IJMMS 2003:9, 587–592
PII. S0161171203207237
http://ijmms.hindawi.com

A REGRESSION CHARACTERIZATION OF INVERSE GAUSSIAN DISTRIBUTIONS AND APPLICATION TO EDF GOODNESS-OF-FIT TESTS

KHOAN T. DINH, NHU T. NGUYEN, and TRUC T. NGUYEN

Received 15 July 2002

We give a new characterization of inverse Gaussian distributions using the regression of a suitable statistic based on a given random sample. A corollary of this result is a characterization of inverse Gaussian distribution based on a conditional joint density function of the sample. Application of this corollary as a transformation in the procedure to construct EDF (empirical distribution function) goodness-of-fit tests for inverse Gaussian distributions is also studied.

2000 Mathematics Subject Classification: 62E10, 62F03.

1. Introduction. A distribution is an inverse Gaussian distribution with parameters $m > 0$ and $\lambda > 0$, denoted $\mathrm{IG}(m,\lambda)$, if it has a density function given by

$$f(x;m,\lambda) = \begin{cases} \left(\dfrac{\lambda}{2\pi x^3}\right)^{1/2} \exp\left\{-\dfrac{\lambda(x-m)^2}{2m^2x}\right\} & \text{for } x > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

(See Tweedie [11].)

The characteristic function of an $\mathrm{IG}(m,\lambda)$ distribution is

$$\varphi(t) = \exp\left\{\frac{\lambda\left[1-(1-2im^2t\lambda^{-1})^{1/2}\right]}{m}\right\}. \tag{1.2}$$

Let X_j, $j = 1,\dots,n$, $n \geq 2$, be a random sample from an $\mathrm{IG}(m,\lambda)$ distribution. Then, the statistics $Y = \sum_{j=1}^n X_j$ and $Z = \sum_{j=1}^n X_j^{-1} - n^2Y^{-1}$ are jointly complete sufficient for m and λ. Y and Z are independently distributed, Y has an $\mathrm{IG}(nm,n^2\lambda)$ distribution, and λZ has a chi-square distribution with $(n-1)$ degrees of freedom. Khatri [4] gave a characterization of the inverse Gaussian distributions based on the independence between Y and Z, then Seshadri [9] gave several characterizations of the inverse Gaussian distributions based on the constant regression of several different statistics given Y. In this note, we give a characterization of the inverse Gaussian distributions based on the regression of a statistic given Y and Z. The corollary of this result is a characterization of the inverse Gaussian distributions based on the conditional joint density function of $X_1,\dots,X_{n-2}$, given Y and Z. The result of this corollary can be used as a transformation in the procedure to construct EDF (empirical distribution function) goodness-of-fit tests for inverse Gaussian distributions.

2. Characterization results. The conditional joint density function of $X_1,\dots,X_{n-2}$, given $Y = y > 0$, $Z = z > 0$, is

$$f_{X_1,\dots,X_{n-2}|Y,Z}(x_1,\dots,x_{n-2}|y,z) = \begin{cases} \dfrac{2\Gamma\left((n-1)/2\right)y^{3/2}}{n\pi^{(n-1)/2}\prod_{j=1}^{n-2}x_j^{3/2}\left(y-\sum_{j=1}^{n-2}x_j\right)^{1/2}z^{(n-1)/2}} \\ \times\left[\left(y-\displaystyle\sum_{j=1}^{n-2}x_j\right)^2\left(z+n^2y^{-1}-\displaystyle\sum_{j=1}^{n-2}x_j^{-1}\right)-4\left(y-\displaystyle\sum_{j=1}^{n-2}x_j\right)\right]^{-1/2} \\ \qquad\text{for } \displaystyle\sum_{j=1}^{n-2}x_j<y,\ \left(y-\displaystyle\sum_{j=1}^{n-2}x_j\right)\left(z+n^2y^{-1}-\displaystyle\sum_{j=1}^{n-2}x_j^{-1}\right)<4, \\ 0 \qquad\text{otherwise.} \end{cases} \tag{2.1}$$

From (2.1), the UMVUE of the density function at a point $x_1 > 0$ is given by

$$f_{X_1|Y,Z}(x_1|y,z) = \begin{cases} \dfrac{(n-1)\Gamma\left((n-1)/2\right)}{n\sqrt{\pi}\,\Gamma\left((n-2)/2\right)} \\ \times\dfrac{y^{3/2}\left[z+n^2y^{-1}-x_1^{-1}-(n-1)^2(y-x_1)^{-1}\right]^{(n-2)/2-1}}{x_1^{3/2}(y-x_1)^{3/2}z^{(n-1)/2}} \\ \qquad\text{for } x_1<y,\ z+n^2y^{-1}-x_1^{-1}-(n-1)^2(y-x_1)^{-1}>0, \\ 0 \qquad\text{otherwise,} \end{cases} \tag{2.2}$$

where $y,z > 0$. (See Chhikara and Folks [1].)

Let $T = X_1[Z+n^2Y^{-1}-X_1^{-1}-(n-1)^2(Y-X_1)^{-1}]$. $E[T|Y,Z]$ can be computed in two different ways. On the one hand,

$$\begin{aligned} E[T|Y,Z] &= E\Big[X_1\Big[Z+n^2Y^{-1}-X_1^{-1}-(n-1)^2(Y-X_1)^{-1}|Y,Z\Big]\Big] \\ &= \frac{YZ}{n}+(n-1)-(n-1)^2E\Big[X_1(Y-X_1)^{-1}|Y,Z\Big] \\ &= \frac{YZ}{n}+n(n-1)-(n-1)^2YE\Big[(Y-X_1)^{-1}|Y,Z\Big]. \end{aligned} \tag{2.3}$$

On the other hand, this expectation can be computed using the conditional density function of X_1 given by (2.2), and the following integral is taken on the support of this conditional density function:

$$E[T|Y,Z] = \int t(x)f_{X_1|Y,Z(x)}\,dx = \int u\,dv, \tag{2.4}$$

where

$$\begin{aligned} u(x) &= \frac{n-1}{n}\times\frac{\Gamma((n-1)/2)y^{3/2}[z+n^2y^{-1}-x^{-1}-(n-1)^2(y-x)^{-1}]^{(n-2)/2}}{\sqrt{\pi}\,\Gamma((n-2)/2)z^{(n-1)/2-1}}, \\ dv &= \frac{dx}{x^{1/2}(y-x)^{3/2}}. \end{aligned} \tag{2.5}$$

Hence,

$$v = \frac{2x^{1/2}}{y(y-x)^{1/2}}. \tag{2.6}$$

Using integration by parts,

$$\begin{aligned} E[T|Y,Z] &= -\left[\frac{n-2}{Y}\right]E[(Y-X_1)-(n-1)^2X_1^2(Y-X_1)^{-1}|Y,Z] \\ &= -n(n-1)(n-2)+(n-1)^2(n-2)YE[(Y-X_1)^{-1}|Y,Z]. \end{aligned} \tag{2.7}$$

Comparing (2.3) and (2.7),

$$E[(Y-X_1)^{-1}|Y,Z] = \frac{nY^{-1}}{n-1}+\frac{Z}{n(n-1)^3}. \tag{2.8}$$

In the following part, we construct a characterization of inverse Gaussian distributions based on regression (2.8).

If X has an inverse Gaussian distribution with the characteristic function $\varphi(t)$ given by (1.2), then take logarithm of this characteristic function following three successive differentiations and several simplifications, then $\varphi(t)$ satisfies the differential equation

$$\varphi'^4(t)-3\varphi(t)\varphi'^2(t)\varphi''(t)-\varphi^2(t)\varphi'(t)\varphi'''(t)+3\varphi^2(t)\varphi''^2(t) = 0. \tag{2.9}$$

Conversely, if $\varphi(t)$ is the characteristic function of a random variable X with finite $E[X^{-1}]$ and $E[X^3]$, that is, a solution of the differential equation (2.9), then, by the continuity of a characteristic function using the reverse procedure for getting (2.9), this characteristic function is (1.2). Hence, the following result is obtained.

Lemma 2.1. *Let X be a nonnegative random variable with a nondegenerate distribution F and with finite $E[X^{-1}]$ and $E[X^3]$. Assume that $E[X]=m$ and $\operatorname{Var}(X)=m^3/\lambda$ for some positive numbers m and λ, then F is an $\mathrm{IG}(m,\lambda)$ if and only if its characteristic function is a solution of the differential equation (2.9).*

The following theorem is the main result of this note.

Theorem 2.2. *Let X_j, $j=1,\dots,n$, $n\ge 2$, be a random sample of n nonnegative random variables from a nondegenerate distribution F with finite $E[X]$ and $\operatorname{Var}(X)$. Then, F is an inverse Gaussian distribution if and only if regression (2.8) holds.*

Proof. We only need to show that if (2.8) holds, then F is an inverse Gaussian distribution.

From (2.8),

$$E\left[e^{itY}\left\{n(n-1)^3(Y-X_1)^{-1}-n^3(n-2)Y^{-1}-\sum_{j=1}^{n}X_j^{-1}\right\}\right]=0. \tag{2.10}$$

From the fact that X is a random variable with finite $E[X^{-1}]$,

$$E\left[X^{-1}e^{itX}\right]=i\int_{-T}^{t}\varphi(u)\,du+\int_R x^{-1}e^{-iTx}\,dF(x), \tag{2.11}$$

for any constant T such that $-T<t$, where φ is the characteristic function of X (Khatri [4]), then

$$\begin{aligned} I_1(t)&=E\left[e^{it(Y-X_1)}(Y-X_1)^{-1}\right]\\ &=i\int_{-T}^{t}\varphi^{n-1}(u)\,du+\int_R x^{-1}e^{-iTx}\,dF^{*(n-1)}(x),\\ I_2(t)&=E\left[e^{itY}Y^{-1}\right]=i\int_{-T}^{t}\varphi^{n}(u)\,du+\int_R x^{-1}e^{-iTx}\,dF^{*(n)}(x),\\ I_3(t)&=E\left[e^{itX}jX_j^{-1}\right]=i\int_{-T}^{t}\varphi(u)\,du+\int_R x^{-1}e^{-iTx}\,dF(x), \end{aligned} \tag{2.12}$$

where $F^{*(k)}$ denotes the k times convolution of F. Substitute (2.12) into (2.10), simplify, and differentiate three times, the differential equation (2.9) is obtained. Then by Lemma 2.1, F is an inverse Gaussian distribution. □

The following characterization of inverse Gaussian distributions based on (2.1) or (2.2) can be obtained directly from Theorem 2.2. This result will be used as a transformation in the procedure to construct EDF goodness-of-fit tests for inverse Gaussian distributions. The application of this result is discussed in Section 3.

Corollary 2.3. *Let X_j, $j = 1,\dots,n$, $n \geq 2$, be a random sample of nonnegative random variables from a nondegenerate distribution F with finite $E[X^{-1}]$ and $E[X^3]$. F is an inverse Gaussian distribution if and only if the conditional joint density function of $X_1,\dots,X_{n-2}$, given $Y = y > 0$ and $Z = z > 0$, is (2.1), or the conditional density function of X_1, given $Y = y > 0$ and $Z = z > 0$, is (2.2).*

3. Application to goodness-of-fit test. Let X_j, $j = 1,\dots,n$, $n \geq 2$, be a sample of nonnegative random variables from a nondegenerate distribution F with finite $E[X^{-1}]$ and $E[X^3]$. To test whether F is an inverse Gaussian distribution, by Corollary 2.3, it is to test the equivalent simple hypothesis that whether the conditional joint density of $X_1,\dots,X_{n-2}$, given $Y = y > 0$ and $Z = z > 0$, is (1.1). The results of Rosenblatt [8] and then of Chhikara and Folks [1] are used to change the X's sample to the U's random sample from a distribution over the interval (0,1), and the equivalent hypothesis now is whether the U's sample is from the uniform distribution over the interval (0,1). Then, any EDF test statistics can be used (D'Agostino and Stephens [2]). Nguyen and Dinh [5] used this transformation and studied the first exact EDF goodness-of-fit tests for inverse Gaussian distributions. In their study, at some alternative distributions, and with reasonable, not large, sample sizes, the exact EDF goodness-of-fit tests based on this transformation behave pretty well comparing with the other approximate EDF goodness-of-fit tests. The other goodness-of-fit tests for inverse Gaussian distributions using EDF statistics were given by Edgeman et al. [3], O'Reilly and Rueda [6], and Pavur et al. [7]. For detailed references, see Seshadri [10].

References

[1] R. S. Chhikara and J. L. Folks, *Estimation of the inverse Gaussian distribution function*, J. Amer. Statist. Assoc. **69** (1974), 250–254.

[2] R. B. D'Agostino and M. A. Stephens (eds.), *Goodness-of-Fit Techniques*, Statistics: Textbooks and Monographs, vol. 68, Marcel Dekker, New York, 1986.

[3] R. L. Edgeman, R. C. Scott, and R. J. Pavur, *A modified Kolmogorov-Smirnov test for the inverse Gaussian density with unknown parameters*, Comm. Statist. Simulation Comput. **17** (1988), no. 4, 1203–1212.

[4] C. G. Khatri, *A characterization of the inverse Gaussian distribution*, Ann. Math. Statist. **33** (1962), 800–803.

[5] T. T. Nguyen and K. T. Dinh, *Exact EDF goodness-of-fit tests for inverse Gaussian distributions*, to appear in Comm. Statist.—Simulation Comput.

[6] F. J. O'Reilly and R. Rueda, *Goodness of fit for the inverse Gaussian distribution*, Canad. J. Statist. **20** (1992), no. 4, 387–397.

[7] R. J. Pavur, R. L. Edgeman, and R. C. Scott, *Quadratic statistics for the goodness-of-fit test of the inverse Gaussian distribution*, IEEE Trans. Reliab. **41** (1992), no. 1, 118–123.
[8] M. Rosenblatt, *Remarks on a multivariate transformation*, Ann. Math. Statistics **23** (1952), 470–472.
[9] V. Seshadri, *The inverse Gaussian distribution: some properties and characterizations*, Canad. J. Statist. **11** (1983), no. 2, 131–136.
[10] ______, *The Inverse Gaussian Distribution. Statistical Theory and Applications*, Lecture Notes in Statistics, vol. 137, Springer-Verlag, New York, 1999.
[11] M. C. K. Tweedie, *Statistical properties of inverse Gaussian distributions. I*, Ann. Math. Statist. **28** (1957), 362–377.

Khoan T. Dinh: US Environmental Protection Agency (EPA), Washington, DC 20460, USA

Nhu T. Nguyen: Department of Mathematical Sciences, New Mexico State University, Las Cruces, NM 88003-8001, USA

Truc T. Nguyen: Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, OH 43403-0221, USA