Score Tests for Inverse Gaussian Mixtures

A. F. Desmond[1]   **and**   **Z. L. Yang**[2]

[1]*Department of Mathematics and Statistics, University of Guelph*
Guelph, Ontario N1G 2W1, Canada.   `tdesmond@uoguelph.ca`

[2]*School of Economics, Singapore Management University*
90 Stamford Road, Singapore 178903.   `zlyang@smu.edu.sg`

July 28, 2010

Abstract

The mixed inverse Gaussian given by Whitmore (1986) provides a convenient way for testing the goodness of fit of a pure inverse Gaussian distribution. The test is a one-sided score test with the null hypothesis being the pure inverse Gaussian (i.e., the mixing parameter is zero) and the alternative a mixture. We devise a simple score test and study its finite sample properties. Monte Carlo results show that it compares favorably to the smooth test of Ducharme (2001).

In practical applications, when the pure inverse Gaussian distribution is rejected, one is interested in making inference about the general values of the mixing parameter. However, as it is well known that the inverse Gaussian mixture is a defective distribution, hence the standard likelihood inference cannot be applied. We propose several alternatives and provide score tests for the mixing parameter. Finite sample properties of these tests are examined by Monte Carlo simulation.

KEY WORDS: Defective Distribution, Inverse Gaussian, Score Tests.

1 Introduction

The inverse Gaussian distribution has become a popular model in a variety of application areas in the past few decades since the review paper of Chhikara and Folks (1978). Three full length monographs, Chhikara and Folks (1989) and Seshadri (1993, 1998) have been devoted to this model while Johnson et al. (Chapter 15, 1994) contains a useful review. This distribution has been used to model equipment life-times (Chhikara and Folks, 1977), strike durations (Lancaster, 1972), labour turnover (Whitmore, 1979), product interpurchase times (Banerjee and Bhattacharyya, 1976) and task completion times (Desmond and Chapman, 1993). More recently it has been used to model positively skewed data in statistical process control (Edgeman, 1989) and Hawkins and Olwell (1997). Hawkins and Olwell (1997) point out that an appeal to the underlying physical process being modelled is a major reason for preferring the inverse Gaussian model to competing positive, right-skewed distributions such as the log-normal and the Weibull. Its genesis as the first passage time distribution for Brownian motion with positive drift make it an attractive model in reliability theory, where failure is due to some cumulative damage process or crack growth phenomenon which may plausibly be modelled as an underlying Brownian motion process. In addition, its hazard rate behaviour is more descriptive of many empirical failure rate phenomena than say the log-normal (which tends to zero asymptotically) as pointed out by Hawkins and Olwell. Finally, tractability of sampling distributions of maximum likelihood estimators and pleasing analogies with normal sampling theory are points in its favour.

In this paper, we consider the assessment of the inverse Gaussian model via goodness of fit, and develop a score test for this assumption which is sensitive to a broad class of mixture alternatives. Desmond and Chapman (1993) found evidence that the inverse Gaussian was inadequate for some data sets on task completion times at a large automobile plant and suggested that this was due to heterogeneity in worker performance which could be modelled as stochastic variation in the drift parameter of the underlying Brownian motion process, or equivalently in the mean parameter of the inverse Gaussian distribution. In that paper using Wald and Likelihood Ratio tests, they concluded that the mixed inverse Gaussian distribution of Whitmore (1979, 1986) provided a substantially better fit for more complex tasks whereas the regular

2

inverse Gaussian proved adequate for simple tasks. The advantage of the score test is that it requires estimation only under the null hypothesis, while the Wald test requires estimation under the alternative. As a result, the score test is computationally more attractive. In addition, it is shown here that the score test applies to a broad class of mixture alternatives which do not require a full parametric specification.

Goodness-of-fit for the inverse Gaussian has been studied previously using methods based on the empirical cumulative distribution function in a series of papers by Edgeman and others; see for example Edgeman (1990) or Pavur et al. (1992). A graphical method for assessing fit based on standardized recursive residuals has been studied by Letac et al. (1985). More recently, a smooth test has been developed and examined by Ducharme (2001). Monte Carlo results presented in this paper show that the proposed test compares favorably to the smooth test in terms of power.

In practical applications, when the pure inverse Gaussian distribution is rejected, one is interested in making inference about the general values of the mixing parameter. However, as it is well known that the inverse Gaussian mixture is a defective distribution, hence the standard likelihood inference cannot be applied. We propose several alternatives and provide score tests for the mixing parameter. Finite sample properties of these tests are examined by Monte Carlo simulation.

2 The Mixed Inverse Gaussian Distributions

We adopt the following parameterization for the inverse Gaussian probability density function

$$f(t; \delta, \lambda) = \left(\frac{\lambda}{2\pi t^3} \right)^{1/2} \exp\left(\frac{-\lambda(\delta t - 1)^2}{2t} \right), \quad t > 0. \tag{1}$$

In terms of the underlying Brownian motion process, with drift and volatility parameters γ and σ^2, and fixed barrier s, we have $\delta = \gamma/s$, and $\lambda = s^2/\sigma^2$. The mean of the random variable T with the above distribution is $E(T) = \mu = 1/\delta$ while its variance is $Var(T) = \mu^3/\lambda = 1/(\delta^3 \lambda)$.

To obtain the mixed inverse Gaussian, we follow Whitmore (1986) and assume that heterogeneity in the drift parameter is modelled by a normal distribution for δ, given λ, with mean d, variance v/λ and density $p(\delta; d, \lambda)$, say. Also, given δ, (1) above

is now a conditional distribution $f(t; \lambda \mid \delta)$ say. The marginal mixture distribution of T is a three parameter model with parameter vector $\boldsymbol{\theta} = (v, d, \lambda)^t$ where ^{t} denotes transpose, and is given by

$$h(t; \boldsymbol{\theta}) = \int f(t; \lambda \mid \delta)p(\delta; d, v)d\delta. \tag{2}$$

Straightforward integration over the entire real line then gives

$$h(t; \boldsymbol{\theta}) = \left(\frac{\lambda}{2\pi t^3(vt + 1)} \right)^{1/2} \exp \left(\frac{-\lambda(dt - 1)^2}{2t(vt + 1)} \right),$$

$$t > 0, \quad \lambda > 0, \quad v \geq 0, \quad d > 0. \tag{3}$$

As pointed out in Desmond and Chapman (1993), there is no reduction in dimensionality, via sufficiency, in (3), unlike the pure inverse Gaussian, so that the inferences about $\boldsymbol{\theta}$ rely on asymptotic likelihood theory. Note that, under $H_0 : v = 0$, (3) reduces to the pure inverse Gaussian given by (1) and if this can be accepted, the exact sampling theory under (1) can be used.

However, there are two problems associated with the function defined in (3). The first is that the value $v = 0$ falls on the boundary of the parameter space, which violates one of the key conditions for the standard likelihood inference methods that requires the true value of the parameter of interest to be an interior point of the parameter space; see Lehmann (1983, Chapter 6). The second problem is that when $v > 0$ the pdf (3) is improper in the sense that it does not integrate to one. It is only close to a proper distribution when v/λ is small relative to d. Plots in Figure 1 show clearly that as v increases, the area under the curve gets smaller and smaller. Some numerical integrations show that the area under the curve can be as small as 0.5. For example, $\int_0^\infty h(t, 1, 1, v)dt = 0.9150, 0.6800$, and 0.5595 for $v = 1, 10$, and 100, respectively. Some authors (Seshadri, 1998, p140; Aalen and Gjessing, 2001, p11) have pointed out this problem, but did not proceed to give methods to deal with it. Clearly, it is inappropriate to apply the standard likelihood inference theory to test $H : v = v_0$. We now elaborate on various methods to deal with these problems and leave the details on inferences to the following sections.

Figure 1 here

The first problem can be avoided by using a score test. This is because (i) v is nonnegative and a one-sided test is desired, (ii) v is a scalar parameter and the

4

score test provides such a possibility, and (iii) the score test does not require the estimation of v at the null (unlike the Wald and the likelihood ratio tests) and hence the technical complications caused by the boundary issue can be avoided. See Verbeke and Molenberghs (2003) for some detailed arguments on this.

To overcome the second problem we propose two approaches. The first approach is to mix over δ by a truncated normal distribution, i.e., instead of assuming δ as $N(d, v/\lambda)$ which is the source of defectiveness as δ cannot be negative, we assume

$$
\begin{aligned}
p(\delta; \boldsymbol{\theta}) &= \frac{1}{P(\delta > 0)} \left(\frac{\lambda}{2\pi v} \right)^{1/2} \exp\left(-\frac{\lambda(\delta - d)^2}{2v} \right) \\
&= \frac{1}{\Phi\left(d\sqrt{\lambda/v}\right)} \left(\frac{\lambda}{2\pi v} \right)^{1/2} \exp\left(-\frac{\lambda(\delta - d)^2}{2v} \right), \ \delta \geq 0,
\end{aligned}
$$

where Φ is the cumulative distribution function (CDF) of the standard normal distribution. On redoing the integration in (3) using the truncated normal mixing distribution, we have after some algebra,

$$
h^{\diamond}(t; \boldsymbol{\theta}) = h(t; \boldsymbol{\theta}) \Phi\left((v + d)\lambda^{\frac{1}{2}}(v^2 t + v)^{-\frac{1}{2}} \right) \Phi^{-1}\left(d\sqrt{\lambda/v} \right). \tag{4}
$$

Essentially, $h(t; \boldsymbol{\theta})$ is converted into a proper density by multiplying by a normalization factor depending on both the variable and the parameters. Though inferences based on (4) are complicated, standard regularity conditions apply, so that asymptotic likelihood methods will be technically valid. Moreover, with the help of modern statistical packages such as GAUSS and S-Plus, maximizing the likelihood based on (4) is, in principle, straightforward. The issue of improper pdf may not arise when testing for departures from the pure inverse Gaussian distribution as in Section 3 (Mardia and Kent, 1991), but surely does when inference concerns the inverse Gaussian mixture as in Section 4. Figure 2 presents plots of the proper and improper inverse Gaussian mixture, from which one sees clearly the need for modification.

Figure 2 here

An alternative way of handling problems with defective mixtures is to allow the mixing distribution to be an arbitrary distribution (not necessary normal) with mean d and variance v/λ, and then use some approximation arguments. If $p(\delta)$ is the pdf

of such a distribution then the marginal distribution of t is

$$h(t; v, d, \lambda) = \int f(t; \delta, \lambda) p(\delta) d\delta$$

$$= E_\delta[f(t; \delta, \lambda)]$$

where $f(t; \delta, \lambda)$ is given by (1).

Expanding (1) in a Taylor series about $E(\delta) = d$, we obtain

$$f(t; \delta, \lambda) = f(t; d, \lambda) \left(1 + \sum_{j=2}^{\infty} \frac{(\delta - d)^j}{j!} \frac{f^{(j)}(t; \delta, \lambda)\big|_{\delta=d}}{f(t; d, \lambda)} \right) \tag{5}$$

where $f^{(j)}$ is the $j^{(\text{th})}$ partial derivative with respect to δ.

Taking expected values with respect to the mixing distribution and neglecting terms of higher order than the quadratic we obtain:

$$h^*(t; \boldsymbol{\theta}) = f(t; d, \lambda) \left[1 + \frac{v}{2} \left(\lambda(dt - 1)^2 - t \right) \right]. \tag{6}$$

For small v, h^* is an approximation to h. Section 4.3 shows that h^* is a proper density under a mild parameter constraint.

3 Score Tests for Inverse Gaussian Distribution

Notice that all the inverse Gaussian mixtures (proper or improper ones) discussed above have the common feature, that, when $v = 0$ the mixture distribution reduces to the pure inverse Gaussian distribution. This provides us with a simple way of testing the most interesting hypothesis: $H_0 : v = 0$, i.e., the pure inverse Gaussian fits the data well, against $H_a : v > 0$, i.e., some kind of inverse Gaussian mixture is appropriate for fitting the data.

We consider two cases for completeness. In the first case, assume that d and λ are known, so that $H_0 : v = 0$ is a simple null hypothesis. There may be occasions in physical or biological applications where subject matter theory implies specified values of the drift and volatility parameters, so that this case may be of more than academic interest. The score test is based on the derivative with respect to v of the log-likelihood based on a random sample $t_1, \cdots, t_n$ from (3), which is given by

$$S(v) = -\sum_{i=1}^{n} \frac{t_i}{2(vt_i + 1)} + \sum_{i=1}^{n} \frac{\lambda(dt_i - 1)^2}{2(vt_i + 1)^2}. \tag{7}$$

The null hypothesis, $H_0 : t_1, \cdots, t_n$ are from an inverse Gaussian with parameters d and λ, is equivalent to

$$H_0 : v = 0.$$

The score test in this case is based on

$$S(0) = \frac{1}{2} \sum_{i=1}^{n} \left(\lambda(dt_i - 1)^2 - t_i \right). \tag{8}$$

The exact variance of (8) can be obtained. Squaring (8) and taking expectations involving the first four moments of the distribution (1) (Johnson et al., 1994) we obtain:

$$Var(S(0)) = \frac{n}{2d^2} + \frac{5n}{2\lambda d^3}. \tag{9}$$

A test pivotal is given by

$$T_1 = \frac{\frac{1}{2} \sum_{i=1}^{n} \left(\lambda(dt_i - 1)^2 - t_i \right)}{\left(\dfrac{n}{2d^2} + \dfrac{5n}{2\lambda d^3} \right)^{1/2}}. \tag{10}$$

Following the arguments of Verbeke and Molenberghs (2003, p256), T_1 converges in distribution to the standard normal. Hence, our one sided test can still be carried out in the usual way with reference to the standard normal distribution even if $v = 0$ falls on the boundary of the parameter space. If one wants to use the likelihood ratio test, the limiting distribution referred to is a modified chi-squared as described in Verbeke and Molenberghs (2003). See also Silvapulle and Silvapulle (1995) and Silvapulle and Sen (2005) for general discussions on statistical inferences under constrains.

While the normal distribution is an approximation here, the mean and variance are exact under the inverse Gaussian assumption. The exact sampling distribution of (10) eludes us, although we note the well-known result that

$$\sum_{i=1}^{n} \frac{\lambda(dt_i - 1)^2}{t_i}$$

is exactly $\chi^2(n)$ and $\sum_{i=1}^{n} t_i$ is inverse Gaussian. The statistic T_1 can be used to test the hypothesis $H_0 : v = 0$ when both parameters are known.

When d and λ are unknown, the above test is no longer valid, and the score test now should be based on the derivative of the log-likelihood in (7) with d and λ

7

estimated by their maximum likelihood estimates under $H_0 : v = 0$. From standard results for the inverse Gaussian we get:

$$\hat{d} = \frac{1}{\bar{t}} \quad \text{and} \quad \hat{\lambda}^{-1} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{1}{t_i} - \frac{1}{\bar{t}} \right) \tag{11}$$

where $\bar{t} = \frac{1}{n} \sum_{i=1}^{n} t_i$. The score test is then based on

$$S(0, \hat{d}, \hat{\lambda}) = \frac{1}{2} \sum_{i=1}^{n} \left(\hat{\lambda}(\hat{d} t_i - 1)^2 - t_i \right). \tag{12}$$

We need to find the (asymptotic) variance of $S(0, \hat{d}, \hat{\lambda})$. From standard asymptotic theory, e.g. Cox and Hinkeley (1974), we obtain the large sample variance of (12) from the expected Fisher information matrix for (3). We find, with $\boldsymbol{\theta} = (v, d, \lambda)^t$ and, under $H_0 : v = 0$, $\mathbf{I}(\boldsymbol{\theta}) = -E\left[\partial^2 l / \partial \boldsymbol{\theta}^2 \right]$ is given by

$$\mathbf{I}(\boldsymbol{\theta})\big|_{v=0} = \begin{bmatrix} \dfrac{n}{2d^2} + \dfrac{5n}{2\lambda d^3} & -\dfrac{n}{d^2} & -\dfrac{n}{2\lambda d} \\[2mm] -\dfrac{n}{d^2} & \dfrac{n\lambda}{d} & 0 \\[2mm] -\dfrac{n}{2\lambda d} & 0 & \dfrac{n}{2\lambda^2} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{11} & \mathbf{I}_{12}^t \\[2mm] \mathbf{I}_{12} & \mathbf{I}_{22} \end{bmatrix}$$

where

$$\mathbf{I}_{11} = \left[\frac{n}{2d^2} + \frac{5n}{2\lambda d^3} \right], \quad \mathbf{I}_{12}^t = \left[-\frac{n}{d^2} \quad -\frac{n}{2\lambda d} \right] \quad \text{and} \quad \mathbf{I}_{22} = \begin{bmatrix} \dfrac{n\lambda}{d} & 0 \\[2mm] 0 & \dfrac{n}{2\lambda^2} \end{bmatrix}$$

are the appropriately partitioned matrices for the parameter of interest v, and the nuisance parameter (d, λ). Thus, the large sample variance of (12) is given by

$$\mathbf{I}_{11} - \mathbf{I}_{12}^t \mathbf{I}_{22}^{-1} \mathbf{I}_{12} = \frac{3n}{2\lambda d^3}.$$

Finally, a score test pivotal is given by:

$$T_2 = \frac{\frac{1}{2} \sum\limits_{i=1}^{n} \left(\hat{\lambda}(\hat{d} t_i - 1)^2 - t_i \right)}{\left(\dfrac{3n}{2\hat{\lambda}\hat{d}^3} \right)^{1/2}} \tag{13}$$

which is asymptotically standard normal, where $\hat{\lambda}$ and $\hat{d}$ are given by (11).

Now we offer some arguments to show that the score test T_2 is valid for a broad class of mixing distributions not necessarily normal. The derivation of T_2 is with respect to a normal mixing distribution for δ. Since the normal assumption for δ is a matter of convenience rather than scientific plausibility, we now assume that δ has an arbitrary distribution $p(\delta)$ with mean d and variance v/λ. The score test is based on $(\partial/\partial v)\ln h^*$ evaluated at $(\hat{d}, \hat{\lambda}, 0)$, where $\hat{d}$ and $\hat{\lambda}$ are the maximum likelihood estimates under $H_0 : v = 0$. We find

$$\frac{\partial \ln h^*}{\partial v} = \frac{\lambda(dt-1)^2 - t}{2 + v\left[\lambda(dt-1)^2 - t\right]}.$$

Evaluating at the constrained MLEs $(\hat{d}, \hat{\lambda}, 0)$ and summing over the sample we get the numerator of T_2. Further, some straightforward calculations show that the asymptotic variance of the concentrated score also has the same expression as that based on the normal mixing distribution. Thus, the score test for $H_0; v = 0$ based on h^* has an identical expression to the score test based on h. The virtue of the above derivation is that it holds for a wide class of mixing distributions. In other words, the test T_2 is valid for any mixing distribution as long as v is small. The numerator of the test statistic T_2 has an interesting intuitive appeal on rewriting it as

$$\mathrm{num}(T_2) = \frac{n}{2\hat{d}}\left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2} - 1\right). \tag{14}$$

In (14), $\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(t_i - \bar{t})^2$ is a method of moments estimator of the population variance, hence independent of the inverse Gaussian distribution, while $\hat{\sigma}^2 = (\hat{\lambda}\hat{d}^3)^{-1}$ is the maximum likelihood estimator of the population variance under the inverse Gaussian assumption. Thus (14) ought to be "*close*" to zero if the inverse Gaussian assumption is true. A similar rationale underlies Fisher's index of dispersion test for the Poisson distribution.

The test is also related to the information matrix test of White (1982), T^* say, which is a general test for model misspecification not dependent on mixture alternatives. To see this, let

$$k(t; \delta, \lambda) = \left(\frac{\partial g}{\partial \delta}\right)^2 + \frac{\partial^2 g}{\partial \delta^2}$$

where $g = \ln f(t; \delta, \lambda)$ and f is the density (1). The basis of White's test is that provided the model is correctly specified, $k(t; \delta, \lambda)$ has zero expectation, from standard

likelihood identities, whereas under misspecification it will, in general, be non-zero. An appropriate test for misspecification, therefore, is given by:

$$T^* = \sum_{i=1}^{n} k(t_i; \hat{\delta}, \hat{\lambda})$$

where $\hat{\delta}$ and $\hat{\lambda}$ are the maximum likelihood estimates of δ and λ under the inverse Gaussian assumption. Elementary calculation shows that T^* and $\mathrm{num}(T_2)$ are equivalent except for a multiplicative factor.

4 Score Tests for Inverse Gaussian Mixtures

When the data rejects the null hypothesis of a pure inverse Gaussian distribution in favor of some form of mixture, it is of interest to carry out further analysis for the mixing parameter v. Now, the regularity conditions for the standard asymptotic likelihood inference theories are satisfied by the models $h^\diamond$ and h^*. Test statistics can be derived based on these two types of mixture. Also, when, v is small such that h is close to a proper density, the standard methods of inference can also be applied. We now introduce the tests associated with these three models bearing in mind the conditions under which they apply.

We will concentrate on the score tests based on the so-called outer product of gradients (OPG) as the other tests are either computationally intensive (likelihood ratio and Wald), or unavailable (score test based on expected information) or problematic (score test based on Hessian matrix, which can give rise to a negative variance estimate). Let $S(v)$ be the score function of v, and G the $n \times 3$ gradient matrix whose ith row is the derivative of the log density with respect to θ corresponding to the ith observation. The OPG-based score statistic is defined as

$$T(v) = S(v)D(v) \tag{15}$$

where $D(v)$ is the square-root of the diagonal element of $(G^t G)^{-1}$ corresponding to v. All quantities are evaluated at $\hat{d}$ and $\hat{\lambda}$, the constrained MLEs of d and λ, given v.

4.1 Score test based on defective mixture

When v is small, h is close to a proper density, and the application of the standard likelihood inference will be correct to a high degree of approximation, the normalizing constant being effectively unity. The log likelihood is, ignoring the constant,

$$\ell(\boldsymbol{\theta}) = \frac{n}{2}\log(\lambda) - \frac{1}{2}\sum_{i=1}^{n}\log(vt_i + 1) - \frac{\lambda}{2}\sum_{i=1}^{n}\frac{(dt_i - 1)^2}{t_i(vt_i + 1)}.$$

The score function for v is

$$S(v) = \frac{\lambda}{2}\sum_{i=1}^{n}\left(\frac{dt_i - 1}{vt_i + 1}\right)^2 - \frac{1}{2}\sum_{i=1}^{n}\frac{t_i}{vt_i + 1}.$$

The ith row of the gradient matrix has elements

$$
\begin{aligned}
g_{di} &= -\lambda\frac{dt_i - 1}{vt_i + 1}, \\
g_{\lambda i} &= \frac{1}{2\lambda} - \frac{1}{2}\frac{(dt_i - 1)^2}{t_i(vt_i + 1)}, \\
g_{vi} &= \frac{\lambda}{2}\left(\frac{dt_i - 1}{vt_i + 1}\right)^2 - \frac{1}{2}\frac{t_i}{vt_i + 1}.
\end{aligned}
$$

Thus, the OPG score test for testing $H_0 : v = v_0$ takes the form

$$T_3(v_0) = S(v_0)D(v_0), \tag{16}$$

where all quantities are evaluated at $(\hat{d}, \hat{\lambda})$ with

$$\hat{d} = \left(\sum_{i=1}^{n}\frac{1}{vt_i + 1}\right)\left(\sum_{i=1}^{n}\frac{t_i}{vt_i + 1}\right)^{-1}, \qquad \hat{\lambda} = \frac{1}{\tilde{t}^{-1} - \hat{d}},$$

and $\tilde{t}$ is the harmonic mean of the t_is. The test is asymptotically standard normal. So, it is seen that implementation of the OPG score test is very simple. Section 5 presents simulation results for its finite sample performance. The test performs reasonably well when v is small.

4.2 Score test based on truncated-normal mixture

When v is not small, it is necessary to modify the defective mixture so that application of the standard likelihood theory is warranted. Based on the proper

11

mixture $h^\diamond(t, \boldsymbol{\theta})$ introduced in Section 2 and using the notation of Section 4.1, we have the log likelihood function,

$$\ell^\diamond(\boldsymbol{\theta}) = \ell(\boldsymbol{\theta}) + \sum_{i=1}^{n} \log \Phi\left(\left\{\frac{(v+d)^2\lambda}{v^2 t_i + v}\right\}^{\frac{1}{2}}\right) - n \log \Phi\left(\left\{\frac{d^2\lambda}{v}\right\}^{\frac{1}{2}}\right),$$

the score function for v is,

$$S^\diamond(v) = S(v) + \frac{\lambda^{\frac{1}{2}}}{2v^{\frac{3}{2}}} \sum_{i=1}^{n} \left(\frac{2v}{(vt_i+1)^{\frac{1}{2}}} - (v+d)\frac{(2vt_i+1)\xi_i}{(vt_i+1)^{\frac{3}{2}}} + d\xi_0\right),$$

where $\xi_i = \phi\big[(v+d)(\lambda/(v^2 t_i + v))^{\frac{1}{2}}\big]/\Phi\big[(v+d)(\lambda/(v^2 t_i + v))^{\frac{1}{2}}\big]$, and $\xi_0 = \phi(d/\sqrt{\lambda/v})$ $/\Phi(d/\sqrt{\lambda/v})$. The gradient matrix $G^\diamond$, corresponding to the ith observation, contains the three elements,

$$
\begin{aligned}
g_{di}^\diamond &= g_{di} + \left(\frac{\lambda}{v}\right)^{\frac{1}{2}}\left(\frac{\xi_i}{(vt_i+1)^{\frac{1}{2}}} - \xi_0\right), \\
g_{\lambda i}^\diamond &= g_{\lambda i} + \frac{1}{2(\lambda v)^{\frac{1}{2}}}\left(\frac{(v+d)\xi_i}{(vt_i+1)^{\frac{1}{2}}} - d\xi_0\right), \\
g_{vi}^\diamond &= g_{vi} - \frac{\lambda^{\frac{1}{2}}}{2v^{\frac{3}{2}}}\left(\frac{2v}{(vt_i+1)^{\frac{1}{2}}} - (v+d)\frac{(2vt_i+1)\xi_i}{(vt_i+1)^{\frac{3}{2}}} + d\xi_0\right).
\end{aligned}
$$

Evaluating all quantities at $(\hat{d}, \hat{\lambda})$ gives the score test

$$T_4(v_0) = S^\diamond(v_0)D^\diamond(v)$$

where $\hat{d}$ and $\hat{\lambda}$ are the constrained MLEs maximizing $\ell^\diamond(\boldsymbol{\theta})$ given $v = v_0$. The test is again asymptotically standard normal.

4.3 Score tests based on general mixtures

The score tests developed in Sections 4.1 and 4.2 are based on normal or truncated normal mixtures. It is desirable to develop a test that is less sensitive to the assumption of the mixing distribution. The following theorem provides a theoretical basis for a general score test.

Theorem. *The function $h^*(t; (4))$ defined in (6) is a proper probability density function provided that*

$$d > \frac{v}{4}\left(1 + \sqrt{1 + \frac{2}{\lambda v}}\right).$$

Furthermore, the score test for testing $H_0 : v = 0$, based on $h^*(t; \boldsymbol{\theta})$ and the expected information, has the same form as T_2 given in Section 2.

Proof: See Appendix.

The results of the theorem provides a basis for general inference on v when the condition is satisfied. It is more robust to the mixing distribution as only the first two moments of the mixing distribution are involved. As the expected information matrix is not available for general v, and the Hessian matrix is quite messy, we derive a score test based on the gradient. Writing the density (6) as $f(t; d, \lambda) m(t; \boldsymbol{\theta})$, and denoting the log likelihood and the gradient elements of the pure inverse Gaussian distribution by attaching a superscript ^{0}, we have the log likelihood,

$$\ell^*(\boldsymbol{\theta}) = \ell^0(d, \lambda) + \sum_{i=1}^{n} \log m(t_i; \boldsymbol{\theta}),$$

the score function for v,

$$S^*(v) = \sum_{i=1}^{n} \frac{m_v(t_i; \boldsymbol{\theta})}{m(t_i; \boldsymbol{\theta})},$$

and the gradient matrix G^* with typical elements of its ith row as

$$
\begin{aligned}
g_{di}^* &= g_{di}^0 + \frac{m_d(t_i; \boldsymbol{\theta})}{m(t_i; \boldsymbol{\theta})} \\
g_{\lambda i}^* &= g_{\lambda i}^0 + \frac{m_\lambda(t_i; d, \lambda, v)}{m(t_i; \boldsymbol{\theta})} \\
g_{vi}^* &= \frac{m_v(t_i; \boldsymbol{\theta})}{m(t_i; \boldsymbol{\theta})}
\end{aligned}
$$

where m_d, m_λ and m_v are the partial derivatives of m. The test statistic for testing $H_0 : v = v_0$ takes the form

$$T_5(v_0) = \sum_{i=1}^{n} \left(\frac{\lambda(\hat{d}t_i - 1)^2 - t_i}{2 + v_0 \big(\hat{\lambda}(\hat{d}t_i - 1)^2 - t_i \big)} \right) D^*(v_0), \tag{17}$$

where $D^*(v_0)$ is the square-root of the vv-element of $(G^{*T} G^*)^{-1}$, and is evaluated at $(\hat{d}, \hat{\lambda}, v_0)$, and $\hat{d}$ and $\hat{\lambda}$ are the constrained MLEs of d and λ at the null, which maximizes $\ell^*(\boldsymbol{\theta})$ for $v = v_0$. The test is asymptotically standard normal.

13

The use of a general function m above allows an easy extension of the results to other types of mixing distribution. For example, suppose δ has a variance v instead of v/λ, then we have

$$m(t; \boldsymbol{\theta}) = 1 + \frac{1}{2} v \lambda [\lambda (dt - 1)^2 - t]$$

and the condition for $f(t; d, \lambda) m(t; \boldsymbol{\theta})$ to be a proper density becomes

$$d > \frac{1}{4} \big(\lambda v + \sqrt{\lambda^2 v^2 + 2} \big).$$

Again, this condition is not restrictive. Note that the assumption that the variance of δ is v/λ by Whitmore (1986) has no clear physical interpretation. It is a matter of convenience. Thus, there is no reason why we could not have other forms of distribution for δ. In fact, this latter form simplifies the matter.

5   Monte Carlo simulations

We carry out Monte Carlo simulations to investigate finite sample properties of the test statistics introduced in Sections 3 and 4, and to compare the tests in Section 3 with the smooth test of Ducharme (2001). Pure inverse Gaussian random variates are generated from an algorithm given in Michael at al. (1976). An inverse Gaussian mixture random variate is generated in two steps: (i) generate a random variate, δ say, from the mixing distribution, (ii) then generate an inverse Gaussian variate using δ and λ as parameters. From each sample generated, the test statistic is calculated and it is checked whether it exceeds the upper 10%, 5% or 1% values of the reference distribution. The proportion of the test statistic values out of 10,000 exceeding 10% (5% or 1%) nominal value gives a Monte Carlo estimate of the actual size of the test.

Table 1 summarizes the size of the test T_1 and Table 2 the size of T_2. First, note that the size of the test can be affected heavily by the skewness of the population. The more skewed the distribution, the poorer the test performs. Note that the measure of skewness for the inverse Gaussian distribution is given by $3/\sqrt{d\lambda}$. A population with a skewness measure of 3 can be considered as a quite skewed population. The parameter configurations given in Tables 1 and 2 are such that the population skewness increases from one row to next. The first set of parameters gives a population skewness measure of 0.3, whereas the last set of parameters gives a population skewness measure of 30.

14

We can see that both T_1 and T_2 perform rather poorly at the last set of parameter values even when the sample size is 200. The tests deteriorates quite rapidly with the increase of the skewness measure.

The test T_2 performs significantly poorer than T_1 even when the population is not so skewed, indicating the cost of estimating parameters is rather high. Increasing the sample size, significantly improves the performance of both statistics. The tests perform better at the 10% level than at the 5% and 1%, in particular at 1%, where the sizes of the tests exceed the nominal levels quite a bit. This indicates that the distributions of T_1 and T_2 both have a long right tail.

Tables 1 and 2 here

Tables 3-5 summarize the size-adjusted empirical power of the tests T_1 and T_2 introduced in Section 3 of the current paper, and the test R_3 of Ducharme (2001), which is in fact what Ducharme recommenced based on the joint considerations of ease of computation and power. As the sizes of the three tests can be different, it is thus only meaningful to compare the size-adjusted power. To this end, we use the simulated critical values which are obtained based on 50,000 Monte Carlo samples under each parameter configuration.

As expected, the results show that the test T_1 is the most powerful one, followed by T_2 and R_3. This is because T_1 is derived under the assumption that the parameters are known, and both T_1 and T_2 are derived based on an inverse Gaussian mixture. In contrast, R_3 is a smooth test without a specific alternative. What is interesting is that our test T_2 can be much more powerful than R_3 of Ducharme (2001) in detecting an inverse Gaussian mixture.

The simulation results (not reported for brevity) show that the above conclusions also hold when the mixing distribution in the inverse Gaussian mixture deviates away from (truncated) normality. This is consistent with the discussions following the introduction of T_2. However, as the test T_2 is one directional against a general inverse Gaussian mixture, it should be well expected that the test T_2 would have a low power compared with the smooth test in detecting other type of alternatives such as gamma, lognormal and the other distributions considered in Ducharme (2001). See the last section for more discussions for this and related issues.

Tables 3-5 here

Table 6 summarizes the simulated sizes of the three score tests for testing the general mixing parameter, with a one-sided alternative on the right. The mixing distribution is chosen to be normal. The parameter configurations are chosen so that essentially there is no truncation. We see from the table that when v is small, the three tests perform very similarly. When v is not small, the test based on a general mixing distribution rejects more frequently.

Table 6 here

6   Discussions

A referee has conjectured that the test based on T_2 may be related to the first component of a smooth test derived from the pure inverse Gaussian. Such a conjecture is plausible, since the smooth test paradigm is based on alternatives of the form $h(t;\theta)w(t;\theta,\beta)$ as in our equation (6). A comprehensive treatment of such smooth tests is given in the recent research monograph by Rayner et al. (2009). A variety of tests may be derived depending on the nature of $w(t;\theta,\beta)$. The original idea is due to Neyman (1937), and was developed to overcome deficiencies in Pearson's chi-squared test. Rayner and co-workers, in a series of papers, have used an alternative formulation to that of Neyman in which $w(t;\theta,\beta) = C(\theta,\beta)\exp\{\sum_{i=1}^{k}\beta_i h_i(t;\theta)\}$ as in Rayner et al. page 95, equation (6.1). Note that we have interchanged θ and β from Rayner et al. to be consistent with our own notation. In (6), for example, $\theta = (\delta,,\lambda)$, $f(t;\theta)$ is the pure inverse Gaussian and $\{h_i(t;\theta) : i = 1,\cdots,k\}$ are orthonormal polynomials with respect to $f(t;\theta)$. A test of $H_0 : \beta_1 = \cdots = \beta_k = 0$, is then a smooth test of goodness of fit for the inverse Gaussian. Using the recurrence relationship (A.1) of Rayner et al. (2009), the first orthonormal polynomial is:

$$h_1(t) = (t - \mu)/\sqrt{\mu^3/\lambda},$$

where $\mu = 1/\delta$, and the second (after some tedious algebra) is:

$$h_2(t) = \frac{t^2/\mu^2 - 3t/\lambda - 2t/\mu + 2\mu/\lambda + 1}{2(\mu/\lambda)^{3/2}(3\mu^2/\lambda + 1)}.$$

16

The first orthogonal component is identically zero when evaluated at the MLEs $\hat{\mu}$ and $\hat{\lambda}$ and does not provide a useful test. The second component gives the test statistic

$$V_2 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} h_2(t_i, \hat{\theta}).$$

We now explore the relationship between our test statistic T_2 and the smooth component V_2. First evaluating the numerator and neglecting the normalizing constant we find that V_2 is proportional to

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (t_i/\bar{t})^2 - \sqrt{n}(\hat{\mu}/\hat{\lambda} + 1) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} [(t_i/\bar{t})^2 - (t_i/\bar{t})^{-1}].$$

Thus the test based on the smooth component is effectively the sample equation version of the moment identity $E(T^2)/E(T)^3 - E(T^{-1}) = 0$ which holds for $T \sim IG(\mu, \lambda)$ distribution. Our test on the other hand motivated by alternatives involving heterogeneity in the mean parameter involves a comparison of the method of moments estimator of the variance and the MLE of the variance. Nevertheless, since it is often the case, for example for the Poisson distribution, that smooth tests result in comparisons between first and second moments under the model, a referee has raised an intriguing question which is worth pursuing in future work. The key question may well revolve around the types of alternative models envisioned and in our case our test seems well motivated for the case of mixture-type alternatives. A further problem with smooth tests for the IG, pointed out by Ducharme (2001), is that they can exhibit inconsistency when the true distribution has $E(1/X) = \infty$, such as for the exponential or the Weibull with shape parameter less than one. This problem does not arise for our tests against alternatives of the mixed inverses Gaussian form, since we have assumed a normal distribution on the drift parameter corresponding to the reciprocal mean of the IG distribution, which ensures that the marginal mean of the reciprocal random variable is finite. This is also true for our truncated normal distributions on the drift parameter and for any distribution with finite first moments. We have not studied the consistency aspect for general alternatives, since we were motivated by heterogeneity in the drift which is important in practice. Nevertheless, in future work we hope to explore this intriguing issue and more general comparisons between the work here and alternative tests of goodness of fit. In this regard we thank the reviewer for pointing out, that in Ducharme (2001), it is noted that tests based

17

on the empirical cumulative distribution function need to be used with caution and that errors have occurred in some published tables of quantiles for these tests. We refer the reader to Ducharme (2001) for references and detailed caveats.

APPENDIX

Proof of the Theorem: Writing $h^*(t; \boldsymbol{\theta}) = f(t; d, \lambda) m(t; \boldsymbol{\theta})$, it suffices to show that i) $\int_0^\infty h^*(t; \boldsymbol{\theta}) dt = 1$, and ii) $m(t; \boldsymbol{\theta}) \geq 0$, $\forall t$, under the condition given in the theorem.

For i), we have,

$$
\begin{aligned}
\int_0^\infty h^*(t; \boldsymbol{\theta}) dt &= \int_0^\infty f(t; d, \lambda) dt + \frac{v}{2} \int_0^\infty [\lambda (dt - 1)^2 - t] f(t; d, \lambda) dt \\
&= 1 + \frac{v\lambda}{2} \int_0^\infty (dt - 1)^2 f(t; d, \lambda) dt - \frac{v}{2} \int_0^\infty t f(t; d, \lambda) dt \\
&= 1 + \frac{v\lambda}{2} \frac{1}{\lambda d} - \frac{v}{2} \frac{1}{d} = 1.
\end{aligned}
$$

For ii), we have,

$$
\frac{\partial m}{\partial t} = \frac{v}{2} [2\lambda d(dt - 1) - 1], \text{ and } \frac{\partial^2 m}{\partial t^2} = v\lambda d^2 > 0,
$$

which show that m has a unique minimum at $t_0 = (1 + 2\lambda d)/(2\lambda d^2)$. Evaluating m at t_0, we obtain,

$$
\begin{aligned}
m(t_0; \boldsymbol{\theta}) &= 1 + \frac{v}{2} \left[\lambda \left(d \frac{1 + 2\lambda d}{2\lambda d^2} - 1 \right)^2 - \frac{1 + 2\lambda d}{2\lambda d^2} \right] \\
&= \frac{8\lambda d^2 - 4\lambda vd - v}{8\lambda d^2}.
\end{aligned}
$$

So, for m to be non-negative, it is sufficient to have

$$
m(t_0; \boldsymbol{\theta}) \geq 0
$$

which is equivalent to having

$$
8\lambda d^2 - 4\lambda vd - v \geq 0.
$$

Since d is non-negative, the admissible solution to the above inequality is

$$
d > \frac{4\lambda v + \sqrt{(4\lambda v)^2 + 32\lambda v}}{16\lambda} = \frac{v}{4} \left(1 + \sqrt{1 + \frac{2}{\lambda v}} \right).
$$

REFERENCES

Aalen, O. O. and Gjessing, H. K. (2001). Understanding the shape of the hazard rate: a process point of view. *Statistical Science*, **16**, 1-22.

Banerjee, A. K. and Bhattacharyya, G. K. (1976). A purchase incidence model with inverse Gaussian interpurchase times. *Journal of the American Statistical Association*, **71**, 823-829.

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

Chhikara, R. S. and Folks, J. L. (1978). The inverse Gaussian distribution and its statistical application - a review. *J. R. Statist. Soc.* B, **40**, 263-289.

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

Cox, D. R. and Hinkeley, D. V. (1974). *Theoretical Statistics.* Chapman and Hall, London.

Desmond, A. F. and Chapman, G. R. (1993). Modelling task completion data with inverse Gaussian mixtures. *Applied Statistics*, **42**, 603-613.

Ducharme, G. R. (2001). Goodness-of-fit tests for the inverse Gaussian and related distributions. *Test*, **10**, 271-290.

Edgeman, R. L. (1989). Inverse Gaussian control charts. *Australian Journal of Statistics*, **31**, 78-84.

Edgeman, R. L. (1990). Assessing the inverse Gaussian distribution assumption. *IEEE Transactions on Reliability.* **R39**, 352-355.

Hawkins, D. M. and Olwell, D. H. (1997). Inverse Gaussian cumulative sum control charts for location and shape. *The Statistician*, **46**, 323-335.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions.* John Wiley & Sons, New York.

Lehmann,E. L. (1983). *Theory of Point Estimation.* John Wiley & Sons, New York.

Lancaster, T. (1972). A stochastic model for the duration of a strike. *J. R. Statist. Soc.* A, **135**, 257-271.

Letac, G., Seshadri, V. and Whitmore, G. A. (1985). An exact chi-squared decomposition theorem for inverse Gaussian variates. *J. R. Statist. Soc.* B, **47**, 476-481.

Mardia K. V. and Kent, J. T. (1991). Rao score tests for goodness of fit and independence. *Biometrika*, **78**, 355-363.

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

Neyman, J. (1937). Smooth test for goodness of fit. *Skandinavisk Aktuarietidskrift* **20**, 149-199.

Pavur, R. J., Edgeman R. L. and Scott, R. C. (1992). Quadratic statistics for the goodness-of-fit test of the inverse Gaussian distribution. *IEEE Transactions on Reliability*, **41**, 118-123.

Rayner, J. C. W., Thas, O. and Best, D. J. (2009). *Smooth Tests of Goodness of Fit Using R, 2nd ed.* John Wiley & Sons, Singapore.

Seshadri, V. (1993). *The inverse Gaussian distribution: A Case Study in Exponential Families.* Oxford University Press, Oxford.

Seshadri, V. (1998). *The inverse Gaussian distribution.* Springer, New York.

Silvapulle, M. J. and Sivapulle, P. (1995). A score test against one-sided alternatives. *Journal of the American Statistical Association* **90**, 342-349.

Silvapulle, M. J. and Sen, P. K. (2005). *Constrained Statistical Inference: Inequality, order and Shape Restrictions.* John Wiley & Sons, Hoboken, New Jersey.

Verbeke, G. and Molenberghs, G. (2003). The use of the score test for inference on variance components. *Biometrics*, **59**, 254-262.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, **50**, 1-25.

Whitmore, G. A. (1979). An inverse Gaussian model for labour turnover. *J. R. Statist. Soc.* A, **142**, 468-478.

Whitmore, G. A. (1986). Normal-gamma mixtures of inverse Gaussian distributions. *Scand. J. Statist.*, **13**, 211-220.

Figure 1: Plots of the PDF of Improper Inverse Gaussian Mixture

Plots of Mixed IG Distribution, $\lambda=1$, $d=1$

Plots of Mixed IG Distribution, $\lambda=5$, $d=1$

Plots of Mixed IG Distribution, $\lambda=10$, $d=1$

22
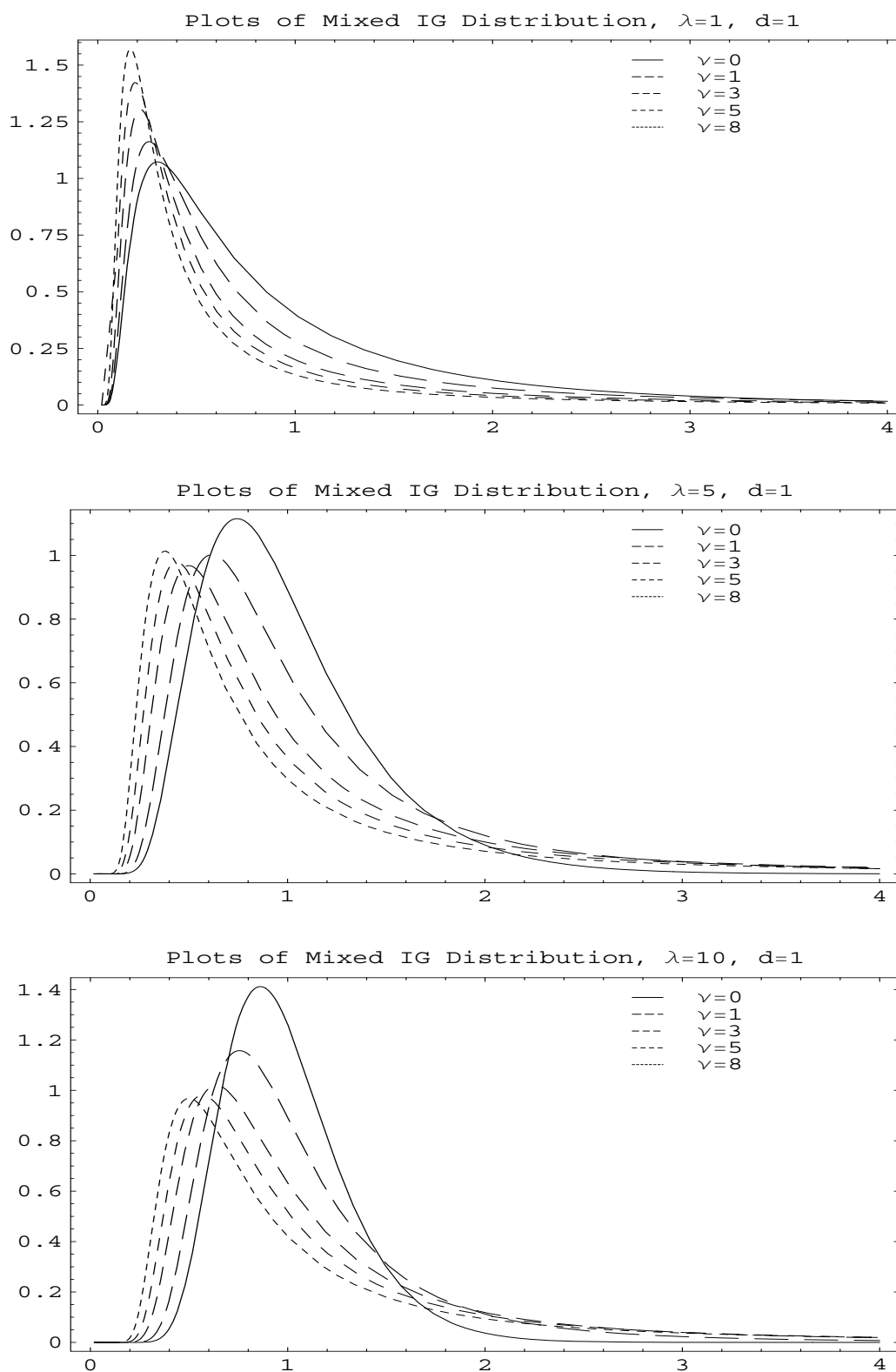
Figure 2: Plots for Comparing Proper and Improper IG Mixtures
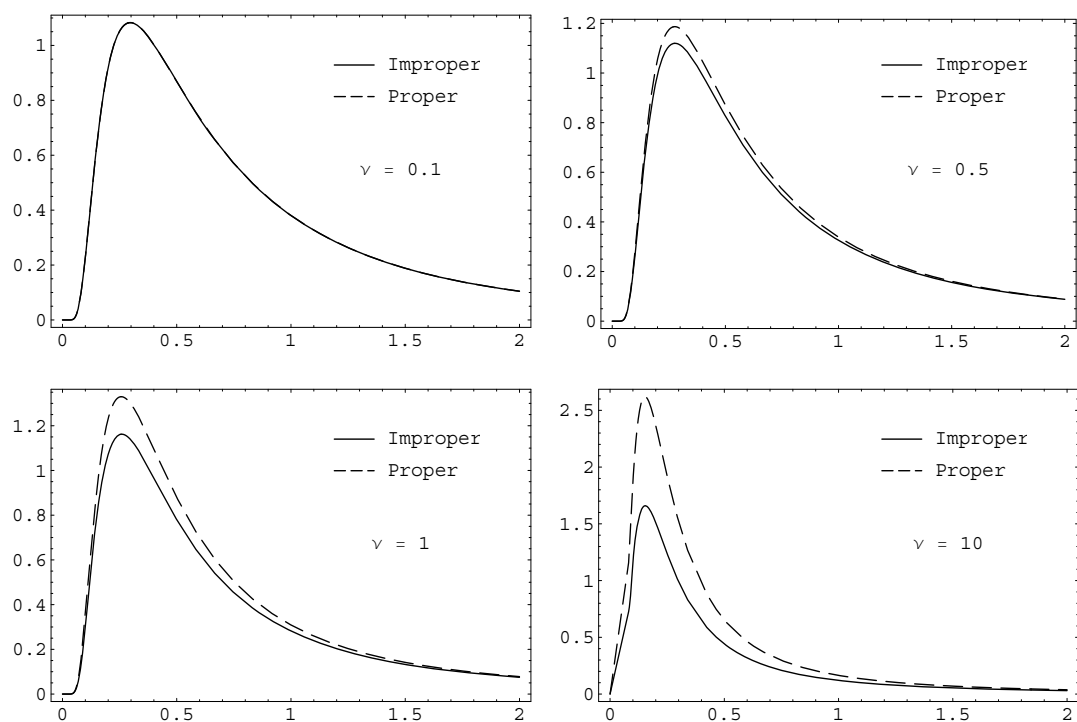
Table 1: Simulated sizes of the score test with known parameters (T_1)

$\lambda\backslash\alpha$	$n = 50$			$n = 100$			$n = 200$			$n = 400$		
	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
10.0	.099	.058	.018	.103	.059	.017	.107	.056	.015	.099	.054	.013
9.0	.102	.059	.019	.104	.057	.017	.105	.056	.015	.103	.055	.013
8.0	.106	.060	.017	.104	.058	.017	.105	.062	.019	.101	.054	.012
7.0	.103	.060	.019	.106	.058	.017	.103	.055	.014	.101	.053	.014
6.0	.110	.064	.023	.104	.059	.016	.100	.058	.016	.105	.055	.012
5.0	.102	.061	.019	.099	.059	.019	.106	.057	.016	.102	.054	.012
4.0	.098	.060	.021	.102	.060	.018	.101	.054	.016	.100	.052	.014
3.0	.103	.061	.021	.100	.056	.017	.104	.056	.016	.101	.054	.013
2.0	.101	.060	.023	.108	.065	.023	.102	.057	.017	.106	.057	.015
1.0	.099	.063	.024	.104	.062	.023	.110	.066	.022	.103	.059	.018
0.8	.099	.062	.025	.104	.061	.020	.107	.065	.022	.100	.054	.017
0.6	.106	.068	.028	.105	.064	.022	.107	.062	.021	.103	.057	.016
0.4	.095	.062	.029	.102	.064	.025	.101	.063	.024	.101	.060	.018
0.2	.095	.062	.030	.097	.064	.030	.103	.064	.026	.102	.062	.023
0.1	.082	.058	.029	.096	.063	.032	.100	.064	.028	.101	.065	.024

Table 2: Simulated sizes of the score test with unknown parameters (T_2), $d = 10$

$\lambda\backslash\alpha$	$n = 50$			$n = 100$			$n = 200$			$n = 400$		
	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01	0.1	0.05	0.01
10.0	.082	.042	.011	.089	.047	.012	.097	.051	.013	.098	.052	.012
9.0	.083	.044	.014	.092	.050	.013	.091	.052	.012	.097	.053	.012
8.0	.083	.044	.012	.097	.052	.014	.096	.050	.012	.100	.054	.012
7.0	.083	.043	.011	.087	.050	.014	.095	.051	.014	.090	.049	.013
6.0	.085	.047	.012	.092	.047	.013	.094	.051	.013	.093	.049	.012
5.0	.084	.046	.013	.091	.048	.014	.093	.052	.015	.095	.049	.012
4.0	.088	.044	.012	.092	.050	.015	.090	.047	.013	.099	.052	.014
3.0	.082	.044	.016	.087	.045	.012	.096	.053	.015	.095	.051	.012
2.0	.083	.045	.014	.091	.052	.016	.093	.052	.014	.094	.050	.012
1.0	.081	.049	.016	.093	.052	.017	.090	.050	.016	.100	.055	.014
0.8	.081	.048	.016	.090	.054	.019	.094	.054	.017	.093	.052	.015
0.6	.075	.043	.017	.081	.047	.016	.092	.054	.018	.092	.053	.016
0.4	.076	.044	.018	.080	.045	.017	.090	.049	.018	.095	.051	.017
0.2	.059	.036	.013	.073	.046	.016	.083	.052	.019	.096	.056	.021
0.1	.048	.029	.012	.067	.042	.018	.075	.047	.018	.084	.049	.019

Table 3: Size-adjusted empirical power of the tests at $\alpha = 0.1$

v	$n=50$			$n=100$			$n=200$			$n=400$		
	T_1	T_2	R_3	T_1	T_2	R_3	T_1	T_2	R_3	T_1	T_2	R_3
$d=10, \lambda=1$												
0	.105	.104	.105	.099	.100	.105	.099	.102	.096	.103	.098	.101
1	.264	.153	.110	.339	.184	.116	.473	.232	.124	.649	.300	.157
2	.444	.212	.138	.619	.277	.161	.817	.395	.217	.961	.555	.327
3	.626	.282	.174	.809	.386	.238	.958	.568	.362	.998	.778	.565
4	.760	.340	.217	.925	.504	.333	.992	.702	.511	1.00	.897	.768
5	.843	.407	.280	.968	.595	.425	.999	.809	.647	1.00	.963	.893
6	.901	.473	.338	.990	.685	.533	1.00	.878	.763	1.00	.986	.956
7	.942	.536	.396	.996	.744	.610	1.00	.930	.857	1.00	.996	.984
8	.965	.592	.459	.999	.805	.692	1.00	.961	.914	1.00	.999	.995
9	.978	.632	.511	.999	.847	.758	1.00	.977	.950	1.00	1.00	.998
10	.986	.675	.557	1.00	.884	.812	1.00	.985	.971	1.00	1.00	1.00
11	.992	.710	.607	1.00	.911	.852	1.00	.993	.981	1.00	1.00	1.00
12	.994	.743	.649	1.00	.933	.888	1.00	.994	.989	1.00	1.00	1.00
$d=4, \lambda=.25$												
.0	.100	.095	.103	.097	.098	.103	.100	.102	.098	.096	.102	.098
.1	.179	.135	.096	.217	.167	.106	.276	.212	.112	.356	.257	.110
.2	.273	.185	.114	.359	.249	.122	.497	.362	.141	.656	.486	.177
.3	.374	.242	.126	.497	.348	.154	.670	.490	.210	.863	.690	.294
.4	.444	.284	.153	.622	.439	.201	.803	.633	.303	.948	.826	.463
.5	.529	.339	.170	.718	.527	.262	.892	.735	.407	.985	.908	.631
.6	.601	.386	.206	.786	.590	.330	.939	.806	.507	.995	.956	.759
.7	.651	.415	.236	.840	.647	.382	.966	.855	.598	.999	.975	.843
.8	.695	.442	.255	.879	.687	.426	.982	.891	.655	.999	.986	.895
.9	.731	.462	.281	.904	.724	.473	.987	.920	.727	1.00	.991	.929
1.0	.769	.484	.298	.919	.746	.502	.991	.928	.754	1.00	.993	.954
1.1	.787	.492	.324	.935	.763	.537	.994	.941	.792	1.00	.996	.964
1.2	.804	.507	.337	.947	.772	.558	.996	.949	.821	1.00	.995	.970
$d=1, \lambda=4$												
.00	.102	.097	.102	.098	.098	.101	.097	.098	.104	.103	.107	.102
.05	.202	.141	.109	.235	.167	.109	.307	.200	.105	.416	.259	.126
.10	.306	.193	.122	.410	.241	.134	.574	.348	.167	.767	.475	.241
.15	.426	.256	.155	.595	.341	.181	.780	.490	.266	.935	.669	.399
.20	.542	.306	.179	.720	.431	.245	.903	.619	.378	.987	.831	.594
.25	.642	.370	.232	.830	.528	.329	.957	.736	.515	.998	.917	.768
.30	.721	.427	.271	.891	.611	.412	.985	.822	.641	1.00	.962	.877
.35	.780	.474	.326	.938	.693	.509	.994	.884	.746	1.00	.987	.948
.40	.836	.530	.384	.961	.750	.579	.998	.930	.826	1.00	.994	.974
.45	.867	.577	.430	.979	.795	.654	.999	.953	.882	1.00	.997	.988
.50	.901	.616	.472	.988	.835	.705	1.00	.971	.926	1.00	.999	.996
.55	.925	.654	.516	.993	.866	.763	1.00	.977	.949	1.00	.999	.998
.60	.940	.679	.554	.995	.893	.806	1.00	.987	.967	1.00	.999	.999

25

Table 4: Size-adjusted empirical power of the tests at $\alpha = 0.05$

	$n = 50$			$n = 100$			$n = 200$			$n = 400$		
v	T_1	T_2	R_3	T_1	T_2	R_3	T_1	T_2	R_3	T_1	T_2	R_3
$d = 10, \lambda = 1$												
0	.052	.050	.051	.051	.052	.053	.051	.050	.048	.052	.049	.050
1	.167	.089	.060	.232	.112	.062	.332	.135	.065	.524	.202	.089
2	.335	.132	.074	.501	.187	.096	.724	.282	.133	.921	.420	.210
3	.514	.189	.107	.724	.283	.154	.924	.440	.243	.996	.660	.421
4	.664	.245	.142	.871	.386	.229	.984	.597	.383	1.00	.831	.647
5	.778	.306	.185	.942	.477	.307	.998	.719	.535	1.00	.928	.819
6	.857	.374	.245	.977	.577	.411	1.00	.817	.666	1.00	.970	.918
7	.901	.418	.288	.991	.664	.506	1.00	.890	.777	1.00	.988	.968
8	.937	.487	.353	.997	.732	.592	1.00	.931	.860	1.00	.996	.987
9	.960	.532	.406	.999	.782	.664	1.00	.954	.909	1.00	.998	.995
10	.975	.586	.456	.999	.824	.731	1.00	.976	.946	1.00	1.00	.999
11	.983	.623	.505	1.00	.865	.780	1.00	.985	.969	1.00	1.00	.999
12	.989	.661	.553	1.00	.894	.832	1.00	.991	.981	1.00	1.00	1.00
$d = 4, \lambda = .25$												
.0	.046	.049	.055	.051	.052	.049	.051	.047	.050	.048	.050	.054
.1	.105	.080	.051	.134	.097	.047	.177	.124	.049	.251	.172	.052
.2	.181	.116	.057	.269	.172	.054	.376	.244	.064	.548	.376	.098
.3	.275	.162	.065	.402	.253	.077	.577	.387	.120	.793	.586	.182
.4	.362	.207	.078	.530	.331	.112	.730	.526	.191	.921	.761	.335
.5	.446	.259	.101	.643	.422	.161	.843	.639	.279	.971	.866	.491
.6	.510	.293	.116	.716	.483	.208	.906	.724	.373	.991	.926	.639
.7	.570	.322	.140	.781	.550	.256	.944	.795	.465	.996	.960	.749
.8	.619	.344	.159	.830	.595	.295	.968	.839	.537	.999	.975	.829
.9	.657	.362	.177	.860	.614	.333	.980	.869	.595	1.00	.983	.868
1.0	.686	.365	.186	.886	.643	.368	.986	.891	.654	.999	.988	.908
1.1	.725	.375	.201	.909	.668	.397	.988	.897	.685	1.00	.989	.930
1.2	.749	.383	.220	.923	.691	.424	.992	.909	.716	1.00	.991	.947
$d = 1, \lambda = 4$												
.00	.047	.051	.051	.048	.052	.054	.053	.049	.048	.052	.050	.051
.05	.123	.078	.051	.152	.096	.056	.202	.111	.057	.300	.163	.067
.10	.210	.112	.057	.292	.159	.070	.439	.221	.091	.664	.347	.133
.15	.329	.162	.077	.468	.240	.105	.685	.369	.155	.897	.561	.277
.20	.439	.207	.106	.618	.323	.156	.845	.509	.255	.974	.732	.440
.25	.547	.270	.139	.744	.419	.219	.928	.633	.379	.995	.857	.643
.30	.623	.320	.180	.838	.505	.298	.972	.747	.516	1.00	.933	.796
.35	.703	.377	.222	.896	.581	.381	.989	.821	.632	1.00	.972	.893
.40	.771	.418	.267	.936	.654	.463	.997	.878	.739	1.00	.988	.950
.45	.819	.473	.314	.958	.713	.533	.999	.920	.816	1.00	.994	.975
.50	.857	.519	.357	.977	.766	.612	.999	.947	.872	1.00	.998	.990
.55	.892	.557	.411	.984	.800	.663	1.00	.964	.915	1.00	.999	.996
.60	.914	.588	.449	.991	.831	.716	1.00	.974	.934	1.00	.999	.999

Table 5: Size-adjusted empirical power of the tests at $\alpha = 0.01$

	$n = 50$			$n = 100$			$n = 200$			$n = 400$		
v	T_1	T_2	R_3	T_1	T_2	R_3	T_1	T_2	R_3	T_1	T_2	R_3
$d = 10, \lambda = 1$												
0	.010	.009	.011	.011	.009	.010	.009	.008	.010	.009	.010	.012
1	.059	.024	.012	.094	.034	.009	.148	.044	.011	.283	.067	.012
2	.166	.046	.015	.287	.068	.014	.516	.119	.020	.801	.217	.042
3	.319	.077	.021	.537	.126	.029	.819	.236	.057	.978	.432	.145
4	.473	.114	.037	.740	.202	.055	.952	.377	.127	.999	.656	.312
5	.619	.159	.055	.867	.284	.099	.991	.514	.215	1.00	.813	.523
6	.714	.205	.080	.936	.376	.151	.998	.643	.347	1.00	.910	.710
7	.804	.242	.103	.972	.466	.216	1.00	.753	.483	1.00	.964	.850
8	.861	.302	.145	.987	.541	.291	1.00	.826	.596	1.00	.984	.929
9	.910	.357	.180	.994	.618	.364	1.00	.890	.711	1.00	.994	.968
10	.938	.401	.219	.998	.675	.436	1.00	.922	.782	1.00	.998	.987
11	.955	.445	.259	.999	.734	.513	1.00	.952	.851	1.00	.999	.995
12	.968	.496	.298	1.00	.779	.570	1.00	.967	.898	1.00	1.00	.998
$d = 4, \lambda = .25$												
.0	.010	.009	.011	.009	.009	.010	.009	.009	.009	.011	.011	.009
.1	.037	.018	.008	.047	.024	.009	.069	.041	.008	.107	.066	.009
.2	.087	.034	.008	.132	.063	.007	.213	.114	.006	.362	.193	.009
.3	.151	.062	.007	.242	.118	.008	.393	.223	.013	.644	.404	.026
.4	.223	.087	.006	.377	.188	.012	.575	.346	.029	.839	.590	.077
.5	.305	.121	.010	.478	.253	.022	.732	.476	.058	.943	.743	.161
.6	.378	.144	.009	.582	.314	.032	.830	.577	.100	.972	.834	.258
.7	.446	.156	.012	.666	.367	.039	.881	.633	.129	.990	.893	.373
.8	.502	.173	.011	.725	.398	.052	.928	.702	.164	.996	.930	.468
.9	.538	.179	.011	.777	.421	.059	.952	.739	.209	.998	.948	.543
1.0	.573	.170	.014	.810	.450	.069	.966	.764	.237	.999	.963	.611
1.1	.603	.178	.017	.833	.475	.072	.976	.795	.274	1.00	.971	.664
1.2	.626	.183	.018	.865	.478	.078	.983	.805	.290	1.00	.975	.708
$d = 1, \lambda = 4$												
.00	.009	.010	.011	.012	.010	.011	.011	.011	.011	.010	.010	.009
.05	.040	.022	.008	.051	.028	.008	.074	.033	.008	.128	.056	.010
.10	.092	.039	.009	.148	.057	.011	.237	.087	.010	.444	.166	.015
.15	.165	.061	.012	.280	.101	.014	.471	.178	.018	.763	.337	.045
.20	.256	.093	.017	.440	.158	.026	.698	.300	.051	.931	.539	.125
.25	.369	.133	.032	.585	.245	.051	.837	.422	.099	.985	.711	.267
.30	.452	.171	.039	.704	.315	.087	.924	.555	.181	.997	.833	.454
.35	.549	.223	.063	.803	.402	.140	.965	.653	.283	1.00	.918	.639
.40	.626	.259	.083	.862	.466	.183	.988	.753	.398	1.00	.960	.775
.45	.689	.294	.102	.913	.547	.248	.995	.818	.501	1.00	.980	.870
.50	.755	.344	.128	.941	.596	.296	.998	.869	.598	1.00	.990	.930
.55	.806	.386	.155	.961	.649	.359	.999	.908	.684	1.00	.996	.962
.60	.834	.404	.170	.975	.692	.404	.999	.930	.742	1.00	.996	.980

Table 6: Simulated sizes of score tests for IG mixture, $\delta \sim N(10, v/\lambda)$

v_0	λ	$\alpha = 0.1$			$\alpha = 0.05$			$\alpha = 0.01$		
		$T_3(v_0)$	$T_4(v_0)$	$T_5(v_0)$	$T_3(v_0)$	$T_4(v_0)$	$T_5(v_0)$	$T_3(v_0)$	$T_4(v_0)$	$T_5(v_0)$
$n = 50$										
0.01	10	0.146	0.144	0.144	0.088	0.086	0.086	0.025	0.024	0.025
0.01	8	0.147	0.143	0.149	0.087	0.086	0.091	0.025	0.023	0.025
0.01	6	0.141	0.142	0.142	0.083	0.079	0.085	0.021	0.022	0.025
0.01	3	0.138	0.135	0.134	0.081	0.077	0.077	0.019	0.020	0.022
0.01	1	0.116	0.114	0.111	0.059	0.061	0.058	0.011	0.014	0.013
0.10	10	0.143	0.150	0.166	0.088	0.091	0.107	0.026	0.027	0.036
0.10	8	0.141	0.142	0.167	0.085	0.087	0.106	0.025	0.024	0.035
0.10	6	0.144	0.142	0.161	0.087	0.083	0.097	0.022	0.021	0.029
0.10	3	0.129	0.128	0.142	0.074	0.070	0.082	0.021	0.021	0.024
0.10	1	0.111	0.119	0.120	0.064	0.062	0.067	0.014	0.015	0.017
1.00	10	0.140	0.146	0.255	0.083	0.089	0.192	0.024	0.027	0.099
1.00	8	0.142	0.146	0.243	0.087	0.082	0.178	0.024	0.025	0.095
1.00	6	0.141	0.141	0.239	0.083	0.085	0.174	0.025	0.024	0.087
1.00	3	0.134	0.133	0.215	0.076	0.075	0.151	0.020	0.020	0.068
1.00	1	0.114	0.110	0.176	0.063	0.061	0.110	0.013	0.014	0.044
$n = 100$										
0.01	10	0.122	0.125	0.124	0.069	0.068	0.067	0.017	0.017	0.017
0.01	8	0.123	0.126	0.125	0.065	0.067	0.069	0.013	0.015	0.017
0.01	6	0.119	0.119	0.124	0.065	0.064	0.065	0.016	0.015	0.014
0.01	3	0.119	0.110	0.113	0.064	0.061	0.059	0.015	0.014	0.012
0.01	1	0.092	0.096	0.102	0.045	0.047	0.052	0.009	0.008	0.008
0.10	10	0.127	0.123	0.139	0.070	0.069	0.082	0.017	0.017	0.025
0.10	8	0.125	0.123	0.139	0.067	0.068	0.084	0.016	0.018	0.023
0.10	6	0.119	0.112	0.123	0.064	0.060	0.070	0.016	0.015	0.019
0.10	3	0.112	0.111	0.116	0.060	0.060	0.066	0.012	0.013	0.015
0.10	1	0.094	0.097	0.100	0.043	0.045	0.051	0.007	0.008	0.010
1.00	10	0.125	0.122	0.195	0.071	0.068	0.141	0.017	0.017	0.064
1.00	8	0.126	0.128	0.193	0.068	0.070	0.130	0.015	0.017	0.063
1.00	6	0.120	0.124	0.194	0.066	0.070	0.135	0.014	0.018	0.060
1.00	3	0.112	0.114	0.174	0.059	0.060	0.119	0.012	0.015	0.047
1.00	1	0.105	0.101	0.147	0.050	0.052	0.089	0.008	0.010	0.029
$n = 200$										
0.01	10	0.115	0.109	0.114	0.060	0.059	0.060	0.013	0.013	0.013
0.01	8	0.110	0.107	0.113	0.055	0.057	0.057	0.012	0.011	0.013
0.01	6	0.108	0.105	0.112	0.057	0.053	0.057	0.011	0.012	0.010
0.01	3	0.105	0.107	0.103	0.053	0.052	0.051	0.009	0.011	0.010
0.01	1	0.093	0.092	0.097	0.044	0.041	0.046	0.007	0.006	0.007
0.10	10	0.111	0.115	0.124	0.057	0.060	0.067	0.014	0.013	0.020
0.10	8	0.112	0.107	0.120	0.057	0.055	0.066	0.014	0.012	0.018
0.10	6	0.111	0.105	0.118	0.056	0.053	0.063	0.013	0.011	0.017
0.10	3	0.106	0.099	0.108	0.052	0.050	0.059	0.010	0.011	0.013
0.10	1	0.093	0.094	0.095	0.043	0.043	0.047	0.007	0.006	0.007
1.00	10	0.113	0.110	0.169	0.059	0.058	0.112	0.014	0.012	0.046
1.00	8	0.110	0.114	0.157	0.057	0.059	0.100	0.014	0.014	0.042
1.00	6	0.108	0.108	0.160	0.058	0.053	0.105	0.012	0.009	0.042
1.00	3	0.102	0.101	0.143	0.048	0.050	0.091	0.011	0.010	0.032
1.00	1	0.094	0.089	0.127	0.044	0.041	0.073	0.007	0.007	0.022