Sankhyā : The Indian Journal of Statistics
2008, Volume 70-A, Part 1, pp. 73-87

Characterizations of the Beta Distributions via Some Regression Assumptions

Wen-Jang Huang
National University of Kaohsiung, Taiwan
Yan-Hau Chen
National University of Kaohsiung, Taiwan

Abstract

Let X and Y be two independent non-degenerate random variables. Also let (U, V) be a bijective map of (X, Y). It is desired to use certain regression assumptions between U and V to characterize the distributions of X and Y, and consequently, the distribution of (U, V). In most of the previous investigations, U and V turn out to be independent too.

Recently, for X, Y valued in $(0, 1)$, Seshadri and Wesołowski (2003) characterize X and Y to be beta distributed based on two constancy of regression assumptions between U and V, where (U, V) is a particular bijective map of (X, Y).

In this work, first we will generalize the results in Seshadri and Wesołowski (2003). It will be proved that for the bijective map given in Seshadri and Wesołowski (2003), X and Y are beta distributed under some more general regression assumptions. Next we illustrate that for some other special bijective maps (U, V), under certain regression assumptions between U and V, X and Y can also be characterized to be beta distributed, yet U and V may not be independent.

AMS (2000) *subject classification.* Primary 60E05; secondary 62E10.
Keywords and phrases. Beta distribution, characterization, conditional expectation, distribution theory, regression assumption.

1 Introduction

Let X and Y be two independent gamma random variables with the same scale parameter, i.e. X is $\Gamma(p, r)$ distributed and Y is $\Gamma(q, r)$ distributed, for some constants $p, q, r > 0$. Let

$$(S, T) = \left(\frac{X}{X+Y}, X+Y \right)$$

be a bijective map of (X, Y). Then it is known that S and T are mutually independent and have $\mathcal{B}e(p, q)$ and $\Gamma(p+q, r)$ distributions, respectively. Denote this as $(S, T) \sim \mathcal{B}e(p, q) \otimes \Gamma(p+q, r)$. Here $\Gamma(\alpha, \beta)$, $\alpha, \beta > 0$, denotes the gamma distribution with the probability density function (p.d.f.)

$$f(x) = \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)\beta^{\alpha}}, \ x > 0,$$

and $\mathcal{B}e(p, q)$, $p, q > 0$, denotes the beta distribution with the p.d.f.

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

In fact, independence of S and T is a property only enjoyed by the gamma distribution. More precisely, Lukacs (1955) proved that if X and Y are independent non-degenerate positive random variables and S and T are mutually independent, then X and Y have gamma distributions with the same scale parameter. Since then many papers considered different extensions. Among others, Huang and Su (1997), and Chou and Huang (2003) obtained similar characterization, under the weaker conditions

$$E(S^{s+1}|T) = aE(S^s|T) \ \text{ and } \ E(S^{r+s+1}|T) = bE(S^{r+s}|T),$$

with $r = 1$ and 2, respectively, s being some fixed integer, and a and b being some constants. Under the so-called dual regression schemes, that is instead of independence of X and Y, independence of S and T together with the regression conditions for X and Y:

$$E(Y^{s+1}|X) = aE(Y^s|X) \ \text{ and } \ E(Y^{s+2}|X) = bE(Y^{s+1}|X), \tag{1.1}$$

for some fixed integer s, where a and b are constants, were assumed, Chou and Huang (2003) proved X and Y are gamma distributed with the same scale parameter. A technique of change of measure for the traditional Laplace transform methods used by Huang and Chou (2004) to extend the above results to that s needs only to be a fixed real number.

Now let X and Y be two independent non-degenerate random variables. Besides the gamma law, there are many characterizations of X and Y, by using the independence, or some weaker regression assumptions, of a bijective map of X and Y. The following are some famous examples.

Matsumoto and Yor (2001) considered the bijective map

$$(M, N) = \left(\frac{1}{X} - \frac{1}{X+Y}, \frac{1}{X+Y}\right),$$

and proved that if X and Y are generalized inverse Gaussian (GIG) and Gamma distributed, respectively, then M and N are also independent and are Gamma and GIG distributed, respectively. The converse of the above result was studied by Letac and Wesołowski (2000), Seshadri and Wesołowski (2001), and Wesołowski (2002). Instead of the independent conditions, by giving two weaker conditions for M, N as in (1.1), Chou and Huang (2004) obtained similar characterizations.

Let X and Y be two independent non-degenerate random variables valued in $(0, 1)$. For the following bijective map

$$(U, V) = (\frac{1-Y}{1-XY}, 1-XY), \tag{1.2}$$

Seshadri and Wesołowski (2003) used constancy of regressions

$$E(U^i|V) = a \quad \text{and} \quad E(U^j|V) = b,$$

where $(i, j) = (1, 2)$ or $(1, -1)$, and a and b are constants, to characterize X and Y, and consequently, U and V, to be beta distributed.

In this paper, first we will generalize Seshadri and Wesołowski (2003) by using the regression assumptions similar to (1.1)

$$E(U^{s+1}|V) = aE(U^s|V), \tag{1.3}$$

and

$$E(U^{s+2}|V) = bE(U^{s+1}|V), \tag{1.4}$$

for some fixed real s, where a and b are constants. In particular $(i, j) = (1, 2)$ and $(1, -1)$ corresponds to $s = 0$ and -1, respectively. Some conditions which are equivalent or not equivalent to the bijective map (1.2) are also studied. Then in Section 3, we will illustrate that for some bijective maps, under suitable regression assumptions, X and Y can be characterized to be beta distributed, yet U and V are not independent.

2 Characterization of the Beta Distribution

In this section, first we extend the results of Seshadri and Wesołowski (2003).

THEOREM 2.1. Let X and Y be two independent non-degenerate random variables valued in $(0,1)$, and (U,V) be defined as in (1.2). Assume that (1.3) and (1.4) hold for some fixed real s such that $E((1-Y)^s) < \infty$, where a and b are constants. Then $0 < a < b < 1$, and there exists $p > 0$ such that $(X,Y) \sim \mathcal{B}e(p,q) \otimes \mathcal{B}e(p+q,r)$, where $q = (1-a)(1-b)/(b-a) > 0$ and $r = a(1-b)/(b-a) - s > 0$. Consequently, $(U,V) \sim \mathcal{B}e(r,q) \otimes \mathcal{B}e(r+q,p)$.

PROOF. First as U is also valued in $(0,1)$, hence $0 < a, b < 1$. On the other hand, $a < b$ follows from Hölder inequality. Next (1.3) and (1.4) imply

$$E((1-Y)^{s+1}|XY) = a(1-XY)E((1-Y)^s|XY), \tag{2.1}$$

and

$$E((1-Y)^{s+2}|XY) = b(1-XY)E((1-Y)^{s+1}|XY), \tag{2.2}$$

respectively. Substituting (2.1) into (2.2) yields

$$E((1-Y)^{s+2}|XY) = ab(1-XY)^2 E((1-Y)^s|XY). \tag{2.3}$$

From (2.1) and (2.3), for every integer $k \geq 0$, we have

$$E((1-Y)^{s+1}(XY)^k) = aE((1-XY)(1-Y)^s(XY)^k), \tag{2.4}$$

and

$$E((1-Y)^{s+2}(XY)^k) = abE((1-XY)^2(1-Y)^s(XY)^k). \tag{2.5}$$

Since X and Y are valued in $(0,1)$ with $E((1-Y)^s) < \infty$, we have $E(X^k) < \infty$, and $E((1-Y)^s Y^k) \leq E((1-Y)^s) < \infty$ for every integer $k \geq 0$. Now denote

$$g(k) = \frac{E(X^{k+1})}{E(X^k)} \quad \text{and} \quad h(k) = \frac{E[(1-Y)^s Y^{k+1}]}{E[(1-Y)^s Y^k]}.$$

Then independence of X and Y, and (2.4) and (2.5) lead to

$$h(k) = \frac{1-a}{1-ag(k)}, \tag{2.6}$$

and

$$1-2h(k)+h(k)h(k+1) = ab-2abg(k)h(k)+abg(k)h(k)g(k+1)h(k+1), \tag{2.7}$$

respectively. Substituting $h(k)$ and $h(k+1)$ into (2.7), we obtain

$$(a-b)g(k)g(k+1)+(1-2a+ab)g(k+1)=(1-2b+ab)g(k)+b-a. \qquad (2.8)$$

By denoting $q=(1-a)(1-b)/(b-a)>0$, (2.8) can be rewritten as

$$g(k+1)=\frac{1+(q-1)g(k)}{q+1-g(k)}. \qquad (2.9)$$

Define p by $g(0)=E(X)=p/(p+q)$. As $0<E(X)<1$, and $q>0$, it can be seen easily $p>0$. Then (2.9) leads to

$$g(k)=\frac{E(X^{k+1})}{E(X^k)}=\frac{k+p}{k+p+q}, \qquad (2.10)$$

for every integer $k\geq 0$. Since X has bounded support, by the celebrated Carleman criterion (see e.g., Chung (2001)), the distribution of X is determined by its moments. We obtain that X is $\mathcal{B}e(p,q)$ distributed.

Next, substituting (2.10) into (2.6), it follows that

$$h(k)=\frac{E[(1-Y)^sY^{k+1}]}{E[(1-Y)^sY^k]}=\frac{k+p+q}{k+p+q/(1-a)}=\frac{k+p+q}{k+p+cq}, \qquad (2.11)$$

where $c=(1-a)^{-1}>1$. Now let F_Y denote the distribution function of Y. As in Chou and Huang (2003), define a new probability measure G on $(0,1)$ by

$$\delta(1-y)^sF_Y(dy)=G(dy), \qquad (2.12)$$

where $\delta^{-1}=E((1-Y)^s)<\infty$. Let Z be a valued $(0,1)$ random variable with the distribution function G. Then (2.11) and (2.12) yield

$$\frac{E(Z^{k+1})}{E(Z^k)}=\frac{k+p+q}{k+p+cq},$$

for every integer $k\geq 0$. Hence Z is $\mathcal{B}e(p+q,(c-1)q)$ distributed. In view of (2.12), we obtain $(c-1)q-s>0$. Consequently, (2.14) implies Y is $\mathcal{B}e(p+q,(c-1)q-s)=\mathcal{B}e(p+q,r)$ distributed, where $r=aq/(1-a)-s>0$, as required.

When the two independent random variables X and Y both are valued in $(0,1)$, except the (U,V) defined in (1.2), we are interested in knowing whether there are other bijective maps of X and Y, which can also be used to characterize both X and Y to be beta distributed. Before answering this question, first we illustrate the connection between Lukacs type characterization for gamma distributions and the present characterization for beta distributions.

Let X_1, Y_1, Z_1 be independent gamma random variables with the same scale parameter, say

$$(X_1, Y_1, Z_1) \sim \Gamma(p,t) \otimes \Gamma(q,t) \otimes \Gamma(r,t),$$

where $p, q, r, t > 0$. Then it is well-known that

$$\left(\frac{X_1}{X_1+Y_1}, \frac{X_1+Y_1}{X_1+Y_1+Z_1}\right) \sim \mathcal{B}e(p,q) \otimes \mathcal{B}e(p+q,r), \tag{2.13}$$

$$\left(\frac{Z_1}{Y_1+Z_1}, \frac{Y_1+Z_1}{X_1+Y_1+Z_1}\right) \sim \mathcal{B}e(r,q) \otimes \mathcal{B}e(r+q,p), \tag{2.14}$$

and

$$\left(\frac{X_1}{X_1+Z_1}, \frac{X_1+Z_1}{X_1+Y_1+Z_1}\right) \sim \mathcal{B}e(p,r) \otimes \mathcal{B}e(p+r,q). \tag{2.15}$$

Let $X = X_1/(X_1+Y_1)$ and $Y = (X_1+Y_1)/(X_1+Y_1+Z_1)$, and substitute X and Y into the right hand side of (1.2), it yields

$$(U,V) = \left(\frac{Z_1}{Y_1+Z_1}, \frac{Y_1+Z_1}{X_1+Y_1+Z_1}\right).$$

In view of (2.14), this explains why U and V defined in (1.2) are independent and have $\mathcal{B}e(r,q)$, $\mathcal{B}e(r+q,p)$ distributions, respectively, when $(X,Y) \sim \mathcal{B}e(p,q) \otimes \mathcal{B}e(p+q,r)$.

It is expected that there is a bijective map of (X,Y) corresponding to (2.15). Indeed this is true. Define

$$(U_1,V_1) = \left(\frac{XY}{1-Y+XY}, 1-Y+XY\right) \tag{2.16}$$

which obviously is a bijective map of (U,V). Again by substituting $(X,Y) = (X_1/(X_1+Y_1), (X_1+Y_1)/(X_1+Y_1+Z_1))$ into the right hand side of (2.16), we obtain

$$(U_1,V_1) = \left(\frac{X_1}{X_1+Z_1}, \frac{X_1+Z_1}{X_1+Y_1+Z_1}\right).$$

Consequently, in view of (2.15), U_1 and V_1 defined in (2.16) are independent, when $(X, Y) \sim \mathcal{B}e(p, q) \otimes \mathcal{B}e(p + q, r)$. Hence it is not surprising that there exists a characterization of beta distributions by using the independent assumption of U_1 and V_1. But this is not a new result as can be seen by first rewriting

$$(U_1, V_1) \quad = \quad \left(1 - \frac{1 - Y}{1 - X'Y}, 1 - X'Y\right),$$

where $X' = 1 - X$, and noting that independence of U_1 and V_1 implies $(1 - Y)/(1 - X'Y)$ and $1 - X'Y$ are independent, hence X' and Y, and consequently X and Y are beta distributed. Yet characterization by using constancy of regression assumptions based on U_1 and V_1 is another story. We state and prove the result in the following Theorem.

Theorem 2.2. *Let X and Y be two independent non-degenerate random variables valued in $(0, 1)$, and (U_1, V_1) be defined as in (2.16). Assume that*

$$E(U_1^{s+1}|V_1) = aE(U_1^s|V_1), \tag{2.17}$$

and

$$E(U_1^{s+2}|V_1) = bE(U_1^{s+1}|V_1), \tag{2.18}$$

hold for some fixed real s such that $E(X^s) < \infty$ and $E(Y^s) < \infty$, where a and b are constants. Then $0 < a < b < 1$, and there exists $q > 0$ such that $(X, Y) \sim \mathcal{B}e(p, q) \otimes \mathcal{B}e(p + q, r)$, where $p = a(1 - b)/(b - a) - s > 0$ and $r = (1-a)(1-b)/(b-a) > 0$. Consequently, $(U_1, V_1) \sim \mathcal{B}e(p, r) \otimes \mathcal{B}e(p+r, q)$.

Proof. As in Theorem 2.1, we obtain $0 < a < b < 1$ immediately. Next (2.17) and (2.18) imply that for every integer $k \geq 0$,

$$E[((1 - X')Y)^{s+1}(X'Y)^k] = aE[(1 - X'Y)((1 - X')Y)^s(X'Y)^k], \tag{2.19}$$

and

$$E[((1 - X')Y)^{s+2}(X'Y)^k] = bE[(1 - X'Y)((1 - X')Y)^{s+1}(X'Y)^k], \tag{2.20}$$

respectively, where $X' = 1 - X$. Since X' and Y are valued in $(0, 1)$ with $E((1 - X')^s) = E(X^s) < \infty$ and $E(Y^s) < \infty$, we have $E((1 - X')^s(X')^k) \leq E((1 - X')^s) < \infty$, and $E(Y^{s+k}) \leq E(Y^s) < \infty$, for every integer $k \geq 0$. Now denote

$$g(k) = \frac{E[(1 - X')^s(X')^{k+1}]}{E[(1 - X')^s(X')^k]} \quad \text{and} \quad h(k) = \frac{E(Y^{s+k+1})}{E(Y^{s+k})}. \tag{2.21}$$

Following similar steps as in Theorem 2.1, (2.19) leads to

$$g(k)=\frac{h(k)-a}{(1-a)h(k)}, \tag{2.22}$$

and (2.20) leads to

$$h(k+1)=\frac{1+(r-1)h(k)}{r+1-h(k)}, \tag{2.23}$$

where $r=(1-a)(1-b)/(b-a)>0$. Let $F_{X'}$ and F_Y denote the distribution functions of X' and Y, respectively. Again define two new probability measures G_1 and G_2 on $(0,1)$ by

$$\delta_1(1-x)^s F_{X'}(dx)=G_1(dx) \quad \text{and} \quad \delta_2 y^s F_Y(dy)=G_2(dy), \tag{2.24}$$

respectively, where $\delta_1^{-1}=E((1-X')^s)<\infty$ and $\delta_2^{-1}=E(Y^s)<\infty$. Let W_1 and W_2 be valued $(0,1)$ random variables with the distribution function G_1 and G_2, respectively. Then (2.24) yields

$$\frac{E(W_1^{k+1})}{E(W_1^k)}=\frac{E[\delta_1(1-X')^s(X')^{k+1}]}{E[\delta_1(1-X')^s(X')^k]}=g(k),$$

and

$$\frac{E(W_2^{k+1})}{E(W_2^k)}=\frac{E[\delta_2 Y^{s+k+1}]}{E[\delta_2 Y^{s+k}]}=h(k),$$

for every integer $k\ge 0$. Define now m by $h(0)=E(W_2)=m/(m+r)$. As $0<E(W_2)<1$, and $r>0$, it can be seen easily $m>0$. Again (2.23) yields

$$h(k)=\frac{E(W_2^{k+1})}{E(W_2^k)}=\frac{k+m}{k+m+r}, \tag{2.25}$$

for every integer $k\ge 0$. Thus we obtain that W_2 is $\mathcal{B}e(m,r)$ distributed, and consequently, Y is $\mathcal{B}e(m-s,r)$ distributed.

Next, substituting (2.25) into (2.22), it follows that

$$g(k)=\frac{E(W_1^{k+1})}{E(W_1^k)}=\frac{k+m-ar/(1-a)}{k+m}<1, \tag{2.26}$$

for every integer $k\ge 0$. Hence $q=m-ar/(1-a)>0$, and W_1 is $\mathcal{B}e(q,n)$ distributed with $n=m-q=ar/(1-a)>0$. Then (2.24) implies X' is

$\mathcal{B}e(q, n-s) = \mathcal{B}e(q,p)$ distributed, where $p = n - s > 0$. Consequently, X is $\mathcal{B}e(p,q)$ distributed, and Y is $\mathcal{B}e(p+q,r)$ distributed, as required.

Next, we consider

$$(U_2, V_2) = \left(\lambda - \frac{1-Y}{1-XY}, 1-XY\right),$$

where λ is a real constant, which is a bijective map of (X,Y) more general than (U_1,V_1). It is desired to use constancy of regression assumptions to characterize the distribution of (X,Y). As the three consecutive moments $E(U_2^s|V_2)$, $E(U_2^{s+1}|V_2)$, and $E(U_2^{s+2}|V_2)$, do not have a simple relationship when $(X,Y) \sim \mathcal{B}e(p,q) \otimes \mathcal{B}e(p+q,r)$, we do not have results similar to Theorem 2.1. Yet when $s = 0$, we can use the following assumptions

$$E(U_2|V_2) = \lambda - c, \tag{2.27}$$

and

$$E(U_2^2|V_2) = \lambda^2 - d\lambda + e, \tag{2.28}$$

where c, d, and e are some constants, to characterize X and Y to be beta distributed. But again as can be seen below, this is not a new result. Substituting $(U_2,V_2) = (\lambda - U, V)$ into (2.27) and (2.28), where (U,V) is defined in (1.2), yields

$$E(U|V) = c,$$

and

$$E(U^2|V) = (2c-d)\lambda + e = b_1.$$

Consequently, conditions (2.27) and (2.28) are the same as the constancy of regression conditions (1.3) and (1.4), when $s = 0$, $a = c$, and $b = b_1/a$. Except this trivial case, for the bijective map (U_2,V_2), we do not have characterization results by using regression assumptions for U_2 and V_2.

Now consider another bijective map

$$(U_3, V_3) = \left(\frac{Y}{X+Y-XY}, X+Y-XY\right).$$

Again it is easy to see that (U_3,V_3) can be rewritten as

$$(U_3, V_3) = \left(\frac{1-Y'}{1-X'Y'}, 1-X'Y'\right),$$

where $X' = 1-X$ and $Y' = 1-Y$. Hence characterization of (X,Y) by using independence, or weaker conditions of constancy of regression, of U_3 and V_3, is equivalent to characterization of (X,Y) by using parallel conditions for U and V. For example, if

$$E(U_3^{s+1}|V_3) = aE(U_3^s|V_3),$$

and

$$E(U_3^{s+2}|V_3) = bE(U_3^{s+1}|V_3),$$

for some fixed real s, where a and b are some constants, then $(X',Y') \sim \mathcal{B}e(p,q)\otimes\mathcal{B}e(p+q,r)$. Consequently, $(X,Y) \sim \mathcal{B}e(q,p)\otimes\mathcal{B}e(r,p+q)$, where p, q and r are given in Theorem 2.1.

3 Characterization of the Beta Distribution Based on non-independent U and V

In this section, we illustrate that for some bijective maps (U,V) of (X,Y), under suitable regression assumptions, X and Y can be characterized to be beta distributed, yet U and V are not independent. First we consider the following bijective map

$$(U_5, V_5) = (\frac{1+\eta Y}{1-XY}, 1-XY), \tag{3.1}$$

where η is a real constant. If $\eta = -1$, then $(U_5,V_5) = (U,V)$, where (U,V) is defined in (1.2). Note that this is the only case that U_5 and V_5 are independent when $(X,Y) \sim \mathcal{B}e(p,q)\otimes\mathcal{B}e(p+q,r)$. We have the following result.

Theorem 3.1. *Let X and Y be two independent non-degenerate random variables valued in $(0,1)$. Let (U_5,V_5) be defined as in (3.1). Assume that*

$$E(U_5|V_5) = (1+\eta)\frac{1}{V_5} - c\eta, \tag{3.2}$$

and

$$E(U_5^2|V_5) = (1+\eta)^2\frac{1}{V_5^2} - 2c\eta(1+\eta)\frac{1}{V_5} + cd\eta^2, \tag{3.3}$$

hold for some constants c, d, where $d<1$, and $\eta \neq 0$. Then $0<c<d$, and there exists $p>0$ such that $(X,Y) \sim \mathcal{B}e(p,q)\otimes\mathcal{B}e(p+q,r)$, where

$q = (1-c)(1-d)/(d-c) > 0$ *and* $r = c(1-d)/(d-c) > 0$.

PROOF. Again $0 < c < d$ can be obtained easily. Also (3.2) and (3.3) imply for every integer $k \geq 0$,

$$E((1+\eta Y)(XY)^k) = E((1+\eta - c\eta(1-XY))(XY)^k), \tag{3.4}$$

and

$$E((1+\eta Y)^2(XY)^k) = E(((1+\eta)^2 - 2c\eta(1+\eta)(1-XY) + cd\eta^2(1-XY)^2)(XY)^k), \tag{3.5}$$

respectively. Now denote

$$g(k) = \frac{E(X^{k+1})}{E(X^k)} \quad \text{and} \quad h(k) = \frac{E(Y^{k+1})}{E(Y^k)}.$$

Then for $\eta \neq 0$, (3.4) leads to

$$g(k) = \frac{h(k) - 1 + c}{ch(k)}, \tag{3.6}$$

and (3.5) leads to

$$h(k+1) = \frac{1 + (r-1)h(k)}{r + 1 - h(k)}, \tag{3.7}$$

which is exactly the same as (2.23) with $r = c(1-d)/(d-c) > 0$ here. Hence Y is $\mathcal{B}e(m,r)$ distributed, where $m > 0$ is defined by $h(0) = E(Y) = m/(m+r)$.

Next, substituting $h(k)$ into (3.6), it follows that

$$g(k) = \frac{k + m - r(1-c)/c}{k+m} < 1. \tag{3.8}$$

Hence $p = m - r(1-c)/c > 0$, and X is $\mathcal{B}e(p,q)$ distributed with $q = m - p = r(1-c)/c > 0$. Consequently, Y is $\mathcal{B}e(p+q,r)$ distributed, as required.

In Theorem 3.1, the joint p.d.f. of (U_5, V_5) is

$$f_{U_5,V_5}(u_5,v_5) = \begin{cases} Cv_5(1-v_5)^{p-1}(\eta v_5-(\eta+1-u_5v_5))^{q-1}(\eta+1-u_5v_5)^{r-1}, \\ \qquad \frac{\eta+1-\eta v_5}{v_5} < u_5 < \frac{\eta+1}{v_5},\ 0<v_5<1, \text{ if } \eta>0, \\ -Cv_5(1-v_5)^{p-1}(\eta v_5-(\eta+1-u_5v_5))^{q-1}(\eta+1-u_5v_5)^{r-1}, \\ \qquad \frac{\eta+1}{v_5} < u_5 < \frac{\eta+1-\eta v_5}{v_5},\ 0<v_5<1, \text{ if } \eta<0, \end{cases} \tag{3.9}$$

where $C=\Gamma(p+q+r)\eta^{1-q-r}/(\Gamma(p)\Gamma(q)\Gamma(r))$. Obviously, from (3.1), the marginal distribution of V_5 is still beta distributed with parameters $r+q$ and p. Yet U_5 is not beta distributed any more, if $\eta \neq -1$. As mentioned it before, $\eta=-1$ is the only case that U_5 and V_5 are independent when $(X,Y) \sim \mathcal{B}e(p,q)\otimes \mathcal{B}e(p+q,r)$. This also can be seen from (3.9). When $\eta=-1$, (3.2) and (3.1) become

$$E(U_5|V_5)=c \quad \text{and} \quad E(U_5^2|V_5)=cd,$$

which corresponds to the case $s=0$ in Theorem 2.1 with $a=c$, and $b=d$.

Next, we state without proof a result based on the following bijective map

$$(U_6,V_6)=\left(\frac{Y}{1-XY},1-XY\right). \tag{3.10}$$

Theorem 3.2. *Let X and Y be two independent non-degenerate random variables valued in $(0,1)$. Let (U_6,V_6) be defined as in (3.10). Assume that*

$$E(U_6|V_6)=\frac{1}{V_6}-c,$$

and

$$E(U_6^2|V_6)=\frac{1}{V_6^2}-2c\frac{1}{V_6}+cd,$$

hold for some constants c and d, where $d<1$. Then $0<c<d$, and there exists $p>0$ such that $(X,Y)\sim \mathcal{B}e(p,q)\otimes \mathcal{B}e(p+q,r)$, where $q=(1-c)(1-d)/(d-c)>0$ and $r=c(1-d)/(d-c)>0$.

Obviously, Theorem 3.2 together with (3.10) imply the marginal distribution of V_6 is still beta distributed with parameters $r+q$ and p, but U_6 is not beta distributed. Also U_6 and V_6 are not independent.

For the above two bijective maps (U_5, V_5) and (U_6, V_6), $V_5 = V_6 = V$, and both the denominators of U_5 and U_6 are the same as that of U. The last bijective map we consider in this work is

$$(U_7, V_7) = \left(\frac{1-Y}{XY}, XY\right), \tag{3.11}$$

which is also a modification of (U, V), where U_7 has the same numerator as that of U, yet V_7 is the same as the denominator of U, but is slightly different from V. For this new map, we can use the general regression assumptions as in Theorem 2.1 to characterize the joint distribution of (X, Y). The proof is still omitted.

Theorem 3.3. *Let X and Y be two independent non-degenerate random variables valued in $(0,1)$. Let (U_7, V_7) be defined as in (3.11). Assume that*

$$E(U_7^{s+1}|V_7) = c\frac{1-V_7}{V_7}E(U_7^s|V_7), \tag{3.12}$$

and

$$E(U_7^{s+2}|V_7) = d\frac{1-V_7}{V_7}E(U_7^{s+1}|V_7), \tag{3.13}$$

hold for some fixed real s such that $E((1-Y)^s) < \infty$, where c and d are constants. Then $0 < c < d < 1$, and there exists $p > 0$ such that $(X, Y) \sim \mathcal{B}e(p, q) \otimes \mathcal{B}e(p+q, r)$, where $q = (1-c)(1-d)/(d-c) > 0$ and $r = c(1-d)/(d-c) - s > 0$.

Again in Theorem 3.3, the marginal distribution of V_7 is $\mathcal{B}e(p, r+q)$ distributed with non-beta distributed U_7. Also U_7 and V_7 are not independent.

4 Conclusion

From Theorem 1.2, it is known that the distribution of (X, Y) can be determined by the regression assumptions with the form such as (1.3) and (1.4), based on either (U, V) or (U_1, V_1). The latter is equal to $(1-U, V)$. But this is not the case for the bijective map (U_7, V_7). For this map, based on $(U_8, V_8) = (1-U_7, V_7)$, we only can use (3.12) and (3.13) with $s = 0$, to obtain $(X, Y) \sim \mathcal{B}e(p, q) \otimes \mathcal{B}e(p+q, r)$.

Finally, it should be mentioned here, not every bijective map of (X, Y) can be used to characterize the distribution of (X, Y) to be beta distributed.

For example, consider the following bijective map

$$(U_9, V_9) = \left(\frac{1-Y}{X}, X\right).$$

Then consider the regression assumptions

$$E(U_9^{s+1}|V_9) = c\frac{1}{V_9}E(U_9^s|V_9), \tag{4.1}$$

and

$$E(U_9^{s+2}|V_9) = d\frac{1}{V_9}E(U_9^{s+1}|V_9), \tag{4.2}$$

which hold when $(X, Y) \sim \mathcal{B}e(p, q) \otimes \mathcal{B}e(p+q, r)$. Obviously, (4.1) and (4.20) are equivalent to

$$E((1-Y)^{s+1}) = cE((1-Y)^s),$$

and

$$E((1-Y)^{s+2}) = dE((1-Y)^{s+1}),$$

which are satisfied by any valued in $(0, 1)$ random variables X and Y.

Acknowledgements. We are grateful to the referee for the valuable comments and suggestions.

References

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

Chou, C.-W. and Huang, W.-J. (2004). On characterizations of the gamma and generalized inverse Gaussian distributions. *Statist. & Probab. Lett.*, **69**, 381–388.

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

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

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

Letac, G. and Wesołowski, J. (2000). An independence property for the product of GIG and gamma laws. *Ann. Probab.*, **28**, 1371–1383.

Lukacs, E. (1955). A characterization of the gamma distribution. *Ann. Math. Statist.*, **26**, 319–324.

Matsumoto, H. and Yor, M. (2001). An analogue of Pitman's $2M - X$ theorem for exponential Wiener functionals. Part II: The role of the generalized inverse Gaussian laws. *Nagoya Math. J.*, **162**, 65–86.

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

Seshadri, V. and Wesołowski, J. (2003). Constancy of regressions for beta distributions. *Sankhyā*, **65**, 284–291.

Wesołowski, J. (2002). The Matsumoto-Yor independence property for GIG and Gamma laws, revisited. *Math. Proc. Camb. Philos. Soc.*, **133**, 153–161.

Wen-Jang Huang
Department of Applied Mathematics
National University of Kaohsiung
Kaohsiung, Taiwan
E-mail: huangwj@nuk.edu.tw

Yan-Hau Chen
Institute of Statistics
National University of Kaohsiung
Kaohsiung, Taiwan
E-mail: hau0630@hotmail.com

Paper received August 2007; revised February 2008.