



HAL
open science

Efficiency combined with simplicity: new testing procedures for Generalized Inverse Gaussian models

Efoevi Angelo Koudou, Christophe Ley

► **To cite this version:**

Efoevi Angelo Koudou, Christophe Ley. Efficiency combined with simplicity: new testing procedures for Generalized Inverse Gaussian models. *Test*, 2014, 23 (4), pp.708-724. 10.1007/s11749-014-0378-2 . hal-01284109

HAL Id: hal-01284109

<https://hal.science/hal-01284109>

Submitted on 11 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Efficiency combined with simplicity: new testing procedures for Generalized Inverse Gaussian models

Angelo Efoevi Koudou · Christophe Ley

Received: date / Accepted: date

Abstract The standard efficient testing procedures in the Generalized Inverse Gaussian (GIG) family (also known as Halphen Type A family) are likelihood ratio tests, hence rely on Maximum Likelihood (ML) estimation of the three parameters of the GIG. The particular form of GIG densities, involving modified Bessel functions, prevents in general from a closed-form expression for ML estimators, which are obtained at the expense of complex numerical approximation methods. On the contrary, Method of Moments (MM) estimators allow for concise expressions, but tests based on these estimators suffer from a lack of efficiency compared to likelihood ratio tests. This is why, in recent years, trade-offs between ML and MM estimators have been proposed, resulting in simpler yet not completely efficient estimators and tests. In the present paper, we do not propose such a trade-off but rather an optimal combination of both methods, our tests inheriting efficiency from an ML-like construction and simplicity from the MM estimators of the nuisance parameters. This goal shall be reached by attacking the problem from a new angle, namely via the Le Cam methodology. Besides providing simple efficient testing methods, the theoretical background of this methodology further allows us to write out explicitly power expressions for our tests. A Monte Carlo simulation study shows that, also at small sample sizes, our simpler procedures do at least as good as the complex likelihood ratio tests. We conclude the paper by applying our findings on two real-data sets.

Angelo Efoevi Koudou
Université de Lorraine and CNRS, Institut Elie Cartan de Lorraine
UMR 7502, Vandoeuvre-lès-Nancy, F-54506, France

Christophe Ley
ECARES and Département de Mathématique, Université libre de Bruxelles
Campus Plaine, CP210, Boulevard du Triomphe, B-1050, Brussels, Belgium
Tel.: +32-2-650-5878
Fax: +32-2-650-5899
E-mail: chrisley@ulb.ac.be

Keywords Asymptotic linearity · GIG distributions · IG distributions · Maximin tests · Uniform local asymptotic normality

Mathematics Subject Classification (2000) Primary 62F03 · Secondary 62F05

1 Introduction.

The *Generalized Inverse Gaussian* (hereafter GIG) distribution with parameters $p \in \mathbb{R}$, $a > 0$, $b > 0$ has density

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

with normalizing constant

$$c(p, a, b) := \frac{(a/b)^{p/2}}{2K_p(\sqrt{ab})}$$

where K_p is the modified Bessel function of the third kind. The parameters a and b regulate both the concentration and scaling of the densities, the former via \sqrt{ab} and the latter via $\sqrt{b/a}$, while the parameter p bears no precise statistical meaning. This is why some authors rather use the parameterization $\theta = \sqrt{ab}$ and $\eta = \sqrt{b/a}$, but we stick here to a and b . The GIG family has been proposed in Good (1953), although Halphen (1941) had already previously discussed such distributions, which is why in some articles one speaks of Halphen Type A distributions instead of GIG distributions (see Seshadri 1999). GIG distributions enjoy several nice probabilistic features such as, e.g., the equivalence $X \sim \text{GIG}(p, a, b) \iff \frac{1}{X} \sim \text{GIG}(-p, b, a)$ (see Barndorff-Nielsen and Halgreen 1977, where some convolution properties and infinite divisibility are pointed out). These distributions have been used in the modelization of diverse phenomena such as, for instance, waiting time (Jørgensen 1982), neural activity (Iyengar and Liao 1997), or, most importantly, hydrologic extreme events (see Chebana *et al.* 2010 and references therein). The GIG family also contains several well-known sub-models such as the Gamma distribution (for $b = 0$ and $p > 0$), the reciprocal Gamma distribution (for $a = 0$ and $p < 0$), the hyperbolic distribution (for $p = 0$), the reciprocal inverse Gaussian (hereafter RIG) distribution (for $p = \frac{1}{2}$) and the *Inverse Gaussian* (hereafter IG) distribution (for $p = -\frac{1}{2}$).

While numerous papers address probabilistic aspects of GIG distributions, relatively few articles discuss their statistical properties. The cornerstone reference in this respect is Jørgensen (1982, Chapter 4-7), complemented by the stream of literature on general Halphen distributions (of which the GIG or Halphen A is one of three representatives), see e.g. Perreault *et al.* (1999a,b) or Chebana *et al.* (2010). Regarding hypothesis testing, the standard tools are likelihood ratio tests based on maximum likelihood (ML) estimation of the parameters both under the null and without constraints. Despite being the

most efficient procedures, ML-based methods suffer from computational complexity as in general no closed-form solutions exist for the likelihood equations in the GIG case (especially when all three parameters are estimated jointly). Improved numerical methods for ML estimation have been put forward by, e.g., Perreault *et al.* (1999b) and Lemonte and Cordeiro (2011). As alternative to the complex ML estimation, Fitzgerald (2000) has given exact expressions for method of moments (MM) estimators, which are thus analytically simple, but tests based on them suffer from a lack of efficiency compared to likelihood ratio tests. In order to find a trade-off between efficiency and simplicity, Chebana *et al.* (2010) propose mixtures of ML and MM estimation.

In the present paper, our aim is not a such a trade-off but rather an optimal combination of both approaches, the resulting tests inheriting efficiency from an ML-like construction and simplicity from MM estimation of the nuisance parameters under the null (no further non-null estimation will be required). We shall achieve this by having recourse to the *Le Cam methodology*, whose first step consists in showing that GIG distributions satisfy the *Uniform Local Asymptotic Normality (ULAN)* property. With this key property in hand, we build optimal test procedures for any null hypothesis involving one or more of the parameters p, a and b , the nuisance parameters remaining unspecified. The resulting test statistics resemble Rao score (or Lagrange Multiplier) statistics, hence are as efficient as the ML-based likelihood ratio tests, but improve on the Rao score tests by the fact that the nuisance parameters need not be ML-estimated under the null, but can be exactly calculated as MM-estimators. Yet another attractive feature of the Le Cam approach consists in the fact that we are able to calculate the power of our tests against sequences of contiguous alternatives.

The paper is organized as follows. In Section 2, we derive the ULAN property for GIG(p, a, b) models and the ensuing crucial *asymptotic linearity*. In Section 3 we show how to construct optimal tests for the null hypothesis $\mathcal{H}_0 : p = p_0$ against $\mathcal{H}_1^\neq : p \neq p_0$, determine their asymptotic behavior both under the null and under a sequence of contiguous alternatives and write out explicitly the asymptotic powers. In Section 4, we then particularize our findings to $p_0 = -1/2$, corresponding to testing for IG sub-models within GIG models (which yields for free a testing procedure for RIG distributions as $X \sim \text{IG}(a, b) \iff \frac{1}{X} \sim \text{RIG}(b, a)$). A Monte Carlo simulation study allows us to investigate the finite-sample behavior of our tests. We then use, in Section 6, our procedures in order to analyze two real-data examples. A brief outlook on future research is provided in Section 7. Finally, an Appendix collects the technical proofs.

2 The ULAN property and asymptotic linearity of general GIG(p, a, b) models.

Let X_1, \dots, X_n be i.i.d. observations following the GIG(p, a, b) distribution with density (1), and consider the GIG-parametric model

$$\mathcal{P}_{GIG}^{(n)} := \left\{ P_{p,a,b}^{(n)} : p \in \mathbb{R}, a \in \mathbb{R}_0^+, b \in \mathbb{R}_0^+ \right\},$$

where $P_{p,a,b}^{(n)}$ stands for the joint distribution of (X_1, \dots, X_n) . The ULAN property of the GIG(p, a, b) model is achieved in the following result, whose proof is given in the Appendix.

Theorem 21 *Let $\boldsymbol{\vartheta} = (p, a, b)'$ for any $p \in \mathbb{R}, a \in \mathbb{R}_0^+$ and $b \in \mathbb{R}_0^+$. Then, for any $\boldsymbol{\vartheta}^{(n)} = (p^{(n)}, a^{(n)}, b^{(n)})' = \boldsymbol{\vartheta} + O(n^{-1/2})$ and for any bounded sequence $\boldsymbol{\tau}^{(n)} = (\tau_1^{(n)}, \tau_2^{(n)}, \tau_3^{(n)})' \in \mathbb{R}^3$ such that $a^{(n)} + n^{-1/2}\tau_2^{(n)} > 0$ and $b^{(n)} + n^{-1/2}\tau_3^{(n)} > 0$, the family of probability distributions $\mathcal{P}_{GIG}^{(n)}$ satisfies*

$$\begin{aligned} \Lambda_{\boldsymbol{\vartheta}^{(n)} + n^{-1/2}\boldsymbol{\tau}^{(n)} / \boldsymbol{\vartheta}^{(n)}}^{(n)} &:= \log(dP_{\boldsymbol{\vartheta}^{(n)} + n^{-1/2}\boldsymbol{\tau}^{(n)}}^{(n)} / dP_{\boldsymbol{\vartheta}^{(n)}}^{(n)}) \\ &= (\boldsymbol{\tau}^{(n)})' \boldsymbol{\Delta}^{(n)}(\boldsymbol{\vartheta}^{(n)}) - \frac{1}{2}(\boldsymbol{\tau}^{(n)})' \boldsymbol{\Gamma}(\boldsymbol{\vartheta}) \boldsymbol{\tau}^{(n)} + o_P(1) \quad (2) \end{aligned}$$

and $\boldsymbol{\Delta}^{(n)}(\boldsymbol{\vartheta}^{(n)}) \xrightarrow{\mathcal{L}} \mathcal{N}_3(\mathbf{0}, \boldsymbol{\Gamma}(\boldsymbol{\vartheta}))$, both under $P_{\boldsymbol{\vartheta}^{(n)}}^{(n)}$ as $n \rightarrow \infty$, where

$$\begin{aligned} \boldsymbol{\Delta}^{(n)}(\boldsymbol{\vartheta}) &:= \begin{pmatrix} \Delta_p^{(n)}(\boldsymbol{\vartheta}) \\ \Delta_a^{(n)}(\boldsymbol{\vartheta}) \\ \Delta_b^{(n)}(\boldsymbol{\vartheta}) \end{pmatrix} := \frac{1}{\sqrt{n}} \sum_{i=1}^n \begin{pmatrix} \frac{\partial_p c(p,a,b)}{c(p,a,b)} + \log(X_i) \\ \frac{\partial_a c(p,a,b)}{c(p,a,b)} - \frac{X_i}{2} \\ \frac{\partial_b c(p,a,b)}{c(p,a,b)} - \frac{1}{2X_i} \end{pmatrix} \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \begin{pmatrix} \frac{1}{2} \log(a/b) - \partial_p \log K_p(\sqrt{ab}) + \log(X_i) \\ \frac{p}{2a} - \partial_a \log K_p(\sqrt{ab}) - \frac{X_i}{2} \\ -\frac{p}{2b} - \partial_b \log K_p(\sqrt{ab}) - \frac{1}{2X_i} \end{pmatrix} \end{aligned}$$

and

$$\boldsymbol{\Gamma}(\boldsymbol{\vartheta}) := \begin{pmatrix} \Gamma_{p,p}(\boldsymbol{\vartheta}) & \Gamma_{p,a}(\boldsymbol{\vartheta}) & \Gamma_{p,b}(\boldsymbol{\vartheta}) \\ \Gamma_{p,a}(\boldsymbol{\vartheta}) & \Gamma_{a,a}(\boldsymbol{\vartheta}) & \Gamma_{a,b}(\boldsymbol{\vartheta}) \\ \Gamma_{p,b}(\boldsymbol{\vartheta}) & \Gamma_{a,b}(\boldsymbol{\vartheta}) & \Gamma_{b,b}(\boldsymbol{\vartheta}) \end{pmatrix},$$

with

$$\begin{aligned} \Gamma_{p,p}(\boldsymbol{\vartheta}) &:= \partial_{pp}^2 \log K_p(\sqrt{ab}), \quad \Gamma_{a,a}(\boldsymbol{\vartheta}) := \frac{p}{2a^2} + \partial_{aa}^2 \log K_p(\sqrt{ab}), \\ \Gamma_{b,b}(\boldsymbol{\vartheta}) &:= -\frac{p}{2b^2} + \partial_{bb}^2 \log K_p(\sqrt{ab}), \quad \Gamma_{p,a}(\boldsymbol{\vartheta}) := -\frac{1}{2a} + \partial_{ap}^2 \log K_p(\sqrt{ab}), \\ \Gamma_{p,b}(\boldsymbol{\vartheta}) &:= \frac{1}{2b} + \partial_{bp}^2 \log K_p(\sqrt{ab}) \quad \text{and} \quad \Gamma_{a,b}(\boldsymbol{\vartheta}) := \partial_{ab}^2 \log K_p(\sqrt{ab}). \end{aligned}$$

In other words, the family $\mathcal{P}_{GIG}^{(n)}$ is ULAN at $\boldsymbol{\vartheta}$ with central sequence $\boldsymbol{\Delta}^{(n)}(\boldsymbol{\vartheta})$ and corresponding Fisher information matrix $\boldsymbol{\Gamma}(\boldsymbol{\vartheta})$.

The central idea of the Le Cam theory is the concept of convergence of statistical models (*experiments* in the Le Cam vocabulary). Quoting Le Cam (1960), “the family of probability measures under study can be approximated very closely by a family of a simpler nature”. The key ingredient in this approximation is the ULAN property, from which we can deduce that (see Le Cam and Yang 2000, page 89 for details) our GIG-parametric model $\mathcal{P}_{GIG}^{(n)}$ is locally (around $(p, a, b)'$) and asymptotically (for large sample sizes) equivalent to a simple Gaussian shift model. Intuitively, this is due to the fact that the likelihood ratio expansion (2), up to the remainder terms, looks like the likelihood ratio of a Gaussian shift model

$$\mathcal{P}_{\boldsymbol{\vartheta}} := \{P_{\boldsymbol{\tau}, \boldsymbol{\vartheta}} = \mathcal{N}_3(\boldsymbol{\Gamma}(\boldsymbol{\vartheta})\boldsymbol{\tau}, \boldsymbol{\Gamma}(\boldsymbol{\vartheta})) \mid \boldsymbol{\tau} \in \mathbb{R}^3\}$$

with a single observation which we denote as $\boldsymbol{\Delta}$. This means that all power functions that are implementable in the local GIG experiments are the power functions that are possible in the Gaussian shift experiment. In view of these considerations, it follows that asymptotically optimal tests in our local models can be derived by analyzing the Gaussian limit model, for which the most efficient procedures are well-known. The detailed construction is described in the next section. We conclude the present section by writing down an immediate consequence of the ULAN property, namely the following *asymptotic linearity* property of the central sequence:

$$\boldsymbol{\Delta}^{(n)}\left(\boldsymbol{\vartheta} + n^{-1/2}\boldsymbol{\tau}^{(n)}\right) = \boldsymbol{\Delta}^{(n)}(\boldsymbol{\vartheta}) - \boldsymbol{\Gamma}(\boldsymbol{\vartheta})\boldsymbol{\tau}^{(n)} + o_{\mathbb{P}}(1), \quad (3)$$

as $n \rightarrow \infty$, under $\mathbb{P}_{\boldsymbol{\vartheta}^{(n)}}^{(n)}$. This asymptotic linearity, combined with the ULAN property, forms the basis for our hypothesis test statistics, as it reveals us the behavior of the central sequence when $\boldsymbol{\vartheta} = (p, a, b)$ is replaced with some estimator $\hat{\boldsymbol{\vartheta}}^{(n)}$. As we shall show in details in the next section, for root- n consistent estimators satisfying some mild regularity assumption, we can replace $\boldsymbol{\tau}^{(n)}$ in (3) with $n^{1/2}(\hat{\boldsymbol{\vartheta}}^{(n)} - \boldsymbol{\vartheta})$, yielding

$$\boldsymbol{\Delta}^{(n)}\left(\hat{\boldsymbol{\vartheta}}^{(n)}\right) = \boldsymbol{\Delta}^{(n)}(\boldsymbol{\vartheta}) - \boldsymbol{\Gamma}(\boldsymbol{\vartheta})n^{1/2}(\hat{\boldsymbol{\vartheta}}^{(n)} - \boldsymbol{\vartheta}) + o_{\mathbb{P}}(1)$$

as $n \rightarrow \infty$ under $\mathbb{P}_{\boldsymbol{\vartheta}^{(n)}}^{(n)}$. This nice asymptotic equality will be put to use in what follows.

3 Construction of efficient testing procedures.

Based on our findings of the previous section, we shall describe in the present section how to tackle hypothesis testing problems in a way as efficient as likelihood ratio tests but with test statistics whose expressions can be explicitly written down, taking advantage of the exact MM expressions. We focus here on the construction of testing procedures for the null hypothesis $\mathcal{H}_0 : p = p_0$

against $\mathcal{H}_1^{\neq} : p \neq p_0$ for $p_0 \in \mathbb{R}$. This will be done in two steps: first, we assume that the nuisance parameters (here a and b) are known, and second, we consider them as unknown and hence they need to be estimated. Of course, the construction below is by no means restricted to the parameter p and can be used for hypothesis tests about a and b , about any of the vectors $(p, a)'$, $(p, b)'$ and $(a, b)'$, and even about the vector $(p, a, b)'$. The latter case is evidently the simplest, as the aforementioned second step will not be needed.

Step 1: the nuisance parameters a and b are known

In this scenario, it suffices to read the ULAN property only in the parameter p , and no asymptotic linearity is needed. By analogy with the Gaussian shift experiment, the optimal testing procedure for $p = p_0$ against $p \neq p_0$ consists in rejecting the null at asymptotic level α whenever the absolute value of the test statistic

$$\begin{aligned} Q_{p_0}^{(n)}(a, b) &:= \frac{\Delta_p^{(n)}(p_0, a, b)}{\sqrt{\Gamma_{p,p}(p_0, a, b)}} \\ &= \frac{n^{-1/2} \sum_{i=1}^n \left(\frac{1}{2} \log(a/b) - (\partial_p \log K_p(\sqrt{ab}))|_{p=p_0} + \log(X_i) \right)}{\sqrt{(\partial_{pp}^2 \log K_p(\sqrt{ab}))|_{p=p_0}}} \end{aligned}$$

exceeds the $\alpha/2$ -upper quantile $z_{\alpha/2}$ of a standard Gaussian distribution. We give at the end of the current section simplified forms for $\partial_p \log K_p(\sqrt{ab})$ and $\partial_{pp}^2 \log K_p(\sqrt{ab})$. The validity of this test, which we denote $\phi_{p_0}^{(n)}(a, b)$, and its optimality follow from the more general result in Theorem 31 below. As stated there, optimality is to be understood in the sense that our test is asymptotically *maximin*, a property derived from the Gaussian limit experiment. For the sake of clarity, we recall that a test ϕ^* is called maximin in the class \mathcal{C}_α of level- α tests for \mathcal{H}_0 against \mathcal{H}_1 if (i) ϕ^* has level α and (ii) the power of ϕ^* is such that

$$\inf_{P \in \mathcal{H}_1} \mathbb{E}_P[\phi^*] \geq \sup_{\phi \in \mathcal{C}_\alpha} \inf_{P \in \mathcal{H}_1} \mathbb{E}_P[\phi].$$

Step 2: the nuisance parameters a and b are unknown

This situation is clearly the less specific and more realistic one, but of course also more complicated. Indeed, since both a and b are unknown, we need to estimate them and plug the corresponding estimators $\hat{a}^{(n)}$ and $\hat{b}^{(n)}$ into the test statistic $Q_{p_0}^{(n)}(a, b)$ to yield $Q_{p_0}^{(n)}(\hat{a}^{(n)}, \hat{b}^{(n)})$. However, such a replacement cannot be achieved without care, and its asymptotic effects need to be calculated. A first observation can be made by looking at the information matrix $\mathbf{\Gamma}(\boldsymbol{\vartheta})$. Would $\mathbf{\Gamma}(\boldsymbol{\vartheta})$ be block-diagonal (in the sense that no correlation exists between the blocks p and $(a, b)'$), then the substitution of $(\hat{a}^{(n)}, \hat{b}^{(n)})'$ for $(a, b)'$ would have no influence, asymptotically, on the behavior of the central sequence $\Delta_p^{(n)}(\boldsymbol{\vartheta})$. However, this is obviously not the case here, and consequently a local

perturbation of a or b has the same asymptotic impact on $\Delta_p^{(n)}(\boldsymbol{\vartheta})$ as a local perturbation of p around p_0 . It follows that the cost of not knowing the true values of the nuisance parameters a and b is strictly positive; the stronger the correlation between the central sequences in p and $(a, b)'$, the larger that cost.

From this short discussion on the information matrix, it becomes clear that a more desirable central sequence for p would enjoy this block-diagonality, in other words, we need a “decorrelated” central sequence. This decorrelation takes into account the aforementioned cost of not knowing a and b and is achieved by the so-called *efficient central sequence* for p

$$\begin{aligned} & \Delta_p^{(n)eff}(\boldsymbol{\vartheta}) \\ := & \Delta_p^{(n)}(\boldsymbol{\vartheta}) - \text{Cov}(\Delta_p^{(n)}(\boldsymbol{\vartheta}), (\Delta_a^{(n)}(\boldsymbol{\vartheta}), \Delta_b^{(n)}(\boldsymbol{\vartheta}))) \left(\text{Var}((\Delta_a^{(n)}(\boldsymbol{\vartheta}), \Delta_b^{(n)}(\boldsymbol{\vartheta}))') \right)^{-1} (\Delta_a^{(n)}(\boldsymbol{\vartheta}), \Delta_b^{(n)}(\boldsymbol{\vartheta}))' \\ = & \Delta_p^{(n)}(\boldsymbol{\vartheta}) - \frac{(\Gamma_{p,a}(\boldsymbol{\vartheta}), \Gamma_{p,b}(\boldsymbol{\vartheta})) \begin{pmatrix} \Gamma_{b,b}(\boldsymbol{\vartheta}) & -\Gamma_{a,b}(\boldsymbol{\vartheta}) \\ -\Gamma_{a,b}(\boldsymbol{\vartheta}) & \Gamma_{a,a}(\boldsymbol{\vartheta}) \end{pmatrix} (\Delta_a^{(n)}(\boldsymbol{\vartheta}), \Delta_b^{(n)}(\boldsymbol{\vartheta}))'}{\Gamma_{a,a}(\boldsymbol{\vartheta})\Gamma_{b,b}(\boldsymbol{\vartheta}) - (\Gamma_{a,b}(\boldsymbol{\vartheta}))^2}. \end{aligned}$$

This efficient central sequence is obtained by projecting $\Delta_p^{(n)}(\boldsymbol{\vartheta})$ onto the subspace orthogonal to $(\Delta_a^{(n)}(\boldsymbol{\vartheta}), \Delta_b^{(n)}(\boldsymbol{\vartheta}))'$, which ensures that the new efficient central sequence $\Delta_p^{(n)eff}(\boldsymbol{\vartheta})$ is asymptotically uncorrelated with the central sequences corresponding to a and b . The Fisher information quantity associated with $\Delta_p^{(n)eff}(\boldsymbol{\vartheta})$ is given by

$$\begin{aligned} \Gamma_{p,p}^{eff}(\boldsymbol{\vartheta}) & := \Gamma_{p,p}(\boldsymbol{\vartheta}) - \frac{(\Gamma_{p,a}(\boldsymbol{\vartheta}), \Gamma_{p,b}(\boldsymbol{\vartheta})) \begin{pmatrix} \Gamma_{b,b}(\boldsymbol{\vartheta}) & -\Gamma_{a,b}(\boldsymbol{\vartheta}) \\ -\Gamma_{a,b}(\boldsymbol{\vartheta}) & \Gamma_{a,a}(\boldsymbol{\vartheta}) \end{pmatrix} (\Gamma_{p,a}(\boldsymbol{\vartheta}), \Gamma_{p,b}(\boldsymbol{\vartheta}))'}{\Gamma_{a,a}(\boldsymbol{\vartheta})\Gamma_{b,b}(\boldsymbol{\vartheta}) - (\Gamma_{a,b}(\boldsymbol{\vartheta}))^2} \\ & = \frac{\det \boldsymbol{\Gamma}(\boldsymbol{\vartheta})}{\Gamma_{a,a}(\boldsymbol{\vartheta})\Gamma_{b,b}(\boldsymbol{\vartheta}) - (\Gamma_{a,b}(\boldsymbol{\vartheta}))^2}. \end{aligned}$$

For further information about efficient central sequences (or efficient score functions) and their associated efficient information matrices, we refer the reader to Bickel *et al.* (1993, Chapter 2) or van der Vaart (1998, Chapter 25).

Now that we have derived this decorrelated efficient central sequence and its information matrix, it becomes of interest to establish their asymptotic linearity. This is achieved in the following Proposition, whose proof heavily relies on the asymptotic linearity (3).

Proposition 31 *For any $\boldsymbol{\vartheta} = (p, a, b)' \in \mathbb{R} \times (\mathbb{R}_0^+)^2$, and for any bounded sequence $\boldsymbol{\tau}^{(n)} = (0, \tau_2^{(n)}, \tau_3^{(n)})' \in \{0\} \times \mathbb{R}^2$ such that $a + n^{-1/2}\tau_2^{(n)} > 0$ and $b + n^{-1/2}\tau_3^{(n)} > 0$, we have that, under $\mathbb{P}_{\boldsymbol{\vartheta}}^{(n)}$ and as $n \rightarrow \infty$,*

$$\Delta_p^{(n)eff}(\boldsymbol{\vartheta} + n^{-1/2}\boldsymbol{\tau}^{(n)}) = \Delta_p^{(n)eff}(\boldsymbol{\vartheta}) + o_{\mathbb{P}}(1) \quad (4)$$

and

$$\Gamma_{p,p}^{eff}(\boldsymbol{\vartheta} + n^{-1/2}\boldsymbol{\tau}^{(n)}) = \Gamma_{p,p}^{eff}(\boldsymbol{\vartheta}) + o_{\mathbb{P}}(1). \quad (5)$$

The proof is deferred to the Appendix. As explained at the end of the previous section, the idea behind this asymptotic linearity result consists in using respectively $a + n^{-1/2}\tau_2^{(n)} = \hat{a}^{(n)}$ and $b + n^{-1/2}\tau_3^{(n)} = \hat{b}^{(n)}$, that is, $\tau_2^{(n)} = n^{1/2}(\hat{a}^{(n)} - a)$ and $\tau_3^{(n)} = n^{1/2}(\hat{b}^{(n)} - b)$. Both sequences remain bounded if the estimators $\hat{a}^{(n)}$ and $\hat{b}^{(n)}$ are root- n consistent, a very natural requirement. However, in order to perform this replacement (which is not evident as the non-random sequences $\tau_1^{(n)}$ and $\tau_2^{(n)}$ are replaced with random sequences), one more condition needs to be imposed on the estimators, and this is summarized in the following

ASSUMPTION A. *The sequence of estimators $\hat{a}^{(n)}$ and $\hat{b}^{(n)}$ is (i) root- n consistent (i.e., $n^{1/2}(\hat{a}^{(n)} - a) = O_P(1)$ and $n^{1/2}(\hat{b}^{(n)} - b) = O_P(1)$ as $n \rightarrow \infty$, under $P_{p,a,b}^{(n)}$) and (ii) locally asymptotically discrete, meaning that, for all $a, b \in \mathbb{R}_0^+$ and all $c > 0$, there exists an $M = M(c) > 0$ such that the number of possible values of $\hat{a}^{(n)}$ and $\hat{b}^{(n)}$ in intervals of the form $\{t \in \mathbb{R} : n^{1/2}|t - a| \leq c\}$ and $\{t \in \mathbb{R} : n^{1/2}|t - b| \leq c\}$ is bounded by M , uniformly as $n \rightarrow \infty$.*

It should be noted that Assumption A(ii) is a purely technical requirement, with little practical implications (for fixed sample size, any estimator indeed can be considered part of a locally asymptotically discrete sequence; see Le Cam and Yang 2000). This condition is however essential as it precisely allows to replace $\tau_2^{(n)}$ with $n^{1/2}(\hat{a}^{(n)} - a)$ and $\tau_3^{(n)}$ with $n^{1/2}(\hat{b}^{(n)} - b)$ in (4) and (5) by Lemma 4.4 in Kreiss (1987), where such replacements have been theoretically worked out in detail. Using this fact in combination with Slutsky's Lemma, we have the following crucial result.

Proposition 32 *Let Assumption A hold and fix $p_0 \in \mathbb{R}$. Then, defining*

$$Q_{p_0}^{(n)eff}(a, b) = \frac{\Delta_p^{(n)eff}(p_0, a, b)}{\sqrt{\Gamma_{p,p}^{eff}(p_0, a, b)}},$$

we have that

$$Q_{p_0}^{(n)eff}(\hat{a}^{(n)}, \hat{b}^{(n)}) - Q_{p_0}^{(n)eff}(a, b) = o_P(1) \quad (6)$$

under $P_{p_0,a,b}^{(n)}$ as $n \rightarrow \infty$.

This finally enables us to derive the optimal testing procedure for $\mathcal{H}_0 : p = p_0$ against $\mathcal{H}_1^\neq : p \neq p_0$ when the nuisance parameters a and b are unknown. This test, which we denote $\phi_{p_0}^{(n)}(\hat{a}^{(n)}, \hat{b}^{(n)})$, rejects the null hypothesis at asymptotic level α whenever the absolute value of the test statistic $Q_{p_0}^{(n)eff}(\hat{a}^{(n)}, \hat{b}^{(n)})$ exceeds $z_{\alpha/2}$, with $\hat{a}^{(n)}$ and $\hat{b}^{(n)}$ satisfying Assumption A. In the next theorem, we formally establish the validity of this test (behavior of $Q_{p_0}^{(n)eff}(\hat{a}^{(n)}, \hat{b}^{(n)})$ under the null), its asymptotic distribution under a sequence of local alternatives and its optimality features; of course, the results also cover the case of the easier test $\phi_{p_0}^{(n)}(a, b)$.

Theorem 31 Fix $p_0 \in \mathbb{R}$ and suppose that $\hat{a}^{(n)}$ and $\hat{b}^{(n)}$ satisfy Assumption A. Then

- (i) the test statistic $Q_{p_0}^{(n)eff}(\hat{a}^{(n)}, \hat{b}^{(n)})$ is asymptotically standard normal under $\bigcup_{a \in \mathbb{R}_0^+} \bigcup_{b \in \mathbb{R}_0^+} \mathbf{P}_{p_0, a, b}^{(n)}$;
- (ii) for any $a \in \mathbb{R}_0^+$ and any $b \in \mathbb{R}_0^+$, $Q_{p_0}^{(n)eff}(\hat{a}^{(n)}, \hat{b}^{(n)})$ is asymptotically normal with mean $\tau_1 \sqrt{\Gamma_{p, p}^{eff}(p_0, a, b)}$ and variance 1 under $\mathbf{P}_{p_0 + n^{-1/2} \tau_1^{(n)}, a, b}^{(n)}$ where $\tau_1^{(n)} \in \mathbb{R}$ is a bounded sequence and $\tau_1 := \lim_{n \rightarrow \infty} \tau_1^{(n)}$;
- (iii) the test $\phi_{p_0}^{(n)}(\hat{a}^{(n)}, \hat{b}^{(n)})$, which rejects the null hypothesis $\mathcal{H}_0 := \bigcup_{a \in \mathbb{R}_0^+} \bigcup_{b \in \mathbb{R}_0^+} \mathbf{P}_{p_0, a, b}^{(n)}$ whenever $|Q_{p_0}^{(n)eff}(\hat{a}^{(n)}, \hat{b}^{(n)})| > z_{\alpha/2}$, has asymptotic level α under \mathcal{H}_0 and is locally and asymptotically maximin for testing $\mathcal{H}_0 : p = p_0$ against $\mathcal{H}_1^\neq := \bigcup_{a \in \mathbb{R}_0^+} \bigcup_{b \in \mathbb{R}_0^+} \bigcup_{p \neq p_0} \mathbf{P}_{p, a, b}^{(n)}$.

Part (iii) of Theorem 31 provides the theoretical proof of efficiency of our test $\phi_{p_0}^{(n)}(\hat{a}^{(n)}, \hat{b}^{(n)})$, which is thus as efficient as the likelihood ratio test $\phi_{p_0}^{\text{LR}}$. However, the advantage of $\phi_{p_0}^{(n)}(\hat{a}^{(n)}, \hat{b}^{(n)})$ over $\phi_{p_0}^{\text{LR}}$ lies in its simplicity, as no ML estimation is required here since any root- n consistent estimators for a and b can be used. Hence we opt for their MM estimators, whose expressions, due to Fitzgerald (2000), are given by $\hat{a}_{\text{MM}}^{(n)} := \frac{\hat{\theta}_{\text{MM}}^{(n)}}{\hat{\eta}_{\text{MM}}^{(n)}}$ and $\hat{b}_{\text{MM}}^{(n)} := \hat{\theta}_{\text{MM}}^{(n)} \hat{\eta}_{\text{MM}}^{(n)}$ where

$$\hat{\eta}_{\text{MM}}^{(n)} := \sqrt{\frac{\bar{X}_{-1} s^2 - \bar{X}(\bar{X} \bar{X}_{-1} - 1)}{\bar{X}(s_{-1})^2 - \bar{X}_{-1}(\bar{X} \bar{X}_{-1} - 1)}}$$

and

$$\hat{\theta}_{\text{MM}}^{(n)} := \frac{2 \left(\frac{1}{\hat{\eta}_{\text{MM}}^{(n)}} \bar{X} - \hat{\eta}_{\text{MM}}^{(n)} \bar{X}_{-1} \right)}{\frac{1}{(\hat{\eta}_{\text{MM}}^{(n)})^2} s^2 - (\hat{\eta}_{\text{MM}}^{(n)})^2 (s_{-1})^2}$$

with $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, $s^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$, $\bar{X}_{-1} = \frac{1}{n} \sum_{i=1}^n \frac{1}{X_i}$ and $(s_{-1})^2 = \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{X_i} - \bar{X}_{-1} \right)^2$. Thus, *in fine*, the computationally simple yet efficient test we propose is $\phi_{p_0}^{(n)} := \phi_{p_0}^{(n)}(\hat{a}_{\text{MM}}^{(n)}, \hat{b}_{\text{MM}}^{(n)})$ which rejects the null hypothesis at asymptotic level α whenever the absolute value of

$$Q_{p_0}^{(n)} := Q_{p_0}^{(n)eff}(\hat{a}_{\text{MM}}^{(n)}, \hat{b}_{\text{MM}}^{(n)}) = \frac{\Delta_p^{(n)eff}(p_0, \hat{a}_{\text{MM}}^{(n)}, \hat{b}_{\text{MM}}^{(n)})}{\sqrt{\Gamma_{p, p}^{eff}(p_0, \hat{a}_{\text{MM}}^{(n)}, \hat{b}_{\text{MM}}^{(n)})}}$$

exceeds $z_{\alpha/2}$. One-sided tests of the form $\mathcal{H}_0 : p = p_0$ against $\mathcal{H}_1^> : p > p_0$ or $\mathcal{H}_1^< : p < p_0$ are of course readily constructed along the same lines and their asymptotic behavior is as well regulated by Theorem 31.

As mentioned in the Introduction, a welcomed feature of the Le Cam approach lies in the fact that it enables us to derive explicit power expressions under sequences of local alternatives of the form $\mathbf{P}_{p_0 + n^{-1/2} \tau_1^{(n)}, a, b}^{(n)}$ for any

$a, b \in \mathbb{R}_0^+$, with $\tau_1^{(n)}$ a bounded sequence with limit τ_1 . These powers are easily determined via Part (ii) of Theorem 31, as the latter provides us with the asymptotic distribution of $Q_{p_0}^{(n)eff}(\hat{a}^{(n)}, \hat{b}^{(n)})$ under that sequence of local alternatives. Denoting by Φ the cumulative distribution function of the standard Gaussian distribution, the asymptotic power of $\phi_{p_0}^{(n)eff}(\hat{a}^{(n)}, \hat{b}^{(n)})$ is given by

$$1 - \Phi \left(z_{\alpha/2} - \tau_1 \sqrt{\Gamma_{p,p}^{eff}(p_0, a, b)} \right) + \Phi \left(-z_{\alpha/2} - \tau_1 \sqrt{\Gamma_{p,p}^{eff}(p_0, a, b)} \right),$$

and by

$$1 - \Phi \left(z_{\alpha} - \tau_1 \sqrt{\Gamma_{p,p}^{eff}(p_0, a, b)} \right) \quad \text{and} \quad \Phi \left(-z_{\alpha} - \tau_1 \sqrt{\Gamma_{p,p}^{eff}(p_0, a, b)} \right)$$

in the respective one-sided tests against $\mathcal{H}_1^> : p > p_0$ and $\mathcal{H}_1^< : p < p_0$.

We conclude this section by providing more user-friendly expressions for the entries of the Fisher information matrix $\mathbf{\Gamma}(\boldsymbol{\theta})$. For j taking values in $\{p-3, p-2, p-1, p, p+1\}$, consider the integrals

$$\begin{aligned} I_j &:= \int_0^\infty u^p \exp \left(-\frac{1}{2}(au + b/u) \right) du, \\ M_j &:= \int_0^\infty u^p \log(u) \exp \left(-\frac{1}{2}(au + b/u) \right) du, \\ N_j &:= \int_0^\infty u^{p-1} (\log(u))^2 \exp \left(-\frac{1}{2}(au + b/u) \right) du. \end{aligned}$$

For the sake of readability, we omit writing $I_j(a, b)$, $M_j(a, b)$ and $N_j(a, b)$ and tacitly assume that each integral depends on a and b . The entries of the matrix $\mathbf{\Gamma}(\boldsymbol{\theta})$ can be written as

$$\begin{aligned} \Gamma_{p,p}(\boldsymbol{\theta}) &= \frac{N_p}{I_{p-1}} - \left(\frac{M_{p-1}}{I_{p-1}} \right)^2, \quad \Gamma_{p,a}(\boldsymbol{\theta}) = -\frac{1}{2} \left(\frac{M_p}{I_{p-1}} - \frac{I_p M_{p-1}}{I_{p-1}^2} \right), \\ \Gamma_{p,b}(\boldsymbol{\theta}) &= -\frac{1}{2} \left(\frac{M_{p-2}}{I_{p-1}} - \frac{M_{p-1} I_{p-2}}{I_{p-1}^2} \right), \quad \Gamma_{a,b}(\boldsymbol{\theta}) = \frac{1}{4} \left(1 - \frac{I_p I_{p-2}}{I_{p-1}^2} \right), \\ \Gamma_{a,a}(\boldsymbol{\theta}) &= \frac{1}{4} (I_{p+1} - I_p^2), \quad \text{and} \quad \Gamma_{b,b}(\boldsymbol{\theta}) = \frac{1}{4} (I_{p-3} - I_{p-2}^2). \end{aligned}$$

This way of writing the information matrix renders its calculation easier in programs such as R or MATHEMATICA; the R code is available from the authors upon request.

4 Efficient tests for (R)IG sub-models within the GIG model.

The most important sub-model nested by the GIG family is the inverse Gaussian distribution, obtained for $p = -\frac{1}{2}$. The IG distribution with its two positive parameters a and b has been introduced to the statistical community by Tweedie (1945, 1956, 1957) and has since been revealed to be quite convenient in modeling and analyzing observations that are right-skewed and positive (see Seshadri 1999). Among the numerous domains of application figure fields such as cardiology, demography, finance, hydrology or pharmacokinetics. For references which give illustrative applications of the IG distribution, we quote Chhikara and Folks (1989), Seshadri (1993) as well as Seshadri (1999). One further reason for the popularity of the IG are the many similarities, in terms of statistical properties, between the Gaussian and the inverse Gaussian families, as pointed out e.g. by Mudholkar and Tian (2002). As its reciprocal distribution, the RIG of course benefits from this popularity and has also awakened the interest of the statistical community (see, e.g., Scaillet 2004).

In view of its relevance, we particularize now our findings from the previous section to $p_0 = -1/2$, the IG sub-model. Thanks to the previously stated property of GIG laws, RIG(b, a) is the law of $1/X$ if $X \sim \text{IG}(a, b)$, hence we shall estimate for free as well a test for $p_0 = 1/2$. Under the IG model, the MM estimates of the parameters a and b admit the nice expressions $\hat{a}_{\text{MM}}^{(n)} = \frac{1}{\bar{X}(\bar{X}\bar{X}_{-1}-1)}$ and $\hat{b}_{\text{MM}}^{(n)} = \frac{\bar{X}}{\bar{X}\bar{X}_{-1}-1}$, which happen to correspond with the ML estimates (see Seshadri 1999, page 7, a result dating back to Tweedie 1957). Further simplifications arise in the Fisher information matrix $\mathbf{\Gamma}(\boldsymbol{\vartheta})$ as

$$\Gamma_{a,b}(-1/2, a, b) = \frac{-1}{4\sqrt{ab}}, \quad \Gamma_{a,a}(-1/2, a, b) = \frac{\sqrt{b}a^{-3/2}}{4}$$

and

$$\Gamma_{b,b}(-1/2, a, b) = \frac{1}{2b^2} + \frac{\sqrt{ab}^{-3/2}}{4}.$$

Next, denote by u , v and w , respectively, the values of the second derivatives $\partial_{pp}^2 \log K_p(\sqrt{ab})$, $\partial_{pa}^2 \log K_p(\sqrt{ab})$ and $\partial_{pb}^2 \log K_p(\sqrt{ab})$ at $p = -1/2$. We then have

$$\mathbf{\Gamma}(-1/2, a, b) = \begin{pmatrix} u & v - \frac{1}{2a} & w + \frac{1}{2b} \\ v - \frac{1}{2a} & \frac{\sqrt{b}a^{-3/2}}{4} & -\frac{1}{4\sqrt{ab}} \\ w + \frac{1}{2b} & -\frac{1}{4\sqrt{ab}} & \frac{1}{2b^2} + \frac{\sqrt{ab}^{-3/2}}{4} \end{pmatrix},$$

yielding $\Gamma_{a,a}(-1/2, a, b)\Gamma_{b,b}(-1/2, a, b) - (\Gamma_{a,b}(-1/2, a, b))^2 = \frac{1}{8(ab)^{3/2}}$ and

$$\Gamma_{p,p}^{eff}(-1/2, a, b) = u - (2av-1)^2 \left(\frac{1}{\sqrt{ab}} + \frac{1}{2} \right) - (2av-1)(2bw+1) - \frac{1}{2}(2bw+1)^2.$$

The expression of $\Gamma_{p,p}^{eff}(-1/2, a, b)$ can be calculated efficiently using the integral expressions of the entries of $\mathbf{\Gamma}(\boldsymbol{\vartheta})$ given at the end of the previous section.

Regarding the efficient central sequence, we find

$$\Delta_p^{(n)eff}(-1/2, a, b) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\log(X_i) - \frac{M_{-3/2}}{I_{-3/2}} + R_i(a, b) \right),$$

with

$$R_i(a, b) = X_i \left[v \left(2a\sqrt{\frac{a}{b}} + a^2 \right) + wab - \sqrt{\frac{a}{b}} \right] + \frac{1}{X_i} (wb^2 + vab) - va(3 + 2\sqrt{ab}) - wb(1 + 2\sqrt{ab}) + 1.$$

Quite interestingly, if a and b are replaced with their MM (or ML) estimates, one gets $\sum_{i=1}^n R_i(\hat{a}_{\text{MM}}^{(n)}, \hat{b}_{\text{MM}}^{(n)}) = 0$, so that

$$\Delta_p^{(n)eff}(-1/2, \hat{a}_{\text{MM}}^{(n)}, \hat{b}_{\text{MM}}^{(n)}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\log(X_i) - \frac{\hat{M}_{-3/2}}{\hat{I}_{-3/2}} \right),$$

where $\hat{I}_{-3/2} = I_{-3/2}(\hat{a}_{\text{MM}}^{(n)}, \hat{b}_{\text{MM}}^{(n)})$ and $\hat{M}_{-3/2} = M_{-3/2}(\hat{a}_{\text{MM}}^{(n)}, \hat{b}_{\text{MM}}^{(n)})$. Summing up, the optimal test $\phi_{-1/2}^{(n)}$ rejects at asymptotic level α the null hypothesis of an IG(a, b) model in favor of a general GIG(p, a, b) model with $p \neq -1/2$ whenever the test statistic

$$|Q_{-1/2}^{(n)}| := \frac{\left| \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\log(X_i) - \frac{\hat{M}_{-3/2}}{\hat{I}_{-3/2}} \right) \right|}{\sqrt{\Gamma_{p,p}^{eff}(-1/2, \hat{a}_{\text{MM}}^{(n)}, \hat{b}_{\text{MM}}^{(n)})}}$$

exceeds $z_{\alpha/2}$.

5 Monte Carlo simulation study.

In this section we study the finite-sample behavior of our testing procedures. More precisely, we focus on testing for IG sub-models within the GIG family as described explicitly in Section 4, and we assess the performance of our optimal test $\phi_{-1/2}^{(n)}$ by comparing it to the likelihood ratio test $\phi_{\text{LR}}^{(n)}$. The Monte Carlo simulation study has been conducted on \mathbf{R} .

We start by investigating in how far the level constraint is met under the null hypothesis. To this end, we have generated $N = 10,000$ independent samples of varying small sample sizes $n = 100, 50$ and even 30 and for distinct values of the parameters a and b , and run both tests $\phi_{-1/2}^{(n)}$ and $\phi_{\text{LR}}^{(n)}$ at the $\alpha = 5\%$ level. The results are summarized in Table 1. We see that for each choice of couple (a, b) , the level is well maintained by our test when $n = 100$ and $n = 50$ (although, of course, less well for 50 than for 100 observations). Regarding the very small sample size $n = 30$, it still gives rise to acceptable levels, although some values are very low. This shows that, even for (very) small sample sizes, $\phi_{-1/2}^{(n)}$ meets the nominal level constraint, especially when

Size	$n = 100$			$n = 50$			$n = 30$		
(a, b)	(1, 5)	(3, 3)	(3, 2)	(1, 5)	(3, 3)	(3, 2)	(1, 5)	(3, 3)	(3, 2)
$\phi_{-1/2}^{(n)}$	0.0443	0.0440	0.0452	0.0407	0.0407	0.0414	0.0370	0.0350	0.0368
$\phi_{LR}^{(n)}$	0.0403	0.0388	0.0398	0.0294	0.0278	0.0314	0.0187	0.0162	0.0171

Table 1 Rejection frequencies (out of $N = 10,000$ replications) of our optimal test $\phi_{-1/2}^{(n)}$ and of the likelihood ratio test $\phi_{LR}^{(n)}$ under three distinct null hypotheses IG(1, 5), IG(2, 2) and IG(3, 2) and, in each setting, for three different sample sizes $n = 100, 50$ and 30. The nominal level is $\alpha = 0.05$.

compared to the likelihood ratio test $\phi_{LR}^{(n)}$ which attains the level not as well as our novel test.

Now, the level constraint being checked, we compare the power of our test $\phi_{-1/2}^{(n)}$ to the likelihood ratio test $\phi_{LR}^{(n)}$ for small and moderate sample sizes $n = 100$ and $n = 200$ and for distinct combinations of parameters (a, b) . To this end, we have once more generated $N = 10,000$ independent samples, this time with $p = -\frac{1}{2} + (\delta - 5)\frac{1}{2}$ for $\delta = 1, \dots, 9$. This choice permits us to test the power of our test against both higher and lower values of the parameter of interest. Note that, for $\delta = 6$, we test against the hyperbolic distribution, whereas for $\delta = 7$ the alternative is of RIG type. The results at the 5% level are reported in Table 2. We clearly see that the two tests detect the deviation from the null hypothesis for any combination (a, b) , the performance increasing logically with the sample size. The differences in performance are quite small, but the more significant differences are always in favor of our test, which is yet another argument that our simple efficient test supersedes the classical test $\phi_{LR}^{(n)}$.

6 Real-data example.

In this section, we analyze two real-data examples in the light of the new test $\phi_{-1/2}^{(n)}$ developed in this paper.

Our first example concerns the traffic data example analyzed in Section 7.3 from Jørgensen (1982). It concerns the length (in seconds) of $n = 128$ intervals between the times at which vehicles pass a point on a road; these values correspond to 2.8, 3.4, 1.4, 14.5, 1.9, 2.8, 2.3, 15.3, 1.8, 9.5, 2.5, 9.4, 1.1, 88.6, 1.6, 1.9, 1.5, 33.7, 2.6, 12.9, 16.2, 1.9, 20.3, 36.8, 40.1, 70.5, 2, 8, 2.1, 3.2, 1.7, 56.5, 23.7, 2.4, 21.4, 5.1, 7.9, 20.1, 14.9, 5.6, 51.7, 87.1, 1.2, 2.7, 1, 1.5, 1.3, 24.7, 72.6, 119.8, 1.2, 6.9, 3.9, 1.6, 3, 1.8, 44.8, 5, 3.9, 125.3, 22.8, 1.9, 15.9, 6, 20.6, 12.9, 3.9, 13, 6.9, 2.5, 12.3, 5.7, 11.3, 2.5, 1.6, 7.6, 2.3, 6.1, 2.1, 34.7, 15.4, 4.6, 55.7, 2.2, 6, 1.8, 1.9, 1.8, 42, 9.3, 91.7, 2.4, 30.6, 1.2, 8.8, 6.6, 49.8, 58.1, 1.9, 2.9, 0.5, 1.2, 31, 11.9, 0.8, 1.2, 0.8, 4.7, 8.3, 7.3, 8.8, 1.8, 3.1, 0.8, 34.1, 3, 2.6, 3.7, 41.3, 29.7, 17.6, 1.9, 13.8, 40.2, 10.1, 11.9, 11 and 0.2. We want to test whether the IG and RIG models fit well the data within the GIG family. For the null hypothesis of an IG model, our test $\phi_{-1/2}^{(n)}$, performed

	$\delta = 1$	$\delta = 2$	$\delta = 3$	$\delta = 4$	$\delta = 5$	$\delta = 6$	$\delta = 7$	$\delta = 8$	$\delta = 9$
	$(a, b) = (1, 2)$ and $n = 100$								
$\phi_{-1/2}^{(n)}$	0.3252	0.2741	0.1697	0.0781	0.0444	0.0910	0.2206	0.3742	0.4926
$\phi_{LR}^{(n)}$	0.3006	0.2505	0.1548	0.0688	0.0406	0.0921	0.2227	0.3762	0.4924
	$(a, b) = (2, 2)$ and $n = 100$								
$\phi_{-1/2}^{(n)}$	0.2213	0.1754	0.1238	0.0635	0.0429	0.0684	0.1264	0.2141	0.2972
$\phi_{LR}^{(n)}$	0.2004	0.1562	0.1116	0.0560	0.0389	0.0668	0.1246	0.2117	0.2933
	$(a, b) = (4, 2)$ and $n = 100$								
$\phi_{-1/2}^{(n)}$	0.1513	0.1094	0.0811	0.0553	0.0461	0.0541	0.0810	0.1190	0.1716
$\phi_{LR}^{(n)}$	0.1330	0.0974	0.0710	0.0476	0.0404	0.0516	0.0759	0.1148	0.1664
	$(a, b) = (1, 2)$ and $n = 200$								
$\phi_{-1/2}^{(n)}$	0.5757	0.4747	0.3051	0.1271	0.0458	0.1377	0.3926	0.6509	0.7872
$\phi_{LR}^{(n)}$	0.5613	0.4584	0.2907	0.1190	0.0435	0.1400	0.3988	0.6538	0.7895
	$(a, b) = (2, 2)$ and $n = 200$								
$\phi_{-1/2}^{(n)}$	0.4098	0.3007	0.1833	0.0885	0.0484	0.0917	0.2134	0.3910	0.5333
$\phi_{LR}^{(n)}$	0.3928	0.2878	0.1714	0.0835	0.0474	0.0923	0.2151	0.3924	0.5339
	$(a, b) = (4, 2)$ and $n = 200$								
$\phi_{-1/2}^{(n)}$	0.2568	0.1860	0.1178	0.0684	0.0486	0.0684	0.1265	0.2140	0.3088
$\phi_{LR}^{(n)}$	0.2426	0.1739	0.1095	0.0634	0.0460	0.0673	0.1254	0.2120	0.3068

Table 2 Rejection frequencies (out of $N = 10,000$ replications) of our optimal test $\phi_{-1/2}^{(n)}$ and the likelihood ratio test $\phi_{LR}^{(n)}$ under $\text{GIG}(-\frac{1}{2} + (\delta - 5)\frac{1}{2}, a, b)$ distributions for $\delta = 1, 2, \dots, 9$ and several choices of (a, b) and, in each setting, for two different sample sizes $n = 100$ and 200 . The nominal level is $\alpha = 0.05$.

under two-sided form, gives the p-value 0.0642, leading to a light non-rejection of the IG model, whereas the null of an RIG model is heavily rejected with a p-value of 1.8232×10^{-5} . We thus conclude that the latter model does not fit the data, whereas the IG model can serve as candidate inside the GIG family. Note that this finding is not in contradiction with the result of Natarajan and Mudholkar (2004) whose test clearly rejects the IG model for this data set, because they do not test IG within the GIG family but against other types of alternatives.

Our second example deals with the repair time data of Section 7.4 in Jørgensen (1982). The data read 0.2, 0.3, 0.5, 0.5, 0.5, 0.5, 0.6, 0.6, 0.7, 0.7, 0.7, 0.8, 0.8, 1, 1, 1, 1, 1.1, 1.3, 1.5, 1.5, 1.5, 1.5, 2, 2, 2.2, 2.5, 2.7, 3, 3, 3.3, 3.3, 4, 4, 4.5, 4.7, 5, 5.4, 5.4, 7, 7.5, 8.8, 9, 10.3, 22 and 24.5 and correspond to $n = 46$ active repair times (in hours) for an airborne communication transceiver. Jørgensen (1982) has concluded on basis of a Kolmogorov-Smirnov test that the IG model is a good underlying distribution for the repair times. Our test $\phi_{-1/2}^{(n)}$ completely agrees with his analysis, with a p-value of 0.4484, whereas the RIG model is here as well rejected (at the 5% level) with a p-value of 0.0207.

7 Future research.

Yet another interesting aspect of the Le Cam methodology consists in the fact that it allows for constructing a simple efficient estimator for $\boldsymbol{\vartheta} = (p, a, b)'$, the so-called one-step estimator. The main idea behind one-step estimation consists in adding to an existing adequate preliminary estimator $\hat{\boldsymbol{\vartheta}}^{(n)}$ a quantity depending on a version of the central sequence for $\boldsymbol{\vartheta}$. More precisely, the one-step estimator takes on the guise

$$\hat{\boldsymbol{\vartheta}}_{Cam}^{(n)} = \hat{\boldsymbol{\vartheta}}^{(n)} + n^{-1/2} \left(\boldsymbol{\Gamma}(\hat{\boldsymbol{\vartheta}}^{(n)}) \right)^{-1} \boldsymbol{\Delta}^{(n)}(\hat{\boldsymbol{\vartheta}}^{(n)}),$$

where $\hat{\boldsymbol{\vartheta}}^{(n)} = (\hat{p}^{(n)}, \hat{a}^{(n)}, \hat{b}^{(n)})'$ is a preliminary estimator of $\boldsymbol{\vartheta}$ fulfilling Assumption A (we would of course choose the MM estimator). The resulting estimator combines simplicity with efficiency, as, like the test statistics developed in this paper, it allows for a closed-form expression but is as efficient as ML estimators. Establishing the asymptotic behavior of such estimators and their optimality features, as well as comparing their performance to the competitors of the literature in a detailed simulation study, is part of ongoing research.

A Appendix.

PROOF OF THEOREM 21. Establishing the ULAN property of $\text{GIG}(p, a, b)$ with respect to all three parameters p, a, b is quite straightforward since we are not working within a semiparametric family of distributions (hence we do not have to deal with an infinite-dimensional parameter); the problem considered involves a parametric family of distributions with densities meeting the most classical regularity conditions. In particular, one readily obtains that

(i) $(p, a, b) \mapsto \sqrt{c(p, a, b)x^{p-1}e^{-(ax+b/x)/2}}$ is continuously differentiable for every $x > 0$ and (ii) the associated Fisher information matrix is well defined and continuous in p, a and b . Thus, by Lemma 7.6 of van der Vaart (1998), $(p, a, b) \mapsto \sqrt{c(p, a, b)x^{p-1}e^{-(ax+b/x)/2}}$ is differentiable in quadratic mean, and the ULAN property follows from Theorem 7.2 of van der Vaart (1998). This completes the proof. \square

PROOF OF PROPOSITION 31. It follows from the asymptotic linearity property of the central sequence given in (3) that, under $\mathbb{P}_{\boldsymbol{\vartheta}}^{(n)}$ and for $n \rightarrow \infty$,

$$\boldsymbol{\Delta}_p^{(n)eff} \left(\boldsymbol{\vartheta} + n^{-1/2} (0, \tau_2^{(n)}, \tau_3^{(n)})' \right) = \boldsymbol{\Delta}_p^{(n)eff}(\boldsymbol{\vartheta}) - \mathbf{P}(\boldsymbol{\vartheta})\boldsymbol{\Gamma}(\boldsymbol{\vartheta}) \begin{pmatrix} 0 \\ \tau_2^{(n)} \\ \tau_3^{(n)} \end{pmatrix} + o_{\mathbb{P}}(1),$$

where

$$\mathbf{P}(\boldsymbol{\vartheta}) = \left(1, -(\Gamma_{p,a}(\boldsymbol{\vartheta}), \Gamma_{p,b}(\boldsymbol{\vartheta})) \begin{pmatrix} \Gamma_{b,b}(\boldsymbol{\vartheta}) & -\Gamma_{a,b}(\boldsymbol{\vartheta}) \\ -\Gamma_{a,b}(\boldsymbol{\vartheta}) & \Gamma_{a,a}(\boldsymbol{\vartheta}) \end{pmatrix} / (\Gamma_{b,b}(\boldsymbol{\vartheta})\Gamma_{a,a}(\boldsymbol{\vartheta}) - (\Gamma_{a,b}(\boldsymbol{\vartheta}))^2) \right)$$

is the projection matrix onto the subspace orthogonal to the central sequences for a and b . One easily checks that the product $\mathbf{P}(\boldsymbol{\vartheta})\boldsymbol{\Gamma}(\boldsymbol{\vartheta})$ is of the form $(\cdot, 0, 0)$, leading to the announced asymptotic linearity (4). The other asymptotic equality in probability follows by continuity. \square

PROOF OF THEOREM 31. The statement in Part (i) is easily proved thanks to the asymptotic linearity in (6) under the null, since $\Delta_p^{(n)eff}(p_0, a, b)$ is asymptotically $\mathcal{N}(0, \Gamma_{p,p}^{eff}(p_0, a, b))$ under $\bigcup_{a \in \mathbb{R}_0^+} \bigcup_{b \in \mathbb{R}_0^+} \mathbb{P}_{p_0, a, b}^{(n)}$ by the central limit theorem.

In order to prove the more delicate Part (ii), observe that, under $\mathbb{P}_{p_0, a, b}^{(n)}$ and for any bounded sequence $\boldsymbol{\tau}^{(n)} = (\tau_1^{(n)}, \tau_2^{(n)}, \tau_3^{(n)})' \in \mathbb{R}^3$, we see that, as $n \rightarrow \infty$,

$$\left(\begin{array}{c} \Delta_p^{(n)eff}(p_0, a, b) \\ \Lambda_{(p_0, a, b)' + n^{-1/2}\boldsymbol{\tau}^{(n)}/(p_0, a, b)'}^{(n)} \end{array} \right) \xrightarrow{\mathcal{L}} \mathcal{N}_2 \left(\begin{array}{c} 0 \\ -\frac{1}{2}\boldsymbol{\tau}'\boldsymbol{\Gamma}(\boldsymbol{\vartheta})\boldsymbol{\tau} \end{array} \right), \left(\begin{array}{cc} \Gamma_{p,p}^{eff}(p_0, a, b) & \tau_1 \Gamma_{p,p}^{eff}(p_0, a, b) \\ \tau_1 \Gamma_{p,p}^{eff}(p_0, a, b) & \boldsymbol{\tau}'\boldsymbol{\Gamma}(\boldsymbol{\vartheta})\boldsymbol{\tau} \end{array} \right),$$

where $\Lambda_{(p_0, a, b)' + n^{-1/2}\boldsymbol{\tau}^{(n)}/(p_0, a, b)'}^{(n)}$ is the log-likelihood ratio and $\boldsymbol{\tau} = \lim_{n \rightarrow \infty} \boldsymbol{\tau}^{(n)}$. We can then apply Le Cam's third lemma which implies that $\Delta_p^{(n)eff}(p_0, a, b)$ is asymptotically $\mathcal{N}(\tau_1 \Gamma_{p,p}^{eff}(p_0, a, b), \Gamma_{p,p}^{eff}(p_0, a, b))$ under $\mathbb{P}_{p_0 + n^{-1/2}\tau_1^{(n)}, a, b}^{(n)}$. Since the asymptotic linearity (6) holds as well under $\mathbb{P}_{p_0 + n^{-1/2}\tau_1^{(n)}, a, b}^{(n)}$ by contiguity, Part (ii) of the theorem readily follows.

As regards Part (iii), the fact that $\phi_{p_0}^{(n)}$ has asymptotic level α follows directly from the asymptotic null distribution given in Part (i), while local asymptotic maximinity is a consequence of the convergence of the local experiments to the Gaussian shift experiment (see Le Cam and Yang 2000). \square

Acknowledgements The research of Christophe Ley is supported by a Mandat de Chargé de Recherche FNRS from the Fonds National de la Recherche Scientifique, Communauté française de Belgique. The authors thank an anonymous referee for helpful comments, and Ivan Nourdin for inviting Christophe Ley for a visit at the Institut Elie Cartan, research stay during which the present work was initiated and the main parts worked out.

References

1. Barndorff-Nielsen, O. E. and Halgreen, C. (1977). Infinite divisibility of the Hyperbolic and generalized inverse Gaussian distribution. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **38**, 309–312.
2. Bickel, P.J., Klaassen, C.A.J., Ritov, Y. and Wellner, J.A. (1993). *Efficient and Adaptive Statistical Inference for Semiparametric Models*. Baltimore, MD: Johns Hopkins University Press.
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. Chhikara, R. S. and Folks, L. (1989). *The inverse Gaussian distribution: theory, methodology, and applications*. Marcel Dekker, New York.
5. Fitzgerald, D. L. (2000). Statistical aspects of Tricomi's function and modified Bessel functions of the second kind. *Stoch. Environ. Res. Risk Assess.* **14**, 139–158.

6. Good, I. J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika* **40**, 237–264.
7. 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.
8. Iyengar, S. and Liao, Q. (1997). Modeling neural activity using the generalized inverse Gaussian distribution. *Biol. Cybern.* **77**, 289–295.
9. Jørgensen B. (1982). *Statistical properties of the generalized inverse Gaussian distribution*. Springer-Verlag, Heidelberg.
10. Kreiss, J-P. (1987). On adaptive estimation in stationary ARMA processes. *Ann. Statist.* **15**, 112–133.
11. Le Cam, L. (1960). Locally asymptotically normal families of distribution. *Univ. Calif. Publ. in Stat.* **3**, 37–98.
12. Le Cam, L. and Yang, G. L. (2000). *Asymptotics in statistics. Some basic concepts*. 2nd ed. Springer-Verlag, New York.
13. Lemonte, A. J. and Cordeiro, G. M. (2011). The exponentiated generalized inverse Gaussian distribution. *Stat. Probab. Lett.* **81**, 506–517.
14. Mudholkar, S. M. and Tian, L. (2002). An entropy characterization of the inverse Gaussian distribution and related goodness-of-fit test. *J. Stat. Plan. Infer.* **102**, 211–221.
15. Natarajan, R. and Mudholkar, G. S. (2004). Moment-based goodness-of-fit tests for the inverse Gaussian distribution. *Technometrics* **46**, 339–347.
16. Perreault, L., Bobée, B. and Rasmussen, P. F. (1999a). Halphen distribution system. I: Mathematical and statistical properties. *J. Hydrol. Eng.* **4**, 189–199.
17. Perreault, L., Bobée, B. and Rasmussen, P. F. (1999b). Halphen distribution system. II: Parameter and quantile estimation. *J. Hydrol. Eng.* **4**, 200–208.
18. Scaillet, O. (2004). Density estimation using inverse and reciprocal inverse Gaussian kernels. *J. Nonparametr. Stat.* **16**, 217–226.
19. Seshadri, V. (1993). *The inverse Gaussian distribution—a case study in exponential families*. Oxford University Press.
20. Seshadri, V. (1999). *The inverse Gaussian distribution*. Springer-Verlag, New York.
21. Tweedie, M. C. K. (1945). Inverse statistical variates. *Nature* **155**, 453–453.
22. Tweedie, M. C. K. (1956). Statistical properties of inverse gaussian distributions. *Virginia Journal of Science* **7**, 160–165.
23. Tweedie, M. C. K. (1957). Statistical properties of inverse gaussian distributions. *Ann. Math. Statist.* **28**, 362–377.
24. van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.