Caractérisations de la loi gaussienne inverse généralisée : deux nouveaux résultats et une revue de la littérature

Angelo Efoévi Koudou [1] & Christophe Ley [2]

[1] *Université de Lorraine et CNRS, Institut Elie Cartan de Lorraine, UMR 7502, Vandoeuvre-lès-Nancy, F-54506, France, Efoevi.Koudou@univ-lorraine.fr*

[2] *ECARES et Département de Mathématique*
Université libre de Bruxelles
Boulevard du Triomphe, CP210, B-1050 Brussels, Belgium, chrisley@ulb.ac.be

Résumé. Nous proposons une revue succincte de la littérature concernant les diverses caractérisations de la loi gaussienne inverse généralisée (GIG) et donnons deux nouvelles caractérisations. La première est basée sur l'estimateur du maximum de vraisemblance du paramètre d'échelle. La deuxième est une caractérisation de type Stein.

Mots-clés. Loi gaussienne inverse généralisée, caractérisation par l'estimateur du maximum de vraisemblance, caractérisation de type Stein.

Abstract. We propose a brief survey of characterizations of the Generalized Inverse Gaussian (GIG) distribution, and establish two new characterizations, one based on maximum likelihood estimation of the scale parameter and the other is a Stein characterization.

Keywords. Generalized inverse Gaussian distribution, MLE characterization, Stein characterization.

1 Introduction

La loi gaussienne inverse généralisée (GIG) de paramètres $p \in \mathbb{R}$, $a > 0$, $b > 0$ a pour densité

$$f_{p,a,b}(x) := \frac{(a/b)^{p/2}}{2K_p(\sqrt{ab})} x^{p-1} e^{-(ax+b/x)/2}, \quad x > 0,$$

où K_p est la fonction de Bessel modifiée de troisième espèce. L'histoire de cette loi, encore appelée loi de Halphen de type A, remonte à Halphen (1941), Good (1953). .

La densité de la loi GIG peut être exprimée (voir Jørgensen, 1982) à l'aide des paramètres $\theta = \sqrt{ab}$ (paramètre de concentration) et $\eta = \sqrt{b/a}$ (paramètre d'échelle) :

$$f_{p,\theta,\eta}(x) := \frac{1}{2\eta K_p(\theta)} \left(\frac{x}{\eta}\right)^{p-1} e^{-\theta(x/\eta+\eta/x)/2}, \quad x > 0. \tag{1}$$

Elle intervient dans la modélisation de nombreux phénomènes réels, notamment la modélisation de données concernant des temps d'attente (Jørgensen, 1982), des phénomènes extrêmes en hydrologie (Chebana *et al*, 2010), l'activité neuronale (Iyengar *et al*, 1997). Concernant l'étude des propriétés statistiques de cette loi , on peut citer par exemple (Jørgensen, 1982), Perreault *et al* 1999a, 1999b).

Koudou et Ley (2013a) ont appliqué la méthodologie de Le Cam (voir Le Cam et Yang, 2000) pour construire des procédures de tests optimaux (au sens maximin) des paramètres de la loi GIG.

Nous proposons dans cette communication une revue succincte de la littérature concernant les diverses caractérisations de la loi GIG, sans prétendre à l'exhaustivité, et nous donnons deux nouvelles caractérisations. La première est basée sur l'estimateur du maximum de vraisemblance du paramètre d'échelle $\eta = \sqrt{b/a}$. La deuxième est une caractérisation de type Stein. Les caractérisations connues sont présentées dans la Section 2, et la Section 3 est consacrée aux deux nouvelles caractérisations. Cette communication est basée sur l'article Koudou et Ley (2013b).

2 Caractérisations existantes de la loi GIG

2.1 Caractérisation en termes de fractions continues

Letac et Seshadri (1983) ont caractérisé la loi GIG comme étant la loi d'une fraction continue constituée par des variables aléatoires indépendantes de loi gamma.

2.2 La propriété de Matsumoto-Yor

Nous notons $\Gamma(p, a/2) = \text{GIG}(p, a, 0)$ la loi gamma de densité proportionnelle à $x^{p-1}e^{-ax/2}$. Soient X et Y des variables aléatoires positives indépendantes telles que $X \sim \text{GIG}(-p, a, b)$ et $Y \sim \Gamma(p, a/2)$ avec $p, a, b > 0$. Selon la propriété de Matsumoto-Yor, les variables

$$U = \frac{1}{X+Y} \quad \text{et} \quad V = \frac{1}{X} - \frac{1}{X+Y}$$

sont indépendantes. Cette propriété a été remarquée par Matsumoto et Yor (2001) dans le cas $a = b$ au cours de l'étude de certaines fonctionnelles exponentielles du mouvement brownien. Letac et Wesołowski (2000) ont remarqué qu'elle est encore vérifiée si $a \neq b$, et ont démontré qu'elle caractérise la loi GIG (ou, plus exactement, le produit tensoriel de la loi GIG et de la loi gamma, avec des paramètres compatibles).

Pour $p = -1/2$, la propriété de Matsumoto-Yor peut être obtenue à partir d'une propriété d'indépendance établie par Barndorff-Nielsen et Koudou (1998) pour un réseau arborescent de résistances suivant la loi gaussienne inverse (voir Koudou, 2006).

Koudou et Vallois (2011, 2012) ont étudié une généralisation de la propriété de Matsumoto-Yor, en recherchant des fonctions régulières f ayant la propriété suivante : il existe des variables aléatoires positives indépendantes X et Y telles que les variables $U = f(X + Y)$ et $V = f(X) - f(X + Y)$ soient indépendantes. Cela a conduit à d'autres propriétés d'indépendance, concernant notamment la loi de Kummer.

2.3 Caractérisations de type régression

Wesołowski (2002), Seshadri et Wesołowski (2001) et Chou et Huang (2004) ont assoupli la condition d'indépendance des variables U et V ci-dessus, en supposant simplement que les espérances conditionnelles $\mathbb{E}(V|U)$ et $\mathbb{E}(1/V|U)$ sont constantes.

Lukacs (1956) a caractérisé la normalité de variables aléatoires $X_1, X_2, \ldots, X_n$ iid par la constance de la régression d'une fonction quadratique de $X_1, X_2, \ldots, X_n$ sur la somme $\Lambda = \sum_{i=1}^n X_i$. Le théorème suivant donne une caractérisation du même genre pour la loi GIG.

Theorem 2.1. *(Pusz, 1997) Soient $X_1, X_2, \ldots, X_n$ des copies iid d'une variable aléatoire $X > 0$ telle que les espérances $\mathbb{E}(1/X^2)$, $\mathbb{E}(1/X)$ et $\mathbb{E}(X)$ soient finies. Etant donnés $q > 0$ et $p \in \mathbb{R}$ tels que $p\mathbb{E}(1/X) + q\mathbb{E}(1/X^2) > 0$, soit $\Lambda = \sum_{i=1}^n X_i$ et*

$$S = \Lambda \left(p \sum_{i=1}^n \frac{1}{X_i} + q \sum_{i=1}^n \frac{1}{X_i^2} \right) - nq \sum_{i=1}^n \frac{1}{X_i}.$$

Alors $\mathbb{E}(S|\Lambda)$ est constante égale à c si et seulement s'il existe $\mu \in \mathbb{R}$, $a, b, \delta > 0$ tels que

$$p = \delta(\mu - 1),\ q = \delta b/2,\ c = \delta n(n\mu - 1)$$

et $X \sim \mathrm{GIG}(p, a, b)$.

2.4 Caractérisation par entropie

La caractérisation d'une loi par le principe de l'entropie maximale remonte à Shannon (1949) qui a montré que les gaussiennes maximizent l'entropie parmi les variables de moyenne et de variance données. D'autres exemples de telles caractérisations sont ensuite apparus dans la littérature. Par exemple, Kagan *et al* (1973) ont caractérisé un certain nombre de lois bien connues (par exemple les lois exponentielle, gamma et Beta) en termes d'entropie maximale sous des contraintes diverses. Par exemple : *(i)* la loi exponentielle de paramètre λ maximise l'entropie parmi toutes les lois d'une variable $X > 0$ sous la contrainte $\mathbb{E}(X) = 1/\lambda$; *(ii)* la loi gamma $\Gamma(p, a)$ maximise l'entropie sous les contraintes $\mathbb{E}(X) = p/a$ et $\mathbb{E}(\log X) = \Psi(p) - \log a$, où Ψ est la fonction Digamma; *(iii)* la loi Beta de paramètres $a, b > 0$ maximise l'entropie parmi toutes les lois d'une variable X de support $[0, 1]$ sous les contraintes $\mathbb{E}(\log X) = \Psi(a) - \Psi(a + b)$ et $\mathbb{E}(\log(1 - X)) = \Psi(b) - \Psi(a + b)$. Ce type de caractérisation existe aussi pour la loi GIG, comme mentionné dans Kawamura et Kōsei (2003).

3 Nouvelles caractérisations des lois GIG

3.1 Une caractérisation par l'estimateur du maximum de vraisemblance

Un théorème célèbre de caractérisation en statistique, dû à Carl Friedrich Gauss (1809), se formule ainsi : la moyenne $\bar{X}$ est, pour tout échantillon i.i.d. $X_1, \ldots, X_n$ de toute taille n, l'estimateur de maximum de vraisemblance (EMV) du paramètre μ dans une famille de densités $\{f(x - \mu), \mu \in \mathbb{R}\}$ si et seulement si l'échantillon est tiré d'une population gaussienne (de variance non fixée). Ce tout premier résultat de caractérisation par EMV a des applications importantes, puisqu'il indique clairement que plus l'on s'éloigne de la situation gaussienne, moins la moyenne est pertinente comme estimateur. D'autres théorèmes de caractérisation ont été établis par le passé, reliant des formes particulières de l'estimateur du maximum de vraisemblance à une loi particulière, par exemple la loi exponentielle, (Poincaré, 1912, Teicher, 1961), la loi de Laplace (Kagan *et al*, 1973), la loi gamma (Marshall et Olkin, 1993) ou la loi harmonique (Hürlimann, 1998). Duerinckx *et al.*(2013) effectue un inventaire de ces résultats et fournit une théorie unifiée.

Dans cette section, nous appliquons le résultat général de Duerinckx *et al.* (2013) pour construire un théorème de caractérisation par EMV de la loi GIG. Pour cela, nous adoptons la reformulation (1) de la densité, car cette paramétrisation fait de la famille $\{f_{p,\theta,\eta}(x),\, \eta > 0\}$ une famille d'échelle. Observons d'abord par un calcul simple que l'EMV de η, à p et θ fixés, est

$$\hat{\eta} = \frac{\sqrt{p^2 + \theta^2 \bar{X}\bar{X}_{-1}} - p}{\theta \bar{X}_{-1}},$$

où $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ et $\bar{X}_{-1} = \frac{1}{n}\sum_{i=1}^n \frac{1}{X_i}$ si $X_1, \ldots X_n$ sont i.i.d. de loi GIG(p,a,b). Nous obtenons le résultat suivant :

Theorem 3.1. *Soit un entier $n \geq 3$. Pour $p \in \mathbb{R}$ et $\theta > 0$, $\frac{\sqrt{p^2+\theta^2\bar{X}\bar{X}_{-1}}-p}{\theta\bar{X}_{-1}}$ est l'estimateur du maximum de vraisemblance du paramètre d'échelle η dans une famille $\left\{\frac{1}{\eta}f(x/\eta), \eta > 0\right\}$ sur $(0,\infty)$ pour tout échantillon $X_1, \ldots, X_n$ de taille fixée n si et seulement s'il existe $d > 0$ tel que la famille soit $\{f_{pd,\theta d,\eta}(x),\, \eta > 0\}$, i.e. la famille des densités des lois* GIG$(pd, \theta d/\eta, \theta\eta d)$*, $\eta > 0$.*

Il est clair que ce théorème de caractérisation par EMV pour la loi GIG contient, entre autres, les caractérisations obtenues en fixant des valeurs du paramètre p, par exemple les caractérisations par EMV des lois gaussienne inverse, hyperbolique ou harmonique. Ainsi, la caractérisation due à Hürlimann (1998) est un cas particulier de notre théorème 3.1.

3.2 Caractérisation à la Stein

La méthode de Stein pour l'approximation normale, introduite par Stein (1972), a été au fil des années adaptée à d'autres loi de probabilité comme, pour n'en citer que quelques unes, la loi de Poisson (Chen, 1975), la loi exponentielle (Chatterjee *et al*, 2011), la loi gamma (Luk, 1994).

La première étape de cette méthode consiste à trouver un "opérateur de Stein" approprié, dont les propriétés déterminent la qualité de l'approximation (voir par exemple Ross, 2011). Cet opérateur vérifie un *théorème de caractérisation de Stein* qui le relie à la loi cible.

Nous proposons cette caractérisation à la Stein pour la loin GIG dans le théorème suivant :

Theorem 3.2. *Une variable aléatoire positive X suit la loi GIG(p,a,b) si et seulement si pour toute fonction dérivable h vérifiant $\lim_{x\to\infty} f_{p,a,b}(x)h(x) = \lim_{x\to 0} f_{p,a,b}(x)h(x) = 0$, on a*

$$\mathbb{E}\left[h'(X) + \left(\frac{p-1}{X} + \frac{b}{2X^2} - \frac{a}{2}\right)h(X)\right] = 0.$$

La fonctionnelle $h \mapsto \mathcal{T}_{f_{p,a,b}}(h)(x) := h'(x) + \left(\frac{p-1}{x} + \frac{b}{2x^2} - \frac{a}{2}\right)h(x)$ est l'opérateur de Stein de la loi GIG.

Ce résultat est un exemple particulier de l'approche par densité des caractérisations de Stein, initiée dans Stein *et al* (2004) et développée par Ley et Swan (2013). On peut remplacer les fonctions $h(x)$ par $h(x)x^2$, et l'opérateur de Stein de la GIG $\mathcal{T}_{f_{p,a,b}}(h)$ prendrait alors la forme

$$\mathcal{T}_{f_{p,a,b}}(h)(x) = x^2 h'(x) + \left(-x^2\frac{a}{2} + (p+1)x + \frac{b}{2}\right)h(x).$$

Bibliographie

[1] Barndorff-Nielsen, O. E. and Koudou, A. E. (1998). Trees with random conductivities and the (reciprocal) inverse Gaussian distribution. *Adv. Appl. Probab.* **30**, 409–424.

[2] Chatterjee, S., Fulman, S. and Röllin, A. (2011). Exponential approximation by Stein's method and spectral graph theory. *ALEA Lat. Am. J. Probab. Math. Stat.* **8**, 197–223.

[3] Chebana, F., El Adlouni, S. and Bobée, B. (2010). Mixed estimation methods for Halphen distributions with applications in extreme hydrologic events. *Stoch. Environ. Res. Risk Assess.* **24**, 359–376.

[4] Chen, L. H. Y. (1975). Poisson approximation for dependent trials. *Ann. Probab.* **3**, 534–545.

[5] Chou, C.-W. and Huang, J.-W. (2004). On characterizations of the gamma and the generalized inverse Gaussian distributions. *Stat. Probab. Lett.* **69**, 381–388.

[6] Duerinckx, M., Ley, C. and Swan, Y. (2013). Maximum likelihood characterization of distributions. *Bernoulli,* to appear.

[7] Gauss, C. F. (1809). *Theoria motus corporum coelestium in sectionibus conicis solem ambientium.* Cambridge Library Collection. Cambridge University Press, Cambridge. Reprint of the 1809 original.

[8] Good, I. J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika* **40**, 237–260.

[9] Halphen, E. (1941). Sur un nouveau type de courbe de fréquence. *Comptes Rendus de l'Académie des Sciences* **213**, 633–635. Published under the name of "Dugué" due to war constraints.

[10] Hürlimann, W. (1998). On the characterization of maximum likelihood estimators for location-scale families. *Comm. Statist. Theory Methods* **27**, 495–508.

[11] Iyengar, S. and Liao, Q. (1997). Modeling neural activity using the generalized inverse Gaussian distribution. *Biol. Cybern.* **77**, 289–295.

[12] Jørgensen B. (1982). *Statistical Properties of the Generalized Inverse Gaussian Distribution.* Springer-Verlag, Heidelberg.

[13] Kagan, A. M., Linnik, Y. V. and Rao, C. R. (1973). *Characterization Problems in Mathematical Statistics.* Wiley, New York.

[14] Kawamura, T. and Kōsei, I. (2003). Characterisations of the distributions of power inverse Gaussian and others based on the entropy maximisation principle. *J. Japan Statist. Soc.* **33**, 95–104.

[15] Koudou, A. E. (2006). A link between the Matsumoto-Yor property and an independence property on trees. *Stat. Probab. Lett.* **76**, 1097–1101.

[16] Koudou, A. E. and Ley, C. (2013). Efficiency combined with simplicity: new testing procedures for Generalized Inverse Gaussian models. `arXiv : 1306.2776`.

[17] Koudou, A. E. and Ley, C. (2013). Characterizations of GIG laws: a survey complemented with two new results. $arXiv : 1312.7142[math.PR]$.

[18] Koudou, A. E. and Vallois, P. (2011). Which distributions have the Matsumoto-Yor property? *Electron. Commun. Prob.* **16**, 556–566.

[19] Koudou, A. E. and Vallois, P. (2012). Independence properties of the Matsumoto-Yor type. *Bernoulli* **18**, 119–136.

[20] Letac, G. and Seshadri, V. (1983). A characterization of the generalized inverse Gaussian distribution by continued fractions. *Z. Wahrsch. Verw. Geb.* **62**, 485–489.

[21] Letac, G. and Wesołowski, J. (2000). An independence property for the product of GIG and gamma laws. *Ann. Probab.* **28**, 1371–1383.

[22] Le Cam, L. and Yang, G. L. (2000). *Asymptotics in statistics. Some basic concepts.* 2nd ed. Springer-Verlag, New York.

[23] Ley, C. and Swan, Y. (2013). Stein's density approach and information inequalities. *Electron. Comm. Probab.* **18**, 1–14.

[24] Luk, H.-M. (1994). Stein's method for the gamma distribution and related statistical applications. Ph.D. thesis. University of Southern California. Los Angeles, USA.

[25] Lukacs, E. (1956). Characterization of populations by properties of suitable statistics. *Proc. Third Berkeley Symp. on Math. Statist. and Prob.*, Vol. 2 (Univ. of Calif. Press), 195–214.

[26] Marshall, A. W. and Olkin, I. (1993). Maximum likelihood characterizations of distributions. *Statist. Sinica* **3**, 157–171.

[27] Matsumoto, H. and Yor, M. (2001). An analogue of Pitman's $2M - X$ theorem for exponential Wiener functional, Part II: the role of the generalized inverse Gaussian laws. *Nagoya Math. J.* **162**, 65–86.

[28] Perreault, L., Bobée, B. and Rasmussen, P. F. (1999a). Halphen distribution system. I: Mathematical and statistical properties. *J. Hydrol. Eng.* **4**, 189–199.

[29] Perreault, L., Bobée, B. and Rasmussen, P. F. (1999b). Halphen distribution system. II: Parameter and quantile estimation. *J. Hydrol. Eng.* **4**, 200–208.

[30] Poincaré, H. (1912). *Calcul des probabilités.* Carré-Naud, Paris.

[31] Pusz, J. (1997). Regressional characterization of the Generalized inverse Gaussian population. *Ann. Inst. Statist. Math.* **49**, 315–319.

[32] Ross, N. (2011). Fundamentals of Stein's method. *Probab. Surv.* **8**, 210–293.

[33] Seshadri, V. and Wesołowski, J. (2001). Mutual characterizations of the gamma and the generalized inverse Gaussian laws by constancy of regression. *Sankhya Ser. A* **63**, 107–112.

[34] Shannon, C. E. (1949). *The Mathematical Theory of Communication.* Wiley, New York.

[35] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (Vol 2, pp. 586-602). Berkeley: University of California Press.

[36] Stein, C., Diaconis, P., Holmes, S. and Reinert, G. (2004). Use of exchangeable pairs in the analysis of simulations. In *Persi Diaconis and Susan Holmes, editors, Stein's method: expository lectures and applications, volume 46 of IMS Lecture Notes Monogr. Ser, pages 1–26. Beachwood, Ohio, USA: Institute of Mathematical Statistics.*

[37] Teicher, H. (1961). Maximum likelihood characterization of distributions. *Ann. Math. Statist.* **32**, 1214–1222.

[38] Wesołowski, J. (2002). The Matsumoto-Yor independence property for GIG and Gamma laws, revisited. *Math. Proc. Camb. Philos. Soc.* **133**, 153–161.