambda > -1$, et où γ_3 et γ_4 désignent respectivement la Skewness et la Kurtosis.

$$c = \frac{k\phi}{2\Gamma(1/k)}, \phi = s\sqrt{\frac{\Gamma(3/k)}{\Gamma(1/k)}}, \delta = -2\lambda As^{-1}$$

$$s = \sqrt{1 + 3\lambda^2 - 4A^2\lambda^2}, A = \frac{\Gamma(2/k)}{\sqrt{\Gamma(1/k)\Gamma(3/k)}}$$

Le paramètre k contrôle l'épaisseur des queues tandis que λ modélise l'asymétrie quand il est différent de 0.

Par ailleurs, il est possible de retrouver plusieurs densités connues comme des cas particuliers. Ainsi, lorsque $k = 2$ et $\lambda = 0$, nous obtenons la densité de la loi Normal standard. Pour $k = 2$ et $\lambda \neq 0$, nous obtenons la Skewed normal d' [Azzalini \(1985\)](#). Enfin pour lorsque le paramètre $\lambda = 0$, nous retombons sur une distribution GED ([Capuccio, Lubian et Raggi, 2004](#))

Rédacteurs : [Chevreau Antoine](#), [Godin Sylvain](#), [Ivanof Mihaela](#) et [Patin Antoine](#)

Correction : [Hurlin Christophe](#)

Master ESA : Novembre 2006

Master Econométrie et Statistique Appliquée

Université d'Orléans

Site Value-at-Risk

http://193.49.79.89/esa_prof/index.php



Fiche 13. Méthode Simulation Historique

1. Présentation de la méthode

La popularité de la VaR Simulation Historique ou **VaR-HS** provient pour une bonne part de la difficulté inhérente à la mise en oeuvre de certaines méthodes de calcul sophistiquées de la VaR (ou risque de marché) d'un portefeuille de valeurs, comme par exemple la VaR Monte Carlo. Ces contraintes d'ordre techniques conditionnent en effet le choix de méthodes plus simples à mettre à oeuvre, dont les VaR non paramétriques en générale (HS ou les méthodes Hybrides).

La VaR-HS est basée sur une hypothèse assez forte, à savoir celle de distributions *i.i.d* (indépendantes et identiquement distribuées). Avec une telle méthode de calcul, on suppose donc que la distribution conditionnelle des rentabilités est invariante et peut alors être simulée par les rentabilités passées. Par conséquent, la VaR (ou de façon équivalente le fractile latent de la distribution conditionnelle des rentabilités) est estimée par sa contrepartie empirique.

Une telle démarche n'est certes pas exempte d'inconvénients, et de nombreux travaux empiriques, mais aussi théoriques ont souligné les insuffisances de la Simulation Historique. On peut citer entre autres Boudoukh et al. (1998) et Pritsker (2001). L'hypothèse contestable dans la démarche telle exposée est celle de distributions *i.i.d*, car il est invraisemblable que les nombreux facteurs microstructurels et macroéconomiques concourant à la formation du prix d'un actif demeurent inchangés dans le temps. Cependant, la méthode reste très usitée de part sa simplicité. Il revient donc à chaque utilisateur d'une telle méthode de s'assurer via des tests appropriés de l'intérêt statistique d'une telle représentation des données, compte tenu des jeux de données disponibles.

2. Méthode de calcul

Formellement l'inférence de la VaR-HS à $\alpha\%$ se ramène au calcul du fractile empirique d'ordre α de la chronique des rentabilités passées sur une fenêtre de taille T_e , soit:

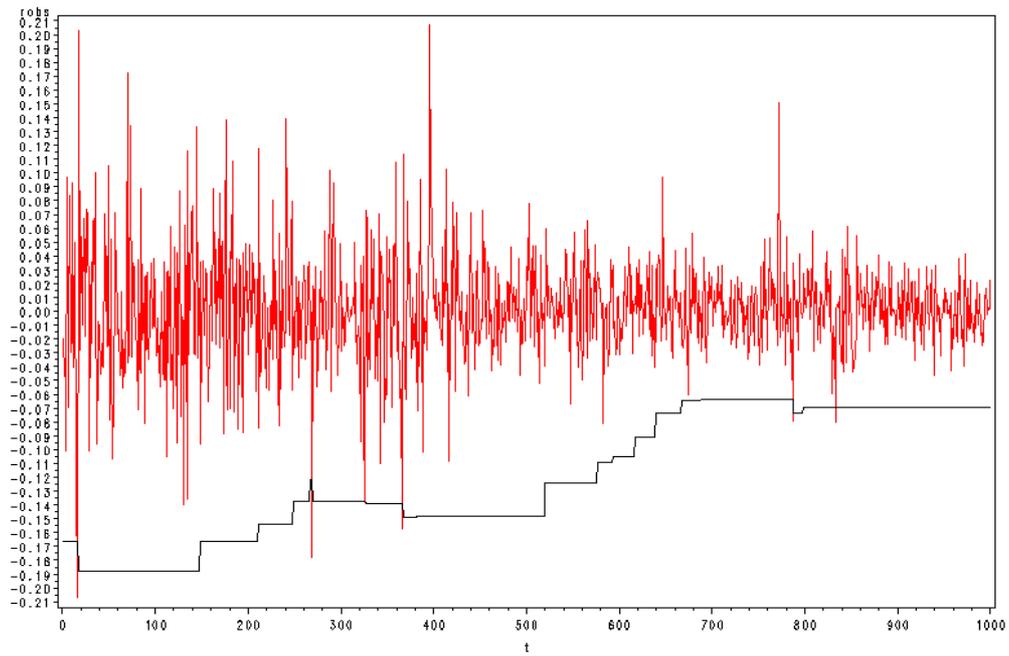
$$VaR_{t+1|t}(\alpha) = \text{percentile}\left(\{r_j\}_{j=t-T_e+1}^t, 100\alpha\right) \quad (1)$$

Généralement on pose T_e à 250 ou 500. De façon pratique pour calculer la VaR HS pour la période $t+1$, compte tenu de l'ensemble de l'information disponible en t , procédez comme suit:

- Choisir l'ensemble de l'information $\Omega_t = \{r_t, r_{t-1}, \dots, r_{t-T_e+2}, r_{t-T_e+1}\}$;
- Ordonner de manière croissante les T_e rentabilités de Ω_t ;
- La VaR à $\alpha\%$, correspond alors à la valeur de la rentabilité de rang $T\alpha\%$. Exemple: $T=500$ et $\alpha=1\%$, la VaR correspond à la valeur de la rentabilité de rang 5. Procédez par interpolation linéaire dans les cas où $T\alpha\%$ n'est pas un entier; Exemple: $T=250$ et $\alpha=1\%$.

3. Un exemple

Le graphique qui suit est une confrontation de la VaR-HS à 1% et des rentabilités observées de l'actif yahoo sur 1000 jours de trading, soit du 13/03/01 au 7/03/05. Pour chaque date appartenant à cette fenêtre, la valeur de la VaR pour la date suivante est générée en utilisant la formule (1). Il s'agit donc de VaR *out-of-sample*. On remarquera l'aspect discontinue de la courbe des VaR, typique des méthodes non paramétriques qui de ce fait contrastent généralement avec des '*VaR smoothed*' issues de méthodes paramétriques comme les modèles Garch.



Rédacteur : [Sessi Tokpavi](#)

Master ESA. 9 janvier 2006

Master Econométrie et Statistique Appliquée

Université d'Orléans

Site Value-at-Risk

http://193.49.79.89/esa_prof/index.php



Fiche 14. Méthode Simulation Historique Pondérée (WHS)

1. Présentation de la méthode

La Simulation Historique Pondérée (Weighted Historical Simulation) est une variante populaire de la Simulation Historique (HS). Rappelons qu'une critique essentielle de la technique HS est l'hypothèse *i.i.d* qu'on suppose implicitement pour la distribution des rentabilités (Voir fiche HS). Cependant comme le souligne Pritsker (2001), on peut toutefois s'accommoder du non respect de l'hypothèse de distributions *i.i.d*, en supposant que les observations les plus récentes de l'ensemble de l'information disponible $\Omega_t = \{r_t, r_{t-1}, \dots, r_{t-Te+2}, r_{t-Te+1}\}$ sont conditionnellement les plus importantes pour une éventuelle prévision :

"When returns are not i.i.d., it might be reasonable to believe that simulated returns from the recent past better represent today portfolio's risk than returns from the distant past...." (Pritsker, 2001, page 4).

La pondération relative des éléments de l'ensemble Ω_t , que suggère ce raisonnement se retrouve chez Boudoukh et al. (1998) qui proposent une variante à la Simulation Historique, à savoir la Simulation Historique Pondérée ou méthode Hybride. L'approche demeure non paramétrique, et à l'instar de la Simulation Historique se résume à assimiler la VaR prévue au fractile empirique des rentabilités identifiées par l'ensemble informationnel retenu. L'avantage ici est qu'on exploite une information supplémentaire, à savoir le caractère plus informatif des rentabilités les plus proches de l'horizon de prévision. Certes, cela peut se traduire par une réduction de l'ensemble de l'information aux seules observations les plus récentes, mais toute précision statistique s'en trouverait perdue, compte tenu de la taille ainsi réduite de Ω_t .

La méthode Hybride est alors issue d'un compromis entre ces deux impératifs (accorder plus d'informations aux observations les plus récentes, tout en considérant le caractère cyclique de la volatilité pour une plus grande précision dans la prévision).

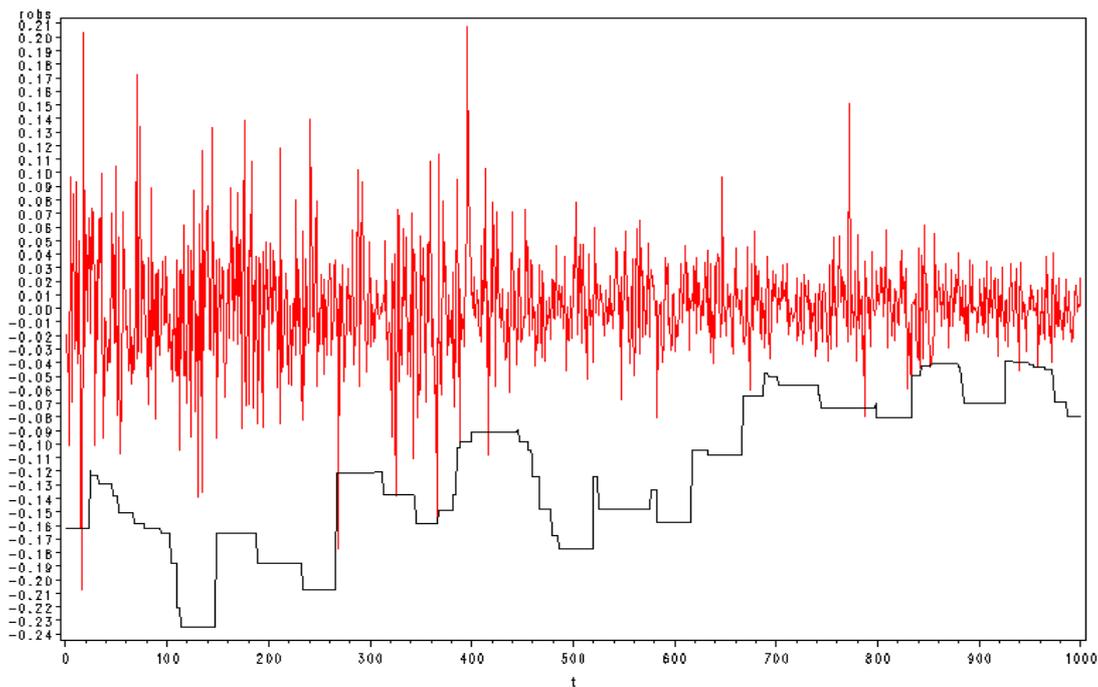
2. Méthode de calcul

Techniquement, le calcul de la VaR Hybride pour un taux de couverture de $\alpha\%$ se fait en trois étapes:

- A chacune des T_e rentabilités les plus récentes $r_t, r_{t-1}, \dots, r_{t-T_e+1}$ constituant l'ensemble Ω_t , est associée respectivement la série de pondérations décroissantes: $\frac{1-\lambda}{1-\lambda^{T_e}}$, $\frac{1-\lambda}{1-\lambda^{T_e}} \lambda$, \dots , $\frac{1-\lambda}{1-\lambda^{T_e}} \lambda^{T_e-1}$, $\frac{1-\lambda}{1-\lambda^{T_e}} \lambda^{T_e}$, avec λ compris entre 0 et 1 ;
- Les rentabilités sont ensuite ordonnées de manière croissante et ;
- Les poids sont sommés jusqu'à $\alpha\%$. La VaR étant égale à la rentabilité correspondant au dernier poids utilisé dans la sommation.

3. Un exemple

Le graphique qui suit est une confrontation de la VaR-Hybride à 1% et des rentabilités observées de l'actif yahoo sur 1000 jours de trading, soit du 13/03/01 au 7/03/05.



Pour chaque date appartenant à cette fenêtre, la valeur de la VaR pour la date suivante est générée en utilisant la démarche exposée ci-contre. On est donc en présence de VaR *out-of-sample*. Le paramètre de pondération λ est égal à 0.98 et $T_e = 250$.

Rédacteur : [Sessi Tokpavi](#)

Master ESA, 9 janvier 2006