Research Article

Comparison of Some Tests of Fit for the Inverse Gaussian Distribution

D. J. Best,[1] J. C. W. Rayner,[1, 2] and O. Thas[2, 3]

[1] *School of Mathematical and Physical Sciences, University of Newcastle, NSW 2308, Australia*
[2] *Centre for Statistical and Survey Methodology, School of Mathematics and Applied Statistics,
   University of Wollongong, NSW 2522, Australia*
[3] *Department of Mathematical Modelling, Statistics and Bioinformatics, 9000 Gent, Belgium*

Correspondence should be addressed to J. C. W. Rayner, john.rayner@newcastle.edu.au

This paper gives an empirical investigation of some tests of goodness of fit for the inverse Gaussian distribution.

1. Introduction

The inverse Gaussian (IG) distribution is an important statistical model for the analysis of positive data. See, for example, Seshadri [1]. In its standard form the distribution, denoted $IG(\lambda, \mu)$, depends on the shape parameter $\lambda > 0$ and the mean $\mu > 0$. Its probability density function is

$$\frac{\sqrt{\lambda}}{\sqrt{2\pi x^3}} \exp\left\{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right\}, \quad \text{for } x > 0, \text{ and zero otherwise.} \tag{1.1}$$

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent observations. We wish to test H_0: X is distributed as $IG(\lambda, \mu)$ for $\lambda > 0$ and $\mu > 0$ against H_A and not H_0.

The maximum likelihood estimators are

$$\widehat{\lambda} = \left\{\frac{\sum_{i=1}^{n}\left(1/X_i - 1/\overline{X}\right)}{n}\right\}^{-1}, \quad \widehat{\mu} = \overline{X}. \tag{1.2}$$

Put $\widehat{\varphi} = \widehat{\lambda}/\widehat{\mu}$.

We consider the following test of fit statistics:

   (i) the smooth test statistic of Ducharme [2], denoted by $\widehat{R}_3$,

   (ii) the Laplace transform test statistic of Henze and Klar [3], denoted by V_0,

   (iii) the empirical likelihood ratio test statistic of Vexler et al. [4], denoted by TK_n,

   (iv) the traditional Anderson-Darling test statistic, denoted by A^2,

   (v) conventional smooth tests following Rayner et al. [5].

We strongly suggest that P values for all tests be found using the parametric bootstrap. A link for smooth test statistics for many distributions and bootstrap P values for their tests of fit is http://www.biomath.ugent.be/~othas/smooth2.

2. Definitions

2.1. Ducharme's $\widehat{R}_3$

The usual approach to constructing a smooth test as outlined, for example, in Rayner et al. [5], produces inconsistent and less powerful tests; see Ducharme [2, page 279] and Henze and Klar [3, Tables 1 and 2]. Ducharme [2] suggests testing H_0 above by first defining $Z_i = \mu/X_i$, $i = 1, 2, \ldots, n$; he then states that Z has a random walk distribution, $RW(\varphi)$ say. The Ducharme [2] statistic is $\widehat{R}_3 = \widehat{V}_2^2 + \widehat{V}_3^2$, where $\widehat{V}_2$ and $\widehat{V}_3$ are, under H_0, asymptotically independent and asymptotically standard normal variables, and hence $\widehat{R}_3$ is asymptotically χ_2^2 distributed. Thus $\widehat{R}_3$ does not depend on φ, as do V_0 (see Section 2.2) and the Anderson-Darling test statistic A^2. Ducharme [2, page 279] notes this implies A^2 loses power for some values φ can take. The components $\widehat{V}_2^2$ and $\widehat{V}_3^2$ are defined by

$$\widehat{V}_2^2 = \frac{n\widehat{\varphi}^4}{24 + 6\widehat{\varphi}} \left\{ \overline{\widehat{Z}}^2 - \left(1 + \frac{3}{\widehat{\varphi}} + \frac{3}{\widehat{\varphi}^2}\right) \right\}^2,$$

$$\widehat{V}_3^2 = \frac{n\widehat{\varphi}^6}{24(4 + \widehat{\varphi}) + (120 + 75\widehat{\varphi} + 15\widehat{\varphi}^2 + \widehat{\varphi}^3)}$$

$$\times \left\{ \overline{\widehat{Z}}^3 (4 + \widehat{\varphi}) - \overline{\widehat{Z}}^2 \left(\frac{60}{\widehat{\varphi}} + 30 + 4\varphi\right) + \frac{120}{\widehat{\varphi}^3} + \frac{195}{\widehat{\varphi}^2} + \frac{123}{\widehat{\varphi}} + 32 + 3\widehat{\varphi}^3 \right\}^2$$

(2.1)

in which $\overline{\widehat{Z}}^r = \sum_{i=1}^{n} (\widehat{\mu}/X_i)^r / n$.

A positive feature of smooth tests is that their components can often shed light on how the data differ from the hypothesised distribution. This is somewhat less evident with Ducharme's test given that he transforms the data. Another positive feature of smooth tests is that their components often give highly focused tests with good power. Ducharme's test has components that are likely to fulfil this role.

2.2. Henze and Klar's V_0

This is defined using the exponentially scaled complementary error function $\mathrm{erfce}(x) = e^{x^2}\mathrm{erfc}(x)$ where $\mathrm{erfc}(x) = 2\int_x^\infty e^{-t^2} dt / \sqrt{\pi}$. Note that in $\mathrm{erfc}(x)$ we divide by $\sqrt{\pi}$ and not

π as in Henze and Klar [3, page 428], as we believe a typographical error was made in their paper. Now let $Z_j = X_j/\widehat{\mu}$ and $Z_{jk} = Z_j + Z_k$. Then

$$V_0 = \frac{1}{n} \sum_{j,k=1}^{n} Z_{jk}^{-1} - 2 \sum_{j=1}^{n} Z_j^{-1} \left\{ 1 + \sqrt{\frac{\pi\widehat{\varphi}}{2Z_j}} \text{erfce}\left(\frac{(Z_j+1)\sqrt{\widehat{\varphi}}}{\sqrt{2Z_j}} \right) \right\} + \frac{n(1+2\widehat{\varphi})}{4\widehat{\varphi}}. \quad (2.2)$$

Tests based on the empirical Laplace transform, as is V_0, have produced powerful tests for other distributions, and so it is useful to compare V_0 with other recently suggested tests.

2.3. The $\log(TK_n)$ Statistic of Vexler et al.

Order the observations so that $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$, and let $Y_{(i)} = 1/\sqrt{X_{(n-i+1)}}$. Then

$$\log(TK_n) = \min_{1 \leq m \leq n^\delta} \left[\frac{n}{2} \log\left(\frac{\pi e}{2\widehat{\lambda}}\right) + \sum_{i=1}^{n} \log\left\{ \frac{2m}{n(Y_{(i+m)} - Y_{(i-m)})} \right\} \right], \quad (2.3)$$

where δ can be taken to be 0.5. Observe that, for small m, such as $m = 1$, the statistic can take an infinite value when there are tied data. Vexler et al. [4] do not appear to note this. For the Poisson alternative in Table 1 and $\delta = 0.5$ the $\log(TK_n)$ statistic is often infinite. Choi and Lim [6] show that an entropy-type statistic-like $\log(TK_n)$ has good power for the Laplace distribution and so it is of interest to see how the entropy method works with a skewed distribution.

2.4. The Anderson-Darling Statistic

Again order the data from smallest to largest to obtain $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ and take $Z_{(i)} = F(X_{(i)}; \widehat{\lambda}, \widehat{\mu})$ where $F()$ is the distribution function for the IG distribution. Then the Anderson-Darling statistic is

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^{n} (2i-1)\left\{ \log Z_{(i)} + \log\left(1 - Z_{(n-i+1)}\right) \right\}. \quad (2.4)$$

The Anderson-Darling has stood the test of time as a useful general option for tests of fit for many distributions. Have newer tests improved on its power performance?

2.5. Conventional Smooth Test Third and Fourth Components

Henze and Klar [3] consider the test based on the conventional second-order component $\widehat{U}_2^2$ and show it has poor power for some alternatives. Ducharme [2] notes that these conventional smooth tests discussed, for example, in Rayner et al. [5], can be inconsistent. However we decided to include tests based on $\widehat{U}_3^2$ and $\widehat{U}_4^2$ in our comparisons.

Table 1: Powers (in percentages) of goodness of fit tests for the inverse Gaussian distribution for $n = 20$ and (a) $\alpha = 0.10$, (b) $\alpha = 0.05$.

(a)

Alternative	$\hat{R}_3$	$\hat{V}_2^2$	$\hat{V}_3^2$	$\log(TK_n)$	V_0	A^2	$\hat{U}_3^2$	$\hat{U}_4^2$	$\tilde{U}_3^2$	$\tilde{U}_4^2$
IG(0.25)	10	10	11	9	9	10	10	10	3	6
IG(0.5)	10	10	9	11	10	10	10	10	4	8
IG(1)	10	10	10	11	10	10	10	8	6	9
IG(3)	10	10	10	10	10	10	9	8	8	10
IG(5)	11	11	10	12	10	11	8	9	8	10
IG(10)	9	9	10	10	10	10	8	10	9	11
IG(20)	10	8	11	9	10	9	9	11	9	9
IG(100)	10	9	9	8	9	10	9	10	11	10
W(0.5)	95	92	83	93	96	95	6	17	0	8
W(0.8)	84	79	72	68	83	82	52	61	6	21
W(1)	74	70	61	55	75	73	65	63	10	30
W(1.2)	67	64	55	49	67	63	61	50	18	36
W(1.5)	64	60	53	42	58	55	52	34	29	47
W(2)	58	54	44	37	54	48	43	12	40	54
W(3)	50	47	37	30	46	39	16	15	50	56
LN(0.6)	15	15	13	9	13	14	13	8	9	12
LN(1)	23	23	19	10	22	22	20	12	4	10
LN(1.4)	42	39	31	18	39	36	23	26	2	7
LN(2)	70	66	57	47	70	68	9	20	0	4
LN(3)	94	90	80	87	95	95	0	1	0	1
G(0.6)	90	87	77	83	91	91	41	51	8	21
G(2)	49	46	37	26	47	41	39	20	21	34
χ_1^2	94	89	86	89	95	95	25	36	4	19
χ_{10}^2	28	26	21	15	24	21	8	5	19	25
HN(0, 1)	79	74	65	65	80	78	74	64	36	59
HC(0, 1)	56	56	41	30	55	49	18	25	2	4
U(0, 1)	90	84	77	93	91	91	88	68	94	98
POI(3)	52	39	49	49	50	92	39	1	40	57

(b)

Alternative	$\hat{R}_3$	$\hat{V}_2^2$	$\hat{V}_3^2$	$\log(TK_n)$	V_0	A^2	$\hat{U}_3^2$	$\hat{U}_4^2$	$\tilde{U}_3^2$	$\tilde{U}_4^2$
IG(0.25)	5	5	5	5	4	5	4	5	2	5
IG(0.5)	5	5	6	5	4	5	5	4	2	5
IG(1)	5	6	5	5	4	6	6	4	2	5
IG(3)	6	5	6	4	4	4	6	6	1	3
IG(5)	5	5	5	4	5	6	4	5	1	2
IG(10)	5	5	5	5	4	6	3	4	0	1
IG(20)	5	4	5	5	5	5	2	4	0	1
IG(100)	5	5	5	5	5	6	2	2	0	0
W(0.5)	93	91	83	87	87	94	1	3	0	2
W(0.8)	76	72	60	61	77	77	26	42	1	10
W(1)	67	65	54	47	70	63	45	49	4	14
W(1.2)	62	59	47	38	63	57	47	42	8	21
W(1.5)	54	53	37	31	50	48	45	28	15	32

(b) Continued.

Alternative	$\widehat{R}_3$	$\widehat{V}_2^2$	$\widehat{V}_3^2$	$\log(TK_n)$	V_0	A^2	$\widehat{U}_3^2$	$\widehat{U}_4^2$	$\widetilde{U}_3^2$	$\widetilde{U}_4^2$
W(2)	45	43	31	22	41	42	37	12	23	39
W(3)	41	39	29	20	34	29	13	3	35	47
LN(0.6)	8	8	7	4	5	8	6	2	3	6
LN(1)	16	18	10	5	15	13	9	2	1	4
LN(1.4)	35	34	24	11	27	27	8	2	1	2
LN(2)	64	61	47	38	55	63	2	1	0	1
LN(3)	90	86	76	81	92	94	0	1	0	0
G(0.6)	87	85	75	78	88	88	20	35	1	11
G(2)	37	37	26	18	36	33	28	15	8	21
χ_1^2	89	87	83	85	93	92	7	17	1	8
χ_{10}^2	17	19	10	6	12	13	6	2	10	18
HN(0, 1)	74	70	60	55	70	72	63	55	19	39
HC(0, 1)	51	51	36	23	46	46	7	14	0	2
U(0, 1)	85	81	67	87	83	88	81	61	80	94
POI(3)	37	30	35	39	34	77	8	0	24	43

In general the rth-order component is $U_r = \sum_{j=1}^{n} \pi_r(x_j)/\sqrt{n}$, in which $\pi_r()$ is an orthonormal polynomial of degree r on the inverse Gaussian distribution. We find

$$\pi_3(z) = \frac{(z^3 - a_3 z^2 - b_3 z - c_3)}{\sqrt{(ne_3)}} \tag{2.5}$$

in which $z = x - \mu$, $w = \lambda + 3\mu$, $y = \lambda^3 + 12\lambda^2\mu + 48\lambda\mu^2 + 60\mu^3$, $a_3 = 3\mu^2(3\lambda + 10\mu)/(w\lambda)$, $b_3 = 3\mu^3(\lambda^2 - \lambda\mu + 15\mu^2)/(w\lambda^2)$, $c_3 = -3\mu^5(2\lambda + 7\mu)/(w\lambda^2)$, and $e_3 = 6\mu^9 y/(w\lambda^5)$.

Moreover

$$\pi_4(z) = \frac{(z^4 + a_4 z^3 + b_4 z^2 + c_4 z + d_4)}{\sqrt{(ne_4)}} \tag{2.6}$$

in which z, w and y are as for $\pi_3(z)$, $a_4 = -6\mu^2(3\lambda^3 + 38\lambda^2\mu + 160\lambda\mu^2 + 210\mu^3)/(y\lambda)$, $b_4 = -3\mu^3(3\lambda^5 + 22\lambda^4\mu - 135\lambda^3\mu^2 - 1800\lambda^2\mu^3 - 5940\mu^4\lambda - 6300\mu^5)/(wy\lambda^2)$, $c_4 = 3\mu^5(17\lambda^5 + 253\lambda^4\mu + 1284\lambda^3\mu^2 + 1980\lambda^2\mu^3 - 2280\lambda\mu^4 - 6300\mu^5)/(wy\lambda^3)$, $d_4 = 6\mu^6(\lambda^5 + 10\lambda^4\mu - 6\lambda^3\mu^2 - 390\lambda^2\mu^3 - 1500\lambda\mu^4 - 1710\mu^5)/(wy\lambda^3)$, and $e_4 = 6\mu^{12}(7\lambda^7 + 198\lambda^6\mu + 2439\lambda^5\mu^2 + 16920\lambda^4\mu^3 + 71100\lambda^3\mu^4 + 180180\lambda^2\mu^5 - 253440\lambda\mu^6 + 151200\mu^7)/(wy\lambda^7)$.

The parameters λ and μ can be estimated by maximum likelihood (ML) estimation using the previous formula for $\widehat{\lambda}$ and $\widehat{\mu}$, thereby giving $\widehat{U}_3$ and $\widehat{U}_4$. We also looked at $\widetilde{U}_3$ and $\widetilde{U}_4$ where the parameters are estimated using method of moments (MOM) estimators $\widehat{\lambda} = n\widetilde{\mu}^3 / \sum_{j=1}^{n}(X_j - \widetilde{\mu})^2$ and $\widetilde{\mu} = \widehat{\mu} = \overline{X}$.

As indicated previously, smooth tests can indicate in terms of moments how data and the hypothesised distribution differ. This feature, and good power in previous studies for testing for other distributions, prompted us to include conventional smooth tests in our comparisons.

3. Sizes and Powers

In this section wherever possible we have used IMSL routines to generate random deviates. Calculations were done using double precision arithmetic and FORTRAN code. For the inverse Gaussian random deviates were found as in Michaels et al. [7].

We examine a similar range of alternatives to that given by Henze and Klar [3] so that comparisons can be made with the other statistics in their Table 1(a). We note that

(i) for the lognormal alternative $\mathrm{LN}(\varphi)$ the probability density function should be $\exp\{-\log^2(x)/2\varphi^2\}/(x\varphi\sqrt{2\pi})$,

(ii) the $W(1)$ alternative is a standard exponential alternative. Note that in Henze and Klar [3, Table 1] $W(1)$, $G(1)$, and χ_2^2 are equivalent.

It appears that the tests based on $\widehat{R}_3$ and V_0 generally do well while that based on $\log(TK_n)$ is only competitive for the symmetric uniform alternative. The smooth tests based on $\widehat{U}_3^2$ and $\widehat{U}_4^2$, like that based on $\widehat{U}_2^2$ in Henze and Klar [3], were not competitive. The tests based on $\widetilde{U}_3^2$ and $\widetilde{U}_4^2$ were generally even less competitive. This is unfortunate as these components help describe the data and this facility is not available with the other tests. All powers were calculated using parametric bootstrap.

The alternatives in Table 1 are defined in Henze and Klar [3]. However note that $\mathrm{IG}(\lambda)$ in Henze and Klar [3] is $\mathrm{IG}(\lambda, 1)$ here. There is, however, one exception, and that is the Poisson-type alternative $\mathrm{POI}(3)$ which has probability function $e^{-\theta}\theta^x/x!$ in which $\theta = 3$ here and if a random x value is zero we take this to be $x = 0.5$. This alternative was suggested by the comment in Henze and Klar [3] that for this shelf life data they examine A^2 and the other EDF statistics have a much smaller P value than V_0 and the other empirical Laplace transform statistics. A feature of the shelf life data was that there were tied observations — not something one would expect for an inverse Gaussian distribution. The $\mathrm{POI}(3)$ alternative gives parametric bootstrap simulated samples with many ties and, as can be seen in Table 1, the power of the test based on A^2 is much greater than those for $\widehat{R}_3$ or V_0. The test based on $\log(TK_n)$ classifies infinite values as rejections of the null hypothesis. We have no explanation as to why the Anderson-Darling test is quite powerful for tied data when the null hypothesis specifies an inverse Gaussian distribution.

In Table 1(a) our powers for the tests based on V_0 and A^2 are very similar to those obtained by Henze and Klar [3]. Table 1(b) gives powers for the same alternatives as Table 1(a) but with $\alpha = 0.05$ as this choice of α is commonly used in practice. The relative performance of the tests in Tables 1(a) and 1(b) is similar. The traditional smooth tests, based on the U statistics, are sometimes particularly poor in Table 1(b).

4. The Approach to χ^2

An advantage of the smooth test statistics and their components is that under the null hypothesis they have asymptotic χ^2 distributions. Thus for larger sample sizes approximate P values can be found using the χ^2 distribution. However for $\widehat{V}_2^2$ and $\widehat{R}_3$ Table 2 indicates that, to give actual test sizes close to the nominal 5%, for $\lambda = 2$ and $\mu = 1$ a sample size of 200 might be needed, while an even greater sample size might be needed for $\widehat{V}_3^2$. This ties in with the suggestion, made in Section 1, to use the parametric bootstrap.

We did not expect the conventional smooth test statistics to be asymptotically χ^2 distributed; see, for example, Rayner et al. [5, Section 9.3]. As an illustration of this Table 2

Table 2: Empirical null 95% points for $\widehat{V}_2^2$, $\widehat{V}_3^2$, $\widehat{R}_3$, $\widetilde{U}_3^2$, and $T^2/v(T)$ for $\lambda = 2$ and $\mu = 1$.

n	$\widehat{V}_2^2$	$\widehat{V}_3^2$	$\widehat{R}_3$	$\widetilde{U}_3^2$	$T^2/v(T)$
20	2.20	1.77	3.89	0.97	0.70
50	2.89	1.96	4.64	1.20	0.83
100	3.26	2.24	4.89	1.58	1.08
200	3.51	2.54	5.44	2.12	1.42
500	3.72	3.07	5.62	2.98	1.99
1000	3.75	3.20	5.76	3.97	2.62
2000	3.82	3.34	5.82	4.34	2.89
5000	3.82	3.54	5.84	5.39	3.58
∞	3.84	3.84	5.99	—	3.84

shows that 95% points of $\widetilde{U}_3^2$ do not converge to 3.84. If $m_3 = \sum_i (X_i - \overline{X})^3/n$, we observe that because of the MOM estimators used in $\widetilde{U}_3^2$ we can write the numerator of $\widetilde{U}_3$ as $\sqrt{n}(m_3 - 3\widetilde{\mu}^5/\widetilde{\lambda}^2) = T$ say. Then $T^2/v(T)$, where $v(T)$ is the variance of T, should be asymptotically χ^2. Here $v(T) = 6\mu^9(\lambda^2 + 12\lambda\mu + 36\mu^2)/(n\lambda^5)$. This formula can be found by the delta method. No powers for $T^2/v(T)$ are shown in Table 1 as they are similar to those for $\widetilde{U}_3^2$.

5. Examples

(i) Failure Times

Proschan [8] has given failure times for air conditioning in Boeing 720 jets. For jet number 7912 the 30 times were

$$23, 261, 87, 7, 120, 14, 62, 47, 225, 71, 246, 21, 42, 20, 5,$$
$$12, 120, 11, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52, 95. \tag{5.1}$$

Does the inverse Gaussian provide a good model for these times? We find $\widehat{\lambda} = 13.76$, $\widehat{\mu} = \widetilde{\mu} = 59.60$, and $\widetilde{\lambda} = 42.38$. With $n^\delta = 5$, $\log(TK_n) = 6.936$ with P value 0.152. Further, $\widehat{R}_3 = 22.07$ with P value 0.0007 and $V_0 = 1.158$ with P value 0.012. If $\alpha = 0.05$, say, it makes a difference which test is used: those based on $\widehat{R}_3$ and V_0 are significant but that based on $\log(TK_n)$ is not. If $\alpha = 0.01$ only the test based on $\widehat{R}_3$ is significant. Further results with P values in parentheses are $\widehat{V}_2^2 = 19.42$ (0.001), $\widehat{V}_3^2 = 2.65$ (0.02), $\widetilde{U}_3^2 = 0.32$ (0.27), $\widetilde{U}_4^2 = 1.02$ (0.002), $\widehat{U}_3^2 = 0.77$ (0.06), and $\widehat{U}_4^2 = 0.71$ (0.06). We see the tests based on $\widehat{V}_2^2$ and $\widetilde{U}_4^2$ are significant at the 5% level; the latter suggests the lack of fit is due to kurtosis differences between the model and the data. See Figure 1. Observe that in Figures 1 and 2 the height of the histogram bars is class frequency/number of observations/class width and that this height is labelled "density" so as to be on the same scale as the probability "density" curve. Figure 1 uses MOM estimators for this curve.

Aside from that, we note that in Henze and Klar [3, Table 3] the value 3.7 should be 3.0. This does not affect the conclusions of Henze and Klar for this data set.
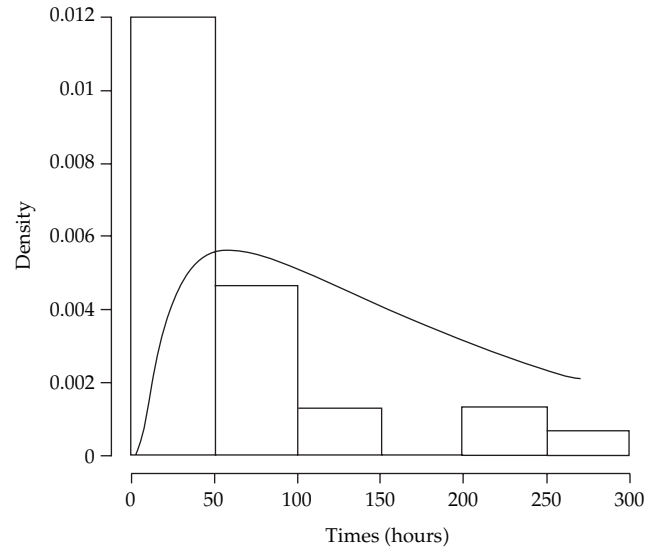
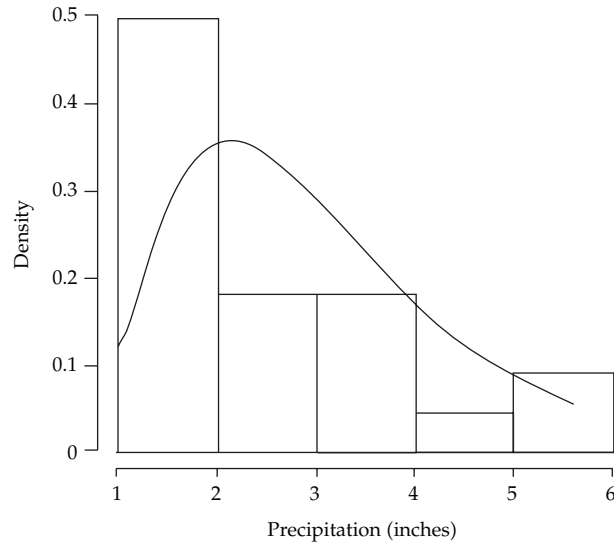**Figure 1:** Air conditioner failure times in Boeing jets.

Figure 2: Storm precipitation at Jug bridge, MD, USA.

(ii) Precipitation at Jug Bridge

Ang and Tang [9, page 266] consider precipitation from storms in inches at the Jug bridge in MD, USA. Their data were

$$1.01.1.11, 1.13, 1.15, 1.16, 1.17, 1.17, 1.2, 1.52, 1.54, 1.54, 1.57, 1.64,$$

$$1.73, 1.79, 2.09, 2.09, 2.57, 2.75, 2.93, 3.19, 3.54, 3.57, 5.11, 5.62.$$

(5.2)

Figure 2 indicates the inverse Gaussian fit is marginal. We find $\hat{\lambda} = 8.08$, $\hat{\mu} = \tilde{\mu} = 2.16$, and $\tilde{\lambda} = 6.72$ with $V_0 = 0.33$, giving a P value of 0.09. Further test statistics with P values in parentheses are $\hat{R}_3 = 2.86$ (0.09), $\log(TK_n) = 9.60$ (0.01), $\hat{V}_2^2 = 1.90$ (0.08), $\hat{V}_3^2 = 0.97$ (0.16), $\tilde{U}_3^2 = 0.05$ (0.78), $\tilde{U}_4^2 = 0.01$ (0.91), $\hat{U}_3^2 = 0.36$ (0.07), and $\hat{U}_4^2 = 0.09$ (0.72). As the test based on $\log(TK_n)$ is most critical of the IG hypothesis, the data may be more symmetric than the IG. The inverse Gaussian curve in Figure 2 uses ML estimators.

In passing we note that the exponential distribution with parameter 0.463 does not provide a good fit to the data. When testing for an exponential distribution $A^2 = 3.226$ with an approximate P value of 0.03. Visual inspection of Figure 2 may have indicated the exponential model may have been appropriate.

6. Conclusion

The tests based on V_0 and $\hat{R}_3$ do well in the power comparisons while that based on $\tilde{U}_4^2$ indicates possible kurtosis differences from the IG distribution in the failure time example. For the precipitation data the test based on $\log(TK_n)$ is most critical of the fit of the IG model. In fact apart from the tests based on $\hat{U}_3^2$ and $\hat{U}_4^2$, all of the tests studied here had something to recommend them: reasonable power or interpretability. No test was uniformly superior to the others.

Acknowledgment

References

[1] V. Seshadri, *The Inverse Gaussian Distribution: Statistical Theory & Applications*, Springer, New York, NY, USA, 1998.

[2] G. R. Ducharme, "Goodness-of-fit tests for the inverse Gaussian and related distributions," *Test*, vol. 10, no. 2, pp. 271–290, 2001.

[3] N. Henze and B. Klar, "Goodness-of-fit tests for the inverse Gaussian distribution based on the empirical Laplace transform," *Annals of the Institute of Statistical Mathematics*, vol. 54, no. 2, pp. 425–444, 2002.

[4] A. Vexler, G. Shan, S. Kim, W. M. Tsai, L. Tian, and A. D. Hutson, "An empirical likelihood ratio based goodness-of-fit test for Inverse Gaussian distributions," *Journal of Statistical Planning and Inference*, vol. 141, no. 6, pp. 2128–2140, 2011.

[5] J. C. W. Rayner, O. Thas, and D. J. Best, *Smooth Tests of Goodness of Fit: Using R*, Wiley, Singapore, 2nd edition, 2011.

[6] B. Choi and K. Kim, "Testing goodness-of-fit for Laplace distribution based on maximum entropy," *Statistics*, vol. 40, no. 6, pp. 517–531, 2006.

[7] J. Michaels, W. Schucany, and R. Haas, "Generating random deviates using transformations with multiple roots," *American Statistician*, vol. 30, no. 2, pp. 88–90, 1976.

[8] F. Proschan, "Theoretical explanation of observed decreasing failure rate," *Technometrics*, vol. 5, no. 3, pp. 375–383, 1963.

[9] A. Ang and W. H. Tang, *Probability Concepts in Engineering*, Wiley, New York, NY, USA, 2nd edition, 2007.