

Inférence bayésienne des paramètres d'une distribution normale d'ordre q : Application à des données financières

Chouik Belmokhtar¹, ESC,
Anes Ouali², ENSSEA, LMSTD,

Résumé : La distribution normale d'ordre q peut présenter des queues plus épaisses que la loi normale classique et permet ainsi de modéliser des données financières à fréquence d'observation élevée. Une inférence statistique concernant le coefficient beta est un facteur important qui aide l'investisseur à prendre une fonction de décision plutôt qu'une autre. Une approche bayésienne est utilisée pour l'estimation des paramètres de position et d'échelle d'une distribution normale d'ordre q , sur la base d'un échantillon et d'une loi a priori gamma normale d'ordre q . A partir d'un cas empirique, nous estimons les moments a posteriori par la méthode de Rao-Blackwell basée sur le procédé d'échantillonnage de Gibbs.

Mots clés : Moments a posteriori, Distribution normale d'ordre q , Echantillonnage de Gibbs. Coefficient beta.

1. Introduction

Dans les modèles de régression linéaire, largement utilisés en économie, les erreurs sont généralement supposées suivre une distribution normale, alors que souvent cette normalité est mise en doute. Par exemple, dans le modèle d'évaluation des actifs financiers, les distributions empiriques des rendements présentent généralement des queues plus épaisses que celles d'une loi gaussienne. Ce phénomène souvent observé sur des grandeurs financières contenant des valeurs extrêmes est relativement difficile à modéliser par une distribution normale classique. Par conséquent, il est plus raisonnable de prendre une loi plus générale qui tient compte de ce phénomène, telle que la distribution normale d'ordre q dont la fonction de densité de probabilité est donnée par,

¹ Enseignant chercheur, ESC d'Alger chouikm@yahoo.fr

² Enseignant chercheur, INPS oualiances@yahoo.fr

$$f(x) = \left(\frac{1}{2\sigma q^q \Gamma(1 + (1/q))} \right)^{-1} \exp\left(-(q\sigma^q)^{-1}|x - \mu|^q\right)$$

(1)

Cette classe de lois, connue sous le nom EPD (Exponential Power Distribution) ou GED (Generalised Error Distribution), est introduite par Subbotin (1923) et a été popularisée par Box et Tiao (1973).

L'inférence statistique des paramètres de position et d'échelle d'une distribution normale d'ordre q sur la base d'un échantillon par les techniques usuelles d'estimation, à savoir maximum de vraisemblance, méthode des moments et l'approche bayésienne, a été abordée par plusieurs auteurs tels que, Box et Tiao (1973), Choy et Walker SG (2003) , Mineo et Ruggieri (2005). Dans un cadre bayésien, l'estimation des paramètres de position et d'échelle sur la base d'un échantillon iid aléatoire gaussien et d'une loi a priori conjointe gamma normale est souvent utilisée. Notre travail consiste à généraliser ce cas normal classique au cas normal d'ordre q afin d'appliquer les résultats obtenus à l'estimation bayésienne du coefficient beta d'un modèle MEDAF pour différentes valeurs de l'ordre q . Les résultats correspondants à la valeur q la mieux adaptée aux données seront choisis. Naturellement, le calcul des moments a posteriori est confronté à des problèmes d'intégration. Le logiciel informatique R est utilisé dans ce cas.

Dans le paragraphe 2, nous utilisons le procédé d'échantillonnage de Gibbs et les techniques de Rao-Blackwell pour estimer l'espérance et la variance a posteriori des paramètres de position et d'échelle. Pour une vérification de nos résultats théoriques, nous traitons des données simulées concernant le cas normal $q=2$ et le cas où l'ordre est différent de 2 dans le paragraphe 3. Enfin, dans le paragraphe 4, nous reprenons les données financières d'un modèle MEDAF citées par R.

J. Butler et co (1990) pour une inférence statistique du coefficient beta par la méthode bayésienne.

2. Moments a posteriori des paramètres de position et d'échelle

Si nous considérons un échantillon aléatoire $(X_1, \dots, X_n)$ issu d'une distribution normale d'ordre q , où q est un paramètre connu et supérieur à 1 et que la distribution a priori des paramètres de position et d'échelle est une gamma normale d'ordre q donnée par,

$$\pi(\mu, \sigma) = \frac{q^{n+1} \omega^{\nu/q}}{2\Gamma(1/q)\Gamma(\nu/q)} \sigma^{-(\nu+2)} \exp\left(-\frac{|\mu - m|^q + \omega}{q\sigma^q}\right),$$

(2) alors les lois a posteriori des paramètres sont les distributions marginales du vecteur aléatoire (μ, r, σ) dont la densité conjointe est donnée par,

$$f(\mu, \sigma, r) = (KJ)^{-1} \sigma^{-(n+\nu+q+2)} \exp\left(-\frac{r+\omega}{q\sigma^q}\right), \text{ pour } \mu \in R, \sigma > 0 \text{ et } S(\mu) < r \text{ et}$$

nulle ailleurs,

$$\text{où } \nu > 0, \quad \omega > 0,$$

$$K = q^{n+1} \Gamma\left(\frac{n+\nu+1}{q}\right), J = \int_{-\infty}^{\infty} (S(\mu) + \omega)^{-\frac{n+\nu+1}{q}} d\mu \text{ et } S(\mu) = \sum_{i=1}^{n+1} |x_i - \mu|^q \text{ avec}$$

$$x_{n+1} = m.$$

Nous notons que les distributions conditionnelles de ce vecteur sont usuelles, à savoir

(i) pour $r \geq r_0$, $\mu | r, \sigma \sim U_{[a(r), b(r)]}$, où, U est la loi uniforme et

$$[a(r), b(r)] = \{\mu \in R : S(\mu) < r\}$$

(ii) $(r - S(\mu)) | \mu, \sigma \sim \gamma\left(1, \frac{\sigma^{-q}}{q}\right)$ où, γ est la distribution gamma

(iii) $\sigma^q | r, \mu \sim i\gamma\left(\frac{n+\nu+1}{q}, \frac{r+\omega}{q}\right)$ où, $i\gamma$ est la distribution gamma inverse

Ceci nous permet d'utiliser le procédé d'échantillonnage de Gibbs afin de simuler selon les distributions a posteriori des paramètres de position et d'échelle. Pour un échantillon simulé, $(\mu^{(t)}, \sigma^{(t)}, r^{(t)})$ $t = 1, \dots, N$, les densités a posteriori et les moments a posteriori sont estimés par la méthode de Rao-Blackwell et sont donnés par,

$$\hat{f}(\mu|x) = N^{-1} \sum_{t=1}^N f_{\mu|r^{(t)}, \sigma^{(t)}}(\mu),$$

$$\hat{E}(\mu|x) = N^{-1} \sum_{t=1}^N \frac{a(r^{(t)}) + b(r^{(t)})}{2}, \quad \hat{V}(\mu|x) = N^{-1} \sum_{t=1}^N \frac{(b(r^{(t)}) - a(r^{(t)}))^2}{12},$$

$$\hat{f}(\sigma|x) = N^{-1} \sum_{t=1}^N f_{\sigma|\mu^{(t)}, r^{(t)}}(\sigma),$$

$$\hat{E}(\sigma|x) = N^{-1} \frac{\Gamma(\frac{1+\nu}{q} + 1)}{\Gamma(\frac{\nu+2}{q} + 1)} \sum_{t=1}^N ((r^{(t)} + \omega)/q)^{1/q},$$

$$\hat{V}(\sigma|x) = N^{-1} \frac{\Gamma(\frac{\nu}{q} + 1)}{\Gamma(\frac{\nu+2}{q} + 1)} \sum_{t=1}^N ((r^{(t)} + \omega)/q)^{2/q} - \hat{E}^2(\sigma|x).$$

3. Application à des données simulées.

Nous présentons le cas classique c'est-à-dire celui d'une loi normale d'ordre 2 puis le cas où l'ordre est différent de 2 à savoir 3.5. Chaque cas est illustré par un exemple.

3.1. Cas normal (q=2) :

Nous supposons que l'échantillon observé de taille 10 est (0.70, 0.32, -0.80, -0.93, 0.37, -0.65, 1.57, -0.32, 2.05, -3.08) et que les paramètres m, ν et ω sont égales à -0.02, 2 et 1 respectivement. Les valeurs exactes de la moyenne et de la variance a posteriori des paramètres de position et d'échelle sont données par, -0.0718 (0.1805) et 1.3743 (0.0966). Nous vérifions ces résultats en utilisant les résultats du paragraphe 2. En effet, nous obtenons à partir de la

méthode de Rao-Blackwell une estimation des moments a posteriori des paramètres $E(\mu|x) = -0.0717$, $V(\mu|x) = 0.1461$, $E(\sigma|x) = 1.3043$ et $V(\sigma|x) = 0.0758$ après 2000 simulations (sans prendre les 500 premières simulations).

3.2. Cas d'ordre 3.5.

Nous supposons que $m = 2.44$, $\nu = 2$ et $\omega = 1$. L'échantillon observé est

(2.13, 2.16, 2.38, 2.41, 2.46, 2.63, 2.73, 2.76, 2.81, 2.83, 2.85, 2.87, 3.15, 3.43).

Nous obtenons à partir de la méthode de Rao-Blackwell une estimation des moments posteriori des paramètres $E(\mu|x) = 2.6985$, $V(\mu|x) = 0.0072$, $E(\sigma|x) = 0.5228$ et $V(\sigma|x) = 0.0046$,

après 1000 simulations (sans prendre les 100 premières simulations).

4. Application à un modèle empirique du marché basé sur MEDAF.

La littérature en finance a montré que le beta d'un titre constitue une caractéristique très importante de toute action ou portefeuille. Ce coefficient permet de mesurer la sensibilité du rendement du titre aux variations du rendement du marché. La valeur de beta peut être estimée par régression mais la statistique mathématique a fourni des travaux récents abordant l'inférence statistique concernant ce coefficient. Par exemple, Butler et co (1990) ont pris un exemple de modèle empirique du marché basé sur MEDAF (Modèle d'évaluation des actifs financiers) à la compagnie AMERICAN CAN dont les données mensuelles sont enregistrées de janvier 1982 jusqu'à décembre 1986 afin de lui appliquer le modèle normal d'ordre q par une méthode du quasi-maximum de vraisemblance.

Ces modèles d'évaluation des actifs financiers s'écrivent sous la forme,

$$(R - R_f)_t = C + \beta(R_m - R_f)_t + \varepsilon_t ,$$

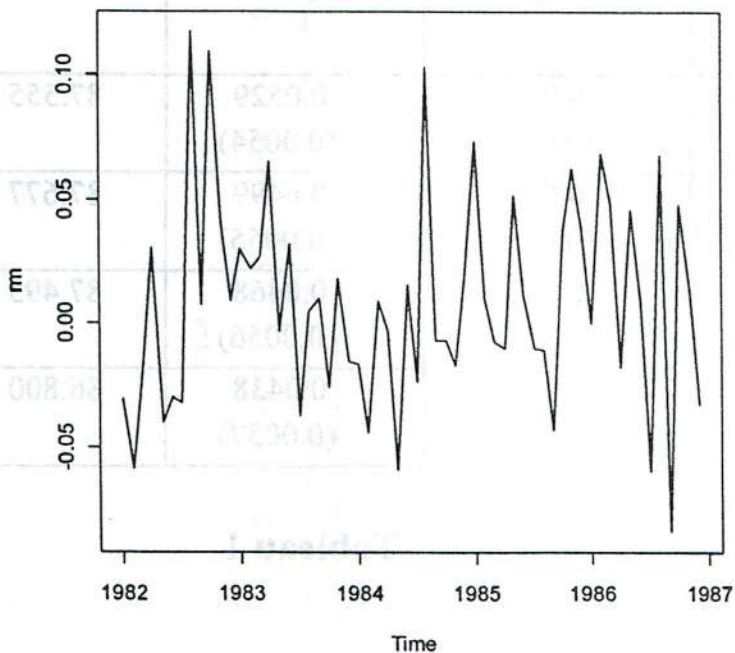
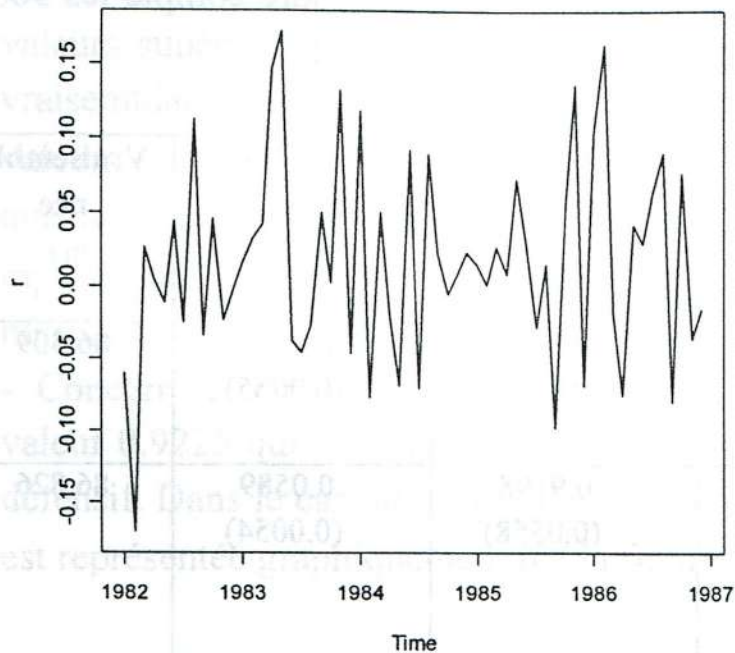
(3)

Où, $(r_i)_t = (R - R_f)_t$ est le rendement excédentaire de la compagnie pour le mois t

et $(r_m)_t = (R_m - R_f)_t$: est le rendement excédentaire du marché pour le mois t

R_f est le taux sans risque.

Les données mensuelles enregistrées de janvier 1982 jusqu'à décembre 1986 sont représentées graphiquement comme suit,



Nous reprenons cet exemple par une approche bayésienne pour une estimation des paramètres dont le coefficient β . Les estimations classiques et bayésiennes des coefficients de régression et du

paramètre d'échelle pour différentes valeurs du paramètre q sont obtenues

après un nombre total de 3000 itérations sans prendre compte les 500 premières itérations et sont données par le tableau 1.

	Constante C	Coefficient beta β	Paramètre d'échelle σ	Vraisemblance $\ln L$
Cas normal d'ordre $q > 2$ ($q = 2.2$)	0.0105 (0.0076)	0.9191 (0.0586)	0.0623 (0.0055)	86.309
Cas normal classique ($q = 2$)	0.0095 (0.0072)	0.9198 (0.0558)	0.0589 (0.0054)	86.826
Cas normal d'ordre $1 \leq q < 2$ ($q = 1.8$)	0.0087 (0.0070)	0.9207 (0.0562)	0.0559 (0.0054)	87.247
($q = 1.6$)	0.0081 (0.0068)	0.9215 (0.0545)	0.0529 (0.0054)	87.555
($q = 1.4$)	0.0078 (0.0067)	0.9224 (0.0548)	0.0499 (0.0055)	87.677
($q = 1.2$)	0.0077 (0.0065)	0.9233 (0.0548)	0.0468 (0.0056)	87.495
($q = 1$)	0.0080 (0.0065)	0.9244 (0.0538)	0.0438 (0.0057)	86.800

Tableau 1

A partir des résultats cités dans le tableau 1, nous donnons quelques commentaires :

- Il n'est pas raisonnable de modéliser ces données financières par un modèle normal d'ordre q où $q > 2$, en effet la valeur de la

vraisemblance correspondante au cas normal classique ($\ln L = 86.826$) est supérieure à celle donnée par le cas où le paramètre q prend des valeurs supérieures à 2. Par contre, nous notons que la valeur de la vraisemblance est supérieure à 86.826 lorsque l'ordre q prend des valeurs inférieures à 2 et supérieures à 1.2, ce qui nous permet de dire que les modèles normales d'ordre q ($1.2 < q < 2$) sont meilleurs par rapport au cas normal. Dans notre cas, nous choisissons le modèle normal d'ordre ($q = 1.4$) car il est le plus adapté aux données.

- Concernant l'estimation du coefficient beta, nous choisissons la valeur 0.9225 qui inférieure à 1 et donc nous notons que le titre est défensif. Dans le cas où $q = 1.4$, la convergence de l'estimateur de beta est représentée graphiquement par la figure 1.

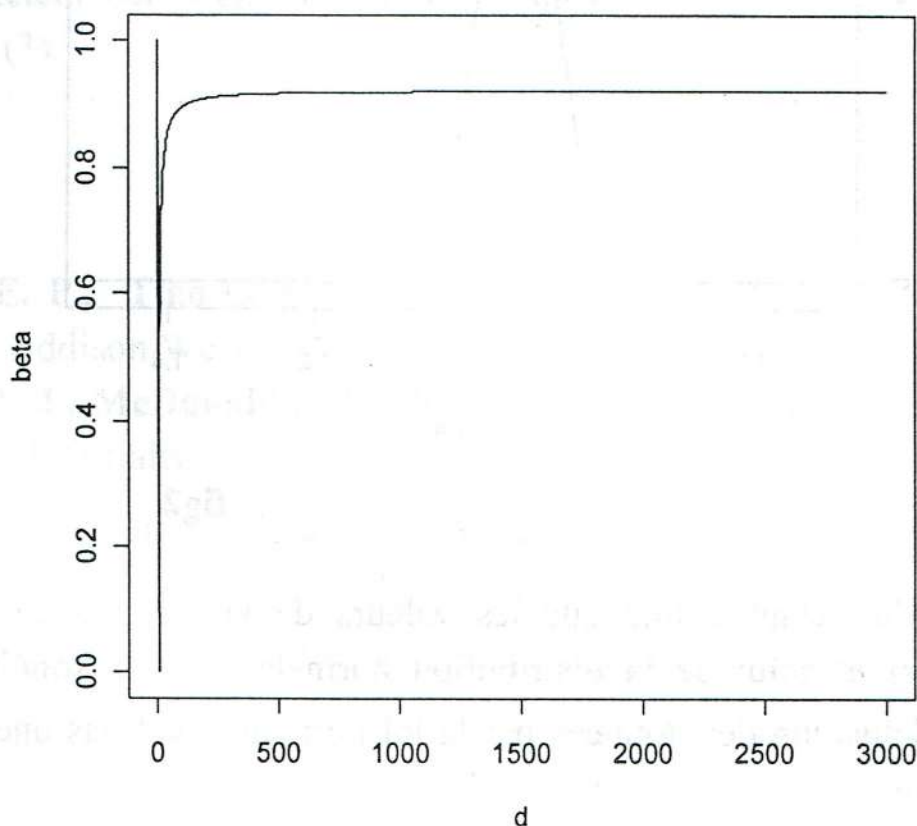


fig1

- L'approche bayésienne nous permet en plus d'estimer les densités marginales des paramètres, construire des intervalles de confiance et des tests sur le coefficient β .

Le graphe de la densité a posteriori estimée du coefficient β ($\text{deb}\beta$) correspondante à $q = 1.4$ est donnée par la figure 2.

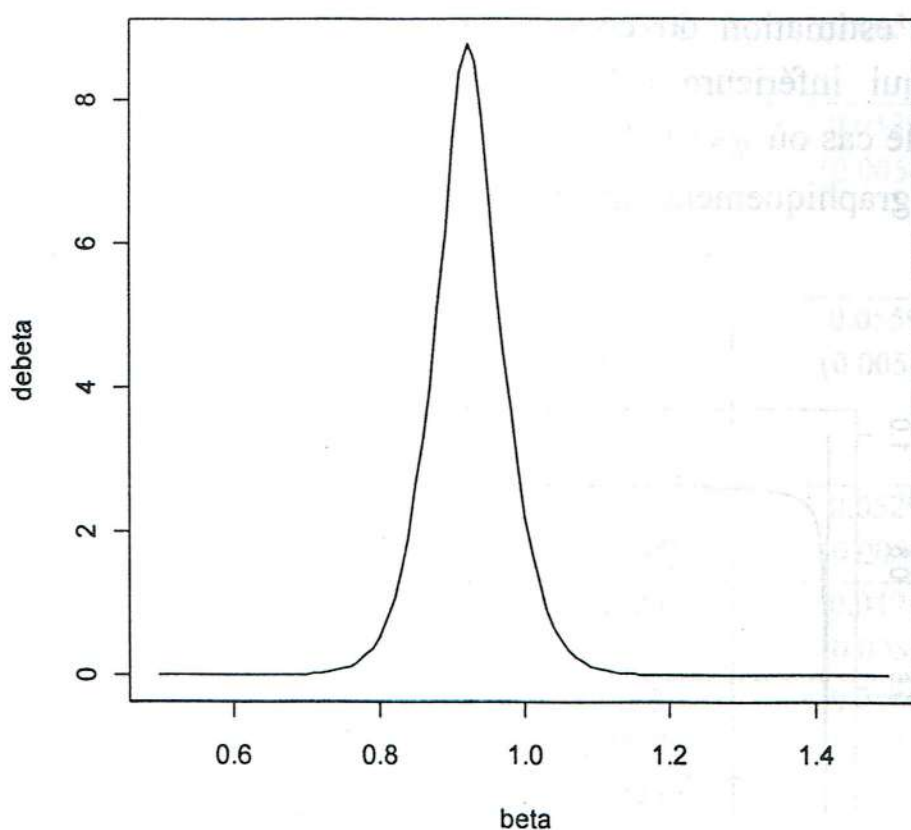


fig2

- Enfin, étant donné que les valeurs de vraisemblance du modèle choisi et celui de la distribution normale ($q = 2$) sont proches, la modélisation des données par la loi normale n'est pas une décision à rejeter.

5. Conclusion.

Dans notre travail, nous avons estimé les paramètres de position et d'échelle sous une distribution normale d'ordre q où q est supérieur à 1 par une approche bayésienne. A partir de l'exemple 3.1, nous avons pu constater que les moments a posteriori des paramètres calculés à partir des résultats du paragraphe 2 coïncident avec les moments a posteriori exacts dans le cas normal. D'autre part, dans l'exemple 3.2, où nous pouvons pas obtenir une expression analytique exacte des moments a posteriori, nous avons déterminé d'une manière similaire une approximation des moments a posteriori. Enfin, selon l'exemple des données financières considéré dans le paragraphe 4, nous avons obtenu une estimation bayésienne du coefficient beta meilleure que celle donnée par le cas classique normal.

Dans de futur recherche, nous considérons un modèle normal d'ordre q où q est inconnu et strictement positif pour mieux modéliser les données financières à valeurs extrêmes et l'appliquer à l'estimation du coefficient beta dans un modèle empirique du marché basé sur MEDAF (3).

Références

Box G. E. P., Tiao G. C. (1973): *Bayesian inference in statistical analysis*, Addison Wesley.

Butler R. J., McDonald J. B., Nelson R. D., White S. B. (1990): Robust and Partially

Adaptive Estimation of Regression Models, *The Review of Economics and Statistics*, 321-327.

Chouik B., Anes O. (2005): Régression linéaire sous une distribution normale d'ordre q .

Colloque Capital Humain, Croissance, Développement et Modélisation, Ecole Supérieure de Commerce d'Alger.

Choy S. T. B, Walker S. G. (2003): The extended exponential power distribution and Bayesian robustness, *Statistics & Probability Letters*, 65, 227-232.

Hamon J. (2005) : *Bourse et Gestion de Portefeuille*. Ed. Economica

Mineo A. M., Ruggieri, M. (2005): A Software Tool for the Exponential Power Distribution: The normal package, *Journal of Statistical Software*, Volume 12, Issue 4.

Références

Box G. E. P., Tiao G. C. (1993). *Bayesian Inference*

analysis, Addison Wesley.

Butler K. A., McDonald J. B., Nelson R. D. (1993). *Robust and Partially*

Adaptive Estimation of Regression Models. *The Annals of Statistics*

and Statistics 21, 1155-1174.

Choi K. B., An H. S. (1993). *Bayesian Inference*

and Statistics 21, 1155-1174.

Choi K. B., An H. S. (1993). *Bayesian Inference*

and Statistics 21, 1155-1174.

Choi K. B., An H. S. (1993). *Bayesian Inference*

and Statistics 21, 1155-1174.