Metrika (2002) 56: 43–54

Metrika

Three dual regression schemes for the Lukacs theorem

Konstancja Bobecka and Jacek Wesołowski

Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska, Plac Politechniki 1, 00-661 Warszawa, Poland

Abstract. In the paper we study regressional versions of Lukacs' characterization of the gamma law. We consider constancy of regression instead of Lukacs' independence condition in three new schemes. Up to now the constancy of regressions of $U = X/(X+Y)$ given $V = X+Y$ for independent X and Y has been considered in the literature. Here we are concerned with constancy of regressions for X and Y while independence of U and V is assumed instead.

Key words and phrases: Characterization of probability distributions, beta distribution, gamma distribution, constancy of regression, method of moments, difference equations

1. Introduction

Denote by $\gamma_{p,b}$ the gamma distribution defined by the density

$$\gamma_{p,b}(dx) = \frac{b^p}{\Gamma(p)} x^{p-1} e^{-bx} I_{(0,\infty)}(x)\,dx,$$

where b, p are positive numbers (scale and shape parameters, respectively). By $\beta_{p,q}$ denote the beta distribution with the density

$$\beta_{p,q}(dx) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} x^{p-1}(1-x)^{q-1} I_{(0,1)}(x)\,dx,$$

where p, q are positive numbers.

Let $\psi : (0,\infty)^2 \to (0,1) \times (0,\infty)$ be a mapping defined by

$$\psi(x,y) = \left(\frac{x}{x+y}, x+y\right).$$

For a random vector (X, Y) with nondegenerate, positive components define $(U, V) = \psi(X, Y)$. It is well known that $(X, Y) \sim \gamma_{p,b} \otimes \gamma_{q,b}$ iff $(U, V) \sim \beta_{p,q} \otimes \gamma_{p+q,b}$ for positive numbers p, q, b ($\otimes$ denotes the product of two measures).

One of the most fascinating results in the area of characterization of probability distributions is the Lukacs (1955) theorem which says that if both (X, Y) and $(U, V) = \psi(X, Y)$ have independent components then X and Y have gamma distributions with the same scale parameter. Since then many authors have considered different extensions and complements of Lukacs' result. Many of them were concerned with regression schemes, i.e. the assumption that X and Y are independent was kept, but instead of independence of U and V conditions of constancy of regressions

$$E(U^r|V) = c, \quad E(U^s|V) = d \tag{1.1}$$

for some real constants c, d, and some fixed pairs of integers (r, s) were imposed. In this way Bolger and Harkness (1965) characterized the gamma law by considering the pair $(r, s) = (1, 2)$. Then Wesołowski (1990) obtained the characterization for the pair $(r, s) = (1, -1)$. In Li, Huang and Huang (1994) it was shown that the characterization holds also in the case $(r, s) = (-1, -2)$. In the meantime Hall and Simons (1969) showed that for the pair $(r, s) = (1, 3)$ the result does not hold. Analogous constancy of regression conditions in more abstract settings were also considered; for instance, bivariate random vectors in Wang (1981), stochastic processes in Wesołowski (1989), positive definite symmetric matrices and Jordan algebras in Letac and Massam (1998). Related regression conditions of the form

$$E(Y^r \,|\, X + Y) = c(X + Y)^r, \quad E(X^r \,|\, X + Y) = d(X + Y)^r$$

were studied in Hall and Simons (1969) for $r = 2$ and in Huang and Su (1997) for $r = -1$.

However it is still not known if the role of U and V in regression conditions (1.1) can be exchanged without influencing the characterization property. Here we would like only to indicate that combining two constancy of regressions conditions with different conditioning as $E(U|V) = c$ and $E(V|U) = d$ does not determine uniquely the gamma law, which is shown in the following:

Example 1. *Take X and Y to be independent identically distributed (iid) random variables (rv's), such that $P(X = a) = P(X = b) = 1/2$ for two distinct and positive numbers a, b. Then, obviously, $E(U|V) = 1/2$, which holds true for any iid integrable rv's, and direct computations lead immediately to: $E(V|U) = a + b$.*

On the other hand it was proved in Khatri and Rao (1968) (see also Kagan, Linnik and Rao (1973), Ch. 6) that if we take three or more variables in the above scheme then the characterization follows. More precisely, these authors, assuming that $X_1, \ldots, X_n$, $n \geq 3$, are independent rv's, considered the condition

$$E(X_1 + \cdots + X_n \,|\, X_1/X_n, \ldots, X_{n-1}/X_n) = const.$$

In the present paper we are concerned with schemes which are in a natural way dual to the regression schemes described by (1.1), i.e. instead assuming that X and Y are independent we assume that U and V are independent, at the same time the constancy of regressions of U given V is changed to constancy of regressions of Y given X or X given Y:

$$E(Y^r|X) = c, \quad E(Y^s|X) = d,$$

or

$$E(X^r|Y) = c, \quad E(X^s|Y) = d$$

for some pairs of (r, s) and some real constants c and d. Thus here the role of X and Y in regression conditions can be interchanged. We are concerned precisely with the same pairs as in the previous setting, i.e. we prove characterizations in three cases: $(r, s) = (1, 2)$, or $(1, -1)$ or $(-1, -2)$. It should be stressed that our argument in the proofs does not make any use of the Laplace transform technique, which has always been the basic tool for dealing with Lukacs' theorem related problems. Instead we rely on the method of moments, which seems to be used for the first time in this area.

Also it is shown, that similarly as in the previous approach, mixing regression conditions by taking different conditioning does not characterize the gamma law.

2. Results

Throughout this section we assume that (X, Y) is a random vector with nondegenerate positive components and $(U, V) = \psi(X, Y)$, where ψ was defined in the previous section. Observe, that since ψ is a bijection, equivalently, one can start with a random vector (U, V) with nondegenerate components, first in $(0, 1)$ and second positive, and define $(X, Y) = \psi^{-1}(U, V)$.

First we consider a result which is dual to the characterization of the gamma law obtained in Bolger and Harkness (1965).

Theorem 1. *Let U and V be independent. Assume that*

$$E[Y|X] = c, \quad E[Y^2|X] = d, \tag{2.1}$$

for some real constants c, d.

Then $d > c^2$ and there exists $a > 0$ such that $(U, V) \sim \beta_{p,q} \otimes \gamma_{a,b}$, where $b = c/(d - c^2) > 0$, $p = a - bc > 0$ and $q = bc > 0$. Moreover $(X, Y) \sim \gamma_{p,b} \otimes \gamma_{q,b}$.

Proof. Conditions (2.1) can be rewritten as

$$E[V \mid UV] = c + UV, \tag{2.2}$$

and

$$E[V^2(1-U)^2 \mid UV] = d. \tag{2.3}$$

Let us note that since $U \in (0,1)$ we have $E(U^k) < \infty$ for $k = 1, 2, \dots$ To show that all the moments $E(V^k)$ are finite we apply induction with respect to k. For $k = 1, 2$ it follows from the conditions (2.2) and (2.3). Now, let us assume that $E(V^k) < \infty$ for some k and observe that if rv's A, B $(B \geq 0)$, $C = AE(B|A)$ are integrable then the product AB is also integrable – see Lemma in Wesołowski (1993). Hence, by taking

$$A = (UV)^k, \quad B = V(1-U), \quad C = AE(B|A) = (UV)^k c,$$

we obtain that $V^{k+1}U^k(1-U)$ is integrable. Thus $E(V^{k+1}) < \infty$ and by the induction argument we get that all the moments of V are finite.

From (2.2) it follows that

$$E[V(UV)^k] = E[(c+UV)(UV)^k], \quad k = 0, 1, \dots, \tag{2.4}$$

which takes the form

$$h(k) = c + g(k)h(k), \quad k = 0, 1, \dots, \tag{2.5}$$

where

$$h(k) = \frac{E[V^{k+1}]}{E[V^k]}, \quad g(k) = \frac{E[U^{k+1}]}{E[U^k]}, \quad k = 0, 1, \dots$$

On the other hand (2.3) implies that

$$E[V^2(1-U)^2(UV)^k] = dE[(UV)^k].$$

Thus, we obtain the equation

$$\begin{aligned} &h(k+1)h(k) - 2g(k)h(k+1)h(k) \\ &\qquad + g(k+1)g(k)h(k+1)h(k) = d, \quad k = 0, 1, \dots \end{aligned} \tag{2.6}$$

Substituting $g(k)h(k)$ and $g(k+1)h(k+1)$ from (2.5) into (2.6) we obtain

$$h(k+1) = h(k) + \frac{d}{c} - c, \quad k = 0, 1, \dots \tag{2.7}$$

Let us observe that since $c = E(Y)$, $d = E(Y^2)$, we have $c > 0$ and $Var(Y) = d - c^2 > 0$. If we denote

$$b = \frac{c}{d - c^2} > 0,$$

the equation (2.7) takes the form

$$h(k) = h(0) + k/b, \quad k = 0, 1, \dots \tag{2.8}$$

Let us define $a = bh(0) = bEV$. Then the above equation leads to

$$E(V^k) = \frac{\Gamma(a+k)}{b^k \Gamma(a)}, \quad k = 1, 2, \ldots$$

Hence, by the uniqueness of the moments sequence for the gamma distribution, we get $V \sim \gamma_{a,b}$.

From (2.5) it follows that

$$g(k) = \frac{h(k) - c}{h(k)} = \frac{a - bc + k}{a + k}, \quad k = 0, 1, \ldots$$

Thus

$$E(U^{k+1}) = E(U^k)\frac{a - bc + k}{a + k}, \quad k = 0, 1, \ldots \tag{2.9}$$

Denote $p = a - bc$, $q = bc$ and observe that $q > 0$, $p = bEV - bc = bEUEV > 0$. Then the equation (2.9) takes the form

$$E(U^k) = \frac{p + k - 1}{p + q + k - 1} E(U^{k-1}) = \frac{\Gamma(p+q)\Gamma(p+k)}{\Gamma(p)\Gamma(p+q+k)}, \quad k = 1, 2, \ldots$$

The uniqueness of the moments sequence for the beta distribution implies that $U \sim \beta_{p,q}$. Now in the standard way we compute the joint density of $(X, Y) = (UV, (1-U)V)$ to conclude that $(X, Y) \sim \gamma_{p,b} \otimes \gamma_{q,b}$. □

The analoguous result with X and Y interchanged follows immediately.

Corollary 1. *If we replace the assumption* (2.1) *in Theorem* 1 *by*

$$E[X|Y] = c, \quad E[X^2|Y] = d \tag{2.10}$$

then $(X, Y) \sim \gamma_{q,b} \otimes \gamma_{p,b}$, *with* b, p, q *defined as in Theorem* 1.

Proof. The conditions (2.10) are equivalent to

$$E[UV \mid V(1-U)] = c,$$

$$E[(UV)^2 \mid V(1-U)] = d.$$

Denote $Z = 1 - U$. It suffices to observe that Z, V are independent and then from Theorem 1 we get $(Z, V) \sim \beta_{q,p} \otimes \gamma_{a,b}$. Hence $(X, Y) \sim \gamma_{q,b} \otimes \gamma_{p,b}$. □

Our next result takes care about the pair $(r, s) = (1, -1)$ and is dual to Wesołowski (1990).

Theorem 2. *Let* U *and* V *be independent. Assume that*

$$E[Y|X] = c, \quad E[Y^{-1}|X] = d \tag{2.11}$$

for some real constants c, d.

Then $cd > 1$ *and there exists* $a > 1$ *such that* $(U, V) \sim \beta_{p,q} \otimes \gamma_{a,b}$, *where* $b = \frac{d}{cd-1} > 0$, $p = a - bc > 0$, $q = bc > 1$. *Moreover* $(X, Y) \sim \gamma_{p,b} \otimes \gamma_{q,b}$.

Proof. Conditions (2.11) can be rewritten as

$$E[V(1-U) \mid UV] = c, \tag{2.12}$$

$$E\left[\frac{1}{V(1-U)} \mid UV\right] = d, \tag{2.13}$$

Let us note that similarly as in Theorem 1 all the moments of U and V are finite. Observe that $E\left(\frac{1}{1-U}\right) < \infty$. Note also that

$$\frac{U^k}{1-U} = \frac{1}{1-U} - \sum_{j=0}^{k-1} U^j = \sum_{j=k}^{\infty} U^j < \frac{1}{1-U},$$

$$E\left(\frac{U^k}{1-U}\right) < E\left(\frac{1}{1-U}\right) < \infty.$$

From (2.13) it follows that

$$E\left[V^{k-1}\frac{U^k}{1-U}\right] = d(UV)^k, \quad k = 0, 1, \ldots$$

The above equation can be written as

$$E(V^{k-1})\sum_{j=k}^{\infty} E(U^j) = dE(V^k)E(U^k), \quad k = 0, 1, \ldots$$

If we denote

$$G(k) = \sum_{j=k}^{\infty} E(U^j),$$

$$h(k-1) = \frac{E(V^k)}{E(V^{k-1})}, \quad k = 0, 1, \ldots$$

(observe that $E(V^{-1}) < \infty$) our equation takes the form

$$G(k) = dh(k-1)[G(k) - G(k+1)], \quad k = 0, 1, \ldots$$

what can be also written as

$$P(k) = 1 - \frac{1}{dh(k-1)}, \quad k = 0, 1, \ldots \tag{2.14}$$

where

$$P(k) = \frac{G(k+1)}{G(k)}, \quad k = 0, 1, \ldots \tag{2.15}$$

The equation (2.12) is the same as in Theorem 1. Thus we get

$$E[U^k]E[V^{k+1}] = cE[U^k]E[V^k] + E[U^{k+1}]E[V^{k+1}],$$

what can be written as

$$[h(k) - c][1 - P(k)] = h(k)P(k)[1 - P(k+1)], \quad k = 0, 1, \ldots \tag{2.16}$$

Substituting $P(k)$ and $P(k+1)$ from (2.14) into (2.16) we obtain:

$$\begin{aligned} h(k) &= h(k-1) + c - \frac{1}{d} \\ &= h(-1) + (k+1)/b, \quad k = 0, 1, \ldots, \end{aligned}$$

where

$$h(-1) = \frac{1}{E(V^{-1})}, \quad b = \frac{d}{cd - 1}.$$

Observe that

$$\begin{aligned} cd &= E[V(1-U)]E([V(1-U)]^{-1}) \\ &= [EVEV^{-1}][E(1-U)E(1-U)^{-1}] > 1. \end{aligned}$$

Hence $b = \dfrac{d}{cd-1} > 0$. Define $a = bEV = bh(0) > 0$. Then we get

$$E(V^k) = \frac{\Gamma(a+k)}{b^k \Gamma(a)}, \quad k = 1, 2, \ldots,$$

and thus $V \sim \gamma_{a,b}$.

Now, from (2.14) we obtain

$$P(k) = \frac{dh(k-1) - 1}{dh(k-1)} = \frac{d[h(-1) + k/b] - 1}{d[h(-1) + k/b]}, \quad k = 0, 1, \ldots \tag{2.17}$$

Let us note that $a = bh(0) = bh(-1) + 1$, hence (2.17) can be rewritten as

$$P(k) = \frac{a - \frac{b}{d} - 1 + k}{a + k - 1} = \frac{p + k}{p + q + k - 1}, \quad k = 0, 1, \ldots, \tag{2.18}$$

where $p = a - 1 - b/d = a - bc$, $q = a - p = 1 + b/d > 1$ and

$$p = \frac{d}{cd-1}EV - \frac{1}{cd-1} - 1 = bEX > 0.$$

On the other hand, from (2.15) we have

$$P(k) = \frac{E\left(\frac{U^{k+1}}{1-U}\right)}{E\left(\frac{U^k}{1-U}\right)}, \quad k = 0, 1, \ldots$$

Hence our equation (2.18) takes the form

$$E\left(\frac{U^{k+1}}{1-U}\right) = \frac{p+k}{p+q+k-1}E\left(\frac{U^k}{1-U}\right), \quad k = 0, 1, \ldots,$$

which implies that

$$E(U^k) = \frac{p+k-1}{p+q+k-1}E(U^{k-1})$$
$$= \frac{\Gamma(p+q)\Gamma(p+k)}{\Gamma(p)\Gamma(p+q+k)}, \quad k = 1, 2, \ldots$$

Hence $U \sim \beta_{p,q}$. Thus, as in Theorem 1, it follows that $(X, Y) \sim \gamma_{p,b} \otimes \gamma_{q,b}$. □

And, again, the roles of X and Y in the previous result can be exchanged.

Corollary 2. *If we replace the assumption* (2.11) *in Theorem* 2 *by*

$$E(X|Y) = c, \quad E(X^{-1}|Y) = d, \tag{2.19}$$

then $(X, Y) \sim \gamma_{q,b} \otimes \gamma_{p,b}$, *where* p, q, b *are defined in Theorem* 2.

Our final scheme is dual to one adopted in Li, Huang, and Huang (1994), and deals with the pair $(r, s) = (-1, -2)$.

Theorem 3. *Let U and V be independent rv's. Assume that*

$$E[Y^{-1}|X] = c, \quad E[Y^{-2}|X] = d \tag{2.20}$$

for some real constants c, d.

Then $d > c^2$ *and there exists* $a > 2$ *such that* $(U, V) \sim \beta_{p,q} \otimes \gamma_{a,b}$, *where* $b = cd/(d - c^2) > 0$, $p = a - 1 - b/c > 0$, $q = 1 + b/c > 2$. *Moreover* $(X, Y) \sim \gamma_{p,b} \otimes \gamma_{q,b}$.

Proof. Conditions (2.20) can be rewritten as

$$E\left[\frac{1}{V(1-U)} \mid UV\right] = c, \tag{2.21}$$

$$E\left[\frac{1}{V^2(1-U)^2} \mid UV\right] = d. \tag{2.22}$$

Since $U \in (0,1)$ we have $E(U^k) < \infty$ for $k = 1, 2, \dots$ To show that all the moments $E(V^k)$ are finite we apply induction with respect to $k = -2, -1, 0, 1, \dots$. Observe that by the regression conditions it follows that $E(1-U)^{-2} < \infty$ and $EV^{-2} < \infty$. Let us assume that $E(V^k) < \infty$ for some k. From (2.21) we have

$$U^{k+1}V^{k+1}c = E\left[V^k \frac{U^{k+1}}{(1-U)} \mid UV\right],$$

hence

$$\infty > E(V^k)E\left(\frac{1}{1-U}\right) > E(V^k)E\left(\frac{U^{k+1}}{1-U}\right) = E(V^{k+1})E(U^{k+1})c$$

Thus $E(V^{k+1}) < \infty$ and by the induction argument we get that all the moments of V are finite.

Let us note that the condition (2.21) is as the one in Theorem 2. Hence we have the equation

$$P(k) = 1 - \frac{1}{ch(k-1)}, \quad k = 0, 1, \dots \tag{2.23}$$

where

$$P(k) = \frac{G(k+1)}{G(k)},$$

$$h(k-2) = \frac{E(V^{k-1})}{E(V^{k-2})}, \quad k = 0, 1, \dots$$

The condition (2.22) implies

$$E(V^{k-2})E\left[\frac{U^k}{(1-U)^2}\right] = dE(V^k)E(U^k), \quad k = 0, 1, \dots$$

which can be written in the form

$$E\left[\frac{U^k}{(1-U)^2}\right] = d[G(k) - G(k+1)]h(k-1)h(k-2), \quad k = 0, 1, \dots$$

where as in the proof of Theorem 2

$$G(k) = \sum_{j=k}^{\infty} E(U^j), \quad k = 0, 1, \dots$$

Let us observe that ($'$ denotes differentiation with respect to U)

$$E\left[\frac{U^k}{(1-U)^2}\right] = E\left[U^k\left(\frac{1}{1-U}\right)'\right] = E\left[U^k\left(\sum_{j=0}^{\infty} U^j\right)'\right]$$
$$= \sum_{j=k}^{\infty}(j-k+1)E(U^j) = \sum_{j=k}^{\infty} G(j), \quad k = 0, 1, \dots$$

Thus we get

$$\sum_{j=k}^{\infty} G(j) = d[G(k) - G(k+1)]h(k-1)h(k-2), \quad k = 0, 1, \dots$$

Then taking differences for k and $k+1$ we obtain

$$G(k) = d[G(k) - G(k+1)]h(k-1)h(k-2)$$
$$- d[G(k+1) - G(k+2)]h(k)h(k-1),$$

which can be written as

$$1 = dh(k-1)\{[1 - P(k)]h(k-2) - P(k)$$
$$[1 - P(k+1)]h(k)\}, \quad k = 0, 1, \dots \tag{2.24}$$

Substituting now $P(k)$ and $P(k+1)$ from (2.23) into (2.24) it follows that

$$h(k-1) = h(k-2) + \frac{1}{c} - \frac{c}{d} = h(k-2) + \frac{1}{b}, \quad k = 0, 1, \dots,$$

where

$$b = \frac{cd}{d - c^2} = \frac{E[Y^{-1}]E[Y^{-2}]}{Var[Y^{-1}]} > 0.$$

Defining $a = bEV = bh(0) > 0$, we obtain, as in the previous proofs, that $V \sim \gamma_{a,b}$. Thus (2.23) takes the form

$$P(k) = \frac{a - \frac{b}{c} - 1 + k}{a + k - 1}, \quad k = 0, 1, \dots,$$

and we get

$$E(U^k) = \frac{p+k-1}{p+q+k-1} E(U^{k-1})$$
$$= \frac{\Gamma(p+q)\Gamma(p+k)}{\Gamma(p)\Gamma(p+q+k)}, \quad k = 1, 2, \ldots,$$

where

$$p = a - \frac{b}{c} - 1 = b[h(-1) - c^{-1}] = b \frac{1}{E(V)} \frac{[E((1-U)^{-1}) - 1]}{E((1-U)^{-1})} > 0,$$

$$q = a - p = 2 + \frac{c^2}{d - c^2} > 2.$$

Hence $U \sim \beta_{p,q}$. Thus, as in Theorem 1, $(X, Y) \sim \gamma_{p,b} \otimes \gamma_{q,b}$. □

The parallel result follows easily, as in the previous cases.

Corollary 3. *If we replace the assumption* (2.20) *in Theorem* 3 *by*

$$E(X^{-1}|Y) = c, \quad E(X^{-2}|Y) = d, \tag{2.25}$$

then $(X, Y) \sim \gamma_{q,b} \otimes \gamma_{p,b}$ *with* p, q, b *defined as in Theorem* 3.

Finally, let us note that the conditions:

$$E(Y|X) = c, \quad E(X|Y) = d,$$

where c and d are constants, are not sufficient for the non-degenerate positive random variables X and Y to be gamma distributed, assuming only that $U = \frac{X}{X+Y}$ and $V = X + Y$ are independent. It is demonstrated by the following:

Example 2. *Suppose that* $U = \frac{X}{X+Y}$ *and* $V = X + Y$ *are non-degenerate independent rv's such that*

$$P(U = a) = P(U = 1 - a) = \frac{1}{2},$$

$$P(V = a) = 1 - a = 1 - P(V = 1 - a)$$

for some $0.5 \neq a \in (0, 1)$.

Then the direct computation shows that

$$E(Y|X) = E(X|Y) = a(1 - a).$$

References

[1] Bolger EM, Harkness WL (1965) A characterization of some distributions by conditional moments. Ann Math Statist 36:703–705
[2] Hall WJ, Simons G (1969) On characterizations of the gamma distribution. Sankhyā Ser A 31:385–390
[3] Huang WJ, Su JC (1997) On a study of renewal process connected with certain conditional moments. Sankhyā Ser A 59:28–41
[4] Kagan AM, Linnik YuV, Rao CR (1973) Characterization problems of mathematical statistics. Wiley, New York
[5] Khatri CG, Rao CR (1968) Some characteristics of the gamma distribution. Sankhyā Ser A 30:157–166
[6] Letac G, Massam H (1998) Quadratic and inverse regressions for Wishart distributions. Ann Statist 26:573–595
[7] Li SH, Huang WJ, Huang MNL (1994) Characterizations of the Poisson process as a renewal process via two conditional moments. Ann Inst Statist Math 46:351–360
[8] Lukacs E (1955) A characterization of the gamma distribution. Ann Math Statist 26:319–324
[9] Wang YH (1981) Extensions of Lukacs' characterization of the gamma distribution. In: Analytical Methods in Probability Theory. Lect Notes in Math 861, Springer, New York, 166–177
[10] Wesołowski J (1989) A characterization of the gamma process by conditional moments. Metrika 36:299–309
[11] Wesołowski J (1990) A constant regression characterization of the gamma law. Adv Appl Prob 22:488–490
[12] Wesołowski J (1993) Stochastic processes with linear conditional expectation and quadratic conditional variance. Probab Math Statist 14:33–44