Recent works on characterizations of the gamma distribution and related studies*

Chao-Wei Chou[1] and Wen-Jang Huang[2]

[1]Department of Applied Mathematics, National Sun Yat-sen University,
Kaohsiung, Taiwan 804, R.O.C.
e-mail: choucw@math.nsysu.edu.tw

[2]Department of Applied Mathematics, National University of Kaohsiung,
Kaohsiung, Taiwan 811, R.O.C.
e-mail: huangwj@nuk.edu.tw

Dedicated to Professor Min-Te Chao for his retirement

Abstract

Characterization of probabilistic distributions has been an important topic in statistical theory for decades. Although there have been many well known results, some new characterizations of commonly used distributions, such as normal or gamma distributions, are found useful in many applications. In practice, such characteristic properties can be used to represent the observed data. In this paper we restrict our attention to the recent works of characterizations of the gamma distribution and the gamma renewal process, as well as to some related studies on the corresponding parameter estimation based on the characterization properties. Simulation studies are presented to demonstrate how these characterization techniques can be used to help determine the distributions of the observed data sets.

AMS (2000) *subject classification.* Primary 62E05, 62E10; secondary 60G55.

Keywords and phrases. Beta distribution, bivariate gamma distribution, change of measure, characterization, constancy of regression, gamma distribution, generalized inverse Gaussian distribution, inverse Gaussian distribution, Laplace transform, mean squared error, method of moments, parameter estimation, Poisson process, power of hypothesis testing, renewal process.

*Support for this research was provided in part by the National Science Council of the Republic of China, Grant No. NSC 92-2118-M-390-001.

1. Introduction

The gamma distribution plays a crucial role in mathematical statistics and many applied areas. A random variable X is said to have a gamma distribution with three parameters α, β, γ, denoted by $X \sim \Gamma(\alpha, \beta, \gamma)$, if X has the probability density function (pdf)

$$f(x) = \frac{\beta^\alpha (x-\gamma)^{\alpha-1} e^{-\beta(x-\gamma)}}{\Gamma(\alpha)} I_{(\gamma,\infty)}(x),$$

where $\alpha > 0, \beta > 0, \gamma$ are the shape, scale and location parameters, respectively, and $\Gamma(\cdot)$ is the gamma function defined by

$$\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx, t > 0.$$

In the classical system of densities introduced by K. Pearson (1894), the gamma density is characterized as Type III. In most cases, the two-parameter gamma distribution with $\gamma = 0$ is considered. We denote it by $X \sim \Gamma(\alpha, \beta)$ and it has the pdf

$$f(x) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} I_{(0,\infty)}(x). \tag{1}$$

In this work we consider only the class of two-parameter gamma distributions in which some well-known distributions are included. When the gamma distribution has an integral parameter α, it is called the Erlang distribution. In particular, if $\alpha = 1$, we have an exponential distribution. The exponential distribution can be characterized by a constant failure rate. That is, the reliability for a given operation interval is the same, no matter from what portion of the useful life of a device the interval is taken. Another important distribution in this class is the chi-squared distribution with degrees of freedom 2α, $\chi^2_{2\alpha}$, when $\beta = 1/2$ and $\alpha > 0$.

Followings are some of the interesting and important properties of the gamma distribution, $X \sim \Gamma(\alpha, \beta)$:
1.The variance of a random sample with size n from a normal population $N(\mu, \sigma^2)$ has a $\Gamma((n-1)/2, (n-1)/(2\sigma^2))$ distribution.
2.The Laplace transform of X is $E(\exp(sX)) = (1 - s/\beta)^{-\alpha}, s \leq 0$.
3.Special transformation $2\beta X \sim \chi^2_{2\alpha}$.
4.The r-th moment of X is $E(X^r) = \Gamma(\alpha + r)/(\Gamma(\alpha)\beta^r), r > -\alpha$.
5.Reproductive property and infinitely divisible property.
These properties provide very useful theoretical tools while studying gamma distributions or using it in real application.

The gamma distribution has been used in the areas such as engineering and business. The applications include queuing systems, reliability assessment, inventory control, computer evaluations, and biological studies in which the occurrence of an event depends on a series of independent sub-events whose occurrence times are independent and identically distributed (iid) exponential random variables. In spite of many uses of the gamma distribution, there have been very few distributional assessment procedures developed. The gamma distribution is one of the most frequently used distributions to model lifetime data. This is due to its flexibility in the choice of the shape and scale parameters. It is also commonly used in the waiting time problems. For example, the waiting time for the k-th occurrence of a Poisson process follows a gamma distribution. In reliability studies and in life testing, the gamma distribution is used as a generalization of the exponential distribution which is also a popular choice for the modeling purpose. Over the last few decades, the gamma distribution has become one of the most important techniques for modeling life-testing situations.

On the other hand, reliability has been an important topic in industry. The gamma distribution has been suggested as the failure time model for a system under continuous maintenance, where the reliability may experience some initial growth or decay but eventually reaches a steady state. The gamma family provides an applicable class of distributions for testing reliability and performing survival analysis. The shape parameter is especially interesting since whether $\alpha - 1$ is negative, zero, or positive corresponds to a decreasing failure rate (DFR), constant failure rate, or increasing failure rate (IFR), respectively.

Scientists have observed that, under repeated observation or sampling, the sample coefficient of variation approaches a deterministic constant. If the underlying model behaves like a gamma distribution, then having the knowledge of the coefficient of variation implies having the knowledge of α.

The situation that the shape parameter is known happens when the number of sub-events needed to activate the occurrence is known. For example, consider a computer buffer which stores six messages before transmitting them to the processor. The waiting time for a transmission then has a gamma distribution with a shape parameter of six if the inter-arrival times of the messages are iid exponential random variables. The case that the shape parameter is unknown sometimes happens. For example, it is usually assumed that the lifetime of a component that fails after k shocks and the time between two shocks are iid exponential random variables, while the value of k is unknown.

In theoretical calculations, the gamma distribution arises as the sum of iid exponential random variables. It can be used for testing the equality of variances for several independent normal distributions. The nice property of the reproductivity and infinitely divisibility of the gamma distribution also makes it much easier to deal with. Johnson and Kotz (1994) provides a good review of the gamma distribution in which several applications in various fields are discussed.

In Section 2 characterizations of the gamma distribution are obtained. In Section 3 characterizations of the Poisson process are given. Characterizations of the bivariate gamma distribution, and GIG and beta distributions are given in Sections 4 and 5, respectively. By using our results we present the parameter estimation for gamma population in Section 6. Finally a brief discussion is given in Section 7.

2. Characterizations of the gamma distribution

It is known that if X and Y are independent gamma random variables with the same scale parameter, *i.e.*, $X \sim \Gamma(\alpha_X, \beta), Y \sim \Gamma(\alpha_Y, \beta)$, then the two random variables

$$X + Y \quad and \quad X/(X+Y)$$

are mutually independent and have $\Gamma(\alpha_X + \alpha_Y, \beta)$ and $Be(\alpha_X, \alpha_Y)$ distributions respectively. The notation $Be(p, q)$ denotes the beta distribution having the pdf

$$f(x) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} x^{p-1}(1-x)^{q-1} I_{(0,1)}(x). \tag{2}$$

Lukacs (1955) showed that this property can be used to characterize the gamma distributions in the following sense. If X and Y are independent non-degenerate positive random variables and $X+Y$ and $X/(X+Y)$ are mutually independent, then X and Y must have gamma distributions with a common scale parameter, but possibly with different values of the shape parameter.

By setting $V = X + Y$ and $U = X/(X+Y)$ in Lukacs type characterization, we get another form of characterization using the independence between U and V and independence between UV and $(1-U)V$. Note that $X = UV$, X, V, have gamma distributions and U has a beta distribution in this case. Since UV is part of V, an interesting characterization result for Poisson distributions is derived by Patil and Seshadri (1964): Let $N = X + Y$ where X and Y are two random variables with common support $\{0, 1, 2, \cdots\}$. Given $N = n, n \geq 1$, and assume that X has a binomial distribution $B(n, \theta), 0 < \theta < 1$, then N has a Poisson distribution if and only if X and Y are independent.

Denote the equivalence in distribution by "$\overset{\mathrm{d}}{=}$". Some related characterizations of the gamma distribution were done by Huang and Chen (1989) using

$$\sum_{i=1}^{M} Y_i \overset{\mathrm{d}}{=} U_1 \sum_{i=1}^{K} Y_i \tag{3}$$

or

$$\sum_{i=1}^{M} Y_i \overset{\mathrm{d}}{=} \sum_{i=1}^{K} U_i Y_i,$$

where $K > M, U_i, i = 1, \cdots, K$, are iid from the common distribution $Be(r, 1), r = M/(K-M)$, and $Y_i, i = 1, \cdots, K$, are iid non-negative random variables. Characterization of multivariate random variables was also discussed in that paper. Further, Huang and Chen (1991) proved that under the condition that

$$Z \stackrel{\mathrm{d}}{=} U_1 X, \tag{4}$$

the distribution of Z can uniquely determine the distribution of X. Furthermore, they gave an implicit relationship between the Laplace tranfoms of Z and that of X. Other related works were done by Yeo and Milne (1991), Alzaid and Al-Osh (1991), Pakes (1992a), Pakes and Khattree (1992). Analogously, Pakes (1992b) considered the continuous version of equation (3) by using

$$Z_u \stackrel{\mathrm{d}}{=} U Z_{u+v},$$

where $\{Z_t, t \geq 0\}$ is a Lévy process and U has all values in $[0, 1]$. Pakes (1994) considered $\{Z_t\}$ as a self-similar process and showed that the solution exists only when U is a constant.

Let X, Y and U be random variables where U is independent of Y and has support on $[0, 1]$. As mentioned by Alzaid and Al-Osh (1991), the formula that $X = UY$ is of paramount importance in many fields. For example, in economic modeling, Y may represent the actual income of an individual and X stands for his reported income.

Similar to Lukacs type characterization, preserving independence under some other transformations are also used to characterize distributions. Among others, Letac and Wesolowski (2000) (LW (2000) in the sequel) proved that $1/(X + Y)$ and $1/X - 1/(X + Y)$ are independent if and only if X and Y have a generalized inverse Gaussian (GIG) distribution and a gamma distribution, respectively. In addition, Seshadri and Wesolowski (2003) characterized beta distributions of X and Y by the independence of $(1 - Y)/(1 - XY)$ and $1 - XY$. Both of the above two characterizations will be discussed in Section 5. Note that if

$$(u_1, u_2) = (f_1(x, y), f_2(x, y)),$$

where

$$f_1(x, y) = 1/(x + y), f_2(x, y) = 1/x - 1/(x + y),$$

then we have that

$$f_1(u_1, u_2) = x, f_2(u_1, u_2) = y.$$

Also, if

$$(v_1, v_2) = (g_1(x, y), g_2(x, y)),$$

where

$$g_1(x, y) = (1 - y)/(1 - xy), g_2(x, y) = 1 - xy,$$

then we have that

$$g_1(v_1, v_2) = x, g_2(v_1, v_2) = y.$$

Extensions of the above Lukacs type characterizations are also developed. For example, assume that X_1 and X_2 are identically distributed with finite second moment, it suffices to show that the regression

$$E[\frac{aX_1^2 + bX_1X_2 + cX_2^2}{(X_1 + X_2)^2}|X_1 + X_2], a + c \neq b,$$

is a constant in order to guarantee that the common distribution of X_1 and X_2 is gamma (see Laha (1964)). Note that here the condition of independence is replaced by constancy of regression, which is the tradeoff with the existence of the second moments.

In addition to the results mentioned above, there are many further investigations. We list some of them in the following. (i) Weakening the independence condition to constancy of regressions (see Bolger and Harkness (1965), Hall and Simon (1969), Wesolowski (1990), Li *et al.* (1994), Huang and Su (1997), Bobecka and Wesolowski (2002a), Chou and Huang (2003), Huang and Chou (2004)). (ii) Considering the renewal process (see Wesolowski (1989), Li *et al.* (1994), Huang and Su (1997), Chou and Huang (2003), Huang and Chou (2004)). (iii) Considering the bivariate cases (see Wang (1981), Bobecka (2002), Pusz (2002), Chou and Huang (2004a)). (iv) Considering the matrix variates (see Olkin and Rubin (1962), Casalis and Letac (1996), Letac and Massam (1998), Bobecka and Wesolowski (2002b)). We will discuss (i) in this section, (ii) in Section 3 and (iii) in Section 4.

Given two independent and non-degenerate positive random variables X and Y, Bolger and Harkness (1965), Wesolowski (1990) and Li *et al.* (1994) characterized X and Y to have gamma distributions, where the condition of independence between $X/(X + Y)$ and $X + Y$ was replaced by

$$E(X^u|X + Y) = a(X + Y)^u$$

and

$$E(X^v|X + Y) = b(X + Y)^v,$$

where $(u, v) = (1, 2), (1, -1)$ or $(-1, -2)$. Huang and Su (1997) generalized these results and obtained a similar characterization under the weaker conditions that

$$E(X^{r+1}|X + Y) = a(X + Y)E(X^r|X + Y)$$

and

$$E(X^{r+s+1}|X+Y) = b(X+Y)E(X^{r+s}|X+Y),$$

where $s = 1$, r is a fixed integer and a and b are constants. Note that $(u, v) = (1, 2), (1, -1)$ and $(-1, -2)$ corresponds to $r = 0, -1$ and -2, respectively. Chou and Huang (2003) proved that this statement remains true when $s = 2$.

A technique of change of measure for the traditional Laplace transform methods was used by Huang and Chou (2004) to extend the result to that r needs only to be a fixed real number and $s = 1$ or 2.

On the other hand, Hall and Simons (1969) and Huang and Su (1997) characterized gamma distributions using

$$E(X^u|X+Y) = a(X+Y)^u$$

and

$$E(Y^u|X+Y) = b(X+Y)^u,$$

where $u = 2$ or -1.

It is easy to see that for the case $u = 1$, the above two equations reduce to only one equation and can not be used to characterize gamma distribution.

Huang and Chou (2004) considered the case that one of X and Y is assumed to be gamma distributed. It seems reasonable that we can reduce the number of equations in the above case to obtain similar characterization results. More specifically, given that X or Y has a gamma distribution and the equation that

$$E(X^{r+1}|X+Y) = a(X+Y)E(X^r|X+Y)$$

or

$$E(X^{r+2}|X+Y) = b(X+Y)^2E(X^r|X+Y),$$

where a, b are some constants, they characterized that Y or X also has a gamma distribution with the same scale parameter.

Another result was shown by Bobecka and Wesolowski (2002) who generalized the Lukacs theorem under the so-called dual regression schemes. In their study the constancy of regressions for X and Y as well as independence of $X + Y$ and $X/(X + Y)$ were assumed. More precisely, they characterized X and Y to have gamma distributions by

$$E(Y^u|X) = a \qquad \textit{and} \qquad E(Y^v|X) = b$$

or

$$E(X^u|Y) = a \qquad and \qquad E(X^v|Y) = b,$$

where $(u, v) = (1, 2), (1, -1)$ or $(-1, -2)$. The theorems were proven using the method of moments since the moments can uniquely determine a gamma distribution and a beta distribution from the sufficient condition due to Carleman (see Chung, 2001, p.103).

Chou and Huang (2003) extended the above results by using

$$E(Y^{r+1}|X) = aE(Y^r|X) \qquad and \qquad E(Y^{r+2}|X) = bE(Y^{r+1}|X),$$

or

$$E(X^{r+1}|Y) = aE(X^r|Y) \qquad and \qquad E(X^{r+2}|Y) = bE(X^{r+1}|Y)$$

for some fixed integer r and constants a and b, to characterize that X and Y have gamma distributions. Again, $(u, v) = (1, 2), (1, -1)$ and $(-1, -2)$ corresponds to $r = 0, -1$ and -2, respectively. However, using change of measure, Huang and Chou (2004) proved that r needs only to be a fixed real number.

3. Characterizations of the Poisson process

The process version of Lukacs type characterizations is studied in this section. Let $A \equiv \{A(t), t \geq 0\}$ be a renewal process and $\{S_k, k \geq 1\}$ be the sequence of arrival times. Li *et al.* (1994) characterized A to be a Poisson process by assuming that

$$E(S_k^u|A(t) = n) = at^u$$

and

$$E(S_k^v|A(t) = n) = bt^v$$

for some fixed integers k and $n, 1 \leq k \leq n$, and constants a and b where $(u, v) = (1, 2), (1, -1)$ or $(-1, -2)$. Huang and Su (1997) extended the above result under the weaker conditions that

$$E(S_k^{r+1}|A(t) = n) = atE(S_k^r|A(t) = n)$$

and

$$E(S_k^{r+s+1}|A(t) = n) = btE(S_k^{r+s}|A(t) = n),$$

where $s = 1$, r, k, n are some fixed integers, $1 \leq k \leq n$ and a, b are some constants. The case that $s = 2$ was proved by Chou and Huang (2003). Again, using change of measure, Huang and Chou (2004) gave an extension to that $s = 1$ or 2 and r is a fixed real number.

4. Characterizations of the bivariate gamma distribution

As mentioned before, if X and Y are two independent gamma random variables with the same scale parameter, then $X+Y$ and $X/(X+Y)$ are independent. However, in the bivariate case such a property does not always hold. Note that a positive random vector $\tilde{X} = (X_1, X_2)$ has a bivariate gamma distribution $BG(p, \tilde{\lambda})$ (denote it by $\tilde{X} \sim BG(p, \tilde{\lambda})$) with shape parameter p and scale parameter $\tilde{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$ if it has the following Laplace transform:

$$\begin{aligned} E(\exp(s_1X_1 + s_2X_2)) &= (1 - \lambda_1 s_1 - \lambda_2 s_2 + \lambda_3 s_1 s_2)^{-p}, s_1, s_2 \leq 0 \quad \textit{and} \\ \lambda_1 s_1 + \lambda_2 s_2 - \lambda_3 s_1 s_2 &< 1, \end{aligned}$$

where $p, \lambda_1, \lambda_2 > 0$ and $\lambda_1\lambda_2 \geq \lambda_3 \geq 0$. The case that $\lambda_1\lambda_2 = \lambda_3$ and $\lambda_3 = 0$ corresponds to the condition that X_1, X_2 are independent and $P(X_2 = \frac{\lambda_2}{\lambda_1}X_1) = 1$, respectively. It is easy to see that this class of bivariate gamma distributions has the reproductive and infinitely divisible property.

Let $\tilde{X} = (X_1, X_2)$ and $\tilde{Y} = (Y_1, Y_2)$ be mutually independent and non-degenerate positive random vectors. Bobecka (2002) gave a bivariate version of Lukacs theorem by showing that $(X_1/(X_1+Y_1), X_2/(X_2+Y_2))$ and (X_1+Y_1, X_2+Y_2) are independent if and only if that $\tilde{X} \sim BG(p, \tilde{\lambda}), \tilde{Y} \sim BG(q, \tilde{\lambda})$ with $P(X_2 = \frac{\lambda_2}{\lambda_1}X_1) = P(Y_2 = \frac{\lambda_2}{\lambda_1}Y_1) = 1$, or $\tilde{X}, \tilde{Y}$ have independent gamma components. Furthermore, she proved that the whole class of BG distributions has the property of constancy for regressions, *i.e.* when $\tilde{X} \sim BG(p, \tilde{\lambda})$ and $\tilde{Y} \sim BG(q, \tilde{\lambda})$, there exists some constants c_r such that

$$E((\frac{X_j}{X_j + Y_j})^r | \tilde{X} + \tilde{Y}) = c_r, j = 1, 2, \tag{5}$$

where $r = 1, 2, -1, -2$, and $r > -p$. However, Chou and Huang (2004a) proved that equation (5) holds for every integer $r > -p$. Conversely, for $(u, v) = (1, 2), (1, -1)$ or $(-1, -2)$ with the assumption that

$$E((\frac{X_j}{X_j + Y_j})^u | \tilde{X} + \tilde{Y}) = a_j \tag{6}$$

and

$$E((\frac{X_j}{X_j + Y_j})^v | \tilde{X} + \tilde{Y}) = b_j \tag{7}$$

hold for some constants a_j and $b_j, j = 1, 2$, Bobecka (2002) characterized $\tilde{X}$ and $\tilde{Y}$ to have bivariate gamma distributions with the same scale parameter. This generalizes the results by Bolger and Harkness (1965), Wesolowski (1990) and Li *et al.* (1994) in which univariate cases were considered. Instead of (6) and (7), the following weaker assumption was given by Chou and Huang (2004a): For some fixed integer r,

$$E((\frac{X_j}{X_j + Y_j})^{r+1}|\tilde{X} + \tilde{Y}) \quad = \quad \alpha_j E((\frac{X_j}{X_j + Y_j})^r|\tilde{X} + \tilde{Y})$$

and

$$E((\frac{X_j}{X_j + Y_j})^{r+2}|\tilde{X} + \tilde{Y}) \quad = \quad \beta_j E((\frac{X_j}{X_j + Y_j})^{r+1}|\tilde{X} + \tilde{Y})$$

hold for some constants $\alpha_j, \beta_j, j = 1, 2$. Note that $(u, v) = (1, 2), (1, -1), (-1, -2)$ in (6) and (7) corresponds to that $r = 0, -1, -2$, respectively. This also genaralizes the result by Huang and Su (1997) for the univariate case.

The following two formulas were used by Bobecka (2002) to characterize the bivariate gamma distribution:

$$E((\frac{X_j}{X_j + Y_j})^2|\tilde{X} + \tilde{Y}) \quad = \quad d_j \tag{8}$$

and

$$E((\frac{Y_j}{X_j + Y_j})^2|\tilde{X} + \tilde{Y}) \quad = \quad e_j \tag{9}$$

for some constants d_j and $e_j, j = 1, 2$. Chou and Huang (2004a) also characterized the bivariate gamma distribution using the constancy of the quadratic regressions

$$E(a_j X_j^2 + b_j X_j Y_j + c_j Y_j^2|\tilde{X} + \tilde{Y}) = 0 \tag{10}$$

and

$$E(d_j X_j^2 + e_j X_j Y_j + f_j Y_j^2|\tilde{X} + \tilde{Y}) = 0 \tag{11}$$

for some constants $a_j, b_j, c_j, d_j, e_j, f_j$, where vectors (a_j, b_j, c_j) are linearly independent of $(d_j, e_j, f_j), j = 1, 2$. It can be found interestingly that equations (10) and (11) are actually equivalent to equations (8) and (9).

The case that $\tilde{X}_i = (X_{i1}, X_{i2}), i = 1, 2, \cdots, n$, are iid has also been considered in literature. For example, Pusz (2002) characterized the bivariate gamma distribution using the assumption that

$$E(\sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik} X_{i1} X_{k1} + \sum_{i=1}^{n} b_i X_{i1}|\sum_{i=1}^{n} \tilde{X}_i) \quad = \quad 0 \tag{12}$$

and

$$E(\sum_{i=1}^{n}\sum_{k=1}^{n} c_{ik}X_{i2}X_{k2} + \sum_{i=1}^{n} d_i X_{i2} | \sum_{i=1}^{n} \tilde{X}_i) \quad = \quad 0, \tag{13}$$

where $a_{ik} = c_{ik}, b_i = d_i$ for all $i, k = 1, \cdots, n$. Chou and Huang (2004a) also gave a similar characterization without assuming that $a_{ik} = c_{ik}, b_i = d_i, i, k = 1, \cdots, n$.

5. Characterizations of the GIG and beta distributions

Similar to Lukacs type characterization, there are some other characterizations using the property of preserving independence under the transformation for independent random variables. Among others, LW (2000) proved that given two independent non-degenerate positive random variables X and Y, $1/(X+Y)$ and $1/X - 1/(X+Y)$ are independent if and only if X and Y have a generalized inverse Gaussian (GIG) distribution and a gamma distribution, respectively. Seshadri and Wesolowski (2003) characterized the beta distributions of two independent and non-degenerate random variables X, Y, valued in $(0,1)$, by the independence of $(1-Y)/(1-XY)$ and $1-XY$. The readers may refer to Pusz (1997) and Matsumoto and Yor (2003) for some related works on GIG distributions.

Extensions of the LW (2000) characterization results are also developed. There are two main directions afterward: (i) Weakening the independence condition to constancy of regressions (see Seshadri and Wesolowski (2001) (SW (2001) in the sequel), Wesolowski (2002), Chou and Huang (2004b)). (ii) Considering the matrix variates (see Wesolowski (2002)).

5.1. Introduction to inverse Gaussian distribution

We first introduce a subclass of GIG distributions, namely the inverse Gaussian (IG) distributions. Let $X(t)$ be a Wiener process starting at x_0 with drift $\nu > 0$ and variance σ^2. Let T be the first time the process hits $a, a > x_0$. That is,

$$X(0) = x_0, X(t) < a, 0 < t < T \qquad \textit{and} \qquad X(T) = a.$$

Then T has an $IG(\theta, \lambda)$ distribution with the pdf

$$f(t) = (\frac{\lambda}{2\pi t^3})^{\frac{1}{2}} \exp(-\frac{\lambda(t-\theta)^2}{2\theta^2 t}) I_{(0,\infty)}(t), \tag{14}$$

where $\theta = (a - x_0)/\nu$ and $\lambda = (a - x_0)^2/\sigma^2 > 0$. The parameter θ is the mean of the distribution and λ is a scale parameter. The Laplace transform of T is then given by

$$E(\exp(sT)) = \exp\{\frac{\lambda}{\theta}[1 - (1 - \frac{2\theta^2}{\lambda}s)^{1/2}]\}, s \leq 0.$$

Given that $T < \infty$, the conditional distribution of T with $\nu < 0$ is $IG(-\theta, \lambda)$. On the other hand, the distribution of T with $\nu = 0$ is a reciprocal of a gamma distribution. The name "inverse Gaussian", named by Tweedie (1945), is based on the inverse relationship between the cumulant-generating functions (cgfs) of the time to cover unit distance (IG law) and that of the distance covered in unit time (normal law) in Wiener process. Note that for a random variable T, the cgf of T is defined by $\varphi(s) = logE(\exp(sT))$. Also two random variables X, Y are called inverse random variables and the corresponding distributions are called a pair of inverse distributions if their cgfs $\varphi_X(s)$ and $\varphi_Y(s)$ satisfy the condition that

$$c_1\varphi_X(s) \quad is \quad the \quad inverse \quad function \quad of \quad c_2\varphi_Y(s)$$

for some fixed constants c_1, c_2. Tweedie (1945) showed three pairs of inverse distributions: (i) Normal and IG; (ii) Binomial and negative binomial; (iii) Poisson and gamma. Apart from the inverse-distribution relationship, IG has many other statistical properties similar to that of normal law. We list some of them in the following. Assume $X \sim IG(\theta, \lambda)$, then (a) $cX \sim IG(c\theta, c\lambda)$ for positive constant c. (b) Linear combination of independent IG random variables with suitable coefficients is also IG distributed. (c) The sample mean $\overline{X}$ and $\sum_{i=1}^{n}(X_i^{-1} - (\overline{X})^{-1})$ are independent. (d) $\lambda(X-\theta)^2/(\theta^2 X) \sim \chi_1^2$. (e) $D = \lambda\sum_{i=1}^{n}(X_i^{-1} - (\overline{X})^{-1}) \sim \chi_{n-1}^2$ and when $n = 2^r, r$ is a positive integer, D can be decomposed into the sum of $n-1$ independent chi-squared random variables with one degree of freedom.

IG distributions can be used in a wide range of statistical methods and seem to be suitable for model fitting especially when the data are skewed. For the details, readers can refer to the works by Chhikara and Folks (1989), Seshadri (1999), *etc.*

5.2. Introduction to GIG distribution

Next, a generalized class of the GIG distributions $\mu_{p,a,b}$ is defined by the pdf

$$f(x) = Cx^{p-1}\exp(-ax - b/x)I_{(0,\infty)}(x),$$

where C is an appropriately chosen constant. The family of the GIG distributions can be partitioned into the following three subclasses:

(i) Class I: $a > 0, b > 0, p \in R$.

(ii) Class II: $a > 0, b = 0, p > 0$.

(iii) Class III: $a = 0, b > 0, p < 0$.

Note that if X is $\mu_{p,a,b}$ distributed, then X^{-1} is $\mu_{-p,b,a}$ distributed. Class I contains the IG (with $p = -1/2$), reciprocal inverse Gaussian (RIG) ($p = 1/2$), hyperbolic

($p = 1$) and hyperbola distributions ($p = 0$). Class II is the class of gamma distributions. Class III is the class of reciprocal gamma distributions. A random variable X is RIG (or reciprocal gamma) distributed, if and only if X^{-1} is IG (or gamma) distributed.

LW (2000) considered only Class II and a subclass of Class I, that is, the gamma distribution $\Gamma(q, c)$ (*i.e.* $\mu_{q,c,0}$) and the GIG distribution $\mu_{-p,a,b}$ with the pdf

$$f(x) = \frac{(a/b)^{-p/2}}{2K_{-p}(2\sqrt{ab})} x^{-p-1} \exp(-ax - b/x) I_{(0,\infty)}(x), \tag{15}$$

where $p, a, b > 0$ and K_{-p} is a modified Bessel function

$$K_{-p}(z) = \frac{1}{2}(\frac{1}{2}z)^{-p} \int_0^\infty u^{p-1} \exp(-u - \frac{z^2}{4u}) du.$$

Note that the pdf of $\mu_{-p,a,b}$ given in equation (15) is the same as that given by Wesolowski (2002), but somewhat different from that given by SW (2001). In fact, it is the same as the pdf of $\mu_{-p,2a,2b}$ given by SW (2001).

5.3. Characterization results of GIG and gamma distributions

Let X and Y be two independent and non-degenerate positive random variables. LW (2000) proved that $U = 1/(X + Y)$ and $V = 1/X - 1/(X + Y)$ are independent if and only if X has a GIG distribution and Y has a gamma distribution with suitable parameters. Similar characterizations were shown by SW (2001), where the independence of U and V is replaced by the constancy of the regression of V (or V^{-1}) on U under suitable moment conditions and the distribution of Y or X is given. More precisely, given $Y \sim \Gamma(p, a)$, SW (2001) characterized that $X \sim \mu_{-p,a,b}$ under the assumption that

$$E(V^{r+1}|U) = c_r E(V^r|U) \tag{16}$$

for $r = 0, -1, r > -p$, and c_r is some constant. Using change of measure, Chou and Huang (2004b) proved that the above result holds for some fixed real number $r > -p$. On the other hand, SW (2001) characterized that $Y \sim \Gamma(p, a)$ if $X \sim \mu_{-p,a,b}$, and (16) holds for $r = 0$ or $-1, r > -p$, with $c_0 = E(V), c_{-1} = 1/E(V^{-1})$. Again, Chou and Huang (2004b) proved that the result is true for some fixed real number $r > -p$.

Simultaneous characterization of the distributions of X and Y were considered by Wesolowski (2002). He characterized X to have a GIG distribution and Y to have a gamma distribution under the assumption that, for $r = -1$, equation (16) and

$$E(V^{r+2}|U) = c_{r+1} E(V^{r+1}|U), \tag{17}$$

hold for some constants c_r and c_{r+1}. The extension for r to be some fixed real number was given by Chou and Huang (2004b).

5.4. Characterization results of beta distributions

Recently, Seshadri and Wesolowski (2003) gave a characterization of beta distributions. It was proved that two independent random variables X and Y, with all values in $(0, 1)$, can be characterized to have beta distributions by using equations (16) and (17), where $U = 1 - XY, V = (1 - Y)/(1 - XY)$ and $r = 0$ or -1.

We believe that the result may hold for some $r \neq 0, -1$, since there are some similar characterization results for other distributions just mentioned before. However, this needs to be further examined.

6. Parameter estimation for gamma populations by using the characterization results

6.1. Parameter estimation on one population

Let $X_1, X_2, \cdots, X_n, n \geq 3$, be a sequence of iid random variables. Denote the sample mean and sample variance by

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \textit{and} \qquad S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

We say that X has a $\Gamma(\alpha, \beta)$ distribution with shape parameter α and scale parameter β, if the pdf of X is as given in (1). Note that if $X_1, X_2, \cdots, X_n$ are $\Gamma(\alpha, \beta)$ distributed, then

$$E(\bar{X}) = \frac{\alpha}{\beta} \tag{18}$$

and

$$E(S_X^2) = \frac{\alpha}{\beta^2}. \tag{19}$$

For a gamma population, Wang (1981) proved that $\bar{X}$ is independent of the statistic $T(X_1, X_2, \cdots, X_n)$ which is invariant under the scale transformation, *i.e.* $T(cX_1, cX_2, \cdots, cX_n) = T(X_1, X_2, \cdots, X_n) \ \ \forall c \neq 0$. This implies that $\bar{X}$ is independent of $S_X/\bar{X}$. Conversely, Hwang and Hu (1999) characterized the common distribution of X_i's to be gamma by using the independence of $\bar{X}$ and $S_X/\bar{X}$. Later on Hwang and Hu (2000) also gave some extensions. More specifically, they characterized the gamma distribution using the independence of $\bar{X}$ and

$$(S_X/\bar{X})\exp(\psi(\lambda_1, \cdots, \lambda_n)), \tag{20}$$

where ψ is a real-valued function defined on a proper domain, and $\lambda_1, \cdots, \lambda_n$ are the studentized order statistics defined by

$$\lambda_i = \frac{X_{(i)} - \bar{X}}{S_X}, 1 \leq i \leq n,$$

where $X_{(1)} \leq \cdots \leq X_{(n)}$ are the order statistics of $X_1, X_2, \cdots, X_n$. Note that the statistic shown in (20) is invariant under the scale transformation. In the rest of this section, we assume that $X_1, X_2, \cdots, X_n$ are iid $\Gamma(\alpha, \beta)$ random variables.

Hwang and Huang (2002) used the independence between $\bar{X}$ and $S_X/\bar{X}$ to obtain that

$$E(\frac{S_X^2}{\bar{X}^2}) = \frac{E(S_X^2)}{E(\bar{X}^2)} = \frac{n}{1+n\alpha}. \tag{21}$$

Combined with equation (18), the estimators for α and β were given by

$$\hat{\alpha}_c = \frac{\bar{X}^2}{S_X^2} - \frac{1}{n} \qquad and \qquad \hat{\beta}_c = \frac{\hat{\alpha}_c}{\bar{X}}. \tag{22}$$

They also compared these estimators with the usual moment estimators

$$\hat{\alpha}_m = \frac{\bar{X}^2}{S_X^2} \qquad and \qquad \hat{\beta}_m = \frac{\hat{\alpha}_m}{\bar{X}} \tag{23}$$

through a simulation study. The results indicate that $(\hat{\alpha}_c, \hat{\beta}_c)$ are better estimators of (α, β) than $(\hat{\alpha}_m, \hat{\beta}_m)$ under the cases that $\alpha = 0.5, 1.0, 1.5, 2.0, \beta = 1, 2, 4$, and $n = 5, 10, 15, 20, 25$ while the mean squared error (MSE) was used to carry out the comparison.

Note that $\hat{\alpha}_m - \hat{\alpha}_c = 1/n$ and $\hat{\beta}_m - \hat{\beta}_c = 1/(n\bar{X})$. Thus, both the difference between $\hat{\alpha}_m$ and $\hat{\alpha}_c$ and that between $\hat{\beta}_m$ and $\hat{\beta}_c$ will converge to zero almost surely as $n \to \infty$. But for a finite n, there are very few analytical comparisons made for these estimators.

Chou and Huang (2004c) compared $\hat{\alpha}_c$ with $\hat{\alpha}_m$ in a theoretical approach. Specifically, they proved that for every α, β and n, $MSE(\hat{\alpha}_c) < MSE(\hat{\alpha}_m)$. This justifies the above simulation results for the parameter α.

6.2. Parameter estimation on two populations

Consider the situation that we have an estimation problem of two gamma populations with a common scale parameter. This may arise when analyzing the lifetime distributions of two independent components, as mentioned at the end of Section 1. We can estimate the parameters using either (22) or the following method given by Chou and Huang (2004c).

Chou and Huang (2004c) presented a scheme for estimating the parameters of the two populations $X \sim \Gamma(\alpha_X, \beta)$ and $Y \sim \Gamma(\alpha_Y, \beta)$. Some estimators were given according to the property that the original gamma populations X and Y can be transformed into a beta population using $Z = X/(X+Y)$. More specifically, they proposed the estimators for α_X, α_Y and β by

$$\hat{\alpha}_{X,b} = \frac{\bar{Z}(\bar{Z} - \overline{Z^2})}{\overline{Z^2} - (\bar{Z})^2}, \quad \hat{\alpha}_{Y,b} = \frac{(1-\bar{Z})(\bar{Z} - \overline{Z^2})}{\overline{Z^2} - (\bar{Z})^2},$$
$$\hat{\beta}_{X,b} = \frac{\hat{\alpha}_{X,b}}{\bar{X}}, \qquad \hat{\beta}_{Y,b} = \frac{\hat{\alpha}_{Y,b}}{\bar{Y}}, \tag{24}$$

where

$$\bar{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i, \qquad \overline{Z^2} = \frac{1}{n}\sum_{i=1}^{n} Z_i^2,$$

and showed that these estimators converge to the true parameters almost surely as $n \to \infty$. This also implies that these estimators are consistent.

Now we compare these estimators with those proposed by Hwang and Huang (2002). The simulation results are based on $\alpha_X, \alpha_Y = 0.5, 1.0, \cdots, 6.0, \alpha_X \leq \alpha_Y, \beta = 1, 2, n = 10, 20, 30, 40$. For each n, we have $(\binom{12}{2} + 12) \times 2 = 156$ cases of $(\alpha_X, \alpha_Y, \beta)$ combination. For each case a random sample (X_i, Y_i) of size $n, n = 10, 20, 30, 40$, is generated and the procedure is repeated $1,000$ times. For an estimator in (24), if the MSE and MAB are both larger than that of Hwang and Huang (2002), then we say that estimator in (24) is worse; if just one of the MSE and MAB is larger and the other is smaller, then we say that the comparison is even; if both MSE and MAB are smaller, then we say that the estimator in (24) is better. The simulation results indicate that: For $n = 10$ the estimators by Hwang and Huang (2002) are in general better, while for $n = 20, 30, 40$, respectively, the estimators in (24) are better in most cases. The larger n, the better the estimators in (24) are. Besides, when the ratio α_Y/α_X is closer to 1, the estimators in (24) seem to be better. To clarify this observation, for each n we extract and analyze among the 156 cases by adding the constrain that either $\alpha_X \leq \alpha_Y < 2\alpha_X$ (84 cases) or $\alpha_X \leq \alpha_Y < 1.5\alpha_X$ (60 cases), respectively, and list the results in the tables. Note that when $n = 10$, the effect of the ratio α_Y/α_X is not very obvious, but when $n = 20, 30$ or 40, the effect is very obvious, especially for the estimators of α_Y and β_Y. To conclude, our estimators are better when:

(i) The true parameters are small, or
(ii) n is large, or
(iii) α_Y/α_X is close to 1.

We list the simulation results in Tables 1 to 4.

Tables 1 to 4: the numbers of cases the estimators given in (24) are worse/even/better.

Table 1: $n = 10$

scope of α-parameter	$\hat{\alpha}_X$	$\hat{\alpha}_Y$	$\hat{\beta}_X$	$\hat{\beta}_Y$
$\alpha_X \leq \alpha_Y$	116/20/20	137/11/8	84/28/44	109/18/29
$\alpha_X \leq \alpha_Y < 2\alpha_X$	67/7/10	69/7/8	45/13/26	57/12/15
$\alpha_X \leq \alpha_Y < 1.5\alpha_X$	49/4/7	48/6/6	33/8/19	41/9/10

Table 2: $n = 20$

scope of α-parameter	$\hat{\alpha}_X$	$\hat{\alpha}_Y$	$\hat{\beta}_X$	$\hat{\beta}_Y$
$\alpha_X \leq \alpha_Y$	32/33/91	106/23/27	14/12/130	70/20/66
$\alpha_X \leq \alpha_Y < 2\alpha_X$	19/19/46	38/19/27	7/6/71	14/16/54
$\alpha_X \leq \alpha_Y < 1.5\alpha_X$	16/14/30	26/10/24	6/5/49	13/7/40

Table 3: $n = 30$

scope of α-parameter	$\hat{\alpha}_X$	$\hat{\alpha}_Y$	$\hat{\beta}_X$	$\hat{\beta}_Y$
$\alpha_X \leq \alpha_Y$	8/10/138	79/23/54	2/5/149	52/12/92
$\alpha_X \leq \alpha_Y < 2\alpha_X$	7/9/68	16/17/51	2/4/78	7/5/72
$\alpha_X \leq \alpha_Y < 1.5\alpha_X$	7/8/45	11/12/37	2/3/55	5/3/52

Table 4: $n = 40$

scope of α-parameter	$\hat{\alpha}_X$	$\hat{\alpha}_Y$	$\hat{\beta}_X$	$\hat{\beta}_Y$
$\alpha_X \leq \alpha_Y$	3/7/146	68/9/79	1/0/155	42/13/101
$\alpha_X \leq \alpha_Y < 2\alpha_X$	3/6/75	5/6/73	1/0/83	0/3/81
$\alpha_X \leq \alpha_Y < 1.5\alpha_X$	3/5/52	2/4/54	1/0/59	0/2/58

The power for testing the hypothesis

$$H_0 : \alpha_Y = \alpha_{Y_{H0}} \qquad v.s. \qquad H_1 : \alpha_Y = \alpha_{Y_{H1}},$$

was also given by Chou and Huang (2004c) based on the simulation. In their study, 18 null hypotheses with the combination of $\beta = 1, 2, \alpha_X = 3, 4, 5$ and $\alpha_{Y_{H0}} = \alpha_X - 1, \alpha_X, \alpha_X + 1$, were considered. For each null hypothesis, there are 7 corresponding alternative hypotheses with $\alpha_{Y_{H1}} = \alpha_{Y_{H0}} - 1.5, \alpha_{Y_{H0}} - 1.0, \alpha_{Y_{H0}} - 0.5, \alpha_{Y_{H0}}, \alpha_{Y_{H0}} + 0.5, \alpha_{Y_{H0}} + 1.0, \alpha_{Y_{H0}} + 1.5$, respectively, where β and α_X remain the same as in the null hypothesis. Hence we have $18 \times 7 = 126$ hypotheses on a set of parameters $(\beta, \alpha_X, \alpha_{Y_{H0}}, \alpha_{Y_{H1}})$. To calculate the power of the test, each time a random sample (X_i, Y_i) of size $n, n = 10, 20, 30, 40, 50, 60$, is generated and the procedure is repeated $1,000$ times. Theoretical 95% confidence intervals are constructed based on $Z_i = X_i/(X_i + Y_i), i = 1, 2, \cdots, n$. According to the Central Limit Theorem, since $\bar{Z}$ is approximately normal with mean $\mu = \alpha_X/(\alpha_X + \alpha_Y)$ and variance $\sigma^2 = \alpha_X\alpha_Y/(n(\alpha_X + \alpha_Y)^2(\alpha_X + \alpha_Y + 1))$ for large n. The constructed 95% confidence intervals for μ is $(\bar{Z} - 1.96\sigma, \bar{Z} + 1.96\sigma)$. The testing powers are computed and part of the results are shown in Tables 5 and 6. Note that when $\alpha_{Y_{H1}} = \alpha_{Y_{H0}}$, the power is referred to the type I error.

We summarize the simulation results in the following:
(i) The change of the scale parameter β seems not influence the power, this is due to that the distribution of Z is irrelevant to the parameter β.
(ii) The power becomes smaller in general as the shape parameters $\alpha_X, \alpha_{Y_{H0}}$ become larger.

Table 5: The powers for the cases that $\beta = 1, \alpha_X = 3$.

				$\alpha_{Y_{H0}} = 2$				
$\alpha_{Y_{H1}} =$		0.5	1.0	1.5	2.0	2.5	3.0	3.5
n	30	1.000	0.984	0.463	0.047	0.292	0.788	0.975
	40	1.000	0.996	0.571	0.056	0.380	0.893	0.996
	50	1.000	1.000	0.662	0.054	0.484	0.959	0.999
	60	1.000	1.000	0.748	0.047	0.561	0.987	1.000
				$\alpha_{Y_{H0}} = 3$				
$\alpha_{Y_{H1}} =$		1.5	2.0	2.5	3.0	3.5	4.0	4.5
n	30	0.996	0.806	0.263	0.058	0.167	0.536	0.848
	40	0.999	0.907	0.340	0.054	0.261	0.662	0.943
	50	1.000	0.956	0.387	0.042	0.289	0.801	0.971
	60	1.000	0.972	0.471	0.048	0.389	0.854	0.994
				$\alpha_{Y_{H0}} = 4$				
$\alpha_{Y_{H1}} =$		2.5	3.0	3.5	4.0	4.5	5.0	5.5
n	30	0.943	0.572	0.173	0.049	0.141	0.394	0.663
	40	0.980	0.714	0.245	0.048	0.165	0.450	0.817
	50	0.991	0.817	0.273	0.042	0.204	0.631	0.895
	60	0.997	0.862	0.300	0.045	0.208	0.684	0.993

Table 6: The powers for the cases that $\beta = 2, \alpha_X = 3$.

$\alpha_{Y_{H0}} = 2$								
$\alpha_{Y_{H1}} =$		0.5	1.0	1.5	2.0	2.5	3.0	3.5
n	30	1.000	0.987	0.464	0.040	0.342	0.784	0.986
	40	1.000	0.998	0.572	0.051	0.392	0.904	0.995
	50	1.000	1.000	0.681	0.046	0.473	0.944	0.999
	60	1.000	1.000	0.706	0.048	0.541	0.975	1.000
$\alpha_{Y_{H0}} = 3$								
$\alpha_{Y_{H1}} =$		1.5	2.0	2.5	3.0	3.5	4.0	4.5
n	30	0.998	0.820	0.261	0.053	0.202	0.542	0.863
	40	1.000	0.905	0.350	0.048	0.223	0.679	0.939
	50	1.000	0.946	0.402	0.050	0.297	0.731	0.980
	60	1.000	0.972	0.454	0.041	0.347	0.848	0.994
$\alpha_{Y_{H0}} = 4$								
$\alpha_{Y_{H1}} =$		2.5	3.0	3.5	4.0	4.5	5.0	5.5
n	30	0.927	0.606	0.181	0.057	0.142	0.350	0.662
	40	0.978	0.765	0.242	0.050	0.169	0.488	0.799
	50	0.994	0.807	0.298	0.046	0.201	0.601	0.905
	60	0.999	0.883	0.350	0.047	0.254	0.674	0.955

7. Concluding remarks

In this paper we investigate some recent works on characterizations of gamma-related distributions and the parameter estimation of gamma distributions. What are left as possible future study are the characterizations of beta distribution using constancy of certain conditional expectations, as mentioned in Section 5. Some other statistical inference problems such as testing the hypothesis of $\alpha_X = \alpha_Y$ given a common scale parameter, or the discriminating method between two gamma populations with the same scale parameters will also be investigated.

Acknowledgement

We would like to thank the referees for their prompt and careful reading of this article as well as the detailed comments which greatly improved the original version.

References

Alzaid, A.A., Al-Osh, M.A. 1991. Characterizations of probability distributions based on the relation $X \stackrel{d}{=} U(X_1 + X_2)$. *Sankhyā B* **53**, 188-190.

Bobecka, K. 2002. Regression versions of Lukacs type characterizations for the bivariate gamma distribution. *J. Appl. Stat. Sci.* **11**, 213-233.

Bobecka, K., Wesolowski, J. 2002a. Three dual regression schemes for the Lukacs theorem. *Metrika* **56**, 55-72.

Bobecka, K., Wesolowski, J. 2002b. The Lukacs-Olkin-Rubin theoRem without invariance of the "quotient". *Studia Math.* **152**, 147-160.

Bolger, E.M., Harkness, W. L. 1965. A characterization of some distributions by conditional moments. *Ann. Math. Stat.* **36**, 703-705.

Casalis , M., Letac, G. 1996. The Lukacs-Olkin-Rubin characterization of Wishart distributions on symmetric cones. *Ann. Stat.* **24**, 763-786.

Chhikara, R.S., Folks, J.L. 1989. *The Inverse Gaussian Distribution: Theory, Methodology and Applications.* Marcel Dekker Inc., New York.

Chou, C.W., Huang, W.J. 2003. Characterizations of the gamma distribution via conditional moments. *Sankhyā* **65**, 271-283.

Chou, C.W., Huang, W.J. 2004a. A note on characterizations of the bivariate gamma distribution. To appear in *J. Stat. Plann. Inference.*

Chou, C.W., Huang, W.J. 2004b. On characterizations of the gamma and generalized inverse Gaussian distributions. To appear in *Stat. Prob. Lett.*

Chou, C.W., Huang, W.J. 2004c. On parameter estimation of two independent gamma population with the same scale parameter. Technical Report. National University of Kaohsiung, Kaohsiung, Taiwan.

Chung, K.L. 2001. *A Course in Probability Theory*, 3rd ed. Academic Press, New York.

Hall, W.J., Simons, G. 1969. On characterizations of the gamma distribution. *Sankhyā A* **31**, 385-390.

Huang, W.J., Chen, L.S. 1989. Note on a characterization of gamma distributions. *Stat. Prob. Lett.* **8**, 485-487.

Huang, W.J., Chen, L.S. 1991. On a study of certain power mixtures. *Chinese J. Math.* **19**, 95-104.

Huang, W.J., Chou, C.W. 2004. Characterizations of the gamma distribution via conditional expectations. Technical Report. National University of Kaohsiung, Kaohsiung, Taiwan.

Huang, W.J., Su, J.C. 1997. On a study of renewal process connected with certain conditional moments. *Sankhyā A* **59**, 28-41.

Hwang, T.Y., Hu, C.Y. 1999. On a characterization of the Gamma distribution: the independence of the sample mean and the sample coefficient of variation. *Ann. Inst. Stat. Math.* **51**, 749-753.

Hwang, T.Y., Hu, C.Y. 2000. On some characterizations of population distributions. *Taiwanese J. Math.* **4**, 427-437.

Hwang, T.Y., Huang, P.H. 2002. On new moment estimation of parameters of gamma distribution using it's characterization. *Ann. Inst. Stat. Math.* **54**, 840-847.

Johnson, N.L., Kotz, S. 1994. *Distributions in Statistics: Continuous Univariate Distributions I*, 2nd ed. John Wiley & Sons, New York.

Laha, R.G. 1964. On a problem connected with beta and gamma distributions, *Transactions of American Mathematical Society* **113**, 287-298.

Letac, G., Massam, H. 1998. Quadratic and inverse regressions for Wishart distributions. *Ann. Stat.* **26**, 573-595.

Letac, G., Wesolowski, J. 2000. An independence property for the product of GIG and gamma laws. *Ann. Prob.* **28**, 1371-1383.

Li, S.H., Huang, W.J., Huang, M.N.L. 1994. Characterizations of the Poisson process as a renewal process via two conditional moments. *Ann. Inst. Stat. Math.* **46**, 351-360.

Lukacs, E. 1955. A characterization of the gamma distribution. *Ann. Math. Stat.* **26**, 319-324.

Matsumoto, H., Yor, M. 2003. Interpretation via Brownian motion of some independence properties between GIG and gamma variables. *Stat. Prob. Lett.* **61**, 253-259.

Olkin, I., Rubin, H. 1962. A characterization of the Wishart distributions. *Ann. Math. Stat.* **33**, 1272-1280.

Pakes, A.G. 1992a. A characterization of gamma mixtures of stable laws motivated by limit theorems. *Statistica Neerlandica* **2-3**, 209-218.

Pakes, A.G. 1992b. On characterizations through mixed sums. *Austral. J. Stat.* **34**, 323-339.

Pakes, A.G., Khattree, R. 1992. Length-biasing, characterizations of laws and the moment problem. *Austral. J. Stat.* **34**, 307-322.

Pakes, A.G. 1994. Necessary conditions for characterization of laws via mixed sums. *Ann. Inst. Stat. Math.* **46**, 797-802.

Patil, G.P., Seshadri, V. 1964. Characterization theorems for some univariate probability distributions. *J. Roy. Stat. Soc. B* **26**, 286-292.

Pusz, J. 1997. Regressional characterization of the generalized inverse Gaussian population. *Ann. Inst. Stat. Math.* **49**, 315-319.

Pusz, J. 2002. Some extensions on the bivariate populations of the Laha-Lukacs characterizations. Technical report, Wydzial Matematyki I Nauk Informacyjnych, Politechnika Warszawska, Warszawa, Poland.

Seshadri, V. 1999. *The Inverse Gaussian Distribution, Statistical Theory and Applications.* Springer-Verlag Inc., New York.

Seshadri, V., Wesolowski, J. 2001. Mutual characterizations of the gamma and generalized inverse Gaussian laws by constancy of regression. *Sankhyā A* **63**, 107-112.

Seshadri, V., Wesolowski, J. 2003. Constancy of regressions for beta distributions. *Sankhyā* **65**, 284-291.

Tweedie, M.C.K. 1945. Inverse statistical variates. *Nature* **155**, 453-453.

Wang, Y.H. 1981. *Extensions of Lukacs' characterization of the gamma distribution. In: Analytical Methods in Probability Theory, Lecture Notes in Mathematics* **861**, Springer, New York, 166-177.

Wesolowski, J. 1989. A characterization of the gamma process by conditional moments. *Metrika* **36**, 299-309.

Wesolowski, J. 1990. A constant regression characterization of the gamma law. *Adv. Appl. Prob.* **22**, 488-490.

Wesolowski, J. 2002. The Matsumoto-Yor independence property for GIG and Gamma laws, revisited. *Math. Proc. Camb. Phil. Soc.* **133**, 153-161.

Yeo, G.F., Milne, R.K. 1991. On characterizations of beta and gamma distributions. *Stat. Prob. Lett.* **11**, 239-242.