GOODNESS–OF–FIT TESTS FOR THE INVERSE GAUSSIAN DISTRIBUTION BASED ON THE EMPIRICAL LAPLACE TRANSFORM [1] [2]

Norbert Henze[3], Bernhard Klar

Institut für Mathematische Stochastik, Universität Karlsruhe,
Englerstr. 2, 76128 Karlsruhe, Germany

Abstract. This paper considers two flexible classes of omnibus goodness-of-fit tests for the inverse Gaussian distribution. The test statistics are weighted integrals over the squared modulus of some measure of deviation of the empirical distribution of given data from the family of inverse Gaussian laws, expressed by means of the empirical Laplace transform. Both classes of statistics are connected to the first nonzero component of Neyman's smooth test for the inverse Gaussian distribution. The tests, when implemented via the parametric bootstrap, maintain a nominal level of significance very closely. A large–scale simulation study shows that the new tests compare favorably with classical goodness–of–fit tests for the inverse Gaussian distribution, based on the empirical distribution function.

Keywords. Goodness–of–fit test, Inverse Gaussian distribution, Empirical Laplace transform, Parametric bootstrap, smooth tests of fit.

---

[1]AMS 1991 subject classification numbers: 62G10, 62G20
[2]Proposed running head: TESTS OF FIT FOR INVERSE GAUSSIAN DISTRIBUTION
[3]Corresponding author

1 Introduction

The inverse Gaussian distribution is a very versatile positive-domain two-parametric probabilistic model having numerous applications in diverse fields (see e.g. Chhikara and Folks (1989), Seshadri (1993) and Seshadri (1999)). It originates as the distribution of the first passage time of Brownian motion with drift. Further applications include lifetime models in connection with repairs (Chhikara and Folks (1977)), accelerated life testing (Bhattacharyya and Fries (1982)), reliability problems (Padgett and Tsoi (1986)) and frailty models (Hougaard (1984)). The naming inverse Gaussian distribution is derived from the fact that its cumulant generating function is the inverse of that of the Gaussian distribution.

A random variable X has an inverse Gaussian distribution with parameters $\mu > 0$ and $\lambda > 0$ (for short: $X \sim IG(\mu, \lambda)$), if X has the density

$$f(x; \mu, \lambda) \quad = \quad \sqrt{\frac{\lambda}{2\pi}}\, x^{-3/2} \, \exp\left(-\frac{\lambda (x-\mu)^2}{2\mu^2 x} \right), \qquad x > 0.$$

This density is unimodal with mean μ and variance μ^3/λ, and its shape depends only on the value of $\varphi = \lambda/\mu$. The distribution function pertaining to $f(x; \mu, \lambda)$ is

$$(1.1) \quad F(x; \mu, \lambda) \quad = \quad \Phi\left[\sqrt{\frac{\lambda}{x}} \left(\frac{x}{\mu} - 1 \right) \right] + e^{2\lambda/\mu}\, \Phi\left[-\sqrt{\frac{\lambda}{x}} \left(1 + \frac{x}{\mu} \right) \right], \qquad x > 0,$$

where $\Phi(t) = (2\pi)^{-1/2} \int_{-\infty}^{t} \exp(-\xi^2/2)\, d\xi$.

Since, for $X \sim IG(\mu, \lambda)$ and $r = 1, 2, \ldots$,

$$E(X^r) \quad = \quad \mu^r \sum_{s=0}^{r-1} \frac{(r-1+s)!}{s!(r-1-s)!} \left(2\,\frac{\lambda}{\mu} \right)^{-s}$$

(see Seshadri (1993), p. 46), skewness and kurtosis of X are $\sqrt{\beta_1} = 3\sqrt{\mu/\lambda}$ and $\beta_2 = 15\,\mu/\lambda + 3$, respectively, showing that the distribution $IG(\mu, \lambda)$ is positively-skewed and leptokurtic. Positive and negative moments of X are related by

$$(1.2) \quad E\left(X^{-r}\right) \quad = \quad \frac{E\left(X^{r+1}\right)}{\mu^{2r+1}}, \quad r = 1, 2, \ldots$$

(see Seshadri (1993), p. 52).

The class of inverse Gaussian distributions is closed with respect to scale transformations since, if $X \sim IG(\mu, \lambda)$ and $c > 0$, the transformed variable cX has the distribution $IG(c\mu, c\lambda)$. In particular, we have $X/\mu \sim IG(1, \varphi)$.

In view of its versatile nature, it is important to know whether the use of the inverse Gaussian model is justified in a given situation. This aspect of testing the goodness-of-fit of data with the class of inverse Gaussian laws has been addressed by several authors, although the existing literature is relatively sparse (see Edgeman et al. (1988), Pavur et al. (1992), O'Reilly, F.J. and Rueda, R. (1992), Gunes et al. (1997), Mergel (1999)). These papers adopt a classical approach to goodness-of-fit testing based on the empirical distribution function. A specific problem when testing the hypothesis H_0 that the underlying distribution is inverse Gaussian with unspecified parameters is that the null distribution of a test statistic usually depends on the unknown shape parameter φ. In order to have a test that maintains a nominal level of significance closely irrespective of the value of φ, the above papers advocate the use of special tables or formulas for critical values, obtained by extensive simulations or by a numerical approximation of the asymptotic null distribution. These formulas differ with the nominal level of significance and the statistic used, and they depend either on the sample size and the estimated value of φ from given data (see Table 2 of Gunes et al. (1997)), or in case that the asymptotic null distribution is used, only on the estimated value of φ (see O'Reilly, F.J. and Rueda, R. (1992), p. 390). A different method which, however, seems to have been overlooked in the context of testing the goodness-of-fit for the inverse Gaussian distribution, is to use a parametric bootstrap. This idea of simulating the null distribution of a test statistic is now well-established (see e.g. Stute et al. (1993)), and it does not only lead to reliable critical values, but also to approximate p-values.

The purpose of the present paper is twofold. First, we introduce two new classes of flexible omnibus tests of fit for the inverse Gaussian distribution. Our approach uses the empirical Laplace transform and is thus in the spirit of previous papers on the problem of testing for exponentiality (see Baringhaus and Henze (1991), Henze (1993) and Henze and Meintanis (2000)), or the testing of goodness-of-fit for discrete distributions based on the empirical probability generating function (see e.g., Baringhaus and Henze (1992), Epps (1995), Gürtler and Henze (2000), Nakamura and Perez-Abreu (1993), and Rueda and O'Reilly (1999)). We strongly advocate the use of the parametric bootstrap in order to obtain critical values or p-values. Since the advent of highspeed computers, the parametric bootstrap should be a standard tool in the context of goodness-of-fit testing. Secondly, we present the results of a large-scale simulation study comprising 14 different tests and various alternatives to the inverse Gaussian model. The new tests compare favorably to the existing procedures of testing for the IG-model.

The paper is organized as follows. Section 2 introduces the new test statistics, and Section 3 gives theoretical results on their limit behavior under H_0 and under local alternatives. Furthermore, the tests are proved to be consistent against general alternatives. Section 4 presents the results of the Monte Carlo study. The paper concludes with some examples.

2 The test statistics

Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of independent observations on a nonnegative random variable X. On the basis of $X_1, \ldots, X_n$, the problem is to test the hypothesis

$$H_0 : X \sim IG(\mu, \lambda) \ \text{ for some } \ \mu > 0, \lambda > 0$$

against general alternatives. The first class of statistics for testing H_0 is motivated by the fact that the Laplace transform $L(t) = E[\exp(-tX)]$ of $X \sim IG(\mu, \lambda)$ is

$$(2.1) \quad L(t) \ = \ \exp\left[\frac{\lambda}{\mu}\left(1 - \sqrt{1 + \frac{2\mu^2 t}{\lambda}}\right)\right], \quad t \geq 0$$

(see Seshadri (1993), p. 41), and thus satisfies the characteristic differential equation

$$(2.2) \quad \mu L(t) \ + \ \left(1 + 2\mu^2 t/\lambda\right)^{1/2} L'(t) \ = \ 0, \quad t > 0,$$

subject to the initial condition $L(0) = 1$. Writing

$$(2.3) \quad L_n(t) \ = \ \frac{1}{n}\sum_{j=1}^{n} e^{-tX_j}$$

for the empirical Laplace transform of $X_1, \ldots, X_n$ and

$$\hat{\mu}_n \ = \ \bar{X}_n \ = \ \frac{1}{n}\sum_{j=1}^{n} X_j, \qquad \hat{\lambda}_n \ = \ \left(\frac{1}{n}\sum_{j=1}^{n}\left(1/X_j - 1/\bar{X}_n\right)\right)^{-1}$$

for the maximum likelihood estimators of μ and λ, respectively, it suggests itself to estimate the left-hand side of (2.2) by

$$\tilde{\epsilon}_n(t) \;=\; \hat{\mu}_n L_n(t) + \left(1 + 2\hat{\mu}_n^2 t/\hat{\lambda}_n\right)^{1/2} L_n'(t)$$

and then use a suitable measure of squared deviation of the random function $\tilde{\epsilon}_n(\cdot)$ from the zero function. The test statistic we propose is the weighted L^2-distance

$$(2.4) \qquad T_{n,a} \;=\; \frac{n}{\hat{\mu}_n} \int_0^\infty \tilde{\epsilon}_n^2(t)\, \exp(-a\hat{\mu}_n t)\, dt,$$

and rejection of H_0 is for large values of $T_{n,a}$. The weight parameter a, the role of which will be discussed in Section 3, is nonnegative and fixed.

Putting $\hat{\varphi}_n = \hat{\lambda}_n/\hat{\mu}_n$, $Y_j = X_j/\hat{\mu}_n$, and using

$$\tilde{\epsilon}_n(t) \;=\; \frac{1}{n}\sum_{j=1}^n e^{-tX_j}\left[\hat{\mu}_n - \left(1 + 2\hat{\mu}_n^2 t/\hat{\lambda}_n\right)^{1/2} X_j\right],$$

a change of variables in (2.4) yields

$$T_{n,a} \;=\; n\int_0^\infty \epsilon_n^2(u)\exp(-au)\, du,$$

where

$$(2.5) \qquad \epsilon_n(u) \;=\; \frac{1}{n}\sum_{j=1}^n e^{-uY_j}\left(1 - Y_j\sqrt{1 + 2u/\hat{\varphi}_n}\right), \qquad u \geq 0.$$

Notice that $T_{n,a}$ is scale invariant, since the values of $Y_1, \ldots, Y_n$ and $\hat{\varphi}_n$ are not affected under the transformation $X_j \to c\, X_j$ $(j = 1, \ldots, n)$, where $c > 0$. Putting

$$(2.6) \qquad \eta_n(u) \;=\; \sqrt{1 + 2u/\hat{\varphi}_n}, \qquad \hat{Z}_{jk} = \hat{\varphi}_n(Y_j + Y_k + a),$$

straightforward computation yields the alternative representation

$$\begin{aligned} T_{n,a} \;&=\; \frac{1}{n}\sum_{j,k=1}^n \int_0^\infty e^{-u(Y_j+Y_k+a)}\left(1 - Y_j\eta_n(u)\right)\left(1 - Y_k\eta_n(u)\right) du \\ &=\; \frac{\hat{\varphi}_n}{n}\sum_{j,k=1}^n \hat{Z}_{jk}^{-1}\left\{1 \;-\; (Y_j + Y_k)\left(1 + \sqrt{\frac{2\pi}{\hat{Z}_{jk}}}\, e^{\hat{Z}_{jk}/2}\left(1 - \Phi\left(\sqrt{\hat{Z}_{jk}}\right)\right)\right)\right. \\ &\qquad\qquad \left. + \left(1 + \frac{2}{\hat{Z}_{jk}}\right) Y_j Y_k\right\}. \end{aligned}$$

With regard to a numerically stable compute routine implementing the test based on $T_{n,a}$, it is advisable to reexpress $T_{n,a}$ by means of the exponentially scaled complementary error function $\mathrm{erfce}(x) = e^{x^2}\mathrm{erfc}(x)$, where $\mathrm{erfc}(x) = 2\int_x^\infty e^{-t^2}dt/\pi$. The function $\mathrm{erfce}(x)$ is closely related to Mill's ratio, defined by $R(x) = (1-\Phi(x))/\phi(x)$, where ϕ denotes the standard normal density. The result is

$$
\begin{aligned}
T_{n,a} &= \frac{\hat{\varphi}_n}{n}\sum_{j,k=1}^n \hat{Z}_{jk}^{-1}\left\{1-(Y_j+Y_k)\left(1+\sqrt{\frac{\pi}{2\hat{Z}_{jk}}}\,\mathrm{erfce}\left(\sqrt{\frac{\hat{Z}_{jk}}{2}}\right)\right)\right.\\
&\qquad\left.+\left(1+\frac{2}{\hat{Z}_{jk}}\right)Y_jY_k\right\}.
\end{aligned}
$$

A second, more direct, way to test H_0 via the empirical Laplace transform is to estimate the function $L(t)$ of (2.1) by the Laplace transform with estimated parameters, i.e., by

$$
\hat{L}_n(t) = \exp\left[\frac{\hat{\lambda}_n}{\hat{\mu}_n}\left(1-\sqrt{1+\frac{2\hat{\mu}_n^2 t}{\hat{\lambda}_n}}\right)\right], \quad t \ge 0,
$$

and to base a test on a measure of deviation between $\hat{L}_n(t)$ and the nonparametric estimator $L_n(t)$ of $L(t)$, defined in (2.3). Our proposal is the statistic

$$
V_{n,a} = n\hat{\mu}_n\int_0^\infty \left(L_n(t)-\hat{L}_n(t)\right)^2 \exp(-a\hat{\mu}_n t)\,dt\,,
$$

which, putting

$$
\delta_n(u) = \frac{1}{n}\sum_{j=1}^n e^{-uY_j} - \exp\left\{\hat{\varphi}_n\left(1-\sqrt{1+2u/\hat{\varphi}_n}\right)\right\}, \quad u \ge 0, \tag{2.7}
$$

takes the form

$$
V_{n,a} = n\int_0^\infty \delta_n^2(u)\,\exp(-au)\,du.
$$

As well as $T_{n,a}$, $V_{n,a}$ is scale invariant. For the case $a>0$, $V_{n,a}$ has the alternative representation

$$
\begin{aligned}
V_{n,a} &= \frac{1}{n}\sum_{j,k=1}^n\int_0^\infty e^{-uZ_{jk}}\,du - 2\sum_{j=1}^n\int_0^\infty e^{-uZ_j}\exp\left\{\hat{\varphi}_n\left(1-\eta_n(u)\right)\right\}\,du\\
&\quad + n\int_0^\infty \exp\left\{2\hat{\varphi}_n\left(1-\eta_n(u)\right)-au\right\}\,du
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{j,k=1}^{n} Z_{jk}^{-1} - 2\sum_{j=1}^{n} Z_j^{-1}\left\{1+\sqrt{\frac{2\pi\hat{\varphi}_n}{Z_j}}\exp\left(\frac{\hat{\varphi}_n(Z_j+1)^2}{2Z_j}\right)\times\right.\\
&\qquad \left.\times\left(\Phi\left(\frac{\hat{\varphi}_n^{1/2}(Z_j+1)}{Z_j^{1/2}}\right)-1\right)\right\}\\
&\quad +\frac{n}{a}\left\{1+\sqrt{\frac{2\pi\hat{\varphi}_n}{a}}\,2\exp\left(\frac{\hat{\varphi}_n(a+2)^2}{2a}\right)\left(\Phi\left(\frac{\hat{\varphi}_n^{1/2}(a+2)}{a^{1/2}}\right)-1\right)\right\},
\end{aligned}
$$

where $Z_j = Y_j + a$ and $Z_{jk} = Y_j + Y_k + a$ (recall $\eta_n(u)$ from (2.6)). With regard to the implementation of a numerical stable computer routine, $V_{n,a}$ can be rewritten as

$$
\begin{aligned}
V_{n,a} &= \frac{1}{n}\sum_{j,k=1}^{n} Z_{jk}^{-1} - 2\sum_{j=1}^{n} Z_j^{-1}\left\{1-\sqrt{\frac{\pi\hat{\varphi}_n}{2Z_j}}\ \mathrm{erfce}\left(\frac{\hat{\varphi}_n^{1/2}(Z_j+1)}{(2Z_j)^{1/2}}\right)\right\}\\
&\quad +\frac{n}{a}\left\{1-\sqrt{\frac{2\pi\hat{\varphi}_n}{a}}\ \mathrm{erfce}\left(\frac{\hat{\varphi}_n^{1/2}(a+2)}{(2a)^{1/2}}\right)\right\}.
\end{aligned}
$$

For the case $a = 0$, $V_{n,a}$ takes the form

$$
\begin{aligned}
V_{n,a} &= \frac{1}{n}\sum_{j,k=1}^{n} Z_{jk}^{-1} - 2\sum_{j=1}^{n} Z_j^{-1}\left\{1-\sqrt{\frac{\pi\hat{\varphi}_n}{2Z_j}}\ \mathrm{erfce}\left(\frac{\hat{\varphi}_n^{1/2}(Z_j+1)}{(2Z_j)^{1/2}}\right)\right\}\\
&\qquad + n\,\frac{1+2\hat{\varphi}_n}{4\hat{\varphi}_n}.
\end{aligned}
$$

3 Weight decay and smooth tests

This section addresses the problem of choosing the parameter a that controls the rate of decay of the weight function figuring in the definition of $T_{n,a}$ and $V_{n,a}$. Since the tail behavior of a probability distribution concentrated on $[0,\infty)$ is reflected by the behavior of its Laplace transform at zero and vice versa (see e.g. Feller (1971), Chapter XIII.5), one may anticipate the following qualitative behavior of the power of a test that rejects H_0 for large values of $T_{n,a}$ or $V_{n,a}$ when varying the decay parameter a: Choosing a small value of a, and thus letting the weight function decay slowly, should give a good safeguard against alternative distributions having a point mass or infinite density at zero. On the other hand, choosing a large value of a means putting most of the mass of the weight function near zero, which should give high power against alternatives with great difference in tail behavior in comparison with the inverse Gaussian distribution.

The case $a \to \infty$ is of particular interest. In fact, letting the rate of decay of the weight function tend to infinity, both $T_{n,a}$ and $V_{n,a}$, after a suitable rescaling, approach a limit that depends only on two empirical moments of $Y_1, \ldots, Y_n$.

Theorem 3.1 *: Let $Y_j = X_j/\overline{X}_n \quad (j = 1, \ldots, n)$, and put*

$$\overline{Y_n^2} \;=\; \frac{1}{n}\sum_{j=1}^{n} Y_j^2, \qquad \overline{Y_n^{-1}} \;=\; \frac{1}{n}\sum_{j=1}^{n} Y_j^{-1}.$$

For fixed n, we have

$$a) \qquad \lim_{a\to\infty} a^3 T_{n,a} \;=\; 2n\left(\overline{Y_n^2} - \overline{Y_n^{-1}}\right)^2 ,$$

$$b) \qquad \lim_{a\to\infty} a^5 V_{n,a} \;=\; 6n\left(\overline{Y_n^2} - \overline{Y_n^{-1}}\right)^2 .$$

Proof. Observe that $T_{n,a} = \int_0^\infty g(u)\exp(-au)\,du$, where $g(u) = n\epsilon_n^2(u)$ and $\epsilon_n(u)$ is defined in (2.5). Since $g(u) = nu^2\left(\overline{Y_n^2} - \overline{Y_n^{-1}}\right)^2 + O(u^3)$ as $u \to 0$ and thus

$$\lim_{u\to 0} \Gamma(3)\,\frac{g(u)}{u^{3-1}} \;=\; 2n\left(\overline{Y_n^2} - \overline{Y_n^{-1}}\right)^2,$$

assertion a) follows from an Abelian theorem on Laplace transforms (see Widder (1959), p.182, or Proposition 1.1 of Baringhaus et al. (2000). b) is proved similarly, since

$$n\delta_n^2(u) \;\sim\; \frac{nu^4}{4}\left(\overline{Y_n^2} - \overline{Y_n^{-1}}\right)^2 \quad \text{as} \quad u \to 0. \;\blacksquare$$

Notice that, from (1.2), we have $E[X^2] - E[X^{-1}] = 0$ if the distribution of X is inverse Gaussian with unit mean. Since the empirical moments $\overline{Y_n^2}$ and $\overline{Y_n^{-1}}$ are computed on the scaled data $Y_j = X_j/\overline{X}_n$, which have unit mean, the difference $\overline{Y_n^2} - \overline{Y_n^{-1}}$ should thus be small under H_0.

It is illuminating to compare the 'limit statistics'

$$T_{n,\infty} \;=\; 2n\left(\overline{Y_n^2} - \overline{Y_n^{-1}}\right)^2, \qquad V_{n,\infty} \;=\; 6n\left(\overline{Y_n^2} - \overline{Y_n^{-1}}\right)^2$$

figuring in the statement of Theorem 3.1 with the first nonzero component of Neyman's smooth test for the inverse Gaussian distribution. These components have the form

$\hat{U}_{n,k} = n^{-1/2} \sum_{j=1}^{n} h_k(X_j; \hat{\vartheta}_n)$, where $\hat{\vartheta}_n = (\hat{\mu}_n, \hat{\lambda}_n)$ is the maximum likelihood estimator of $\vartheta = (\mu, \lambda)$, and $\{h_0(\cdot;\vartheta) \equiv 1, h_1(\cdot;\vartheta), h_2(\cdot;\vartheta), \ldots\}$ are orthonormal polynomials with respect to $IG(\mu, \lambda)$, that is, we have

$$\int_0^\infty h_k(t;\vartheta) h_l(t;\vartheta)\, f(t;\vartheta)\, dt \;=\; \delta_{kl} \quad (k, l \geq 0),$$

where δ_{kl} denotes Kronecker's delta and $f(t;\vartheta) = f(t;\mu,\lambda)$. Orthogonal polynomials of degree 1 and 2 are given by $\tilde{h}_1(x;\vartheta) = x - \mu$ and

$$\tilde{h}_2(x;\vartheta) \;=\; \frac{x^2}{\mu^2} - \left(\frac{3\mu}{\lambda} + 2\right)\frac{x}{\mu} + \frac{2\mu}{\lambda} + 1.$$

Hence $\hat{U}_{n,1} = 0$ and

$$\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \tilde{h}_2(X_j; \hat{\vartheta}_n) &= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} Y_j^2 - \sqrt{n}\left(\frac{\hat{\mu}_n}{\hat{\lambda}_n} + 1\right) \\
&= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} Y_j^2 - \frac{1}{\sqrt{n}} \sum_{j=1}^{n} Y_j^{-1} \;=\; \sqrt{n}\left(\overline{Y_n^2} - \overline{Y_n^{-1}}\right).
\end{aligned}$$

The normalized orthogonal polynomial of degree 2 is $h_2(x;\vartheta) = \tilde{h}_2(x;\vartheta)/s_2(\vartheta)$, where

$$s_2(\vartheta) \;=\; \left(E\left[\tilde{h}_2^2(X;\vartheta)\right]\right)^{1/2} \;=\; \sqrt{2}\,\frac{\mu}{\lambda}\left(3\,\frac{\mu}{\lambda} + 1\right)^{1/2}.$$

Therefore, the first nonzero component is

$$\hat{U}_{n,2} \;=\; \sqrt{n}\left(\overline{Y^2} - \overline{Y^{-1}}\right) \Big/ s_2\left(\hat{\vartheta}_n\right), \tag{3.1}$$

where

$$s_2\left(\hat{\vartheta}_n\right) \;=\; \sqrt{2}\,\frac{\hat{\mu}_n}{\hat{\lambda}_n}\left(3\,\frac{\hat{\mu}_n}{\hat{\lambda}_n} + 1\right)^{1/2} \;=\; \sqrt{2}\left(\overline{Y^{-1}} - 1\right)\left(3\overline{Y^{-1}} - 2\right)^{1/2}.$$

Thus, apart from a factor that converges in probability, $\hat{U}_{n,2}^2$ coincides with $T_{n,\infty}$ and $V_{n,\infty}$. For further examples on the connection between weighted integral test statistics and components of smooth tests of fit, see Baringhaus et al. (2000).

Typically, the second component of a smooth test of fit, which is based on a polynomial of degree 2, reflects a relationship between the first and the second moment under the hypothetical model. In case of the inverse Gaussian distribution, however, the second component is essentially the empirical counterpart of the equation $EX^{-1} - EX^2/(EX)^3 = 0$. This somewhat surprising fact is due to the use of the maximum likelihood estimator $\hat{\lambda}_n$ of λ, which does not coincide with the method of moments estimator. If one uses the latter, the second component is also zero. This can be verified by direct computation or using Lemma 2.2 of Klar (2000).

4 Asymptotic distribution theory

Throughout this section, we assume that the parameter a in the weight function has a fixed nonnegative value. To derive the asymptotic null distribution of $T_{n,a}$ and $V_{n,a}$ as $n \to \infty$, we work in the separable Hilbert space $\mathcal{L}_2 = \mathcal{L}_2(\mathbb{R}_+, \mathcal{B}_+, P_a)$ of measurable functions on $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$, equipped with the Borel σ-algebra $\mathcal{B}_+$, that are square integrable with respect to P_a, the exponential distribution with parameter a. P_0 stands for Lebesgue measure on $\mathbb{R}_+$. Notice that ϵ_n and δ_n, which were defined in (2.5) and (2.7), respectively, are $\mathcal{L}_2$-valued random elements, and that $T_{n,a}$ and $V_{n,a}$ are continuous functionals of ϵ_n and δ_n, respectively. The inner product and the norm in $\mathcal{L}_2$ are denoted by $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$, respectively.

If $a > 0$, a complete orthonormal set in $\mathcal{L}_2(\mathbb{R}_+, \mathcal{B}_+, P_a)$ is $(\pi_n(a \cdot t))_{n \geq 0}$, where $\pi_n(\cdot)$ denotes the n^{th} (normalized) Laguerre polynomial. If $a = 0$, a complete orthonormal set in $\mathcal{L}_2(\mathbb{R}_+, \mathcal{B}_+, P_0)$ is $(\exp(-t/2)\pi_n(t))_{n \geq 0}$ (see Courant and Hilbert (1953), p. 93). This orthonormal set will be used in the proof of Theorem 4.1.

The maximum likelihood estimator $\hat{\vartheta}_n = (\hat{\mu}_n, \hat{\lambda}_n)$ of $\vartheta = (\mu, \lambda)$ has the representation

$$(4.1) \qquad \sqrt{n}(\hat{\vartheta}_n - \vartheta) \;=\; \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \tilde{l}(X_j, \vartheta) \;+\; \tilde{r}_n,$$

where $\tilde{l}(x, \vartheta) = (\tilde{l}_1(x, \vartheta), \tilde{l}_2(x, \vartheta)) = [I(\vartheta)]^{-1} \cdot \nabla_\vartheta \log f(x, \vartheta)$, $I(\vartheta)$ is the Fisher information matrix, and $\tilde{r}_n = o_P(1)$. Since

$$\frac{\partial \log f}{\partial \mu} = \frac{\lambda(x-\mu)}{\mu^3}, \qquad \frac{\partial \log f}{\partial \lambda} = \frac{1}{2\lambda} - \frac{1}{2x}\frac{(x-\mu)^2}{\mu^2}$$

and

$$I(\vartheta) = \begin{pmatrix} \lambda\mu^{-3} & 0 \\ 0 & (2\lambda^2)^{-1} \end{pmatrix},$$

we have $\tilde{l}_1(x, \vartheta) = x - \mu$ and $\tilde{l}_2(x, \vartheta) = \lambda - \lambda^2 (x-\mu)^2/(x\mu^2)$. Note that $E_\vartheta[\tilde{l}(X, \vartheta)] = 0$ and $E_\vartheta[\|\tilde{l}(X, \vartheta)\|^2] < \infty$. Combining (4.1) and the equality

$$\sqrt{n}(\hat{\varphi}_n - \varphi) = \frac{\mu}{\hat{\mu}_n}\left(\frac{\sqrt{n}(\hat{\lambda}_n - \lambda)}{\mu} - \frac{\sqrt{n}(\hat{\mu}_n - \mu)\lambda}{\mu^2} \right)$$

yields

$$\sqrt{n}(\hat{\varphi}_n - \varphi) \;=\; \frac{1}{\sqrt{n}}\frac{\lambda}{\mu}\sum_{j=1}^{n}\left(1 - \frac{\lambda}{\mu^2}\frac{(X_j-\mu)^2}{X_j} - \frac{X_j-\mu}{\mu}\right) + o_P(1).$$

In the following, we assume $\mu = 1$ without loss of generality. Hence,

$$\sqrt{n}(\hat{\varphi}_n - \varphi) \;=\; \frac{1}{\sqrt{n}}\sum_{j=1}^{n} l(X_j,\varphi) + r_n, \tag{4.2}$$

where $l(x,\varphi) = \varphi\left(1 - \varphi(x-1)^2/x - (x-1)\right)$ and $r_n = o_P(1)$.

Since the test for the inverse Gaussian distribution will be carried out by means of a parametric bootstrap procedure (see Section 5), we have to show the weak convergence (denoted by $\Rightarrow$ in what follows) of the test statistics under a triangular array $X_{n1}, X_{n2}, \ldots, X_{nn}, n \geq 1$, of rowwise independent random variables having a common inverse Gaussian distribution $IG(1,\varphi_n)$, where $0 < \varphi = \lim_{n\to\infty}\varphi_n$ exists. In the bootstrap procedure, ϑ_n is $\hat{\vartheta}_n(\omega)$ for a fixed ω of the underlying probability space that generates the realizations $X_1(\omega), X_2(\omega), \ldots$. Therefore, also the convergence in distribution of $\sqrt{n}\,\delta_n$ to a limiting Gaussian process must be established for triangular arrays. To this end, we use the following Hilbert space Central Limit Theorem of Kundu et al. ((2000), Theorem 1.1). Therein, $\mathcal{H}$ denotes a real separable infinite dimensional Hilbert space.

Lemma 4.1 *Let $\{e_k : k \geq 0\}$ be an orthonormal basis of $\mathcal{H}$. For each $n \geq 1$, let $W_{n1}, W_{n2}, \ldots, W_{nn}$ be a finite sequence of independent $\mathcal{H}$-valued random elements with zero means and finite second moments, and put $W_n = \sum_{j=1}^{n} W_{nj}$. Let C_n be the covariance operator of W_n. Assume that the following conditions hold:*

a) $\lim_{n\to\infty}\langle C_n e_k, e_l\rangle = a_{kl}$ *(say) exists for all $k \geq 0$ and $l \geq 0$.*

b) $\lim_{n\to\infty}\sum_{k=0}^{\infty}\langle C_n e_k, e_k\rangle = \sum_{k=0}^{\infty} a_{kk} < \infty$.

c) $\lim_{n\to\infty} L_n(\varepsilon, e_k) = 0$ *for every $\varepsilon > 0$ and every $k \geq 0$, where, for $b \in \mathcal{H}$,*
$L_n(\varepsilon, b) = \sum_{j=1}^{n} E\left(\langle W_{nj}, b\rangle^2\, \mathbf{1}\{|\langle W_{nj}, b\rangle| > \varepsilon\}\right)$.

Then $W_n \Rightarrow \mathcal{N}(0, C)$ in $\mathcal{H}$, where the covariance operator C is characterized by $\langle Ch, e_l\rangle = \sum_{j=0}^{\infty}\langle h, e_j\rangle a_{jl}$, for every $l \geq 0$.

Theorem 4.1 *Let $\{X_{n1}, X_{n2}, \ldots, X_{nn}\}$, $n \geq 1$, be a triangular array of rowwise independent and identically distributed random variables, such that X_{n1} has the inverse Gaussian distribution $IG(1, \varphi_n)$, where $0 < \varphi = \lim_{n\to\infty} \varphi_n$ exists. Then*

$$(4.3) \qquad \sqrt{n}\,\delta_n \implies W$$

in $\mathcal{L}_2$, where W is a centered Gaussian process with covariance function

$$c(u,v) \;=\; Cov(W(u), W(v)) \;=\; E[g(u,X)\,g(v,X)], \quad 0 \leq s,t < \infty.$$

Here, $X \sim IG(1,\varphi)$ and

$$(4.4) \qquad g(u,x) \;=\; e^{-xu} - L(u,\varphi) + (x-1)u\left(\frac{\varphi}{\varphi+2u}\right)^{1/2} L(u,\varphi) - l(x,\varphi)\,\frac{\partial L(u,\varphi)}{\partial\varphi},$$

where $L(u,\varphi) = \exp\left[\varphi\left(1 - \sqrt{1+2u/\varphi}\right)\right]$. Furthermore,

$$(4.5) \qquad V_{n,a} \;=\; n\|\delta_n\|^2 \implies \|W\|^2.$$

PROOF: A Taylor expansion yields

$$\begin{aligned}
\sqrt{n}\,\delta_n(u) \;&=\; \sqrt{n}\left(\frac{1}{n}\sum_{j=1}^{n}\exp(-uX_{nj}/\bar{X}_n) - L(u,\hat{\varphi}_n)\right)\\
&=\; \sqrt{n}\Bigg(L_n(u) \;+\; \left(\bar{X}_n - 1\right)\frac{1}{n}\sum_{j=1}^{n} uX_{nj}\,\exp(-uX_{nj})\\
&\qquad - L(u,\varphi_n) \;-\; (\hat{\varphi}_n - \varphi_n)\,\frac{\partial L(u,\varphi)}{\partial\varphi}\bigg|_{\varphi=\varphi_n} \;+\; R_{n,1}\Bigg),
\end{aligned}$$

where

$$\begin{aligned}
R_{n,1} \;&=\; \frac{\left(\bar{X}_n - 1\right)^2}{n}\sum_{j=1}^{n}\frac{u^2X_{nj}^2 - 2\mu_n^* uX_{nj}}{2(\mu_n^*)^4}\,\exp(-uX_{nj}/\mu_n^*)\\
&\qquad - \frac{\left(\hat{\varphi}_n - \varphi_n\right)^2}{2}\,\frac{\partial^2 L(u,\varphi)}{\partial\varphi^2}\bigg|_{\varphi=\varphi_n^*},
\end{aligned}$$

with μ_n^* between $\bar{X}_n$ and 1 and φ_n^* between $\hat{\varphi}_n$ and φ_n. Using (4.2), we have

$$\begin{aligned}
\sqrt{n}\,\delta_n(u) \;&=\; \sqrt{n}\left(L_n(u) - L(u,\varphi_n)\right) \;+\; \sqrt{n}\left(\bar{X}_n - 1\right)\frac{1}{n}\sum_{j=1}^{n} uX_{nj}\,\exp(-uX_{nj})\\
&\qquad - \frac{\partial L(u,\varphi)}{\partial\varphi}\bigg|_{\varphi=\varphi_n}\frac{1}{\sqrt{n}}\sum_{j=1}^{n} l(X_{nj},\varphi_n) - r_n\,\frac{\partial L(u,\varphi)}{\partial\varphi}\bigg|_{\varphi=\varphi_n} + R_{n,1}.
\end{aligned}$$

Hence,

$$\sqrt{n}\,\delta_n(u) \;=\; W_n(u) \,+\, R_n(u), \tag{4.6}$$

where

$$\begin{aligned} W_n(u) \;&=\; \sqrt{n}\,(L_n(u) - L(u,\varphi_n)) \,+\, u\,E(X_{n1}\,\exp(-uX_{n1}))\,\frac{1}{\sqrt{n}}\sum_{j=1}^{n}(X_{nj}-1) \\ &\quad - \left.\frac{\partial L(u,\varphi)}{\partial\varphi}\right|_{\varphi=\varphi_n}\,\frac{1}{\sqrt{n}}\sum_{j=1}^{n} l(X_{nj},\varphi_n) \end{aligned}$$

and

$$\begin{aligned} R_n(u) \;&=\; \sqrt{n}\left(\bar{X}_n - 1\right)\left(\frac{1}{n}\sum_{j=1}^{n}(uX_{nj}\,\exp(-uX_{nj})) \,-\, u\,E\left(X_{n1}\,\exp(-uX_{n1})\right)\right) \\ &\quad - r_n\,\left.\frac{\partial L(u,\varphi)}{\partial\varphi}\right|_{\varphi=\varphi_n} \,+\, R_{n,1}. \end{aligned} \tag{4.7}$$

Using $E\left(X_{n1}\exp(-uX_{n1})\right) = (\varphi_n/(\varphi_n+2u))^{1/2}L(u,\varphi_n)$, and putting

$$\begin{aligned} g_n(u,x) \;&=\; e^{-xu} + L(u,\varphi_n)\left[-1 + \frac{(x-1)u}{\kappa(u,\varphi_n)}\right. \\ &\qquad\left. -l(x,\varphi_n)\left(1-\kappa(u,\varphi_n) + \frac{u}{\varphi_n\kappa(u,\varphi_n)}\right)\right], \end{aligned}$$

where, generically,

$$\kappa(u,\xi) \;=\; \left(1+\frac{2u}{\xi}\right)^{1/2}, \tag{4.8}$$

we obtain the representation $W_n(\cdot) = \sum_{j=1}^{n} W_{n,j}(\cdot)$, where $W_{n,j}(\cdot) = n^{-1/2}\,g_n(\cdot,X_{nj})$ are $\mathcal{L}_2$-valued random elements, which have zero mean (recall that $El(X_{n1},\varphi_n) = 0$) and finite second moments.

We now verify conditions a) - c) of Lemma 4.1 for $W_{n1},\ldots,W_{nn}$. To this end, let C_n be the covariance operator of W_n which, by independence and symmetry, is the covariance operator of $g_n(\cdot,X_{n1})$, and put $c_n(u,v) \;=\; E\left[g_n(u,X_{n1})\cdot g_n(v,X_{n1})\right]$. Note that $\lim_{n\to\infty} g_n(u,x) = g(u,x)$, where $g(u,x)$ is defined in (4.4). Put $c(u,v) \;=\; E\left[g(u,X)\cdot g(v,X)\right]$, where X is a random variable having the distribution $IG(1,\varphi)$, and write C for the covariance operator of $g(\cdot,X)$.

Using the relations $E(X) = 1$, $E(X^2) = 1 + 1/\varphi = E(1/X)$, $E(1/X^2) = E(X^3) = 1 + 3/\varphi + 3/\varphi^2$, $E(\exp(-uX)) = L(u, \varphi)$, $E(X \exp(-uX)) = L(u, \varphi)/\kappa(u, \varphi)$ (recall $\kappa(u, \varphi)$ from (4.8)), $E(\exp(-uX)/X) = L(u, \varphi)(\kappa(u, \varphi) + 1/\varphi)$, some algebra gives

$$\begin{aligned} c(u,v) \;=\; & L(u+v,\varphi) + L(u,\varphi)L(v,\varphi)\Bigg(\varphi\,\kappa(u,\varphi)\,\kappa(v,\varphi) \;+ \\ & + \varphi^2\,[\kappa(u,\varphi)+\kappa(v,\varphi)] + \varphi(\varphi+1)\left[\frac{1}{\kappa(u,\varphi)} + \frac{1}{\kappa(v,\varphi)}\right] \\ & + \left[\left(2+\frac{4}{\varphi}\right)uv + (\varphi+2)(u+v)\right]\frac{1}{\kappa(u,\varphi)\kappa(v,\varphi)} - (1+\varphi+\varphi^2) \\ & - [u(2+\varphi)+\varphi(\varphi+1)]\,\frac{\kappa(v,\varphi)}{\kappa(u,\varphi)} - [v(2+\varphi)+\varphi(\varphi+1)]\,\frac{\kappa(u,\varphi)}{\kappa(v,\varphi)}\Bigg). \end{aligned}$$

The same expression is valid for $c_n(u, v)$ with φ replaced throughout by φ_n. It follows that $\lim_{n\to\infty} c_n(u, v) = c(u, v)$ pointwise on $\mathbb{R}_+ \times \mathbb{R}_+$ and, uniformly in n,

$$(4.9)\qquad |c_n(u,v)| \;\le\; \bar{c}(u,v)$$

for some function $\bar{c}$ satisfying

$$(4.10)\qquad \int_0^\infty \int_0^\infty [\bar{c}(u,v)]^r\, P_a(du)P_a(dv) \;<\; \infty \text{ for } r = 1, 2.$$

Since, by (4.10) and the Cauchy-Schwarz inequality,

$$\int_0^\infty \int_0^\infty |\bar{c}(u,v)e_\kappa(u)e_l(v)|P_a(du)P_a(dv) \;<\; \infty,$$

(4.9) and dominated convergence yield

$$\begin{aligned} \lim_{n\to\infty} \langle C_n e_k, e_l\rangle \;&=\; \lim_{n\to\infty} \int_0^\infty \int_0^\infty c_n(u,v)e_\kappa(u)e_l(v)P_a(du)P_a(dv) \\ &=\; \int_0^\infty \int_0^\infty c(u,v)e_\kappa(u)e_l(v)P_a(du)P_a(dv) \\ &=\; \langle Ce_k, e_l\rangle \end{aligned}$$

which, setting $a_{kl} = \langle Ce_k, e_l\rangle$, proves condition a) of Lemma 4.1.

To verify condition b) of Lemma 4.1, use monotone convergence, Parseval's equality and dominated convergence to show

$$\begin{aligned} \lim_{n\to\infty} \sum_{k=0}^\infty \langle C_n e_k, e_k\rangle \;&=\; \lim_{n\to\infty} \sum_{k=0}^\infty E\left(\langle e_k, g_n(\cdot, X_{n1})\rangle^2\right) \\ &=\; \lim_{n\to\infty} E\|g_n(\cdot, X_{n1})\|^2 \end{aligned}$$

$$
\begin{aligned}
&= \int_0^\infty \lim_{n\to\infty} c_n(u,u) P_a(du) \\
&= \int_0^\infty c(u,u) P_a(du) \\
&= E\|g(\cdot, X)\|^2 \\
&= \sum_{k=0}^\infty a_{kk} < \infty.
\end{aligned}
$$

To prove condition c) of Lemma 4.1, notice that, by symmetry and the definition of W_{nj}, $L_n(\varepsilon, e_k) \;=\; E\,(V_{nk}^2\ \mathbf{1}\{|V_{nk}| > \varepsilon\sqrt{n}\})$ where

$$
V_{nk} \;=\; \langle g_n(\cdot, X_{n1}), e_k\rangle \;=\; \int_0^\infty g_n(u, X_{n1})\, e_k(u)\ P_a(du).
$$

From continuity and the fact that $EX_{n1}^k < \infty$ for each $k = \pm 1, \pm 2, \ldots$, it follows that $\sup_{n\geq 1} E|V_{nk}|^3 < \infty$. Since $E\,(V_{nk}^2\ \mathbf{1}\{|V_{nk}| > \varepsilon\sqrt{n}\}) \;\leq\; E|V_{nk}|^3/(\varepsilon\sqrt{n})$, we are done.

By Lemma 4.1, $W_n \Rightarrow \mathcal{N}(0, \tilde{C})$ in $\mathcal{H}$, where the covariance operator $\tilde{C}$ is characterized by $\langle \tilde{C}h, e_l\rangle = \sum_{j=0}^\infty \langle h, e_j\rangle a_{jl}$, for every $l \geq 0$. Since

$$
\begin{aligned}
\langle Ch, e_l\rangle &= \int_0^\infty \int_0^\infty c(u,v) \left(\sum_{j=0}^\infty \langle h, e_j\rangle e_j(u)\right) e_l(v) P_a(du) P_a(dv) \\
&= \sum_{j=0}^\infty \langle h, e_j\rangle a_{jl},
\end{aligned}
$$

we have $\tilde{C} = C$ and thus $W_n \Rightarrow W$, where W is given in Theorem 4.1. On the other hand, since straightforward algebra shows that $\|R_n\| = o_P(1)$ (recall the definition of R_n from (4.7)), (4.3) follows from (4.6) and Theorem 4.1 of Billingsley (1968), and (4.5) is a consequence of the continuous mapping theorem. ■

Next, we consider the asymptotic behavior of $V_{n,a}$ under contiguous alternatives to the inverse Gaussian distribution. To this end, let $X_{n1}, \ldots, X_{nn}$, $n \geq 1$, be a triangular array of rowwise independent random variables having the Lebesgue density

$$
f_n(x) \;=\; f_0(x) \cdot \left(1 + \frac{h(x)}{\sqrt{n}}\right),
$$

where f_0 is the density of the inverse Gaussian distribution $IG(1, \varphi)$, and h is a bounded measurable function such that $\int_0^\infty h(x) f_0(x) dx = 0$. To guarantee that f_n is nonnegative, we tacitly assume n to be large enough.

Theorem 4.2 *Under the triangular array $X_{n1}, \ldots, X_{nn}$ and the standing assumptions, we have*

$$\sqrt{n}\, \delta_n \implies W + \Delta$$

in $\mathcal{L}_2$, where W is the centered Gaussian process figuring in Theorem 4.1, and the shift function Δ is given by

$$\Delta(u) = \int_0^\infty g(u,x) h(x) f_0(x) dx,$$

where $g(u,x)$ is defined in (4.4). Furthermore,

$$V_{n,a} = n\|\delta_n\|^2 \implies \|W + \Delta\|^2.$$

PROOF: Since the reasoning, mutatis mutandis, follows the proof of Theorem 3.1 of Henze and Wagner (1997), it will be omitted. ■

Remark 4.1 To show the weak convergence of $T_{n,a} = n\|\epsilon_n\|^2$ under H_0 and under contiguous alternatives, one can proceed along the same lines, using the fact that

$$\sqrt{n}\, \epsilon_n(u) = \tilde{W}_n(u) + \tilde{R}_n(u),$$

where

$$\begin{aligned} \tilde{W}_n(u) = {} & \frac{1}{\sqrt{n}} \sum_{j=1}^n \left(\exp(-uX_{nj})\,(1 - \rho_n(u) X_{nj})\right) + \frac{1}{\sqrt{n}} \sum_{j=1}^n (X_{nj} - 1) \\ & \times \left((u + \rho_n(u))\, E\left(X_{n1} \exp(-uX_{n1})\right) - u\rho_n(u) E\left(X_{n1}^2 \exp(-uX_{n1})\right)\right) \\ & + u\rho_n^{-1}(u) \varphi_n^{-2} E\left(X_{n1} \exp(-uX_{n1})\right)\ \frac{1}{\sqrt{n}} \sum_{j=1}^n l(X_{nj}, \varphi_n), \end{aligned}$$

$\rho_n(u) = (1 + 2u/\varphi_n)^{1/2}$, and $\|\tilde{R}_n\| = o_P(1)$. The details are omitted.

As a consequence of Theorem 4.1 and Remark 4.1, a test that rejects H_0 for large values of $T_{n,a}$ or $V_{n,a}$, carried out via the parametric bootstrap procedure described in Section 5, attains a given nominal level in the limit as both the sample size and the bootstrap sample size tend to infinity. The proof runs along the lines of Henze (1996).

From Theorem 4.2 and Remark 4.1, we conclude that the asymptotic level of the proposed tests is above the nominal level for sequences of alternatives that converge to the inverse Gaussian distribution at the rate $n^{-1/2}$. Of course, more work needs to be done to understand the dependence of power on the parameter a, particularly if one has in mind some kind of adaptive test for the inverse Gaussian distribution.

We now consider the behavior of $T_{n,a}$ and $V_{n,a}$, $a \geq 0$, under fixed alternatives to H_0, with the aim of showing the consistency of the corresponding tests. To this end, suppose that the distribution of X has finite positive expectation μ and finite negative moment $\lambda = E[1/X]$. We then have $\hat{\mu}_n \to \mu$ almost surely and $\hat{\lambda}_n \to \lambda$ almost surely, whence $\hat{L}_n(t) \to L(t)$ almost surely, where $L(t)$ is given in (2.1). By Fatou's Lemma, it follows that

$$\liminf_{n\to\infty} \frac{V_{n,a}}{n} \;\geq\; \mu \int_0^\infty (E[\exp(-tX)] - L(t))^2 \exp(-at)dt \tag{4.11}$$

almost surely. Since the right-hand side of (4.11) is positive if the distribution of X is not inverse Gaussian, a test that rejects H_0 for large values of $V_{n,a}$, carried out via the parametric bootstrap, is consistent against any such alternative. In the same way, start with (2.4) and use Fatou's Lemma to get

$$\liminf_{n\to\infty} \frac{T_{n,a}}{n} \geq \frac{1}{\mu} \int_0^\infty \left(\mu E[e^{-tX}] \;+\; \left(1 + 2\mu^2 t/\lambda\right)^{1/2} E[Xe^{-tX}]\right)^2 \exp(-at)dt \tag{4.12}$$

almost surely. In view of (2.2), the right-hand side of (4.12) is positive if the distribution of X is not inverse Gaussian. Consequently, also the parametric bootstrap test that rejects H_0 for large values of $T_{n,a}$ is consistent against any alternative with the properties stated above. It is not difficult to see that the property of consistency of both tests continues to hold under the condition $E[1/X] = \infty$.

5 Simulations

To assess the power of the new goodness-of-fit tests for the inverse Gaussian distribution in comparison with classical procedures that are based on the empirical distribution function (so-called edf tests), a large–scale simulation study has been conducted. Among the edf tests, we considered the *Kolmogorov-Smirnov test*, the *Cramér-von Mises test*, the *Anderson-Darling test*, and the *Watson test*, which were also studied by Gunes et al. (1997). Putting $\hat{F}(x) = F(x; \hat{\mu}_n, \hat{\lambda}_n)$, where $F(x; \mu, \lambda)$ is given in

(1.1), and writing $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$ for the order statistics of $X_1, \ldots, X_n$, the Kolmogorov-Smirnov statistic is

$$KS \;=\; \max(D^+, D^-),$$

where $D^+ = \max_{i=1,\ldots,n}(i/n - \hat{F}(X_{(i)}))$, $D^- = \max_{i=1,\ldots,n}(\hat{F}(X_{(i)}) - (i-1)/n)$. The Cramér-von Mises statistic is

$$CM \;=\; \frac{1}{12n} \;+\; \sum_{j=1}^{n} \left(\hat{F}(X_{(j)}) - \frac{2j-1}{2n} \right)^2,$$

whereas the Watson statistic is given by

$$WA \;=\; CM \;-\; n \left(\sum_{j=1}^{n} \frac{\hat{F}(X_{(j)})}{n} - \frac{1}{2} \right)^2.$$

Finally, the Anderson-Darling statistic takes the form

$$AD \;=\; -n - \frac{1}{n} \sum_{j=1}^{n} \Big((2j-1) \log \hat{F}(X_{(j)}) + (2(n-j)+1) \log(1 - \hat{F}(X_{(j)})) \Big).$$

Among the new procedures based on the empirical Laplace transform, we chose four tests from each of the two different classes, namely $T_{n,0}$, $T_{n,0.25}$, $T_{n,1}$, $T_{n,10}$ and $V_{n,0}$, $V_{n,0.25}$, $V_{n,1}$, $V_{n,10}$. Although not leading to consistent tests, we also considered the square of the second component of Neyman's smooth test, i.e. $\hat{U}_{n,2}^2$, where $\hat{U}_{n,2}$ is defined in (3.1), and the statistic

$$S_{n,\infty} \;=\; n \left(\overline{Y_n^2} - \overline{Y_n^{-1}} \right)^2$$

which, apart from a factor, is the limit of $a^3 T_{n,a}$ and $a^5 V_{n,a}$ as $a \to \infty$.

Writing $W_n = W_n(X_1, \ldots, X_n)$ for any of the statistics described above, the corresponding level-α test rejects H_0 if $W_n(x_1, \ldots, x_n)$ exceeds some critical value c on given data $x_1, \ldots, x_n$. Since W_n is not distribution–free under H_0, c will be estimated from the data by a parametric bootstrap, which avoids doubtful reliance upon asymptotic critical values. The parametric bootstrap runs as follows:

Let $H_{n,\vartheta}(t) := P_\vartheta(W_n \le t)$ be the distribution function of the null distribution of W_n under $\vartheta = (\mu, \lambda)$. The parametric bootstrap estimates the natural critical value, which is the unknown $(1-\alpha)$–quantile of $H_{n,\hat{\vartheta}_n}$, by the following Monte Carlo procedure: Conditionally on the observed value of $\hat{\vartheta}_n = \hat{\vartheta}_n(X_1,\ldots,X_n) = (\hat{\mu}_n,\hat{\lambda}_n)$, generate B pseudo–random samples of size n with the distribution $IG(\hat{\mu}_n,\hat{\lambda}_n)$, i.e. generate $X^*_{j1},\ldots,X^*_{jn}$, $j=1,\ldots,B$, i.i.d. according to $IG(\hat{\mu}_n,\hat{\lambda}_n)$. Then calculate $W^*_{j,n} := W_n(X^*_{j1},\ldots,X^*_{jn})$ for $j=1,\ldots,B$. Writing $H^*_{n,B}(t) := B^{-1}\sum_{j=1}^{B}\mathbf{1}\{W^*_{j,n}\le t\}$ for the empirical distribution function of $W^*_{1,n},\ldots,W^*_{B,n}$ and $W^*_{1:B}\le\ldots\le W^*_{B:B}$ for their order statistics, the empirical $(1-\alpha)$–quantile $c^*_{n,B}$ of $H^*_{n,B}$ is

$$c^*_{n,B} := \begin{cases} W^*_{B(1-\alpha):B} & , \text{ if } B(1-\alpha) \text{ is an integer} \\ W^*_{[B(1-\alpha)]+1:B} & , \text{ otherwise,} \end{cases}$$

where $[y]$ is the largest integer not greater than y. We used the modified critical value

$$\tilde{c}_{n,B} := W^*_{\alpha_n:B} + (1-\gamma_n)(W^*_{\alpha_n+1:B} - W^*_{\alpha_n:B})$$

with $\alpha_n := B - [\alpha(B+1)]$, $\gamma_n := \alpha(B+1) - [\alpha(B+1)]$, suggested by Baringhaus and Henze (1992), which leads to an accurate empirical level of the test even for a fairly moderate bootstrap sample size B.

Each of the tests under discussion was implemented via the parametric bootstrap as described above with a bootstrap sample size of $B := 200$. The nominal level is $\alpha = 0.1$, the sample size is 20 and 50, and each power estimate is based on 10 000 Monte Carlo replications. Calculations were done on an IBM RS/6000 SP parallel computer at the Rechenzentrum of the University of Karlsruhe, using high precision arithmetic in FORTRAN 90 and routines from the NAG and the IMSL libraries, whenever available.

To check the actual level of the test, the statistics were simulated under the null hypothesis $IG(1,\varphi)$ for a wide range of values of the shape parameter φ, namely $\varphi \in \{0.25, 0.5, 1, 3, 5, 10, 20, 100\}$. In a second step, we considered the following alternatives to the inverse Gaussian law, bearing in mind that each of the test statistics is scale-invariant:

- the Weibull distribution $W(\varphi)$ with density $\varphi x^{\varphi-1}\exp(-x^{\varphi})$, $x>0$, for $\varphi \in \{0.5,\ 0.8,\ 1,\ 1.2,\ 1.6,\ 2,\ 3\}$,
- the Lognormal distribution $LN(\varphi)$ with density $\exp(-\log^2(x)/\varphi^2)/(\varphi x(2\pi)^{1/2})$, $x>0$, for $\varphi \in \{0.6,\ 1,\ 1.4,\ 2,\ 3,\ 5\}$,

- the Gamma distribution $G(\varphi)$ with density $x^{\varphi-1}\exp(-x)/\Gamma(\varphi)$, $x > 0$, for $\varphi \in \{0.6, 1, 2\}$,
- the Half-normal distribution with HN density $(2/\pi)^{1/2}\exp(-x^2/2)$, $x \geq 0$,
- the Half-Cauchy distribution HC with density $2/(\pi(1+x^2))$, $x \geq 0$,
- the uniform distribution $\mathcal{U}(0,1)$.

Inverse Gaussian random variates were generated using the 'transformations with multiple roots method' of Michael et al. (1976); see, e.g. Seshadri (1993), p.203. Standard routines of the IMSL library were used to generate random numbers from the remaining distributions.

Power estimates of the tests under discussion are given in Tables 1 and 2. All entries are the percentages of 10 000 Monte Carlo samples that resulted in rejection of H_0, rounded to the nearest integer. An asterisk denotes power 100%. Further simulation results for other alternatives and sample sizes are available from the authors upon request.

The main conclusions that can be drawn from the simulation results are the following:

1. The tests based on $\hat{U}^2_{n,2}$ and $\hat{S}^2_{n,\infty}$ have low power compared with the other procedures and should not be recommended as omnibus procedures for the testing problem under discussion. Furthermore, the actual level is far below the nominal level for small values of the shape parameter φ.

2. The tests from the new classes as well as the edf tests, when implemented via the parametric bootstrap, maintain the nominal level very closely, even for the sample size $n = 20$. The same behavior was observed for the nominal level $\alpha = 0.05$.

3. Among the group of edf tests, CM and AD outperform the tests of Watson and Kolmogorov-Smirnov, with AD having a slight edge over CM in some cases.

4. The new tests based on $T_{n,a}$ and $V_{n,a}$ behave fairly similar, with $T_{n,0}$ and $V_{n,0}$ performing best. $V_{n,0}$ has a slight edge over $T_{n,0}$.

5. Over the whole range of alternatives considered, the test based on $V_{n,0}$ is at least as powerful as the Anderson-Darling test, which is the best test from the group of edf tests, and it clearly dominates AD in many cases. We thus conclude that $V_{n,0}$ yields a strong omnibus test for the Inverse Gaussian distribution.

6 Examples

We applied the tests under discussion to several data sets. The first set, which was also considered by Gunes et al. (1997), refers to $n = 46$ active repair times for an airborne transceiver (in hours). The data are given in Table 3.

The second example was also considered by Pavur et al. (1992). The results recorded in Table 4 are the millions of revolutions to failure of $n = 23$ ball bearings in a life test study.

The third data set is given in Table 5. It consists of $n = 16$ intervals in operating hours between successive failures of airconditioning equipment in a Boeing 720 aircraft (see Edgeman et al. (1988)).

Our final examples are taken from O'Reilly, F.J. and Rueda, R. (1992). Table 6 shows the days of shelf life of a food product (sample size $n = 26$), and Table 7 exhibits precipitation data ($n = 25$), measured in inches, from Jug Bridge, Maryland.

Table 8 shows the estimated p-values for $T_{n,a}$, $V_{n,a}$ ($a = 0, 0.25, 1$) and the edf statistics on the five data sets. For each statistic T and each data set $x_1, \ldots, x_n$ (say), the p-value was obtained by first calculating $T_n = T(x_1, \ldots, x_n)$ and then calculating $T_b^* = T(x_{b,1}^*, \ldots, x_{b,n}^*)$ for $b = 1, 2, \ldots, 999$. Here, conditionally on $x_1, \ldots, x_n$, the bootstrap samples $x_{b,1}^*, \ldots, x_{b,n}^*$, $1 \leq b \leq 999$, are independent with the distribution $IG(\hat{\mu}_n, \hat{\lambda}_n)$, where $\hat{\mu}_n$ and $\hat{\lambda}_n$ are given in (2.4). The p-value of T_n is then one plus the number of those T_b^* ($1 \leq b \leq 999$) that exceed T_n, divided by 1000.

For the transceiver data, the ball bearing data and the airconditioning equipment data, none of the tests rejects the Inverse Gaussian model at the 10% level. The results for the first two data sets agree with the findings of Gunes et al. (1997) and Pavur et al. (1992). However, the p-value of 40% for the Kolmogorov-Smirnov statistic and the airconditioning equipment data contradicts the result of Edgeman et al. (1988); they obtained a p-value between 0.05 and 0.10. This discrepancy is due to an error in the computation of the Kolmogorov-Smirnov statistic: Edgeman et al. (1988), p. 1210, reported a value of 0.2641, whereas the correct value is 0.162. Using this value and proceeding along the lines of Edgeman et al., one arrives at the conclusion that the hypothesis of an IG model is not rejected at the 20% level.

For the shelf life data set, there is a remarkable difference between the edf tests and the tests based on the empirical Laplace transform. Whereas the p-values for the latter group are roughly 20%, each of the edf tests rejects the Inverse Gaussian model at the 5% level.

For the precipitation data, the p-values are between 8% and 11% for the statistics based on the Laplace transform. Except for the Kolmogorov-Smirnov statistic, which yielded a p-value of 16%, the p-values for the edf statistics are below 6%. The p-values of 1.3% and 3.6% for the Anderson-Darling statistic and the last two examples are in agreement with the findings of O'Reilly, F.J. and Rueda, R. (1992).

Acknowledgements

The authors thank the referees for their constructive comments and Nora Gürtler for helpful remarks.

References

Baringhaus, L., Gürtler, N., and Henze, N. (2000). Weighted integral test statistics and components of smooth tests of fit. *Austral. New Zeal. J. Statist.*, *42*(2), 179–192.

Baringhaus, L., and Henze, N. (1991). A class of consistent tests for exponentiality based on the empirical Laplace transform. *Ann. Inst. Statist. Math.*, *43*, 551–564.

Baringhaus, L., and Henze, N. (1992). A goodness of fit test for the Poisson distribution based on the empirical generating function. *Statist. & Prob. Lett.*, *13*, 269–274.

Bhattacharyya, G., and Fries, A. (1982). Fatigue failure models - Birnbaum-Saunders vs. inverse Gaussian. *IEEE Trans. Reliab.*, *R-31*, 439–441.

Billingsley, P. (1968). *Convergence of probability measures.* New York: Wiley.

Chhikara, R., and Folks, J. (1977). The inverse Gaussian distribution as a lifetime model. *Technometrics*, *19*, 461–468.

Chhikara, R., and Folks, J. (1989). *The inverse Gaussian distribution, theory, methodology and applications.* New York: Marcel Dekker.

Courant, R., and Hilbert, D. (1953). *Methods of Mathematical Physics* (Vol. I). New York: Interscience.

Edgeman, R., Scott, R., and Pavur, R. (1988). A modified Kolmogorov-Smirnov test for the inverse Gaussian density with unknown parameters. *Commun. Statist. – Simula.*, *17*, 1203–1212.

Epps, T. (1995). A test of fit for lattice distributions. *Commun. Statist. – Th. Meth.*, *24*(6), 1455–1479.

Feller, W. (1971). *An introduction to probability theory and its applications* (Vol. II, 2^{nd} ed.). New York: Wiley.

Gunes, H., Dietz, D., Auclair, P., and Moore, A. (1997). Modified goodness-of-fit tests for the inverse Gaussian distribution. *Computational Statistics and Data Analysis*, *24*, 63–77.

Gürtler, N., and Henze, N. (2000). Recent and classical goodness–of–fit tests for the Poisson distribution. *J. Statist. Plann. Inf.*, *90*, 207–225.

Henze, N. (1993). A new flexible class of omnibus tests for exponentiality. *Commun. Statist. – Th. Meth.*, *22*, 115–133.

Henze, N. (1996). Empirical–distribution–function goodness–of–fit tests for discrete models. *Canad. J. Statist.*, *24*(1), 81–93.

Henze, N., and Meintanis, S. (2000). Tests of fit for exponentiality based on the empirical Laplace transform. *Statistics.* (to appear)

Henze, N., and Wagner, T. (1997). A new approach to the BHEP tests for multivariate normality. *J. Multiv. Anal.*, *62*(1), 1–23.

Hougaard, P. (1984). Life table methods for heterogeneous populations: Distributions describing the heterogeneity. *Biometrika*, *71*, 75–83.

Klar, B. (2000). Diagnostic smooth tests of fit. *Metrika*, *52*, 237–252.

Kundu, S., Majumdar, S., and Mukherjee, K. (2000). Central Limit Theorems revisited. *Statist. & Prob. Lett.*, *47*, 265–275.

Mergel, V. (1999). Test of goodness of fit for the inverse-gaussian distribution. *Math. Commun.*, *4*, 191–195.

Michael, J., Schucancy, W., and Haas, R. (1976). Generating random variates using transformations with multiple roots. *Amer. Statist.*, *30*(2), 88–90.

Nakamura, M., and Pérez-Abreu, V. (1993). Use of an empirical probability generating function for testing a Poisson model. *Canad. J. Statist.*, *21*(2), 149–156.

O'Reilly, F., and Rueda, R. (1992). Goodness of fit for the inverse Gaussian distribution. *Canad. J. Statist.*, *20*, 387–397.

Padgett, W., and Tsoi, S. (1986). Prediction intervals for future observations from the inverse Gaussian distribution. *IEEE Trans. Reliab.*, *R-35*, 406–408.

Pavur, R., Edgeman, R., and Scott, R. (1992). Quadratic statistics for the goodness-of-fit test of the inverse Gaussian distribution. *IEEE Trans. Reliab.*, *41*, 118–123.

Rueda, R., and O'Reilly, F. (1999). Tests of fit for discrete distributions based on the probability generating function. *Commun. Statist. – Simula.*, *28*(1), 259–274.

Seshadri, V. (1993). *The inverse Gaussian distribution - a case study in exponential families.* Oxford: Clarendon Press.

Seshadri, V. (1999). *The inverse Gaussian distribution - statistical theory and applications.* New York: Springer.

Stute, W., Manteiga, W., and Quindimil, M. (1993). Bootstrap based goodness–of–fit tests. *Metrika*, *40*, 243–256.

Widder, D. (1959). *The Laplace transform* (5th ed.). Princeton: Princeton University Press.

alternative	T_0	$T_{0.25}$	T_1	T_{10}	V_0	$V_{0.25}$	V_1	V_{10}	CM	W	AD	KS	$\hat{U}^2_{n,2}$	$\hat{S}^2_{n,\infty}$
$IG(0.25)$	9	9	10	8	10	12	12	10	11	10	11	10	6	2
$IG(0.5)$	9	9	10	8	10	11	11	10	10	10	10	10	9	4
$IG(1)$	9	10	10	9	10	11	11	10	11	10	10	10	10	6
$IG(3)$	10	10	10	9	10	10	10	9	10	10	10	10	11	8
$IG(5)$	10	10	10	9	10	10	10	10	10	10	10	10	12	9
$IG(10)$	10	10	10	10	10	10	10	10	10	10	10	10	12	9
$IG(20)$	10	10	10	10	10	10	10	10	10	10	10	10	12	9
$IG(100)$	10	10	10	10	10	10	10	10	10	10	10	10	12	10
$W(0.5)$	94	95	95	91	96	97	96	94	96	93	96	95	1	21
$W(0.8)$	81	82	83	77	83	85	85	80	81	74	82	79	30	4
$W(1.0)$	73	74	74	68	75	76	77	71	72	63	72	67	53	12
$W(1.2)$	66	67	67	62	68	69	69	64	64	55	64	59	59	23
$W(1.6)$	58	58	58	53	59	60	60	55	53	45	54	48	56	38
$W(2.0)$	53	53	53	49	54	54	55	50	47	40	48	42	53	41
$W(3.0)$	48	47	47	44	48	48	48	45	40	35	42	36	49	41
$LN(0.6)$	12	12	13	11	13	13	13	12	12	11	12	12	13	9
$LN(1.0)$	19	19	19	15	22	22	21	17	19	15	19	17	15	7
$LN(1.4)$	32	33	33	25	39	39	37	30	35	27	36	32	13	2
$LN(2)$	61	64	62	47	70	72	67	58	69	57	70	65	3	0
$LN(3)$	93	94	93	85	96	94	89	88	95	91	96	94	0	40
$G(0.6)$	90	91	92	88	91	93	93	90	91	86	91	89	19	6
$G(1.0)$	73	74	74	68	75	76	77	71	72	63	72	67	53	12
$G(2.0)$	44	44	44	38	46	46	46	40	39	32	40	36	40	24
χ^2_1	94	95	95	93	95	96	96	95	95	91	95	93	8	13
χ^2_2	73	74	74	68	75	76	77	71	72	63	72	67	53	12
χ^2_{10}	21	21	21	18	22	22	22	19	18	16	19	16	21	15
$HN(0,1)$	79	80	80	77	80	81	81	79	76	69	77	72	73	37
$HC(0,1)$	48	46	42	31	55	51	45	35	52	46	52	49	10	3
$\mathcal{U}(0,1)$	91	92	92	92	90	91	92	92	90	86	92	86	90	70

Table 1: Percentage of 10 000 Monte Carlo samples declared significant by various tests for the Inverse Gaussian distribution ($\alpha = 0.1$, $n = 20$)

alternative	T_0	$T_{0.25}$	T_1	T_{10}	V_0	$V_{0.25}$	V_1	V_{10}	CM	W	AD	KS	$\hat{U}^2_{n,2}$	$\hat{S}^2_{n,\infty}$
$IG(0.25)$	9	10	10	9	10	11	11	10	11	10	10	10	9	4
$IG(0.5)$	9	10	10	9	10	10	11	10	10	10	10	10	10	6
$IG(1)$	9	10	10	9	9	10	10	9	10	10	10	10	10	8
$IG(3)$	9	9	10	10	9	10	10	10	10	10	10	10	10	9
$IG(5)$	9	9	9	9	10	10	10	9	10	10	10	10	10	9
$IG(10)$	9	9	9	9	9	9	9	9	10	10	10	10	10	9
$IG(20)$	9	9	9	9	9	9	9	9	10	10	10	10	10	9
$IG(100)$	9	9	9	9	9	9	9	9	10	10	10	10	10	9
$W(0.5)$	*	*	*	*	*	*	*	*	*	*	*	*	3	21
$W(0.8)$	99	99	99	98	99	99	99	98	99	97	99	98	76	31
$W(1.0)$	97	97	97	95	97	98	97	96	96	92	96	94	90	70
$W(1.2)$	95	95	95	93	95	95	95	93	93	87	93	90	91	83
$W(1.6)$	91	91	91	89	91	91	91	89	86	78	87	80	88	84
$W(2.0)$	88	88	88	86	87	88	88	86	80	72	81	74	86	83
$W(3.0)$	83	83	83	82	83	83	83	83	72	64	75	65	83	81
$LN(0.6)$	15	15	15	13	16	16	16	14	13	12	13	12	14	12
$LN(1.0)$	31	30	28	21	35	34	30	23	28	22	29	26	20	14
$LN(1.4)$	60	59	54	38	66	63	56	41	61	49	61	56	29	9
$LN(2)$	92	92	88	71	95	94	89	74	95	89	95	93	11	0
$LN(3)$	*	*	*	97	*	99	97	93	*	*	*	*	0	35
$G(0.6)$	*	*	*	*	*	*	*	*	*	*	*	*	47	16
$G(1.0)$	97	97	97	95	97	98	97	96	96	92	96	94	90	70
$G(2.0)$	78	78	77	71	79	79	79	73	70	60	72	65	69	63
χ^2_1	*	*	*	*	*	*	*	*	*	*	*	*	21	15
χ^2_2	97	97	97	95	97	98	97	96	96	92	96	94	90	70
χ^2_{10}	41	41	40	35	42	42	41	36	30	25	32	27	35	32
$HN(0,1)$	99	99	99	99	99	99	99	99	98	96	98	97	96	91
$HC(0,1)$	81	74	61	39	82	71	59	41	80	78	81	77	19	6
$\mathcal{U}(0,1)$	*	*	*	*	*	*	*	*	*	*	*	*	99	95

Table 2: Percentage of 10 000 Monte Carlo samples declared significant by various tests for the Inverse Gaussian distribution ($\alpha = 0.1$, $n = 50$)

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	1.0	1.0	1.0	1.0	1.1	1.3	1.5	1.5	1.5	1.5	2.0
2.0	2.2	2.5	2.7	3.7	3.0	3.3	3.3	4.0	4.0	4.5	4.7
5.0	5.4	5.4	7.0	7.5	8.8	9.0	10.3	22.0	24.5		

Table 3: Repair times (in hours) for airborne transceivers

17.88	28.92	33.00	41.52	42.12	45.60	48.48	51.84
51.96	54.12	55.56	67.80	68.64	68.64	69.88	84.12
93.12	98.64	105.12	105.84	127.92	128.04	173.40	

Table 4: Number of revolutions (in millions) to failure of ball bearings

102	209	14	57	54	32	67	59	134	152	27	14	230	66	61	34

Table 5: Intervals between failures of airconditioning equipment

24	24	26	26	32	32	33	33	33	35	41	42	43
47	48	48	48	50	52	54	55	57	57	57	57	61

Table 6: Days of shelf life of a food product

1.01	1.11	1.13	1.15	1.16	1.17	1.17	1.20	1.52	1.54	1.54	1.57	1.64
1.73	1.79	2.09	2.09	2.57	2.75	2.93	3.19	3.54	3.57	5.11	5.62	

Table 7: Precipitation (in inches) from Jug Bridge, Maryland

	n	$T_{n,0}$	$T_{n,0.25}$	$T_{n,1}$	$V_{n,0}$	$V_{n,0.25}$	$V_{n,1}$	CM	W	AD	KS
transceiver	46	94	94	94	95	95	93	86	81	87	91
ball bearing	23	47	46	42	43	43	41	87	88	88	93
air condition	16	54	52	46	53	51	45	56	54	53	40
shelf life	26	18	18	17	21	20	18	2.1	1.9	1.3	3.0
precipitation	25	9.6	9.7	11	8.0	8.7	10	5.0	5.7	3.6	16

Table 8: p-values (in percent) of the test statistics under discussion on the given examples, based on 999 bootstrap replications