Robust Tests for Homogeneity of Inverse Gaussian Scale Parameters

Rajeshwari Natarajan

Department of Statistical Science, Southern Methodist University, Dallas, Texas, 75275

Govind S. Mudholkar

Department of Statistics, University of Rochester, Rochester, New York 14627

Abstract

The two parameter inverse Gaussian family $IG(\mu, \lambda)$ with expectation μ and scale parameter λ is known for its remarkable similarities with the Gaussian family; see Mudholkar and Natarajan (2002). In this paper, we consider inference methods associated with the inverse Gaussian scale parameter. The assumption of homogeneity of the scale parameters (λ) is as basic in the analysis of reciprocals for testing equality of inverse Gaussian means as the homoscedasticity assumption is in the normal theory analysis of variance. For testing this assumption we propose and examine IG analogues of the classical tests due to Bartlett (1937), Cochran (1941) and Hartley (1950). It is seen that the validity of these analogues is violated when there are arbitrary departures from the inverse Gaussian assumption just as the validity of their normal counterparts is compromised in case of non-normality. Miller (1968) considered robust alternatives to the classical methods for testing normal variances. We construct and study similar robust alternatives to testing homogeneity of IG scale parameters. Specifically, we investigate an analogue of Box-Andersen's (1955) test and a jackknife based test for testing homogeneity of IG scale parameters. It is seen that the robustness properties of these tests are very similar to those of their normal theory counterparts.

1 Introduction

The inverse Gaussian distribution known as the first passage time distribution of Brownian motion with positive drift, derived by Schrödinger (1915) and Smoluchowsky (1915), is now widely used to model positively skewed data. Wald (1947) derived it as the asymptotic distribution of average sample number (ASN) in the context of sequential probability ratio test. Tweedie (1947) promoted its use in statistics by discovering several of its similarities with the Gaussian distributions. Numerous additional analogies between the two families are now known; see Folks and Chhikara (1978), Iyengar and Patwardhan (1988), Mudholkar and Natarajan (2002), and for a recent overview see Natarajan (2002). Seshadri (1993, 1999) gives a comprehensive review of the theory and applications of the IG distribution. One of its latest application is in the area of internet communication; see Huberman *et al.* (1998) where an argument similar to Wald (1947) is used to show that the number of links an internet user follows before the page value first reaches the stopping threshold has an asymptotic inverse Gaussian distribution; the empirical fit to the data is seen to be excellent.

Inferences on location and scale parameters serve as guiding paradigms for a broad array of statistical decisions. Of these, problems involving location parameters are the most basic and common in applied statistics whereas those of scale parameters most commonly appear in secondary roles in the context of assumptions involving the nuisance parameters.

The scale parameters can be of intrinsic importance in their own right and problems involving them can play a pivotal role. Risk in equity markets is measured through price volatility and thus measures of variability capture the notion of risk. Quality of measurement devices such as bathroom scales is assessed in terms of the reproducibility of their readings and a scale parameter captures this variability. Additionally in quality improvement settings, the goal is to control production of defective items in a manufacturing process. This can be attained by minimizing the variability as measured by the scale parameter.

The scale parameter of the normal distribution has been central to the development of basic ideas such as the small sample exact distribution and test (the student's t), biased estimates, biased likelihood ratio tests, sufficiency, completeness, and critical regions with Neyman structure. In the second half of the twentieth century Box's (1953) analysis of the optimal test for the hypothesis of equality of variances inaugurated the modern study and theory of robust inference.

In the normal theory analysis of variance the homoscedasticity assumption is ubiquitous. Similarly, in testing homogeneity of inverse Gaussian means using the analysis of reciprocals, the analogue of analysis of variance, equality of the scale parameters is commonly assumed.

The rescaled sampling distributions of the maximum likelihood estimates of λ^{-1} and σ^2, for the inverse Gaussian and normal populations, respectively, are both chi-square. Interestingly, inference procedures concerning λ are also strikingly similar to those concerning the normal scale parameter. Box (1953) showed that normal theory inferences about variances are very sensitive to the assumption of normality. Miller (1968) studied various approaches with the jackknife technique showing some promise, for a recent survey, see Shao and Tu (1995). In Section 2, analogues of tests due to Bartlett (1937), Cochran (1941) and Hartley (1950) for testing homogeneity of λ are shown to be non-robust in ways similar to the normal case. Construction of robust methods using the Box-Andersen (1955) approach and using analysis of variance on jackknife pseudovalues, similar to Miller's (1968) approach is also discussed in Section 2. Section 3 presents a Monte Carlo study that evaluates the validity and power properties of these tests. Robustness properties of the procedures are examined for variety of alternatives including lognormal and the recently introduced contaminated inverse Gaussian distributions, see Mudholkar and Natarajan (2002).

2 Inferences Regarding Inverse Gaussian Scale Parameters

Let $X_{ij},\ j = 1, 2, \ldots n_i$ be random samples of size n_i from inverse Gaussian populations $IG(\mu_i, \lambda_i)$, $i = 1, 2, \ldots, k$. Let $\sum_i n_i = N$. We consider the problem of testing the hypothesis $H_0 : \lambda_1 = \lambda_2 = \ldots = \lambda_k$ against the alternative that H_0 is false. The maximum likelihood estimates of the λ_i^{-1} are given by $V_i = \{\sum_{j=1}^{n}(1/X_{ij} - 1/\bar{X}_i)\}/(n_i - 1),\ (n_i - 1)\lambda_i V_i \sim {\chi^2}_{n_i-1},\ i = 1, 2, \ldots, k$. Let $\nu_i = n_i - 1$ and $V = \sum_i \nu_i V_i/\nu$. The IG analogues for testing homogeneity of inverse Gaussian scale parameters corresponding to the three classical statistics for testing homogeneity of normal variances due to Bartlett (1937), Cochran (1941) and Hartley (1950), respectively are:

$$M_{IG} = \nu \log V - \sum_{i=1}^{k} \nu_i \log V_i, \tag{2.1}$$

$$G_{IG} = \frac{\max_{1\le i\le k} V_i}{\sum_{i=1}^{k} V_i}, \tag{2.2}$$

$$F_{IG} = \frac{\max_{1\le i\le k} V_i}{\min_{1\le i\le k} V_i}. \tag{2.3}$$

These can be unified as test statistic $T_{r,s}$:

$$T_{r,s} = \frac{\{\sum_{i=1}^{k}(V_i)^r\}^{1/r}}{\{\sum_{i=1}^{k}(V_i)^s\}^{1/s}}, \tag{2.4}$$

where (2.1)-(2.3) correspond to $\log T_{1,0},\ T_{\infty,-\infty},\ T_{\infty,1}$, respectively. The null distributions of $M_{IG},\ G_{IG},\ F_{IG}$ coincide with their normal counterparts, the percentiles of which are widely available including in Biometrika Tables for Statisticians (1966).

Following an argument similar to that of Bartlett (1937), it can be shown that for moderate size samples from inverse Gaussian populations M_{IG} is approximately $(1+A){\chi^2}_{k-1}$, where A is given by $A = \sum_i \{{\nu_i}^{-1} - \nu^{-1}\}/3(k-1)$. Also, analogous to the observation in Box (1953), each of the modified statistics given in equations (2.1)-(2.3) is nonrobust. To see this, note that the asymptotic distribution of V_i from a single population with the first two positive and first two negative moments finite, see Mudholkar and Natarajan (2002), as $n_i \to \infty$, is given by

$$\sqrt{n_i}\ \ (V_i - [\nu_i - (1/\mu_i)])\ \xrightarrow{d}\ N(0,\ {\eta_i}^2), \tag{2.5}$$

where

$$\nu_i = E(X_{ij}^{-1}),\ \mu_i = E(X_{ij}),\ {\tau_i}^2 = Var({X_{ij}}^{-1}),\ {\sigma_i}^2 = Var(X_{ij}),$$

$${\eta_i}^2 = {\tau_i}^2 + 2(1-\mu_i\nu_i)/{\mu_i}^2 + {\sigma_i}^2/{\mu_i}^4.$$

Consequently, the logarithmic transformation yields

$$\sqrt{n_i}\ \ (\log V_i - \log[\nu_i - (1/\mu_i)])\ \xrightarrow{d}\ N\left(0,\ \frac{{\eta_i}^2{\mu_i}^2}{(\nu_i\mu_i-1)^2}\right), \tag{2.6}$$

or $N(0, \delta_{2i}-1)$, where δ_{2i} is the IG analogue of the coefficient of kurtosis β_{2i}, see Mudholkar and Natarajan (2002). It is also shown that for the IG family, $\delta_{2i} = 3$. Furthermore, the moment estimator d_{2i} of δ_{2i} has exactly the same asymptotic distribution as in the normal case, i.e. $N(3, 24/n)$. Under inverse Gaussian assumption, the right hand side of (2.6) reduces to $N(0, 2)$. When $\delta_{2i} = \delta_2$ is assumed to be equal across the k populations, then M_{IG} can be shown to be asymptotically $(1 + 0.5(\delta_2 - 3))\chi^2_{k-1}$ for any parent distribution having finite first two moments; both positive and negative. Thus, the Type I error of M_{IG} is not even asymptotically controlled if δ_2 is unequal to three. In particular, in large samples the asymptotic mean of M_{IG} would be $(1+0.5(\delta_2-3))(k-1)$ instead of $k-1$ and asymptotic standard deviation $(1 + 0.5(\delta_2 - 3))\sqrt{2(k-1)}$ instead of $\sqrt{2(k-1)}$. Hence, as in the case of normal variances, the asymptotic sampling distribution of V_i may be used in constructing the analogue of the robust Box-Andersen (1955) test. The analogue of the well-known jackknife test of Miller (1968) for homogeneity of the variances of normal populations will also be described.

Analogue of Box-Andersen Test. The Box-Andersen (1955) test is a robust modification of the F-test test for equality of variances from two normal populations. This is obtained by adjusting the degrees of freedom so that the asymptotic variance of the F statistic under normal theory matches the asymptotic variance of F under sampling distribution with general kurtosis. Thus, the analogue of this test for testing equality of scale parameters from two inverse Gaussian populations can be constructed as follows. We adjust the degrees of freedom of the F statistic based on the asymptotic distribution of $V_i,\ i = 1, 2$, see Mudholkar and Natarajan (2002), using the fact that, as $n \to \infty$:

$$\sqrt{\frac{n_i}{0.5(\delta_{2i}-1)}}\,(V_i - (1/\lambda_i)) \xrightarrow{d} N\left(0, \frac{2}{{\lambda_i}^2}\right), \tag{2.7}$$

where $1/\lambda_i = (\nu_i\mu_i - 1)/\mu_i$. Also,

$$F_{IG} = \frac{V_2}{V_1} \sim F[\delta(n_2-1), \delta(n_1-1)], \tag{2.8}$$

under H_0, where $\delta = [0.5(\delta_2 - 1)]^{-1}$. As in the normal case, the sample IG-kurtosis d_2

$$d_2 \;=\; \sum_{i=1}^{2} \frac{\left(m_{2i}'/\bar{X}_i^{\,2} + m_{-2i}'\bar{X}_i^{\,2} - 3\bar{X}_i^{\,2}\bar{Y}_i^{\,2} + 2\bar{X}_i\bar{Y}_i - 1\right)}{2(\bar{X}_i\bar{Y}_i - 1)^2} + 3. \tag{2.9}$$

based on corresponding sample moments, i.e., $m_{-2i} = \sum_{j=1}^{n}\{1/X_{ij}\}^2/n$, $\bar{Y}_i = \sum_{j=1}^{n}\{1/X_{ij}\}/n$ and $m_{2i}{}' = \sum_{j=1}^{n} X_{ij}{}^2/n$, would be used in place of δ_2 above when conducting the test. When the samples are of different sizes from the two populations, then a weighted average would have to be considered. This analogue can also be generalized to $k > 2$ populations by using a Bartlett's statistics with the degrees of freedom adjusted, similar to its normal counterpart.

Jackknife test. We considered $\log \lambda_i{}^{-1}$ as the pivotal quantities appropriate for jackknifing since, as in the normal case, the logarithmic transformation stabilizes the asymptotic variance, see Cressie (1981). Further, the 'delete-one' jackknife yields the following pseudovalues:

$$W_{ij} = n_i \log V_i - (n_i - 1)\log V_{i(-j)}, \tag{2.10}$$

where $V_{i(-j)}$ is the sample estimate of λ^{-1} from group i with observation j deleted, $i = 1, 2, \ldots, k,\ j = 1, \ldots, n_i$. It is well-known from the jackknife theory, e.g., Thorburn (1977), the asymptotic distribution of $\bar{W}_i$, the mean of the pseudovalues in group i, is asymptotically $N(\log \lambda_i{}^{-1}, \delta_{2i} - 1),\ i = 1, \ldots, k$. Hence, a robust test may be based on the analysis of variance of these pseudovalues.

3 A Monte Carlo Study

In this section we present a Monte Carlo experiment performed in order to understand the operating characteristics of the robust tests proposed in this paper. In the robustness literature involving normal variances, symmetric populations such as contaminated normal, uniform, double exponential, and t-distributions have been used as alternatives. In an analogous manner, in this experiment we use IG-symmetric populations such as the lognormal, and *contaminated* inverse Gaussian, a special case of the IG scale-mixtures introduced in

Mudholkar and Natarajan (2002). The density functions of these populations appear very similar to the inverse Gaussian densities. However, we note that the inverse Gaussian and the lognormal populations are hard to distinguish from each other; see Jorgensen (1982) or Dhulesia *et al.* (1991).

Fifty thousand pairs of samples each of sizes $n = 10, 20, 25, 40$ were generated from each of the above populations. The inverse Gaussian samples and their scale mixtures were generated using the algorithm by Michael *et al.* (1976) and the IMSL C subroutine library was used to obtain the lognormal samples. We present only the Bartlett's analogue or equivalently the F test amongst the analogues of Bartlett, Cochran and Hartley (2.1) - (2.3) as they gave similar results. Three tests including the Bartlett (2.1), the Box-Andersen analogue given in (2.8), and the jackknife test based on the pseudovalues given in (2.10), were applied to each pair and the number of significant results were counted. We considered the parameter values $\Delta = \lambda_2/\lambda_1 = 1, 3, 5, 10$, where $\Delta = 1$ corresponds to the null hypothesis. A selection of the proportions of the tests which rejected the null hypothesis $H_0 : \lambda_1 = \lambda_2$ versus $H_1 :$ Not H_0, for various populations, $n = 25, 40$ and α, the level of significance, is presented in Tables I and II.

The following are some of the observations gleaned from the results of the simulation study:

***IG*-Analogue of Classical Tests:** As in the normal case, the analogue of Bartlett's test is very non-robust, in the sense that it is highly sensitive to departures from the IG assumptions; even Type I error is not controlled. For example in Table I for the lognormal distribution and for sample of size 25, the Type I errors are 12.2% for the level 5% test.

Robust Alternatives: The IG-analogue of the Box-Andersen test performs well in general. It shows the best Type I error control and corresponding power in our empirical study. Since the test involves four moments; positive as well as negative moments, it is expected to perform reasonably for moderate sized samples such as $n \geq 20$ as confirmed in our study. The jackknife test has much better Type I error control than the Bartlett analogue. In general, it appears anti-conservative for the alternatives; all of which have

TABLE I. Monte Carlo Power* Function of Competing Tests

for $\mathbf{H_0} : \lambda_1 = \lambda_2$ vs. $\mathbf{H_1} : \lambda_1 \neq \lambda_2$ for $\mathbf{n = 25}$

	$\alpha = 0.05$				$\alpha = 0.01$			
$\Delta = \lambda_2/\lambda_1$	1	3	5	10	1	3	5	10
Inverse Gaussian Distribution								
Bartlett	0.054	0.749	0.969	0.999	0.012	0.513	0.892	0.999
Box-Andersen	0.051	0.678	0.935	0.997	0.008	0.361	0.737	0.960
Jackknife	0.054	0.704	0.950	0.998	0.015	0.476	0.847	0.991
Contaminated IG Distribution: IG(1,1) + 10% IG(1, 10)								
Bartlett	0.065	0.737	0.962	0.999	0.016	0.511	0.879	0.996
Box-Andersen	0.054	0.644	0.913	0.994	0.097	0.338	0.696	0.943
Jackknife	0.059	0.674	0.933	0.997	0.016	0.448	0.813	0.985
Contaminated IG Distribution: IG(1,1) + 20% IG(1, 10)								
Bartlett	0.085	0.725	0.951	0.998	0.023	0.511	0.866	0.998
Box-Andersen	0.053	0.583	0.874	0.987	0.010	0.281	0.615	0.987
Jackknife	0.059	0.618	0.899	0.994	0.017	0.389	0.755	0.994
Lognormal Distribution								
Bartlett	0.122	0.707	0.933	0.997	0.040	0.517	0.839	0.988
Box-Andersen	0.047	0.585	0.893	0.995	0.075	0.293	0.687	0.963
Jackknife	0.068	0.572	0.838	0.968	0.019	0.366	0.679	0.920

* Based on 50,000 replications for each entry, $SE \leq 0.002$

TABLE II. Monte Carlo Power* Function of Competing Tests

for $H_0 : \lambda_1 = \lambda_2$ vs. $H_1 : \lambda_1 \neq \lambda_2$ for $\mathbf{n = 40}$

	$\alpha = 0.05$				$\alpha = 0.01$			
$\Delta = \lambda_2/\lambda_1$	1	3	5	10	1	3	5	10
Inverse Gaussian Distribution								
Bartlett	0.050	0.923	0.998	1.000	0.0099	0.784	0.990	1.000
Box-Andersen	0.052	0.903	0.996	0.999	0.0097	0.693	0.966	0.999
Jackknife	0.055	0.911	0.998	0.999	0.014	0.767	0.985	0.999
Contaminated IG Distribution: IG(1,1) + 10% IG(1, 10)								
Bartlett	0.065	0.909	0.997	1.000	0.015	0.771	0.986	0.999
Box-Andersen	0.050	0.861	0.992	0.999	0.009	0.619	0.937	0.997
Jackknife	0.055	0.873	0.994	1.000	0.013	0.705	0.969	0.999
Contaminated IG Distribution: IG(1,1) + 20% IG(1, 10)								
Bartlett	0.085	0.895	0.612	0.875	0.023	0.754	0.980	0.999
Box-Andersen	0.054	0.808	0.436	0.698	0.0094	0.538	0.893	0.994
Jackknife	0.057	0.827	0.492	0.758	0.014	0.634	0.945	0.999
Lognormal Distribution								
Bartlett	0.129	0.872	0.988	0.999	0.047	0.738	0.964	0.999
Box-Andersen	0.048	0.787	0.978	0.999	0.0077	0.530	0.905	0.996
Jackknife	0.067	0.750	0.945	0.993	0.018	0.553	0.996	0.980

* Based on 50,000 replications for each entry, $SE \leq 0.002$

$\delta_2 \geq 3$, i.e. analogous to the heavy tailed alternatives considered in Miller's (1968) empirical study such as double exponential, skew double exponential and sixth power distributions with $\beta_2 \geq 3$.

In conclusion, as in the normal case, the IG-analogue of Bartlett's test is hypersensitive to the assumption of the distribution being from the inverse Gaussian universe. Interestingly, the IG-analogue of Box-Andersen's test provides a robust alternative to the IG-analogue due to Bartlett for testing homogeneity of IG scale parameters. The jackknife test albeit anti-conservative seems to adequately have Type I error control for samples of sizes over 20.

Acknowledgements

The authors are thankful to Bill Schucany and Lynne Stokes for their constructive comments.

References

Bartlett, M.S. (1937) Properties of sufficiency and statistical tests. *Proceedings in Royal Statistical Society, A*, **160**, 268-282.

Biometrika Tables for Statisticians (1966), Volume I. Edited by Pearson and Hartley, Third Edition, Cambridge University Press, London.

Box, G.E.P. (1953) Non normality and tests on variances. *Biometrika*, **40**, 318-335.

Box, G.E.P. and Anderson, S.L. (1955) Permutation theory in the derivation of robust criteria and the study of departures from assumption. *Journal of the Royal Statistical Society Series B*, **17**, 1-26.

Chhikara, R.S. and Folks, J.L. (1989) *The Inverse Gaussian Distribution.* Dekker, New York.

Cochran, W.G. (1941) The distribution of the largest of a set of estimated variances as a fraction of their total. *The Annals of Eugenics*, **11**, 47-52.

Cressie, N. (1981) Transformations and the Jackknife. *Journal of the Royal Statistical Society Series B*, **43**, 177-182.

Dhulesia, H., Bernicot, M. and Deheuvels, P. (1991) Statistical analysis and modelling of slug lengths. *Multi-Phase Production* (A.P. Burns, editor). Elsevier, London, 88-112.

Efron, B. (1982) *The jackknife, the bootstrap and other resampling plans.* CBMS Regional Conference Series in Applied Mathematics, 38.

Folks, J.L. and Chhikara, R.S. (1978) The inverse Gaussian distribution and its statistical application - a review. *Journal of the Royal Statistical Society Series B*, **40**, $263-289$.

Frangos, C.C. (1987) An updated bibliography on the jackknife method. *Communications in Statistics, A*, **16**, 1543-1584.

Hartley, H.O. (1939) Testing homogeneity of a set of variances. *Biometrika*, **31**, 249-255.

Huberman, B.A., Pirolli, P.L.T., Pitkow, J.E. and Lukose, R.M. (1998) Strong regularities in World Wide Web surfing. *Science*, **280**, 95-97.

Iyengar, S. and Patwardhan, G. (1988) Recent developments in the inverse Gaussian distribution. *Handbook of Statistics, P.R. Krishnaiah and C.R. Rao editors*, **7**, 479-480.

Jørgensen, B. (1982) *Statistical properties of the Generalized Inverse Gaussian Distribution.* Springer-Verlag, New York.

Levene, H. (1960) Robust tests for equality of variances. *Contributions to Probability and Statistics* (I. Olkin, et al, editors), 278-292.

Michael, J.R., Schucany, W.R. and Haas, R.W. (1976) Generating random variables using transformation with multiple roots. *American Statistician*, **30**, $88-90$.

Miller, R.G. Jr. (1968) Jackknifing variances. *The Annals of Mathematical Statistics*, **39**, 567-582.

Mudholkar, G.S. and Natarajan, R. (2002) The inverse Gaussian analogues of symmetry, skewness and kurtosis. *Annals of the Institute of Statistical Mathematics*; **54**, 138-154.

Natarajan, R. (2002) Inverse Gaussian and Gaussian Analogies. *The Encyclopedia of Statistical Sciences, Second Edition*; to appear

Schrödinger, E. (1915) Zür Theorie der Fall-und Steigversuche an Teilchenn mit Brownsche Bewegung. *Physikalische Zeitschrift*, **16**, 289-295.

Seshadri, V. (1993) *The Inverse Gaussian distribution. A case study in exponential fami-*

lies. Clarendon Press, Oxford.

Seshadri, V. (1999) *The Inverse Gaussian distribution: Statistical Theory and Applications.* Springer-Verlag, New York.

Shao, J. and Tu, D. (1995). *The Jackknife and Bootstrap.* Springer -Verlag, New York.

Smoluchowsky, M.V. (1915) Notiz über die Berechning der Brownshen Molkular-bewegung bei des Ehrenhaft-millikanchen Versuchsanordnung. *Physikalische Zeitschrift*, **16**, 318-321.

Thorburn, D. (1977) On the asymptotic normality of the Jackknife. *Scandinavian Jounal of Statistics*, **4**, 113 – 118.

Tweedie, M.C.K. (1947) Functions of a statistical variate with given means, with special reference to Laplacian distributions. *Proceedings of the Cambridge Philosophical Society*, **43**, 41-49.

Wald, A .(1947) *Sequential Analysis.* Wiley, New York.